AMERICAN JOURNAL OF PHYSICAL ANTHROPOLOGY 139:512–522 (2009)

Genetic Heterogeneity in Regional Populations of Quebec—Parental Lineages in the Gaspe Peninsula

Claudia Moreau,[1] Hélène Vézina,[2] Vania Yotova,[1] Robert Hamon,[1] Peter de Knijff,[3] Daniel Sinnett,[1,4] and Damian Labuda[1,4]*

[1]*Centre de Recherche, CHU Sainte-Justine, Montréal, PQ, Canada H3T 1C5*
[2]*Interdisciplinary Research Group in Demography and Genetic Epidemiology, Université du Québec à Chicoutimi, Chicoutimi, PQ, Canada G7H 2B1*
[3]*Department of Human Genetics, Center of Human and Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands*
[4]*Département de Pédiatrie, Université de Montréal, Montréal, PQ, Canada H3T 1C5*

KEY WORDS parental lineages; colonization of Quebec; haplotype frequency classes

ABSTRACT Stable colonization of the Gaspe Peninsula by Europeans started in the middle of the 18th century at the time of the British conquest of New France. The earliest settlers were Acadians, escaping British deportation policies, followed by Loyalists from the US, who preferred to remain under British rule after the Declaration of Independence. In the 19th century, the developing fishing industry attracted French Canadians from the St. Lawrence Valley and newcomers from Europe including Channel Islanders from Jersey and Guernsey. We analyzed parental lineages of the self-declared descendants of these four groups of settlers by mtDNA D-loop sequencing and Y-chromosome genotyping and compared them with French, British, and Irish samples. Their representation in terms of haplotype frequency classes reveals different signatures of founder effects, such as a loss of rare haplotypes, modification of intermediate frequency haplotypes, reduction in genetic diversity (seen in Acadians), but also enrichment by admixture. Parental lineages correlate with group identity. Descendants of early settlers, Acadians and Loyalists, preserved their identity more than those of French Canadian and Channel Islander "latecomers." Although overall genetic diversity among Gaspesians is comparable with their European source populations, F_{ST} analysis indicated their greater differentiation. Distinct settlement history, a limited number of founders and relative genetic isolation contributed to the regionalization of the Quebec gene pool that appears less homogenous than usually anticipated. Am J Phys Anthropol 139:512–522, 2009.

Significant effort has been deployed to characterize the genetic structure of post-Columbian populations of the New World (Tian et al., 2006; Price et al., 2007; Wang et al., 2008). One of the underlying motivations is the practical use of such information, defining shared ancestry and/or admixture, to refine association studies and help map genetic diseases and susceptibilities (Price et al., 2008). Less attention was given to populations of European descent who colonized North America via the St. Lawrence River. Seventeenth century settlements along this route led to the foundation of New France (Charbonneau et al., 2000), later renamed Canada with its French-speaking province of Quebec. In medical genetics, Quebec is known for the presence of a number of specific genetic disorders, evoking an initial founder effect as a cause, especially because most of the population can trace its ancestry back to the early founders (Bouchard and De Braekeleer, 1991). However, in spite of the resulting belief that Quebec is genetically homogenous, the distribution of its particular Mendelian disorders and the underlying mutations is uneven across regions (Scriver, 2001; Laberge et al., 2005). Likewise, genealogical studies indicate a differential regional contribution of the early founders, suggesting genetic stratification across the province (Tremblay et al., 2003; Vezina et al., 2005). The underlying genetic population structure can be examined by the analysis of parental lineages (Shriver and Kittles, 2004). In this article, we focus on the Gaspe Peninsula region, which we will refer to as Gaspesia. Its name likely derives from the Mi'kmaq's (Algonquian language) word Gespeg meaning "land's end." Gaspesia is bordered by the estuary of the St. Lawrence River to the north, the Gulf of St. Lawrence to the east, and the Baie-des-Chaleurs to the south. The native Mi'kmaqs still inhabit the peninsula today.

French colonization of the St. Lawrence Valley started in 1608 with the foundation of Quebec City. Settlers subsequently spread along the shores of the St. Lawrence

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

Grant sponsors: The Réseau de Médecine Génétique Appliquée of the Fonds de la Recherche en Santé du Québec, the Canadian Institutes of Health Research, and the National Science and Engineering Research Council of Canada.

*Correspondence to: Damian Labuda, Centre de Recherche, CHU Sainte-Justine, 3175 Côte Sainte-Catherine, Montréal (Québec), Canada H3T 1C5. E-mail: damian.labuda@umontreal.ca

Received 12 June 2008; accepted 1 December 2008

DOI 10.1002/ajpa.21012
Published online 18 February 2009 in Wiley InterScience (www.interscience.wiley.com).

and Trois-Rivières was founded in 1634 and Montreal in 1642. Although most French pioneers came to the valley before 1680 (Charbonneau et al., 2000), a second important wave of immigrants of French ancestry occurred just before the British conquest, in the middle of the 18th century. Among those immigrants were the Acadians, descendants of French pioneers from Acadia (presently Nova Scotia and New Brunswick), escaping the British deportation campaign in Nova Scotia that started in 1755. A group of Acadians then settled in Gaspesia. Subsequently, in 1784, a group of English-speaking United Empire Loyalists, loyal to the British Crown, came to Gaspesia escaping the American Revolution (War of Independence). After the British conquest in the 1760s, French immigration to New France ceased, but its French-speaking population continued to grow, reaching 700,000 in the middle of the 19th century (Henripin and Peron, 1973). This increase in population size contributed to range expansion into new regions that were peripheral to the initial settlements. At that time, French Canadians from the Upper St. Lawrence moved to Gaspesia, attracted by its developing fishing, naval and lumber industries. There they were joined by Channel Islanders, mostly men, from the English Channel islands of Jersey and Guernsey, and a small number of immigrants who came directly from England and Ireland but who no longer represent a distinctly identifiable group (Desjardins et al., 1999). The principal groups of European origin still discernible in Gaspesia are Acadians, Loyalists, French Canadians, and Channel Islanders. Although their relative contributions to the contemporary population are not well defined, English-speakers represent 9% of the population and 5% are Protestants (2001 Canadian census http://www.stat.gouv.qc.ca/).

We sampled the population of Gaspesia specifically targeting these four populations. This article presents the analysis of their patrilineal and matrilineal lineages and compares them among themselves and with the published data on the three contemporary European samples: French, English, and Irish. Starting with Ewens formula (Ewens, 1972; Chakraborty and Weiss, 1991), we introduce plots of haplotype frequency classes. They facilitate visual analysis of the data in the framework of the infinitely-many-alleles model (Kimura and Crow, 1964) by direct comparison with theoretical expectations under the mutation-drift equilibrium. Our analysis reveals important demographic factors that shaped the genetic population structure of Gaspesians and provides us with insight to understanding the genetics of post-Columbian populations in terms of the consequences of their colonization history. In Gaspesia, parental lineages correlate well with group identities, preserved more among descendants of early groups of settlers than among those who came later. As a result, these groups appear genetically further apart than the populations of Western Europe, suggesting an important genetic structure among different regions of Quebec and less overall genetic homogeneity than generally believed.

MATERIALS AND METHODS

Samples

Peripheral blood samples were obtained from 397 unrelated (at least until the third generation, i.e., not allowing first cousins) individuals from Gaspesia of self-declared ethnic affiliation as French Canadians (n = 106), Acadians (n = 104), Loyalists (n = 94), or Channel Islanders (n = 93). The participants provided informed consent and their genealogical information, which were sent directly to the Groupe de recherche interdisciplinaire en démographie et épidémiologie génétique (GRIG) at the Université du Québec à Chicoutimi (UQAC). At the same time, blood samples were coded and separately sent to the CHU Sainte-Justine, Montreal, Canada. Upon completion of genealogical reconstructions, all nominative data on participants and their ancestors were coded before analyses were performed. The research protocol was approved by the respective Institutional Review Boards. Our data on Gaspesian populations were compared with the literature data on French (n = 1,127), British (n = 380), and Irish (n = 300) samples (Piercy et al., 1993; Richards et al., 1996; Rousselet and Mangin, 1998; Helgason et al., 2001; Dubut et al., 2004; McEvoy et al., 2004; Richard et al., 2007).

