Mitogenomic and microsatellite variation in descendants of the founder population of Newfoundland: high genetic diversity in an historically isolated population

A.M. Pope, S.M. Carr, K.N. Smith, and H.D. Marshall

Abstract: The island of Newfoundland, the first of England's overseas colonies, was settled from the 17th century onward by restricted numbers of English, Irish, and French immigrants, in small "outport" communities that have maintained geographic, religious, and linguistic isolation to the latest generations. To measure the extent of modification and loss of genetic variation through founder effect, drift, and inbreeding in this historically isolated population, we analyzed the complete mitochondrial DNA (mtDNA) genomes and 14 microsatellite loci from each of 27 individuals with matrilineal ancestries extending to the colonial period. Every individual has a unique mtDNA genome sequence. All but one of these genomes are assignable to one of five major (H,J,K,T, and U) or minor (I) European haplogroups. The possibility of homoplasy at single nucleotide polymorphism (SNP) sites that define subtypes within the H haplogroup is discussed. Observed haplogroup proportions do not differ significantly from those of western Europeans or between English and Irish Newfoundlanders. The exceptional individual is a member of haplogroup A2, who appears to be the descendant of a Mi'kmaq First Nations mother and a French father, a common marriage pattern in the early settlement of Newfoundland. Microsatellite diversity is high (H_E = 0.763), unstructured with respect to mtDNA haplotype or ethnicity, and there is no evidence of linkage disequilibrium. There is a small but significant degree of inbreeding (F_{IS} = 0.0174). Collection of whole mtDNA genome data was facilitated by the use of microarray sequencing, and we describe a simple algorithm that is 99.67% efficient for sequence recovery.

Key words: inbreeding, isolated populations, microarray DNA sequencing, microsatellites, Mi'kmaq First Nations, mitogenomics, mtDNA.

Résumé : L'île de Terre-Neuve, la première des colonies anglaises outre-mer, a été colonisée à compter du 17ième siècle par un nombre limité d'immigrants anglais, irlandais et français qui ont vécu au sein de petites communautés portuaires en isolement géographique, religieux et linguistique jusqu'à récemment. Pour mesurer les degrés de modification et de perte de variation génétique dus à l'effet fondateur, à la dérive génétique et à la consanguinité au sein de cette population historiquement isolée, nous avons analysé le génome mitochondrial (ADNmt) complet et 14 locus microsatellites chez 27 individus dont la généalogie maternelle remontait à l'époque des colonies. Chaque individu avait une séquence génomique unique d'ADNmt. Il s'est avéré possible d'assigner tous ces génomes, sauf un, à l'un de cinq haplogroupes européens majeurs (H,J,K,T et U) ou mineur (I). Nous discutons de la possibilité d'observer l'homoplasmie à des sites SNP qui distinguent les différents sous-types au sein de l'haplogroupe H. Les proportions observées pour les différents haplogroupes ne diffèrent pas significativement de celles rapportées chez les Européens occidentaux, ni entre Terre-neuviens d'origine anglaise et irlandaise. L'individu exceptionnel est membre de l'haplogroupe A2 et serait un descendant d'une mère autochtone de la nation Micmac et d'un père Français, un type de mariage répandu au début de la colonisation de Terre-Neuve. La diversité des microsatellites est élevée (H_E = 0,763), sans structure en lien avec l'haplotype d'ADNmt ou l'origine ethnique, et il n'y a aucune indication de déséquilibre de liaison. Il y a un faible, mais significatif, degré de consanguinité (F_{IS} = 0,0174). La collection de données génomiques sur l'ADNmt a été facilitée par le recours au séquençage sur puces à ADN, et nous décrivons un algorithme simple qui est efficace à 99,67 % pour récupérer les séquences.

Mots-clés : consanguinité, populations isolées, séquençage sur puces à ADN, microsatellites, première nation Micmac, mitogénomique, ADNmt.

[Traduit par la Rédaction]

Received 16 March 2010. Accepted 25 September 2010. Published on the NRC Research Press Web site at genome.nrc.ca on 20 January 2011.

Corresponding Editor: B. Golding.

A.M. Pope, S.M. Carr,[1] K.N. Smith, and H.D. Marshall. Department of Biology, Memorial University of Newfoundland, St. John's, NL A1B 3X9, Canada.

[1]Corresponding author (e-mail: scarr@mun.ca).

Introduction

The importance of Newfoundland as a place to study human genetics stems from the history of its early settlement. After European rediscovery in 1497 and until confederation with Canada in 1949, the island of Newfoundland was the oldest of England's overseas colonies (Prowse 1895; Mannion 1977; Rowe 1980). English fisherman initially conducted a nonresident "migratory" seasonal fishery, but overwintering establishments in support of the fishery occurred as early as 1527 at St. John's (Handcock 1989), and permanent settlement was attempted beginning in 1621 (Prowse 1895). By 1675, English Protestants inhabited more than 30 sites between Bonavista and Trepassey along the Avalon Peninsula of eastern Newfoundland (Fig. 1). Almost all of these "outport" communities were founded by fewer than five families, drawn from highly localized regions of southwest England (Mannion 1977). Irish Catholic immigrants are documented by 1675 in St. John's, and by 1753 all major communities on the Avalon Peninsula had Irish majorities (Rowe 1980; O'Neill 2003). The French also maintained small settlements in support of an extensive migratory fishery from 1500 onwards, on the north coast westward from Bonavista to Point Riche on the Northern Peninsula (the "Petit Nord"), the Port au Port Peninsula, and St. George's Bay, and subsequently as far south as Cape Ray (the "French Shore") (Winsor 1887).

Throughout the 17th–20th centuries, intermarriage among these diverse settlements was limited by religious, linguistic, socioeconomic, and geographic barriers. Outport communities have typically remained small and isolated, and their progeny remained near the original settlements, such that current regional populations are relatively homogeneous genetically (Crawford et al. 1995; Young et al. 1999; Martin et al. 2000; Parfrey et al. 2002). A study of outport populations found that only 1%–8% of the population were immigrants to the area, and that 60% of births were to parents originating from the same community (Bear et al. 1987). Estimated inbreeding coefficients (F_{IS}) for two such communities on the Avalon and Northern Peninsulas were 0.0032 and 0.0171, respectively (Bear et al. 1988).

Expected genetic consequences of these historical phenomena for the present-day descendants of the early Newfoundland settlers are those characteristic of small, isolated populations: founder effect, genetic drift, and inbreeding. In consequence, a number of genetic diseases show high prevalence in the island portion of the province with respect to its source populations, including Bardet–Biedel syndrome (Woods et al. 1999; Moore et al. 2005), nonpolyposis colorectal cancer (Woods et al. 2005), and late infantile neuronal ceroid lipofuscinosis (Moore et al. 2008). For the maternally inherited mitochondrial DNA (mtDNA) genome, predictable stochastic consequences include loss or skewed representation of European haplogroups and their subtypes among the small founding groups, subsequent fluctuation of haplogroup frequencies within and among persistently small communities, and ultimately loss of heterogeneity with respect to the source populations.

