Hum Genet (2005) 117: 177–187
DOI 10.1007/s00439-005-1298-8

ORIGINAL INVESTIGATION

**Vania Yotova · Damian Labuda · Ewa Zietkiewicz
Dominik Gehl · Alan Lovell · Jean-François Lefebvre
Stéphane Bourgeois · Émilie Lemieux-Blanchard
Marcin Labuda · Hélène Vézina · Louis Houde
Marc Tremblay · Bruno Toupance · Evelyne Heyer
Thomas J. Hudson · Claude Laberge**

Anatomy of a founder effect: myotonic dystrophy in Northeastern Quebec

Received: 26 October 2004 / Accepted: 18 February 2005 / Published online: 10 May 2005

Abstract Founder effects are largely responsible for changes in frequency profiles of genetic variants in local populations or isolates. They are often recognized by elevated incidence of certain hereditary disorders as observed in regions of Charlevoix and Saguenay-Lac-Saint-Jean (SLSJ) in Northeastern Quebec. Dominantly transmitted myotonic dystrophy (DM1) is highly prevalent in SLSJ where its carrier rate reaches 1/550, compared with 1/5,000 to 1/50,000 elsewhere. To shed light on the origin of DM1 in this region, we have screened 50 nuclear DM1 families from SLSJ and studied the genetic variation in a 2.05 Mb (2.9 cM) segment spanning the site of the expansion mutation. The markers analyzed included 22 biallelic SNPs and two microsatellites. Among 50 independent DM1 chromosomes, we distinguished ten DM1-associated haplotypes and grouped them into three haplotype families, A, B and C, based on the relevant extent of allele sharing between them. To test whether the data were consistent with a single entry of the mutation into SLSJ, we evaluated the age of the founder effect from the proportion of recombinant haplotypes. Taking the prevalent haplotype A1_21 (58%) as ancestral to all the disease-associated haplotypes in this study, the estimated age of the founder effect was 19 generations, long predating the colonization of Nouvelle-France. In contrast, considering A1_21 as ancestral to the haplotype family A only, yielded the estimated founder age of nine generations, consistent with the settlement of Charlevoix at the turn of 17th century and subsequent colonization of SLSJ. We conclude that it was the carrier of haplotype A (present day carrier rate of 1/730) that was a "driver" of the founder effect, while minor haplotypes B and C, with corresponding carrier rates of 1/3,000 and 1/10,000, respectively, contribute DM1 to the incidence level known in other populations. Other studies confirm that this might be a general scenario in which a major "driver" mutation/haplotype issued from a founder effect is found accompanied by distinct minor mutations/haplotypes occurring at background population frequencies.

Keywords SNPs · Founder effect · Myotonic dystrophy · Northeastern Quebec

Electronic Supplementary Material Supplementary material is available for this article at http://dx.doi.org/10.1007/s00439-005-1298-8

V. Yotova · D. Labuda · E. Zietkiewicz · D. Gehl
A. Lovell · J.-F. Lefebvre · S. Bourgeois
É. Lemieux-Blanchard · M. Labuda
Centre de recherche, Hôpital Sainte-Justine,
3175 Côte Sainte-Catherine,
Montréal, PQ, Canada, H3T 1C5

D. Labuda (✉)
Département de Pédiatrie,
Université de Montréal, Montréal, PQ, Canada
E-mail: damian.labuda@umontreal.ca
Tel.: +1-514-3454931
Fax: +1-514-3454731

H. Vézina · L. Houde · M. Tremblay
Interdisciplinary Research Group in Demography and
Genetic Epidemiology, Université du Québec à Chicoutimi,
Chicoutimi, PQ, Canada

B. Toupance · E. Heyer
Unité d'Eco-Anthropologie,
UMR CNRS/Université Paris, Paris, France

T. J. Hudson
McGill University and Genome Quebec Innovation Centre,
Montreal, PQ, Canada

C. Laberge
Centre Hospitalier Université Laval,
Québec, PQ, Canada

E. Zietkiewicz
Institute of Human Genetic,
Polish Academy of Sciences, Poznan, Poland

Introduction

Myotonic dystrophy type I (DM1; MIM160900) or Steinert disease, a dominantly inherited autosomal

disorder with uneven penetrance, is the most common form of muscle disease, characterized by progressive muscle weakness, wasting and myotonia (Harper 1989). It is caused by an expansion of CTG-trinucleotide repeats in the 3′-untranslated region of the serine/threonine protein kinase gene (*DMPK*, for dystrophia myotonica protein kinase) on chromosome 19q13.3 (Brook et al. 1992; Fu et al. 1992). As opposed to the normal range of 5–35 repeats, the number of CTG-repeats in alleles of DM1 patients ranges from 80 up to several thousand (Brook et al. 1992; Fu et al. 1992; Harley et al. 1992; Martorell et al. 2001). The disease-causing alleles are believed to originate from so-called protomutated alleles (of between 50 and 80 repeats) that are unstable and occasionally expand (Brunner et al. 1992). The predisposing event(s) that led to a protomutated allele could have occurred (1) once as an increase of the size of the CTG-repeat to a critical protomutated length, or (2) through the acquisition of other modifications in *cis* promoting multiple independent expansions to the protomutated state and beyond (e.g. Labuda et al. 2000).

The DM1 predisposing mutation seems to have originated only once for both the disease-causing alleles and the protomutated state were found to be in tight association with the same set of surrounding markers, suggesting a single common ancestral chromosome (Abbruzzese et al. 2002; Brunner et al. 1989; Goldman et al. 1995, 1996; Junghans et al. 2001; Yamaoka et al. 1990). Due to its single origin, the ancestral haplotype represents a genetic signature of DM1 or its protomutated alleles. However, the length of the common, disease-associated haplotype decreases with time due to recombination events. On the other hand, each time a single copy of the predisposing chromosome enters a new population, a founder effect is repeated, extending the genetic signature of the disease over all marker alleles found on this de novo ancestral chromosome.

DM1 is particularly frequent in Saguenay-Lac-Saint-Jean (SLSJ) and in Charlevoix (in Northeastern Quebec), where its carrier rate in the range of 1/500 to 1/600 (Mathieu et al. 1990) greatly exceeds that of 1/50,000 to 1/7,000 found in most other populations worldwide (Harper 1989), including the rest of Quebec (~1/10,000; Jack Puymirat, personal communication). This dramatic increase in the DM1 carrier frequency resembles those known for a number of other diseases in these populations and is likely to reflect a founder effect that has occurred during the colonization of the Charlevoix and SLSJ regions (hereafter collectively referred to as Ch–SLSJ) in the 17th to 19th centuries (Gauvreau et al. 1991; Jetté et al. 1991; Labuda et al. 1996; Scriver 2001). The SLSJ population (278,000 inhabitants as of 2004) is relatively young. Sustained settlement in this region started in the second quarter of the 19th century (Pouyez and Lavoie 1983), with the majority of the first settlers coming from the adjacent region of Charlevoix, whose colonization began in the late 17th century, just a few decades after the arrival of the first French settlers in Canada (Jetté et al. 1991). In spite of a limited number of founders, the average annual growth rate of around 3% from 1750 to 1850 led to demographic pressures that induced strong emigration during the 19th century. As a result, between 1838 and 1911, nearly 14,000 people from Charlevoix emigrated to SLSJ (Gauvreau et al. 1991), such that approximately 65% of the contemporary SLSJ gene pool can be attributed to founders from Charlevoix (Tremblay et al. 2002).