Mitochondrial DNA sequencing and genotyping

DNA was extracted using the Puregene DNA Purification kit (Gentra). PCR and sequencing reactions were as previously described (Heyer et al., 2001) except for the sequence TTGAGGAGGTAAGCTACATA of the reverse primer MTH00580. One microlitre of the reaction mixture was directly used for sequencing using the thermosequenase cycle sequencing kit (USB Corporation, Cleveland, Ohio). The products were revealed by electrophoresis in 5.5% polyacrylamide gel in the Li-Cor IR^2 sequencing system (LiCor Biosciences). Although we determined the HVS1 sequence between positions 16,069 and 16,383, only the portion from position 16,090 to 16,365 was considered in the comparative analyses. MtDNA haplogroups were primarily determined based on the mutations within HVS1 and HVS2 (positions from 58 to 370) (Horai et al., 1993; Torroni et al., 1996; Richards et al., 1998). To distinguish among haplogroups HV, H, and U, coding region positions 7,028, 14,766, and 12,308 were additionally typed by allele-specific-oligonucleotide hybridization (Bourgeois and Labuda, 2004). Sequences were verified for phantom mutations by reads of both strands, by two independent persons and additionally checking for mutations causing reticulations in haplotype networks (Bandelt et al., 2002).

Y-chromosome SNP and STR typing

Seven short tandem repeats, STRs (DYS19, DYS389 I, DYS389 II, DYS390, DYS391, DYS392, DYS393), and 13 simple polymorphisms (M181, $RPS4Y_{711}$, $SRY_{10831.1}$, SRY_{4064}, M170, M213, M172, M9, M175, M45, M173, $SRY_{10831.2}$, M17) of the Y chromosome were typed as described (de Knijff et al., 1997; Kayser et al., 1997). Haplotypes were formed from the STR alleles, whereas the haplogroups were determined based on simple polymorphisms according to (YCC 2002).

Statistical analysis

The infinite allele model was used to analyze variations among the Y-chromosome and mtDNA haplotypes (Chakraborty, 1990; Chakraborty and Weiss, 1991; Helgason et al., 2003). This was justified by the complexity of these haplotypes combining stable polymorphisms characterizing haplogroups with rapidly changing loci defining haplotypes. The haplotypes are defined by sets of microsatellites in the Y-chromosome or by numerous polymorphic sites within HVS1 and HVS2 in mtDNA. In

fact, complex sets of stepwise mutating markers reveal very low levels of homoplasy (Estoup et al., 2002; Yotova et al., 2007), as also shown in the case of Y-chromosome haplotypes (Pereira et al., 2003). Summary statistics and other population parameters were estimated using ARLEQUIN v. 3.11 (http://cmpg.unibe.ch/software/arlequin3/) (Excoffier et al., 2005). Gene (haplotype) diversity, which represents the probability that two haplotypes randomly chosen from a sample are different, was estimated as

$$G = [n/(n-1)]\left(1 - \sum_{i=1}^{k} p_i^2\right) \quad (1)$$

where n is the sample size, k is the number of distinct haplotypes, and p_i is the frequency of each haplotype. Both k and G can be compared in terms of a common population mutation parameter $\theta = 2N_e\mu$, where N_e is the effective female or male population size and μ is the mutation rate/generation per the analyzed portion of the mtDNA or Y-chromosome, respectively. Because μ should be the same in all populations, the differences in θ can be accounted for by historical fluctuations in the population size, reflecting these populations' demographic past. Furthermore, the indices k and G, and their derived θ_k and θ_G estimates, are expected to be differentially affected by demographic processes causing departures from the mutation–genetic drift equilibrium (migration, founder effect, population growth or admixture), justifying their joint use in describing population diversity. From the observed k and G, the corresponding θ_k and θ_G are obtained using the formulas

$$E(k) = \theta_k \sum_{i=0}^{n-1} 1/(\theta_k + i) \quad (2)$$

by Ewens (1972) and

$$E(G) = \theta_G/(1 + \theta_G) \quad (3)$$

by Kimura and Crow (1964), respectively, as implemented in ARLEQUIN, where θ_G, there calledθ_{hom}, is estimated according to Chakraborty and Weiss (1991). The observed distribution of haplotypes can be presented in terms of frequency classes r, either by counting the number of distinct haplotypes k_r, each occurring r-times in a sample ($r = 1, 2, \ldots, n$), such that $k = \sum_{r=1}^{n} k_r$, or by counting the number of chromosomes $r \cdot k_r$ in each frequency class (i.e., sample occupancy), such that $n = \sum_{r=1}^{n} r \cdot k_r$. The expected sample configuration [k_1, k_2,... k_n] given θ can be obtained from

$$E(k_r) = \frac{\theta}{r} \cdot \frac{n!}{(n-r)!} \cdot \frac{\Gamma(n+\theta-r)}{\Gamma(n+\theta)} \quad (4)$$

The equation above is the same as Eq. (7) from Chakraborty (1990), who noted "that because $k = \sum_{r=1}^{n} k_r$, if θ is estimated by θ_k, even though the expected value of k will agree with the observed k, there is no guarantee that, for each r, the observed k_r will agree with expected k_r," here given by Eq. (4). To identify the frequency classes that are responsible for differences between the expected and observed distribution, the expected configurations based on the estimated values of θ_k and θ_G, can be confronted with the observed values on the histograms of the haplotype occupancy $r \cdot k_r$ at each frequency class r (see Figs. 1 and 2 below). We chose the histograms of $r \cdot k_r$ rather than of k_r for practical reasons; in this way the data points at high r values remain noticeable on the plot. The overall significance of these differences can be tested (Ewens, 1972) using i) the Ewens-Watterson homozygosity test (Watterson, 1978), confronting the expected value G, given the number of haplotypes, with its observed value, ii) Slatkin exact test confronting observed and expected sample configurations, given k and n (Slatkin, 1996), thus testing mutation-drift equilibrium, and iii) the Chakraborty test, comparing observed k with their expected number, given observed G (Chakraborty, 1990). These tests, often referred to as neutrality tests, were carried out as implemented in the Arlequin package (Excoffier et al., 2005).

RESULTS

Mitochondrial DNA lineages

HVS1 and HVS2 of mtDNA were sequenced in 394 Gaspesian samples of Acadian, French Canadian, Channel Islander, and Loyalist ancestry (Supporting Information Table S1; see also S10). Table 1 presents the analysis of Gaspesian HVS1 haplotypes in comparison with the data on French, British, and Irish samples (see Materials and Methods and Supporting Information Table S2). Except for Acadians, the genetic diversity of the Gaspesians when expressed by G and the corresponding θ_G is comparable with that of the three European samples. In contrast, θ_k estimates differ dramatically between Europeans and Gaspesians reflecting noticeable differences in k, even taking the variation in sample size into account (for example, when all Gaspesians are considered together, they have fewer haplotypes than the Irish represented by a smaller number of individuals). Clearly, different measures of genetic diversity provide different views on the relationship between local New World populations of European descent and their European populations of origin. To understand these results, it is useful to directly examine the data by comparing observed haplotype distributions with the expected distributions given the estimated values of θ. In Figure 1, these are presented as the histograms of the haplotype frequency classes obtained by grouping haplotypes of the same multiplicity r (frequency class) and plotting them versus the number of genomic (mtDNA) copies $r \cdot k_r$ within each of these classes (class occupancy). In other words, these histograms stand for the observed sample configurations. The theoretical distributions of $r \cdot k_r$ as a function of r were computed using Eq. (4), given θ_k or θ_G estimate. The resulting plots reveal the source of the deviation from the theoretical curve (see Fig. 1).

