Hum Genet (2005) 117: 119–132
DOI 10.1007/s00439-005-1297-9

ORIGINAL INVESTIGATION

**Hélène Vézina · Francine Durocher · Martine Dumont
Louis Houde · Csilla Szabo · Martine Tranchant
Jocelyne Chiquette · Marie Plante · Rachel Laframboise
Jean Lépine · BCLC Haplotype Group · Heli Nevanlinna
Dominique Stoppa-Lyonnet · David Goldgar
Peter Bridge · INHERIT BRCAs · Jacques Simard**

Molecular and genealogical characterization of the R1443X *BRCA1* mutation in high-risk French-Canadian breast/ovarian cancer families

Received: 26 November 2004 / Accepted: 6 February 2005 / Published online: 10 May 2005

Abstract The Quebec population contains about six-million French Canadians, descended from the French settlers who colonized "Nouvelle-France" between 1608 and 1765. Although the relative genetic contribution of each of these founders is highly variable, altogether they account for the major part of the contemporary French-Canadian gene pool. This study was designed to analyze the role of this founder effect in the introduction and diffusion of the *BRCA1* recurrent R1443X mutant allele. A highly conserved haplotype, observed in 18 French-Canadian families and generated using 17 microsatellite markers surrounding the *BRCA1* locus, supports the fact that the R1443X mutation is a founder mutation in the Quebec population. We also performed haplotyping analysis of R1443X carriers on 19 other families from seven different nationalities; although the same alleles are shared for three markers surrounding the *BRCA1* gene, distinct haplotypes were obtained in four families, suggesting multiple origins for the R1443X mutation. Ascending genealogies of the 18 French Canadian families and of controls were reconstructed on an average depth of 10 generations. We identified the founder couple with the highest probability of having introduced the mutation in the population. Based on the descending genealogy of this couple, we detected the presence of geographical concentration in the diffusion pattern of the mutation. This study demonstrates how molecular genetics and demogenetic analyses can complement each

Other members of the BCLC Haplotype Group involved in this study are listed in Appendix 1

Other members of INHERIT BRCAs involved in this study are listed in Appendix 2

H. Vézina and F. Durocher contributed equally to this work and should be regarded as joint first author

H. Vézina · L. Houde
Interdisciplinary Research Group on Demography and Genetic Epidemiology (GRIG), University of Quebec at Chicoutimi, Chicoutimi, Canada, G7H 2B1

F. Durocher (✉) · M. Dumont · M. Tranchant · J. Simard
Cancer Genomics Laboratory,
CHUL Research Center, CHUQ, Laval University,
T2-53, 2705 Laurier Blvd, Sainte-Foy, QC, Canada, G1V 4G2
E-mail: Francine.Durocher@crchul.ulaval.ca
Tel.: +1-418-6542296
Fax: +1-418-6542761

C. Szabo · D. Goldgar
Genetic Epidemiology Unit,
International Agency for Research on Cancer,
Lyon, Cedex 08, France,

J. Chiquette
Clinique des maladies du sein Deschênes-Fabia,
Hôpital du Saint-Sacrement, QC, Canada, G1S 4L8

M. Plante
Service de gynécologie, CHUQ,
L'Hôtel-Dieu de Québec,
QC, Canada, G1R 2J6

R. Laframboise
Service de médecine génétique, CHUQ,
Pavillon CHUL, QC, Canada, G1V 4G2

J. Lépine
Center hospitalier régional de Rimouski,
Rimouski, Canada, G5L 5T1

H. Nevanlinna
Department of Obstetrics and Gynecology,
Helsinki University Central Hospital,
Helsinki, 00029, Finland,

D. Stoppa-Lyonnet
Department of Genetics, Institut Curie,
Paris, Cedex 05, France,

P. Bridge
Molecular Diagnostic Laboratory,
Alberta Children's Hospital,
Calgary, Canada, T2T 5C7

J. Simard
Canada Research Chair in Oncogenetics,
Faculty of Medicine,
Laval University, Canada

other to provide findings that could have an impact on public health. Moreover, this approach is certainly not unique to breast cancer genetics and could be used to understand other complex traits.

Introduction

Since the discovery of the *BRCA1* and *BRCA2* genes (Miki et al. 1994; Wooster et al. 1995; Tavtigian et al. 1996), extensive mutation detection has been performed in various populations [see Breast Cancer Information Core database (BIC): http://www.nhgri.nih.gov/Intramural research/Lab_transfer/Bic/]. Current knowledge indicates that *BRCA1* mutations account for the majority of familial breast/ovarian cancer cases (Ford et al. 1998). Many different mutations in *BRCA1* have been identified in high-risk families with multiple cases of breast and ovarian cancer, and it is now well known that some recurrent mutations occur in specific populations (Simard et al. 1994; Tonin et al. 1995; Johannesdottir et al. 1996; Thorlacius et al. 1996; Gayther et al. 1997; Peelen et al. 1997; Petrij-Bosch et al. 1997; Ramus et al. 1997; Szabo and King 1997; Thorlacius et al. 1997; Tonin et al. 1998; Gorski et al. 2000; Sarantaus et al. 2001). In many instances, haplotype analysis of recurrent mutations has shown evidence of a founder effect, thus explaining the higher frequency of these mutations (Simard et al. 1994; Berman et al. 1996; Neuhausen et al. 1996, 1998; Dorum et al. 1997; Liede et al. 2000; Sarantaus et al. 2000; Barkardottir et al. 2001). Studies based on these founder populations have also proven useful to achieve characterization of these mutations. Establishing the origin of population-specific mutations in *BRCA1* and *BRCA2* is a critical step towards providing accurate counseling and developing inexpensive mutation detection in a specific founder population.

The Quebec population (7.5 million) contains about six-million French Canadians, who are descendants of 10,000 immigrants, coming mostly from France, who settled in ''Nouvelle-France'' between 1608 and 1760. Although the relative genetic contribution of each of these founders is highly variable, altogether they account for the major part of the contemporary French-Canadian gene pool (Charbonneau et al. 2000). Following the British conquest of 1759, the population expanded in a context of relative isolation and rapid growth, which contributed to amplify the consequences of the founder effect (Bouchard and de Braekeleer 1990).

This founder effect is also detectable through the observed frequency of some otherwise rare disorders, which are found at an elevated frequency in the Quebec population, particularly in some of the eastern regions of the province (see for a review Scriver 2001). In many instances, an increase in frequency of one or a few mutations associated to these disorders accounts for the observed incidence among French Canadians (Davignon and Roy 1993; Engert et al. 2000; Dupré et al. 2003; Mootha et al. 2003). There were numerous emigrants among French Canadians, especially at the end of the 19th and at the beginning of the 20th centuries, due to industrialization and high rates of unemployment that forced thousands to emigrate to other parts of Canada and to the United States (McInnis 2000). Nowadays several millions of North Americans have ancestry among French Canadians. It is thus not surprising that some of these mutations have also been detected elsewhere in Canada and in the United States and, that in most cases, the carriers have been found to be of French-Canadian descent (Triggs-Raine et al. 1995; Bayleran et al. 1996; Couture et al. 2001).

As part of an ongoing research program started in 1996, aimed at the identification and characterization of *BRCA1* and *BRCA2* mutant alleles in the French-Canadian population of Quebec, more than 250 high-risk families ascertained on the basis of: (1) three first-degree relative or (2) four first- or second-degree relative affected with breast and/or ovarian cancer cases were recruited. Among the 62 families in which a mutation in *BRCA1* or *BRCA2* was identified, 18 families were found to carry the same *BRCA1* R1443X mutation, suggesting the possibility of a founder mutation. This mutation was first reported by Mary-Claire King's group in family 81, which is of French-Canadian descent (Friedman et al. 1995). In order to get a better understanding of the origins of the R1443X mutation and of the role of the founder effect in the introduction and spread of this mutation in the French-Canadian population, we reconstructed haplotypes using 17 microsatellite markers surrounding the *BRCA1* region. We also compared the extent of these haplotypes obtained in various families carrying the R1443X allele in different populations of European ancestry. Using genealogical data, we attempted identification of the founders who are the most likely to have introduced this mutant allele and we investigated the diffusion pattern of the mutation on the Quebec territory to detect the presence of a geographical concentration of the mutation in one or more regions of the province. In this paper, we report on the extent and divergence of haplotypes in a total of 37 R1443X families with different nationalities. Nineteen of them are of French-Canadian ancestry, one is a Hutterite family from Canada having Russian ancestry, ten families are from France, three from Finland, two from the United Kingdom, one from Belgium and one from the Netherlands. We also present the results of extensive genealogical analyses on 18 French-Canadian families.\

Subjects and methods

Ascertainment of families and DNA extraction

As part of an ongoing study to estimate the frequency and penetrance of mutations in the *BRCA1* and *BRCA2* genes, a total of 251 (counted as 256 in total because five families had a history of cancer on both sides that met the criteria of eligibility mentioned below) high-risk French-Canadian breast/ovarian cancer families were recruited

in all socio-sanitary regions of Quebec according to strict criteria: (1) individuals from high-risk families with four cases of breast and/or ovarian cancer diagnosed in first- or second-degree relatives; (2) families presenting three cases of breast and/or ovarian cancer cases in 1st degree relatives; (3) families bearing a mutation already identified in the *BRCA1* or *BRCA2* gene. All individuals participating in the present study had to be at least 18 years of age and mentally capable. Whenever possible, the diagnoses of breast and/or ovarian cancer were confirmed by obtaining a pathology report. This research program is composed of a network of referring clinicians across the Quebec province. Moreover, clinicians from seven participating hospitals were directly involved in this integrated clinical research program and were responsible for the disclosure of genetic test results. The study was approved by eight ethics committees, corresponding to the different institutions participating in this research project. More details regarding experimental and clinical procedures as well as the INHERIT BRCAs research program have been described elsewhere (Godard et al. 2003, Simard et al., manuscript in preparation). After obtaining a signed informed consent from each participant 40 ml of blood was drawn and genomic DNA was extracted using the guanidine hydrochloride-proteinase k method (Jeanpierre 1987) and the QIAmp Maxi blood kit according to the manufacturer's instructions (QIAGEN, Mississauga, Canada).

