

Origines et contributions génétiques des fondatrices et des fondateurs de la population québécoise

Hélène VÉZINA, Marc TREMBLAY,
Bertrand DESJARDINS et Louis HOUDE *

Cette étude présente une analyse des origines géographiques et de la contribution génétique des ancêtres fondateurs de la population du Québec. À l'aide de données tirées du fichier BALSAC et du Registre de la population du Québec ancien, un corpus de 2223 généalogies ascendantes couvrant l'ensemble du territoire québécois a été construit. Ces généalogies remontent jusqu'au 17^e siècle et couvrent en moyenne plus de neuf générations. Tous les liens généalogiques unissant l'ensemble des individus identifiés dans les ascendances ont été établis. Les résultats indiquent qu'environ 81 pour cent du pool génétique québécois s'explique par des fondatrices et fondateurs venus de France au 17^e siècle. Des différences importantes existent toutefois selon le lieu d'origine en France et selon le sexe des fondateurs. Les résultats ont aussi mis en évidence la présence de nombreux liens d'apparentement éloigné dans cette population.

English abstract, p. 258.

* Hélène Vézina, Marc Tremblay et Louis Houde : Groupe de recherche interdisciplinaire en démographie et épidémiologie génétique, Université du Québec à Chicoutimi. Bertrand Desjardins : Programme de recherche en démographie historique, Université de Montréal. Une version préliminaire de cette étude a été présentée au XXV^e Congrès international de la population de l'UIESP (Tours, 18-23 juillet 2005). Les auteurs remercient les assistants de recherche Ève-Marie Lavoie, Lise Gobeil, Michèle Jomphe, Diane Brassard, France Néron, Frédéric Payeur et Denis Duval pour leur soutien technique, et le Laboratoire de géographie historique du Centre interuniversitaire d'études québécoises (Université Laval) pour la production des cartes. Ils remercient également les lecteurs anonymes pour leurs suggestions et commentaires pertinents. Cette étude a reçu l'appui financier du Conseil de recherches en sciences humaines du Canada (CRSHC), du Réseau de médecine génétique appliquée du Québec (RMGA-FRSQ) et des Instituts de recherche en santé du Canada (ECOGENE-21).

Le peuplement d'origine européenne au Québec remonte à l'arrivée de pionniers français, au début du 17^e siècle (Charbonneau et al., 1987). Jusqu'au milieu du 18^e siècle, soit pendant toute la durée du Régime français, environ 10 000 immigrants se sont établis en famille dans la vallée du Saint-Laurent (Charbonneau et al., 2000). Durant cette période, l'accroissement naturel devint rapidement le principal moteur de la croissance de la population. L'épisode d'immigration le plus intense eut lieu entre 1663 et 1673, lorsque Louis XIV parraina la venue de quelque 800 « Filles du Roy », afin d'atténuer le déséquilibre du marché matrimonial et d'encourager les soldats du régiment de Carignan à prendre épouse et à s'établir en Nouvelle-France (Landry, 1992). Après cet effort de colonisation, l'immigration française devint moins importante, et surtout masculine. Après la conquête anglaise, en 1760, l'immigration française fut réduite à néant, si l'on fait exception des Acadiens, descendants des colons français établis en Acadie au 17^e siècle, qui vinrent s'installer au Québec suite aux déportations dont ils furent victimes (Dickinson, 1994). Durant la deuxième moitié du 18^e siècle, la plupart des nouveaux arrivants provenaient des îles et colonies britanniques, exception faite d'un contingent de mercenaires allemands qui vinrent s'établir durant les années 1780 (Wilhelmy, 1984). La population canadienne-française étant essentiellement catholique, il y eut très peu de mariages avec les Anglo-protestants. Au 19^e siècle, le Royaume-Uni continua à fournir la plus grande partie de l'immigration au Québec (McInnis, 2000; Beaujot et Kerr, 2004). Des milliers d'immigrants en provenance de l'Angleterre, de l'Irlande et de l'Écosse s'établirent principalement dans les zones urbaines de la province (et surtout dans les villes de Montréal et Québec). La présence d'Irlandais catholiques, à partir du milieu du 19^e siècle, entraîna un certain nombre de mariages interethniques avec les Canadiens français. Au tournant du 20^e siècle, l'immigration devint beaucoup plus diversifiée, avec notamment des arrivants du sud et de l'est de l'Europe (Piché, 2003). Enfin, depuis les années 1970, l'immigration en provenance de l'Asie, de l'Amérique du sud et des Antilles est prédominante (Piché, 2003; Duchesne, 2004).

Ces vagues successives de mouvements migratoires et l'accroissement naturel qui s'ensuivit sont à l'origine de la plus grande partie de la population contemporaine du Québec. Autrement dit, le « pool génique » de la population québécoise, soit l'ensemble des gènes portés par les individus de cette population, provient des immigrants qui ont laissé des descendants

jusqu'à aujourd'hui¹. Cette étude vise à caractériser les origines et la diversité de ce pool génique, à l'aide de données généalogiques issues des fichiers BALSAC et BALSAC-RETRO (Projet BALSAC, 2005) et du Registre de la population du Québec ancien (Programme de recherche en démographie historique, 2005). Les données très riches contenues dans ces fichiers ont été exploitées dans le cadre de nombreux travaux de recherche qui ont permis, notamment, de faire ressortir la variabilité de la contribution démographique et génétique des fondateurs de la population québécoise (Charbonneau et al., 1987; Bouchard et al., 1995; Heyer et al., 1997; Tremblay et al., 2000; Gagnon et Heyer, 2001; Gagnon et al., 2001; Tremblay et al., 2001). Cependant, ces études ont principalement mis l'accent sur une approche régionale de la population du territoire québécois ou sur la période du Régime français (17^e-18^e siècles). La présente étude est la première, à notre connaissance, à couvrir l'ensemble de la population contemporaine du Québec. Elle s'inscrit dans un programme de recherche dont l'objectif est de développer les connaissances sur la formation, la stratification et la composition contemporaine des bassins génétiques régionaux du Québec (Groupe de recherche interdisciplinaire en démographie et épidémiologie génétique, 2005).

SOURCE DES DONNÉES : LES FICHIERS DE POPULATION BALSAC, BALSAC-RETRO ET RPQA

Les données généalogiques utilisées pour cette étude proviennent de trois sources principales, soit le fichier de population BALSAC, le fichier généalogique BALSAC-RETRO et le Registre de la population du Québec ancien (RPQA). Le fichier BALSAC contient des données démographiques et généalogiques sur la population du Québec des 19^e et 20^e siècles (Bouchard, 2004; Projet BALSAC, 2005). La plus grande partie de l'information a été tirée d'actes de mariage. À ce jour, le fichier contient environ deux millions d'actes, dont un peu plus de la moitié ont été jumelés. Le fichier BALSAC-RETRO a été développé dans le cadre de divers projets de recherche nécessitant des reconstructions généalogiques (Jomphe et Casgrain, 1997; Bouchard, 2004; Projet BALSAC, 2005). Au cours des dernières années, ce fichier s'est enrichi considérablement grâce aux travaux du Groupe de recherche interdisciplinaire en démographie et épidémiologie génétique (2005). Divers corpus généalogiques couvrant l'ensemble du territoire québécois ont notamment été constitués

(Vézina et al., 2005). Le fichier contient présentement des données portant sur quelque 360 000 individus dont les dates de mariage remontent jusqu'au début du 17^e siècle. La plupart des données de BALSAC-RETRO portant sur les 19^e et 20^e siècles ont été tirées du fichier BALSAC, alors que les données couvrant les 17^e et 18^e siècles proviennent du RPQA (Programme de recherche en démographie historique, 2005). Le RPQA a été construit à partir d'environ 700 000 actes de baptême, mariage et sépulture (Charbonneau et al., 1987; Desjardins, 1998).

