

Sur la distribution spatiale des gènes délétères dans la région du Saguenay (XIXe et XXe siècles)

Gérard Bouchard

Volume 32, Number 85, 1988

URI: <https://id.erudit.org/iderudit/021926ar>

DOI: <https://doi.org/10.7202/021926ar>

[See table of contents](#)

Publisher(s)

Département de géographie de l'Université Laval

ISSN

0007-9766 (print)

1708-8968 (digital)

[Explore this journal](#)

Cite this article

Bouchard, G. (1988). Sur la distribution spatiale des gènes délétères dans la région du Saguenay (XIXe et XXe siècles). *Cahiers de géographie du Québec*, 32(85), 27–47. <https://doi.org/10.7202/021926ar>

Article abstract

Based on the Saguenay Genealogical Database, various sets of maps were drawn in order to bring out models or structures of genes distribution across the Saguenay territory. From a genetic epidemiology standpoint, this research had a very practical interest in that it was aimed at validating spatial data as potentially useful risk indicators for the prevention of recessive genetic disorders in this particular population. The answer seems to be negative. With a few exceptions, the mutant genes causing the disease do not show patterns of distribution sharp enough to allow a quite safe prediction of genetic phenomena. This finding is somewhat surprising since a set of indices — like the distribution of endogamy, consanguinity, kinship, etc. — suggest the existence of some spatial genes clustering.

SUR LA DISTRIBUTION SPATIALE DES GÈNES DÉLÉTÈRES DANS LA RÉGION DU SAGUENAY (XIX^e-XX^e SIÈCLES)

par

Gérard BOUCHARD

*SOREP¹, Université du Québec à Chicoutimi,
555, boul. de l'Université, Chicoutimi, G7H 2B1*

RÉSUMÉ

À l'aide du fichier-réseau de la population du Saguenay et par le biais de diverses reconstitutions cartographiques, la présente recherche vise à faire ressortir la structure spatiale du bassin génétique saguenayen. Du point de vue de l'épidémiologie génétique, cet exercice poursuit un objectif très concret : il s'agit de voir si la référence spatiale peut constituer un indicateur de risque utile pour la prévention des maladies récessives parmi cette population. Il semble que ce ne soit pas le cas. Sauf exceptions très localisées, les gènes mutants qui sont à l'origine de ces maladies ne présentent pas une distribution géographique qui permettrait de prédire avec une certaine fiabilité l'occurrence des phénomènes. Ce résultat est un peu surprenant dans la mesure où des indices comme la distribution de l'endogamie, de la consanguinité, de la parenté, etc. suggèrent l'existence d'une structure ou d'une stratification spatiale.

MOTS-CLÉS : Région du Saguenay, gènes récessifs, stratification spatiale, indicateur de risque, épidémiologie génétique, fichier de population.

ABSTRACT

Spatial Distribution of Defective Genes in the Saguenay Region (19th-20th century)

Based on the Saguenay Genealogical Database, various sets of maps were drawn in order to bring out models or structures of genes distribution across the Saguenay territory. From a genetic epidemiology standpoint, this research had a very practical interest in that it was aimed at validating spatial data as potentially useful risk indicators for the prevention of recessive genetic disorders in this particular population. The answer seems to be negative. With a few exceptions, the mutant genes causing the disease do not show patterns of distribution sharp enough to allow a quite safe prediction of genetic phenomena. This finding is somewhat surprising since a set of indices — like the distribution of endogamy, consanguinity, kinship, etc. — suggest the existence of some spatial genes clustering.

KEY WORDS : Saguenay region, recessive genes, spatial stratification, risk indicator, genetic epidemiology, genealogical database.

*

*

*

ESPACE ET GÉNÉTIQUE²

Les régions du Nord-Est québécois, tout particulièrement le Saguenay³ et Charlevoix, présentent des prévalences et incidences élevées pour certaines maladies héréditaires très rares ou inexistantes dans d'autres populations. Ce phénomène n'est pas unique puisqu'on en trouve l'équivalent notamment en Finlande (Norio *et al*, 1973) et dans certaines régions de Suède⁴. Il a cependant attiré depuis plus de vingt ans l'attention des chercheurs, l'un des premiers étant le pédiatre-généticien Claude Laberge (1967, 1969). Depuis quelques années, dans le cadre de son Programme de recherches en génétique humaine, le centre interuniversitaire SOREP a voulu prolonger ces efforts en effectuant une série d'enquêtes sur les populations du Saguenay et de Charlevoix. Relevant à la fois de l'histoire, de la démographie et de la génétique, ces études mettent à profit le fichier-réseau de la population dont la constitution a débuté en 1972 à l'Université du Québec à Chicoutimi⁵. Elles visent deux objectifs généraux. Le premier consiste à rendre compte de la formation et de l'évolution du bassin génétique dans ces deux régions. Dans ce cas, il s'agit surtout d'expliquer la diffusion de génopathies rares ou inexistantes ailleurs. Le deuxième objectif, auquel se rattache le présent article, relève de la recherche appliquée. Il consiste à éclairer le modèle de distribution des gènes mutants ou délétères⁶ dans la population, de manière à rendre plus efficace la prévention des maladies génétiques.

Problématique spatiale du bassin génétique saguenayen

Les gènes sont portés par des individus qui les transmettent d'une génération à l'autre. Mais cette transmission se fait rarement au hasard et son cours peut être infléchi de diverses façons. Par exemple, certaines descendance se reproduisent plus vite que d'autres, soit parce que leur fécondité est plus élevée, soit parce qu'elles sont moins affectées par l'émigration ou pour toute autre raison. Dans un cas comme dans l'autre, on verra croître la fréquence de certains patronymes qui, éventuellement, pourraient devenir des indicateurs utiles dans une perspective épidémiologique, s'il est prouvé que ces descendance sont associées à tel ou tel gène mutant.

De la même façon, la transmission des gènes récessifs peut donner lieu à des subdivisions dans l'espace. La composition du bassin génétique s'exprime alors par une stratification ou une structure spatiale qui fait que, par exemple, certaines aires géographiques deviennent plus homogènes que d'autres et présentent des traits spécifiques. À une échelle plus fine, il se peut que la répartition géographique d'un gène mutant en particulier épouse des cloisonnements de ce genre.

Durant les quinze dernières années, cet axe de recherche a donné lieu à d'importantes contributions en génétique des populations et en épidémiologie génétique. Les gènes se diffusent dans l'espace au gré de mouvements migratoires qui ne sont pas nécessairement aléatoires, en particulier lorsqu'ils sont influencés par les solidarités de parenté. La notion de subdivisions génétiques est ici fondamentale (Li, 1976, chap. 25). Les formes qu'elles prennent sont évidemment influencées par les facteurs géographiques (reliefs, cours d'eau, etc.). L'une des formes les plus caractéristiques est le « cline », sorte de gradient spatial créé par des migrations en chaîne qui diffusent les gènes par vagues successives dans une même direction. Très tôt, les généticiens se sont donc intéressés aux effets de différenciation entraînés par les migrations (Fix, 1978 ; Harrison et Boyce, 1972 ; Jorde, 1980 ; Gradie et Gauvreau, 1987). À l'aide de tests statistiques comme les mélanges de distributions ou l'autocorrélation spatiale (Cliff et Ord, 1981 ;

Sokal et Oden, 1978a, 1978b), il s'agit de faire ressortir des stratifications parmi une population, donc d'y distinguer des sous-groupes.

Ces enquêtes poursuivent deux buts. Sur un premier plan, d'ordre fondamental, on cherche à comprendre les processus qui commandent la structuration du *pool* génique. L'étude des matrices migratoires parents-enfants en est une bonne illustration (cf. Jorde, 1982; Gradie *et al*, 1988). Sur un second plan, qui relève de la recherche appliquée, on s'intéresse en particulier à la diffusion de gènes délétères afin de localiser le risque des génopathies dans une population. Il peut arriver que ce risque soit très facile à localiser; c'est souvent le cas avec des minorités et des isolats ethniques ou religieux comme les Amish (McKusick *et al*, 1985), les Huttérites (Morgan *et al*, 1983) ou les communautés ethniques, juives notamment. Ces dernières sont en effet étroitement associées à la maladie de Tay-Sachs, ce qui a donné lieu à des démarches préventives de type « community approach » (cf. Clow et Scriver, 1977; Motulsky, 1980; Goodman, 1980). Dans ce cas, le travail de l'épidémiologue est facilité du fait que le risque est clairement circonscrit par un caractère ethnique. La thalassémie, qui est une maladie récessive très répandue parmi les populations méditerranéennes, en est un autre exemple, en particulier lorsqu'elle est dépistée parmi des minorités ethniques établies dans des métropoles européennes ou américaines (Mouzouras *et al*, 1980; Scriver *et al*, 1984; Wagener *et al*, 1978).

Il en va bien autrement lorsque le risque se présente sous une forme diffuse parmi une grande population dont les subdivisions ne sont pas d'emblée visibles. Il faut alors se livrer à des reconstitutions cartographiques qui visent à spatialiser le risque, en l'occurrence les malades, les porteurs hétérozygotes et leur dynamique migratoire. Un travail de ce genre a été effectué dans certains pays, notamment en Finlande (Norio *et al*, 1973). Au Québec, des maladies comme la tyrosinémie (Laberge, 1969) et la dystrophie oculo-pharyngée (Barbeau, 1966) ont fait l'objet de recherches. Tous ces travaux ont cependant été réalisés à l'échelle interrégionale. L'enquête que nous présentons ici se situe à une échelle plus fine et plus difficile puisque nous essayons de déceler une variabilité du risque à l'échelle intrarégionale. Hormis l'étude réalisée par Ginter *et al* (1980), on connaît peu d'essais de ce genre appliqués à des gènes récessifs.

