Social transmission of reproductive behavior increases frequency of inherited disorders in a young-expanding population

FRÉDÉRIC AUSTERLITZ* AND EVELYNE HEYER[†][‡]

*Laboratoire Evolution et Systématique, Université Paris-Sud, 91405 Orsay, France; and [†]Centre National de la Recherche Scientifique, Laboratoire d'Anthropologie Biologique, Musée de l'Homme, 75016 Paris, France

ABSTRACT The observation of high frequencies of certain inherited disorders in the population of Saguenay–Lac Saint Jean can be explained in terms of the variance and the correlation of effective family size (EFS) from one generation to the next. We have shown this effect by using the branching process approach with real demographic data. When variance of EFS is included in the model, despite its profound effect on mutant allele frequency, any mutant introduced in the population never reaches the known carrier frequencies (between 0.035 and 0.05). It is only when the EFS correlation between generations is introduced into the model that we can explain the rise of the mutant alleles. This correlation is described by a c parameter that reflects the dependency of children's EFS on their parents' EFS. The c parameter can be considered to reflect social transmission of demographic behavior. We show that such social transmission dramatically reduces the effective population size. This could explain particular distributions in allele frequencies and unusually high frequency of certain inherited disorders in some human populations.

In this paper, we will show that a high variance of effective family size (EFS) cannot alone account for high frequencies of inherited disorders in certain human populations but that sociodemographic phenomena should be included in the mechanism leading to high frequencies of rare mutations. More specifically, we will focus on intergeneration correlation of effective family size.

By using a branching process and a coalescent approach, Thompson and Neel (1–3) studied the impact of family size on gene frequencies. Assuming a geometric distribution of a number of offspring per family, they showed that a single copy allele introduced in a rapidly expanding population can reach high frequencies over a short period of time.

Regarding the intergeneration correlation in family size, this phenomenon had already been brought up by Fisher (4). Huestis and Maxwell (5) gave a correlation of 0.12 for a population of the USA. Nei and Murata (6) indicated that the mother–daughter correlation of sibship size is in the range 0.1–0.2 in several human populations, leading to a rather strong reduction in the effective population size. Other studies (7) indicated that this correlation reflected mostly cultural phenomena and that there was no clear evidence of any genetic component in family size. From a genetic point of view, what matters is not family size but effective family size, i.e., the number of children that reproduce in the population per reproducing individual. None of these previous studies calculated correlation of EFS.

This correlation can be computed in the French-Canadian population of Quebec, and, therefore, this population provides a good framework for the study of the impact of such social transmission on the evolution of the gene pool. Founded in the 17th century by up to 5,000 immigrants, it remained almost isolated, and, after 1765, it grew rapidly to >5 million today. We focus our work on the Saguenay–Lac Saint Jean region (SLSJ), a region in northeastern Quebec where carriers of certain rare recessive inherited disorders can be found at frequency of almost 5%. From molecular data and genealogical analyses, it is likely that each mutant allele was introduced by no more than one individual in the population (8). The demographic history of the SLSJ population is well documented in a number of studies (9) of the Interuniversity Institute for Population Research. Strong variation in effective family size (10) and the correlation of the average effective family size from one generation to the next already have been documented (11).

By using a branching process method, we evaluate the impact of the variance and intergeneration correlations in the EFS on the fate of an allele introduced by a single founder in the population. We use a geometric distribution that fits, better than the classically used Poisson distribution (12, 13), the real distribution of EFS, and we include a factor c for the intergeneration correlation in EFS. The results of this model then are compared with the known allele frequencies of some inherited disorders in the population and to the results of previous genealogical analysis (14). Then, by using gene dropping along the known genealogical paths, we verify the validity of the model.

MATERIALS AND METHODS

Demographic Data. The population of SLSJ numbers ≈300,000 inhabitants. They descend essentially from early founders among the first 5,000 settlers of Quebec at the end of the 17th century (see Fig. 1) The SLSJ region was opened for settlement only in the 19th century, colonized mainly by immigrants from Charlevoix and the eastern part of Quebec. The population is defined to include contemporary SLSJ individuals and all of their ancestors back to 17th-century founders; thus, our definition is not closely geographic, including residents of other parts of Quebec before the opening of SLSJ for colonization (Fig. 1). This yields a geometric population growth rate per generation of 1.41.