CONSTRUCTION ET STRUCTURE DE L'ÉCHANTILLON GÉNÉALOGIQUE

Un ensemble de 2223 généalogies ascendantes a été construit à partir de sujets (points de départ des généalogies) mariés au Québec entre 1945 et 1965, dont les parents se sont aussi mariés au Québec. Ces sujets ont été choisis au hasard parmi les données disponibles dans le fichier BALSAC-RETRO (mariages catholiques essentiellement). Nous avons effectué le choix de la période de mariage des sujets de façon à maximiser la taille de l'échantillon, tout en cherchant à sélectionner les points de départ les plus proches possible de la période contemporaine. Dans son état actuel, le fichier RETRO contient en effet relativement peu d'informations tirées de mariages survenus après 1965. En outre, nous avons effectué l'échantillonnage de façon à obtenir une représentation proportionnelle à la distribution géographique de la population du Québec durant la période visée. Les données publiées du recensement canadien de 1956 ont été utilisées à cette fin. À cette époque, 88 pour cent de la population québécoise était catholique (Henripin et Péron, 1973).

Nous avons construit les généalogies en utilisant toutes les informations disponibles dans les trois fichiers sources. Les branches généalogiques remontent donc aussi loin que les données le permettent, c'est-à-dire, essentiellement, jusqu'aux premiers arrivants en sol québécois ayant été identifiés dans les généalogies. Grâce au jumelage des données, tous les liens généalogiques existant parmi l'ensemble des individus retracés dans les 2223 ascendances ont été établis.

La plus grande partie de ces généalogies remonte jusqu'au début du 17^e siècle, mais certaines branches s'interrompent plus tôt en raison du manque d'informations. Ainsi, les diverses branches généalogiques n'atteignent pas toutes la même pro-

fondeur. L'indice de complétude (C_g) fournit une mesure de cette profondeur généalogique atteinte dans les ascendances. Cet indice correspond, pour une profondeur donnée, au rapport du nombre total d'ancêtres identifiés (connus) à cette profondeur au nombre maximum (théorique) d'ancêtres à cette même profondeur :

$$C_g = A_g / (N \cdot 2^g)$$

où g est la profondeur généalogique (celle des parents des sujets étant la première),

A_g est le nombre d'ancêtres connus à la profondeur g et

N est le nombre de générations.

Cette mesure fait ressortir la quantité d'informations disponible à chaque profondeur générationnelle. La valeur maximale de 1 signifie que tous les ancêtres ont pu être identifiés à cette profondeur. Cette valeur maximale est atteinte pour les deux premières générations (parents et grands-parents des sujets), tel qu'attendu en raison des critères de sélection des sujets. À partir de la troisième génération (celle des arrière-grands-parents), certaines branches généalogiques s'interrompent (tableau 1). En moyenne toutefois, les générations sont au moins à 90 pour cent complètes jusqu'à la septième génération. Ce n'est qu'à partir de la neuvième génération que la complétude commence à

TABLEAU 1. Complétude généalogique, coefficient moyen de parenté et proportion d'apparentés dans l'échantillon des 2223 générations

Génération	Complétude généalogique (%)	Coefficient moyen de parenté (x 10 000)	Proportion d'apparentés (%)
1	100,00	0,00	0,00
2	100,00	0,02	0,003
3	97,94	0,05	0,02
4	95,71	0,11	0,13
5	93,81	0,22	0,87
6	92,09	0,42	4,92
7	90,07	0,82	22,01
8	86,55	1,67	69,39
9	79,70	3,13	96,08
10	58,78	4,49	97,65
11	24,86	5,11	97,77
12	5,32	5,24	97,78
13	0,65	5,25	97,78

décroître rapidement, ce qui correspond à la période d'arrivée de la plupart des immigrants français du 17^e siècle (Charbonneau et al., 1987). Au delà de la treizième génération, seuls quelques ancêtres ont pu être identifiés, jusqu'à une profondeur maximale de 17 générations. La profondeur généalogique moyenne, qu'on obtient en sommant les complétudes jusqu'à la profondeur maximale, est de 9,3 générations.

Au total, les 2223 généalogies contiennent plus de cinq millions de mentions d'ancêtres, mais plusieurs de ces mentions concernent les mêmes individus. En effet, un grand nombre d'ancêtres apparaissent dans plus d'une généalogie et même aussi à plusieurs reprises dans une même généalogie. En comptant chaque ancêtre une seule fois, on obtient 155 363 ancêtres distincts, pour une moyenne de 32,2 mentions par ancêtre. Le nombre d'hommes (77 005) est un peu moins élevé que le nombre de femmes (78 358), ce qui signifie que les remariages étaient plus fréquents pour les hommes que pour les femmes dans cette population, en raison notamment de la mortalité différentielle aux âges adultes.

L'APPARENTEMENT DES SUJETS, SELON LA PROFONDEUR GÉNÉRATIONNELLE

Deux individus sont biologiquement apparentés s'ils partagent au moins un ancêtre. En termes génétiques, ces individus auront donc une probabilité non nulle de partager deux copies identiques d'un gène provenant de cet ancêtre commun (Hartl, 1994). Ainsi, les reconstructions généalogiques permettent d'estimer l'intensité de l'apparentement dans une population. Plus grande est la profondeur généalogique, plus grande sera la probabilité de retrouver au moins un ancêtre commun à deux sujets donnés. Le niveau d'apparentement entre ces deux sujets dépendra du nombre d'ancêtres communs identifiés et de la distance généalogique (nombre de générations) séparant chacun des deux sujets de chaque ancêtre commun.

Ce niveau d'apparentement est mesuré à l'aide du coefficient de parenté (Φ) qui peut être défini comme la probabilité qu'un allèle choisi au hasard chez un sujet i soit identique à l'allèle correspondant (situé au même locus) chez un autre sujet j (Thompson, 1986):

$$\Phi_{i,j} = \sum_A \sum_P (1/2)^k (1 + F(A))$$

- où A est l'ensemble des ancêtres communs à i et j ,
 P est l'ensemble des chemins généalogiques reliant i à j et passant par l'ancêtre A ,
 k est le nombre d'individus dans le chemin P et
 $F(A)$ est le coefficient de consanguinité de A (c'est-à-dire le coefficient de parenté des parents de A).

À titre d'exemple, deux individus ayant une paire d'arrière-grands-parents comme ancêtres communs auront un coefficient de parenté d'au moins 0,015625.

Près de 2,5 millions de coefficients de parenté (un coefficient par paire de sujets distincts) ont ainsi été calculés pour chaque profondeur générationnelle, soit $(2223 * 2222)/2 = 2\,469\,753$. On calcule le coefficient moyen de parenté pour l'ensemble des 2223 sujets en divisant la somme de tous les coefficients par le nombre de coefficients.