To examine mechanisms of the DM1 frequency increase in the SLSJ region, we analyzed a representative sample of disease-carrying chromosomes in this population. We found that an SNP-based and microsatellite-enriched haplotype may serve as a genetic signature of DM1-associated chromosomes in SLSJ. Indeed, assaying an SNP-based haplotype may replace technically demanding direct analysis of CTG-expansion and facilitate evaluation of the carrier/predisposition status at a population scale. Our analysis of the extent of decay of linkage disequilibrium between the disease and the marker alleles suggests that while the founder effect during the settlement of Charlevoix (and subsequently SLSJ) may easily explain the high frequency of the disease observed today, more than one carrier has introduced the disease to SLSJ.

Materials and methods

DNA samples

Nuclear DM1 families of French-Canadian origin from the region of SLSJ were taken from a previous study (Thibault et al. 1989); the presence of CTG-expansion was subsequently confirmed. DNA samples of both parents and two to three children, at least one of which was affected by DM1, were analyzed in 37 families. DNA of one affected parent and two affected children, or of both parents and a single affected child, were analyzed in 13 additional families. Unrelated population samples (nine French-Canadian and nine Polish individuals), used for the searching of new DNA variants, were obtained on a non-nominative basis from consenting adults. The research protocol was approved by the relevant Institutional Review Boards.

Polymorphisms and genotyping

Twenty-nine segments, to be screened for the presence of polymorphisms, were selected from ten contigs surrounding the *DMPK* gene on chromosome 19q13.3, as identified in the NCBI database (radiation-hybrid map as of August 2000). Two to four amplicons of ~300 bp were designed in each of these segments using Primer 3 Software (Rozen and Skaletsky 2000). Electronic supplementary material (ESM) Table 1S lists the primers used in these assays. DNA fragments were amplified by standard PCR using 10 ng of the genomic template per

reaction. The products were analyzed by denaturing high-pressure liquid chromatography (dHPLC) (using the WAVE system, Transgenomic) at three different temperatures. The purported heteroduplexes were sequenced in an ABI 377 Automated Sequencer (Perkin-Elmer, Applied Biosystems, Foster City, Calif., USA). Fourteen of the 24 substitution or small insertion/deletion polymorphisms found by dHPLC were retained as markers for subsequent genotyping of the DM1 families.

In addition, we genotyped four previously described substitution polymorphisms from the *ERCC2* locus (Shen et al. 1998), three polymorphisms from introns 5, 9 and 14 of the *DMPK* gene (Mahadevan et al. 1993), the *Taq*I RFLP polymorphism (D19S463) (Neville et al. 1994), and two dinucleotide repeats: D19S219 and D19S412 (Gyapay et al. 1994). To resolve whether variation in D19S412 on the affected chromosomes was due to mutations or recombinations of this flanking microsatellite (see Results), an ad hoc, potentially polymorphic, CA-repeat marker, was typed in a subgroup of DM1 patients. This marker was located approximately 100 kb in a centromeric direction from D19S412 and defined by primers 99960F, AGCCTGAGCGACAGAATGAG and 99960R, ATGTGTCTCCTCCTTGGTGC.

Allelic state of substitution and small insertion/deletion polymorphisms were determined by allele-specific oligonucleotide (ASO) hybridization (Zietkiewicz et al. 1997) (ESM Table 1S). The *Taq*I RFLP assay was performed according to Neville et al. (1994). Standard denaturing electrophoresis in a polyacrylamide gel was used to type the CA-repeats. Partial RFLP data from a previous study by Thibault et al. (1989) were added for the *APOC2* (*Taq*I/*Nco*I), *CKMM3* (*Taq*I/*Nco*I) and *Pd10* (*Pvu*2) loci.

Genomic positions of the markers constituting the haplotype, according to the UCSC browser assembly from April 2003 (http://genome.ucsc.edu/) (Kent et al. 2002), are given in Fig. 1. The relative order and distances of these positions changed frequently during the period of 2000–2002, stabilizing only recently; this explains the uneven distribution of SNP markers analyzed on either side of the *DMPK* gene. The resulting haplotype covers a region of 2.05 Mb, based on the UCSC browser, or 2.9 cM, according to the deCode genetic map (Kong et al. 2002).

The reported haplotypes were obtained directly from pedigrees and subsequently confirmed by a computer-assisted analysis using Cyrillic 2 (http://www.cyrillicsoftware.com). Population frequencies of the haplotype-forming

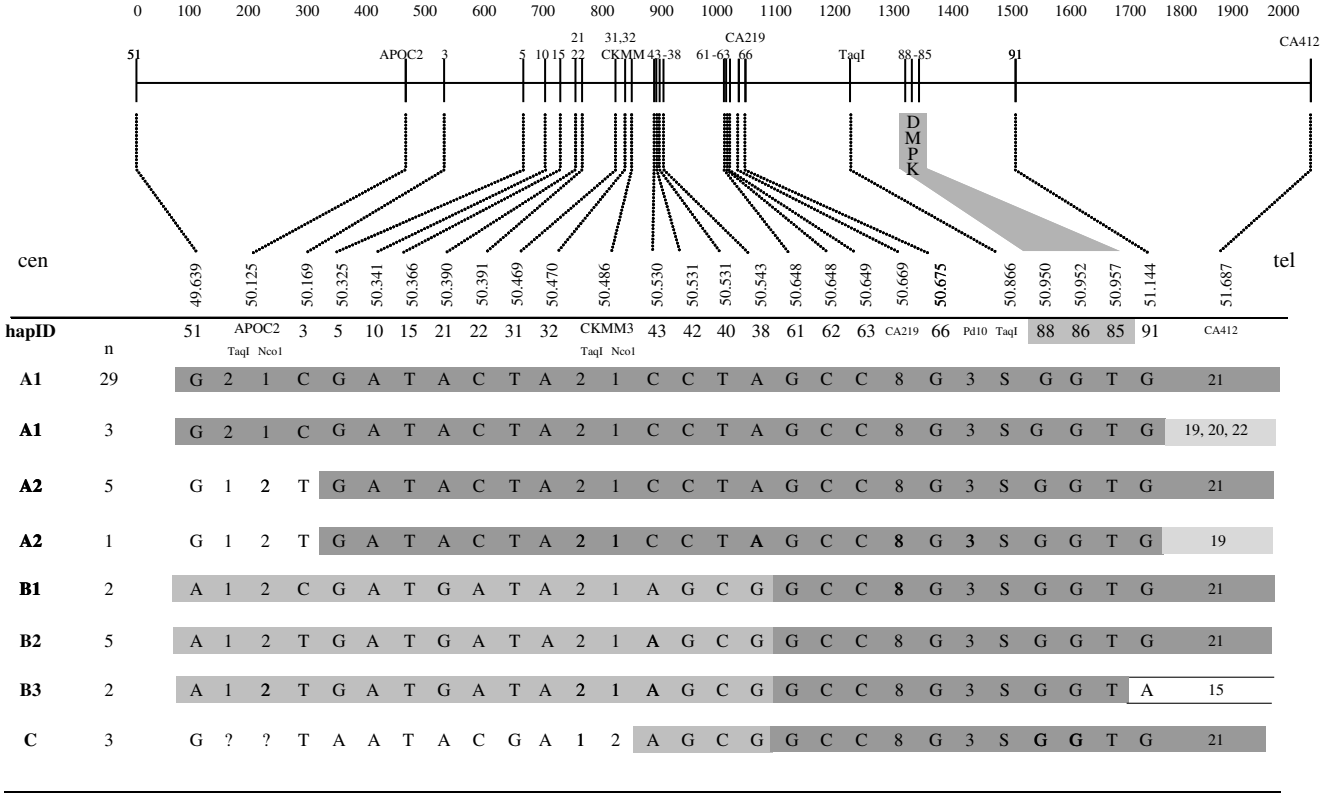

Fig. 1 DM1-associated haplotypes (*lower*) and the distribution of the contributing markers with their names *on top* of the physical map (*upper*), and with their UCSC map positions given at the level where the orientation, centromeric to telomeric, is indicated. Haplotype IDs given in the *left column* do not include the diversity of the distal CA412 repeat, whose allelic state is indicated in the *rightmost column*. Alleles of substitution polymorphisms are represented by the corresponding *nucleotide residue* in the DNA sequence; microsatellite polymorphisms by *numbers* describing their repeat length; and RFLP polymorphisms by *letters* (e.g., *S* for short in *Taq*I RFLP), as used in the literature