The EFS is the number of children that reproduce in the population per reproducing woman. The distribution of married children per married woman is known for this population for the 19th century (10). It gives a mean number of 3.4 married children per married woman and so a net reproduction rate of 1.7. This value differs from the growth rate of 1.41 because the
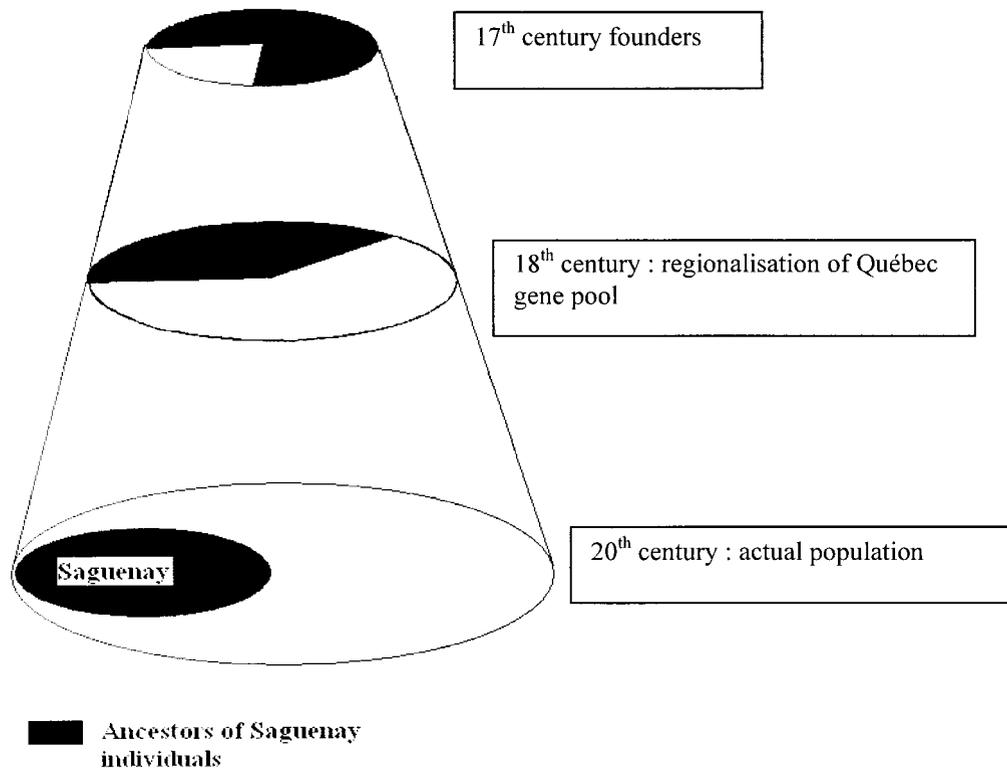
15140

www.manaraa.com

FIG. 1. Schematic representation of the "SLSJ population." The SLSJ population is defined as all individuals who live or lived in the SLSJ region and all of their ancestors back to the 17th century.

reproduction period of a woman overlaps more than a generation. With the same demographic data set ($n = 10,589$ individuals), we calculate the EFS distribution and correlation in EFS from one generation to the next. We used all of the pairs (EFS of an individual, EFS of his parents) to compute this correlation.

Genetic Data. The high incidence of autosomal recessive disorders specific to SLSJ reflects their relatively high carrier frequency as illustrated in Table 1. We will only focus on the four most frequent disorders with a carrier frequency between 0.035 and 0.050. Because they were introduced as a single copy, their carrier frequency increased from the putative 1/5,000 among the first Quebec settlers to ≈1/25 in the SLSJ region within 12 generations.

Genealogical Data. We used 891 individuals born around 1930 whose data we found in the Interuniversity Institute for Population Research ascending genealogies database. These individuals are presumed Alzheimer cases documented within the IMAGE project by Algène Biotechnologies (Montreal). These 891 individuals trace back to 2,631 founders who settled in Nouvelle, France before 1700 (15). The

genetic contributions of the 17th-century founders to contemporary individuals are similar whether these founders are carriers of an inherited disorder or not (14). We did the same calculation for these 891 individuals, yielding the same results. Therefore, the fact that these 891 individuals have been recruited as Alzheimer patients should not bias our results, and they are taken as representative of the contemporary SLSJ population.