A panel of 24 mutations, including: 16 frameshift mutations, five nonsense mutations and three sequence variants for which clinical significance was not well-defined or still under investigation (two missense mutations and one in-frame deletion), observed and/or reported in the French-Canadian population was tested systematically in our cohort. For each family, or branch of a family that met criteria and which received an inconclusive result, at least one DNA sample of an individual with breast and/or ovarian cancer or of an obligate carrier has been sent to Myriad Genetic Laboratories, Salt Lake City, Utah, USA for full-length *BRCA1*/*BRCA2* sequencing. Testing services were performed according to the Memorandum of Understanding (MOU) with the National Cancer Institute (NCI) for NCI-funded research testing services for *BRCA1* and *BRCA2* (project no. NCI 173). A confirmation test for all conclusive results (carrier or non-carrier of a known familial mutation) was performed on a 2nd blood sample by the Molecular Diagnostic Laboratory, Alberta Children's Hospital, Calgary. The identification of *BRCA1* R1443X mutation in families of other nationalities included in this study was performed by a variety of techniques, primarily direct sequencing, PTT, DGGE or SSCP analyses, which was always confirmed by direct sequencing.

Microsatellite genotyping and haplotype construction

A total of 37 families of seven different nationalities presenting a family history of breast and/or ovarian cancer and carrying the *BRCA1* nonsense R1443X mutation were genotyped for 14 markers surrounding the *BRCA1* region and three markers within the *BRCA1* gene. The 17 microsatellite markers genotyped are located on chromosome 17q and span approximately 13 kb of genomic DNA. Genotyping was carried out at two centers. The families collected by the Cancer Genomics Laboratory in Quebec City (family members of kindreds 4616, 4049, 5582, 5071, 6079, 3258, 2551, 4609, 3594, 5344, 3762, 4315, 3587, 5099, 3531, 4966, 4910 and 3356), the Molecular Diagnostic Laboratory of Alberta Children's Hospital in Alberta (the Hutterite family), the Research Laboratory in Helsinki (families 286, 425 and 5004) and the Curie Institute in Paris (families F347, F390, F1203, F1375, F1531 and F1628) were genotyped in Quebec City, while other families were genotyped in Lyon, at the International Agency for Research on Cancer (F1, F31, F32, F35, F85, B35, B88 and D99). A similar protocol was followed by both centers, using several internal controls to assure accuracy of allele calling. All 17 polymorphic microsatellite repeat markers genotyped were assayed by PCR using standard procedures. Briefly, PCRs were carried out on 100 ng genomic DNA using 20 μM of dCTP, dGTP, dTTP, 4 μM dATP, 2.5% final DMSO, 1.5 mM $MgCl_2$ and 0.5 μM of forward and reverse primers at 96°C, 8 min and 94°C for 2 min. Following a hot start, 1.5 U Taq DNA polymerase (Perkin-Elmer, Mass., USA), 1.5 μCi of ^{35}S-dATP and 10× PCR buffer were added for a final volume of 20 μl. Amplification conditions on a Perkin-Elmer, model 9700 were as follows: 35 cycles of 94°C for 40 s and 56°C for 30 s, followed by 8 min at 72°C. For markers D17S1859, 776, 855, 1323, 1327, 1326, EDH17B2 and 1185, the annealing was performed at 58°C, while the annealing temperature was 59°C for marker D17S1325. Radiolabeled PCR products were electrophoresed on standard 6% polyacrylamide denaturing gels. Allele sizes were matched according to the genotype of CEPH reference individual 1347–02 and a French-Canadian individual carrying the R1443X *BRCA1* mutation. Both these samples were used as controls on each gel. Moreover, for each marker, two PCR products coming from individuals showing homozygous patterns, were sub cloned and sequenced to determine the nature of repeats and size-range of alleles observed in French-Canadian families. In order to have an idea of the allele frequency in the general population, a series of non-carriers, known not to carry the haplotype associated with the R1443X mutation, was selected ($n = 33$ individuals) among the 18 French-Canadian families.

The average heterozygosity of microsatellite markers genotyped is 0.73, the lowest being 0.40 for D17S776. Of the 17 markers used in this study, three fall within intronic portions of the *BRCA1* gene: D17S1323, D17S1322 and D17S855 are located in intron 12, 19 and 20, respectively, while the mutation R1443X lies in exon 13. Haplotype construction was done manually. In most instances, the haplotype associated with the mutation could be inferred accurately as an average of three carriers per family were

genotyped as specified in Fig. 1. In the French-Canadian families, approximately six individuals (carriers and known non-carriers) per family were genotyped.

Genealogical reconstruction

In each of the 18 French-Canadian families carrying the R1443X mutant allele, we selected one individual who was a confirmed carrier. We considered as carriers participants found to be carriers through DNA analysis and those with a carrier child and with a confirmed non-carrier spouse. Each case was matched to three different controls based on place and year (± 2 years) of marriage of their parents. Fifty-four controls were thus selected in the BALSAC-RETRO database (described below) for comparison purposes. Our goal was to determine whether or not the genealogical characteristics of the carrier individuals were related to their status or if they merely depicted the underlying population structure. As controls were not actual participants to the project, we could not verify their carrier status and must consider that the carrier frequency in this group is that of their population of origin. However, since that frequency, albeit unknown, is low and the sample size is quite small, the probability of having a carrier in our control group is also quite low.

Genealogical reconstruction relied on the BALSAC population register and its peripheral databases such as the BALSAC-RETRO database which contains linked genealogical information on nearly 340,000 individuals married in Quebec between the beginning of the 17th century and the present day (Bouchard 2003). Complementary sources, such as marriage repositories and family dictionaries, were also consulted. Reconstruction was performed as far back as sources allowed: in most instances, lineages were traced back to the first immigrants to Canada (12 generations ago). All genealogies were

Fig. 1 Haplotypes of the *BRCA1* region for the different families harboring the R1443X mutation

Comparison of haplotypes segregating with the R1443X mutation in BRCA1 in French-Canadian and European families