Prenant la population du Saguenay pour objet, nous voulons donc établir si le bassin génétique, et plus spécialement les gènes mutants récessifs qui en font partie, se distribuent géographiquement selon un modèle ou une structure quelconque qui permettrait de mieux cerner le risque d'occurrence des génopathies. On perçoit aisément l'enjeu pratique d'une telle démarche: rappelons qu'en matière d'épidémiologie génétique, la prévention est le mot-clé puisque ces maladies sont présentement incurables. Celle-ci exige toutefois le recours à divers indices permettant d'identifier ce qu'on pourrait appeler les zones à haute susceptibilité.

Les indicateurs utilisés

Deux séries d'indicateurs sont ici mis à l'essai. Les uns sont des indicateurs d'inférence qui témoignent indirectement de la structure générale du *gene pool*, dans la mesure où ils révèlent des lignes ou des aires préférentielles de transmission génique. C'est le cas par exemple des mesures d'endogamie, qui expriment la propension à trouver son conjoint dans sa propre paroisse ou à l'extérieur de celle-ci. C'est le cas des indicateurs d'apparentement comme la consanguinité (le fait que deux individus pris au hasard dans une population aient un ou des ancêtres communs). C'est aussi le

cas des mesures de distance, notamment entre les lieux de naissance ou de résidence de deux conjoints.

À une autre échelle, l'analyse s'appuie sur une deuxième série d'indicateurs reflétant directement la distribution géographique des gènes mutants. Pour diverses périodes et à l'échelle de micro-régions, les résidences des porteurs (hétérozygotes) connus de gènes récessifs ont été cartographiées de manière à produire une représentation longitudinale de leur répartition. Six maladies ont été retenues à cette fin. Deux sont des maladies neuro-musculaires (agénésie du corps calleux, ataxie spastique Charlevoix — Saguenay) ; deux autres affectent les enfants en très bas âge et peuvent entraîner un décès après quelques années (tyrosinémie, fibrose kystique) ; une autre compromet la croissance osseuse chez l'enfant (rachitisme vitamino-dépendant) et la dernière consiste dans une accumulation de fer dans le sang, à partir de l'âge adulte (hémochromatose).

Dans un premier temps, pour chacune de ces maladies, il est utile de savoir d'où provenaient les immigrants ayant introduit le gène mutant dans la population du Saguenay⁷. À partir de l'ensemble des dossiers de patients (N = 239 en 1984-1985), les ascendances ont été reconstruites par ordinateur. Dans chacun des six fichiers ainsi constitués, on appelle fondateurs les ancêtres les plus lointains, c'est-à-dire ceux au-delà desquels l'information généalogique s'arrête. Certains de ces ancêtres fondateurs apparaissent à la tête de deux, trois, quatre ascendances ou plus et, pour cette raison, ils ont donc une probabilité beaucoup plus élevée d'avoir introduit et transmis le gène. Nous considérons comme présumés porteurs ces ancêtres fondateurs qui, dans chacun des six fichiers, apparaissent dans deux ascendances ou plus. Le tableau 1 permet de retracer leur région de provenance. On constate que plus le nombre d'occurrences est élevé, plus la région de Charlevoix est représentée. Autrement dit, les fondateurs les plus « lourds », ceux qui ont le plus contribué à diffuser les gènes mutants au Saguenay, venaient presque tous de Charlevoix.

À la lumière de cette importante indication, le problème consiste à éclairer les modalités, notamment spatiales, de cette transmission génétique au Saguenay, à partir du milieu du XIX^e siècle jusqu'à récemment.

LA STRUCTURE SPATIALE DU BASSIN GÉNÉTIQUE : LES INDICATEURS D'INFÉRENCE

Pour les fins de la cartographie, le territoire saguenayen a été découpé en fonction de diverses échelles. Ces dernières permettent une spatialisation plus ou moins fine selon la question posée et le volume des effectifs à représenter. Dans les cartes qui suivent, trois unités ou divisions spatiales sont utilisées.

- Les unités de résidence de base (URB) : au nombre de 66, elles correspondent aux limites des municipalités actuelles, rurales et urbaines, dont plusieurs sont issues de fusions plus ou moins récentes. Il a été fait exception à cette règle pour 3 ou 4 localités en raison de leur importance historique ou démographique.
- Les regroupements municipaux (RM) : ce sont 19 unités spatiales, comprenant de 2 à 7 petites municipalités, apparentées sur la base de la taille et des activités économiques.
- Les micro-régions (MR) : entre les regroupements municipaux et les trois grandes sous-régions (cf. note 3), un découpage intermédiaire s'imposait. Il

prend la forme de 8 divisions délimitées en fonction de : a) certaines caractéristiques géographiques (barrières naturelles, voies de communication), b) l'existence d'un pôle urbain⁸.

Tableau 1

Régions de provenance des couples fondateurs présumés porteurs *

Maladies/régions de provenance	Nombre d'ascendances dans lesquelles le couple fondateur est représenté									
	1		2		3		4 et plus		Total	
	Nombre absolu	%	Nombre absolu	%	Nombre absolu	%	Nombre absolu	%	Nombre absolu	%
Tyrosinémie										
Charlevoix	281	60,7	110	84,6	37	90,2	27	87,1	455	68,4
Non-Charlevoix	182	39,3	20	15,4	4	9,8	4	12,9	210	31,6
Ataxie spastique										
Charlevoix	320	79,6	67	87,0	38	88,4	16	94,1	441	81,8
Non-Charlevoix	82	20,4	10	13,0	5	11,6	1	5,9	98	18,2
Agénésie du corps calleux										
Charlevoix	296	75,5	78	88,6	35	89,7	23	85,2	432	79,1
Non-Charlevoix	96	24,5	10	11,4	4	10,3	4	14,8	114	20,9
Rachitisme vitamino-dépendant										
Charlevoix	186	77,8	26	96,3	6	100	—	—	216	80,0
Non-Charlevoix	53	22,2	1	3,7	—	—	—	—	54	20,0
Fibrose kystique										
Charlevoix	267	66,3	68	83,0	28	100	20	83,3	383	71,3
Non-Charlevoix	136	33,7	14	17,0	—	—	4	16,7	154	28,7
Hémochromatose										
Charlevoix	81	72,3	7	77,8	6	100	2	100	96	74,4
Non-Charlevoix	31	27,7	2	22,2	—	—	—	—	33	25,6

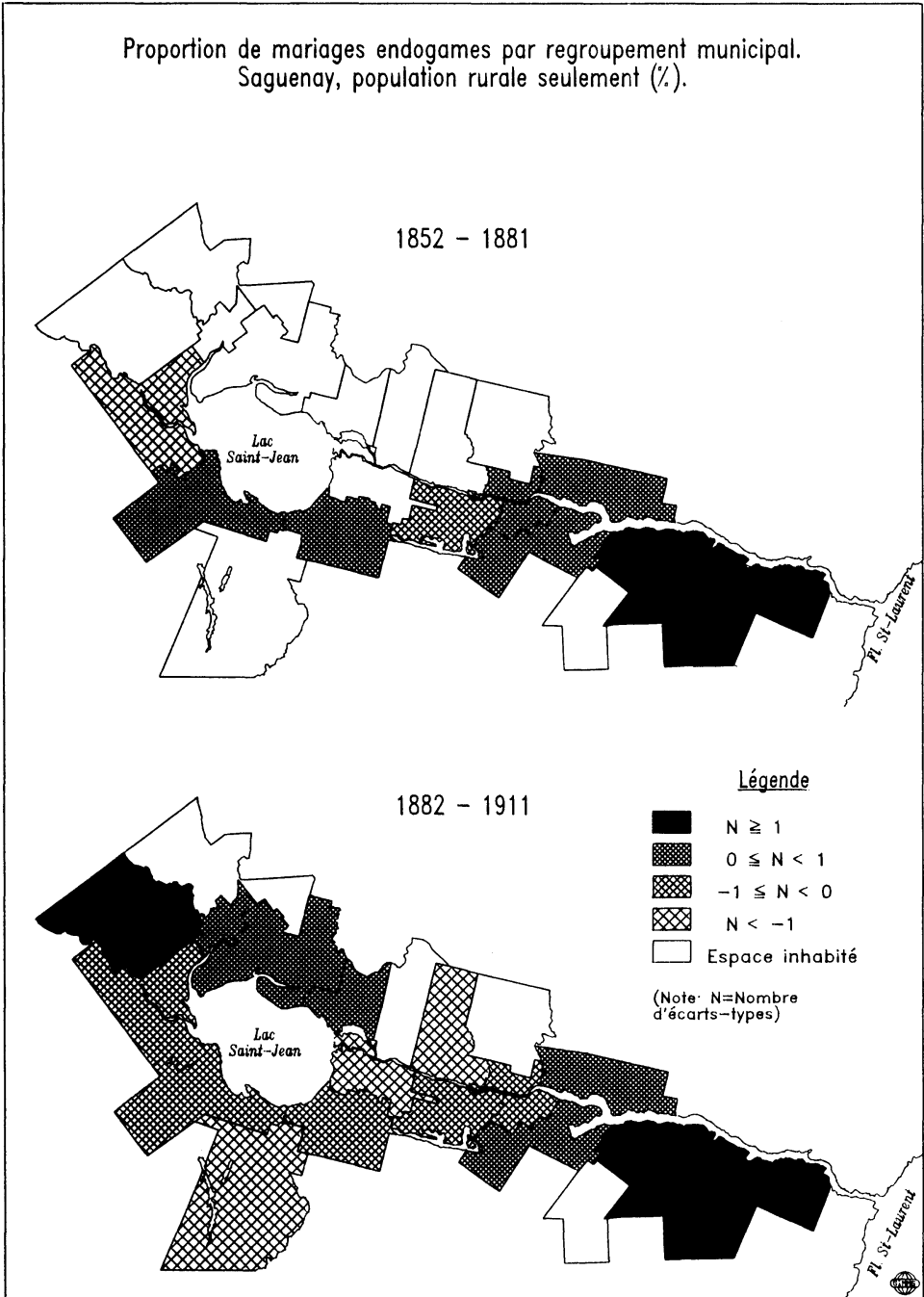
(SOREP)

* Selon le nombre d'ascendances dans lesquelles ils apparaissent : Saguenay, 6 maladies récessives.

