

L'apport des données génétiques à la mesure généalogique des origines amérindiennes des Canadiens français

**HÉLÈNE VÉZINA^{*}, MICHÈLE JOMPHE^{*}, ÈVE-MARIE LAVOIE^{*},
CLAUDIA MOREAU^{**} ET DAMIAN LABUDA^{**}**

Dans le cadre d'un programme de recherche sur la diversité génétique des populations régionales du Québec, nous avons mesuré la contribution des ancêtres amérindiens au pool génique de 794 participants résidant dans les régions de Montréal, du Saguenay—Lac-Saint-Jean (SLSJ), de la Gaspésie et de la Côte-Nord. Les ancêtres amérindiens ont été identifiés à partir de reconstructions généalogiques (fichier BALSAC) et de données génétiques (ADN mitochondrial) à l'aide d'une approche qui repose sur l'exploitation conjointe de ces deux types de données. Les résultats indiquent que plus de la moitié des participants ont au moins un ancêtre amérindien dans leur généalogie. Ceci veut dire que la majorité des participants sont porteurs de gènes reçus de fondateurs amérindiens. Cependant la contribution génétique totale de ces ancêtres aux quatre groupes régionaux demeure faible. En effet, elle est de moins de 1 % au SLSJ et à Montréal alors qu'elle dépasse à peine 1 % sur la Côte-Nord et en Gaspésie.

English abstract, p. 105

INTRODUCTION

Les Canadiens français sont les descendants des pionniers venus s'établir en Nouvelle-France aux xvii^e et xviii^e siècles. Tout au cours de

* Projet BALSAC, Université du Québec à Chicoutimi, Chicoutimi

** Centre de recherche de l'Hôpital Ste-Justine, Université de Montréal, Montréal

Les auteurs tiennent à remercier les personnes qui ont généreusement accepté de participer à ce projet de recherche ainsi que le Réseau de médecine génétique appliquée du Fonds de recherche en santé du Québec et les Instituts de recherche en santé du Canada qui ont contribué au financement. Nos remerciements vont aussi à Catherine Laprise, grâce à qui nous avons pu obtenir les données pour la région du Saguenay—Lac-Saint-Jean, ainsi qu'à Danielle Gauvreau et aux deux évaluateurs anonymes pour leurs suggestions et commentaires fort pertinents.

cette période et des siècles qui ont suivi, des immigrants d'origines autres que française se sont aussi établis sur le territoire de ce qui est aujourd'hui la province de Québec et se sont intégrés à des degrés divers à la population canadienne-française. Les mesures réalisées sur divers groupes de fondateurs font ressortir l'importance — prévisible — de la contribution française, mais elles montrent aussi que celle des autres immigrants, quoique bien inférieure, n'est pas négligeable. Ainsi Bherer et ses collègues (2011) ont analysé un échantillon de 2221 généalogies de sujets mariés au Québec entre 1945 et 1965 et dont les quatre grands-parents s'étaient mariés au Québec. Ils ont trouvé que chaque sujet avait en moyenne dans sa généalogie des ancêtres de 6,5 origines différentes. Les travaux indiquent aussi que la contribution des fondateurs à la population contemporaine varie selon leur sexe, selon leur période d'arrivée ainsi que selon la région du Québec où sont effectuées les mesures (Bherer et collab., 2011 ; Tremblay, Vézina, Desjardins et Houde, 2008 ; Vézina, Tremblay, Desjardins et Houde, 2005). Cette variabilité interrégionale de l'apport des fondateurs a été étudiée de façon spécifique pour certains groupes comme les Acadiens (Bergeron, Vézina, Houde et Tremblay, 2008), les Irlandais (Tremblay, Letendre, Houde et Vézina, 2009) et les immigrants d'origine germanique (Tremblay, 2010). Enfin, les études génétiques montrent aussi une diversité importante au sein de la population d'origine canadienne-française (De Braekeleer, 1990 ; Moreau et collab., 2007 ; Scriver, 2001).

Au cours des siècles, des Amérindiens ont épousé des Canadiens français et se sont ainsi intégrés à la population d'origine européenne. La mesure de ce métissage et de la contribution des Amérindiens au bassin génétique canadien-français a depuis plusieurs années suscité l'intérêt des chercheurs (Charbonneau, 1990 ; D'Avignon, 2009 ; Desjardins, 1990 ; Savoie, 2000) et soulevé bien des controverses. Les avis sont partagés, certains estimant que cette contribution se situerait à moins de 1 % du pool génique (Hubert Charbonneau cité dans Beauregard, 1993), alors que d'autres considèrent qu'elle est certainement sous-estimée (Denys Delâge cité dans Dubé, 2008) et qu'elle pourrait même être de l'ordre de 5 à 10 % (Jacques Beaugrand cité dans Dubé, 2008). Dans leur étude, Bherer et ses collègues (2011) ont trouvé que 47 % des 2221 généalogies analysées comportaient au moins un fondateur amérindien. Au total, les fondateurs amérindiens représentaient 1,2 % de l'ensemble des fondateurs, mais leur contribution génétique n'était que de 0,2 %.

Une chose est certaine : l'ampleur du métissage entre Amérindiens et Européens (ou leurs descendants) s'avère très difficile à appréhender par les sources démographiques telles que les registres paroissiaux, l'état civil

et les recensements car elles sont toutes incomplètes en ce qui concerne les Amérindiens (Beauregard, 1993 ; Lachance et Savoie, 1996). La mesure de la contribution amérindienne basée sur des reconstructions généalogiques ne peut donc conduire qu'à une sous-estimation et ne peut prétendre être exhaustive ou exacte. En revanche, les avancées récentes de la génétique moléculaire donnent accès à des données d'une autre nature et offrent la possibilité de développer de nouvelles approches pour approfondir les connaissances sur le sujet.