Origin of Families		Family number	Number of R1443X-carriers genotyped	Haplotypes: Microsatellite Markers(Cen ▶ Tel) (D17S...) — 933	800	846	776	1185	EDH17B2	1320	1328	855	1322 ▼ R1443X	1323	1327	1326	1325	1329	791	1859
CEPH Mother 1347-2				[189/193]	[170/176]	[229/233]	120	[221/225]	209	[173/175]	[199/229]	[145/149]	[122/125]	[150/156]	[132/156]	[88/98]	[212/214]	[186/190]	[181/191]	[177/183]
Canada	French-Canadian	4616	4	193	172	241	120	213	205	171	199	149	122	150	132	102	202	184	191	181
		3258	5	193	172	241	120	213	205	171	199	149	122	150	132	102	202	184	191	181
		2551	2	193	172	241	120	213	205	171	199	149	122	150	132	102	202	184	191	181
		4910	1	[191/193]	[172/174]	[229/241]	[118/120]	[213/225]	[205/209]	171	199	149	122	[150/154]	132	[98/102]	[186/202]	[184/186]	[191/193]	[177/181]
		4315	1	[191/193]	[168/172]	[233/241]	[118/120]	[209/213]	[205/209]	[171/179]	199	[149/153]	122	150	132	[98/102]	[198/202]	[184/190]	[187/191]	[181/183]
		3587	3	193	172	241	120	213	205	171	199	149	122	150	132	102	202	184	[189/191]	[181/183]
		3594	1	-	172	[229/241]	[120/126]	[213/217]	[205/209]	[171/177]	199	[143/149]	122	150	132	[100/102]	[198/202]	[184/192]	[187/191]	[177/181]
		3356	2	193	172	241	120	213	205	171	199	149	122	150	132	102	202	184	191	181
			1	189	172	241	120	213	205	171	199	149	122	150	132	102	202	184	191	181
		5099	1	[189/191]	[172/176]	[233/241]	[120/122]	[209/213]	[205/209]	171	[199/205]	[145/149]	[122/131]	[150/156]	[132/168]	[88/102]	202	[184/190]	[183/191]	181
		4049	4	197	172	241	120	213	205	171	199	149	122	150	132	102	202	184	191	181
		6079	7	195	172	241	120	213	205	171	199	149	122	150	132	102	202	184	191	181
		4609	3	199	172	241	120	213	205	171	199	149	122	150	132	102	202	184	191	181
		5582	6	193	172	241	120	213	205	171	199	149	122	150	132	102	202	184	191	181
			2	[189/203]	172	241	120	213	205	171	199	149	122	150	132	102	202	184	191	181
			2	197	172	241	120	213	205	171	199	149	122	150	132	102	202	184	191	181
			1	193	172	241	120	213	205	171	199	149	122	150	132	102	202	184	163	177
		5071	5	193	172	241	120	213	205	171	199	149	122	150	132	102	202	184	195	181
		5344	3	[193/199]	172	241	[118/120]	213	[205/209]	[171/175]	199	[149/153]	122	150	132	[98/102]	[198/202]	184	183	[177/181]
		3531	1	189	172	241	120	[205/213]	205	171	199	149	122	150	132	102	202	184	191	181
			1	197	172	241	120	[205/213]	205	171	199	149	122	150	132	102	204	184	191	181
		3762	1	[189/197]	178	[241/253]	[118/120]	[209/213]	[205/209]	[171/179]	[199/205]	149	122	150	132	102	198	[186/188]	[177/183]	181
		4966	3	193	176	241	120	213	205	171	199	149	122	150	132	102	214	184	185	181
United States	French-Canadian	B88	3	189	172	-	120	213	[205/209]	171	199	[149/151]	122	150	132	102	[198/202]	[184/186]	[189/191]	181
			1	189	172	-	[114/120]	213	209	171	199	[149/151]	122	150	132	102	[196/198]	[184/186]	[189/191]	181
France	Brittany	347	1	197	172	241	120	221	209	171	199	149	122	150	132	98	194	188	189	181
			2	197	172	241	120	221	209	171	199	149	122	150	132	98	194	188	191	177
			1	193	172	241	120	221	209	171	199	149	122	150	132	98	194	188	191	177
		1203	2	189	172	241	120	221	209	171	199	149	122	150	132	98	194	188	191	181
		1375	1	189	172	[241/253]	120	221	209	171	199	149	122	150	132	[98/102]	[194/196]	[188/190]	[179/191]	-
		1628	2	[191/193]	172	229	124	[209/217]	209	175	199	149	122	150	132	98	196	188	191	177
	Normandy — Rouen	F35	1	[191/193]	174/176	-	[120/130]	[217/221]	[209/211]	171	[199/205]	[143/151]	[122/125]	[150/152]	132	[98/102]	[192/200]	[184/188]	[185/191	[177/181]
		1531	1	[189/191]	[172/178]	[233/241]	118	[213/221]	[205/209]	[175/179]	[199/223]	[143/147]	[122/125]	[150/156]	[132/154]	[90/98]	[188/192]	188	165	183
		F31	3	189	172	-	114	213	209	179	199	[143/149]	122	150	132	98	192	188	167	183
	Clermont-Ferrand	F1	1	[189/203]	[170/172]	-	[114/120]	-	209	[171/179]	199	143	122	150	132	[98/102]	[192/204]	188	[167/189	183
	Belgium	F32	2	191	[172/176]	-	114	213	209	179	199	143	122	150	132	98	192	188	189	181
	Netherlands	D99	1	[189/193]	[168/170]	-	114	[205/217]	209	[171/179]	199	[143/147]	122	150	132	98	[192/200]	[186/188]	[189/197	[177/181]
	Canadian Hutterites (Russian ancestry)		1	189	[170/182]	237	[120/130]	225	209	[175/179]	199	[143/151]	122	[148/150]	132	[98/100]	[198/202]	[186/188]	[163/185]	[181/183]
	Finland	286	3	193	176	241	120	221	209	173	223	147	122	150	132	98	196	186	189	177
		5004	2	203	176	241	120	221	209	173	223	147	122	150	132	98	196	186	189	177
		425	2	197	176	[233/241]	120	221	209	173	223	147	122	150	132	98	196	186	185	181
France	Brittany	390	3	[189/193]	174	229	120	217	209	171	199	143	125	152	132	100	196	184	165	177
	Normandy	F85	2	191	170	-	120	217	209	171	199	[143/151]	125	152	132	100	[198/200]	184	191	181
United-Kingdom	Wales	B35	3	193	172	-	118	221	209	[171/175]	[199/205]	145	125	158	158	88	204	184	185	[181/183]
	United-Kingdom	B155	1	[189/191]	[174/178]	-	120	[209/225]	209	[169/175]	[201/207]	[145/149]	[125/128]	[150/156]	[154/158]	88	[214/222]	[184/186]	[179/189]	179

submitted to various validation and consistency checks in order to minimize the occurrence of false links due for example to illegitimacy or adoptions (Bouchard 1986).

Since deep-rooted pedigrees have very complex structures, descriptive parameters are used to ensure that the groups of genealogies have structures allowing for calculations and comparisons. These parameters include the completeness index, which gives, for each generation level, the proportion of ancestors, which have been successfully identified. Theoretically, within a genealogy, at a given generation x, 2^x ancestors can be found (for instance generation 2 is the grandparents' generation and contains 2^2 or four ancestors). However, due to availability of genealogical sources, it is not always possible to locate all ancestors and this loss of information is measured by the completeness index. This index is also used to calculate the mean generation level at which the lineages in the genealogies come to an end (average genealogical depth).

For each of the 27,227 ancestors found in the genealogies, we calculated the number of carriers and controls to which they were related, the total number of times they appeared in the genealogies of each group—since many of them appear more than once in a genealogy—and their genetic contribution to the carriers' group and to the three control groups. Calculation of the genetic contribution of an ancestor to a set of probands S is given by

$$\sum_S \sum_P (1/2)^g$$

where P is the set of paths from the ancestor to the proband and g is the length of the path (number of generations).

Demogenetic analyses

Based on the observed haplotypes, we hypothesized that the R1443X mutation was introduced by a single founder (or founder couple) and that the sharing of this common ancestor by the carriers should lead to an increase in their kinship level. To verify this, the mean kinship coefficient of carriers and of each control group was calculated by identifying, for each pair of individuals, all known common ancestors over all known genealogical links and dividing the sum of these coefficients by the total number of possible pairs (for a group of 18 subjects, this number is 153). Calculations were performed for each generation level, starting at the third generation and were based on the formula described by Thomson (1986). As we made sure no controls were related among themselves or with carriers before the third generation, no kinship links could be found before that. Statistical comparisons of the results were performed using the statistic proposed by Hauck and Martin (1984). The P values for unilateral testing were obtained by the bootstrap method (Efron and Tibshirani 1993) with 5,000 repetitions to provide a sufficient precision for a 5% level.

The next set of analyses aimed at the identification of the most likely introducer of the mutation. Here, based on our knowledge of the population structure, our hypothesis was that this ancestor: (1) should have a high frequency among carriers in terms of number of appearances and of genetic contribution and (2) should be more specific to the carriers, but not exclusive to them, as we have to keep in mind that at each generation the mutation is transmitted on average to half of the children of a carrier individual and that therefore the introducer may have more descendants in the population who are non-carriers than carriers. To identify the ancestors complying with these criteria, we selected those who appeared in more than 80%, that is more than 14 of the 18 carriers' genealogies. Of course, based on the one mutation-one founder hypothesis, the introducer must be related to all carriers and therefore be present in all genealogies. However, we cannot discard the possibility that the introducer would not be present in one or a few carriers' genealogies because of a false or missing genealogical link due either to an error in the reconstruction process, to a mistake in the consulted source of information, or to a false paternity or an undeclared adoption in a lineage. After selecting these ancestors, genealogical links among them were examined in order to cluster them in ancestral family groups. Finally, we evaluated the frequency and the specificity of these groups by comparing their number of appearances and their genetic contribution to carriers and to controls.

Results

Haplotype analysis

Results of haplotype analysis are shown in Fig. 1. For families in which more than one haplotype is present, all the haplotypes are listed along with the number of people within the family bearing that haplotype. Also, for families originating from France, it was possible to ascertain from which region of France they came from, which greatly facilitated haplotype analysis. A total of 99 carriers were genotyped, representing nearly three individuals per family. Also, for 15 of the 18 French-Canadian families, 60 non-carriers were also genotyped and therefore haplotypes could be phased accurately. Four different haplotypes seem to emerge quite clearly for families of French-Canadian, Finnish and French (Breton) origins. Indeed, all French-Canadian families share a haplotype that extends over 9.3 cM (~12,956 kb). This highly conserved haplotype observed in 60 R1443X carriers from 18 French-Canadian families confirms that the R1443X mutation is a founder mutation in the Quebec population. The family B88 (equivalent to family 81 in Friedman et al. 1995) from the United States reported to be of French-Canadian descent also shares a long haplotype with the other French-Canadian families. The same phenomenon is observed for the three families from Finland, consistent with a previous finding of a core

haplotype of 7.9 cM (Sarantaus et al. 2000), as well as for three of the four families from Brittany (families 347, 1203 and 1375). To some extent, two families from Normandy (families 1531 and F31), family F1 from Clermont-Ferrand, the Belgian family F32 and family D99 from the Netherlands also display a shared haplotype. However, for the majority of families analyzed, a core haplotype consisting of three markers, namely D17S1322, D17S1323 and D17S1327, where the R1443X mutation lies between the first two markers, is well conserved. These results are consistent with a common origin for this mutant allele in these populations. The only outliers are two French families (one from Brittany and one from Normandy) and two families from the United Kingdom (families B35 and B155), which do not share these three conserved markers. Distinct haplotypes are indeed obtained in these four families, thus suggesting multiple origins for the R1443X mutation.