alleles were obtained by counting them on all independent chromosomes "entering" each analyzed pedigree, excluding those associated with the disease. Haplotypes were given arbitrary names, reflecting their structural similarity at the level of biallelic markers; haplotypes associated with different alleles of the CA412 microsatellite were distinguished by adding the relevant extension, e.g. A1_21 would correspond to haplotype A1 with 21 repeats at CA412 (see Fig. 1).

Statistical analysis

When a disease mutation (or a protomutation) arises de novo or is introduced into a population by an immigrating carrier, all descendant individuals initially carry the same disease-associated founder haplotype. The fraction of the mutation-carrying chromosomes harboring the full-length founder haplotype decreases over time, at a proportion $(1–r)$ per generation, where r is the recombination rate over the genetic distance considered. After g generations, the expected proportion of a complete (i.e. non-recombined) founder haplotype will be $P=(1–r)^{g}$. At a small r, the age of the founder effect g can be estimated from the equation

$$\ln P = -rg \tag{1}$$

However, some haplotypes, identical by state with the bona fide ancestral one, could have arisen by recombination, especially if alleles of the markers at the edges of this haplotype occur at significant frequencies on normal chromosomes. Consequently, the real proportion of recombinants, corresponding to $(1–P)$, differs from the observed one, R, such that $R=(1–P)(1–p_{n})$, where p_{n} describes the frequency of the linked allele present on normal chromosomes. The observed proportion of recombinant haplotypes can be also expressed as $R=1–p_{d}$, where p_{d} is the frequency of the corresponding allele on the disease chromosomes. Substituting $P=(p_{d}–p_{n})/(1–p_{n})$ into Eq. 1, we obtain

$$-\ln\left(\frac{p_{d}-p_{n}}{1-p_{n}}\right)=rg \tag{2}$$

from which g, the age of the founder effect, can be evaluated (Bengtsson and Thomson 1981).

If a population grows rapidly (Hastbacka et al. 1992; Kaplan and Weir 1995; Luria and Delbrück 1943), the number of generations since the founder effect would be better approximated by adding the correction g_{0} to the above estimate (Labuda et al. 1996, 1997). The correction is

$$g_{o}=-\frac{1}{d}\ln\left(\frac{rf_{d}}{1+rf_{d}}\right) \tag{3}$$

where $f_{d}=e^{d}/(e^{d}–1)$ in a population growing at a rate d per generation. We note that a d of 0.8, estimated for the Ch–SLSJ population assuming an average generation time of 25 years (Labuda et al. 1996), would correspond to $d=0.96$ if a generation span of 30 years was assumed, or to $d=0.89$ per generation using an annual growth rate of 3%.

Pairwise linkage disequilibrium in both disease and non-disease chromosomes was measured by the parameter D according to Lewontin and Kojima (1960), using the Arlequin Software (Schneider et al. 2000).

Genealogical reconstruction

Ascending genealogies of DM1 carriers were reconstructed using the BALSAC-RETRO database (Jomphe et al. 2001). All genealogies were completed up to the first immigrants, going back as far as the early 17th century. Mean kinship coefficients among DM1 carriers were calculated by identifying all known common ancestors over all known genealogical links between each pair of carriers, for each generation level (up to 13 generations). For all ancestors found in the genealogies, we calculated the number of carriers to which they were related. Ancestors' places and dates of marriage were obtained from the BALSAC-RETRO marriage records.

Results

We genotyped 50 nuclear DM1 families for 22 bi-allelic and two microsatellite markers along a 2.05 Mb region surrounding the *DMPK* gene. The resulting haplotype extends over 2.9 cM, 1.85 cM towards the centromere, and 1.05 cM towards the telomere upstream from the gene (Kong et al. 2002). On 50 independent carrier chromosomes, we observed ten distinct haplotypes, as shown in Fig. 1. The markers 61–85, including the *DMPK* locus, constituted a core haplotype, common to all the disease-chromosomes analyzed in this study. More variability was observed among the marker alleles at the left and right segments of the haplotype. For convenience, we divided DM1 haplotypes into three groups, denoted family A, B and C, which emphasize their structural similarities at the level of the biallelic markers. Family A was subdivided into two subgroups, A1 and A2, which additionally differed in their allelic states for CA412 microsatellite. Note that the DM1-associated haplotypes included partial information on the genotyped status of the *APOC2*, *CKMM3* and *Pd10* (D19S63) alleles, available from the previous study by Thibault et al. (1989). These earlier results, while consistent with our findings, were incomplete and therefore could not be used in all the analyses reported below. However, these data did help to phase chromosomes in two DM1 families, and to distinguish between the A1 and A2 haplotype subgroups. Altogether, we observed 31 unambiguously phased A1 and five A2 haplotypes in the disease chromosomes. Two additional A haplotypes could have been either A1_21 or A2_21. For convenience, in the analyses presented below, we have

assumed having observed 32 A1 and six A2 haplotypes (Fig. 1).

Looking at the markers downstream (centromeric) of the *DMPK* locus (Fig. 1), haplotype A2_21 can be derived from A1_21 by a recombination event exchanging segment 51–3, 0.78 Mb downstream of the *DMPK* gene. The B and C haplotypes differ from haplotype A in segments 51–38, 0.4 Mb downstream from *DMPK*. Haplotype B1_21 can be derived from B2_21 by a recombination event, again exchanging segment 51–3. In turn, haplotype B3_15 seems to be derived from B2_21 by a recombination event involving segment 91-CA412, 0.18 Mb upstream of *DMPK*. Haplotype C_21 can be derived either from B1/2_21 by recombination occurring between polymorphisms CKMM3 and 43, or from A1/2_21 by recombination between the more distant markers 38 and 61. Alternatively, A, B and C can be ancestrally linked to one or more distinct ancestral haplotypes.