L'objectif de ce travail est double : il s'agit d'abord de présenter les résultats d'une étude visant à mesurer, à l'aide de données généalogiques, la contribution génétique amérindienne à un groupe d'individus résidant dans quatre régions du Québec. Deuxièmement, sur le plan méthodologique, nous voulons montrer en quoi les données génétiques (provenant de l'analyse de l'ADN mitochondrial et du chromosome Y) et démographiques (provenant de reconstructions généalogiques) peuvent être complémentaires et quel peut être l'intérêt d'en faire une analyse conjointe. En particulier, il s'agit de vérifier si les données génétiques peuvent permettre d'évaluer l'ampleur et de corriger au moins partiellement la sous-estimation de la contribution amérindienne mesurée à l'aide des sources généalogiques.

Ainsi, notre démarche vise à exploiter de façon optimale les sources génétiques et démographiques afin d'en arriver à une mesure plus complète mais aussi plus fine de la part de la contribution amérindienne au pool génique de la population canadienne-française du Québec. En effet, nous faisons l'hypothèse que cette contribution, en plus d'être globalement sous-estimée, diffère selon que l'on considère les ancêtres masculins et féminins, et qu'elle varie d'une région à l'autre.

DONNÉES ET MÉTHODES

Nous avons démarré en 2002 un programme de recherche dont l'objectif central est de mesurer la diversité génétique des populations régionales du Québec¹. Notre approche repose sur l'analyse conjointe de données moléculaires et généalogiques. À ce jour, près de 800 individus âgés de 18 ans et plus et résidant dans quatre régions du Québec (île de Montréal,

1. Pour plus d'information, voir le site de l'Échantillon de référence québécois — Épidémiologie génétique et génétique des populations du Québec à l'adresse : www.quebecgenpop.ca

Saguenay—Lac-Saint-Jean, Gaspésie et Côte-Nord)² ont accepté de participer à notre étude et le recrutement est en cours dans trois autres régions (région de Québec, Abitibi-Témiscamingue et Lanaudière). Les critères de participation varient légèrement afin de prendre en compte certaines caractéristiques de l'histoire du peuplement propre à chaque région mais, globalement, une personne est éligible si on peut reconstruire sa généalogie au moins jusqu'aux quatre grands-parents et si elle n'a pas d'apparentement proche avec un autre participant³. Tous les participants ont fourni les informations nécessaires à la reconstruction de leur généalogie et un échantillon de sang ou de salive permettant l'extraction et l'analyse de leur ADN⁴.

Analyses généalogiques

La reconstitution des généalogies des participants a été effectuée en remontant aussi loin que les sources le permettaient c'est-à-dire la plupart du temps jusqu'aux premiers immigrants sur le territoire québécois ou jusqu'aux ancêtres dont on a pu retracer les parents. Les actes de mariage informatisés et jumelés contenus dans le fichier BALSAC (Vézina, 2010) constituent la source principale des données généalogiques. En plus des noms et prénoms, les actes de mariage comportent des informations sur les lieux de résidence et dans certains cas sur les lieux d'origine des époux et de leurs parents. On retrouve aussi parfois des annotations du célébrant lorsque ce dernier a jugé important de consigner certains renseignements comme par exemple le fait d'être Amérindien. Toutes ces informations sont utilisées pour documenter le statut migratoire (natif ou immigrant) ainsi que l'origine des ancêtres présents dans les généalogies. Ceci nous a permis d'identifier des ancêtres amérindiens qui ont contribué au pool génique canadien-français en épousant un membre de ce groupe.

Afin de maximiser notre capacité à identifier ces ancêtres, des sources complémentaires ont été utilisées, telles que, pour la période du régime français, le Registre de population du Québec ancien (RPQA) du Programme de recherche en démographie historique de l'Université de

2. Les analyses présentées ici portent sur 794 individus : 160 résidant sur l'île de Montréal, 173 au Saguenay—Lac-Saint-Jean, 384 en Gaspésie et 77 sur la Côte-Nord.

3. Le niveau d'apparentement accepté entre deux sujets ne devait pas dépasser la valeur de 0,125, soit l'équivalent d'un lien oncle|tante — neveu|nièce.

4. Les généalogies ont été reconstruites au Projet BALSAC à l'Université du Québec à Chicoutimi et les échantillons d'ADN ont été analysés au laboratoire de Damian Labuda à l'Hôpital Ste-Justine à Montréal.

Montréal, qui contient les actes de baptême, mariage et sépulture de cette période (Desjardins, 1998). On a aussi eu recours (mais de façon non systématique) aux recensements canadiens ainsi qu'à des dictionnaires et à des sites internet généalogiques. Ces derniers ont été utilisés avec prudence car la qualité de leur contenu est variable. Toutes les généalogies ont été vérifiées et validées en portant une attention particulière aux données provenant d'autres sources que le fichier BALSAC ou le RPQA.

Dans chaque généalogie, nous avons identifié les fondateurs (immigrants ou individus au-delà desquels il n'est pas possible de poursuivre la généalogie) et nous avons calculé la contribution génétique de chacun de ces fondateurs aux participants de chacune des régions en utilisant la formule suivante :

$$CG = \sum_s \sum_p (1/2)^g$$

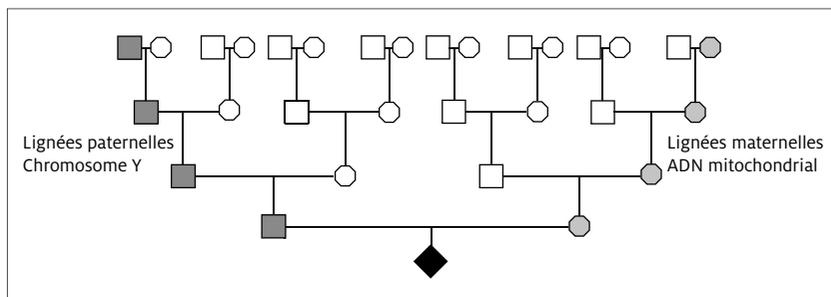
où s est le nombre de sujets reliés à un ancêtre,

p est le nombre de chemins généalogiques entre le sujet et l'ancêtre,

g est le nombre de générations entre le sujet et l'ancêtre pour chacun des chemins généalogiques.

La contribution génétique d'un fondateur représente donc la part du bassin génétique d'un groupe de sujets provenant de ce fondateur. On peut ensuite sommer ces contributions pour estimer la part du pool génique provenant d'un ensemble de fondateurs, par exemple ici les fondateurs d'origine amérindienne.