Descriptive parameters of the genealogies and kinship measurements

As shown in Table 1, genealogies were reconstructed to an average depth of ten generations. Between 63,000 and 69,000 ancestors were identified in each of the four groups (carriers and three control groups). Because of the population structure, we have to take into consideration the fact that many ancestors appear more than once in the genealogies. In fact, the number of distinct ancestors is between 11,000 and 14,000 in each group corresponding to a mean of five appearances for each ancestor. Moreover, only one third of the distinct ancestors of a group are specific to this group, while the remaining two thirds are found in at least one other group, reflecting once again the complexity of the population structure. Finally, as one would expect, an inverse relationship is observed between the number of ancestors and the number of genealogies in which they appear: within each group, the majority (about 60%) of distinct ancestors appear in only one genealogy while less than 1% appear in more than 15 genealogies.

Table 1 Descriptive parameters of carriers' and controls' genealogies

	Cases	Controls		
		1	2	3
Number of genealogies	18	18	18	18
Mean depth of lineages (generations)	9.9	9.9	9.8	9.9
Total ancestors	65,085	61,965	61,511	65,327
Distinct ancestors	13,210	11,788	11,784	13,168
Ancestors specific to the group (%)	32.6	30.8	29.1	32.6
Ancestors non-specific (%)	67.4	69.2	70.9	67.4
Ancestors (%) appearing				
In one genealogy	60.9	62.9	61.8	61.9
In 2–14 genealogies	38.0	36.5	37.5	37.6
In 15–18 genealogies	1.1	0.6	0.7	0.6

Source: BALSAC, RETRO file

Figure 2 shows kinship curves for R1443X carriers and controls. Kinship coefficients are higher among carriers than among each of the three control groups for all generations. Since the three control groups were selected in the same population using the same criteria, the variability observed in the controls' kinship curves simply reflects sampling distribution and the mean values for the three groups were used for comparison. The difference between carriers and controls builds up for the most part from the fifth to the seventh generation, indicating that it is mainly explained by kinship links attributable to common ancestors found at these genealogical depths. The observed differences remain more or less constant thereafter and are statistically significant (P values ranging from 0.003 to 0.032) from the sixth to the 12th generation. We stopped calculations at the 12th generation since only a few more genealogical links could be identified corresponding to a negligible increase in the kinship coefficient.

Identification of the most likely introducers of the R1443X mutation

Genealogical links among the 102 ancestors appearing in more than 15 of the 18 carriers' genealogies were examined and these ancestors were clustered in 17 ancestral groups. The number of genealogies where these ancestral groups appear were then compared for carriers and for each of the three control groups. Results are shown on the left side of Table 2. Within each ancestral group, in an attempt to identify the most likely introducer, we then selected the ancestor (or couple ancestor) with the highest genetic contribution (GC) to cases and we compared this GC to the mean GC of this same ancestor to the three control groups combined. GCs to cases and controls, as well as the difference between these two values, are on the right side of Table 2. Based on frequency and specificity, we can see quite clearly that ancestral group 1C emerges as the most likely founder group for the R1443X mutation as: (1) it appears in all of the 18 carriers' genealogies but in 8–10 of the genealogies of the control groups, and (2) it has the third highest contribution to carriers, while showing the highest difference between carriers and controls in terms of GC.

In order to obtain more information on the ancestors and on the genealogical paths linking the 18 carriers and the founder couple of ancestral group 1C, we drew the descending pedigree shown in Fig. 3. The founder couple married in Quebec City in 1671. The husband migrated to Nouvelle-France coming from Portugal while his wife was an immigrant from France. The 18 carriers are linked to this founder couple through three of their offspring, one of them being linked to 16 of the 18 carriers. For all couples of the pedigree, regions and dates of marriage are indicated. Regions were grouped according to their localization in the province except for the eastern Quebec region of Bas-Saint-Laurent (*in red*),

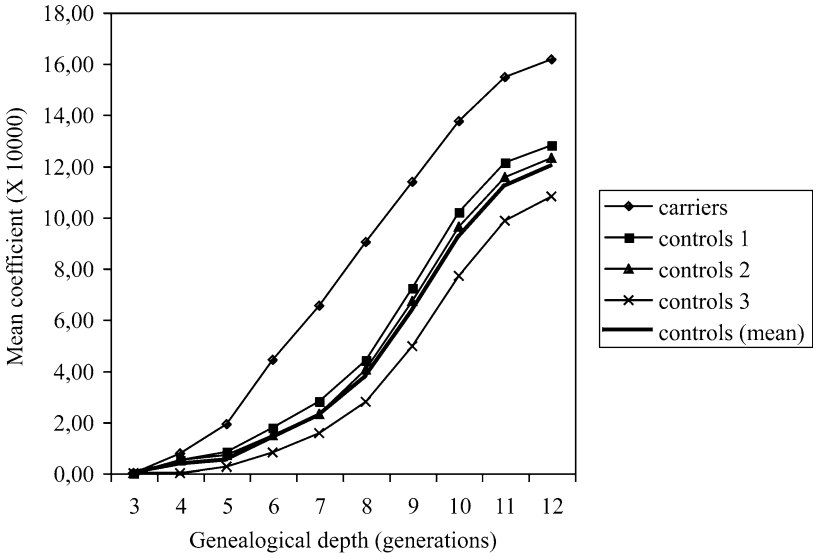

Fig. 2 Mean kinship coefficients for the R1443X carriers and for the controls. The mean value for the three control groups is shown by a *solid line*. Kinship coefficients were calculated among all pairs ($n = 153$) of individuals within each group and for all generations from the third to the 12th

which was singled out because of its high number of appearances in the pedigree.

Diffusion of the mutation in the Quebec population

We identified the founder couple with the highest probability of having introduced the R1443X mutation in the Quebec population. The 18 carriers can be traced to this couple through one son and two daughters, who can therefore be considered as obligate carriers. In Fig. 4, we measured the genetic contributions of these offspring to the regional populations of Quebec in order to get a better understanding of the spread in time and space of the mutation, and to detect the presence of geographical concentration in the diffusion pattern of the mutation. Genealogical data for these calculations come from an ongoing project on the demogenetic characteristics of the regional populations of Quebec (Tremblay et al. 2001). In the course of this project, 2,600 genealogies of individuals married in the various regions of Quebec between 1935 and 1974 were reconstructed. It can be seen that the mean genetic contributions of the son and the two daughters are quite variable

Table 2 Description of the 17 ancestral groups where founders have the highest probability of having introduced the R1443X mutation in the French-Canadian population

Ancestral group	Number of genealogies where group is found				Genetic contribution[a] (×10,000)		
	Among carriers	Among controls			To carriers	To controls[b]	Difference
		1	2	3			
1A	18	15	15	16	1060.2	959.1	101.1
1B	18	17	18	18	914.5	1189.2	−274.7
1C	18	8	8	10	827.6	315.8	511.9
1D	18	17	18	17	818.5	743.3	75.2
1E	18	17	16	17	751.0	792.3	−41.3
1F	18	17	18	15	691.5	671.8	19.7
1G	17	15	17	17	659.2	653.9	5.3
1H	16	13	15	13	618.3	453.9	164.4
1I	16	10	14	12	592.0	924.1	−332.0
1J	18	14	14	12	587.8	450.0	137.7
1L	17	17	18	17	574.6	822.1	−247.5
1M	16	12	13	10	499.3	625.8	−126.5
1N	17	16	15	18	451.0	525.9	−74.9
1P	17	13	18	15	429.1	623.6	−194.5
1R	16	12	11	11	329.6	154.6	175.0
1S	16	10	10	10	287.5	145.3	142.2
1V	17	11	12	11	141.0	77.5	63.5

Source: BALSAC, RETRO file
[a]Based on the ancestor with the highest contribution to carriers
[b]Average contribution to the three control groups