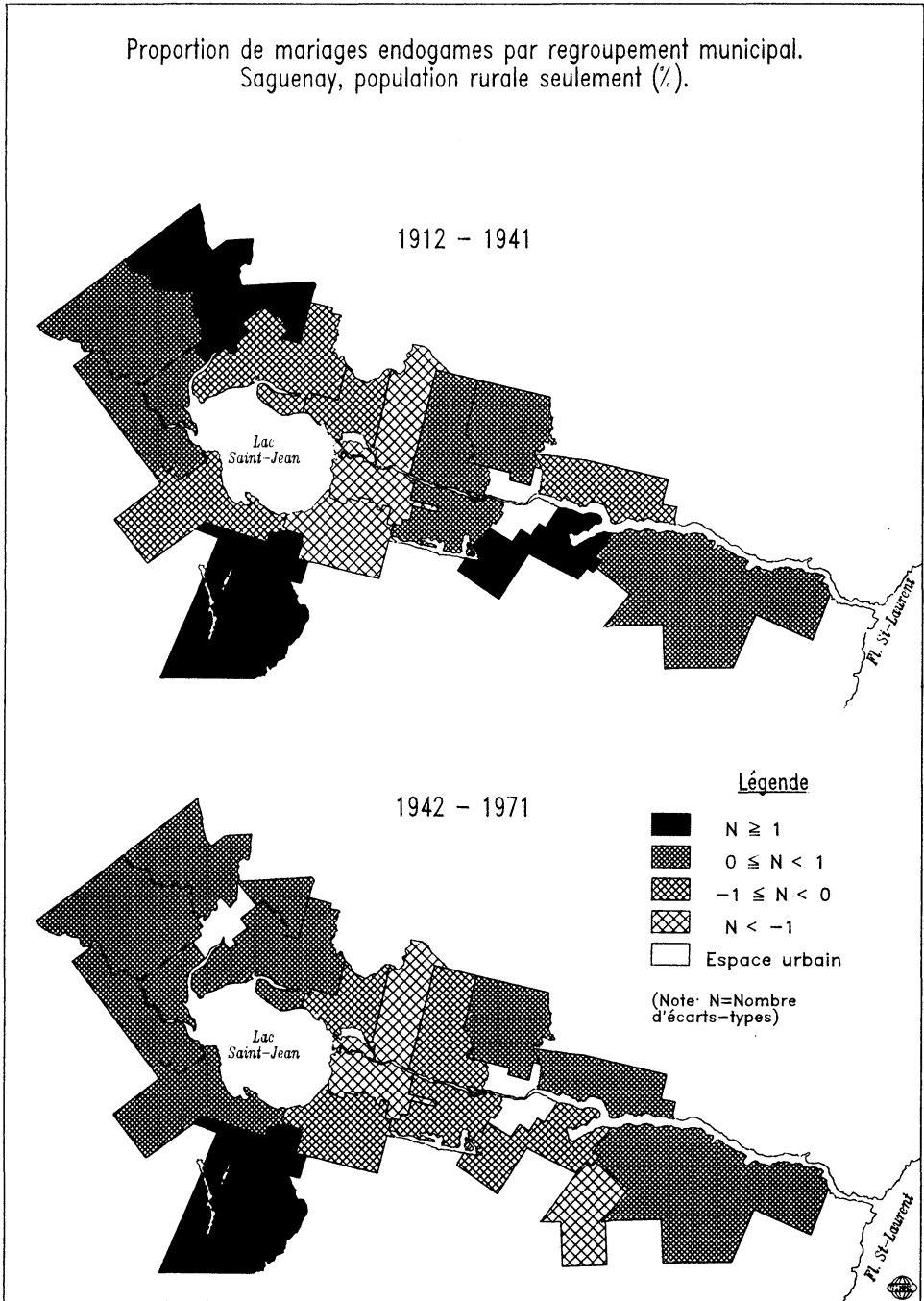
Au lieu d'effectuer toute l'étude à l'échelle relativement large des micro-régions, il a semblé utile d'intégrer un cadre plus fin d'analyse lorsque cela était possible, soit celui des RM ou des URB. En effet, chaque série cartographique doit être vue comme une tentative spécifique pour faire ressortir des clivages, d'où l'intérêt de leur diversité. Ceci dit, les trois échelles utilisées sont compatibles puisque chaque micro-région se décompose en regroupements municipaux, et chacun de ceux-ci en unités de résidence de base.

Par ailleurs, il convient de rappeler que les indicateurs d'inférence présentés ici sont tous susceptibles de biais. Par exemple, les mesures d'endogamie sont sensibles à la taille des paroisses ou autres unités spatiales utilisées ; les coefficients de consanguinité et de parenté varient en fonction de la profondeur des généalogies ; les distances entre lieux de résidence ne tiennent pas toujours pleinement compte de la mobilité ultérieure, etc. Ces indicateurs sont en outre limités, chacun d'eux ayant été conçu pour mesurer

Figure 1



Proportion de mariages endogames par regroupement municipal.
Saguenay, population rurale seulement (%).



des phénomènes très spécifiques. C'est pourquoi il a semblé prudent de recourir à plusieurs indices, nous ménageant ainsi la possibilité de dégager des convergences.

L'endogamie

La proportion des mariages endogames (les deux conjoints résident dans la même paroisse) célébrés dans une paroisse est révélatrice des limites géographiques circonscrivant le flux génique, dans la mesure où celui-ci est influencé par les migrations matrimoniales. Plus de 90 000 mariages ont été célébrés dans la région du Saguenay entre 1842 et 1971, mais pour la mesure de l'endogamie, il a fallu exclure les paroisses urbaines. À cause de leur densité démographique, celles-ci ne peuvent être comparées aux paroisses rurales.

Au cours des quatre sous-périodes étudiées, l'indice d'endogamie diminue sensiblement (figure 1)⁹. Les zones à forte intensité se concentrent d'abord à l'est, sur la rive sud. Au gré du peuplement, elles se déplacent progressivement vers l'ouest, en particulier au sud et au nord-ouest du lac Saint-Jean (1912-1941). L'indice s'uniformise ensuite (1942-1971), sauf dans le RM Lac-Bouchette, au sud du lac Saint-Jean. En résumé, à la lumière de ce premier indicateur, trois unités spatiales se démarquent : à l'est, le RM Bas-Saguenay — Rive-sud, celui du Lac-Bouchette au sud et, à un degré moindre, deux RM du Nord-Ouest (Normandin et Dolbeau — Mistassini). On retient que, dans ces trois unités spatiales et durant quelques décennies au moins, les alliances matrimoniales ont pu favoriser la mise en place d'un cloisonnement génétique¹⁰.

Le coefficient de consanguinité¹¹

La distance isonymique (ou distance de Lasker) entre les 19 regroupements municipaux (RM) du Saguenay a été calculée pour quatre sous-périodes. Fondé sur une statistique des patronymes (ou noms de famille) apparaissant dans les 90 000 actes de mariage saguenayens, cet indicateur exprime le degré de similarité entre des unités spatiales — ici les RM¹². Concrètement, le calcul consiste à créer des couples théoriques avec des individus de deux RM distincts et, sur la base des patronymes, de vérifier l'existence d'un ancêtre commun. La distance mesurée équivaut donc en réalité à un coefficient de consanguinité. L'indice suppose que deux personnes portant le même patronyme ont un ancêtre commun, plus ou moins éloigné, ce qui a pu être vérifié pour les patronymes les plus fréquents au Saguenay (Jetté, 1987). Les mesures sont toutefois légèrement affectées par les changements de nom, phénomène commun à toutes les populations (naissances illégitimes, adoptions, etc.).

Dans l'ensemble, les trois RM du Bas-Saguenay, à l'est, de même que le RM Lac-Bouchette, au sud du lac Saint-Jean, se démarquent encore une fois (figure 2). Apparaissent également comme marginaux le RM Saint-Honoré — Falardeau et le RM Notre-Dame-du-Rosaire, situés un peu à l'ouest du précédent. Ces aires géographiques partagent les mêmes caractéristiques. Elles sont faites de paroisses agro-forestières, établies en marge des grosses agglomérations et des principales voies de communication.