Mentionnons enfin qu'il est possible de mesurer la contribution génétique des fondateurs en considérant l'ensemble des données généalogiques ou en regardant de façon particulière les lignées paternelles (constituées exclusivement d'hommes) et les lignées maternelles (constituées exclusivement de femmes). Ces lignées (figure 1) permettent entre autres de mesurer l'impact différentiel des ancêtres masculins et féminins et de comparer les résultats avec ceux obtenus à partir des données génétiques provenant du chromosome Y (pour les lignées paternelles) et de l'ADN mitochondrial (pour les lignées maternelles).

FIGURE 1 Identification des lignées paternelles et maternelles dans une généalogie

Analyses génétiques

À partir d'un échantillon sanguin (Gaspésie et Saguenay—Lac-Saint-Jean) ou de salive (Montréal et Côte-Nord) recueilli auprès de chaque participant, l'ADN a été extrait et analysé comme le décrivent Moreau et collab. (2009). Les analyses génétiques présentées ici ont porté sur des séquences provenant du chromosome Y, qui est transmis de père en fils, et de l'ADN mitochondrial, qui est transmis de la mère à ses enfants (donc de mère en fille au fil des générations puisque les fils le reçoivent de leur mère mais ne peuvent le transmettre). Une séquence correspond à un groupe d'allèles situés côte à côte sur un même chromosome et généralement transmis ensemble à la génération suivante. Ces séquences ne sont pas identiques parmi les participants, car au cours de l'histoire des mutations se sont produites chez certains ancêtres. Ceci permet le regroupement des séquences au sein d'haplogroupes qui sont des lignages différenciés par la présence ou l'absence de ces mutations. Chaque séquence peut ainsi être classée au sein d'un haplogroupe qui constitue la signature génétique d'ancêtres très anciens⁵, et on peut considérer que les individus qui appartiennent au même haplogroupe possèdent une origine géographique commune (à l'échelle d'un continent ou d'un grand ensemble géographique). Il a donc été possible de déterminer, parmi nos échantillons, les séquences d'origine amérindienne introduites dans la population canadienne-française par des ancêtres présents dans les lignées paternelles et maternelles.

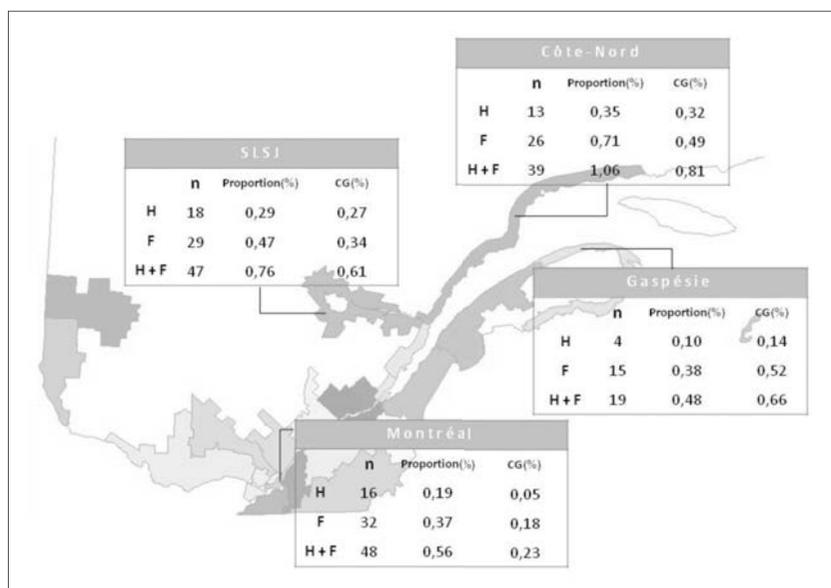
5. Pour une description plus détaillée de l'utilisation des marqueurs du chromosome Y et de l'ADN mitochondrial en généalogie génétique, voir Johnston et Thomas (2003).

RÉSULTATS

Fondateurs amérindiens retracés à l'aide des sources généalogiques

On retrouve à la figure 2 le nombre de fondateurs et de fondatrices d'origine amérindienne identifiés à l'aide des sources généalogiques dans chacune des quatre régions à l'étude. Le nombre de fondatrices amérindiennes est plus élevé que le nombre de fondateurs de cette origine dans toutes les régions, avec un facteur de 1,5 au Saguenay—Lac-Saint-Jean, de 2 à Montréal et sur la Côte-Nord et de 4 en Gaspésie. C'est dans cette dernière région que la proportion des fondateurs amérindiens (hommes et femmes ensemble) est la plus faible par rapport à l'ensemble des fondateurs suivi de Montréal, du Saguenay—Lac-Saint-Jean et de la Côte-Nord. Dans les quatre régions, cette proportion est peu élevée, puisqu'il n'y a que sur la Côte-Nord qu'elle dépasse légèrement le 1 %. Enfin, dans toutes les régions sauf la Gaspésie, la contribution génétique relative des fondateurs d'origine amérindienne est moindre que leur fréquence relative. Rappelons que dans le calcul de la fréquence tous les fondateurs ont le même poids, alors que dans celui de la contribution génétique le poids de chaque fondateur est modulé par son nombre d'apparitions dans la généalogie et

FIGURE 2 Fondateurs d'origine amérindienne d'après les sources généalogiques, selon la région