Demographic Simulations. By using a branching process method, we simulated the genealogies of the population. The EFS of each couple was drawn from a distribution, with mean m given by (i) the growth rate λ of the population, (ii) the mean between the EFS of the wife's parents and the husband's parents, and (iii) a given level r of intergeneration correlation in EFS. For a given mean m, this distribution, denoted D(m), is Poisson or geometric.

We let N_t, W_t, and M_t denote, respectively, the total number of individuals, of women, and of men in the population at time t ($N_t = W_t + M_t$). At $t = 0$, initial values for EFS were drawn from the distribution D(2λ). At each generation t, couples were formed at random (i.e., we assumed no assortative

Table 1. Common autosomal recessive disorders specific to the Saguenay

Disorder	Carrier frequency in SLSJ	Carrier frequency elsewhere	Frequency of most common mutations in SLSJ
Spastic ataxia	1/21	Unknown form	
Tyrosinemia I	1/22	1/165 Norway, Sweden	96%
Sensorimotor polyneuropathy	1/23	Unknown form	100% (from haplotypes)
Pseudovitamin D deficient rickets	1/26	Very rare	100%
Cytochrome C oxidase deficiency	1/31	Very rare	
Cystinosis	1/39	1/200	
Histidinemia histidase deficiency	1/32	1/50 to 1/150	
Lipoprotein lipase deficiency	1/43	Very rare	100%
Pyruvate kinase deficiency	1/64	Rare	90%

Source: Conseil génétique–Complexe Hospitalier de la Sagamie (Chicoutimi Hospital).

mating), and each couple's EFS was drawn according to the following distribution:

$$D\left\{\alpha_t\left[(1-c)\lambda + c\left(\frac{n_{w_i} + n_{m_i}}{2}\right)\right]\right\},$$

where n_{w_i} and n_{m_i} denote, respectively, the EFS of the parents of a woman, w_i ($1 \le i \le W_t$) and EFS of the parents of her husband, m_i.

The parameter c indicates the strength of the intergeneration dependency. For $c = 0$, there is no dependency: the EFSs of all couples in a given generation are drawn according to the same distribution. The higher the value of c, the more the EFS of a couple depends on the EFS of their parents, and their reproductive behavior is closer to that of their parents. α_t is a normalization parameter whose value is calculated numerically at each generation so that the total growth of the population remains equal to λ.

$$\alpha_t = \frac{\lambda N_t}{\sum_{i=1}^{W_t}\left[(1-c)\lambda + c\left(\frac{n_{w_i} + n_{m_i}}{2}\right)\right]}$$

We obtained empirically from simulations the value of r that corresponds to a given value of c. The relation between the values of c and r depends on the distribution of EFS (D). The parameters are as follows: initial population of 5,000 individuals (2,500 women and 2,500 men) and a growth rate λ of 1.41, during 12 generations. We used either a Poisson or a geometric distribution of EFS and various levels of intergeneration correlation. For each set of parameters, we performed 100 simulations that yielded complete simulated genealogies in the population.

Genealogical Simulations. Either on the simulated genealogies or on the real ascending genealogies of the 891 individuals, we simulated transmission of alleles along the genealogical paths. For each individual in the genealogies, we chose at random one of the two alleles carried by his or her father and one of the two carried by his or her mother. This Mendelian segregation was performed starting from the founders and proceeding along the genealogical paths. We attributed two unique alleles to each founder. This process was carried out 1,000 times for every simulated genealogy and 50,000 times for the real ones. All alleles were assumed to be neutral, but research has shown (8) that lethality has very little influence on changes in allele frequencies because the appearance of homozygotes for an allele introduced as a single copy remains a rare event.

For each founder, we obtained the probability that one of his or her alleles would reach a frequency in the range of frequencies (0.035–0.050) observed for the disease genes in the population. As in Thompson and Neel (1, 2), we computed prior probabilities; that is, we did not compute the probability for a gene to reach a given frequency conditional on its survival because we were interested in the evolution of the initial gene pool.