Both in Gaspesian and European populations, there exists at least one haplotype of unusually high frequency. It appears as an outlier, far to the right from the rest of the data, in a tail of the theoretical distributions traced either using the θ_k or the θ_G estimate. Otherwise, the theoretical curve based on θ_k describes relatively well the remaining data in Gaspesians, suggesting that a single (or at most a few) high frequency haplotypes "distort" the distribution. The theoretical curve based on θ_G fits neither the high nor the low haplotype frequency classes. This is understandable, because a single high frequency haplotype is sufficient to dramatically affect the value of G and thus θ_G. In contrast, θ_k would be

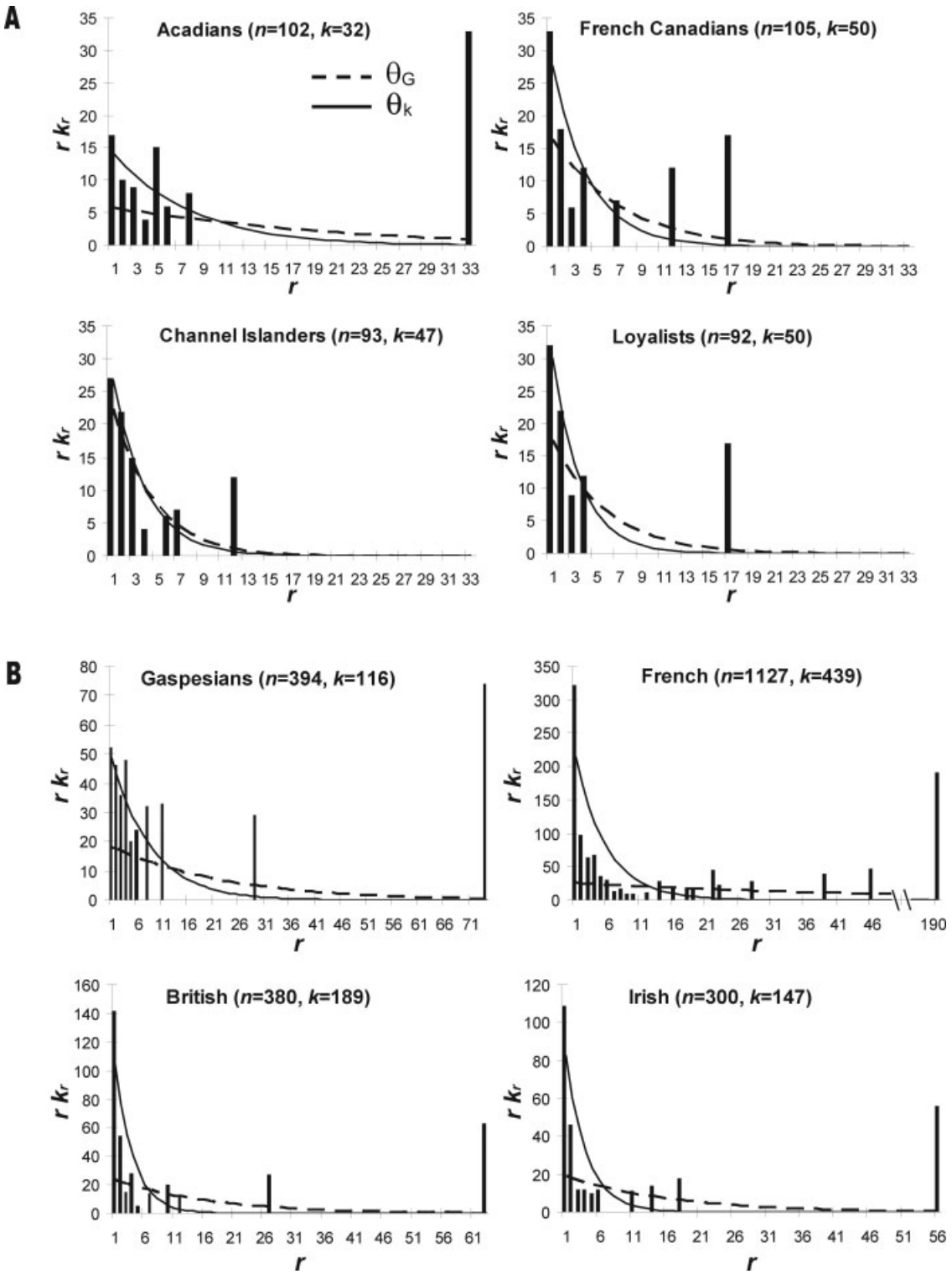

Fig. 1. Distribution of the frequency classes of mtDNA HVS1 haplotypes for the four Gaspesian populations (**A**) compared with three contemporary European samples and the composite/overall Gaspesian sample (**B**). Lines represent theoretical curves expected under mutation-drift equilibrium computed using Eq. (4) and θ_k (solid lines) and θ_G (dashed lines) estimates (Table 1).

similarly affected by the addition or subtraction of one haplotype, irrespective of whether it is a singleton or a highly frequent variant.

This rightmost outlier in all the plots in Figure 1, except for Channel Islanders, corresponds to the same HVSI haplotype of the Cambridge reference sequence, CRS (Anderson et al., 1981). This haplotype represents a heterogeneous pool of different mtDNA subclades (Loogvali et al., 2004) regrouping variants of the most prevalent European haplogroup H (Supporting Information