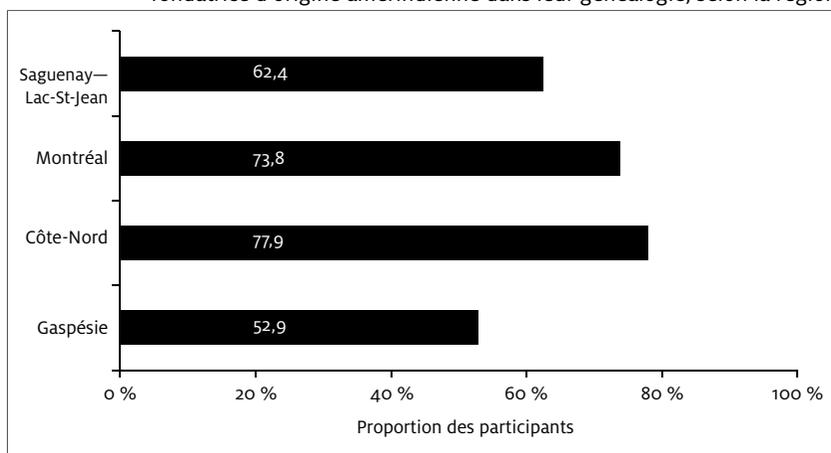


Abréviations: CG, contribution génétique; F, femmes; H, hommes; n, nombre

par la distance générationnelle qui le sépare des sujets auxquels il est relié. Une contribution génétique relative plus faible indique donc un plus petit nombre d'apparitions ou une plus grande distance générationnelle chez les fondateurs amérindiens par rapport à l'ensemble des fondateurs.

La figure 3 indique, pour chacune des régions, la proportion des généalogies où l'on a identifié au moins un fondateur ou une fondatrice d'origine amérindienne. On constate que cette proportion dépasse les 50 % dans toutes les régions, et atteint 78 % sur la Côte-Nord. C'est donc dire que dans les régions de Montréal et de la Côte-Nord les trois quarts des sujets ont au moins un fondateur d'origine amérindienne qui a contribué à leur bagage génétique. Une comparaison avec les résultats de la figure 2 nous permet aussi d'observer qu'il n'existe pas de lien clair entre la contribution à l'ensemble du groupe et la proportion des individus de ce groupe qui ont au moins un ancêtre d'origine amérindienne.

FIGURE 3 Proportion de participants ayant au moins un fondateur ou une fondatrice d'origine amérindienne dans leur généalogie, selon la région



Lignées amérindiennes identifiées à l'aide des analyses génétiques

L'analyse de l'ADN mitochondrial (pour l'ensemble des participants) et du chromosome Y (pour les participants masculins) a permis de déterminer à quel haplogroupe chaque participant appartenait, et donc d'identifier ceux qui ont une origine amérindienne via la lignée respectivement maternelle et paternelle. Pour l'ADN mitochondrial, on distingue 5 haplogroupes qui définissent l'origine amérindienne (A, B, C, D et X2a). Pour le chromosome Y, les haplogroupes Q et C3 sont associés à une origine amérindienne.

Au total, 47 des 794 participants ont un ADN mitochondrial, et donc une lignée maternelle, d'origine amérindienne alors qu'un seul individu (originaire de la région gaspésienne) sur les 305 participants masculins possède un chromosome Y possiblement d'origine amérindienne⁶. Compte tenu de ces résultats, nous avons choisi pour la suite de l'étude de nous concentrer sur les lignées maternelles, dont la répartition par région et par haplogroupe est présentée au tableau 1. Cette distribution varie énormément passant de 0,6 % des lignées au Saguenay à 10 % en Gaspésie. C'est l'haplogroupe C qui est le plus représenté, avec 31 des 47 lignées, et c'est en Gaspésie que l'on retrouve 29 de ces 31 lignées.

TABEAU 1 Lignées amérindiennes identifiées à l'aide de l'ADN mitochondrial et fondatrices généalogiques, selon l'haplogroupe et la région

Régions	Haplogroupes									
	A		C		D		X2a		Total	
	LIG	FON	LIG	FON	LIG	FON	LIG	FON	LIG	FON
SLSJ			1	1					1 (0,6%)	1
Montréal	4	3					1	1	5 (3,1%)	4
Côte-Nord	2	2	1	1					3 (3,9%)	3
Gaspésie	5	3	29	3	4	3			38 (9,9%)	9

Abréviations : FON, fondatrice généalogique ; LIG, lignée amérindienne, SLSJ, Saguenay—Lac-Saint-Jean.

Le tableau 1 indique aussi pour chaque région le nombre de fondatrices généalogiques identifiées dans les lignées maternelles amérindiennes. En effet, les reconstructions généalogiques nous permettent de retracer les liens qui unissent — à la profondeur temporelle correspondant aux sources généalogiques évidemment — les participants qui partagent le même haplogroupe. Dans la région de Montréal, deux participants ont pu être reliés par leur branche généalogique maternelle, alors qu'en Gaspésie

6. En ce qui concerne le chromosome Y, les analyses effectuées permettent de distinguer entre une origine européenne et une origine asiatique ou amérindienne. Pour déterminer avec certitude si l'origine est asiatique ou amérindienne, il faut faire des analyses complémentaires.

les 39 participants descendent de 9 fondatrices, dont 28 participants qui descendent de 2 fondatrices⁷. Si l'on considère ensemble les 47 participants issus d'une lignée mitochondriale amérindienne, on retrouve en fait 15 fondatrices distinctes, puisque des participants nord-côtiers, saguenéens et gaspésiens porteurs de l'haplogroupe C sont reliés à la même fondatrice.

L'utilisation des données génétiques mitochondriales permet aussi d'effectuer une validation de l'exhaustivité et de la qualité des sources généalogiques en ce qui concerne l'identification des ancêtres amérindiens. Ainsi, d'après les résultats des analyses de l'ADN mitochondrial des participants, 15 fondatrices ont une origine amérindienne confirmée. Parmi celles-ci, cinq sont également confirmées par les sources généalogiques soit par une mention directe de leur appartenance amérindienne dans l'acte de mariage, soit à cause d'un patronyme clairement amérindien.

Pour tenter d'expliquer pourquoi nous n'avions pu retracer l'origine amérindienne des dix autres fondatrices, nous sommes retournés aux informations contenues dans le fichier BALSAC — et donc tirées des actes de mariage. Nous avons également consulté à nouveau les autres sources pour vérifier si certaines informations pouvaient être erronées ou encore nous avoir échappé. Nous avons pu observer que sept fondatrices apparaissent dans des lignées généalogiques interrompues, pouvant donc cacher des fondatrices amérindiennes. En effet, la reconstruction de ces lignées a dû être interrompue à cause d'un mariage introuvable, de l'absence du nom des parents dans un acte ou de l'illégitimité d'un enfant. Finalement, dans les trois derniers cas, les fondatrices sont des immigrantes d'origine française selon les sources généalogiques et il n'y a pas de raison de croire que ces données sont fausses. Il subsiste donc une contradiction entre les informations provenant des sources généalogiques et celles provenant des sources génétiques, ce qui nous amène à conclure qu'un lien généalogique est erroné dans la lignée, vraisemblablement à cause d'un cas d'adoption non déclarée.