La probabilité d'origine des gènes

Dans les limites chronologiques du peuplement saguenayen (en ne remontant donc pas au-delà de 1840), les ascendances des 180 000 personnes mariées au Saguenay ont été construites, ce qui a permis d'identifier leurs ancêtres fondateurs. Pour chacun de ces conjoints, il s'agissait de calculer la probabilité qu'il ait reçu des gènes de l'un ou l'autre de ses fondateurs¹³. L'origine extra-régionale de ces derniers nous était donnée par le fichier-réseau dans une proportion variant entre 15% et 20%. Après analyse, il ne semble pas que ce prélèvement forcé soit affecté de biais pouvant compromettre sa représentativité. Il a donc été possible de cartographier les 180 000 conjoints selon leur micro-région de résidence et selon la probabilité qu'ils aient reçu des gènes provenant d'ancêtres de Charlevoix — on se rappellera que la plupart des immigrants ayant introduit des gènes mutants récessifs au Saguenay étaient originaires de cette région (tableau 1).

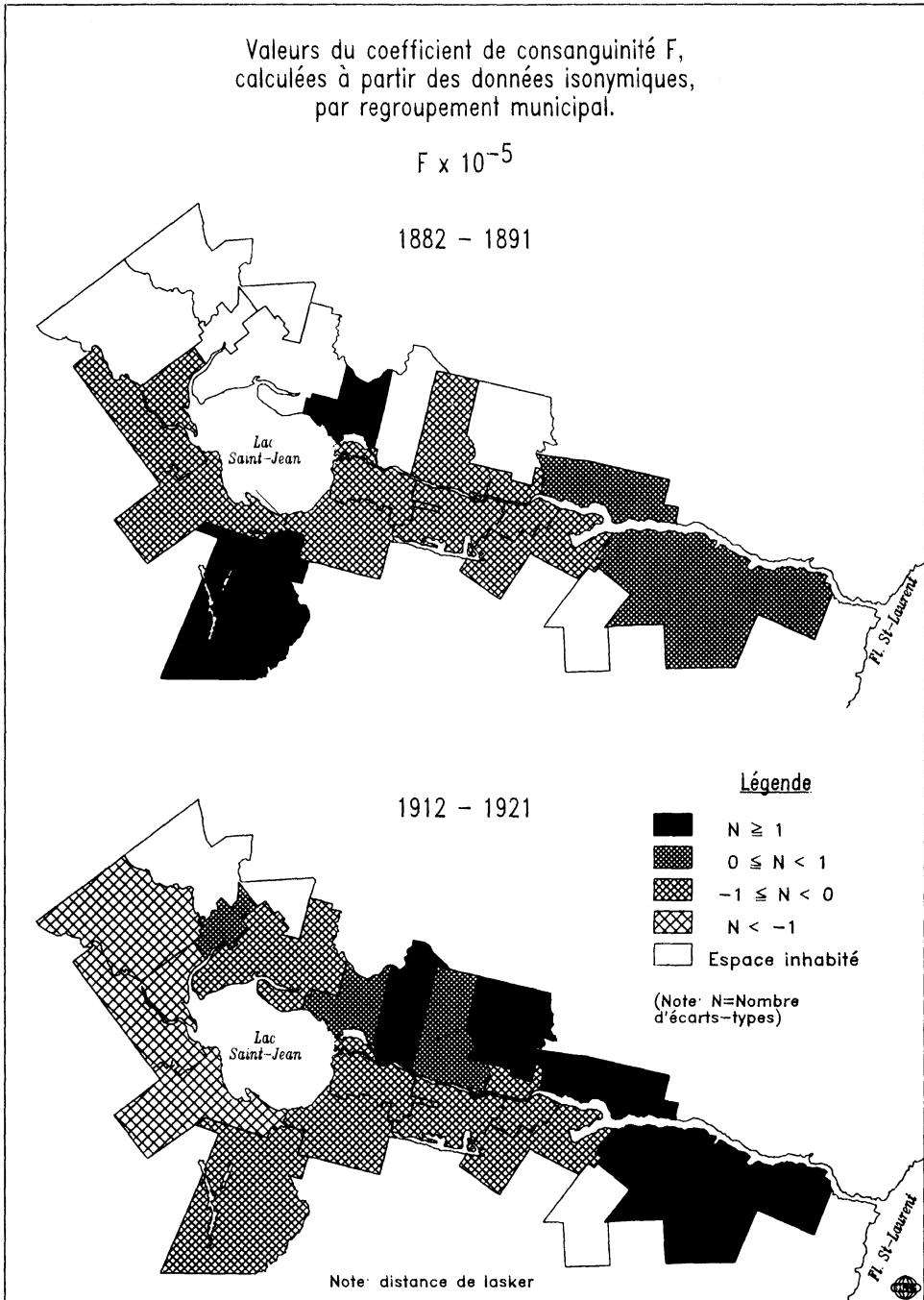
Ce troisième indicateur fait ressortir d'une manière très claire un gradient est-ouest, déjà perceptible dans les figures précédentes (figure 3). Le facteur d'uniformisation génétique représenté par l'immigration de Charlevoix serait surtout concentré dans l'est, les RM du Bas-Saguenay se démarquant encore une fois. Cette conclusion est conforme à ce que l'on sait de la dynamique migratoire saguenayenne : la majorité des pionniers venaient d'abord de Charlevoix et ils ont commencé les défrichements dans l'Est du Saguenay. Progressivement, les points de chute se sont déplacés vers le Lac-Saint-Jean, tandis que la part de Charlevoix parmi les régions de provenance diminuait rapidement (Gauvreau, 1987).

Les coefficients de consanguinité et de parenté

Pour chacune des 66 URB du Saguenay, un échantillon a été construit comprenant les individus mariés entre 1956 et 1971 et résidant dans l'URB. La taille de ces échantillons varie entre 24 et 100 individus en fonction de la taille de l'URB, la moyenne se situant à 76. L'ascendance de ces 4 998 conjoints a été construite par ordinateur, puis les coefficients de consanguinité (F) et de parenté (Φ) ont été calculés pour chaque URB (figure 4). Ces deux coefficients font ressortir respectivement dans quelle mesure : a) on se mariait entre apparentés et, b) les habitants de la même URB descendent des mêmes ancêtres. Ils offrent donc — surtout le Φ — une sorte d'image-synthèse des comportements démographiques qui ont forgé la population de l'URB et balisé la transmission génétique. Or ils prennent ici des valeurs généralement convergentes, faisant tous deux ressortir le statut particulier du Bas-Saguenay et du sud du Lac-Saint-Jean (RM Lac-Bouchette). Le Φ fait en outre apparaître quelques îlots secondaires, en particulier autour du lac Saint-Jean.

Dans l'ensemble, l'étude de ces quelques distributions (figures 1 à 4) suggère donc une convergence presque parfaite autour d'une structure spatiale caractérisée, sous l'angle du cloisonnement, par l'existence d'un gradient (décroissant) est-ouest et de deux zones marginales. D'un point de vue génétique, on s'attend à ce que, d'une façon générale, les unités spatiales les plus refermées sur elles-mêmes, et par conséquent les moins ouvertes aux mouvements des personnes et des gènes, soient celles où les mêmes gènes aient eu le plus de chance de se reproduire et se concentrer localement. S'il arrivait qu'un gène délétère s'introduise dans ces unités spatiales, le risque de génopathies dans les sous-populations concernées pourrait être significativement

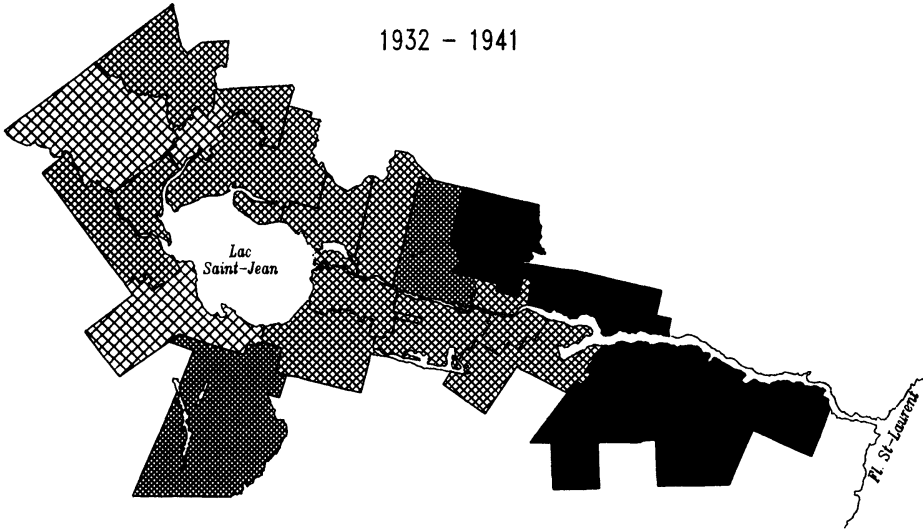
Figure 2



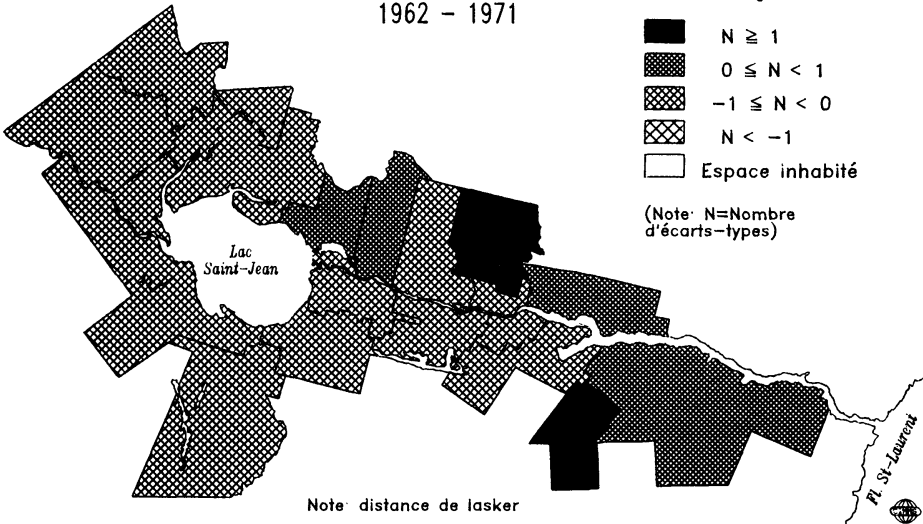
Valeurs du coefficient de consanguinité F ,
calculées à partir des données isonymiques,
par regroupement municipal.