Le tableau 1 présente les valeurs des coefficients moyens de parenté et celles des proportions des paires de sujets apparentés. Jusqu'à la septième génération, les coefficients sont relativement faibles. À cette profondeur, 22 pour cent des paires de sujets partagent au moins un ancêtre. La plus grande partie de la croissance des coefficients moyens s'effectue entre les septième ($\Phi = 0,000082$) et dixième ($\Phi = 0,000449$) générations, où presque 98 pour cent des paires de sujets ont un coefficient d'apparentement non nul. Après la onzième génération, les valeurs se stabilisent (jusqu'à une valeur maximale de 0,000525), en raison du manque d'informations généalogiques à cette profondeur. À la douzième génération, la valeur la plus élevée des coefficients de parenté atteint presque 0,10000, mais la plupart des coefficients ont une valeur beaucoup plus basse. Il faut noter que la distribution des coefficients de parenté ne suit pas une loi normale (près des trois quarts des coefficients sont en dessous de la valeur moyenne).

Ces résultats confirment encore une fois l'importance de l'apparentement éloigné dans la population du Québec (Tremblay et al., 2001; Vézina et al., 2005). Le nombre d'ancêtres communs après la septième génération atteint des proportions telles que presque chaque sujet de la population est apparenté à presque tous les autres sujets. Ces résultats peuvent paraître surprenants a priori, mais ils sont le reflet de la structuration de la population québécoise sur près de quatre siècles (il faut aussi préciser que rares sont les populations où l'on peut mesurer le coefficient d'apparentement à de telles profondeurs généalogiques).

LES ORIGINES ET PÉRIODES DE MARIAGE DES ANCÈTRES FONDATEURS

Les liens d'apparentement qui unissent les sujets de la population contemporaine du Québec peuvent être davantage caractérisés par l'analyse des origines géographiques et des périodes de mariage des ancêtres fondateurs identifiés dans les ascendances. Ces ancêtres fondateurs correspondent aux premiers arrivants en terre québécoise retracés dans l'ensemble des branches généalogiques. Leurs origines géographiques ont été déterminées à partir des informations disponibles sur leurs lieux de mariage, de naissance ou d'émigration².

Au total, 6808 fondateurs ont ainsi été identifiés, dont 1689 femmes et 5119 hommes. Ce déséquilibre hommes-femmes parmi les fondateurs s'explique par la nature même de l'immigration fondatrice au Québec (Charbonneau et al., 1987). La distribution de ces fondateurs selon leurs origines géographiques et leurs périodes de mariage est présentée aux tableaux 2 (femmes) et 3 (hommes). La grande majorité de ces ancêtres fondateurs proviennent de France, mais on note une différence assez importante entre les femmes (61,9 pour cent) et les hommes (79,8 pour cent). La plus grande partie de cette différence s'explique par la proportion assez élevée de fondatrices acadiennes, soit 22,4 pour cent des fondatrices (la plupart d'entre elles s'étant mariées entre 1700 et 1765), comparativement à seulement 7,8 pour cent chez les hommes. Chacune des autres origines figurant dans ces tableaux est celle de moins de 6 pour cent des fondatrices et fondateurs. La distribution selon la période de mariage fait ressortir l'importance des fondateurs mariés au 17^e siècle, soit 58,4 pour cent des fondatrices et 40,5 pour cent des fondateurs. L'écart entre hommes et femmes est attribuable ici à la part importante des fondateurs mariés entre 1700 et 1765 (44,7 pour cent, comparativement à seulement à 25,3 pour cent des fondatrices); ce groupe est composé en bonne partie de militaires français (Charbonneau et al., 2000).

La forte proportion des fondateurs et fondatrices françaises s'exprime encore davantage lorsqu'il s'agit de repérer, parmi l'ensemble des ascendances, celles qui comportent au moins une fondatrice (tableau 4) ou au moins un fondateur (tableau 5) d'une origine donnée. En effet, pratiquement toutes les générations de l'échantillon contiennent au moins une fondatrice et un fondateur d'origine française. Malgré la faible importance numérique des fondateurs d'origine britannique (voir les tableaux 2

TABLEAU 2. Distribution des 1689 immigrantes fondatrices selon l'origine et la période de mariage (%)

Origine	Période de mariage			Ensemble
	Avant 1700	1700- 1765	Après 1765	
France	56,7	3,8	1,4	61,9
Grande-Bretagne	0,2	0,4	2,6	3,1
Allemagne	0,1	0,2	0,5	0,7
Irlande	0,1	0,2	2,3	2,5
Autres pays européens	0,3	0,1	0,2	0,5
<i>Sous-total Europe</i>	<i>57,3</i>	<i>4,6</i>	<i>6,9</i>	<i>68,9</i>
Acadie	0,4	18,0	4,0	22,4
Reste de l'Amérique	0,7	2,1	2,8	5,6
<i>Sous-total Amérique</i>	<i>1,1</i>	<i>20,1</i>	<i>6,8</i>	<i>28,0</i>
Inconnue	0,0	0,5	2,6	3,1
Toutes les origines	58,4	25,3	16,3	100,0

TABLEAU 3. Distribution des 5119 immigrants fondateurs selon l'origine et la période de mariage (%)

Origine	Période de mariage			Ensemble
	Avant 1700	1700- 1765	Après 1765	
France	39,9	35,9	4,0	79,8
Grande-Bretagne	0,1	0,7	2,9	3,7
Allemagne	0,0	0,3	2,1	2,3
Irlande	0,1	0,2	1,3	1,5
Autres pays européens	0,3	0,7	0,6	1,6
<i>Sous-total Europe</i>	<i>40,4</i>	<i>37,8</i>	<i>10,8</i>	<i>89,0</i>
Acadie	0,0	5,9	1,9	7,8
Reste de l'Amérique	0,1	0,8	0,8	1,7
<i>Sous-total Amérique</i>	<i>0,1</i>	<i>6,7</i>	<i>2,8</i>	<i>9,6</i>
Inconnue	0,0	0,2	1,2	1,4
Toutes les origines	40,5	44,7	14,8	100,0

et 3), on retrouve au moins un de ces fondateurs dans plus de 90 pour cent des généalogies, alors que cette proportion n'atteint que 25 pour cent pour les fondatrices britanniques. Près d'un sujet sur cinq compte au moins un fondateur irlandais dans son ascendance, et 14 pour cent des ascendances comportent au moins un fondateur d'origine allemande. Les proportions correspondantes sont nettement plus faibles du côté des fondatrices, qui présentent toutefois une plus forte composante acadienne : près de 72 pour cent des ascendances contiennent au moins une fondatrice d'origine acadienne (60 pour cent pour les fondateurs acadiens). Par ailleurs, l'avantage de l'ancienneté ressort encore une fois assez nettement dans ces résultats, puisque 99 pour cent des généalogies comportent au moins un fondateur et une fondatrice mariés avant 1700. Cette proportion se maintient pour ce qui concerne les fondateurs de la période 1700-1765, mais elle baisse à 78 pour cent environ pour les fondatrices. Moins de la moitié des ascendances comportent un fondateur marié après 1765 (20 pour cent seulement pour les fondatrices).