The sample of 891 ascending genealogies allowed us to incorporate all of the demographic parameters that influence the evolution of allele frequencies. It should be emphasized that here, rather than simulating a population, we simulated the allelic transmission in a real population. This real population can be seen as one realization of the demographic process we simulated with the branching process. Detailed results from this method called gene dropping (16) are presented in ref. 8.
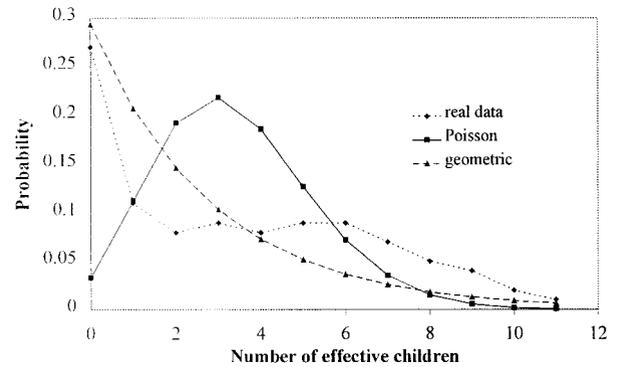

FIG. 2.   Observed distribution of EFS per married woman in the 19th-century SLSJ population compared with the Poisson and geometric distributions of same mean.

RESULTS

Demographic Analysis. The EFS distribution of the 19th-century SLSJ individuals ($n = 10,589$) shows a variance greater than a Poisson. The general shape of this curve was much more like a geometric distribution, with a high number of women with no effective children and many with high EFS (see Fig. 2). EFS was correlated from one generation to the next. Overall, the correlation was 0.161, with 0.144 for men and 0.174 for women, viewing the sexes separately. These data are for individuals married from 1870 to 1930. For the first 30 years period, the correlations are higher (0.176 for men, 0.194 for women, and 0.186 for both) and thereafter decreased slightly.

We also measured this correlation from 18th-century individuals in the ascending genealogies. The correlation in this case was much higher: 0.34. Because EFS measures the number of children who reproduced in the population, the discrepancy between these two values can be explained by more differential migration in the 18th century (17).

Changes in Gene Frequencies. Table 2 gives the mean number of founders with a given probability P that one of the two alleles that they introduced in the population reached the observed disease carrier frequency in the contemporary population. The distribution of EFS and the correlations both had an impact on allelic frequencies. The observed frequency of disease genes only was reached with a high probability for a geometric distribution and a relatively high level of intergeneration correlation (c, 0.6; r, 0.27).

The average number of founders with a P value >0.1 was lower than the number obtained with the gene-dropping method on genealogies, where nine founders were found. But there was high variance from one simulation to another. For the geometric distribution with the highest intergeneration dependency ($c = 1$), in 5 simulations of 100, nine or more founders had a P value >0.1. Even with $c = 0.5$, this was the case in 1 simulation of 100. However, even with $c = 1$, in 29 simulations of 100, no founders had a P value >0.1. The percentage of genes that were lost increased considerably (from 33 to 77%) when the geometric distribution was used instead of the Poisson distribution and also when the correlations were introduced into the system (Table 3).

DISCUSSION

Few, but some, founders' alleles reached the observed carrier frequencies for some inherited disorders in SLSJ. Assuming that every individual carries 4 to 5 recessive lethal alleles in its genome (see ref. 18, p. 499), we can conclude that demography itself is sufficient to explain why so many disease mutations have reached relatively high frequency in the population of SLSJ.

Table 2. Average number of founders ($\pm$SD) for which one of their alleles reached the frequency (0.035–0.05) with a probability P in a given range for a particular distribution of number of offspring per family and level of the c parameter (the corresponding correlation rate r also is given)