In contrast to the disease-associated haplotypes, those found on non-affected chromosomes, and thus representing the general population, were highly diversified. Consequently, allele frequencies of all contributing markers differed between these two groups of chromosomes (Fig. 2). The difference was remarkable at the level of biallelic markers, and even more pronounced in the multiallelic microsatellites. From nine alleles of CA219 microsatellite, only the shortest one (found in the general population at a frequency of 28%) was seen on the DM1 chromosomes (Fig. 2b), suggesting their identity by descent at this particular locus. Nine alleles of CA412, the most telomeric flanking microsatellite, exhibited a bimodal population frequency distribution (Fig. 2c), similar to that reported for CEPH families (GDB: 246049). In this marker, one frequency peak was at allele 21 (found on 37% of non-DM1 chromosomes) surrounded by minor alleles, one or two repeats apart (frequencies of about 5%). The second mode peaked at alleles 16 and 15, representing 19 and 14% of the non-DM1 chromosomes, respectively. Forty-four out of the 50 DM1 chromosomes (88%) were associated with the CA412 allele 21 (Fig. 1). There was one clear case of recombination, leading to haplotype B3_15, where replacement of allele 21 by the frequent allele 15 at the CA412 locus was accompanied by an allele replacement at marker 91, 0.5 Mb closer to *DMPK*. The remaining non-21 CA412 alleles found on the disease-associated chromosomes were associated with haplotypes A1 and A2. They all represented minor alleles separated from the modal peak of 21 by one (alleles 20 or 22) or two repeats (allele 19), as if they were due to stepwise mutations (Valdes et al. 1993) rather than to crossover events. This, however, is an unlikely possibility, implying an unusually fast mutation rate, of the order of $\sim 10^{-2}$ per generation, which is comparable to or even faster than the recombination rate implied by the genetic distance between CA412 and the disease locus. If mutations were responsible for the CA412 variability in haplotypes A1 and A2, we would expect other markers, adjacent to this repeat polymorphism, to be identical in the modal (A1_21 and A2_21) and in the derived haplotypes (A1_22, 20, 19 and A2_19), respectively. To test this, we searched in silico for potentially polymorphic microsatellite sequences adjacent to the CA412 locus. One informative CA-repeat marker, located approximately 100 kb centromeric from CA412, was analyzed in 24 DM1 patients (due to shortage of the genomic material we could not systematically test all DM1 families). We

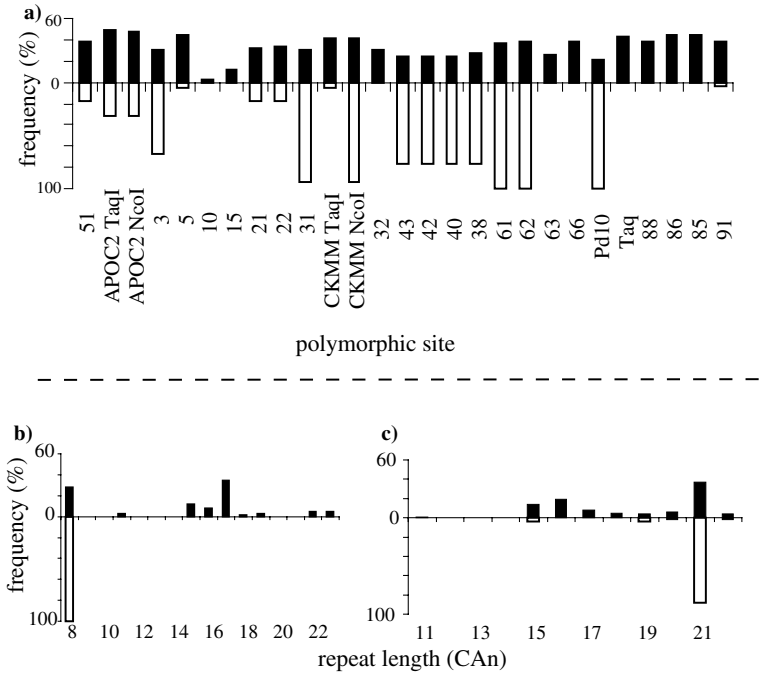

Fig. 2a–c Allele frequencies of the markers contributing to the haplotype in the general population (*upper*) and on disease-associated chromosomes (*lower part of the plot*). Minor allele frequencies of non-microsatellite markers are given in (**a**) and the allelic distribution of D19S219 and D19S412 CA repeats are given in (**b**) and (**c**), respectively

found that all 15 tested patients with the CA412 allele 21 shared the same common "upper allele" of the centromeric marker. On the contrary, patients with the CA412 allele 19 showed an allele of the ad hoc marker that was two CA-repeats below the common "upper allele," and patients with CA412 alleles 15, 20 and 22 inherited yet another common "lower allele". These findings are thus consistent with the interpretation that recombinations, and not stepwise mutations, were responsible for the CA412 allele diversity on the DM1 chromosomes.

The age of the founder effect

The decay of linkage disequilibrium due to recombinations (genetic clock) can be used to date the age of the disease-associated mutation or that of its introduction into the analyzed population on a particular haplotype (see Materials and methods). Assuming a single founder event, we considered the most common haplotype, A1_21 (29 copies out of 50), as ancestral (Fig. 3a). Using the recombination rate, $r = 0.029$, as estimated from Kong et al. (2002), we obtain (see Eq. 1) 18.8 generations as the age of the founder effect (Table 1). Using a generation time of about 30 years (Tremblay and Vezina 2000), this corresponds to 560 years, placing the founder effect at the end of fourteenth century, well before the colonization of Nouvelle-France and the foundation of the Charlevoix population in the 17th/18th century. Similar age estimates were obtained considering markers to the left or right of the disease locus separately (Table 1). Note that in the case of the telomeric part of the haplotype, allele 21 of the CA412 marker occurs at a significant proportion also on normal chromosomes ($p_n = 0.37$) and thus Eq. 2 rather than Eq. 1 applies. In the left part of the haplotype, the presence of intermediate markers spread evenly between DMPK and the most centromeric marker makes each recombination virtually unique and thus Eq. 2 reduces to Eq. 1. In addition, the probability of recurrent recombinational exchange of the most centromeric allele is no longer defined by its distance from DM1, but it rather reflects the probability of its recombination with the immediately adjacent markers.

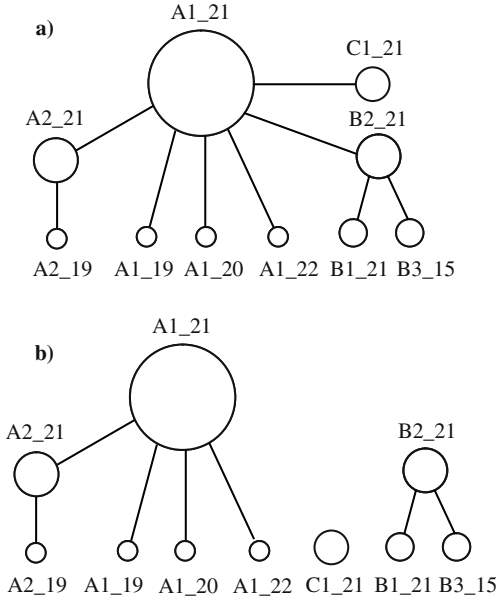

Fig. 3 Network of DM1-associated haplotypes assuming a unique founder effect and A1_21 as an ancestral haplotype (**a**) or a major founder effect of haplotypes A and separate entries of haplotypes B and C (**b**). *Size of the circle* reflects relative frequency of the haplotype

Table 1 Age of the founder effect estimated assuming single entry for different groups of DM1-associated haplotypes

Founder effect scenario—assuming single entry of:	Age in generations		
	Total haplotype	Centromeric markers	Telomeric markers
All haplotypes	18.8	24.1	20.1
As only	9.3	9.3	17.4
Bs and C	30.0	29.1	29.3

Estimates were obtained from Eq. 1 (total haplotype or that consisting of centromeric markers) or Eq. 2 (telomeric markers)

Since estimations of the age of the ancestral haplotype were inconsistent with a single founder effect during the colonization of Nouvelle France, we considered multiple entry scenarios. Considering haplotype A1_21 to be ancestral for only the A family of haplotypes (which represents 76% of the analyzed sample) leads to an estimate of the founder effect of about nine generations, coinciding with the colonization of Charlevoix at the end of the 17th century. Correcting for rapid population growth (using Eq. 3 and $d = 0.9$) increases the age of the founder effect to 12.8 generations, still plausible given the uncertainty of the estimate. Similar estimations, assuming haplotype B1_21 as ancestral for the group of haplotypes B and C taken together (Table 1), yielded founder age estimates of ~30 generations.