Étant donné la grande importance des fondateurs et fondatrices provenant de France, les origines géographiques de ces immigrants français ont été examinées de plus près, soit à l'échelle des provinces d'origine (figures 1 et 2). Ce découpage géographique du territoire français correspond à peu de chose près à celui des provinces de l'Ancien Régime (Charbonneau et al., 1987). De façon générale, les résultats font bien ressortir l'importance plus grande des provinces de l'ouest et du nord du pays, par rapport à celles du sud ou de l'est. Il existe cependant des différences notables entre les fondatrices et les fondateurs. Chez les femmes, c'est la province de l'Île-de-France qui domine nettement, avec un peu plus de 31 pour cent des 930 fondatrices françaises, en raison de l'importance de la ville de Paris dans le recrutement des Filles du roi (Landry, 1992). Chez les hommes, la plus grande part revient à la Normandie (16 pour cent des 3779 fondateurs français), suivie d'assez près par le Poitou (12 pour cent). La Normandie (17 pour cent) et l'Aunis (13 pour cent) ont également fourni un important contingent de fondatrices. Ainsi, 61 pour cent des fondatrices françaises proviennent de seulement trois provinces, alors que chez les hommes, les trois principales provinces d'origine (Normandie, Poitou et Île-de-France) ont fourni 36 pour cent des fondateurs. Les origines des fondateurs français sont donc plus diversifiées que celles des fondatrices, et on remarque d'ailleurs que pratiquement tou-

TABLEAU 4. Proportion des généralogies où l'on retrouve au moins une fondatrice d'une origine et d'une période de mariage données (%)

Origine	Période de mariage			Ensemble
	Avant 1700	1700- 1765	Après 1765	
France	99,0	31,5	3,3	99,0
Grande-Bretagne	21,1	1,8	2,8	25,0
Allemagne	4,6	0,5	0,7	5,9
Irlande	0,6	0,7	2,4	3,6
Autres pays européens	47,9	0,1	0,1	48,0
<i>Sous-total Europe</i>	99,0	33,9	8,5	99,7
Acadie	39,6	57,6	7,6	71,9
Reste de l'Amérique	26,9	20,8	3,2	40,2
<i>Sous-total Amérique</i>	58,9	67,2	10,2	86,4
Inconnue	0,0	3,7	5,0	7,8
Toutes les origines	99,0	77,8	20,4	100,0

TABLEAU 5. Proportion des généralogies où l'on retrouve au moins un fondateur d'une origine et d'une période de mariage données (%)

Origine	Période de mariage			Ensemble
	Avant 1700	1700- 1765	Après 1765	
France	99,0	98,4	23,4	99,0
Grande-Bretagne	86,0	20,3	10,4	90,7
Allemagne	3,2	1,6	9,9	14,0
Irlande	10,4	6,1	4,7	19,1
Autres pays européens	92,6	13,9	2,2	93,3
<i>Sous-total Europe</i>	99,0	98,5	39,6	99,7
Acadie	4,6	56,5	10,7	60,1
Reste de l'Amérique	6,7	33,0	3,5	38,8
<i>Sous-total Amérique</i>	10,8	71,3	13,6	75,8
Inconnue	0,0	3,5	5,7	8,1
Toutes les origines	99,0	99,2	48,4	100,0

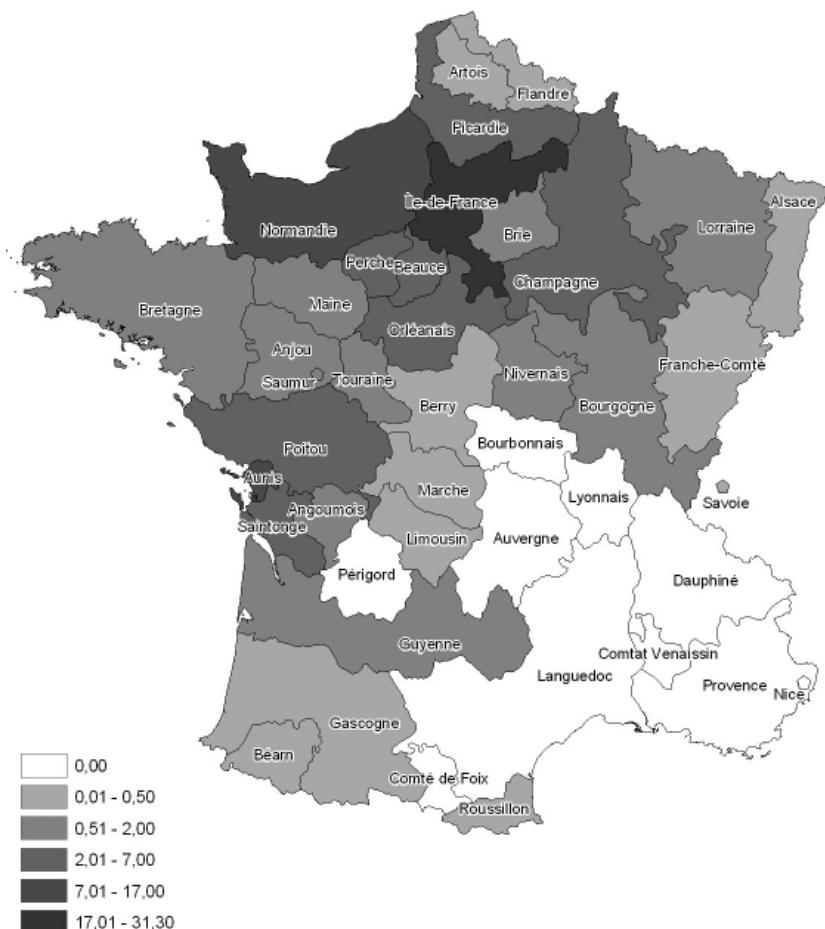


FIGURE 1. Origine des fondatrices françaises (%)

tes les provinces françaises ont fourni au moins un fondateur, ce qui n'est pas le cas du côté des fondatrices (la plupart des provinces du sud-est ne comportent aucune fondatrice).

LA CONTRIBUTION GÉNÉTIQUE DES FONDATRICES ET DES FONDATEURS

Plusieurs fondateurs apparaissent plus d'une fois dans les généalogies, et c'est particulièrement fréquent parmi les plus anciens arrivants. Ainsi, 87,6 pour cent de l'ensemble des fon-

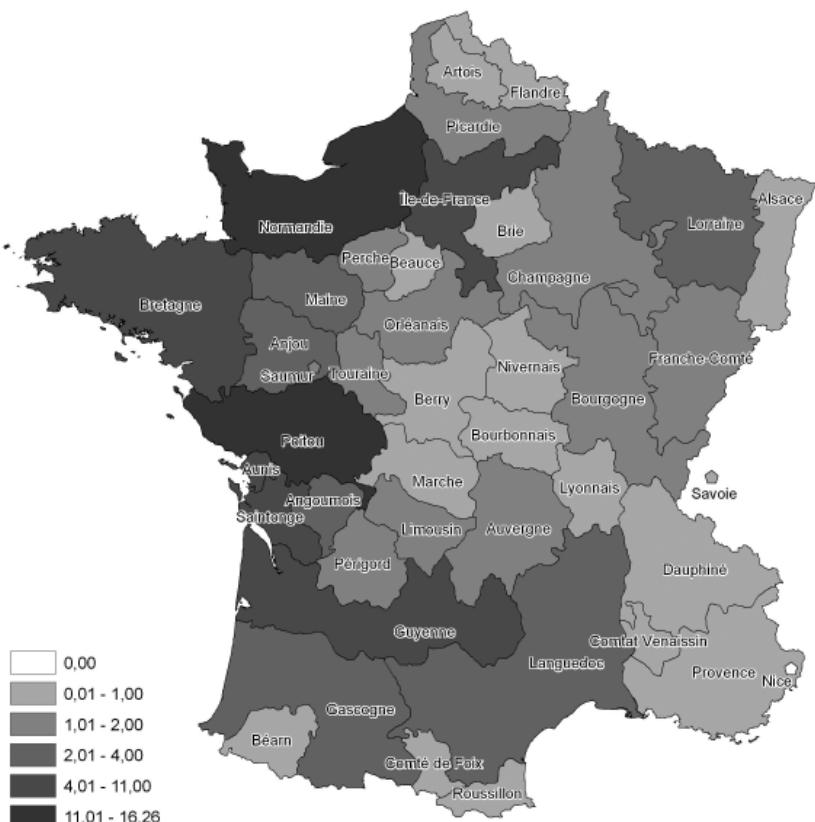


FIGURE 2. Origine des fondateurs français (%)