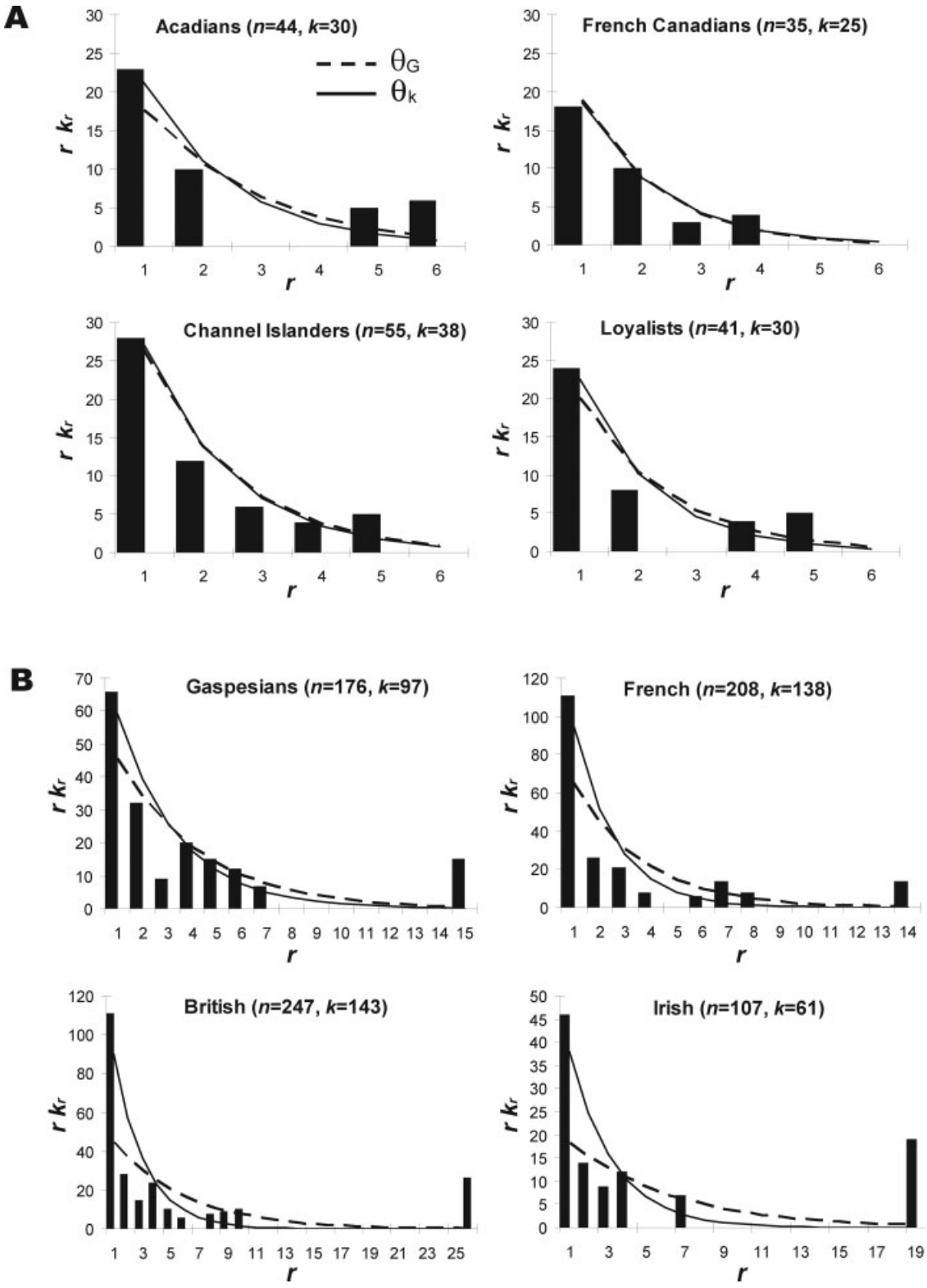

Fig. 2. Distribution of the frequency classes of 7 STR Y-chromosome haplotypes for the four Gaspesian populations (**A**) compared with three contemporary European samples and the composite/overall Gaspesian sample (**B**). Lines represent theoretical curves expected under mutation-drift equilibrium computed using Eq. (4) and θ_k (solid lines) and θ_G (dashed lines) estimates (Table 2).

Table S3). With the predominance of CRS HVS1, Gaspesians seem to reflect the haplotype configurations of the European source populations. Yet, we note important quantitative and qualitative differences in the underlying distributions (Fig. 1 and Supporting Information Table S4). In Acadians, the proportion of CRS is approxi-

TABLE 1. Summary statistics of mtDNA HVS1 haplotypes in Gaspesian and European populations

HVS1 (16,090–16,365)	*n*	*k*	*G*	θ_k	θ_G	*k* Expected[a]	*P* (Chakraborty)	*G* expected	*P* (Watterson)	*P* (Slatkin)
Acadians	102	32	0.88	16	6	18	10^{-4}	0.93	0.99	0.99
French Canadians	105	50	0.95	37	19	36	10^{-3}	0.96	1	1
Channel Islanders	93	47	0.97	37	29	42	0.16	0.96	0.89	0.78
Loyalists	92	50	0.96	44	21	36	10^{-3}	0.97	1	0.97
Gaspesians	394	116	0.95	55	19	59	$<10^{-5}$	0.98	1	0.99
French[b]	1,127	439	0.97	264	26	99	$<10^{-5}$	1	1	1
British[b]	380	189	0.96	149	25	70	$<10^{-5}$	0.99	1	1
Irish[b]	300	147	0.96	113	20	57	$<10^{-5}$	0.99	1	1

n, sample size; *k*, no of haplotypes; *G*, haplotype diversity.
[a] Based on θ_G.
[b] Data were drawn from (Piercy et al., 1993; Richards et al., 1996; Rousselet and Mangin, 1998; Helgason et al., 2001; Dubut et al., 2004; Richard et al., 2007; McEvoy et al., 2004).

mately double that observed in other populations. In Channel Islanders, the most frequent HVS1 haplotype is not CRS, but one that belongs to the Amerindian haplogroup C. The same haplotype of haplogroup C is found among French Canadians as second in frequency to the CRS haplotype. Thus beside CRS, the common Gaspesian and European haplotypes belong to different lineages. In British and Irish populations, the second in frequency is haplogroup J, in the French haplogroup K, and in Gaspesians C. The appearance of a haplotype of haplogroup C (in Acadians it occurs once and in four copies among Loyalists) along with less frequent haplotypes of haplogroups A and D (Supporting Information Table S1) is necessarily due to the Amerindian admixture (Lorenz and Smith, 1996; Smith et al., 1999; Malhi et al., 2001). Furthermore, comparing the data with the θ_k based theoretical distributions, suggests a marked excess of singleton haplotypes in European populations, which is not seen among Gaspesians. In other words, significant *P*-values for the tests reported in Table 1 for the European populations are due to the presence of outlier high frequency haplotypes on the right of the plot and to a surplus of singleton haplotypes seen on the left. In principle, because the estimates of θ take into account the sample size they also allow direct comparison with populations whose samples are bigger. On the other hand, Chakraborty et al. (1988) noted that when sampling heterogeneous populations-either due to recent population agglomeration or to artificial pooling of samples—the allele frequency spectrum of the pooled sample can substantially deviate from theoretical expectation even if each sampling unit exhibited variability in apparent agreement with expectations of mutation-drift equilibrium. This so called effect of population amalgamation may inflate the number of rare alleles when sample size increases, and in consequence affect θ_k. To correct for this effect, when comparing samples of different sizes, we can "adjust" the number of haplotypes *k* and the resulting θ_k to the lowest sample size according to Chakraborty (Chakraborty et al., 1988; Chakraborty, 1990). Here, even after adjustment, the "adjusted" number of haplotypes and the corresponding estimates of θ_k of European populations remain greater than in Gaspesians (Supporting Information Table S5). In Gaspesians, the disparity between θ_k-based theoretical curve and the data in the plots in Figure 1 appears "one-sided", only due to the presence of one or two unusually frequent haplotypes. When compared with their source populations in Europe, Gaspesian populations do not show excess of low frequency variants. In the same time, they carry more distinct mtDNA lineages, enriched with the additional Amerindian haplogroups A, C, and D. Thus, the population amalgamation that took place in the peopling of the Gaspe Peninsula, due not only to Amerindian but also to diverse European contributions, did not leave here a signature in the form of supernumerary singleton haplotypes. We note that among Gaspesians, Acadians appear least diverse with the lowest number of distinct haplotypes ($k = 32$) and the lowest haplotype diversity ($G = 0.88$) as also reflected in the two corresponding θ estimates.

Y-chromosome diversity

In 176 Gaspesians of the four investigated groups, Y-chromosome variation was analyzed for seven STRs and thirteen segregating sites of unique event polymorphisms, used to partition STR haplotypes within major Y-chromosome haplogroups (Materials and Methods, Supporting Information Table S6, see also Supporting Information Table S11). The data on the French, British and Irish samples were from Roewer et al. (2005) (Supporting Information Table S7). Gaspesian groups do not appreciably differ in their indices of genetic diversity (Table 2). In contrast to the analysis of matrilineal lineages, both θ estimates θ_G and θ_k are similar, the neutrality tests are consistent with mutation-drift equilibrium (Table 2) and, accordingly, theoretical distributions match the data in the plots of haplotype frequency classes (Fig. 2A). In this respect, Gaspesians differ from the Europeans (Fig. 2B), who show a slight excess of singleton haplotypes and the presence of an outlier high frequency class haplotype on the right. This high frequency haplotype (14-13-29-24-11-13-13) is identical in the three European groups (Supporting Information Table S6), as was the case for the common maternal CRS haplotype in Figure 1. The 14-13-29-24-11-13-13 haplotype is the most frequent in Channel Islanders and French Canadians and appears as an extra, high frequency peak ($r = 15$) in a plot of all Gaspesians (Fig. 2B). Interestingly, common Acadian haplotypes are either rare or nonexistent in our European samples; this is also the case for haplotypes of the haplogroup I among Channel Islanders and Loyalists. Acadians also carry a haplotype that belongs to non-European haplogroup C that is likely of Amerindian origin (Zegura et al., 2004). In contrast to differences in relative haplotype content, all major European haplogroups are represented and occur at similar frequencies in the European and overall Gaspesian samples (Supporting Information Table S8). Overall, the diversity measured by *k* is lower in Gaspesian populations than in Europeans, even after adjustment (Supporting Information Table S9). This difference disappears when all Gaspesians are pooled together.