Using a different approach, which simultaneously evaluates the demographic growth rate and the Luria–Delbruck correction (Austerlitz et al. 2003), we obtained the founder effect age estimate of 28.8 (22.9–36.8) generations for the full data, and 15.7 (12–21.1) for the haplotypes A. Taken together, these analyses suggest that all three haplotype families, A, B and C, shared a common ancestral haplotype well before their carriers settled in North America (Fig. 3b).

Genealogical analysis

Genealogies were analyzed based on groups A, B and C defined by the haplotype data. Distribution of the places of marriage of the closest ancestors reflects the expected concentration in Ch–SLSJ and does not show any significant differences between the three groups (Table 2). Parents and grandparents of A and B haplotype carriers

Table 2 Places of marriage (counts and proportions *in parentheses*) of the closest ancestors for haplotype carriers A, B and C

Generation	Place of marriage			Total
	SLSJ	Charlevoix	Other	
Group A				
Parents	36 (0.95)	2 (0.05)	0	38
Grandparents	55 (0.72)	13 (0.17)	8 (0.11)	76
Great-grandparents	73 (0.48)	60 (0.39)	19 (0.13)	152
Group B				
Parents	6 (0.67)	2 (0.22)	1 (0.11)	9
Grandparents	12 (0.67)	6 (0.33)	0	18
Great-grandparents	15 (0.47)	15 (0.47)	2 (0.06)	32
Group C				
Parents	2 (0.67)	0	1 (0.33)	3
Grandparents	2 (0.33)	2 (0.33)	2 (0.33)	6
Great-grandparents	1 (0.08)	8 (0.66)	3 (0.25)	12

were predominantly married in the SLSJ region: 95 and 72%, respectively, for group A, and 67 and 67% for group B. A large fraction of great-grandparents of A (48%) and B (47%) carriers were also married in the SLSJ, but many others came from Charlevoix (39% for group A and 47% for group B). For C carriers, a somewhat greater proportion of marriages outside SLSJ and Charlevoix was noticed, but the numbers were very small (three pairs of parents and six pairs of grandparents).

The genealogical impact of all common ancestors can be summarized using the mean kinship coefficients (Fig. 4). The effect of multiple introducers of DM1 mutations could be detected as a difference between the mean kinship coefficients within and among groups. However, the results revealed a difference only between group B and the other groups, and this difference was mainly explained by kinship links at the third generation in group B. The kinship values at intermediate generations indicated the population growth from a limited number of founders, and the leveling off at the highest generations suggests that carriers share a great number of ancestors at these genealogical depths. Any of these ancestors, were they carriers, could have introduced the mutation in the population, albeit with highly variable probabilities. On the other hand, rare migrants from outside the historically defined region (Table 2) could have introduced the mutation on a different haplotype, without affecting the pattern seen in Fig. 4, if their offspring intermarried with the "local" population.

Haplotype diversity and linkage disequilibrium among biallelic markers

Among the 126 non-disease chromosomes fully typed for 22 biallelic markers (haplotype extending over 1.5 Mb and 2.1 cM), we observed 114 haplotypes, of which 104 were observed once, eight twice and two occurred three times. None of these were identical with any of the DM1-associated haplotypes. Thus the latter, once defined, represent signatures of the DM1 mutation. Knowing that certain SNP-based haplotypes are uniquely associated with the DM1 mutation suggests that technically demanding population screening for CTG-expansion could be assisted by searching for the unique SNP-based haplotype(s) that mark the majority of mutation-carrying chromosomes. This approach, amenable to automatic screening, could be effective in evaluating risk in the majority of the population; it would also reduce the task of direct mutation detection to the fraction of genotypes where the presence of a marker haplotype(s) could not be excluded. However, in a 126×126 matrix of possible genotypes that could be

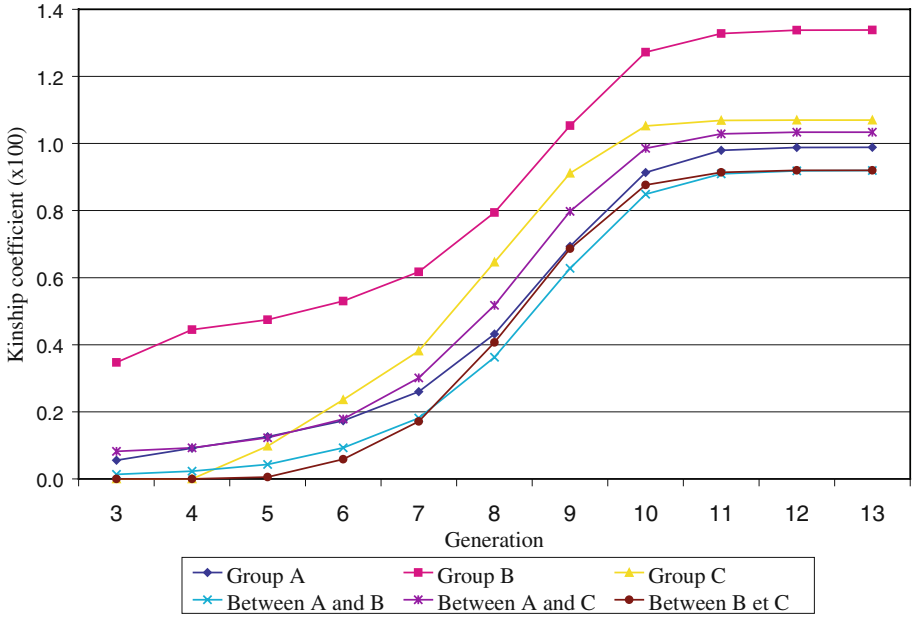

Fig. 4 Kinship coefficients within and between haplotype groups A, B and C from the 3rd to the 13th generation

constructed using all the non-disease haplotypes determined in this study, as many as 19.3% (3,071 genotypes) could have contained one of the six DM1-associated haplotypes (A1/2, B1/2/3 or C), thus representing the rate of false positives due to the full-length haplotype being found in a healthy population. Markers closer to *DMPK* should, therefore, reduce the probability of a false positive if such a test was to be considered for population screening.

Analysis of linkage disequilibrium (LD) provides another way of looking at the complexity of these haplotypes. The coefficient D' decays exponentially as a function of time and of the genetic distance between recombining markers. The markers analyzed in this study displayed almost complete linkage disequilibrium in the disease-associated haplotypes, while a rapid D' decay with the pairwise distance between the markers was observed in haplotypes from non-affected individuals (Fig. 5). While there are islands of LD around clusters of adjacent markers, D' in the general population does not generally exceed 60% at distances of more than 50 kb. Extensive LD is seen only in the disease chromosomes that were collected through the clinical bias, while no strong signal of the founder effect is observed in the general population represented by the non-disease chromosomes.