datrices et 84,1 pour cent de l'ensemble des fondateurs apparaissent dans plus d'une généalogie, et ces proportions atteignent respectivement 98,9 pour cent et 98,4 pour cent pour les fondatrices et fondateurs mariés avant 1700. Les deux tiers des fondatrices et la moitié des fondateurs mariés avant 1700 apparaissent dans plus de 100 générations (soit plus de 5 pour cent des générations), les plus prolifiques d'entre eux atteignant jusqu'à 2050 générations (92 pour cent des sujets). À l'opposé, 98 pour cent des fondatrices et fondateurs mariés après 1765 apparaissent dans 10 générations ou moins.

Les fondateurs qui reviennent le plus souvent auront donc une plus forte probabilité d'avoir transmis des copies de leurs gènes, à travers leurs descendants, à la population contemporaine.

raîne, que ceux qui n'apparaissent qu'une seule fois ou un petit nombre de fois. La mesure de la contribution génétique des fondateurs permet d'illustrer ce phénomène. La contribution génétique (CG) d'un fondateur à un groupe de sujets se calcule de la façon suivante :

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

où S est l'ensemble des sujets liés au fondateur,

P est l'ensemble des chemins généalogiques entre le fondateur et chaque sujet, et

g est le nombre de générations, dans chaque chemin, séparant le fondateur et le sujet.

Par exemple, la contribution génétique d'un grand-parent à ses 10 petits-enfants sera égale à 2,5 et on pourra considérer qu'en moyenne 25 pour cent des gènes dans le groupe des petits-enfants provient de ce grand-parent.

Théoriquement, cette mesure peut aussi être interprétée comme le nombre attendu de copies, dans une cohorte de descendants, d'un gène particulier provenant d'un fondateur donné (Roberts, 1968; O'Brien et al., 1994; Heyer, 1995). Le nombre de copies de ce gène dans l'échantillon de sujets dépendra donc du nombre d'ascendances dans lesquelles se retrouve le fondateur, du nombre de chemins généalogiques entre ce fondateur et chacun des sujets qui y sont reliés et de la distance, en termes de générations, dans chacun de ces chemins généalogiques (dans un chemin donné, la probabilité de transmission sera plus élevée si la distance générationnelle est plus courte). En sommant les valeurs de contribution génétique pour l'ensemble des fondateurs d'une origine donnée, et en divisant le résultat par le nombre total de sujets, on obtient la proportion du pool génique de la population qui provient des fondateurs de cette origine.

La figure 3 montre la grande variabilité de la contribution génétique des fondateurs à la population québécoise. Dans ce graphique, la contribution génétique relative est cumulée, en commençant par les fondateurs ayant les plus fortes contributions. Le premier 1 pour cent des fondateurs explique ainsi 10 pour cent du pool génique, et la moitié de ce pool génique est attribuable à seulement 9 pour cent des fondateurs. À l'opposé, la contribution de 70 pour cent des fondateurs est inférieure à 19 pour cent du pool génique de cette population.

Les valeurs de la contribution génétique selon l'origine des fondateurs sont présentées aux tableaux 6 (femmes) et 7

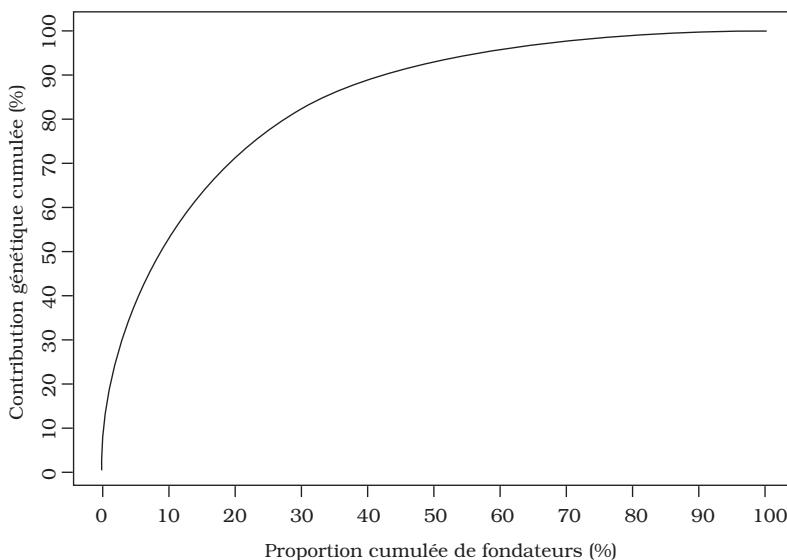


FIGURE 3. Contribution génétique cumulée de l'ensemble des fondateurs (%)

(hommes). Ces résultats font clairement ressortir l'importance de la contribution française à la population du Québec. En effet, 88,7 pour cent de la contribution totale des fondatrices est française, et cette proportion est encore plus élevée parmi les hommes (90,8 pour cent). Les fondatrices acadiennes expliquent 6,4 pour cent de la contribution féminine, tandis que 3,6 pour cent de la contribution masculine provient des Acadiens. Les fondatrices et fondateurs d'autres origines expliquent moins de 2 pour cent de la contribution génétique totale. Signalons que les fondatrices et fondateurs français sont les seuls à avoir une contribution génétique dont la part relative est plus élevée que leur poids relatif en termes de nombres de fondateurs (voir les tableaux 2 et 3). Cette différence relative est encore plus frappante selon les périodes de mariage : la contribution génétique relative des fondateurs mariés avant 1700 est de 78,8 pour cent chez les hommes et de 88,8 pour cent chez les femmes, alors que la proportion des fondateurs et fondatrices mariés durant cette période est de 40,5 et 58,4 pour cent respectivement. Les fondatrices et fondateurs mariés après 1700 ont une contribution génétique proportionnellement trois fois plus faible que la valeur « attendue » selon leurs poids respectifs.