Utilisation des données génétiques pour la mesure généalogique de la contribution amérindienne

En plus de nous donner les moyens de détecter des erreurs ou des imprécisions dans les sources généalogiques, les données d'ADN mitochondrial montrent l'existence d'une ancêtre commune et donc d'un apparentement entre les participants par leur lignée maternelle à une profondeur tempo-

7. Le détail des résultats sur la Gaspésie est présenté dans Moreau et collab. (2011).

relle éloignée (figure 4A). Les résultats, dans la mesure où il n’y a pas eu d’erreur dans la production et l’analyse des données en laboratoire, sont certains au sens où ils confirment qu’il existe bel et bien un apparentement biologique entre les individus qui partagent le même haplogroupe. Par contraste, les données généalogiques demeurent probabilistes, à cause du risque de fausse paternité ou maternité et d’adoption non déclarée, cas qui ne peuvent être détectés lorsque l’on établit les liens généalogiques.

Malgré ces risques d’erreurs (qui sont jugés très faibles), les données généalogiques permettent quant à elle d’identifier les ancêtres communs et les liens précis d’apparentement entre les participants à la profondeur historique qu’offrent les sources disponibles (figure 4B).

Enfin, l’utilisation conjointe des deux types de données rend possible l’identification de l’ensemble des participants qui sont les descendants d’une fondatrice identifiée dans une lignée maternelle à l’aide des données génétiques. En effet, grâce aux sources généalogiques, on peut retracer tous les liens présents dans une généalogie et non pas seulement ceux qui se trouvent au sein des lignées maternelles (figure 4C).

FIGURE 4 Apport conjoint des données génétiques et des données généalogiques dans la construction des lignées

A) Apport des données génétiques provenant de l’ADN mitochondrial pour établir les liens ancestraux anciens entre des participants via les lignées maternelles

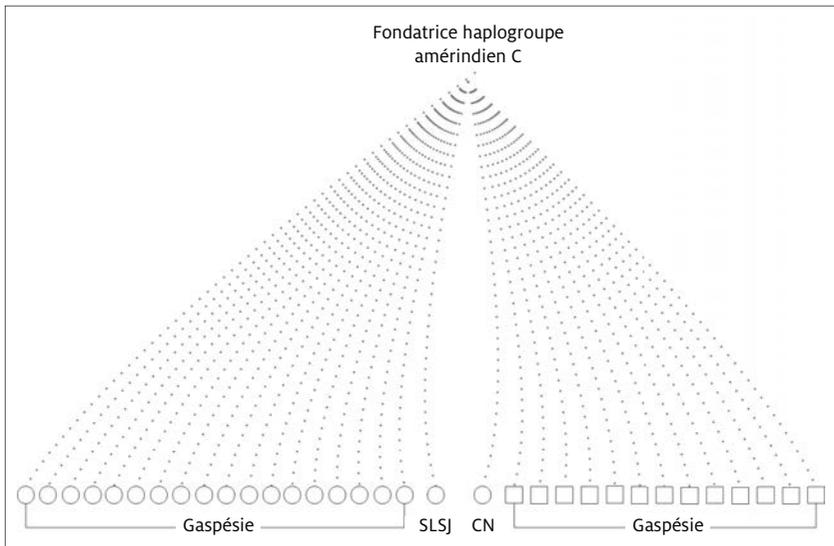
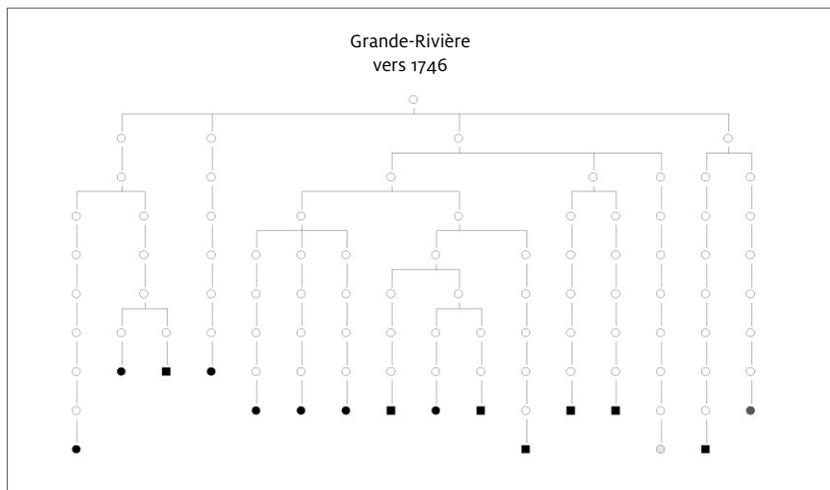
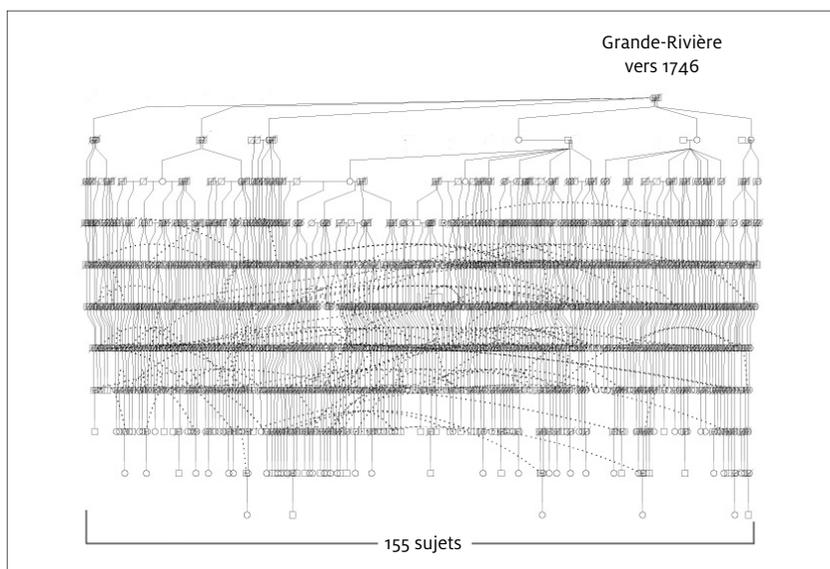