Discussion

Numerous previous DM1 studies have used RFLP markers that were often identified by a clone name rather than its sequence, and hence today it is often impossible to localize them on the human genome. In addition, the physical and genetic map of the chromosome 19q changed many times before stabilizing in 2002, such that it is difficult to compare results across studies. This is also true for the study of Whiting et al. (1995), who analyzed DM1 haplotypes in Canadian patients. Interestingly, two major haplotypes (in addition to their derivatives) were observed in French-Canadians from Northeastern Quebec and from the Outaouais region, which were distinct from the major haplotype characteristic of Canadian DM1 patients of non-French descent. An example of the diversification of a founder effect in an isolated population is provided by Segel et al. (2003), who studied the prevalence of DM1 in three Israeli-Jewish populations: the Ashkenazim, Sepharadim and Yemenites. They found that DM1 prevalence varied widely, ranging from as low as 1/17,544 in the Ashkenazim to 1/5,000 in the Sepharadim and up to 1/2,114 in the Yemenites. Furthermore, strong LD was found for two CA repeat markers (D19S112 and D19S207) in the Sepaharidic and Yemenite Jews, but not in the Ashkenazim population. Such a result would suggest a recent bottleneck or founder effect in the Sepharadic and Yemenite populations. An extragenic marker D19S63 has been studied for a number of worldwide DM1 populations, and its "allele 3" was found to range from 32% in the French DM1 population (Lavedan et al. 1994), to 43% in the Japanese (Yamagata et al. 1992), 58% in Britain (Harley et al. 1991) and up to 75% in Finland (Nokelainen et al. 1993) and 93% in DM1 chromosomes from South Africa (Goldman et al. 1996), while it was found at only 18% worldwide on non-DM1 chromosomes. In our population we found that D19S63 was still in complete linkage disequilibrium with the disease allele, consistent with a recent founder event (Thibault et al. 1989). However, the founder effect that led to an increase in the number of rare mutations in Ch–SLSJ did not markedly affect the genetic diversity of non-disease chromosomes. No significant increase in linkage disequilibrium among markers on non-disease chromosomes was observed (Fig. 5) and frequencies of the microsatellite alleles were very similar to those characteristics of CEPH families (Fig. 2).

The DM1 chromosomes abundant in SLSJ and Charlevoix share the same core haplotype, extending from markers 61 and 85 (Fig. 1), some of which are also characteristic of other DM1 chromosomes found worldwide. This characteristic of the haplotypes in the immediate vicinity of the *DMPK* gene indicates that the DM1-predisposing chromosomes, rather than being due to recurrent mutations, are identical by descent and thus originated in the same ancestral event favoring CTG-expansion. Thus, the elevated frequency of DM1 in SLSJ and Charlevoix, greater by more than an order of magnitude than that in other Eurasian populations (Harper 1989) or even other regions of Quebec (Jack Puymirat, personal communication), must be due to a local founder effect rather than to de novo mutations. However, examination of the full-length DM1 haplotypes analyzed herein suggests that their common ancestor is older than the European colonization of Nouvelle-France and of the Charlevoix and SLSJ regions in particular. Among the three structural families of the DM1 associated haplotypes, only those of group A occur at a considerably elevated frequency of 1/730; those of group B occur at a frequency of 1/3,060, and those of group C at only 1/9,170 (as estimated from their respective contributions to our sample of 50 DM1 chromosomes, and using the overall DM1 frequency of

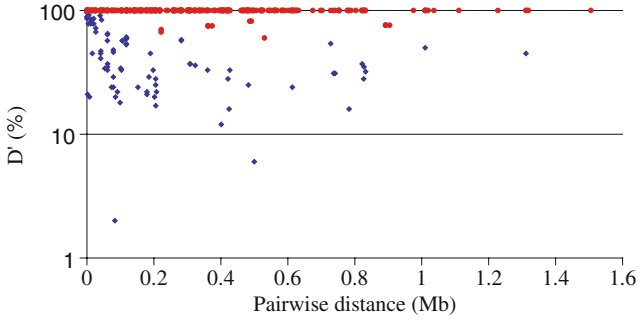

Fig. 5 Statistically significant pairwise D' values (in %) between biallelic markers in the general population (*in blue*) and on DM1-chromosomes (*in red*) (note the logarithmic ordinate scale)

1/550). If, for example, group C alone was present in Ch–SLSJ population, the resulting DM1 frequency of about 1/10,000 would be observed, no different from most other populations worldwide. The contribution of group B haplotypes is threefold higher than that of group C, but their frequency is still not markedly different from that of the European average. Furthermore, an increased kinship coefficient among B haplotypes carriers within the last three generations (Fig. 4) suggests either a very recent increase in their frequency (because of a larger number of children in previous generations) or simply their over-sampling among the 50 DM1 families analyzed.

Thus, it is only the group A of carrier-haplotypes, contributing between 70 and 80% of the affected chromosomes, which is found at the significantly increased frequency above that of the background population, and thus could be considered responsible for the founder effect of DM1 in Northeastern Quebec. In this group, the A1_21 appears as the founder haplotype, with the segment of common ancestral alleles extending for 0.5 Mb between markers 61 and 91 (Fig. 1) and the shared allele 21 of CA412 microsatellite, pointing to a recent common ancestry and reflecting the common origin of the settlers of Nouvelle France.

There are of course caveats to our conclusion about the age of the founder effect and about what it ultimately implies. For example, the map of Kong et al. (2002) shows that the *DMPK* gene is in a region with varying rates of recombination, and hence the judgment of an accurate estimation of the genetic distance can be problematic. The sample of chromosomes is relatively small, increasing the variance of the estimates even further. In addition, it could be argued that the 30-year generation time used to estimate the historical time of the DM1 introduction should be modified for muscular dystrophy sufferers, given the parental disease and possibility of a decreasing age of onset with the increasing CTG-expansion. Assuming a possible shortening of the intergeneration distance, the age of the founder effect given above would be an overestimate. On the other hand, in rapidly growing populations, estimates of the age of the founder effect using Eq. 1 will tend to underestimate the time since the introduction of a mutation (Hastbacka et al. 1992; Kaplan and Weir 1995; Labuda et al. 1997; Thompson and Neel 1997). For example, the age of the founder effect in the Acadians and in Ch–SLSJ, calculated from the proportion of non-recombined ancestral haplotypes of pseudo vitamin-D-deficiency rickets (PDDR), was 9.9 and 6.2 generations, respectively, i.e. less than the historically plausible 10–12 generations (Labuda et al. 1996). The founding of the major mutation of the autosomal recessive spastic ataxia of Charlevoix–Saguenay (ARSACS) can be estimated at 5–12 generations, based on the conserved proportions of 5.1 and 11.1 cM associated haplotypes, respectively (Richter et al. 1999). The founding of the SLSJ cytochrome oxidase deficiency was estimated at less than eight generations, based on the data from a 4 cM haplotype (Lee et al. 2001). Using Eq. 1, we estimated the founding of oculopharyngeal muscular dystrophy (OPMD) at an average of 11 generations, as calculated from the proportion of shared haplotypes extending between 0.9 and 9.5 cM to the left and the right of the disease locus, respectively (Brais 1998). In contrast to the diseases specific to Ch–SLSJ, OPMD is relatively prevalent in French-Canadians all over Quebec (Brais et al. 1998). These estimates are therefore consistent with the population history of Nouvelle-France, with the overall founder effect starting with the colonization in the 17th century, followed by regional founder effects such as that caused by the population bottleneck during the settlement of Charlevoix at the turn of the 17th century. Severe bottlenecks (Labuda et al. 1996), reinforced by intergenerational correlation in the effective family size (Austerlitz and Heyer 1998), can explain carrier rates in the order of 1/20 to 1/30 that characterize the autosomal recessive diseases enriched in Ch–SLSJ. For carrier frequencies in the range of 1/500 to 1/1,000, however, the introduction of the mutated chromosome does not have to coincide with the initial bottleneck, but rather could have occurred at a later time when the population grew in numbers.