In this study, we seek to determine whether historical events have significantly altered mtDNA haplogroup proportions in this culturally unique and isolated Canadian population with respect to the European source populations. We make parallel analyses of variation in mtDNA genomes and in a panel of biparentally inherited microsatellite loci for a set of individuals with matrilineal ancestries documented at least as far back as the 19th century colonial period. To maximize resolution of differences among individuals, their assignment to subhaplogroups, and the presence of unusual haplotypes with respect to the extant database, we obtained complete mtDNA genome sequences. To determine any consequences of the genetics of small populations, we also ask whether there is significant partitioning of nuclear variance, either with respect to mtDNA haplogroups or ethnic groups, or among individuals with respect to the total population (inbreeding).

Materials and methods

Choice of individuals

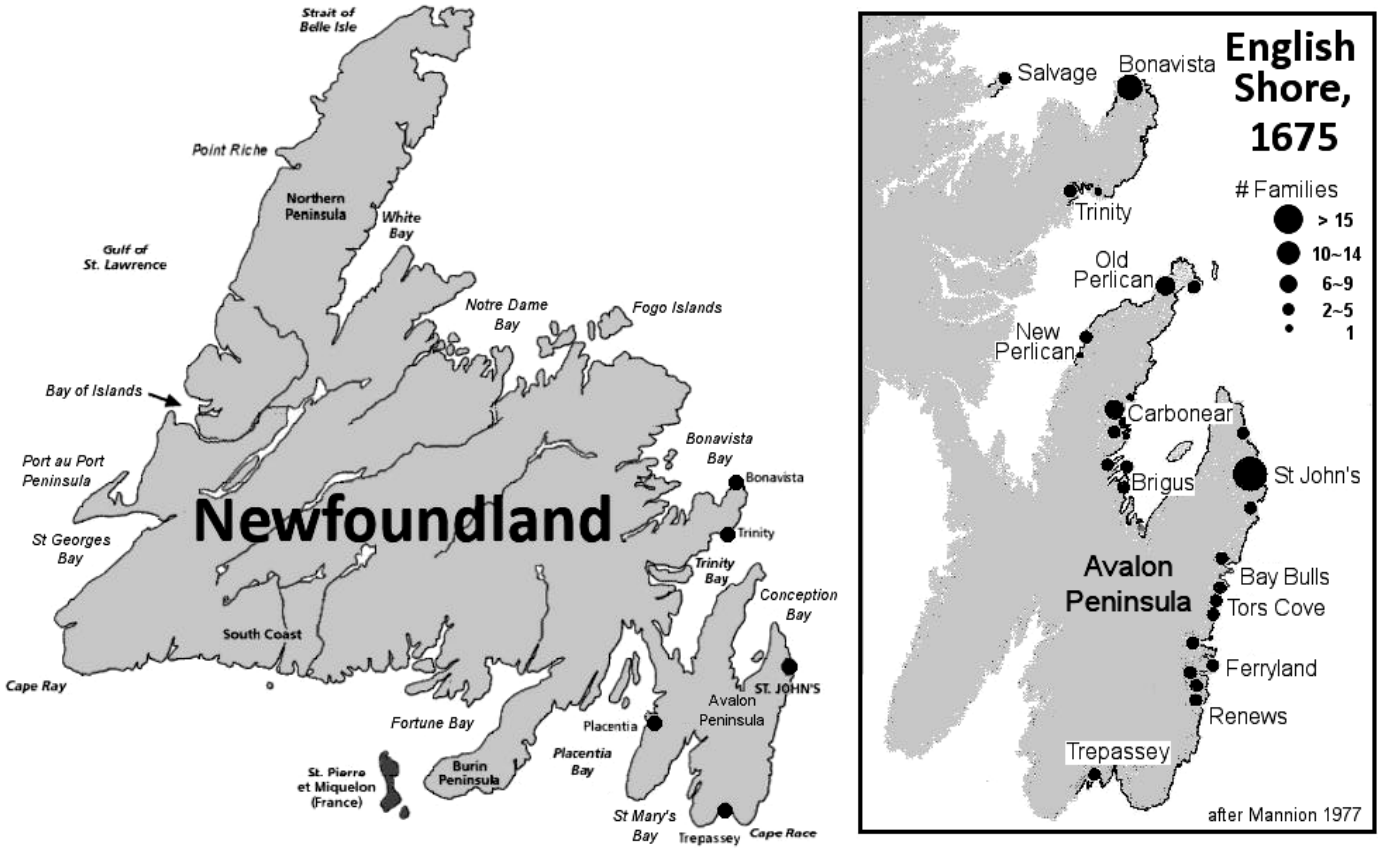
Individuals included in this study were selected from an extant DNA library at the Memorial University Faculty of Medicine, St. John's, Newfoundland. Individuals were screened on the basis of a questionnaire about matrilineal ancestry, with informed consent. Among these, we were able to identify a total of 27 individuals whose genealogical records documented their matrilineal English, Irish, or French ancestries in Newfoundland at least as far back as the 19th century colonial period, including six prior to 1850 (Table 1). Because these individuals and their ancestors are drawn from throughout the island, rather than from any one community, analysis of these matrilines can only identify founder, drift, and inbreeding effects in descendants of the ancestral population as a whole.

Polymerase chain reaction: dideoxy and microarray DNA sequencing

The complete mtDNA genomes of the 27 individuals were enzymatically amplified by the polymerase chain reaction in either 24 (for dideoxy sequencing) or 14 (for microarray sequencing) overlapping segments (Rieder et al. 1998; Pope 2003, 2007). The mtDNA sequences of 19 individuals were determined with an ABI 377 automated DNA sequencer (Applied Biosystems, Inc., Richmond, California). ABI Sequencing Analysis v.3.2 software was used to extract sequence data. Sequences were assembled and edited with the Sequencher 4.9 software (Gene Codes Corporation, Ann Arbor, Michigan). The mtDNA genomes of 10 individuals were determined by a combination of microarray sequencing with a commercially available MitoChip (v.1.1) (Maitra et al. 2004) (Affymetrix, Inc., Santa Clara, California) for 15 452 bp of the mtDNA genome, not including the Control Region (CR), and conventional dideoxy sequencing to complete the remaining sequence. Two individuals (Fr1 and Fr2) were sequenced by both methods (Table 1).

Microarray sequencing procedures followed those of the GeneChip CustomSeq Resequencing Array v.2.0 protocol (Affymetrix, Inc.). Data for each position consisted of four fluorescent intensity values, one for each of the four possible bases (A, C, G, or T), from both DNA strands. For each position for either strand, our algorithm (Flynn and Carr 2007; Carr et al. 2008, 2009; cf. Maitra et al. 2004; Hartmann et al. 2009) assigns the presumptive call as that of the

Fig. 1. The island of Newfoundland, with towns and geographical features mentioned in the text and Table 1. The "French Shore" runs northwestward from Bonavista around the Northern Peninsula to Point Riche, and historically as far south as Cape Ray. St. John's is the first settlement, capital, and largest city, and is part of the English Shore that runs southward from Bonavista along the eastern shore of the Avalon Peninsula to Trepassey. The right-hand panel shows English settlements as of 1675: note that most were founded by five or fewer families (after Mannion 1977).



highest signal, and calculates a confidence score (d*S*/*N*) as the difference between the highest and second-highest signals, divided by the sum of the signals at all four bases. Based on our previous analysis (Flynn and Carr 2007) of Fr2, we used two thresholds, d*S*/*N* ≥ 0.10 and ≥ 0.13, for lower and higher confidence single nucleotide polymorphism (SNP) calls, respectively (see Supplementary data, Table S1).[2]

Phylogenetic analysis

Maximum parsimony (MP) trees were constructed with PAUP*4.0 (Swofford 2002) based on the complete genomes with 100 random sequence additions and tree bisection and reconnection (TBR) branch swapping. Analyses were run either with all sites weighted equally or with known hypervariable sites in the CR effectively weighted as zero by exclusion from the initial parsimony analysis. Confidence in branch nodes was determined by 10 000 bootstrap replications, with two random addition orders each and TBR branch swapping. Chi-square tests were performed by a Monte Carlo method designed for small cell values (Roff and Bentzen 1989).