Demographic law	$0.001 < P < 0.01$	$0.01 < P < 0.1$	$P > 0.1$
Poisson ($c = 0$, $r = 0$)	0.00 ($\pm$0.00)	0.00 ($\pm$0.00)	0.00 ($\pm$0.00)
Poisson ($c = 0.5$, $r = 0.21$)	0.00 ($\pm$0.00)	0.00 ($\pm$0.00)	0.00 ($\pm$0.00)
Poisson ($c = 1$, $r = 0.36$)	0.00 ($\pm$0.00)	0.00 ($\pm$0.00)	0.00 ($\pm$0.00)
Geometric ($c = 0$, $r = 0$)	0.00 ($\pm$0.00)	0.00 ($\pm$0.00)	0.00 ($\pm$0.00)
Geometric ($c = 0.2$, $r = 0.12$)	0.00 ($\pm$0.42)	0.00 ($\pm$0.00)	0.00 ($\pm$0.00)
Geometric ($c = 0.3$, $r = 0.12$)	0.06 ($\pm$0.42)	0.00 ($\pm$0.00)	0.00 ($\pm$0.00)
Geometric ($c = 0.4$, $r = 0.2$)	0.61 ($\pm$1.10)	0.52 ($\pm$1.56)	0.06 ($\pm$0.34)
Geometric ($c = 0.5$, $r = 0.24$)	6.12 ($\pm$9.92)	3.26 ($\pm$6.07)	0.58 ($\pm$1.79)
Geometric ($c = 0.6$, $r = 0.27$)	17.50 ($\pm$20.02)	5.27 ($\pm$6.85)	0.73 ($\pm$2.05)
Geometric ($c = 0.7$, $r = 0.29$)	27.33 ($\pm$30.34)	12.44 ($\pm$9.11)	1.54 ($\pm$2.44)
Geometric ($c = 0.8$, $r = 0.30$)	42.41 ($\pm$35.77)	20.86 ($\pm$22.42)	2.05 ($\pm$2.65)
Geometric ($c = 0.9$, $r = 0.31$)	53.42 ($\pm$34.85)	26.12 ($\pm$24.15)	2.46 ($\pm$2.86)
Geometric ($c = 1$, $r = 0.32$)	67.82 ($\pm$45.68)	33.00 ($\pm$20.80)	3.38 ($\pm$3.48)
allele dropping	7	21	9

Values obtained with the allele-dropping method on genealogies also are shown (no SD can be calculated in this case because we only have one repetition of the demographic process).

Rapid demographic growth alone, however, cannot explain the high frequency of disease carriers in this population. For example, when using Thompson and Neel's (1) model on SLSJ population, the probability that a unique variant reaches a frequency between 0.035 and 0.050 in the contemporary population was only 2.35×10^{-31}, with $m = 1.41$, $r = 0.3$, and $g = 12$ from ref. 10, using Thompson and Neel's notations. In the context of another study on linkage disequilibrium in this population, we used a branching process (19) with a higher growth rate (1.8), and the probability of reaching the observed number of carrier chromosomes was also very small.

Similarly, a very large variance in the number of effective children per woman was also not sufficient to explain the high frequencies of inherited disorders. On the other hand, adding a low level of correlation ($r = 0.27$) between the effective family sizes for successive generations gave a good fit to the data on the frequency of genetic disorders in SLSJ. This correlation is included in our branching process through a c parameter (see Eq. **1**).

The factor c represents the partial dependence of demographic behavior of an individual on that of his parents. The c value is a summary of the vertical transmission of all demographic behaviors important for the effective reproduction of an individual. These include mortality, nuptiality, fertility, and also migration. Studies on the transmission of any one of these factors in the Quebec population did not show a strong link between generations (20, 21): correlations were very low and even null. But when all of these factors were taken together, correlations from one generation to the next appeared. For example, in the 19th century, the correlation between one generation and the next for family size, i.e., children born, is very low (0.07). Therefore, the observed correlations in EFS reflect mainly a cultural phenomenon with additional behavioral components, our population being characterized by very strong differential migration, the children of some families remaining mostly in SLSJ with others emigrating

Table 3. Average proportion (%) of genes introduced by a single founder that were lost

Demographic law	Proportion
Poisson, $c = 0$	33%
Poisson, $c = 1$	41%
Geometric, $c = 0$	57%
Geometric, $c = 0.4$	66%
Geometric, $c = 0.6$	70%
Geometric, $c = 1$	77%

massively, mostly toward western Quebec (Montreal) or to the U.S. during the later part of the 19th century (9, 10)

Our results clearly imply that the high frequency of inherited disorders in the SLSJ population is consistent with a correlation. Also, even if weak, this correlation had a strong impact on the evolution of the gene pool of the population. We have no evidence that this aspect of the demographic behavior is determined genetically, but this behavior strongly influences allele frequencies.