The scenario where more than one carrier introduced the DM1 mutation into the population of Ch–SLSJ is not unexpected. For example, two distinct mutations in vitamin D1 alpha hydrolase are responsible for the unusually high frequency of the otherwise extremely rare PDDR in two Canadian populations of French descent, the Acadians and French-Canadians from Ch–SLSJ (Labuda et al. 1996; Wang et al. 1998). Two mutations also underlie autosomal recessive tyrosinemia type I that is highly prevalent in Ch–SLSJ, whereas three additional mutations were found elsewhere in Quebec (Demers et al. 1994; Scriver 2001). ARSACS in Ch–SLSJ is due to two distinct mutations (Richter et al. 1999). Similar observations have been made for other hereditary diseases (Scriver 2001). An increase in the frequency of cystic fibrosis (CF) in SLSJ (overall carrier rate of 1/15) is primarily due to the mutation $621 + 1G \rightarrow T$; carriers of this particular mutation occur in SLSJ at a frequency of 1/35 compared with only 1/500 in Quebec outside SLSJ (Rozen et al. 1992). As a result, the usually most frequent CF mutation, delta F508, contributes only 55% of all CF carriers in SLSJ, and its frequency in this area (1/27) is comparable to the European average (1/35).

In summary, the founder effect associated with an increase in the frequency of one "lucky" carrier chromosome is often accompanied by other distinct carrier chromosomes that remain at their usual background frequency.

A 4% contribution of a minor mutation to the overall carrier rate of ~1/22 (i.e. ~1/44 overall mutation frequency, as for example tyrosinemia type I, ARSACS, or cytochrome oxydase deficiency in Ch–SLSJ) corresponds to this mutation frequency of 1/1,100, and thus, a disease incidence due to this mutation alone is less than one in a million. Consequently, in a population with an

annual number of births of 10,000, one expects one patient per 100 years, in contrast to five such cases per year in a population with the overall mutation frequency of 1/44. In the case of a dominant disorder, the incidence increases linearly with the carrier frequency and therefore even rare single gene disorders have a chance to be recognized as unique clinical entities.

Acknowledgements We are thankful to Jack Puymirat for estimating DM1 prevalence in Quebec, to André Lescault for samples and earlier RFLP data, Julie Fortin and Pierre Lepage for sequencing of dHPLC variants, and Dominika Kozubska for secretarial assistance. S.B. had a studentship from the Fondation de l'Hôpital Sainte-Justine, while B.T. a postdoctoral fellowship of the Association Française contre les Myopathies. This study was supported by the Réseau de Médecine Génétique Appliquée of the Fonds de la Recherche en Santé du Québec (FRSQ).

References

Abbruzzese C, Costanzi Porrini S, Mariani B, Gould FK, McAbney JP, Monckton DG, Ashizawa T, Giacanelli M (2002) Instability of a premutation allele in homozygous patients with myotonic dystrophy type 1. Ann Neurol 52:435–441

Austerlitz F, Heyer E (1998) Social transmission of reproductive behavior increases frequency of inherited disorders in a young-expanding population. Proc Natl Acad Sci USA 95:15140–15144

Austerlitz F, Kalaydjieva L, Heyer E (2003) Detecting population growth, selection and inherited fertility from haplotypic data in humans. Genetics 165:1579–1586

Bengtsson BO, Thomson G (1981) Measuring the strength of associations between HLA antigens and diseases. Tissue Antigens 18:356–363

Brais B (1998) Oculopharyngeal muscular dystrophy: from phenotype to genotype. McGill University, Montreal

Brais B, Bouchard JP, Xie YG, Rochefort DL, Chretien N, Tome FM, Lafreniere RG, Rommens JM, Uyama E, Nohira O, Blumen S, Korczyn AD, Heutink P, Mathieu J, Duranceau A, Codere F, Fardeau M, Rouleau GA, Korcyn AD (1998) Short GCG expansions in the *PABP2* gene cause oculopharyngeal muscular dystrophy. Nat Genet 18:164–167

Brook JD, McCurrach ME, Harley HG, Buckler AJ, Church D, Aburatani H, Hunter K, Stanton VP, Thirion JP, Hudson T et al (1992) Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3′ end of a transcript encoding a protein kinase family member. Cell 68:799–808

Brunner HG, Korneluk RG, Coerwinkel-Driessen M, MacKenzie A, Smeets H, Lambermon HM, van Oost BA, Wieringa B, Ropers HH (1989) Myotonic dystrophy is closely linked to the gene for muscle-type creatine kinase (CKMM). Hum Genet 81:308–310

Brunner HG, Nillesen W, van Oost BA, Jansen G, Wieringa B, Ropers HH, Smeets HJ (1992) Presymptomatic diagnosis of myotonic dystrophy. J Med Genet 29:780–784

Demers SI, Phaneuf D, Tanguay RM (1994) Hereditary tyrosinemia type I: strong association with haplotype 6 in French Canadians permits simple carrier detection and prenatal diagnosis. Am J Hum Genet 55:327–333

Fu YH, Pizzuti A, Fenwick RG Jr, King J, Rajnarayan S, Dunne PW, Dubel J, Nasser GA, Ashizawa T, de Jong P et al (1992) An unstable triplet repeat in a gene related to myotonic muscular dystrophy. Science 255:1256–1258

Gauvreau D, Guerin M, Hamel M (1991) De Charlevoix au Saguenay: mesure et caractéristiques du mouvement migratoire avant 1911. In: Bouchard G, de Braekeleer M et al (eds) Histoire d'un génome. Population et génétique dans l'est du Québec. Presses de l'Université du Québec, Sillery, pp 145–159

Goldman A, Ramsay M, Jenkins T (1995) New founder haplotypes at the myotonic dystrophy locus in southern Africa. Am J Hum Genet 56:1373–1378

Goldman A, Krause A, Ramsay M, Jenkins T (1996) Founder effect and prevalence of myotonic dystrophy in South Africans: molecular studies. Am J Hum Genet 59:445–452

Gyapay G, Morissette J, Vignal A, Dib C, Fizames C, Millasseau P, Marc S, Bernardi G, Lathrop M, Weissenbach J (1994) The 1993–94 Genethon human genetic linkage map. Nat Genet 7:246–339

Harley HG, Brook JD, Floyd J, Rundle SA, Crow S, Walsh KV, Thibault MC, Harper PS, Shaw DJ (1991) Detection of linkage disequilibrium between the myotonic dystrophy locus and a new polymorphic DNA marker. Am J Hum Genet 49:68–75

Harley HG, Rundle SA, Reardon W, Myring J, Crow S, Brook JD, Harper PS, Shaw DJ (1992) Unstable DNA sequence in myotonic dystrophy. Lancet 339:1125–1128

Harper PS (1989) Gene mapping and the muscular dystrophies. Prog Clin Biol Res 306:29–49

Hastbacka J, de la Chapelle A, Kaitila I, Sistonen P, Weaver A, Lander E (1992) Linkage disequilibrium mapping in isolated founder populations: diastrophic dysplasia in Finland [published erratum appears in Nat Genet 1992 Dec 2(4):343]. Nat Genet 2:204–211

Jetté R, Gauvreau D, Guerin M (1991) Aux origines d'une région: le peuplement fondateur de Charlevoix avant 1850. In: Bouchard G, de Braekeleer M et al (eds) Histoire d'un génome. Population et génétique dans l'est du Québec. Presses de l'Université du Québec, Sillery, pp 75–106

Jomphe M, Tremblay M, Vézina H (2001) Analyses généalogiques à partir du fichier RETRO. Projet BALSAC, Chicoutimi, pp Document 1-C-204