Microsatellite fragment analysis

A panel of 15 dinucleotide microsatellite loci was chosen from the Marshfield Clinic Screening Set 10 list of loci (Rosenberg et al. 2002). Fluorescently labeled primers were either custom-prepared by Applied Biosystems Inc. or obtained from the ABI PRISM Human Linkage Mapping Set v2.5. PCR amplification was performed in 10–20 μL volumes containing 1× PCR Master mix (Promega Corp., Madison, Wisconsin). For loci *D7S247* and *D11S969*, 1 μmol/L concentrations of each primer were used (Smith 2009). For all other loci, 0.3 μmol/L primer concentrations were used and the final $MgCl_2$ concentration was increased by 1 mmol/L. Thermal cycling was done in a GeneAmp PCR System 9700 thermal cycler (Applied Biosystems Inc.) with the following profile: 95 °C for 2 min; 10 cycles of 94 °C for 15 s, 55 °C for 15 s, and 72 °C for 30 s; 20 cycles of 89 °C for 15 s, 55 °C for 15 s, and 72 °C for 30 s; 72 °C for 10 min. For loci *D7S247* and *D11S969*, an alternative profile was used: 95 °C for 4 min; 35 cycles of 95 °C for 30 s, 50 °C for 30 s, and 72 °C for 1 min; 72 °C for 7 min. Fragment analysis was performed with the Applied Biosystems Inc. 3730 DNA analyzer. The size of each allele at

[2] Supplementary data for this article are available on the journal Web site (genome.nrc.ca).

Table 1. Samples codes (ethnicity: En, English; Fr, French; Ir, Irish), case numbers, matrilineal lineage history (origin and birth date of oldest documented matrilineal ancestor), GenBank accession numbers, method of sequencing (MC, MitoChip; dd, Sanger dideoxy), and assigned haplogroup (see Results).

					Method		
Code	Case No.	Matrilineal origin	Birth date	GenBank accession	MC	dd	Haplogroup
En1	157	Notre Dame Bay	1835	HQ287893		X	H1a
En2	1017	Fortune Bay	~1846	HQ287875		X	T2a
En3	1208	Bonavista Peninsula	1892	HQ287896	X		H
En4	1402	Fogo Islands District	1861	HQ287886		X	H1c
En5	1524	Fortune Bay	~1861	HQ287894		X	H5
En6	10354	Conception Bay	~1874	HQ287885	X		H1c
En7	10656	Bonavista Peninsula	~1830s	HQ287883		X	H16
En8	10670	Notre Dame Bay	1877	HQ287882	X		I2
En9	11269	Placentia Bay	1875	HQ287882	X		H16
En10	11469	Bonavista Bay	1837	HQ287873		X	J1c
En11	11528	Fortune Bay	~1872	HQ287876	X		T2a
En12	11645	Trinity Bay	1877	HQ287891		X	H1a
En13	11785	Fogo Islands District	1874	HQ287874		X	J1c
En14	802	Fortune Bay	1875	HQ287889		X	H3
En15	12765	Conception Bay	1873	HQ287892		X	H1a
En16	13016	Fortune Bay	~1835	HQ287888	X		H16
En17	12218	Port au Port Peninsula	1880	HQ287887		X	H1c
Ir1	1351	Conception Bay	1896	HQ287898		X	H
Ir2	10744	Conception Bay	1885	HQ287895		X	H5a
Ir3	10796	Fogo Islands District	1897	HQ287890	X		H3
Ir4	10799	Avalon Peninsula	~1850s	HQ287897		X	H
Ir5	11727	Placentia Bay	~1869	HQ287879		X	U5b
Ir6	11983	Avalon Peninsula	1882	HQ287878		X	U5b
Ir7	12127	St. Mary's Bay	1852	HQ287877	X		T2
Ir8	13136	Avalon Peninsula	1883	HQ287881		X	K1a
Fr1	12204	St. George's Bay	~1840s	HQ287880	X	X	U6
Fr2	13392	Port au Port Peninsula	1862	HQ287872	X	X	A2

Note: Where a matrilineal lineage is known to have existed in Newfoundland one or two generations prior to the most recent maternal ancestor with a documented birth date, the age of lineages marked ~ are estimates made by subtracting 20 years (one generation) from the oldest documented birth date, and those of the form ~1830s by subtracting 40 years (two generations) and rounding to the interval of the closest decade.

each locus was determined with Peak Scanner (Applied Biosystems Inc.).

Microsatellite description statistics and population structure analysis

For each locus, we calculated the number of alleles, expected and observed heterozygosity (H_E and H_O), and F_{IS} with Genepop (Raymond and Rousset 1995). Linkage disequilibrium among all pairs of loci and a global estimate of F_{IS} and conformance to Hardy–Weinberg expectations were also evaluated with Genepop (http://kimura.univ-montp2.fr/~rousset/Genepop.htm). F_{ST} and R_{ST} among subgroups of individuals partitioned either by English or Irish ethnicity or inclusive mtDNA clade (H + I, J + T, and K + U) were calculated with Arlequin (http://cmpg.unibe.ch/software/arlequin3/) (Excoffier et al. 2005). An independent estimate of the number of distinct genetic clusters (*K*, evaluated from 1 to 27) within the sample was made with Structure (http://pritch.bsd.uchicago.edu/structure.html) (Pritchard et al. 2000), with the admixture model, a burn-in period of 10 000, and 100 000 Markov chain Monte Carlo (MCMC) repetitions.