TABLE 2. Summary statistics of Y-STR haplotypes in Gaspesian and European populations

Y-STR haplotypes[a]	n	k	G	θ_κ	θ_G	k Expected[b]	P (Chakraborty)	G expected	P (Watterson)	P (Slatkin)
Acadians	44	30	0.97	40	29	27	0.21	0.95	0.95	0.94
French Canadians	35	25	0.98	38	40	25	0.64	0.95	0.56	0.56
Channel Islanders	55	38	0.98	53	50	37	0.49	0.96	0.72	0.74
Loyalists	41	30	0.98	49	38	28	0.32	0.96	0.91	0.91
Gaspesians	176	97	0.98	88	61	83	0.01	0.98	0.98	0.99
French[c]	208	138	0.99	178	94	110	$<10^{-4}$	0.99	1	1
British[c]	247	143	0.98	141	54	93	$<10^{-5}$	0.99	1	1
Irish[c]	107	61	0.96	58	22	40	$<10^{-5}$	0.97	1	1

n, sample size; k, no of haplotypes; G, haplotype diversity.
[a] DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393.
[b] Based on θ_G.
[c] Data were drawn from (Roewer et al., 2005).

Relations between Gaspesian groups

Gaspesian F_{ST} of 2.0% for HVS1 haplotypes compare with an F_{ST} of 0.5% for the French (Richard et al., 2007) and an F_{ST} of 0 among Irish, British, and French. For patrilineal 7-STR Y-chromosome haplotypes, the F_{ST} in Gaspesians was 1.2%, again higher than in Western Europeans ($F_{ST} = 0.3\%$). Greater population differentiation in Gaspesians is essentially due to Acadians and Loyalists who show the highest pairwise F_{ST}s with their neighbors (Table 3—upper right for mtDNA and lower left for Y haplotypes). Given very similar sample sizes, genetic similarities among Gaspesian groups can be evaluated by comparing their haplotypes. Table 4 lists the proportions of the shared haplotypes (k) and of the shared chromosomes (n). The sharing is evaluated for the population on the left with those indicated on the top, whereas the percent of population-specific haplotypes/chromosomes can be found in the diagonal. For example, 65% of the Acadian HVS1&2 haplotypes are specific to this population. These haplotypes represent 43% of the Acadian mtDNA copies. At the same time, 31% of Acadian haplotypes, representing 55% of their mtDNA copies, are found in French Canadians. At the level of patrilineal lineages, Acadians share the greatest proportion of their Y-chromosomes with French Canadians (27%), even if they have more haplotypes in common with Channel Islanders than with French Canadians (7 versus 5, corresponding to 23% and 17%, respectively). In turn, French Canadians share most of their paternal and maternal lineages with Channel Islanders, and Channel Islanders with French Canadians, whereas Loyalists share the most with Channel Islanders. In all these comparisons, the haplotypes in common and the proportions of shared chromosomes point in the same direction. Acadians share the fewest with Loyalists and Loyalists with Acadians, whereas French Canadians the most with Channel Islanders and vice versa.

TABLE 3. Pairwise F_{st}'s in Gaspesians using HVS1 mtDNA (upper right) and 7 STRs Y-chromosome haplotypes (lower left)

Pairwise F_{st} (%)	Acadians	French Canadians	Channel Islanders	Loyalists
Acadians	0	2.7[a]	4.9[a]	2.1[a]
French Canadians	1.6[a]	0	0.1	0.6[b]
Channel Islanders	1.6[a]	0.1	0	1.3[a]
Loyalists	2.1[a]	0.9[b]	0.8[b]	0

[a] $P < 10^{-5}$.
[b] $P < 0.05$.

DISCUSSION

We used the infinitely-many-alleles model, Ewens' formalism (Ewens, 1972; Watterson, 1978; Chakraborty, 1990) and plots of haplotype frequency classes, to examine our data. Although mutation-drift equilibrium is not necessarily expected in contemporary human populations, the underlying standard population model provides a useful framework to identify changes in the haplotype frequency spectra. These changes can be due to selection, population growth or decline, founder effects, admixture or sampling of a heterogeneous population resulting from population amalgamation (Chakraborty et al., 1988; Chakraborty and Weiss, 1991). The significance of the departure from the mutation drift-equilibrium, seen in the difference between the observed G and k and the corresponding θ_k and θ_G estimates, is evaluated using the tests by Ewens-Watterson, Chakraborty and Slatkin (Materials and Methods). The data from Western Europe (Figs. 1B and 2B) clearly deviate from this simple model and the difference is significant (Tables 1 and 2). This is also true for maternal lineages in Acadians, Loyalists, and French Canadians (Fig. 1A) where, however, the deviation is only seen on the right of the plot and no excess of the rare haplotypes is observed when the theoretical curve is modeled using θ_k.

Founder effects are likely to modify allelic frequencies in the colonies. A reduction in population size, implicit in a founder effect (Mayr, 1963; Chakraborty and Nei, 1977; Allendorf, 1986; Clegg et al., 2002) alters the number of alleles/haplotypes more profoundly than the overall heterozygosity and this should be reflected in the corresponding estimates of θ_k and θ_G. However, when comparing Gaspesians and Europeans an important decrease in G and θ_G of maternal lineages is only seen in Acadians. In Figure 3, we compared the HVS1 data of Gaspesian groups with the regions of France (Richard et al., 2007). We used four statistics: G, both estimates of θ and the ratio of k/n to partially correct for differences in sample size. Here again, only Acadians consistently appear below the cluster of values representing different French regions. At the level of paternal lines, Acadians and French Canadians exhibit the lowest diversity (Table 2). This tendency is even more pronounced when full information on parental lines is considered to estimate the corresponding θ's, that is, using 11 Y-chromosome STRs and the extended mtDNA HVS1&2 haplotypes (Supp. Info. Table S1 and data not shown).

The lower genetic diversity of Acadians, particularly on the maternal side, can be due to the relatively low number of first Acadian settlers when compared with

TABLE 4. Sharing of HVS1&2 mtDNA and 7 STRs Y-chromosome haplotypes between population pairs

	mtDNA HVS1&2 (16,069–16,383; 58–370)					7 STRs Y-chromosome haplotypes				
	k	Acadians	French Canadians	Channel Islanders	Loyalists	*k*	Acadians	French Canadians	Channel Islanders	Loyalists
Acadians	49	65	31	16	10	30	67	17	23	7
French Canadians	62	24	52	32	15	25	20	56	32	12
Channel Islanders	57	14	35	51	23	38	18	21	55	18
Loyalists	60	8	15	22	72	30	7	10	23	70
	n					*n*				
Acadians	99	43	55	31	25	44	50	27	23	16
French Canadians	104	37	39	50	28	35	26	49	40	20
Channel Islanders	92	25	49	38	38	55	24	35	40	27
Loyalists	91	14	20	32	64	41	12	17	24	68

Proportions (%) of haplotypes shared between the populations on the left with these indicated on the top were evaluated based on the number of distinct haplotypes k and on the number of copies n. Diagonal provides percentages (underlined) of population specific haplotypes.

other Gaspesians and the fact that they already included extended families (Desjardins et al., 1999). Moreover, judging from F_{ST} and haplotype sharing, Acadians exchanged less with their neighbors than other groups (Tables 3 and 4). Their immediate neighbors on the southern side of the peninsula were Loyalists, who spoke a different language, practiced a different religion and were of opposite historical allegiance to the British Crown. Other Gaspesians do not show similar signs of a demographic bottleneck. In a compilation of θ_k and G estimates for HVS1, in a broader collection of European populations by Helgason (Helgason et al., 2000), Gaspesian groups, keeping Acadians apart, fall very close to Portuguese ($\theta_k = 50$), Swedes (40), Bulgarians (36), Welsh (34), or Finns (48). This may suggest that peripheral populations tend to have lower diversity than those geographically placed at the crossroads of population movements or those composed of distinct regional groups.