The c value required in the model to give observed carrier frequencies was quite high ($c = 0.6$, $r = 0.27$). This c value corresponds to a correlation rate (r) higher than that measured in the SLSJ population in the 19th century but slightly smaller than the measured value in the population for the 18th century ($r = 0.34$). Higher c in the model could be explained by two factors: (*i*) we used a constant value throughout the history of the population, but higher correlations of EFS during the first generations have a much stronger impact than lower values afterward; and (*ii*) we did not include in our model any assortative mating regarding the EFS, which presumably would lower the level of correlation that has to be entered in the model.

As we have shown, there is a great variance in the number of founders who have a high probability of giving the frequencies of disease alleles observed in the studied population. This indicates that the same demographic events, even on a short time scale, have a highly variable impact on the fate of rare alleles.

The factor c reduces the effective population size, yielding the loss of many alleles. When calculated from classical formulae by using the change in allelic frequencies obtained with the branching process simulations for the geometric distribution of EFS with the maximum intergenerational correlation, we found an effective population size $\approx$1,000. This is far lower than the expected value estimated from demographic data (the harmonic mean of the population is $\approx$17,000). This result is consistent with that of Nei and Murata (6), who found that intergenerational correlation in offspring number strongly reduces effective size and therefore increases the drift. Although rapid population growth usually prevents the loss of variation during the growth period (22, 23), variance in the number of children per woman and intergenerational correlations act in the opposite way.

SLSJ proved to be a very useful population for this study, thanks to the high quality of both demographic and genetic data that are available. For instance, correlations in progeny size can be calculated not only for women but also for men. This allowed us to test the validity of our model, notably by

comparing our results with the data on the frequencies of genetic disease carriers.

Our results are also consistent with those obtained by using real genealogies (8). With gene-dropping simulations on these genealogies, nine founders had a probability of >0.1 to see one of their alleles reach the observed carrier frequencies. Because of the high variance, we obtained this high number with the simulated populations in only a few cases. Nevertheless, other factors, like assortative mating suggested above, also could account in part for this higher number. Results (not shown) are also consistent with a previous study on founder's genetic contribution (14).

Using another population (the Valserine Valley in Eastern France), we also observed a similar process (24): the population is characterized by a core surrounded by a fringe of immigrants and their descendants. Core individuals have children that also belong to the core; they produce more effective children than do fringe individuals. This differential reproduction, transmitted from one generation to the next, may explain the existence of a high frequency of a rare inherited disorder in this valley.

Because demographic behavior of human populations has a clear impact on gene frequencies, it should be taken into account in using population genetics data to estimate parameters like mutation or recombination rate. Social transmission of EFS could explain high frequencies of mutant alleles in other populations, without necessarily invoking any heterozygous advantage. Further studies are needed in which the correlation in EFS, and not simply fertility correlation, is measured (see ref. 25, for example, on cystic fibrosis frequency).

In Indian tribes of Central America, it was shown that an intergenerational correlation in male fertility increased the expected amount of inbreeding in the population (26, 27). In a recent study on mitochondrial diversity among the Maori of New Zealand (28), the authors also proposed a cultural transmission of offspring number to explain the unequal distribution of the four haplotypes found in their sample. This process, among other explanations, would explain the high frequency of one haplotype (47/54) if it was carried by high ranking Maori females in the founding population. This process also could have contributed to the increase in frequency of idiopathic torsion dystonia in the Ashkenazi Jewish population from an initially lower frequency (29–31).

Other social factors may also have a significant effect on the population genetics of rare alleles. For example, McKusick *et al.* (32) derived a model based on nonrandom mating. In our model, we used a simple parameter c and restricted the transmission of demographic behavior to "vertical transmission" (33). This yielded good concordance between model results and data from real populations, but it does not exclude the occurrence of other modes of transmission.