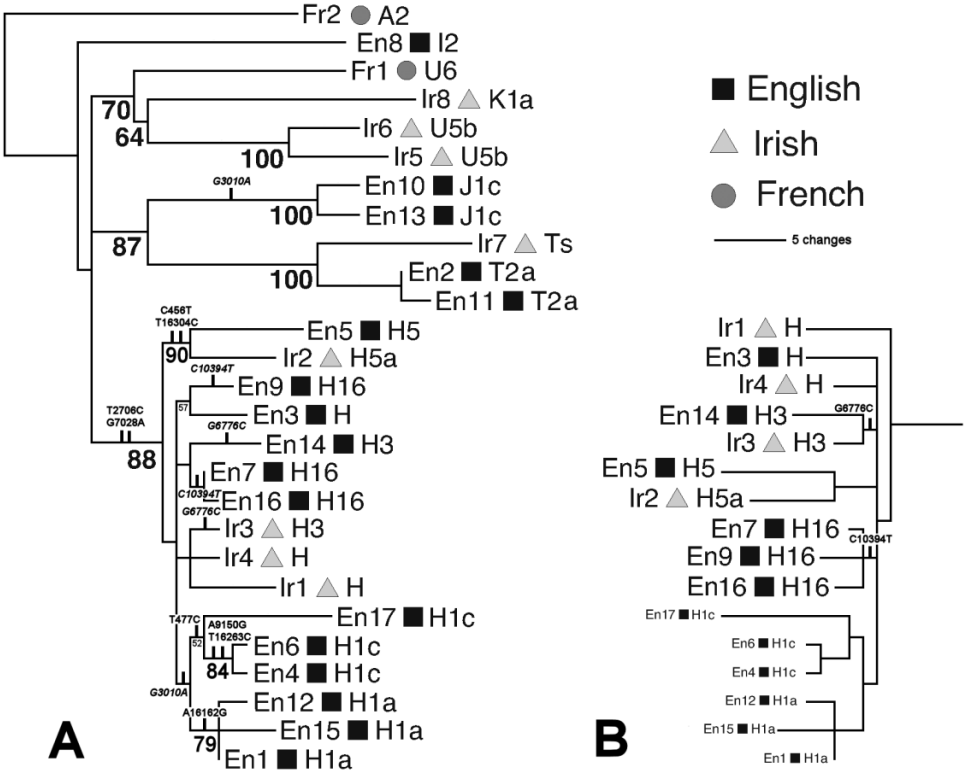
Results

DNA sequence differentiation, haplogroup assignments, and ethnicity

All sequences were submitted to GenBank and assigned the accession numbers HQ287872–HQ287898. A total of 197 SNP sites were identified, of which 82 were phylogenetically informative sensu Nei (1987) (see Supplementary data, Table S2). The mean pairwise difference among genomes is 24.9 ± 12.1, with a range of 1–56 differences. Genomes here differ by up to 15 substitutions from the next most similar in GenBank (accessed August 2010). Five pairs of individuals have identical CR sequences.

Haplogroup assignments were initially made (Pope 2003, 2007) according to the presence or absence of canonical restriction sites and characteristic SNP signatures (Macaulay et al. 1999; Richards et al. 1998, 2000; Maca-Meyer et al. 2001; Brandstätter et al. 2006) as subsequently confirmed with the Mitomaster tool (Brandon et al. 2009) and the PhyloTree database (van Oven and Kayser 2009). A majority of Newfoundlanders examined are assignable to the most abundant haplogroup, H, which has been subdivided into more

Fig. 2. (A) Equally weighted maximum parsimony (MP) analysis of the phylogenetic relationships among 27 Newfoundlanders based on complete mitochondrial genomes. The tree shown is that with minimum length (231 events, consistency index = 0.84). Numbers below branches are bootstrap support (10 000 replicates) for the node to the right. Assignments to haplogroups are based on canonical single nucleotide polymorphism (SNP) sites (van Oven et al. 2009; Brandon et al. 2009). Labels show patterns of inferred changes at 11 canonical sites that define haplogroup H (A2706G and C7028T) and subhaplogroups H1a (A16162G), H1c (T477C), H1c1 (A9150G and T16263C), H3 (T6776), H5 (C456T and T16311C), and H16 (C10394T). Sites labeled in italics are subject to homoplasy: note that H3 and H16 are not monophyletic. Subgroup H5 is also supported by T16513C, but this site is characterized in PhyloTree as "recurrent" and thus not reliable. (For a complete list of informative sites with diagnoses of homoplasy, see Table S2 of Supplementary data). (B) As above, with five hypervariable Control Region (CR) sites (16129, 16183, 16189, 16311, and 16519) excluded from the parsimony analysis. The single minimum-length tree with the topology shown requires 216 events at the remaining sites: with changes at the five excluded sites restored in the most parsimonious manner, the tree requires an additional 18 changes for a total of 234 events (consistency index = 0.83). Topology and branch lengths are identical to those in A, except within the H haplogroup as indicated; patterns within H1a and H1c are unchanged. Note that the H3 and H16 subhaplogroups are monophyletic as defined by T6776 and C10394T, respectively.



than two dozen subgroups on the basis of one or a few SNP markers each (Mishmar et al. 2003; Loogväli et al. 2004; Torroni et al. 2006; Roostalu et al. 2007). Of these, H1 (including three subtypes), H3, H5, and H16 are identifiable in Anglo–Irish Newfoundlanders by nine canonical SNPs (see Fig. 2). Another three individuals are not unambiguously assignable to any previously defined H haplogroup. Non-H individuals were also assigned to subhaplogroups based on Mitomaster and PhyloTree.

Proportions of haplogroups among Newfoundlanders collectively (Table 2) do not differ significantly from those expected for western Europeans (Töpf et al. 2006), either with (χ^2 = 6.75, $p > 0.67$) or without (χ^2 = 5.77, $p > 0.64$) haplotype A2. Proportions do not differ significantly between the two larger ethnic samples, English and Irish (χ^2 = 8.15, $p > 0.10$), although a slight shift in the count of individuals with this haplotype would produce statistical significance. Similar nonsignificant differences were obtained upon comparison with separate estimates of "British" (English) and Irish haplogroup frequencies in Moreau et al. (2009) (results not shown).

Phylogenetic analysis

Initial analysis with genomes in haplogroup M identified all of the present individuals as members of haplogroup N (Carr et al. 2008), and individual Fr2 as the outgroup to the remaining Newfoundland individuals within haplogroup N. The MP analysis of these 27 complete genomes identifies a single minimum-length tree (Fig. 2A). The tree requires 231 events; 13 nodes are supported by more than half of bootstrap replications, including six by 90% or more. These include clades corresponding to the established H, J, K, T, and U haplogroups, and the pairing of J + T (van Oven et al. 2009; Brandon et al. 2009). The phylogenetic position of the individual En8 assigned to haplogroup I outside the remaining clades is not confirmed by a majority of bootstrap replications.

The MP tree shows extensive homoplasy, especially with respect to the H group. The two genomes assigned to H3 by T6776C are not each other's closest relatives, owing to SNP variation in the CR at positions 16 129 and 16 189 that associates each of them with various other H haplogroup sequen-

Table 2. Observed counts and proportions of haplogroups among Newfoundland ethnic groups: observed frequencies (Töpf et al. 2006) and expected counts for western Europeans.

Haplogroup	English	Irish	French	Total	Newfoundland (%)	European (expected)	European (%)
H	12	4	0	16	59	12	47
I	1	0	0	1	4	0	0
J	2	0	0	2	7	4	17
K	0	1	0	1	4	2	6
T	2	1	0	3	11	2	9
U	0	2	1	3	11	3	11
V	0	0	0	0	0	1	5
X	0	0	0	0	0	2	6
A	0	0	1	1	4	0	0
Total	17	8	2	27		26	

ces. The pairwise sequence difference between the two H3 haplotypes (8) is slightly greater than that between them and two H16 haplotypes (5–6). Likewise, the assignment of three genomes to H16 by C10394T contrasts with the patterns implied by SNPs in the CR at 16 189 and 16 311. There are also 12 SNP sites that indicate homoplasic change between one or more individual sequences in the H group with non-H haplogroups (see Supplementary data, Table S3).