It has been estimated that a founding population of less than one hundred individuals and/or a long subsequent isolation would be required to see a marked decrease in genetic diversity of the population isolate (Clegg et al., 2002). However, successive colonizations could also erase the consequences of an initial demographic bottleneck. As compared with Acadians, Loyalist migrants were more numerous and of greater diversity, although this difference may be subtle, because they also included families (Matthews and Gentilcore, 1993). French Canadians and Channel Islanders arrived later and their immigration was extended over time (Desjardins et al., 1999). It also involved more settlers and more contacts between the newcomers. For example, Roman Catholic Irish and Scottish settlers were easily accepted among French Canadians, and male fishermen from Jersey and Guernsey often married French Canadian women, which partly explains the greater genetic similarities between French Canadians and Channel Islanders when compared with Acadians and Loyalists. The prosperity derived from the development of the fishing industry greatly benefited the northern and eastern part of the peninsula with the port of Gaspe in the Gulf of the St. Lawrence, creating conditions conducive to demographic growth as well as gene flow. In contrast, the original Acadian and Loyalist settlements were on the southern side of the Gaspesian Peninsula. This could have caused a relative isolation of these earliest groups of founders, contributing at the same time to the preservation of their traditions and identity. We can envisage that these traditions and group identity were preferably

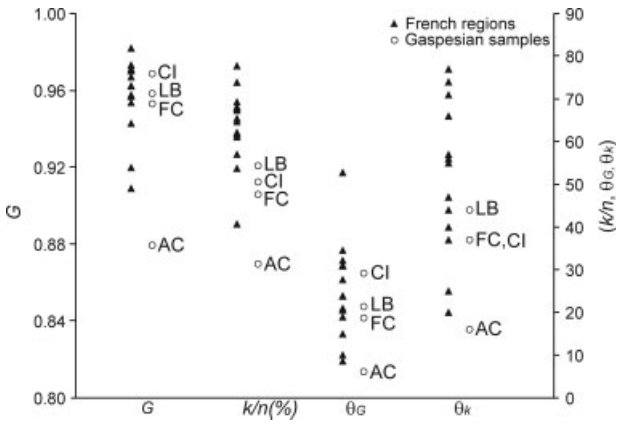

Fig. 3. HVS1 diversity indices of Gaspesian groups (AC, Acadians; LB, Loyalists; CI, Channel Islanders; FC, French Canadians) compared with 14 regional samples from France (Richard et al., 2007). The scale for haplotype diversity G is on the left, that for the number of haplotypes k, normalized by the sample size n (times 100), and the estimates of θ on the right.

preserved among those who remained at initial settlement sites. If so, in our sampling scheme based on self-identification, this would reinforce the isolation of these groups with respect to younger settlers, such as Channel Islanders and French Canadians in the East and Northeast of the peninsula. It could also have affected the haplotype sharing between the groups as well as the pairwise F_{st}'s (Tables 3 and 4). Fuller analysis and interpretation of these data would however require additional knowledge about what is due to genetic differences (or variant sharing) between these groups' founders and what to subsequent drift and gene flow.

The composition of the population sample and the way participants were recruited are also important issues and different sampling schemes would affect the results. For example, Europeans in Figures 1 and 2 represent collective samples of France, Great Britain, and Ireland, in which the contribution of their different localities is usually unknown and most certainly nonproportional. As mentioned earlier, amalgamation of local groups into a single population may artificially inflate the number of rare variants (Chakraborty et al., 1988). This can indeed be observed with the "artificial" sample of Acadians, Channel Islanders, French Canadians, and Loyalists pooled together (Figs. 1 and 2; Tables 1 and 2), but the

effect is relatively modest. In contrast, the results we obtained for the regional French groups of Richard et al. (2007) (Fig. 3 and data not shown) are qualitatively similar to the collective French sample, which is marked by excess of singletons as well as the presence of high frequency haplotypes. In Europe, these high frequency "founder" haplotypes can be ascribed to population bottlenecks and subsequent expansions in the history of these populations, from Upper Paleolithic, through the population refugia during last glacial maximum to Neolithic revolution (Richards et al., 1996; Richards et al., 1998; Dubut et al., 2004). In turn, an excess of singleton haplotypes, which is seen also in regional population samples from France, suggests star-like phylogeny, consistent with demographic growth, independently documented in paleontological and archeological records (Livi-Bacci, 2001). Relative loss of the rare-haplotypes category in Gaspesia can be explained by lower probability of singleton variants to be carried away by migrant population. If sampled such variants are either lost or would have a tendency to become relatively common. Furthermore, in spite of the rapid population growth in Nouvelle France, the time elapsed since colonization has been too short to permit accumulation of new singleton haplotypes by mutation. Thus, apart for the presence of frequent variants that have high probability to remain similarly frequent in the colonies, such as CRS HVS1, the population of migrants appears to fit better the mutation-drift equilibrium expectation than its parental populations (see Fig. 2).

Observation of the apparent loss of rare haplotypes is interesting, because in human/medical genetics, the founder effect is usually associated with the appearance of a hereditary disorder (or a particular underlying mutation) that is nonexistent or very rare elsewhere. This implies a frequency increase of a particular deleterious allele which (considering the very low probability of new mutations) must have been very rare in the source population (e.g., Labuda et al., 1996; Labuda et al., 1997; Carter et al., 1998; Kere, 2001; Scriver, 2001; Yotova et al., 2005). However, there is no conflict between these two observations. Increase in the frequency of a rare deleterious allele is expected to occur very rarely and such events are only picked up because of the associated clinical phenotypes. In agreement with the overall tendency, some hereditary diseases become rarer or disappear from a founder population (Kere, 2001; Scriver, 2001; Laberge et al., 2005).

Indeed, our study provides many examples of the modification of the haplotype frequency spectra. Some haplotypes that are rare elsewhere occur in Gaspesians at relatively high or intermediate frequencies (Supporting Information Tables S1, S3, and S4). For example, this is the case for the two most frequent Y-chromosome haplotypes in Loyalists. In Channel Islanders, the most frequent mtDNA HVS1&2 lineage belongs to the Amerindian haplogroup C, thus replacing the CRS haplotype that is the most frequent haplotype in virtually all populations of European origin. This is something new. So far the presence of the CRS haplotype as an outstanding outlier in the plots of haplotype frequency classes in Gaspesians could be seen as a reflection of the European distribution. In Channel Islanders, we observe the establishment of a new lineage that similarly dominates the distribution (see Fig. 1). Its high frequency in Channel Islanders but also in French Canadians strongly argues that this lineage was introduced at the beginning of the colony. In principle, any of the mtDNA lineages present among the early founders could have undergone such a founder effect, but the prevalence of males among first immigrants certainly favored a matrilineal Native admixture. On the other hand, the presence of three distinct mtDNA haplotypes from haplogroups A and D indicate that this was not a single event and the finding of Y-chromosome haplogroup C in Acadians suggest a Native admixture from the male side as well.