FIGURE 4 Apport conjoint des données génétiques et des données généalogiques dans la construction des lignées (suite)

B) Apport des données généalogiques provenant des lignées maternelles pour l'identification des fondatrices amérindiennes et des liens d'apparementement à une profondeur historique



C) Exploitation conjointe des deux types de données (génétiques et généalogiques) pour l'identification de l'ensemble des lignées généalogiques descendantes d'une fondatrice amérindienne identifiée par l'ADN mitochondrial



En utilisant cette approche, nous avons recalculé la contribution génétique des fondateurs amérindiens ainsi que la proportion des participants qui ont au moins un fondateur d'origine amérindienne dans leur généalogie en ajoutant les nouvelles fondatrices amérindiennes identifiées par l'ADN mitochondrial et en reconstituant tous les liens généalogiques qui les unissent aux participants. Les résultats qui sont présentés à la figure 5 montrent l'augmentation observée dans chaque région en lien avec le gain d'information associé à notre approche. En ce qui concerne la contribution génétique, c'est sur la Côte-Nord et en Gaspésie que ces gains sont les plus marqués, puisque dans les deux cas la contribution totale a augmenté de 50 %. Elle passe ainsi à 1,2 % du bassin génétique des participants sur la Côte-Nord, et à 1,1 % en Gaspésie. On constate par ailleurs que la proportion des participants qui ont au moins un fondateur amérindien dans leur généalogie n'augmente pas ou très peu dans ces deux régions. Ceci indique que dans la plupart des cas les nouvelles fondatrices identifiées par l'ADN mitochondrial sont présentes dans des généalogies où l'on retrouvait déjà au moins un fondateur amérindien. Sur l'île de Montréal, l'augmentation de la contribution génétique est moins marquée (0,07 %), cette dernière atteignant ainsi 0,3 %. C'est dans cette région que l'ajout des nouvelles lignées amérindiennes a cependant un impact plus marqué avec une proportion de participants ayant au moins un fondateur amérindien qui passe de 74 à 85 %. Enfin, au Saguenay—Lac-Saint-Jean on n'observe aucun changement dans les résultats puisque la fondatrice identifiée par l'ADN mitochondrial était déjà connue par les sources généalogiques.

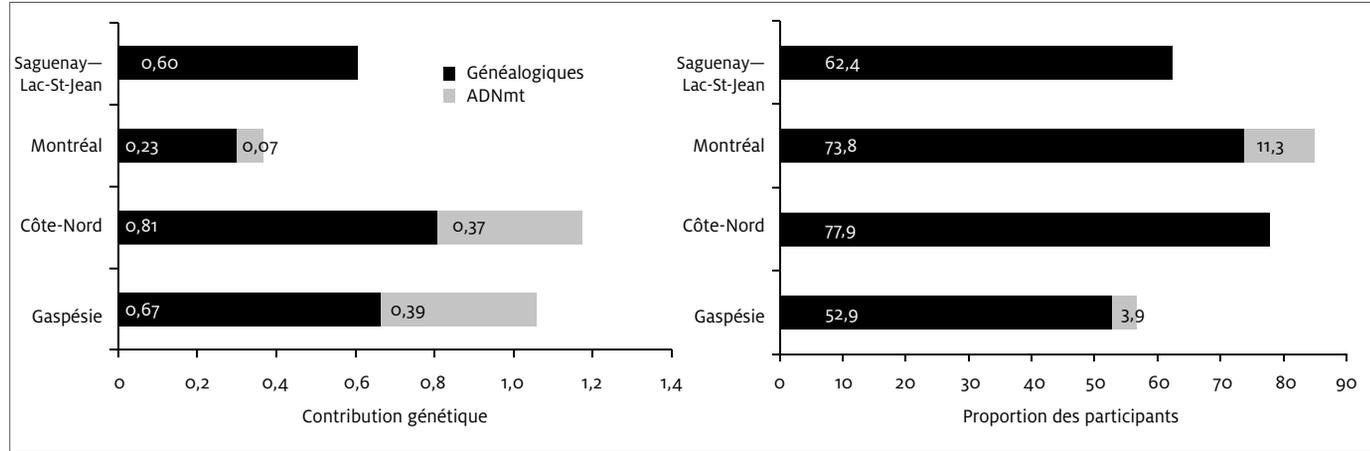
DISCUSSION

Dans le cadre d'un programme de recherche sur le patrimoine génétique des populations régionales du Québec, nous avons mesuré la contribution des ancêtres amérindiens aux généalogies de 794 participants résidant sur l'île de Montréal, au Saguenay—Lac-Saint-Jean, en Gaspésie et sur la Côte-Nord. Les ancêtres amérindiens ont été identifiés à partir de sources généalogiques et des données génétiques provenant de l'analyse de l'ADN mitochondrial des participants. Les résultats indiquent que, dans chacune des régions, plus de la moitié des participants ont au moins un ancêtre amérindien dans leur généalogie, et cette proportion atteint même 85 % sur l'île de Montréal. Bien que la majorité des participants soient porteurs de gènes reçus de fondateurs amérindiens, la contribution génétique totale de ces ancêtres aux quatre groupes régionaux demeure cependant assez faible. En effet, elle est de moins de 1 % au Saguenay—Lac-Saint-Jean et

FIGURE 5 Mesures de l'apport amérindien en combinant données mitochondriales et généalogiques

A) Contribution génétique des fondateurs amérindiens (%)

B) Proportion des participants ayant au moins un fondateur amérindien (%)



Abréviations : ADNmt, ADN mitochondrial; SLSJ, Saguenay—Lac-Saint-Jean.

sur l'île de Montréal et dépasse à peine cette valeur sur la Côte-Nord et en Gaspésie.