Positions 16 129, 16 189, and 16 311, as well as 16 183 and 16 519, are known to be subject to recurrent mutations such that that their phylogenetic implications are considered less reliable than those inferred from other, non-CR SNP sites (cf. van Oven et al. 2009; Brandon et al. 2009). Accordingly, we performed a second parsimony analysis with these five sites excluded, saved the tree topology, and superimposed the changes at the excluded sites on this topology in the most parsimonious manner. The resulting tree (Fig. 2B) requires 18 changes at the five CR sites (a net increase of three for an overall length of 234) and alters relationships among H-haplogroup sequences such that the two canonical SNPs T6776C and C10394T are synapomorphies that define H3 and H16, respectively.

Microsatellite structure

The mean number of alleles per locus is 7.9 (range 6–10) (Table 3). The mean H_E was 0.763 (range 0.600–0.864) as compared with a mean H_O of 0.748 (range 0.478–0.926). Although only one locus (*D20S851*) showed a significant F_{IS} ($p < 0.05$), over all loci there was a slight but significant departure from Hardy–Weinberg expectations, indicative of inbreeding ($F_{IS} = 0.0174$, $p = 0.033$). Among 91 pairwise comparisons, only two pairs of loci showed significant linkage disequilibrium ($p < 0.05$), no more than would be expected by chance. Measures of population differentiation were not significantly different from zero for groups defined either by English versus Irish ethnicity ($F_{ST} = 0.00525$, $p = 0.243$; $R_{ST} = –0.0107$; $p = 0.637$) or H versus non-H mitochondrial haplogroup ($F_{ST} = 0.0171$, $p = 0.973$; $R_{ST} = –0.0148$, $p = 0.762$). The Structure analysis gave $K = 1$ as the most likely number of genetic clusters in the sample. That is, there is no allocation of individuals to subsamples that would minimize either departure from Hardy–Weinberg expectations or linkage disequilibrium with respect to the sample as a whole.

Discussion

European haplogroup variation in Newfoundland

The European population includes seven mtDNA haplogroups (H, J, K, T, U, V, and X) that occur at frequencies >5%, plus several minor haplogroups (including I), all within the N supergroup (van Oven et al. 2009; Brandon et al. 2009). Of these, six (H, J, K, T, U, and I) were identified among 27 Newfoundlanders. A majority of Newfoundlanders examined are assignable to one or another subtype of the most abundant haplogroup H (Loogväli et al. 2004; Brandstätter et al. 2006; Roostalu et al. 2007). The possibility that some of these H subtypes may represent multifurcation nodes is well recognized (Torroni et al. 2006). Attempts to classify cladistic haplogroups by phenetic signatures based on single SNPs may be hindered by the occurrence of mutational hot spots or more rapidly evolving nucleotide sites (Richards et al. 2000; Loogväli et al. 2004), as occurs also in other nonhuman species (Carr et al. 2008; Carr and Marshall 2008*a*, 2008*b*).

In the present data for example, the G3010A SNP that defines H1 shows homoplasy with respect to haplogroup J. Homoplasy is also evident at 12 other sites between sequences in the H and non-H haplogroups (see Supplementary data, Table S3). Within H, three individuals are not definitely assignable to any previously defined subgroup, although each has at least one SNP associating it with one or another H subtype; their basal placement in Fig. 2B suggests these are not novel types originating within the Newfoundland population. The hierarchal arrangement of the PhyloTree classification tacitly assumes that the H-group signature SNPs define monophyletic lineages and are not subject to homoplasy, as in Fig. 2B. However, in Fig. 2A, such an assumption of irreversibility of the single canonical coding-region SNPs for either H3 (T6776C) seen in En14 and Ir3, or for H16 (C10394T) seen in En9, contrasts with a more parsimonious reconstruction based on three other sites in the CR, which suggests an alternative hypothesis where the H3 and H16 signature SNPs have evolved in parallel. Although H subgroups defined by one or two SNPs are widely used as analytical units in human mitogenomic phylogeography (Loogväli et al. 2004; Brandstätter et al. 2006; Roostalu et al. 2007), we suggest that there is no a priori reason to assume that the defining third-position silent transitions are not subject to homoplasic reversal. The possibil-

Table 3. Variation at 14 microsatellite loci among 27 Newfoundlanders.

Locus	N_A	H_E	H_O	F_{IS}
D3S3630	10	0.805	0.840	–0.045
D4S403	7	0.773	0.630	0.188
D5S408	6	0.769	0.680	0.118
D6S305	9	0.864	0.778	0.102
D7S2477	9	0.762	0.846	–0.113
D8S503	8	0.833	0.880	–0.056
D9S1779	7	0.600	0.478	0.207
D11S969	6	0.723	0.913	–0.271
D12S1638	6	0.605	0.625	–0.034
D13S285	10	0.824	0.800	0.030
D15S165	7	0.636	0.696	–0.097
D16S403	9	0.839	0.926	–0.106
D20S851	8	0.862	0.630	0.274*
D22S1169	8	0.785	0.750	0.046
Mean	7.9	0.763	0.748	0.0174

Note: Note: Numbers of alleles per locus (N_A), expected and observed heterozygosity (H_E and H_O), and inbreeding coefficient (F_{IS}) are given. The one F_{IS} value significant at $p < 0.05$ is indicated by *.

ity that independent patterns of nucleotide substitution may evolve in other, less well-characterized populations should be considered. In the present analysis, the most parsimonious explanation of the data requires three events fewer than the alternative with the monophyly of H3 and H16 preserved.

The two missing (or as yet unidentified) haplogroups, V and X, constitute, respectively, 4%–5% and 1%–2% of modern Europeans (Richards et al. 2000; Torroni et al. 2006). Frequencies of haplogroup V in England and western Ireland are estimated as 2.4% and 5.7%, respectively (Torroni et al. 2001); the estimated frequency for X in northwestern Europe and Scandinavian are both ~1% (Reidla et al. 2003).

Scandinavian populations did not contribute directly to the founding population of Newfoundland. The Viking settlement from Iceland ca. CE 1000 at L'Anse aux Meadows on the Northern Peninsula of Newfoundland was withdrawn after less than 100 years, did not involve intermarriage with local First Nations or aboriginal peoples, and left behind no known descendants (Wallace 2003). Viking (i.e., Norse) contribution to the Anglo–Irish gene pool during the same period and thence to Newfoundland immigrants would be indirect (Helgason et al. 2001).

First Nations haplogroups in Newfoundland

Phylogenetic analysis associates individual Fr2 with haplogroup A2 (Carr et al. 2008): a GenBank BLAST search (accessed August 2010) indicates differences of five substitutions from the next most similar individuals, all described as Native Americans. Haplogroup A2 is associated with northeastern Eurasian natives and North American First Nations and aboriginal peoples (Malhi et al. 2001; Mishmar et al. 2003). One such is the Mi'kmaq, a nation of the Algonquian First Nations language group who occupied most of Atlantic Canada by the middle of the 18th century and who established permanent settlements on the island of Newfoundland by the early 1800s (Bartels and Janzen 1990; Martijn 2003). The Mi'kmaq population is known to include individuals assigned to the A2, C, and X CR haplotypes (Malhi et al. 2001). Individual Fr2 shares 12 characteristic SNPs in the HVS-1 (Supplementary data, Table S3) and HVS-2 of the CR with two haplogroup A2 Mi'kmaq from Nova Scotia, and differs from either at only one site each in this region. We previously hypothesized that this individual is the matrilineal descendant of a Mi'kmaq mother and a French father (Carr et al. 2008). Individual Fr2 has French matrilineal ancestry to at least 1862 from the Port au Port Peninsula of the "French Shore" on the west coast of the island, a region with a long history of Mi'kmaq–French intermarriage: offspring of such unions are noted before 1850 (Story et al. 1999). Study of isolated European founder populations in the Gaspé Peninsula of Quebec, another region originally inhabited by the Mi'kmaq and other Algonquian First Nations, also identified the presence of eight mtDNA HVS-1 A, C, and X haplotypes, as well as the rarer D haplotype (Moreau et al. 2009). Individual Fr2 is distinguishable from both Gaspasian haplogroup A haplotypes (Supplementary data, Table S3).