$F \times 10^{-5}$

1932 - 1941



1962 - 1971

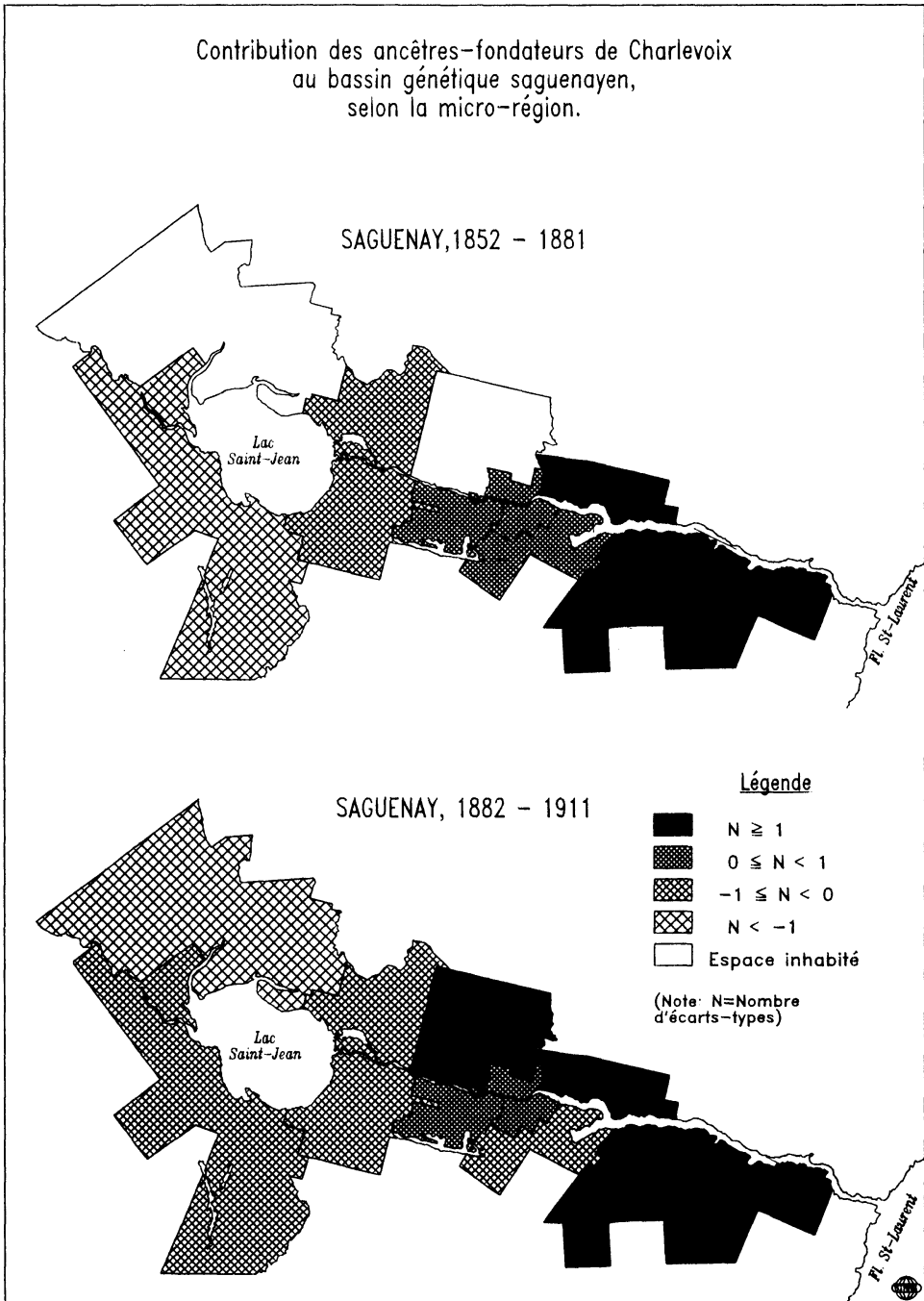


Légende

- $N \geq 1$
- $0 \leq N < 1$
- $-1 \leq N < 0$
- $N < -1$
- Espace inhabité

(Note: N=Nombre d'écart-types)

Note: distance de Hasker

Figure 3

Contribution des ancêtres-fondateurs de Charlevoix
au bassin génétique saguenayen.
selon la micro-région.