1. Thompson, E. A. & Neel, J. V. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 1442–1445.
2. Thompson, E. A. & Neel, J. V. (1996) *Mol. Phylogenet. Evol.* **5,** 220–231.
3. Thompson, E. A. & Neel, J. V. (1997) *Am. J. Hum. Genet.* **60,** 197–204.
4. Fisher, R. A. (1930) *The Genetical Theory of Natural Selection* (Clarendon, Oxford).
5. Huestis, R. R. & Maxwell, A. (1932) *J. Hered.* **23,** 77–79.
6. Nei, M. & Murata, M. (1966) *Genet. Res.* **8,** 257–260.
7. Williams, L. A. & Williams, B. J. (1974) *Soc. Biol.* **21,** 225–231.
8. Heyer, E. (1999) *Hum. Biol.* **71,** 91–101.
9. Bouchard, G., Charbonneau, H., Desjardins, B., Heyer, E. & Tremblay, M. (1993) in *Les Chemins de la Migration en Belgique et au Québec, XVIIe-XXe Siècles* (Editions Académia, Louvain-la-Neuve, Canada), pp. 51–59.
10. Tremblay, M. & Heyer, E. (1993) *Cah. Québécois de démographie* **22,** 263–283.
11. Tremblay, M. (1997) *Cah. québécois de démographie* **26,** 129–145.
12. Kaplan, N. L., Hill, W. G. & Weir, B. S. (1995) *Am. J. Hum. Genet.* **56,** 18–32.
13. Kaplan, N. L. & Weir, B. S. (1995) *Am. J. Hum. Genet.* **57,** 1486–1498.
14. Heyer, E. & Tremblay, M. (1995) *Am. J. Hum. Genet.* **56,** 970–978.
15. Heyer, E., Tremblay, M. & Desjardins, B. (1997) *Hum. Biol.* **69,** 209–225.
16. MacCluer, J. W., VandeBerg, J. L., Read, B. & Ryder, O. A. (1986) *Zoo Biol.* **5,** 147–160.
17. Charbonneau, H., Desjardins, B., Guillemette, A., Landry, Y., Légaré, J. & Nault, F. (1987) *Naissance d'une Population: Les Français Établis au Canada au XVIIè Siècle.* (Presses de l'Université de Montréal et Presses Universitaires de France, Montreal and Paris).
18. Vogel, F. & Motulsky, A. (1986) *Human Genetics* (Springer, Berlin).
19. Austerlitz, F. & Heyer, E. *Genet. Epidemiol.*, in press.
20. Desjardins, B. & Charbonneau, H. (1990) *Population* **45,** 603–615.
21. Desjardins, B., Bideau, A., Heyer, E. & Brunet, G. (1991) *J. Biosoc. Sci.* **23,** 49–54.
22. Kojima, K.-I. & Kelleher, T. M. (1962) *Am. Nat.* **96,** 329–346.
23. Nei, M., Maruyama, T. & Chakraborty, R. (1975) *Evolution* **29,** 1–10.
24. Heyer, E. (1993) *Ann. Hum. Biol.* **20,** 565–573.
25. Pritchard, D. (1991) *Hum. Genet.* **87,** 671–676.
26. Neel, J. V. (1970) *Science* **170,** 815–822.
27. MacCluer, J. W., Neel, J. V. & Chagnon, N. A. (1971) *Am. J. Phys. Anthropol.* **35,** 193–207.
28. Murray-McIntosh, R. P., Scrimshaw, B. J., Hatfield, P. J. & Penny, D. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 9047–9052.
29. Risch, N., de Leon, D., Ozelius, L., Kramer, P., Almasy, L., Singer, F., Fahn, S., Brakefield, X. & Bressman, S. (1995) *Nat. Genet.* **9,** 152–159.
30. Risch, N., de Leon, D., Fahn, S., Bressman, S., Ozelius, L., Brakefield, X., Kramer, P., Almasy, L. & Singer, B. (1995) *Nat. Genet.* **11,** 14–15.
31. Labuda, D., Zietkiewicz, E. & Labuda, M. (1997) *Am. J. Hum. Genet.* **61,** 768–771.
32. McKusick, K. B., Schach, S. R. & Koeslag, J. H. (1990) *Am. J. Med. Genet.* **36,** 178–182.
33. Cavalli-Sforza, L. L. & Feldman, M. W. (1991) *Cultural Transmission and Evolution: A Quantitative Approach.* (Princeton Univ. Press, Princeton, NJ).