En plus d'un essai de mesure de la contribution amérindienne, ce travail avait aussi comme objectif de montrer la complémentarité des données génétiques et généalogiques et l'intérêt d'en faire une exploitation conjointe. En particulier nous voulions vérifier si les données génétiques permettaient de corriger au moins partiellement la sous-estimation de la contribution amérindienne mesurée à l'aide des sources généalogiques. L'analyse des séquences de l'ADN mitochondrial a conduit à l'identification de 47 lignées maternelles d'origine amérindienne chez les 794 participants (soit 6 %) et permis d'attribuer une origine amérindienne à sept fondatrices dont l'origine n'avait pu être déterminée par les recherches généalogiques. Nous avons ainsi pu corriger nos estimations de la contribution amérindienne.

Un seul chromosome Y sans doute d'origine amérindienne a été détecté à l'aide des données génétiques chez les 305 participants masculins (soit 0,3 %). Un autre fondateur amérindien a pu être repéré avec les données généalogiques provenant des lignées paternelles. Ce sont des valeurs nettement plus faibles que celles que nous avons trouvées dans les lignées maternelles avec les deux types de données. Cette différence entre fondateurs et fondatrices amérindiens au sein des lignées est aussi beaucoup plus marquée que celle observée parmi les fondateurs retracés dans les généalogies complètes (voir la figure 2). On considère généralement que le nombre d'unions entre des femmes amérindiennes et des hommes européens a été plus important que celui entre des hommes amérindiens et des femmes européennes (Lachance et Savoie, 1996). Cependant, l'ampleur des différences observées ici commande une étude plus approfondie, ce que nous projetons de faire à une étape ultérieure de ce projet.

Les données génétiques utilisées permettent de cibler uniquement les fondateurs apparaissant dans les lignées maternelles ou paternelles. Par conséquent, tous les fondateurs amérindiens qui auraient pu laisser des traces dans d'autres branches généalogiques (ou autres régions du génome) sont manquants. Il est aussi fort probable que des fondateurs amérindiens présents dans les autres branches n'aient pas été identifiés par les sources généalogiques. Les résultats obtenus en termes de nombre d'ancêtres et de contribution génétique doivent donc être considérés comme des estimations sûres au sens de la qualité des sources utilisées, mais minimales au sens où elles ne couvrent pas toutes les sources d'identification possibles.

D'autres raisons peuvent expliquer que nos résultats ne soient pas complets. D'abord, nous n'avons couvert que quatre régions québécoises

et, comme chaque région a une histoire de peuplement qui lui est propre, en particulier en ce qui concerne l'importance des mariages mixtes entre Amérindiens et Européens (et leurs descendants), on ne peut prétendre que nos résultats soient représentatifs de l'ensemble du Québec du point de vue de la répartition spatiale. Déjà, parmi les quatre régions étudiées, on constate une importante variabilité interrégionale, variabilité que nous avons aussi observée dans une étude précédente sur la composition génétique de ces populations régionales (Roy-Gagnon et collab. 2011). De plus, à cause de nos critères de recrutement, notre échantillon de participants est formé presque complètement d'individus d'origine canadienne-française⁸. On ne peut donc prétendre non plus couvrir toute la population québécoise sous l'angle de l'origine ethnoculturelle. Enfin, il importe de rappeler que ce que nous avons mesuré donne un aperçu des conséquences génétiques de l'intégration des Amérindiens à la population canadienne-française mais que nous ne pouvons par contre rien dire sur la contribution européenne aux pools géniques des divers groupes amérindiens. Cette contribution est généralement considérée comme étant plus importante (surtout du côté masculin), mais elle demeure impossible à mesurer avec des sources démographiques comme celles de l'état civil (Lachance et Savoie, 1996).

À cet égard, les données génétiques offrent des possibilités extrêmement intéressantes, qui ont entraîné le développement d'une nouvelle approche pour connaître ses origines et rechercher ses ancêtres qu'on appelle la généalogie génétique⁹. En effet, chaque individu contemporain porte dans son ADN les traces de son histoire génétique, ce qui permet d'en comprendre certaines facettes même en l'absence de sources historiques ou archéologiques. Dans cette étude, les analyses effectuées avec les données génétiques provenant de l'ADN mitochondrial ont permis la détection de nouvelles fondatrices et donc de nouvelles lignées amérindiennes. Elles ont aussi permis de confirmer en partie les conclusions tirées des sources généalogiques. De leur côté, les données généalogiques, en plus de fournir des informations sur les origines ancestrales et géographiques « récentes » des participants, rendent possible l'identification de liens

8. L'échantillon gaspésien est plus diversifié sur le plan de l'origine ethnoculturelle des participants, et on y a observé des différences importantes entre les groupes étudiés concernant la contribution amérindienne (Moreau et collab., 2011)

9. La généalogie génétique est devenue très populaire ces dernières années. Malgré son grand intérêt, elle soulève tout de même un certain nombre de questions. Voir par exemple Johnston et Thomas (2003) et Wolinsky (2006).

précis d'apparemment entre ceux qui partagent une même fondatrice. Par l'utilisation conjointe des deux sources d'information, nous avons pu améliorer les estimations de la contribution amérindienne à la population québécoise.

En conclusion, nos travaux se poursuivent et nous n'avons pas épuisé toutes les sources disponibles pour identifier des ancêtres amérindiens. Nous pouvons cependant déjà avancer qu'une grande partie des Québécois comptent des Amérindiens parmi leurs ancêtres et sont donc porteurs de gènes d'origine amérindienne. Cependant, si on considère l'ensemble des fondateurs qui ont participé à la constitution du pool génique des Québécois d'origine canadienne-française, l'impact de la contribution amérindienne est faible et le chiffre précédemment avancé de 1 % pourrait s'avérer assez juste.