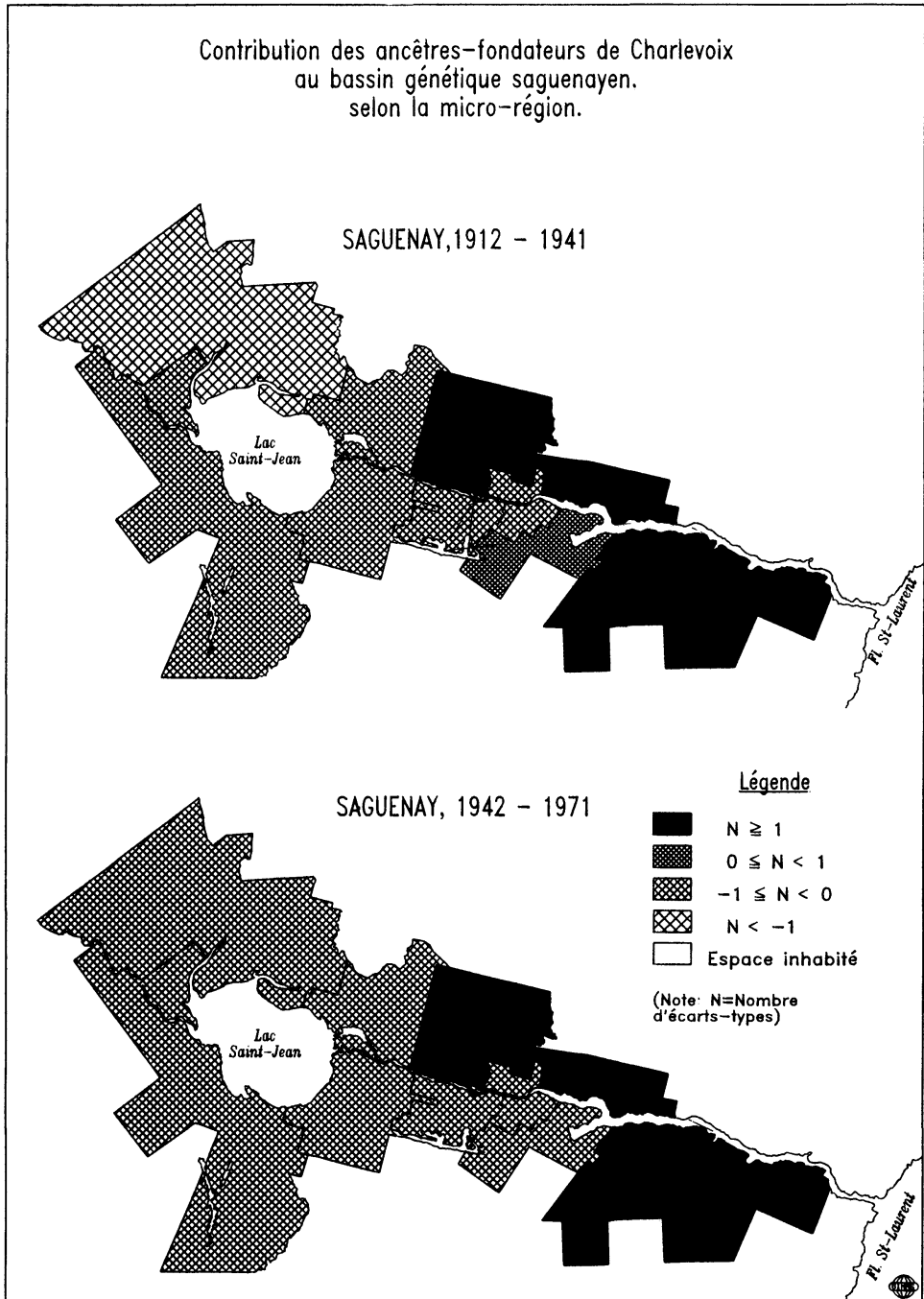
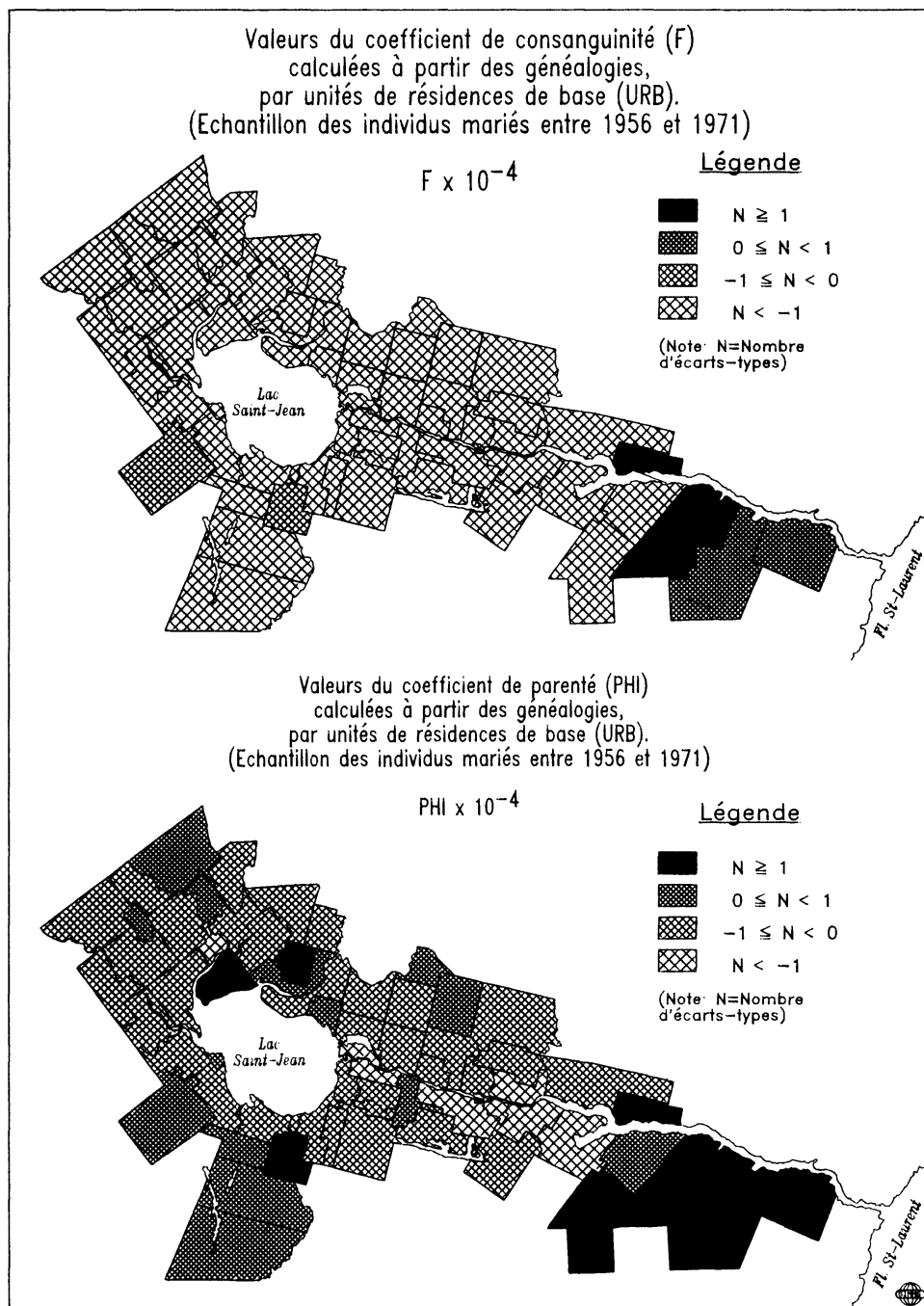


Figure 4



plus élevé que dans l'ensemble de la population saguenayenne, d'où la pertinence d'actions préventives prioritaires.

Les données disponibles sur les principales maladies récessives dans la région offrent la possibilité de tester cette hypothèse, lourde de conséquences, sous l'angle de l'épidémiologie génétique. La question peut être posée en ces termes : la cartographie des porteurs connus de gènes mutants se superpose-t-elle à la cartographie issue des indicateurs d'inférence ?

DISTRIBUTION SPATIALE DE GÈNES MUTANTS

Les six maladies récessives présentées antérieurement sont prises à témoin. Afin de détecter toute forme possible de stratification, elles ont été soumises à trois types de cartographie : premièrement, la répartition des parents des malades ; deuxièmement la répartition, par période quinquennale, des grands-parents des malades et de leurs ancêtres présumés porteurs et, troisièmement, une analyse comparée, sur des cas et des témoins, de la distance entre le lieu de naissance des conjoints au sein de chaque couple.

Les parents des malades (1960-1985)

À l'échelle des 8 micro-régions du Saguenay, la distribution géographique des 478 pères et mères de malades a été étudiée à partir de la résidence qu'ils ont déclarée à leur mariage. Il est à noter que, puisqu'il s'agit ici de porteurs et non des malades, on aurait pu tout aussi bien retenir la résidence à un autre moment de leur existence — leur naissance, par exemple. Pour chaque unité spatiale, le tableau 2 présente un taux de porteurs, lequel rapporte le nombre de parents à la population de la micro-région. Compte tenu de la grande sensibilité du taux utilisé, on espérerait des variations beaucoup plus fortes. En fait, on observe certains accidents très ponctuels, comme pour l'ataxie spastique ou la tyrosinémie dans le Bas-Saguenay, mais dans l'ensemble, les écarts observés sont peu ou non significatifs. Un test statistique (mélanges de distributions) adapté au traitement des petits nombres a été appliqué afin d'établir si l'une ou l'autre des micro-régions constituait un ensemble ou une sous-population distincte des autres. Il s'agit en fait d'un test de vraisemblance permettant de vérifier si on est en présence d'une seule ou de deux distributions distinctes (cf. Symons *et al*, 1983). Il appert que dans l'ensemble, aucune micro-région ne se démarque d'une manière vraiment significative. Mentionnons toutefois, aux deux extrémités de la région, le statut particulier de la MR Dolbeau — Mistassini, à l'ouest, où les prévalences sont statistiquement basses, et la MR du Bas-Saguenay à l'est, où certains résultats sont élevés — mais le poids de l'ataxie spastique est dans ce dernier cas déterminant puisque 3 des 6 maladies n'y sont même pas représentées.

Les grands-parents des malades

Comme il s'agit ici de maladies récessives, il s'ensuit que dans chaque couple de grands-parents de malades, au moins un conjoint était nécessairement porteur d'un gène mutant. Ces 478 couples ont donc pu être traités comme s'ils étaient des individus porteurs (hétérozygotes). Leur distribution a été étudiée non seulement à partir de leur

Tableau 2

**Taux de porteurs de gènes récessifs par micro-région : parents des malades
(1 pour 10 000)**

Micro-régions	Population totale en 1985	Rachitisme vita- mino-dépendant	Agénésie du corps calleux	Fibrose kystique	Sous-fichiers		Ataxie spastique Charlevoix - Saguenay	Ensemble
					Hémochromatose	Tyrosinémie		
Bas-Saguenay	7 300	0,0 (0)	6,8 (5)	0,0 (0)	0,0 (0)	8,2 (6)	19,2 (14)	34,2 (25)
La Baie	25 080	1,2 (3)	1,2 (3)	2,0 (5)	0,0 (0)	4,8 (12)	2,8 (7)	12,0 (30)
Chicoutimi	63 740	1,4 (9)	2,5 (16)	1,9 (12)	1,6 (10)	5,5 (35)	3,8 (24)	16,6 (106)
Jonquière	61 120	0,5 (3)	3,9 (24)	1,6 (10)	1,8 (11)	2,9 (18)	1,3 (8)	12,1 (74)
Saint-Ambroise	13 620	2,9 (4)	4,4 (6)	2,9 (4)	0,7 (1)	7,3 (10)	4,4 (6)	22,8 (31)
Alma	52 590	2,3 (12)	6,3 (33)	3,6 (19)	1,3 (7)	3,4 (18)	5,1 (27)	22,1 (116)
Roberval	32 800	0,3 (1)	1,8 (6)	3,0 (10)	0,3 (1)	4,3 (14)	5,5 (18)	15,2 (50)
Dolbeau-Mistassini	31 571	2,2 (7)	2,5 (8)	5,1 (16)	1,3 (4)	1,3 (4)	2,2 (7)	14,6 (46)
Région	287 821	1,4 (39)	3,5 (101)	2,6 (76)	1,2 (34)	4,1 (117)	3,9 (111)	16,6 (478)

Note : les chiffres absolus apparaissent entre parenthèses.

résidence au début de leur vie conjugale mais jusqu'à la fin de leur union. Cela a été possible en dressant une cartographie sur une base quinquennale, car le fichier-réseau permet en effet de suivre à une échelle très fine l'histoire migratoire des couples, tant qu'ils ne quittent pas la région. Il n'est pas possible de reproduire ici, pour chacune des six maladies, les nombreuses distributions ainsi reconstituées au cours de la période s'étendant de 1900 à 1960. Disons simplement que les écarts observés — sauf encore une fois pour l'ataxie dans le Bas-Saguenay — ne font ressortir aucun modèle particulier. Cette conclusion a été validée par une comparaison avec deux groupes-témoins qui ont été soumis aux mêmes exercices cartographiques.