The Beothuk were the only aboriginal group residents in Newfoundland at the time of first European contact, and became extinct in the 1820s (Marshall 1996). The HVS-1 sequence of individual Fr2 is distinct from those of two historical Beothuk individuals, who have been assigned to the C and X haplogroups (Kuch et al. 2007). The Beothuk X HVS-1 motif differs from that of six other known eastern Mi'kmaq and (or) Gaspasian X haplotypes (Supplementary data, Table S3), and the Beothuk C HVS-1 motif differs markedly from the other known eastern Mi'kmaq C motif, such that a close relationship between the two First Nations groups is not supported (cf. Kuch et al. 2007). Comparison of three additional Newfoundland Mi'kmaq C and X CR sequences confirms this (Collier 2010). The Beothuk C HVS-1 motif is indistinguishable from the single Gaspasian C motif, but the comparison includes several uncertain sites.

Mito-nuclear genome structure of Newfoundlanders

Neither loss of mtDNA heterogeneity nor skewed haplogroup frequencies with respect to the European source population can be demonstrated in the present data. If a particular haplogroup were over represented in the maternal founding population of a particular community, the initial association of this haplogroup with a set of nuclear alleles might be expected to persist for a number of generations. Similarly, if there were a significant shift in haplogroup frequencies between the English and Irish founders, partitioning of the nuclear variance would be expected. Although the microsatellite data do not provide evidence of linkage disequilibrium, they do detect slight but significant F_{IS}, on the same order as previously calculated from genealogical data in one local population (Bear et al. 1988). Given the broad geographic sampling here, F_{IS} might arise from "Wahlund effect" in an admixture of two or more differentiated populations. The absence of a significant F_{ST} suggests that this does not arise from differences between English and Irish ethnic groups per se. Another recent study of Newfoundlanders also found no marked founder effect, as measured by degree of linkage disequilibrium on one nuclear chromosome, with respect to other isolated populations and a European control population (Service et al. 2006). Molecular genetic studies focused on single small isolated communities in Newfoundland may yet identify founder, drift, and inbreeding phenomena on local scales (cf. Bear et al. 1987, 1988; Crawford et al. 1995; Martin et al. 2000).

The mtDNA results contrast with those from other European-derived Canadian founder populations, including the Acadian population on the Gaspé Peninsula of Quebec (Moreau et al. 2009) that showed loss of rare haplotypes, changes in the relative frequency of non-H haplogroups, and loss of diversity, as well as more extensive admixture of First Nations haplogroups. The Gaspé study assigned individuals to haplogroups based on a few hundred base pairs from the HVS-1 and HVS-2 regions and allele-specific oligonucleotides (ASOs) for SNPs elsewhere in the genome, which would tend to underestimate genomic diversity. In our data, where all individuals have unique whole-genome sequence, analysis of just the CR region would fail to resolve differences between five pairs of individuals.

Haplogroup proportions in English and Irish Newfoundlanders are not significantly differentiated (Table 2), nor are proportions between the much larger sample of European "British" (English) and Irish individuals compiled by Moreau et al. (2009) (Supplementary data, Table S3); both differ from their continental French compilation (analysis not shown). English Newfoundlanders examined belong predominantly (71%) to haplogroup H. A significant shift in haplogroup frequencies between pre-1000 CE and contemporary English populations has been documented (Töpf et al. 2006), including an increase in the frequency of haplogroup H from ~30% to 53%. The elevated frequency of H in the Newfoundland English data is suggestive, although the differences are not significant from either the historical or contemporary European samples in Töpf et al. (2006) or in Moreau et al. (2009). Should more extensive or focused sampling of Newfoundland outports identify any such shift as significant, the question would then arise whether these reflect founder and drift effects over the history of the colonial period, precolonial patterns arising from nonrandom sampling of the British populations, shifts in the English population since the colonial period, or combinations of any of these. Given the near-400 year insularization of the oldest Newfoundland lineages, identification of distinctive lineages that have persisted in the Newfoundland population while going extinct elsewhere would be genealogically and historically illuminating. With the rich tradition of family history and well-documented genealogy available in Newfoundland (Rahman et al. 2003), these questions are amenable to further study.

Acknowledgements

The mtDNA data were produced as part of the BSc (hons) and MSc dissertations of A.M. Pope, who thanks B.H. Younghusband for provision of a graduate fellowship, and B.H. Younghusband and T.L. Young for comments on the theses. The microsatellite data were produced as part of the BSc (hons) dissertation of K.N. Smith. We thank J. Green for access to samples and help with the analysis of genealogical records of the participants; J. Green, A.A. Collier, B.L. Pilgrim, and several anonymous reviewers for comments on this and earlier drafts; A.A. Collier for access to data on Newfoundland Mi'kmaq in advance of publication; J.T.A. Labour for the comparative analysis of GenBank sequences (accessed August 2010); S. Coady and L. Peddle of Newfound Genomics Inc. for access to their Affymetrix Chip Reader; E.A. Perry and E.S. Yaskowiak at the Genomics and Proteomics Facility at Memorial University of Newfoundland for assistance with the microsatellite analysis. The staff of the Centre for Newfoundland Studies, J. Hewson, and R. Rompkey at Memorial University provided valuable information on the history of the Mi'kmaq and French in Newfoundland, as have members of the Family History Society of Newfoundland and Labrador on the founding European population. Costs of the mtDNA sequencing and microarray experiments were defrayed by a University Research Support agreement with the Canadian Department of Fisheries and Oceans to S.M. Carr. The microsatellite experiments were supported by an NSERC Discovery grant to H.D. Marshall.

References

Bartels, D.A., and Janzen, O.U. 1990. Micmac migration to western Newfoundland. Can. J. Native Stud. **10**: 71–94.

Bear, J.C., Nemec, T.F., Kennedy, J.C., Marshall, W.H., Power, A.A., Kolonel, V.M., et al. 1987. Persistent genetic isolation in outport Newfoundland. Am. J. Med. Genet. **27**(4): 807–830. doi:10.1002/ajmg.1320270410. PMID:3425597.

Bear, J.C., Nemec, T.F., Kennedy, J.C., Marshall, W.H., Power, A.A., Kolonel, V.M., and Burke, G.B. 1988. Inbreeding in outport Newfoundland. Am. J. Med. Genet. **29**(3): 649–660. doi:10.1002/ajmg.1320290324. PMID:3377008.

Brandon, M.C., Ruiz-Pesini, E., Mishmar, D., Procaccio, V., Lott, M.T., Nguyen, K.C., et al. 2009. MITOMASTER: a bioinformatics tool for the analysis of mitochondrial DNA sequences. Hum. Mutat. **30**(1): 1–6. doi:10.1002/humu.20801. PMID:18566966.