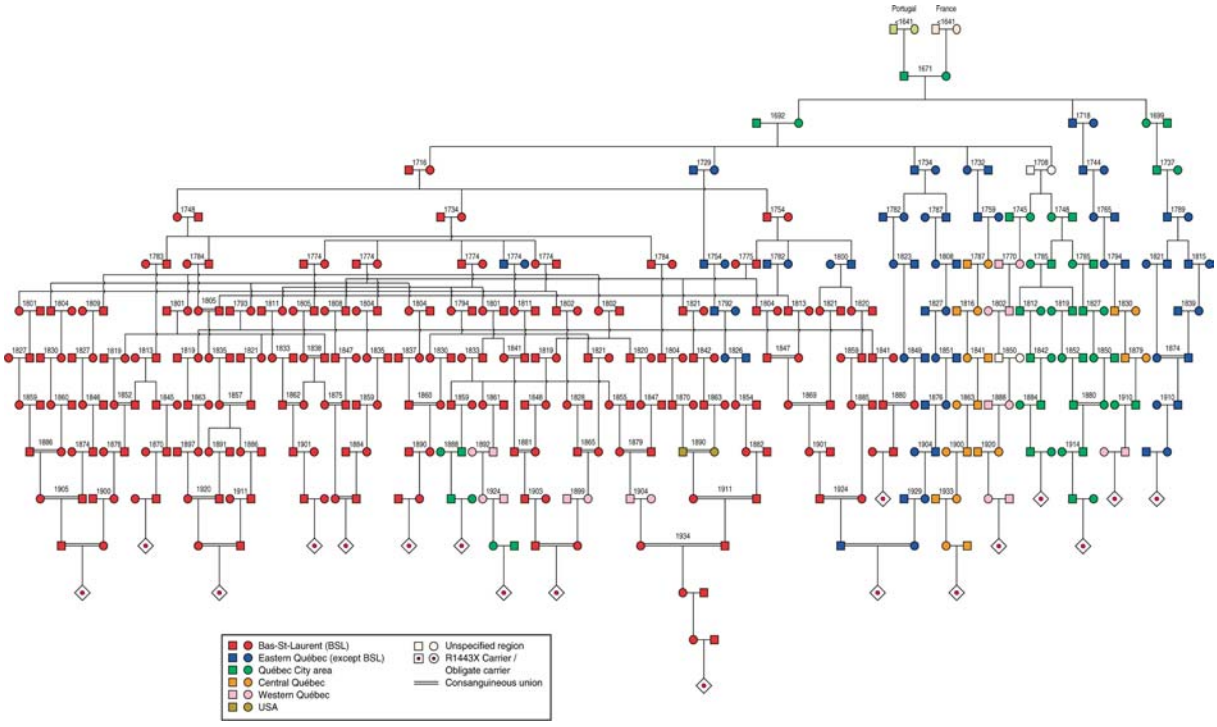

Fig. 3 Introduction and diffusion of the R1443X mutation in the French-Canadian population. The founder couple married in Quebec City in 1671. The husband migrated from Portugal, while his wife had come from France. None of their parents ever came to North America. For each couple, dates (starting from the probands' grandparents) and places of marriage are indicated. Regions were grouped according to their localization in the province except for the eastern Quebec region of Bas-Saint-Laurent (*in red*), which was singled out because of its high number of appearances in the pedigree. For confidentiality purposes, sex of carriers as well as dates of their parents' marriage are not indicated

among the different regions. In agreement with Fig. 3, the highest genetic contribution of these three obligate carriers is found in the Bas-St-Laurent region followed by the two adjacent regions of Gaspe peninsula and Cote-du-Sud. The other regions do not display any clear geographical gradient or grouping except that the lowest values are found mostly among the southwest regions of Quebec.

Discussion

Haplotype analysis

The extent of haplotype data among the 18 French-Canadian families analyzed indicates that there was a single introduction of the mutation for this population and therefore the observed elevated frequency of this specific mutation is explained by a founder effect. Although the origin of the founders of the Quebec population has been quite extensively studied, information on the level of kinship between founders originating from the same village or region is scarce and we cannot rule out the possibility of more than one introduction by individuals sharing the same haplotype (mutational event); however as previously shown, genealogical data make us confident that the mutation was introduced by a single founder.

The French-Canadian families share a haplotype extending over 9.3 cM in most families. Only at two markers, D17S933 and D17S1859, the most centromeric and telomeric, respectively, can we see a difference in haplotypes. However, at least two of the changes observed are recombination events, while the other events are either marker mutations or 2nd recombination events within the same family (family 5582). The Finnish families also share a very long haplotype and differ only at the most telomeric and centromeric markers. A core haplotype of 7.9 cM has already been reported in this region for these three kindreds and this particular mutation has been associated with a south-eastern region in southern Karelia, near the Russian border (Sarantaus et al. 2000). Because of its defined ethnic and geographic location, the Finnish population is also considered a founder population. Finland is inhabited by a population distinguishable from other European populations by its language, which stems from the Uralic language family (Kere 2001). It is estimated that approximately 1,000 founders populated the coastal and southern regions of Finland, nearly 2000 years ago. A smaller subgroup of the Finnish founder population migrated to the more eastern and northern areas 20–25 generations ago (de la Chapelle 1993; Peltonen et al. 1999). Three families from Brittany (347, 1203 and 1375) also display a large portion of a shared haplotype and to a lesser extent with family 1628, also of Breton origin. Even though Brittany is not considered a founder population per se, the Bretons represent a linguistic minority

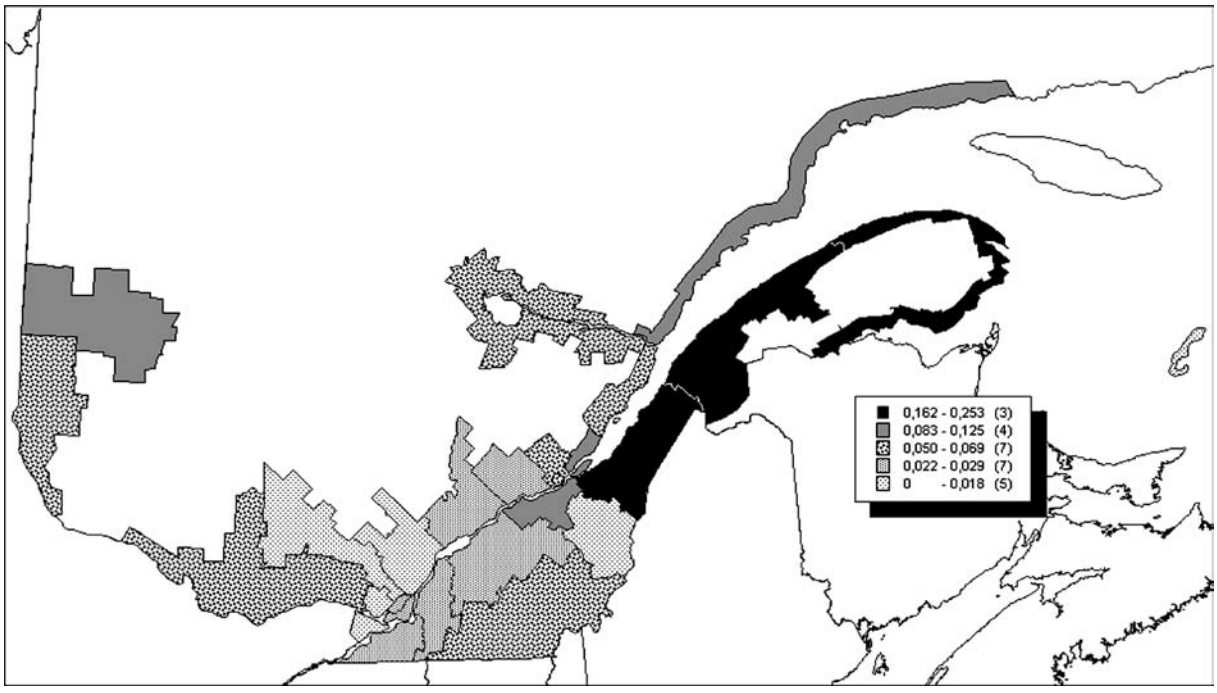

Fig. 4 Map of the mean genetic contribution (%) of the founders' obligate carrier children to the contemporary regional populations. All regions are represented except for northern Quebec. The highest genetic contribution of these three obligate carriers is found in the Bas-St-Laurent region and in the two adjacent regions of Gaspe peninsula and Cote-du-Sud (*black color*). *Numbers in brackets* refer to the number of regions pertaining to each class of values

in France. A Celtic language is spoken, which was reintroduced from the British Isles (Cavalli-Sforza et al. 1994). This linguistic difference helped to create some degree of isolation of this part of the country.

Haplotypes extending over long distances have also been reported for other founder populations (Neuhausen et al. 1996). However, when we compare the haplotype data obtained from the French-Canadian families with the 19 families from other populations and bearing the same mutation, it becomes probable that there is more than one origin of the *BRCA1* R1443X mutation, which is in agreement with a previous study where only four families with this mutation were analyzed (Neuhausen et al. 1996). Recurrent mutation is not surprising since R1443X results from mutation of the hypermutable CG dinucleotide to TG. Indeed, even though a core haplotype of three markers (D17S1322, D17S1323 and D17S1327) encompassing the R1443X mutation is observed for the great majority of families, four are clearly different. The outliers are two families from France (family 390 from Brittany and family F85 from Normandy) and two British families (B35 and B155 from Wales and United Kingdom, respectively). However, families 390 and F85 do share a haplotype between them spanning more than 1,200 kb. Two alternative explanations may account for the distinctive haplotype observed in these French families. One possibility might be an independent mutational event in the Scandinavian settlers of Normandy (ninth century Norwegian and Danish Vikings, as described in Cavalli-Sforza 1994). On the other hand, marker mutations within the core haplotype in the lineage giving rise to the most closely related haplotype among other French families (F35) could equally well explain the observed haplotype and may be more plausible, given F35 is also from Normandy and that R1443X has not been reported in Scandinavian populations. Nevertheless, the British family from Wales (family B35) is entirely unrelated to any other haplotypes and thus likely represents an independent mutational event. This family is possibly of Celtic origin.