BIBLIOGRAPHIE

- BEAUREGARD, Y. 1993. « Mythe ou réalité. Les origines amérindiennes des Québécois : entrevue avec Hubert Charbonneau », *Cap-aux-diamants*, 34 : 38-42.
- BERGERON, J., H. VÉZINA, L. HOUDE et M. TREMBLAY. 2008. « La contribution des Acadiens au peuplement des régions du Québec », *Cahiers québécois de démographie*, 37,1 : 181-204.
- BHERER, C., D. LABUDA, M. H. ROY-GAGNON, L. HOUDE, M. TREMBLAY et H. VÉZINA. 2011. « Admixed ancestry and stratification of Quebec regional populations », *American Journal of Physical Anthropology*, 144, 3 : 432-441.
- CHARBONNEAU H. 1990. « Le caractère français des pionniers de la vallée laurentienne », *Cahiers québécois de démographie*, 19, 1 : 49-62.
- D'AVIGNON, M. 2009. « La place des "autres" dans l'histoire du Québec », dans D'AVIGNON, MATHIEU et C. GIRARD (dir.), *A-t-on oublié que jadis nous étions « frères » ?*, Québec, Les Presses de l'Université Laval : 9-29.
- DE BRAEKELEER, M. 1990. « Homogénéité génétique des Canadiens français du Québec : mythe ou réalité ? », *Cahiers québécois de démographie*, 19, 1 : 29-48.
- DESJARDINS, B. 1990. « Homogénéité ethnique de la population québécoise sous le Régime français », *Cahiers québécois de démographie*, 19, 1 : 63-76.
- DESJARDINS, B. 1998. « Le Registre de population du Québec ancien », *Annales de démographie historique*, 2 : 215-226.
- DUBÉ, C. 2008. « Un génome bien accommodant », *Québec Sciences*, été 2008 : 41-43.
- JOHNSTON, J. et M. THOMAS. 2003. « Summary : the science of genealogy by genetics », *Developing World Bioethics*, 3, 2 : 103-108.
- LACHANCE, A. et S. SAVOIE. 1996. « Les Amérindiens sous le régime français », dans A. LACHANCE (dir.), *Les Marginaux, les exclus et l'autre au Canada aux XVII^e et XVIII^e siècles*, Montréal, Fides : 181-200.

- MOREAU, C., H. VÉZINA, M. JOMPHE, E.-M. LAVOIE, M.-H. ROY-GAGNON et D. LABUDA. 2011. « When genetics and genealogies tell different stories—maternal lineages in Gaspesia », *Annals of Human Genetics* 75, 2 : 247-254.
- MOREAU, C., H. VÉZINA et D. LABUDA. 2007. « Founder effects and genetic variability in Quebec », *Médecine sciences MS*, 23, 11 : 1008-1013.
- MOREAU, C., H. VÉZINA, V. YOTOVA, R. HAMON, P. DE KNIJFF, D. SINNETT et D. LABUDA. 2009. « Genetic heterogeneity in regional populations of Quebec—parental lineages in the Gaspé Peninsula », *American Journal of Physical Anthropology*, 139, 4 : 512-522.
- ROY-GAGNON, M.-H., C. MOREAU, C. BHERER, P. ST-ONGE, D. SINNETT, C. LAPRISE, H. VEZINA et D. LABUDA. 2011. « Genomic and genealogical investigation of the French Canadian founder population structure », *Human Genetics*, 129,5 : 521-531.
- SAVOIE, S. 2000. « Métissage et alliances à l'époque de la Nouvelle-France », Conférence présentée à la *Société historique de St-Romuald*, novembre.
- SCRIVER, C. R. 2001. « Human genetics : lessons from Quebec populations », *Annual Review of Genomics and Human Genetics*, 2 : 69-101.
- TREMBLAY, M. 2010. « La contribution des immigrants d'origine germanique au peuplement des régions de Lanaudière, de la Mauricie, de la Montérégie, de Chaudière-Appalaches et du Bas-Saint-Laurent », *Cahiers québécois de démographie*, 39, 2 : 179-200.
- TREMBLAY, M., M. LETENDRE, L. HOUDE et H. VEZINA. 2009. « The Contribution of Irish Immigrants to the Quebec (Canada) Gene Pool : An Estimation Using Data from Deep-Rooted Genealogies », *European Journal of Population Revue européenne de Démographie*, 25, 2 : 215-233.
- TREMBLAY, M., H. VÉZINA, B. DESJARDINS et L. HOUDE. 2008. « Distant kinship and founder effects in the Quebec population », dans T. BENGTSOON et G.P. MINEAU (dir.), *Kinship and demographic behavior in the past*, Springer : 259-277.
- VÉZINA, H. 2010. *Projet BALSAC : Rapport annuel 2009-2010*, 43 p.
- VÉZINA, H., M. TREMBLAY, B. DESJARDINS et L. HOUDE. 2005. « Origines et contributions génétiques des fondatrices et des fondateurs de la population québécoise », *Cahiers québécois de démographie*, 34, 2 : 235-258.
- WOLINSKY, H. 2006. « Genetic genealogy goes global », *EMBO Reports*, 7 : 1072-1074.

ABSTRACT

Hélène VÉZINA, Michèle JOMPHE, Ève-Marie LAVOIE, Claudia MOREAU and Damian LABUDA

The contribution of genetic data to the genealogical measurement of the Amerindian origins of French Canadians

In the context of a research programme on the genetic diversity of regional populations in Quebec, we measured the contribution of Amerindian ancestors to the gene pool of 794 participants living in the Montreal, Saguenay-Lac-St-Jean (SLSJ), Gaspé Peninsula and North Shore regions. Amerindian ancestors were identified through genealogical reconstructions (BALSAC database) and genetic data (mitochondrial DNA) using an approach relying on the joint exploitation of the two types of data. Results show that more than half of the participants have at least one Amerindian ancestor in their genealogy. This means that a majority of participants carry genes received from Amerindian founders. However the total genetic contribution of these ancestors to the four regional groups is low, reaching less than 1% in SLSJ and Montreal and slightly more than 1% on the North Shore and in the Gaspé Peninsula.