Brandstätter, A., Salas, A., Niederstätter, H., Gassner, C., Carracedo, A., and Parson, W. 2006. Dissection of mitochondrial superhaplogroup H using coding region SNPs. Electrophoresis,

27(13): 2541–2550. doi:10.1002/elps.200500772. PMID: 16721903.

Carr, S.M., and Marshall, H.D. 2008*a*. Intraspecific phylogeographic genomics from multiple complete mtDNA genomes in Atlantic cod (*Gadus morhua*): origins of the "codmother," transatlantic vicariance and midglacial population expansion. Genetics, **180**(1): 381–389. doi:10.1534/genetics.108.089730. PMID:18716332.

Carr, S.M., and Marshall, H.D. 2008*b*. Phylogeographic analysis of complete mtDNA genomes from Walleye Pollock (*Gadus chalcogrammus* Pallas, 1811) shows an ancient origin of genetic biodiversity. Mito. DNA, **19**(6): 490–496. doi:10.1080/19401730802570942.

Carr, S.M., Marshall, H.D., Duggan, A.T., Flynn, S., Johnstone, K., Pope, A., and Wilkerson, C. 2008. Phylogeographic genomics of mitochondrial DNA: highly-resolved patterns of intraspecific evolution and a multi-species, microarray-based DNA sequencing strategy for biodiversity studies. Comp. Biochem. Physiol., D: Genomics Proteomics, **3**(1): 1–11. doi:10.1016/j.cbd.2006.12.005.

Carr, S.M., Duggan, A.T., and Marshall, H.D. 2009. Iterative DNA sequencing microarrays: a high-throughput NextGen technology for ecological and evolutionary mitogenomics. Lab. Focus, **13**: 8–12.

Collier, A.A. 2010. Mitochondrial haplotype analyses of Canadians of aboriginal descent, with particular emphasis on the Mi'kmaq of the Miawpukek First Nation, Newfoundland. BSc. (hons) thesis, Memorial University of Newfoundland, St. John's, Nfld.

Crawford, M.H., Koertevlyessy, T., Huntsman, R.G., Collins, M., Duggirala, R., Martin, L., et al. 1995. Effects of religion, economics, and geography on genetic structure of Fogo Island, Newfoundland. Am. J. Hum. Biol. **7**(4): 437–451. doi:10.1002/ajhb.1310070405.

Excoffier, L., Laval, G., and Schneider, S. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol. Bioinf. Online, **1**: 47–50. PMID:19325852.

Flynn, S.M.C., and Carr, S.M. 2007. Interspecies hybridization on DNA resequencing microarrays: efficiency of sequence recovery and accuracy of SNP detection in human, ape, and codfish mitochondrial DNA genomes sequenced on a human-specific MitoChip. BMC Genomics, **8**(1): 339. doi:10.1186/1471-2164-8-339. PMID:17894875.

Handcock, W.G. 1989. Soelonge as there comes noe women: origins of English settlement in Newfoundland. Breakwater Books, St. John's, Nfld.

Hartmann, A., Thieme, M., Nanduri, L.K., Stempfl, T., Moehle, C., Kivisild, T., and Oefner, P.J. 2009. Validation of microarray-based resequencing of 93 worldwide mitochondrial genomes. Mutat. Res. **30**(1): 115–122. doi:10.1002/humu.20816.

Helgason, A., Hickey, E., Goodacre, S., Bosnes, V., Stefánsson, K., Ward, R., and Sykes, B. 2001. mtDna and the islands of the North Atlantic: estimating the proportions of Norse and Gaelic ancestry. Am. J. Hum. Genet. **68**(3): 723–737. doi:10.1086/318785. PMID:11179019.

Kuch, M., Gröcke, D.R., Knyf, M.C., Gilbert, M.T., Younghusband, B., Young, T., et al. 2007. A preliminary analysis of the DNA and diet of the extinct Beothuk: a systematic approach to ancient human DNA. Am. J. Phys. Anthropol. **132**(4): 594–604. doi:10.1002/ajpa.20536. PMID:17205549.

Loogväli, E.L., Roostalu, U., Malyarchuk, B.A., Derenko, M.V., Kivisild, T., Metspalu, E., et al. 2004. Disuniting uniformity: a pied cladistic canvas of mtDNA haplogroup H in Eurasia. Mol. Biol. Evol. **21**(11): 2012–2021. doi:10.1093/molbev/msh209. PMID:15254257.

Maca-Meyer, N., González, A.M., Larruga, J.M., Flores, C., and Cabrera, V.M. 2001. Major genomic mitochondrial lineages delineate early human expansions. BMC Genet. **2**(1): 13. doi:10.1186/1471-2156-2-13. PMID:11553319.

Macaulay, V., Richards, M., Hickey, E., Vega, E., Cruciani, F., Guida, V., et al. 1999. The emerging tree of west Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. Am. J. Hum. Genet. **64**(1): 232–249. doi:10.1086/302204. PMID:9915963.

Maitra, A., Cohen, Y., Gillespie, S.E., Mambo, E., Fukushima, N., Hoque, M.O., et al. 2004. The Human MitoChip: a high-throughput sequencing microarray for mitochondrial mutation detection. Genome Res. **14**(5): 812–819. doi:10.1101/gr.2228504. PMID:15123581.

Malhi, R.S., Schultz, B.A., and Smith, D.G. 2001. Distribution of mitochondrial DNA lineages among Native American tribes of northeastern North America. Hum. Biol. **73**(1): 17–55. doi:10.1353/hub.2001.0008. PMID:11332644.

Mannion, J.J. (*Editor*). 1977. The peopling of Newfoundland: essays in historical geography. Institute of Social and Economic Research, Memorial University of Newfoundland, St. John's, Nfld.

Marshall, I. 1996. A history and ethnography of the Beothuk. McGill-Queen's University Press, Montreal, Que.

Martijn, C.A. 2003. Early Mi'kmaq presence in southern Newfoundland: an ethnohistorical perspective, c.1500–1763. Nfld. Labrador Stud., **19**: 44–102.

Martin, L.J., Crawford, M.H., Koertvelyessy, T., Keeping, D., Collins, M., and Huntsman, R. 2000. The population structure of ten Newfoundland outports. Hum. Biol. **72**(6): 997–1016. PMID:11236869.

Mishmar, D., Ruiz-Pesini, E., Golik, P., Macaulay, V., Clark, A.G., Hosseini, S., et al. 2003. Natural selection shaped regional mtDNA variation in humans. Proc. Natl. Acad. Sci. U.S.A. **100**(1): 171–176. doi:10.1073/pnas.0136972100. PMID:12509511.

Moore, S.J., Green, J.S., Fan, Y., Bhogal, A.K., Dicks, E., Fernandez, B.A., et al. 2005. Clinical and genetic epidemiology of Bardet–Biedl syndrome in Newfoundland: a 22-year prospective, population-based, cohort study. Am. J. Med. Genet. A. **132**(4): 352–360. PMID:15637713.

Moore, S.J., Buckley, D.J., MacMillan, A., Marshall, H.D., Steele, L., Ray, P.N., et al. 2008. The clinical and genetic epidemiology of neuronal ceroid lipofuscinosis in Newfoundland. Clin. Genet. **74**(3): 213–222. doi:10.1111/j.1399-0004.2008.01054.x. PMID:18684116.