Extensive linkage disequilibrium in the *BRCA1* region has been reported in previous studies (Durocher et al. 1996; Dunning et al. 1997), as well as in the EDH17B2 and *RNU2* regions (Normand et al. 1993; Liao et al. 1997). Interestingly, the region on chromosome 17q21 spanning *RNU2-BRCA1* has shown evidence for suppression of recombination (Liu and Barker 1999). This could explain why, in refining the location of *BRCA1*, the meiotic breaks usually observed involved markers no closer than 0.5–1.0 Mb (Albertsen et al. 1994; Tonin et al. 1994). A lack of recombination in this region could also account for the homozygous haplotypes observed by Tonin et al. (1998) for that specific mutation over a shorter distance, highlighting the importance and implications of extending a haplotype especially for the estimation of the time of origin of *BRCA1* mutations in a founder population.

Thus, analysis of haplotypes associated with the R1443X mutation in individuals of French-Canadian, French, Finnish, Dutch, Belgian and Russian ancestry, suggests existence of a relatively old mutation that has spread through migration. The novel haplotypes asso-

ciated with R1443X in families from United Kingdom point towards independent origins in different populations. The R1443X has been reported not only in patients from the populations studied herein, but also of Italian, Ashkenazi and central/eastern European descent (BIC database—http://research.nhgri.nih.gov/bic/). Other, rare reports of the R1443X mutation include patients of native American and latin America/Carribean ancestry. It would, of course, be interesting to determine whether these mutations derive from the described founder mutation, or represent additional independent mutational events.

Identification of the most likely introducers of the R1443X mutation in the Quebec population

We have identified the founder couple with the highest probability of having introduced the R1443X mutation in the Quebec population. Using criteria of frequency and specificity, this couple is the most likely, as it appears in the 18 carriers' genealogies but in only 26 of the 54 controls' genealogies. Moreover, this ancestral group has the third highest contribution to carriers and the greatest difference when compared to controls. Even if ancestral groups 1A and 1B had a higher contribution to carriers, we could exclude them as they were also very frequent among the control groups, indicating that they have a great overall contribution to the whole population.

The founder couple married in Quebec City in 1671. They are both the only members of their sibship who migrated to Nouvelle-France (PRDH 2004). The wife came from France and the husband came from Portugal, which is interesting since most 17th century founders of the Quebec population came from France. As the R1443X mutation has been reported in French families (Stoppa-Lyonnet et al. 1997) and, to our knowledge, has not been found in Portuguese families (Soares et al. 2000), we can therefore hypothesize that the mutation was introduced from the mother's side, but of course we cannot be sure. This founder couple had five children, who all married (PRDH 2004), and based on our pedigree, three of them are obligate carriers. Among these three children, one daughter is related to 16 of the 18 carriers. An explanation for this could be that she had ten children, eight of which married (five appearing in the pedigree), whereas her sister had five children, only two of which married (one appears in the pedigree) and her brother had only one married child. We must also keep in mind that we do not have complete ascertainment of R1443X carrier families in the Quebec population.

Another interesting feature of the pedigree is related to the spatial distribution of the ancestors based on their place of marriage. The color code *red* representing the Bas-St-Laurent region, which is located on the southeast shore of the St. Lawrence River is clearly over-represented and this concerns more specifically the descendants of the daughter with the most children. The other ancestors come mostly from the adjacent regions of eastern Quebec (shown in *blue*), whereas some are from central Quebec and Quebec city area and very few are from western Quebec.

The pedigree also displays the complexity of the transmission paths of the mutation, which in turn is but one illustration of the French-Canadian genealogical structure. For instance, five of the 18 carriers could theoretically have inherited R1443X from either their mother or father because they are related to the founder couple through both their maternal and paternal genealogies. In principal, this could even lead to homozygotes, except that such a genotype would probably be embryonic lethal, as mice embryos homozygous for a *BRCA1* mutation do not survive past 10–13 days (Gowen et al. 1996).

Diffusion of the mutation in the Quebec population

Based on the descending pedigree and on the genetic contributions of the three obligate carriers to the regional populations of Quebec, we see a regional clustering of the mutation in the Bas-St-Laurent region. These results indicate that the consequences of the founder effect in the Quebec French-Canadian population show variability at the regional level. This stratification is due to differential demographic behaviors of founders and their descendants in terms of migration and settlement patterns on the territory as well as characteristics of spouse choice, fertility and pre-reproductive mortality. It has been observed previously both in demographic studies (Heyer et al. 1997) and in studies on the distribution of mutations involved in various inherited disorders (Carter et al. 1998; Couture et al. 1999).

In a population genetics perspective, we believe our combination of molecular, haplotypic and genealogical approaches for the study of a specific mutation has allowed us to perform one of the most in-depth studies of this type conducted so far. Although the identification of the founders who introduced the mutation will always remain probabilistic we are confident that our method, based on the use of control groups and criteria of frequency and specificity to select ancestors gives us a very high probability of having pinpointed the right couple. On more theoretical grounds, the combined use of genealogical, historical and molecular information from the same population contributes to a better understanding of the impact of demographic factors on the genetic structure of populations. The French-Canadian population of Quebec is particularly well suited for this type of study.

In an epidemiological setting, our findings cannot be interpreted as an indication that breast cancer should be elevated or more frequent in the contemporary population of Bas-St-Laurent, not even that the proportion of hereditary cancers among overall breast cancers should be increased in this region. We have to

be very cautious while interpreting genealogical information in the context of risk factor distribution. One has to keep in mind that genealogical data on ancestors yield genetic information on the whole genome of these ancestors and not on a specific allele they carry. Genetic contribution of an ancestor to the gene pool of an individual or of a group of individuals represents the proportion of the genome of this individual or this group that is expected to have come from this ancestor. We do not believe it should be considered as the expected carrier rate of a specific mutation, at least not in the practical context of genetic counseling on a day-to-day basis in the clinics. However, it could be that among affected individuals pertaining to high-risk families, the probability of identifying the R1443X mutation would be higher if this family has ancestry in the Bas-St-Laurent area.

Nonetheless, it is worth mentioning that our findings are in accordance with a study on cancer incidence in Quebec, which shows the highest prevalence of breast cancer cases in the Bas-St-Laurent area (Louchini et al. 2004). We believe these results deserve further investigation for a better understanding of the distribution and epidemiology of the mutation. The prevalence of the R1443X mutation in the Quebec population is not known. However, an estimate for the Bas-Saint-Laurent could be obtained by screening a series of breast tumors from this specific region. Moreover, as we know the mean age at onset of breast cancer for women with a *BRCA1* mutation in the French-Canadian cohort is in the early forties, we are also intending to obtain information on the age of onset of breast cancer and other *BRCA1*-associated cancers, such as ovarian, among affected women in this region.

We strongly believe that identifying founder mutations and establishing their origin and spread is a critical step towards providing accurate counseling and developing inexpensive mutation detection in the concerned populations. Moreover, as the majority of R1443X carriers analyzed in the French-Canadian families turned out to be unaffected, this suggests that even among families with founder mutations, there appear to be differences in susceptibility and in age of onset of cancer, highlighting the role and importance of risk-modifying factors, such as diet, lifestyle and modifier genes, especially in the presence of a strong family history of breast and/or ovarian cancer. By identifying those at high-risk (i.e., carriers), but unaffected, it is possible to offer to these healthy individuals a better follow-up for good risk management.

By providing a clearer representation of the population structure and of the role of demographic history in shaping this structure, this type of study could also facilitate the identification of new deleterious alleles and therefore impact on those affected with breast cancer in other populations as well. Such studies are certainly not unique to breast cancer genetics and could be used to understand other diseases. Here, with breast cancer as a model for other complex traits, we have demonstrated how molecular genetics and demogenetic analyses can complement each other and provide findings that could have an impact on public health.

Acknowledgments The authors are indebted to the participants and their families for providing DNA samples. We would also like to thank Claire Brousseau, Sylvie Délos, Marie-Andrée Lajoie, Pascale Léger, Josée Rhéaume, Gilles Leblanc, Carolle Samson, Hélène Malouin, Nathalie Bolduc, Andrée McMillan and Tina Babineau of the Cancer Genomics Laboratory (Québec) and Jillian Parboosingh and Ross Mc Leod from the Molecular Diagnostic Laboratory (Alberta) for sample management and skillful technical assistance, Michèle Jomphe and France Néron from the BALSAC Project (Chicoutimi) as well as Lise Gobeil from the Interdisciplinary Research Group on Demography and Genetic Epidemiology (Chicoutimi) for genealogical reconstruction and analyses and Laurent Richard and Nicolas Lanouette from the Historical Geography Laboratory at Laval University for cartography work. We also appreciated advice received from the different ethics committees. This work was supported by the Fonds de la Recherche en Santé du Québec (FRSQ)/ Réseau de Médecine Génétique Appliquée (RMGA), the Canadian Breast Cancer Research Alliance and the Canadian Institutes of Health Research (CIHR) for the INHERIT BRCAs research program. F.D. is a recipient of a Research Career Award in the Health Sciences by IRSC/Rx&D HRF, J.S. is Chairholder of the Canada Research Chair in Oncogenetics.