Les présumés porteurs

Comme les grands-parents, les 509 ancêtres fondateurs présumés porteurs ont donné lieu à des cartographies quinquennales pour la période comprise entre 1850 et 1905, et ils ont été soumis au même type d'analyse. Un phénomène très net doit être relevé ici : dans les premières décennies du peuplement au XIX^e siècle, on observe une surreprésentation marquée des présumés porteurs dans les paroisses du Bas-Saguenay et de la baie des Ha! Ha! (à l'est de Chicoutimi). Nous savons (Bouchard et Larouche, 1987a) qu'au XIX^e siècle, l'Est du Saguenay accueillait surtout des pionniers de Charlevoix, région où se concentraient quatre des gènes mutants étudiés ici (jusqu'à preuve du contraire, il n'y a pas de raisons de penser que la fibrose kystique et l'hémochromatose y aient été plus fréquents qu'ailleurs). Par contre, les immigrants en provenance d'autres régions s'établissaient davantage vers l'ouest, d'où les distributions observées parmi nos présumés porteurs. Toutefois, cette surreprésentation à l'est disparaît pratiquement avec la fin du siècle, sous l'effet des déplacements migratoires intra-régionaux vers le Lac-Saint-Jean.

Les distances géographiques

Enfin, recourant à un indicateur utilisé avec profit ailleurs (cf. Norio, 1981 ; Bois *et al.*, 1976), nous avons calculé la distance moyenne entre les lieux de naissance des pères et mères des malades, tout en effectuant parallèlement un contrôle à l'aide de groupes-témoins. Il faut supposer ici que, sous le rapport de la mobilité géographique, les hétérozygotes ne sont pas désavantagés et qu'ils se comportent comme les non-porteurs — postulat que rien ne permet d'invalidier. Dès lors, si les porteurs étaient concentrés localement, on devrait s'attendre à ce que les distances moyennes dans les groupes-cibles soient substantiellement inférieures à celles relevées dans les groupes-témoins. Ce n'est pas le cas. En général, les écarts observés vont dans le sens attendu mais ils sont très faibles. Ainsi, les distances moyennes dans chaque groupe-cible et ses 3 groupes-témoins sont respectivement de 20 et 22 km pour l'ataxie, de 18 et 25 km pour la tyrosinémie, de 18 et 17,7 km pour l'agénésie. Appliqués aux distances entre les lieux de résidence des conjoints à leur mariage, les mêmes calculs donnent respectivement : 8 et 12 km, 11 et 13 km, 8 et 11 km (cf. Bouchard *et al.*, 1987b).

Ces écarts ne sont pas révélateurs de cloisonnements spatiaux très accusés et ils s'avèrent à première vue peu utiles dans la perspective de l'épidémiologie génétique.

CONCLUSION

C'est sans doute la conclusion à tirer de l'ensemble de nos données saguenayennes : la référence spatiale paraît constituer ici un indicateur de risque trop peu discriminant pour appuyer efficacement des activités de prévention des maladies récessives. On peut s'en étonner, dans la mesure où la cartographie appuyée sur les indicateurs d'inférence laissait présager une certaine stratification spatiale des gènes mutants. Or nous avons vu que la cartographie des porteurs (dûment attestés ou présumés) recoupait imparfaitement celle des indicateurs d'inférence, ce qu'il faut maintenant expliquer. Certes, sur un plan strictement méthodologique, il faut signaler quelques lacunes. Ainsi, notre cartographie repose sur des données de prévalence alors que des données d'incidence (difficiles à produire) auraient en principe fourni des mesures un peu plus fines. La façon dont le dépistage et le recensement des malades ont été réalisés depuis les années soixante est également importante ; le même modèle n'a pas prévalu d'une maladie à l'autre, ce qui est une source potentielle de biais — difficilement remédiable toutefois. Par ailleurs, ce dernier facteur tend sans doute à amplifier la stratification des gènes mutants et non à l'atténuer ; il confirme d'une certaine manière nos résultats.

Il faut donc rechercher dans la dynamique de la population elle-même l'explication de la discordance observée. Une hypothèse retient pour l'instant notre attention. Il se pourrait que l'ensemble du bassin génétique actuel présente un certain cloisonnement spatial, comme le suggèrent les indicateurs d'inférence, sans que la diffusion de nos six gènes récessifs n'épouse cette stratification. Ainsi, à cause de la mobilité des familles au sein de la région depuis le milieu du XIX^e siècle, les gènes mutants auraient eu la possibilité de s'introduire et de se reproduire dans presque toutes les unités spatiales. Pour la période 1852-1911, cette hypothèse a été validée par une étude comparée des itinéraires migratoires parents-enfants (Gradie *et al*, 1988) et par une étude de la mobilité géographique intrarégionale (Bouchard, 1987).

Il sera utile toutefois de poursuivre l'étude de la structure spatiale du *pool* génique saguenayen, mais cette fois en utilisant des marqueurs de gènes non mutants et en intégrant des analyses d'autocorrélation spatiale, destinées à mesurer la proximité des unités géographiques semblables¹⁴.

NOTES

¹ Cet article a été rédigé dans le cadre des travaux de SOREP (Centre interuniversitaire de recherches sur les populations). Fruit d'une entente de coopération entre l'Université du Québec à Chicoutimi, l'Université Laval (Québec) et l'Université McGill (Montréal), SOREP est reconnu officiellement comme centre de recherche interinstitutionnel. Le centre réunit une équipe pluridisciplinaire d'informaticiens, de démographes, de médecins, de généticiens, de sociologues, d'anthropologues et d'historiens appartenant à diverses institutions québécoises. Il a pour but de constituer et d'exploiter, pour les régions et sous-régions du Nord-Est du Québec et pour la période allant des débuts du peuplement jusqu'à nos jours, un registre ou fichier de population informatisé, à caractère universel, fondé sur le jumelage automatique de données nominatives. Le registre prend la forme d'un fichier-réseau (ou *data base*), contient des données de nature économique, sociale, culturelle, démographique, génétique, médicale... et donne lieu à des analyses et enquêtes dans chacune de ces directions de recherche. La base de données, dont la construction a débuté en 1972, est présentement exploitée dans le cadre de deux programmes relevant de SOREP, soit le Programme de recherches sur les sociétés régionales et le Programme de recherches en génétique humaine.

² Les recherches qui ont permis la réalisation de cet article ont été financées par le Fonds FCAR (Québec), le Conseil de recherche en sciences humaines (Ottawa), l'Université du Québec

à Chicoutimi (UQAC) et la Fondation de l'UQAC. L'auteur remercie Kevork Kouladjian pour la partie statistique et informatique : Alain Simard, Marc Saint-Hilaire et Michel Guérin qui ont produit les cartes ; M.-Josée Huot, Anne Vigneault et Michèle Jomphe qui ont contribué à la compilation des données épidémiologiques.

³ Selon la définition étendue retenue pour les fins de nos recherches, le Saguenay comprend les trois sous-régions du Lac-Saint-Jean, du Haut-Saguenay et du Bas-Saguenay. Il désigne donc toute la population établie à l'ouest de Tadoussac, le long du Saguenay et autour du lac Saint-Jean (cf. figures dans le texte).

⁴ Voir la série d'articles publiée à l'initiative de L. Beckman entre 1971 et 1983 dans les revues *Hereditas* et *Human Heredity*, sous le titre général : « Population studies in Northern Sweden ».

⁵ Ce fichier informatisé, qui contient tout l'état civil du Saguenay (1840-1986) et une partie de celui de Charlevoix, permet d'accéder instantanément aux biographies familiales et individuelles, de même qu'aux généalogies ascendantes et descendantes. Il permet, en particulier, de reconstituer les itinéraires migratoires dans un cadre intra ou intergénérationnel (Bouchard *et al.*, 1985).

⁶ Rappelons qu'un gène délétère — ou un gène mutant — est un fragment d'ADN (acide désoxyribonucléique) qui, par suite d'une mutation ou modification de ses composantes, ne produit plus l'enzyme nécessaire à telle ou telle fonction métabolique. Il peut être dominant ou récessif. Dans le premier cas, il suffit de recevoir une copie du gène de l'un de ses deux parents pour être atteint de la maladie. Dans le deuxième cas, il faut que les deux parents transmettent une copie du gène pour que l'enfant soit atteint. On dit alors de cet enfant qu'il est un homozygote. L'individu qui ne porte qu'une copie d'un gène récessif est un hétérozygote et il ne souffre pas de la maladie.

⁷ Ouverte au peuplement d'ascendance européenne à partir de 1840, les premiers immigrants qui s'implantèrent dans la région du Saguenay provenaient de Charlevoix et du Bas-Saint-Laurent (Bouchard, 1983).

⁸ Ces trois découpages sont illustrés sur les figures dans le texte. Pour une présentation plus détaillée, voir Lachance *et al.*, 1985.

⁹ Noter que sur cette figure et celles qui suivent, les unités spatiales sont classées sur la base non pas des valeurs absolues mais du nombre d'écart-types par rapport à la moyenne. Cette disposition a l'avantage de rendre toutes les données cartographiques compatibles.