TABLEAU 6. Contribution génétique des immigrantes fondatrices selon l'origine et la période de mariage (%)

Origine	Période de mariage			Ensemble
	Avant 1700	1700- 1765	Après 1765	
France	87,4	0,8	0,5	88,7
Grande-Bretagne	0,2	0,1	0,8	1,0
Allemagne	0,0	0,0	0,1	0,2
Irlande	0,0	0,0	0,5	0,5
Autres pays européens	0,4	0,0	0,1	0,6
<i>Sous-total Europe</i>	<i>88,0</i>	<i>1,0</i>	<i>2,0</i>	<i>90,9</i>
Acadie	0,5	5,2	0,7	6,4
Reste de l'Amérique	0,4	0,6	0,9	1,9
<i>Sous-total Amérique</i>	<i>0,9</i>	<i>5,8</i>	<i>1,7</i>	<i>8,3</i>
Inconnue	0,0	0,2	0,6	0,8
Toutes les origines	88,8	6,9	4,3	100,0

TABLEAU 7. Contribution génétique des immigrants fondateurs selon l'origine et la période de mariage (%)

Origine	Période de mariage			Ensemble
	Avant 1700	1700- 1765	Après 1765	
France	77,3	12,4	1,1	90,8
Grande-Bretagne	0,6	0,3	1,0	1,9
Allemagne	0,0	0,0	0,4	0,5
Irlande	0,1	0,1	0,5	0,6
Autres pays européens	0,8	0,2	0,3	1,4
<i>Sous-total Europe</i>	<i>78,7</i>	<i>13,0</i>	<i>3,4</i>	<i>95,1</i>
Acadie	0,0	3,0	0,6	3,6
Reste de l'Amérique	0,0	0,4	0,4	0,8
<i>Sous-total Amérique</i>	<i>0,0</i>	<i>3,4</i>	<i>1,0</i>	<i>4,4</i>
Inconnue	0,0	0,1	0,4	0,5
Toutes les origines	78,8	16,5	4,8	100,0

L'importance des provinces du nord-ouest de la France en termes de contribution au peuplement du Québec ressort encore lorsque l'on considère la contribution génétique relative des fondatrices et fondateurs français (figures 4 et 5). Les fondatrices d'Île-de-France et les fondateurs de Normandie dominent ceux

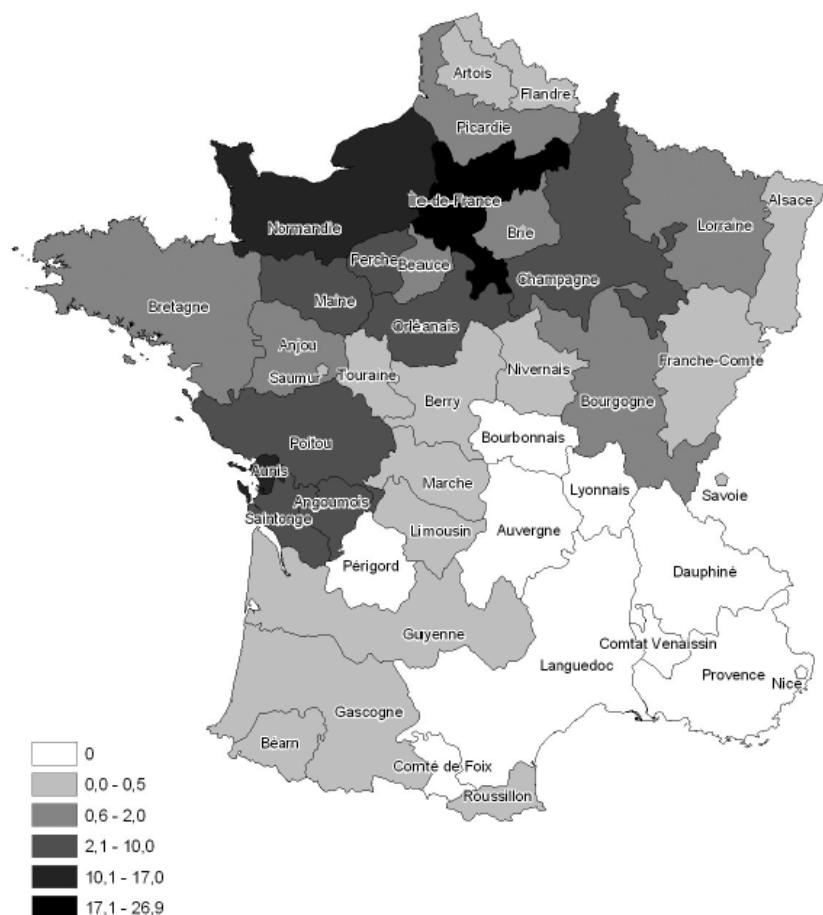


FIGURE 4. Contribution génétique des fondatrices françaises (%)

des autres provinces pour ce qui concerne leur contribution génétique, avec respectivement 27 pour cent et 21 pour cent de la contribution totale des Françaises et des Français. Dans ce cas, toutefois, la contribution des fondatrices d'Île-de-France est proportionnellement moindre que leur poids dans l'ensemble des fondatrices françaises, alors qu'on observe le contraire pour les fondateurs normands. Parmi les autres provinces à forte contribution génétique, notons, chez les fondatrices, l'Aunis (17 pour cent), la Normandie (16 pour cent) et le Perche (9 pour cent) et, chez les fondateurs, le Poitou (13 pour cent), l'Aunis (10 pour cent) et le Perche (9 pour cent). Le cas de la province du Perche est particulièrement intéressant. En effet, bien qu'ils ne repré-

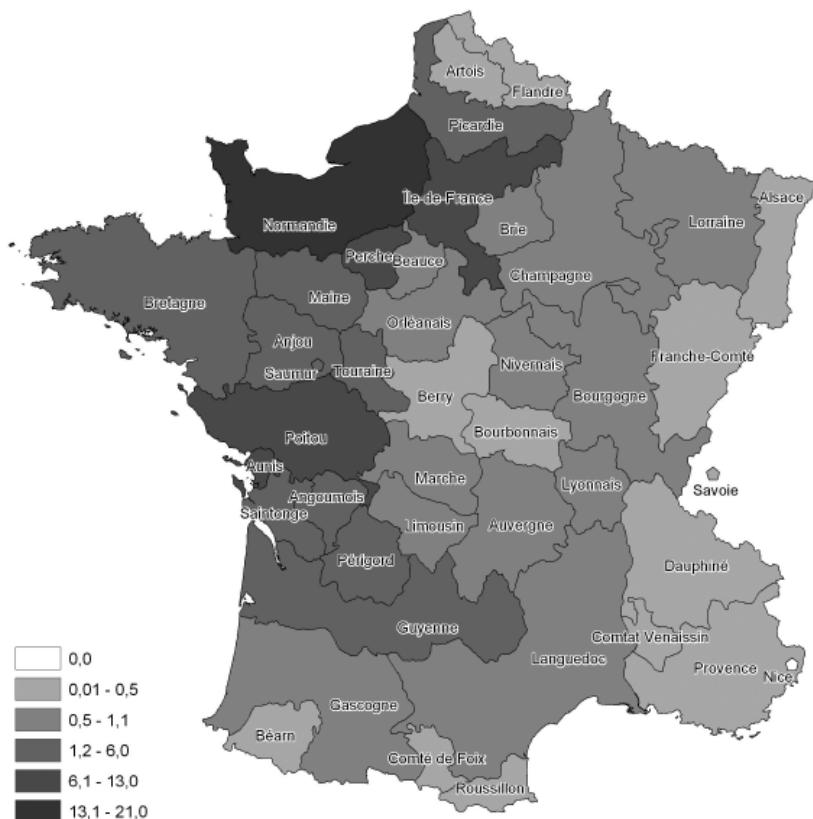


FIGURE 5. Contribution génétique des fondateurs français (%)

sentent que 3 pour cent des fondatrices et moins de 2 pour cent des fondateurs français, la contribution génétique relative des Percheronnes et des Percherons atteint 9 pour cent dans les deux cas, soit trois à quatre fois plus que leurs poids respectifs. En fait, la contribution génétique moyenne des fondateurs et fondatrices du Perche est la plus élevée de toutes. Ensemble, les 93 fondateurs (64 hommes et 29 femmes) originaires de cette petite région rurale du nord-ouest de la France expliquent 7 pour cent du pool génique québécois. Cette importante contribution des fondateurs percherons avait déjà été constatée pour certaines populations de l'est du Québec, notamment celles de Charlevoix et du Saguenay-Lac-Saint-Jean (Heyer et al., 1997; Tremblay et al., 2000).