Appendix 1

- Yves-Jean Bignon, Centre Jean-perrin, Clermont-Ferrand, France
- Agnès Hardouin, Centre Régional Francois Baclesse, Caen, Normandy
- Ludwine Messiaen Universitair Ziekenhuis, Gent, Belgium
- Mike Stratton, Institute for Cancer Research, Sutton, Surrey, UK
- Ans MW van den Ouweland, Erasmus Universiteit (Rotterdom, The Netherlands)

Appendix 2

- Paul Bessette, Service de gynécologie, Centre hospitalier Universitaire de Sherbrooke, Fleurimont.
- Bartha Maria Knoppers, Chaire de recherche en droit et médecine, Centre de recherche en droit public, Université de Montréal, Montréal.
- Bernard Lespérance, Roxane Pichette, Service d'hémato-oncologie, Hôpital du Sacré-Cœur, Montréal.
- Louise Provencher, Clinique des maladies du sein Deschênes-Fabia, Hôpital du Saint-Sacrement, Québec.
- Marc Tremblay, Département des sciences humaines, Université du Québec à Chicoutimi, Chicoutimi.
- Patricia Voyer, Clinique des maladies du sein, Carrefour de Santé de Jonquière, Jonquière.

Appendix 3

Table A Details of microsatellite markers and their relative map positions

Marker, *BRCA1*	Position, 41570–41650 (kb)[a]	PCR primer sequences	
		Forward (5′-3′)	Reverse (5′-3′)
D17S933	33343	ACTCACTGGGGTCCTGG	TGTGGTTTCCTTATAGACTGTAGA
D17S800	39430	GGTCTCATCCATCAGGTTTT	ATAGACTGTGTACTGGGCATTGA
D17S846	40266	TGCATACCTGTACTACTTCAG	TCCTTTGTTGCAGATTTCTTC
D17S776	40583	TCCATGACCCTATGGACCA	AAACCTCTGTCTCTTTGCAG
D17S1185	40927	GGTGACAGAACAAGACTCCATC	GGGCACTGCTATGGTTTAGA
EDH17B2	41079	CAGTACTAAAGGCCCTATTATCAAA	AGGCTGCAGTGAGTCCAGAT
D17S1320	41153	ACTTTCCAGAAAATCTCTGCTC	CCACGTCTTTTCTGTGTTCC
D17S1328	~41533	TCTGTAGGTGCTAAGCAGTGG	GTTGCAGTGAGCCGAGAATG
D17S855	41578	GGATGGCCTTTTAGAAAGTGG	ACACAGACTTGTCCTACTGCC
D17S1322	~41584	CTAGCCTGGGCAACAAACGA	GCAGGAAGCAGGAATGGAAC
D17S1323	41611	TAGGAGATGGATTATTGGTG	AAGCAACTTTGCAATGAGTG
D17S1327	41851	CTAAGGAGGTTTCTCTGGAC	TTCACAACTCAAGGTAAGATAGG
D17S1326	41860	CAGCTGATATTTCACAGGACT	AGAGCAAAACTCCATCTCAAACA
D17S1325	42011	AAAGGTGGCAATTCACAGTTG	GTGATAAAACTCAGTGGTACTC
D17S1329	42098	GACTCTGAAGGTAAAGAGCAA	CTCCCCTGCCTTGGGAGTAG
D17S791	45331	GTTTTCTCCAGTTATTCCCC	GCTCGTCCTTTGGAAGAGTT
D17S1859	46299	CATTAACATTAGGTGGGTTAGACA	CCTTCATTTCACTTGAGAGTTTG

[a]According to NCBI

References

Albertsen HM, Smith SA, Mazoyer S, Fujimoto E, Stevens J, Williams B, Rodriquez P et al (1994) A physical map and candidate genes in the *BRCA1* region on chromosome 17q12-21. Nat Genet 7:472–479

Barkardottir RB, Sarantaus L, Arason A, Vehmanen P, Bendhal PO, Kainu T, Syrjakoski K et al (2001) Haplotype analysis in Icelandic and Finnish *BRCA2* 999del5 breast cancer families. Eur J Hum Genet 9:773–779

Bayleran JK, Yan H, Hopper CA, Simpson EM (1996) Frequencies of cystic fibrosis mutations in the Maine population: high proportion of unknown alleles in individuals of French-Canadian ancestry. Hum Genet 98:207–209

Berman DB, Wagner-Costalas J, Schultz DC, Lynch HT, Daly M, Godwin AK (1996) Two distinct origins of a common *BRCA1* mutation in breast-ovarian cancer families: a genetic study of 15 185delAG-mutation kindreds. Am J Hum Genet 58:1166–1176

Bouchard G (1986) The processing of ambiguous links in computerized family reconstruction. Historical Methods 19:9–19

Bouchard G (2003) Projet BALSAC—Rapport annuel 2002–2003. Chicoutimi (Quebec) (http://www.balsac.uqac.ca)

Bouchard G, de Braekeleer M (1990) Homogénéité ou diversité? L'histoire de la population du Québec revue à travers ses gènes. Histoire Soc 23:325–361

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

Cavalli-Sforza LL, Menozzi P, Piazza A (1994) The history and geography of human genes. Princeton University Press, Princeton

Charbonneau H, Desjardins B, Légaré J, Denis H (2000) The population of the St. Lawrence valley, 1608–1760. In: Haines MR, Steckel RH (eds) A population history of North America. Cambridge University Press, New York, pp 99–142

Couture P, Morissette J, Gaudet D, Vohl MC, Gagne C, Bergeron J, Despres JP, Simard J (1999) Fine mapping of low-density lipoprotein receptor gene by genetic linkage on chromosome 19p13.1-p13.3 and study of the founder effect of four French-Canadian low-density lipoprotein receptor gene mutations. Atherosclerosis 143:145–151

Couture P, Bovill EG, Demers C, Simard J, Delage R, Scott BT, Valliere JE, Callas PW, Jomphe M, Rosendaal FR, Aiach M, Long GL (2001) Evidence of a founder effect for the protein C gene 3363 inserted C mutation in thrombophilic pedigrees of French origin. Thromb Haemost 86:1000–1006

Davignon J, Roy M (1993) Familial hypercholesterolemia in French-Canadians: taking advantage of the presence of a "founder effect". Am J Cardiol 72:6D–10D

de la Chapelle (1993) Disease gene mapping in isolated human populations: the example of Finland. J Med Genet 30:857–865

Dorum A, Moller P, Kamsteeg EJ, Scheffer H, Burton M, Heimdal KR, Maehle LO et al (1997) A *BRCA1* founder mutation, identified with haplotype analysis, allowing genotype/phenotype determination and predictive testing. Eur J Cancer 33:2390–2392

Dunning AM, Chiano M, Smith NR, Dearden J, Gore M, Oakes S, Wilson C et al (1997) Common *BRCA1* variants and susceptibility to breast and ovarian cancer in the general population. Hum Mol Genet 6:285–289

Dupre N, Howard HC, Mathieu J, Karpati G, Vanasse M, Bouchard JP, Carpenter S, Rouleau GA (2003) Hereditary motor and sensory neuropathy with agenesis of the corpus callosum. Ann Neurol 54:9–18

Durocher F, Shattuck-Eidens D, McClure M, Labrie F, Skolnick MH, Goldgar DE, Simard J (1996) Comparison of *BRCA1* polymorphisms, rare sequence variants and/or missense mutations in unaffected and breast/ovarian cancer populations. Hum Mol Genet 5:835–842

Efron B, Tibshirani RJ (1993) An introduction to the bootstrap. Chapman & Hall, London

Engert JC, Berube P, Mercier J, Dore C, Lepage P, Ge B, Bouchard JP, Mathieu J, Melancon SB, Schalling M, Lander ES, Morgan K, Hudson TJ, Richter A (2000) ARSACS, a spastic ataxia common in northeastern Quebec, is caused by mutations in a new gene encoding an 11.5-kb ORF. Nat Genet 24:120–125

Ford D, Easton DF, Stratton M, Narod S, Goldgar D, Devilee P, Bishop DT et al (1998) Genetic heterogeneity and penetrance analysis of the *BRCA1* and *BRCA2* genes in breast cancer families. Am J Hum Genet 62:676–689

Friedman LS, Szabo CI, Ostermeyer EA, Dowd P, Butler L, Park T, Lee MK et al (1995) Novel inherited mutations and variable expressivity of *BRCA1* alleles, including the founder mutation 185delAG in Ashkenazi Jewish families. Am J Hum Genet 57:1284–1297