Taken together, Gaspesians show different signatures of founder effects. Generally, a decrease in genetic diversity was seen as a loss of unique, singleton haplotypes. This does not affect overall heterozygosity, whose clear decrease was only seen in Acadian maternal lineages. In all groups, we observed a modification of frequencies of certain common or intermediate variants; some of these variants were new, acquired by admixture thus enriching genetic diversity. As a result of the demographic processes related to the colonization of new territories, such as the initial sampling of migrants and subsequent population growth in isolation or including additional immigration and exchange between neighboring groups, Gaspesian groups appear genetically further apart than populations in Europe. The early settlers, Acadians and Loyalists, exhibit a stronger identity with respect to the French Canadian and Channel Islander latecomers. Our findings support earlier work on the regionalization of the Quebec genetic pool, based on genealogical analyses (Gagnon and Heyer, 2001; Vezina et al., 2005) and on the distribution of genetic diseases (Laberge et al., 2005; Scriver, 2001). They also improve our understanding of the mechanisms and consequences of founder effects in human populations (Labuda et al., 1996; Labuda et al., 1997; Yotova et al., 2005).

ACKNOWLEDGMENTS

The authors are grateful to Jules Bélanger, Marthe Bourdages, Louise Cyr, Hélène Cyr, Helena Fournier, George-Edison Langlois, Jean-Claude Lebreux, Diane Sawyer, and Gérard-Raymond Thériault for their assistance in sample collection and to all the Gaspesians who agreed to participate in this study. They also thank Jean-Francois Lefebvre, Cathy Cauvier, and Antonio Torroni for their help at different stages of this study. The reviewers and the associate editor provided insightful comments helping to improve the earlier version of this article.

LITERATURE CITED

Allendorf FW. 1986. Genetic drift and the loss of alleles versus heterozygosity. Zoo Biol 5:181–190.

Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, Young IG. 1981. Sequence and organization of the human mitochondrial genome. Nature 290:457–465.

Bandelt HJ, Quintana-Murci L, Salas A, Macaulay V. 2002. The fingerprint of phantom mutations in mitochondrial DNA data. Am J Hum Genet 71:1150–1160.

Bouchard G, De Braekeleer M. 1991. Histoire d'un Génome. Population et Génétique dans l'Est du Québec. Presses de l'Université du Québec, Québec City.

Bourgeois S, Labuda D. 2004. Dynamic allele-specific oligonucleotide hybridization on solid support. Anal Biochem 324: 309–311.

Carter KC, Byck S, Waters PJ, Richards B, Nowacki PM, Laframboise R, Lambert M, Treacy E, Scriver CR. 1998.

Mutation at the phenylalanine hydroxylase gene (PAH) and its use to document population genetic variation: the Quebec experience. Eur J Hum Genet 6:61–70.

Chakraborty R. 1990. Mitochondrial DNA polymorphism reveals hidden heterogeneity within some Asian populations. Am J Hum Genet 47:87–94.

Chakraborty R, Nei M. 1977. Bottleneck effects on average heterozygosity and genetic distance with the stepwise mutation model. Evolution 31:347–356.

Chakraborty R, Smouse PE, Neel JV. 1988. Population amalgamation and genetic variation: observations on artificially agglomerated tribal populations of Central and South America. Am J Hum Genet 43:709–725.

Chakraborty R, Weiss KM. 1991. Genetic variation of the mitochondrial DNA genome in American Indians is at mutation-drift equilibrium. Am J Phys Anthropol 86:497–506.

Charbonneau H, Desjardins B, Légaré J, Denis H. 2000. The population of the St. Lawrence Valley, 1608–1760. In: A population history of North America. Cambridge, UK: Cambridge University Press.

Clegg SM, Degnan SM, Kikkawa J, Moritz C, Estoup A, Owens IP. 2002. Genetic consequences of sequential founder events by an island-colonizing bird. Proc Natl Acad Sci USA 99: 8127–8132.

de Knijff P, Kayser M, Caglia A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, Herrmann S, Herzog B, Hidding M, Honda K, Jobling M, Krawczak M, Leim K, Meuser E, Oesterreich W, Pandya A, Parson W, Penacino G, Perez-Lezan A, Piccinini A, Prinz M, Roewer L, et al. 1997. Chromosome Y microsatellites: population genetic and evolutionary aspects. Int J Legal Med 110:134–149.

Desjardins M, Frenette Y, Bélanger J, Hétu B. 1999. Histoire de la Gaspésie. IQRC, editor. Québec: Les presses de l'Université Laval, Quebec City.

Dubut V, Chollet L, Murail P, Cartault F, Beraud-Colomb E, Serre M, Mogentale-Profizi N. 2004. mtDNA polymorphisms in five French groups: importance of regional sampling. Eur J Hum Genet 12:293–300.

Estoup A, Jarne P, Cornuet JM. 2002. Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. Mol Ecol 11:1591–1604.

Ewens WJ. 1972. The sampling theory of selectively neutral alleles. Theor Popul Biol 3:87–112.

Excoffier L, Laval G, Schneider S. 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50.

Gagnon A, Heyer E. 2001. Fragmentation of the Quebec population genetic pool (Canada): evidence from the genetic contribution of founders per region in the 17th and 18th centuries. Am J Phys Anthropol 114:30–41.

Helgason A, Hickey E, Goodacre S, Bosnes V, Stefansson K, Ward R, Sykes B. 2001. mtDNA and the islands of the North Atlantic: estimating the proportions of Norse and Gaelic ancestry. Am J Hum Genet 68:723–737.

Helgason A, Hrafnkelsson B, Gulcher JR, Ward R, Stefansson K. 2003. A populationwide coalescent analysis of Icelandic matrilineal and patrilineal genealogies: evidence for a faster evolutionary rate of mtDNA lineages than Y chromosomes. Am J Hum Genet 72:1370–1388.

Helgason A, Sigurdardottir S, Gulcher JR, Ward R, Stefansson K. 2000. mtDNA and the origin of the Icelanders: deciphering signals of recent population history. Am J Hum Genet 66:999–1016.

Henripin J, Peron Y. 1973. ≪La transition démographique de la province de Québec≫ dans in La population du Québec: études rétrospectives. Montréal: Boréal Express. p 23–44.

Heyer E, Zietkiewicz E, Rochowski A, Yotova V, Puymirat J, Labuda D. 2001. Phylogenetic and familial estimates of mitochondrial substitution rates: study of control region mutations in deep-rooting pedigrees. Am J Hum Genet 69:1113–1126.

Horai S, Kondo R, Nakagawa-Hattori Y, Hayashi S, Sonoda S, Tajima K. 1993. Peopling of the Americas, founded by four major lineages of mitochondrial DNA. Mol Biol Evol 10: 23–47.

Kayser M, Caglia A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, Herrmann S, Herzog B, Hidding M, Honda K, Jobling M, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, Perez-Lezan A, Piccinini A, Prinz M, Schmitt C, Roewer L. 1997. Evaluation of Y-chromosomal STRs: a multicenter study. Int J Legal Med 110:125–133,141–129.

Kere J. 2001. Human population genetics: lessons from Finland. Annu Rev Genomics Hum Genet 2:103–128.

Kimura M, Crow JF. 1964. The number of alleles that can be maintained in a finite population. Genetics 49:725–738.

Laberge AM, Michaud J, Richter A, Lemyre E, Lambert M, Brais B, Mitchell GA. 2005. Population history and its impact on medical genetics in Quebec. Clin Genet 68:287–301.

Labuda D, Zietkiewicz E, Labuda M. 1997. The genetic clock and the age of the founder effect in growing populations: a lesson from French Canadians and Ashkenazim. Am J Hum Genet 61:768–771.

Labuda M, Labuda D, Korab-Laskowska M, Cole DE, Zietkiewicz E, Weissenbach J, Popowska E, Pronicka E, Root AW, Glorieux FH. 1996. Linkage disequilibrium analysis in young populations: pseudo-vitamin D-deficiency rickets and the founder effect in French Canadians. Am J Hum Genet 59: 633–643.

Livi-Bacci M. 2001. A concise history of world population, 3rd ed. Malden, MA and Oxford, UK: Blakwell Publishers.

Lorenz JG, Smith DG. 1996. Distribution of four founding mtDNA haplogroups among native North Americans. Am J Phys Anthropol 101:307–323.