Junghans RP, Ebralidze A, Tiwari B (2001) Does (CUG)n repeat in DMPK mRNA 'paint' chromosome 19 to suppress distant genes to create the diverse phenotype of myotonic dystrophy? A new hypothesis of long-range *cis* autosomal inactivation. Neurogenetics 3:59–67

Kaplan NL, Weir BS (1995) Are moment bounds on the recombination fraction between a marker and a disease locus too good to be true? Allelic association mapping revisited for simple genetic diseases in the Finnish population. Am J Hum Genet 57:1486–1498

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D (2002) The human genome browser at UCSC. Genome Res 12:996–1006

Kong A, Gudbjartsson DF, Sainz J, Jonsdottir GM, Gudjonsson SA, Richardsson B, Sigurdardottir S, Barnard J, Hallbeck B, Masson G, Shlien A, Palsson ST, Frigge ML, Thorgeirsson TE, Gulcher JR, Stefansson K (2002) A high-resolution recombination map of the human genome. Nat Genet 31:241–247

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

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 [letter]. Am J Hum Genet 61:768–771

Labuda M, Labuda D, Miranda C, Poirier J, Soong BW, Barucha NE, Pandolfo M (2000) Unique origin and specific ethnic distribution of the Friedreich ataxia GAA expansion. Neurology 54:2322–2324

Lavedan C, Hofmann-Radvanyi H, Boileau C, Bonaiti-Pellie C, Savoy D, Shelbourne P, Duros C, Rabes JP, Dehaupas I, Luce S et al (1994) French myotonic dystrophy families show expansion of a CTG repeat in complete linkage disequilibrium with an intragenic 1 kb insertion. J Med Genet 31:33–36

Lee N, Daly MJ, Delmonte T, Lander ES, Xu F, Hudson TJ, Mitchell GA, Morin CC, Robinson BH, Rioux JD (2001) A genomewide linkage-disequilibrium scan localizes the Saguenay-Lac-Saint-Jean cytochrome oxidase deficiency to 2p16. Am J Hum Genet 68:397–409

Lewontin RC, Kojima K (1960) The evolutionary dynamics of complex polymorphisms. Evolution 14:458–472

Luria SE, Delbrück M (1943) Mutations of bacteria from virus sensitivity to virus resistance. Genetics 28:491–511

Mahadevan MS, Amemiya C, Jansen G, Sabourin L, Baird S, Neville CE, Wormskamp N, Segers B, Batzer M, Lamerdin J et al (1993) Structure and genomic sequence of the myotonic dystrophy (DM kinase) gene. Hum Mol Genet 2:299–304

Martorell L, Monckton DG, Sanchez A, Lopez De Munain A, Baiget M (2001) Frequency and stability of the myotonic dystrophy type 1 premutation. Neurology 56:328–335

Mathieu J, De Braekeleer M, Prevost C (1990) Genealogical reconstruction of myotonic dystrophy in the Saguenay-Lac-Saint-Jean area (Quebec, Canada). Neurology 40:839–842

Neville CE, Mahadevan MS, Barcelo JM, Korneluk RG (1994) High resolution genetic analysis suggests one ancestral predisposing haplotype for the origin of the myotonic dystrophy mutation. Hum Mol Genet 3:45–51

Nokelainen P, Shelbourne P, Shaw D, Brook JD, Harley HG, Johnson K, Somer H, Savontaus ML, Peltonen L (1993) The DM mutation; diagnostic applications in the Finnish population. Clin Genet 43:190–195

Pouyez C, Lavoie Y (1983) Les Saguenayens. Introduction à l'histoire des populations du Saguenay. Les Presses de l'Université du Québec, Sillery

Richter A, Rioux JD, Bouchard JP, Mercier J, Mathieu J, Ge B, Poirier J, Julien D, Gyapay G, Weissenbach J, Hudson TJ, Melancon SB, Morgan K (1999) Location score and haplotype analyses of the locus for autosomal recessive spastic ataxia of Charlevoix–Saguenay, in chromosome region 13q11. Am J Hum Genet 64:768–775

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132:365–386

Rozen R, De Braekeleer M, Daigneault J, Ferreira-Rajabi L, Gerdes M, Lamoureux L, Aubin G, Simard F, Fujiwara TM, Morgan K (1992) Cystic fibrosis mutations in French Canadians: three CFTR mutations are relatively frequent in a Quebec population with an elevated incidence of cystic fibrosis. Am J Med Genet 42:360–364

Schneider MP, Erdmann J, Delles C, Fleck E, Regitz-Zagrosek V, Schmieder RE (2000) Functional gene testing of the Glu298Asp polymorphism of the endothelial NO synthase. J Hypertens 18:1767–1773

Scriver CR (2001) Human genetics: lessons from Quebec populations. Annu Rev Genomics Hum Genet 2:69–101

Segel R, Silverstein S, Lerer I, Kahana E, Meir R, Sagi M, Zilber N, Korczyn AD, Shapira Y, Argov Z, Abeliovich D (2003) Prevalence of myotonic dystrophy in Israeli Jewish communities: inter-community variation and founder premutations. Am J Med Genet 119A:273–278

Shen MR, Jones IM, Mohrenweiser H (1998) Nonconservative amino acid substitution variants exist at polymorphic frequency in DNA repair genes in healthy humans. Cancer Res 58:604–608

Thibault MC, Mathieu J, Moorjani S, Lescault A, Prevost C, Gaudet D, Morissette J, Laberge C (1989) Myotonic dystrophy: linkage with apolipoprotein E and estimation of the gene carrier status with genetic markers. Can J Neurol Sci 16:134–140

Thompson EA, Neel JV (1997) Allelic disequilibrium and allele frequency distribution as a function of social and demographic history. Am J Hum Genet 60:197–204

Tremblay M, Vezina H (2000) New estimates of intergenerational time intervals for the calculation of age and origins of mutations. Am J Hum Genet 66:651–658

Tremblay M, Lavoie EM, Houde L, Vézina H (2002) Demogenetic study of three populations within a region with strong founder effect. Eur J Hum Genet 10:183

Valdes AM, Slatkin M, Freimer NB (1993) Allele frequencies at microsatellite loci: the stepwise mutation model revisited. Genetics 133:737–749

Wang JT, Lin CJ, Burridge SM, Fu GK, Labuda M, Portale AA, Miller WL (1998) Genetics of vitamin D 1alpha-hydroxylase deficiency in 17 families. Am J Hum Genet 63:1694–1702

Whiting EJ, Tsilfidis C, Surh L, MacKenzie AE, Korneluk RG (1995) Convergent myotonic dystrophy (DM) haplotypes: potential inconsistencies in human disease gene localization. Eur J Hum Genet 3:195–202

Yamagata H, Miki T, Ogihara T, Nakagawa M, Higuchi I, Osame M, Shelbourne P, Davies J, Johnson K (1992) Expansion of unstable DNA region in Japanese myotonic dystrophy patients. Lancet 339:692

Yamaoka LH, Pericak-Vance MA, Speer MC, Gaskell PC Jr, Stajich J, Haynes C, Hung WY, Laberge C, Thibault MC, Mathieu J et al (1990) Tight linkage of creatine kinase (CKMM) to myotonic dystrophy on chromosome 19. Neurology 40:222–226

Zietkiewicz E, Yotova V, Jarnik M, Korab-Laskowska M, Kidd KK, Modiano D, Scozzari R, Stoneking M, Tishkoff S, Batzer M, Labuda D (1997) Nuclear DNA diversity in worldwide distributed human populations. Gene 205:161–171