DISCUSSION

Cette étude sur les origines et la structure généalogique de la population québécoise repose sur l'analyse des généalogies ascendantes de 2223 sujets mariés au Québec entre 1945 et 1965. Le fondateur-immigrant de chaque branche généalogique a été identifié et, dans la mesure du possible, son lieu d'origine et sa période d'arrivée ont été précisés. Les résultats indiquent que les effets fondateurs à l'origine du peuplement de souche européenne sont encore nettement perceptibles dans la population contemporaine. L'analyse des liens d'apparentement éloigné montre que presque tous les Québécois d'ascendance française partagent au moins un ou même plusieurs ancêtres. Certains de ces ancêtres apparaissent beaucoup plus fréquemment dans les généalogies, ce qui signifie que la contribution génétique est extrêmement variable au sein du groupe des immigrants fondateurs. Cette variabilité est apparue dès les toutes premières générations ayant suivi le peuplement initial de la Nouvelle-France. Ainsi, 20 pour cent des pionniers établis entre 1608 et 1680 expliquaient 50 pour cent du pool génique de la population en 1730 (Charbonneau et al., 1987). Parmi les fondateurs identifiés dans les généalogies des individus mariés entre 1780 et 1799, les 8 pour cent ayant la contribution la plus élevée ont une contribution totale de 50 pour cent (Gagnon et Heyer, 2001); ce dernier résultat est pratiquement identique à celui obtenu dans notre corpus (9 pour cent). On constate donc que l'avantage pris par certains des fondateurs initiaux est demeuré insurmontable et que les ancêtres communs les plus fréquents se retrouvent, indéniablement, parmi les premiers immigrants venus de France au cours du 17^e siècle. En effet, ces Français expliquent à eux seuls 81 pour cent de la contribution génétique totale de tous les immigrants-fondateurs identifiés au sein de notre corpus. L'origine de ces fondateurs à l'échelle des provinces, plus diversifiée chez les hommes que chez les femmes, est néanmoins très concentrée, puisque les fondateurs provenant de cinq provinces (Normandie, Île-de-France, Aunis, Poitou et Perche) représentent 65 pour cent de la contribution génétique de tous les fondateurs français. L'importance relative de ces provinces d'origine avait été observée par les chercheurs du PRDH pour l'ensemble des immigrants établis durant le Régime français (Charbonneau et al., 1987; Charbonneau et Robert, 1987).

Par ailleurs, il est intéressant de constater la présence relativement grande (en termes de proportions de généalogies recouvertes) de fondateurs et fondatrices d'origine britannique, irlan-

daise, allemande et acadienne. À notre connaissance, c'est la première fois qu'une mesure aussi précise de la contribution des fondateurs de ces origines au peuplement du Québec est effectuée. En ce qui concerne les ancêtres acadiens, on pourrait bien sûr associer ces individus à des origines françaises, mais il faut souligner que la provenance géographique précise des immigrants qui se sont établis en Acadie ne suit pas nécessairement la même distribution que celle des Français et Françaises qui ont immigré dans la vallée du Saint-Laurent (Charbonneau et al., 1987; Charbonneau et al., 2000).

Un des objectifs de nos travaux est d'estimer la proportion des immigrants qui ont encore des descendants dans la population contemporaine du Québec. Nous avons ainsi effectué une comparaison entre les immigrants identifiés dans notre échantillon et l'ensemble des immigrants retracés dans le Registre de population du Québec ancien du PRDH (Programme de recherche en démographie historique, 2005). Les résultats de cette comparaison, présentés au tableau 8, indiquent que 44 pour cent de l'ensemble des immigrants identifiés par le PRDH sont présents dans notre échantillon. On constate aussi que ce nombre varie fortement selon l'origine et la période de mariage des fondateurs. En effet, presque les deux tiers des immigrants mariés au 17^e siècle ont des descendants parmi nos généalogies, comparativement à 39 pour cent des immigrants mariés entre 1700 et 1765 et à 22 pour cent des immigrants mariés entre la Conquête et la fin du 18^e siècle. Ici encore, ces nombres témoignent de l'avantage pris par les premiers immigrants (presque exclusivement français) venus au 17^e siècle. Des analyses

TABLEAU 8. Comparaison immigrants/immigrants-fondateurs : rapport entre le nombre d'immigrants identifiés dans les 2223 généalogies et le nombre total d'immigrants, selon l'origine et la période de mariage (%)

Origine	Période de mariage			Ensemble
	Avant 1700	1700- 1765	1766- 1799	
France	65	40	33	51
Grande-Bretagne	25	17	22	20
Allemagne	75	29	26	27
Autres pays européens	65	31	12	31
Acadie	44	40	11	26
Autres	68	63	60	62
Toutes les origines	64	39	22	44

plus approfondies devraient nous permettre de distinguer les immigrants qui n'ont pas laissé de descendance et ceux dont la descendance n'est pas représentée dans notre échantillon. On peut facilement supposer que, dans ce dernier cas, il s'agira surtout d'immigrants dont la descendance est moindre que celle des immigrants déjà identifiés.

Il importe enfin de mentionner que, comme notre échantillon généalogique porte sur des sujets mariés entre 1945 et 1965, les résultats obtenus diffèrent sans doute un peu de ceux qui seraient obtenus à partir d'un échantillon de sujets mariés plus récemment. Cependant, ces différences s'expliqueraient principalement par l'apport des immigrants (et de leurs descendants) arrivés au Québec après 1965. Étant donné l'avantage conféré par l'ancienneté dans la probabilité de diffusion des gènes dans la population, les effets sur le pool génique ne pourraient donc qu'être relativement faibles. D'autre part, rappelons qu'environ 12 pour cent de la population québécoise de la période 1945-1965 n'a pas été incluse dans cette étude car nos reconstructions généalogiques reposent essentiellement sur les actes de mariage de la population catholique, qui formait alors environ 88 pour cent de la population. On peut penser que la plupart de ces exclusions concernent des individus d'ascendance britannique (Angleterre, Écosse, Pays de Galles et Irlande du Nord).

La suite des analyses visera à examiner la contribution différentielle des fondatrices et fondateurs aux populations des diverses régions du Québec afin de préciser les conséquences démogénétiques de l'histoire du peuplement suite aux effets fondateurs initiaux. Une attention particulière sera aussi portée aux lignées maternelles et paternelles au sein de chaque généalogie. L'étude de ces lignées généalogiques permettra d'effectuer des comparaisons avec les travaux de génétique moléculaire qui portent sur la transmission de l'ADN mitochondrial (lignées maternelles) et du chromosome Y (lignées paternelles). À terme, ces travaux permettront une caractérisation fine de la contribution génétique des fondateurs de la population québécoise en apportant, entre autres choses, un éclairage nouveau sur les ramifications spatiales des effets fondateurs. En plus d'accroître les connaissances sur l'histoire démographique de la population québécoise, ces résultats pourront être utilisés dans les domaines de la génétique des populations et de l'épidémiologie génétique, dans le cadre de travaux visant à identifier et à comprendre la distribution sur le territoire des déterminants génétiques associés à diverses maladies.