Gayther SA, Harrington P, Russell P, Kharkevich G, Gargavtseva RF, Ponder BAJ (1997) Frequently occurring germ-line mutations of the *BRCA1* gene in ovarian cancer families from Russia. Am J Hum Genet 60:1239–1242

Godard B, Simard J, INHERIT BRCAs (2003) Les enjeux éthiques de l'identification d'une susceptibilité génétique au cancer du sein dans un contexte de recherche clinique intégrée. In: Herve C, Knoppers BM, Molinari PA (eds) Les pratiques de recherche biomédicale visitées par la bioéthique. Dalloz, pp 113–136

Gorski B, Byrski T, Huzarski T, Jakubowska A, Menkiszak J, Gronwald J, Pluzariska A et al (2000) Founder mutations in the *BRCA1* gene in Polish families with breast-ovarian cancer. Am J Hum Genet 66:1963–1968

Gowen LC, Johnson BL, Latour AM, Sulik KK, Koller MH (1996) *BRCA1* deficiency results in early embryonic lethality characterized by neuroepithelial abnormalities. Nat Genet 12:191–194

Hauck WW, Martin AO (1984) A statistical test for detection of ancestral genetic contributions to disease occurrence in finite populations. Genet Epidemiol 1:383–400

Heyer E, Tremblay M, Desjardins B (1997) The 17th century European origins of hereditary diseases in the Saguenay population (Quebec, Canada). Hum Biol 69:209–225

Jeanpierre M (1987) A rapid method for the purification of DNA from blood. Nucleic Acids Res 15:9611

Johannesdottir G, Gudmundsson J, Bergthorsson JT, Arason A, Agnarson BA, Eiriksdottir G, Johannsson OT et al (1996) High prevalence of the 999del5 mutation in Icelandic breast and ovarian cancer patients. Cancer Res 56:3663–3665

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

Liao D, Pavelitz T, Kidd JR, Kidd KK, Weiner AM (1997) Concerted evolution of the tandemly repeated genes encoding human U2 snRNA (the *RNU2* locus) involves rapid intrachromosomal homogenization and rare interchromosomal gene conversion. EMBO J 16:588–598

Liede A, Cohen B, Black DM, Davidson RH, Renwick A, Hoodfar E, Olopade O et al (2000) Evidence of a founder *BRCA1* mutation in Scotland. Br J Cancer 82:705–711

Liu X, Barker DF (1999) Evidence for effective suppression of recombination in the chromosome 17q21 segment spanning *RNU2-BRCA1*. Am J Hum Genet 64:1427–1439

Louchini R, Beaupré M, Bouchard C, Goggin P (2004) La prévalence du cancer au Québec en 1998. Institut National de Santé Publique du Québec, Octobre 2004

McInnis M (2000) The population of Canada in the 19th century. In: Haines MR, Steckel RH (eds) A population history of North America. Cambridge University Press, New York, pp 371–402

Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q et al (1994) A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*. Science 266:66–71

Mootha VK, Lepage P, Miller K, Bunkenborg J, Reich M, Hjerrild M, Delmonte T, Villeneuve A, Sladek R, Xu F, Mitchell GA, Morin C, Mann M, Hudson TJ, Robinson B, Rioux JD, Lander ES (2003) Identification of a gene causing human cytochrome c oxidase deficiency by integrative genomics. Proc Natl Acad Sci USA 100:605–610

Neuhausen SL, Mazoyer S, Friedman L, Stratton M, Offit K, Calido A, Tomlinson G et al (1996) Haplotype and phenotype analysis of six recurrent *BRCA1* mutations in 61 families: result of an international study. Am J Hum Genet 58:271–280

Neuhausen SL, Godwin AK, Gershoni-Baruch R, Schubert E, Garber J, Stoppa-Lyonnet D, Olah E et al (1998) Haplotype and phenotype analysis of nine recurrent *BRCA2* mutations in 111 families: results of an international study. Am J Hum Genet 62:1381–1388

Normand T, Narod S, Labrie F, Simard J (1993) Detection of polymorphisms in the estradiol 17β-hydroxysteroid dehydrogenase II gene at the EDH17B2 locus on 17q11-q21. Hum Mol Genet 2:479–483

Peelen T, van Vliet M, Petruj-Bosch A, Mieremet C, van den Ouweland AMW, Hoervorst F, Brohet R et al (1997) A high proportion of novel mutations in *BRCA1* with strong founder effects among Dutch and Belgian hereditary breast and ovarian cancer families. Am J Hum Genet 60:1041–1049

Peltonen L, Jalanko A, Varilo T (1999) Molecular genetics of the Finnish disease. Hum Mol Genet 8:1913–1923

Petrij-Bosch A, Peelen T, van Vliet M, van Eijk R, Olmer R, Drusedau M, Hogervorst FBL et al (1997) *BRCA1* genomic deletions are major founder mutations in Dutch breast cancer patients. Nat Genet 17:341–345

PRDH (2004) Programme de recherche en démographie historique, http://www.genealogie.umontreal.ca/en/

Ramus SJ, Kote-Jarai Z, Friedman LS, van der Looij M, Gayther SA, Csokay B, Ponder BAJ, Olah E (1997) Analysis of *BRCA1* and *BRCA2* mutations in Hungarian families with breast or breast-ovarian cancer. Am J Hum Genet 60:1242–1246

Sarantaus L, Huusko P, Eerola H, Launonen V, Vehmanen P, Rapakko K, Gillanders E et al (2000) Multiple founder effects and geographical clustering of *BRCA1* and *BRCA2* families in Finland. Eur J Hum Genet 8:757–763

Sarantaus L, Vahteristo P, Bloom E, Tamminen A, Unkila-Kallio L, Butzow R, Nevanlinna H (2001) *BRCA1* and *BRCA2* mutations among 233 unselected Finnish ovarian carcinoma patients. Eur J Hum Genet 9:424–430

Scriver C (2001) Human genetics: Lessons from the Quebec population. Annu Rev Genomics Hum Genet 2:69–101

Simard J, Tonin P, Durocher F, Morgan K, Rommens J, Gingras S, Samson C et al (1994) Common origins of *BRCA1* mutations in Canadian breast and ovarian cancer families. Nat Genet 8:392–398

Soares R, Amendoeira I, Monteiro P, Lopes CS, Schmitt FC (2000) Analysis of mutations in the *BRCA1* gene in patients with cancer of the breast and/or the ovary in Portugal. Acta Med Port 13:273–276

Stoppa-Lyonnet D, Laurent-Puig P, Essioux L, Pages S, Ithier G, Ligot L, Fourquet A et al (1997) *BRCA1* sequence variations in 160 individuals referred to a breast/ovarian family cancer clinic. Am J Hum Genet 60:1013–1020

Szabo CI, King M-C (1997) Population genetics of *BRCA1* and *BRCA2*. Am J Hum Genet 60:1013–1020

Tavtigian SV, Simard J, Rommens J, Couch F, Shattuck-Eidens D, Neuhausen S, Merajver S et al (1996) The complete *BRCA2* gene and mutations in 13q-linked kindreds. Nat Genet 12:333–337

Thompson EA (1986) Pedigree analysis in human genetics. Johns Hopkins University Press, Baltimore

Thorlacius S, Olafsdattir G, Tryggvadottir L, Neuhausen S, Jonasson JG, Tavtigian S, Tulinius H et al (1996) A single *BRCA2* mutation in male and female breast cancer families from Iceland with varied cancer phenotypes. Nat Genet 13:117–119

Thorlacius S, Sigurdsson S, Bjarnadottir H, Olafsdottir G, Jonasson JG, Tryggvadottir L, Tulinius H et al (1997) Study of a single *BRCA2* mutation with high carrier frequency in a small population. Am J Hum Genet 60:1079–1084

Tonin P, Serova O, Simard J, Lenoir G, Feunteun J, Morgan K, Lynch H et al (1994) The gene for hereditary breast-ovarian cancer, *BRCA1*, maps distal to EDH17B2 in chromosome region 17q12-q21. Hum Mol Genet 3:1679–1682

Tonin P, Serova O, Lenoir G, Lynch H, Durocher F, Simard J, Morgan K et al (1995) *BRCA1* mutations in Ashkenazi Jewish women. Am J Hum Genet 57:189

Tonin PN, Mes-Masson AM, Futreal PA, Morgan K, Mahon M, Foulkes WD, Cole DE, Provencher D, Ghadirian P, Narod SA (1998) Founder *BRCA1* and *BRCA2* mutations in French Canadian breast and ovarian cancer families. Am J Hum Genet 63:1341–1351

Tremblay M, Jomphe M, Vézina H (2001) Comparaison de structures patronymiques et génétiques dans la population québécoise. In: Brunet G, Darlu P, Zei G (eds) Le patronyme: histoire, anthropologie, société. CNRS Éditions, Paris, pp367–389

Triggs-Raine B, Richard M, Wasel N, Prence EM, Natowicz MR (1995) Mutational analyses of Tay-Sachs disease: studies on Tay-Sachs carriers of French-Canadian background living in New England. Am J Hum Genet 56:870–879

Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, Collins N et al (1995) Identification of the breast cancer susceptibility gene *BRCA2*. Nature 378:789–792