Loogvali EL, Roostalu U, Malyarchuk BA, Derenko MV, Kivisild T, Metspalu E, Tambets K, Reidla M, Tolk HV, Parik J, Pennarun E, Laos S, Lunkina A, Golubenko M, Barac L, Pericic M, Balanovsky OP, Gusar V, Khusnutdinova EK, Stepanov V, Puzyrev V, Rudan P, Balanovska EV, Grechanina E, Richard C, Moisan JP, Chaventre A, Anagnou NP, Pappa KI, Michalodimitrakis EN, Claustres M, Golge M, Mikerezi I, Usanga E, Villems R. 2004. Disuniting uniformity: a pied cladistic canvas of mtDNA haplogroup H in Eurasia. Mol Biol Evol 21:2012–2021.

Malhi RS, Schultz BA, Smith DG. 2001. Distribution of mitochondrial DNA lineages among Native American tribes of northeastern North America. Hum Biol 73:17–55.

Matthews G, Gentilcore R. 1993. Historical atlas of Canada volume 2: the land transformed, 1800–1891. Toronto: University of Toronto Press.

Mayr E. 1963. Animal species and evolution. Cambridge, MA: Harvard University Press.

Piercy R, Sullivan KM, Benson N, Gill P. 1993. The application of mitochondrial DNA typing to the study of white Caucasian genetic identification. Int J Legal Med 106:85–90.

McEvoy B, Richards M, Forster P, Bradley DG. 2004. The Longue Duree of genetic ancestry: multiple genetic marker systems and Celtic origins on the Atlantic facade of Europe. Am J Hum Genet 75:693–702.

Pereira L, Prata MJ, Amorim A. 2003. An evaluation of the proportion of identical Y-STR haplotypes due to recurrent mutation. Int Congr Ser 1239:57–60.

Price AL, Butler J, Patterson N, Capelli C, Pascali VL, Scarnicci F, Ruiz-Linares A, Groop L, Saetta AA, Korkolopoulou P, Seligsohn U, Waliszewska A, Schirmer C, Ardlie K, Ramos A, Nemesh J, Arbeitman L, Goldstein DB, Reich D, Hirschhorn JN. 2008. Discerning the ancestry of European Americans in genetic association studies. PLoS Genet 4:e236.

Price AL, Patterson N, Yu F, Cox DR, Waliszewska A, McDonald GJ, Tandon A, Schirmer C, Neubauer J, Bedoya G, Duque C, Villegas A, Bortolini MC, Salzano FM, Gallo C, Mazzotti G, Tello-Ruiz M, Riba L, Aguilar-Salinas CA, Canizales-Quinteros S, Menjivar M, Klitz W, Henderson B, Haiman CA, Winkler C, Tusie-Luna T, Ruiz-Linares A, Reich D. 2007. A genomewide admixture map for Latino populations. Am J Hum Genet 80:1024–1036.

Richard C, Pennarun E, Kivisild T, Tambets K, Tolk HV, Metspalu E, Reidla M, Chevalier S, Giraudet S, Lauc LB, Pericic M, Rudan P, Claustres M, Journel H, Dorval I, Muller C,

Villems R, Chaventre A, Moisan JP. 2007. An mtDNA perspective of French genetic variation. Ann Hum Biol 34:68–79.
Richards M, Corte-Real H, Forster P, Macaulay V, Wilkinson-Herbots H, Demaine A, Papiha S, Hedges R, Bandelt HJ, Sykes B. 1996. Paleolithic and neolithic lineages in the European mitochondrial gene pool. Am J Hum Genet 59:185–203.
Richards MB, Macaulay VA, Bandelt HJ, Sykes BC. 1998. Phylogeography of mitochondrial DNA in Western Europe. Ann Hum Genet 62 (Part 3):241–260.
Roewer L, Croucher PJ, Willuweit S, Lu TT, Kayser M, Lessig R, de Knijff P, Jobling MA, Tyler-Smith C, Krawczak M. 2005. Signature of recent historical events in the European Y-chromosomal STR haplotype distribution. Hum Genet 116:279–291.
Rosser ZH, Zerjal T, Hurles ME, Adojaan M, Alavantic D, Amorim A, Amos W, Armenteros M, Arroyo E, Barbujani G, et al. 2000. Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. Am J Hum Genet 67:1526–1543.
Rousselet F, Mangin P. 1998. Mitochondrial DNA polymorphisms: a study of 50 French Caucasian individuals and application to forensic casework. Int J Legal Med 111:292–298.
Scriver CR. 2001. Human genetics: lessons from Quebec populations. Annu Rev Genomics Hum Genet 2:69–101.
Shriver MD, Kittles RA. 2004. Genetic ancestry and the search for personalized genetic histories. Nat Rev Genet 5:611–618.
Slatkin M. 1996. A correction to the exact test based on the Ewens sampling distribution. Genet Res 68:259–260.
Smith DG, Malhi RS, Eshleman J, Lorenz JG, Kaestle FA. 1999. Distribution of mtDNA haplogroup X among Native North Americans. Am J Phys Anthropol 110:271–284.
Tian C, Hinds DA, Shigeta R, Kittles R, Ballinger DG, Seldin MF. 2006. A genomewide single-nucleotide-polymorphism panel with high ancestry information for African American admixture mapping. Am J Hum Genet 79:640–649.
Torroni A, Huoponen K, Francalacci P, Petrozzi M, Morelli L, Scozzari R, Obinu D, Savontaus ML, Wallace DC. 1996. Classification of European mtDNAs from an analysis of three European populations. Genetics 144:1835–1850.
Tremblay M, Vézina H, Houde L. 2003. Les probabilités de transmission des gènes fondateurs dans cinq populations régionales du Québec. Population 58:403–423.
Vezina H, Durocher F, Dumont M, Houde L, Szabo C, Tranchant M, Chiquette J, Plante M, Laframboise R, Lepine J, Nevanlinna H, Stoppa-Lyonnet D, Goldgar D, Bridge P, Simard J. 2005. Molecular and genealogical characterization of the R1443X BRCA1 mutation in high-risk French-Canadian breast/ovarian cancer families. Hum Genet 117:119–132.
Wang S, Ray N, Rojas W, Parra MV, Bedoya G, Gallo C, Poletti G, Mazzotti G, Hill K, Hurtado AM, Camrena B, Nicolini H, Klitz W, Barrantes R, Molina JA, Freimer NB, Bortolini MC, Salzano FM, Petzl-Erler ML, Tsuneto LT, Dipierri JE, Alfaro EL, Bailliet G, Bianchi NO, Llop E, Rothhammer F, Excoffier L, Ruiz-Linares A. 2008. Geographic patterns of genome admixture in Latin American Mestizos. PLoS Genet 4:e1000037.
Watterson GA. 1978. The homozygosity test of neutrality. Genetics 88:405–417.
Weale ME, Weiss DA, Jager RF, Bradman N, Thomas MG. 2002. Y chromosome evidence for Anglo-Saxon mass migration. Mol Biol Evol 19:1008–1021.
Yotova V, Labuda D, Zietkiewicz E, Gehl D, Lovell A, Lefebvre JF, Bourgeois S, Lemieux-Blanchard E, Labuda M, Vezina H, Houde L, Tremblay M, Toupance B, Heyer E, Hudson TJ, Laberge C. 2005. Anatomy of a founder effect: myotonic dystrophy in Northeastern Quebec. Hum Genet 117:177–187.
Yotova V, Lefebvre JF, Kohany O, Jurka J, Michalski R, Modiano D, Utermann G, Williams SM, Labuda D. 2007. Tracing genetic history of modern humans using X-chromosome lineages. Hum Genet 122:431–443.
Zegura SL, Karafet TM, Zhivotovsky LA, Hammer MF. 2004. High-resolution SNPs and microsatellite haplotypes point to a single, recent entry of Native American Y chromosomes into the Americas. Mol Biol Evol 21:164–175.