NOTES

- ¹ À ces immigrants s'ajoutent bien sûr les Autochtones qui étaient déjà sur place avant l'arrivée des premiers Européens.
- ² Les fondateurs d'origine autochtone identifiés dans les généalogies ont été regroupés dans la catégorie « Reste de l'Amérique ».

RÉFÉRENCES BIBLIOGRAPHIQUES

- BEAUJOT, R., et D. KERR. 2004. *Population Change in Canada*. Don Mills, Oxford University Press, 2^e édition.
- BOUCHARD, G. 2004. *Projet BALSAC. Rapport annuel 2003-2004*. Université du Québec à Chicoutimi, Projet BALSAC.
- BOUCHARD, G., H. CHARBONNEAU, B. DESJARDINS, É. HEYER et M. TREMBLAY. 1995. « Mobilité géographique et stratification du pool génétique canadien-français sous le Régime français », dans Y. LANDRY, J. A. DICKINSON, S. PASLEAU et C. DESAMA, éd. *Les chemins de la migration en Belgique et au Québec, du XVII^e au XX^e siècle*. Louvain-La-Neuve, Éditions Academia : 51-59.
- CHARBONNEAU, H., et N. ROBERT. 1987. « Origines françaises de la population canadienne, 1608-1759 », dans *Atlas historique du Canada*, vol. I, *Des origines à 1800*. Les Presses de l'Université de Montréal.
- CHARBONNEAU, H., B. DESJARDINS, A. GUILLEMETTE, Y. LANDRY, J. LÉGARÉ et F. NAULT. 1987. *Naissance d'une population : les Français établis au Canada au XVII^e siècle*. Paris et Montréal, Institut national d'études démographiques et Presses de l'Université de Montréal.
- CHARBONNEAU, H., B. DESJARDINS, J. LÉGARÉ et H. DENIS. 2000. « The population of the St-Lawrence Valley, 1608-1760 », dans M. R. HAINES, R. H. STECKEL et al. *A Population History of North America*. Cambridge University Press : 99-142.
- DESJARDINS, B. 1998. « Le Registre de la population du Québec ancien », *Annales de démographie historique*, 2 : 215-226.
- DICKINSON, J. A. 1994. « Les réfugiés acadiens au Québec, 1755-1775 ». *Études canadiennes/Canadian Studies*, 37 : 51-61.
- DUCHESNE, L. 2004. *La situation démographique au Québec, bilan 2004*. Les Publications du Québec.
- GAGNON, A., et É. HEYER. 2001. « Fragmentation of the Quebec population genetic pool (Canada): Evidence from the genetic contribution of founders per region in the 17th and 18th centuries ». *American Journal of Physical Anthropology*, 114, 1 : 30-41.
- GAGNON, A., H. VÉZINA et B. BRAIS. 2001. « Histoire démographique et génétique du Québec », *Pour la Science*, 287 : 62-69.
- GROUPE DE RECHERCHE INTERDISCIPLINAIRE EN DÉMOGRAPHIE ET ÉPIDÉMIOLOGIE GÉNÉTIQUE. 2005. <http://www.uqac.ca/grig>.

- HARTL, D. L. 1994. *Génétique des populations*. Flammarion, Médecine-Sciences.
- HENRIPIN, J., et Y. PÉRON. 1973. « La transition démographique de la province de Québec », dans H. CHARBONNEAU, éd. *La population du Québec : études rétrospectives*. Éditions du Boréal Express : 23-44.
- HEYER, É. 1995. « Genetic consequences of differential demographic behaviour in the Saguenay region, Québec », *American Journal of Physical Anthropology*, 98 : 1-11.
- HEYER, É., M. TREMBLAY et B. DESJARDINS. 1997. « Seventeenth-century European origins of hereditary diseases in the Saguenay population (Quebec, Canada) ». *Human Biology*, 69 : 209-225.
- INSTITUT DE LA STATISTIQUE DU QUÉBEC. 2005. <http://www.stat.gouv.qc.ca/>.
- JOMPHE, M., et B. CASGRAIN. 1997. *Base de données généalogiques RETRO : structure des données*. Chicoutimi, IREP, Programme de recherches en génétique des populations, Document III-C-97.
- LANDRY, Y. 1992. *Orphelines en France, pionnières au Canada. Les Filles du roi au XVII^e siècle*. Leméac.
- MCINNIS, M. 2000. « The population of Canada in the nineteenth century », dans M. R. HAINES, R. H. STECKEL et al. *A Population History of North America*. Cambridge, Cambridge University Press : 371-432.
- O'BRIEN, E., R. A. KERBER, L. B. JORDE et A. R. ROGERS. 1994. « Founder effect: Assessment of variation in genetic contributions among founders », *Human Biology*, 66 : 185-204.
- PICHÉ, V. 2003. « Un siècle d'immigration au Québec : de la peur à l'ouverture », dans V. PICHÉ et C. LE BOURDAIS, éd. *La démographie québécoise. Enjeux du XXI^e siècle*. Les Presses de l'Université de Montréal : 225-263.
- PROGRAMME DE RECHERCHE EN DÉMOGRAPHIE HISTORIQUE. 2005. <http://www.genealogie.umontreal.ca>.
- PROJET BALSAC. 2005. <http://www.uqac.ca/BALSAC>.
- ROBERTS, D. F. 1968. « Genetic effects of population size reduction ». *Nature*, 220 : 1084-1088.
- THOMPSON, E. A. 1986. *Pedigree Analysis in Human Genetics*. Baltimore, Johns Hopkins University Press.
- TREMBLAY, M., N. GAGNON et É. HEYER. 2000. « A genealogical analysis of two eastern Quebec populations », *Canadian Studies in Population*, 27,2 : 307-327.
- TREMBLAY, M., M. JOMPHE et H. VÉZINA. 2001. « Comparaison de structures patronymiques et génétiques dans la population québécoise », dans G. BRUNET, P. DARLU et G. ZEI, dir. *Le patronyme : histoire, anthropologie, société*. Paris, CNRS Éditions : 367-389.
- VÉZINA, H., M. TREMBLAY et L. HOUDE. 2005. « Mesures de l'apparentement biologique au Saguenay-Lac-St-Jean (Québec, Canada) à

partir de reconstitutions généalogiques », *Annales de démographie historique*, 2 : 67-84.

WILHELMY, J.-P. 1984. *Les mercenaires allemands au Québec du XVIII^e siècle et leur apport à la population*. Beloeil, Maison des mots, 331 p.

ABSTRACT

Hélène VÉZINA, Marc TREMBLAY, Bertrand DESJARDINS and Louis HOUDE
ORIGINS AND GENETIC CONTRIBUTIONS OF THE QUÉBEC POPULATION
FOUNDERS

This study analyzes the geographic origins and genetic contributions of the founding ancestors of the Québec population. The authors used data from the BALSAC population register and the Early Quebec population register to build a corpus of 2,223 ascending genealogies covering the entire territory of Québec. These genealogies go back to the 17th century and on average include more than 9 generations. Genealogical links between all individuals identified in the genealogies were established. Results show that approximately 81% of Québec's gene pool derives from the founders who came from France in the 17th century. There are however important differences in founders' genetic contribution according to gender and place of origin in France. Genealogical analyses also reveal the existence of many distant kinship ties within this population.