Moreau, C., Vézina, H., Yotova, V., Hamon, R., de Knijff, P., Sinnett, D., et al. 2009. Genetic heterogeneity in regional populations of Quebec—Parental lineages in the Gaspé Peninsula. Am. J. Phys. Anthropol. **139**(4): 512–522. doi:10.1002/ajpa.21012. PMID:19226649.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York.

O'Neill, P. 2003. The oldest city: the story of St. John's, Newfoundland. Boulder Publications, Portugal Cove – St. Philip's, Nfld.

Parfrey, P.S., Davidson, W.S., and Green, J.S. 2002. Clinical and genetic epidemiology of inherited renal disease in Newfoundland. Kidney Int. **61**(6): 1925–1934. doi:10.1046/j.1523-1755.2002.00305.x. PMID:12028433.

Pope, A.M. 2003. An investigation of the ethnic composition of the Newfoundland population based on whole mitochondrial genomes. BSc. (hons) thesis, Memorial University of Newfoundland, St. John's, Nfld.

Pope, A.M. 2007. An investigation into the phylogeography of the founder population of Newfoundland – the importance of whole mitochondrial genome sequencing. M.Sc. thesis, Memorial University of Newfoundland, St. John's, Nfld.

Pritchard, J.K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genetics, **155**(2): 945–959. PMID:10835412.

Prowse, D.W. 1895. A history of Newfoundland from the English, Colonial, and foreign records. MacMillan and Company, New York.

Rahman, P., Jones, A., Curtis, J., et al. 2003. The Newfoundland population: a unique resource for genetic investigation of complex diseases. Hum. Mol. Genet. **12**(suppl 2): R167–R172. PMID:12915452.

Raymond, M., and Rousset, F. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J. Hered. **86**(3): 248–249.

Reidla, M., Kivisild, T., Metspalu, E., Kaldma, K., Tambets, K., Tolk, H.V., et al. 2003. Origin and diffusion of mtDNA haplogroup X. Am. J. Hum. Genet. **73**(5): 1178–1190. doi:10.1086/379380. PMID:14574647.

Richards, M.B., Macaulay, V.A., Bandelt, H.J., and Sykes, B.C. 1998. Phylogeography of mitochondrial DNA in western Europe. Ann. Hum. Genet. **62**(3): 241–260. doi:10.1046/j.1469-1809.1998.6230241.x. PMID:9803269.

Richards, M., Macaulay, V., Hickey, E., Vega, E., Sykes, B., Guida, V., et al. 2000. Tracing European founder lineages in the Near Eastern mtDNA pool. Am. J. Hum. Genet. **67**(5): 1251–1276. PMID:11032788.

Rieder, M.J., Taylor, S.L., Tobe, V.O., and Nickerson, D.A. 1998. Automating the identification of DNA variations using quality-based fluorescence re-sequencing: analysis of the human mitochondrial genome. Nucleic Acids Res. **26**(4): 967–973. doi:10.1093/nar/26.4.967. PMID:9461455.

Roff, D.A., and Bentzen, P. 1989. The statistical analysis of mitochondrial DNA polymorphisms: chi-2 and the problem of small samples. Mol. Biol. Evol. **6**(5): 539–545. PMID:2677600.

Roostalu, U., Kutuev, I., Loogväli, E.L., Metspalu, E., Tambets, K., Reidla, M., et al. 2007. Origin and expansion of haplogroup H, the dominant human mitochondrial DNA lineage in West Eurasia: the Near Eastern and Caucasian perspective. Mol. Biol. Evol. **24**(2): 436–448. doi:10.1093/molbev/msl173. PMID:17099056.

Rosenberg, N.A., Pritchard, K., Webert, J.L., et al. 2002. Genetic structure of human populations. Science, **298**(5602): 1381–1385. doi:10.1126/science.1078311. PMID:12399543.

Rowe, F.W. 1980. A history of Newfoundland and Labrador. McGraw-Hill Ryerson, Toronto, Ont.

Service, S., DeYoung, J., Karayiorgou, M., Roos, J.L., Pretorious, H., Bedoya, G., et al. 2006. Magnitude and distribution of linkage disequilibrium in population isolates and implications for genome-wide association studies. Nat. Genet. **38**(5): 556–560. doi:10.1038/ng1770. PMID:16582909.

Smith, K.N. 2009. Genetic diversity within the human population of Newfoundland: do microsatellites recover mitochondrial haplogroup structure? BSc. (hons) thesis, Memorial University of Newfoundland, St. John's, Nfld.

Story, G.M., Kirwin, W.J., and Widdowson, J.D.A. 1999. Dictionary of Newfoundland English. 2nd ed. University of Toronto Press, Toronto, Ont.

Swofford, D. 2002. PAUP*: Phylogenetic Analysis Using Parsimony (and other methods). v. 4.0 Beta. Florida State University.

Töpf, A.L., Gilbert, M.T.P., Dumbacher, J.P., and Hoelzel, A.R. 2006. Tracing the phylogeography of human populations in Britain based on 4th–11th century mtDNA genotypes. Mol. Biol. Evol. **23**(1): 152–161. doi:10.1093/molbev/msj013.

Torroni, A., Bandelt, H.-J., Macaulay, V., Richards, M., Cruciani, F., Rengo, C., et al. 2001. A signal, from human mtDNA, of postglacial recolonization in Europe. Am. J. Hum. Genet. **69**(4): 844–852. doi:10.1086/323485. PMID:11517423.

Torroni, A., Achilli, A., Macaulay, V., Richards, M., and Bandelt, H.J. 2006. Harvesting the fruit of the human mtDNA tree. Trends Genet. **22**(6): 339–345. doi:10.1016/j.tig.2006.04.001. PMID:16678300.

van Oven, M., and Kayser, M. 2009. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum. Mutat. **30**(2): E386–E394. doi:10.1002/humu.20921 . PMID:18853457.

Wallace, B. 2003. The Norse in Newfoundland: L'Anse aux Meadows and Vinland. Nfld. Stud., **19**: 5–43.

Winsor, J. (*Editor*). 1887. Narrative and critical history of America: the English and French in North America 1689–1763. Vol. V. Houghton, Mifflin & Co., Boston, Mass.

Woods, M.O., Young, T.L., Parfrey, P.S., Hefferton, D., Green, J.S., and Davidson, W.S. 1999. Genetic heterogeneity of Bardet–Biedl syndrome in a distinct Canadian population: evidence for a fifth locus. Genomics, **55**(1): 2–9. doi:10.1006/geno.1998.5626. PMID:9888993.

Woods, M.O., Hyde, A.J., Curtis, F.K., Stuckless, S., Green, J.S., Pollett, A.F., et al. 2005. High frequency of hereditary colorectal cancer in Newfoundland likely involves novel susceptibility genes. Clin. Cancer Res. **11**(19): 6853–6861. doi:10.1158/1078-0432.CCR-05-0726. PMID:16203774.

Young, T.L., Woods, M.O., Parfrey, P.S., Green, J.S., Hefferton, D., and Davidson, W.S. 1999. A founder effect in the Newfoundland population reduces the Bardet–Biedl syndrome I (*BBS1*) interval to 1 cM. Am. J. Hum. Genet. **65**(6): 1680–1687. doi:10.1086/302686. PMID:10577922.