¹⁰ On s'attend à ce que ce genre de cloisonnement soit un facteur de diversification ou de stratification spatiale. La notion de diversité doit être entendue ici au sens suivant : si l'on prélève 2 gènes au hasard chez 2 individus dans une unité géographique donnée, la probabilité que ces 2 gènes soient identiques sera plus élevée que s'ils étaient prélevés chez des individus habitant 2 unités géographiques distinctes. Dans le cas contraire, on parle d'homogénéité.

¹¹ Le coefficient de consanguinité est calculé à l'aide de données isonymiques.

¹² Les unités spatiales sont comparées 2 à 2 ; pour chaque nom de famille observé, on calcule une distance ou similitude fondée sur la comparaison des fréquences relatives de ce patronyme dans les 2 RM. Le score attribué à chaque RM est la moyenne de ses scores ou distances par rapport aux autres RM. La formule utilisée est la suivante :

$$\frac{1}{\sum_{i=1}^1 \frac{n_{ia} n_{ib}}{4 Na Nb}}$$

où : Na = population de l'unité spatiale a

Nb = population de l'unité spatiale b

n_{ia} = nombre de personnes portant le nom i dans l'unité a

n_{ib} = nombre de personnes portant le nom i dans l'unité b

1 = nombre total de noms différents observés

Cette méthode a été mise au point par Crow et Mange (1965). Voir aussi Kouladjian (1984).

¹³ L'analyse s'appuie ici sur un programme informatique de K. Kouladjian, inspiré de la méthode mise au point par Albert Jacquard (1972).

¹⁴ Sur le modèle proposé entre autres par Moran (1948) et Geary (1954).

SOURCES CITÉES

- BARBEAU, A. (1966) The Syndrome of Hereditary Late Onset Ptosis and Dysphagia in French Canada, in Kuhn, E., éd. *Symposium über progressive Muskel-dystrophie*. Berlin, Springer-Verlag, p. 102-109.
- BOIS, E., FEINGOLD, J., FRENAY, P. et BRIARD, M.-L. (1976) Infantile Cystinosis in France: Genetics, Incidence, Geographic Distribution. *Journal of Medical Genetics*, 13: 434-438.
- BOUCHARD, Gérard (1987) Sur la reproduction familiale en milieu rural: systèmes ouverts et systèmes clos. *Recherches sociographiques*, XXVIII (2-3): 229-251.
- _____ (1983) Le peuplement blanc, in Pouyez, C. et al., *Les Saguenayens. Introduction à l'histoire des populations du Saguenay, XVI^e-XX^e siècles*. Québec, Presses de l'Université du Québec, p. 125-180.
- BOUCHARD, G. et LAROUCHE, J. (1987a) Dynamique des populations locales: la formation des paroisses rurales au Saguenay (1840-1911). Texte à paraître dans la *Revue d'histoire de l'Amérique française*.
- BOUCHARD, G., JOMPHE, M. et VIGNEAULT, A. (1987b) *Indicateurs de distances géographiques appliqués à l'étude de porteurs de gènes récessifs et de témoins*. Document de SOREP n° III-C-55, 14p.
- BOUCHARD, G., ROY, R. et CASGRAIN, B. (1985) *Reconstitution automatique des familles. Le système SOREP*. Université du Québec à Chicoutimi, Dossier n° 2, 2 vol., 745p.
- CLIFF, A.D. et ORD, J.K. (1981) *Spatial Processes*. London, Pion.
- CLOW, C.L. et SCRIVER, C.R. (1977) Knowledge About and Attitudes Toward Genetic Screening among High-school Students; the Tay-Sachs Experience. *Pediatrics*, 59: 86-91.
- CROW, J.F. et MANGE, A.P. (1965) Measurement of Inbreeding from the Frequency of Marriages between Persons of the Same Surname. *Eug. Quart.*, 12: 199-205.
- FIX, A.G. (1978) The Role of Kin-structured Migration in Genetic Micro-differentiation. *Annals of Human Genetics*, 41: 329-339.
- GAUVREAU, Danièle (1987) Le peuplement du Saguenay au 19^e siècle: mesure et caractéristiques du mouvement d'immigration jusqu'en 1911. Communication présentée au congrès de la Société historique du Canada, Hamilton (juin).
- GEARY, R.C. (1954) The Continuity Ratio and Statistical Mapping. *The Incorporated Statistician*, 5: 115-145.
- GINTER, E.K., GARKAVTZEVA, R.G. et REVAZOV, A.A. (1980) Population Structure and Hereditary Diseases in Uzbekistan, in Eriksson, A.W. *Population Structure and Genetic Disorders*. New York, Academic Press, p. 327-337.
- GOODMAN, R.M. (1980) Thoughts on Genetic Diseases among the Jewish People, in Eriksson, A.W. *Population Structure and Genetic Disorders*. New York, Academic Press, p. 337-352.
- GRADIE, M.I. et GAUVREAU D. (1987) Migration and Hereditary Disease in the Saguenay Population of Eastern Québec, *IMR (International Migration Review)*, 21: 592-608.
- GRADIE, M.I., JORDE, L.B. et BOUCHARD, G. (1988) The Genetic Structure of the Saguenay, 1852-1911: Evidence from Parents-Offspring and Isonymy Matrices. Texte soumis pour publication.
- HARRISON, G.A. et BOYCE A.J. (1972) Migration, Exchange and the Genetic Structure of Populations in Harrison, G.A. et Boyce, A.J., éd. *The Structure of Human Populations*. Oxford, Clarendon, p. 128-145.
- JACQUARD, Albert (1972) Évolution du patrimoine génétique des Kel Kummer. *Population*, 27 (4-5): 783-800.
- JETTÉ, René (1987) *La formation de la population de Charlevoix. Données préliminaires*. SOREP, document n° II-C-127, 64p.
- JORDE, L.B. (1980) The Genetic Structure of Subdivided Human Populations: a Review, in Mielke, J.H. et Crawford, M.H., éd. *Current Developments in Anthropological Genetics*. New York, Plenum, vol. 1, p. 135-208.
- _____ (1982) The Genetic Structure of the Utah Mormons: Migration Analysis. *Human Biology*: 54 (3), p. 583-597.
- KOULADJIAN, Kevork (1984) *Note méthodologique sur le calcul de la consanguinité par l'étude des mariages isonymes*. SOREP, document n° III-C-36, 6p.
- LABERGE, Claude (1967) La consanguinité des Canadiens français. *Population*, 22 (5): 861-896.
- _____ (1969) Hereditary Tyrosinemia in a French Canadian Isolate. *The American Journal of Human Genetics*, 21 (1): 36-45.
- LACHANCE, M., BOUCHARD, G. et ROY, R. (1985) *Nouvelle table synchronique des équivalences et divisions territoriales de la région du Saguenay*. SOREP, document n° II-C-107, 53p.
- LI, Ching Chun (1976) *First Course in Population Genetics*. California, Pacific Grove, 631p.

- McKUSICK, Victor A. *et al* (1985) *Medical Genetics Studies of the Amish*. Baltimore/London, The Johns Hopkins University Press, 525p.
- MORAN, P.A.P. (1948) The Interpretation of Statistical Maps. *Journal of the Royal Statistical Society*, B, 10: 243-251.
- MORGAN, K., HOLMES, T.M., GRACE, M. *et al* (1983) Patterns of C Cancer in Geographic and Endogamous Subdivisions of the Hutterite Brethren of Canada. *American Journal of Physical Anthropology*, 62: 3-10.
- MOTULSKY, A.G. (1980) Ashkenazi Jewish Gene Pools: Admixture, Drift, and Selection, in Eriksson, A.W. éd. *Population Structure and Genetic Disorders*. New York, Academic Press, p. 353-366.
- MOUZOURAS, M., CAMBA, L., IOANNOU, P., MODELL, B. *et al* (1980) Thalassemia as a Model of Recessive Genetic Disease in the Community. *Lancet*, II (8): 574-578.
- NORIO, Reijo (1981) Diseases of Finland and Scandinavia, in Rothschild, H. éd. *Biocultural Aspects of Disease*. New York, Academic Press, 653p.
- NORIO, R., NEVANLINNA, H.R. et PERHEENTUPA, J. (1973) Hereditary Diseases in Finland; Rare Flora in Rare Soil. *Annals of Clinical Research*, 5: 109-141.
- SOKAL, R.R. et ODEN, N.L. (1978a) Spatial Autocorrelation in Biology. 1. Methodology. *Biol. J. Linn. Soc.*, 10: 199-228.
- _____ (1978b) Spatial Autocorrelation in Biology. 2. Some Biological Implications and Four Applications of Evolutionary and Ecological Interest. *Bio. J. Linn. Soc.*, 10: 229-249.
- SCRIVER, C.R., BARDANIS, M. *et al* (1984) Thalassemia Disease Prevention: Genetic Medicine Applied. *American Journal of Human Genetics*, 36: 1024-1038.
- SYMONS, M.J., GRIMSON, R.C. et YUAN, Y.C., (1983) Clustering of Rare Events. *Biometrics*, 39.
- WAGENER, D., CAVALLI-SFORZA, L.L. et BARAKAT, R. (1978) Ethnic Variation of Genetic Diseases: Roles of Drift for Recessive Lethal Genes. *American Journal of Human Genetics*, 20: 262-270.

(Acceptation définitive en février 1988)