Received: 26 October 2019 | Revised: 13 January 2020 | Accepted: 4 February 2020

DOI: 10.1002/ajpa.24024

RESEARCH ARTICLE

American Journal of PHYSICAL ANTHROPOLOGY | WILEY

Historical human remains identification through maternal and paternal genetic signatures in a founder population with extensive genealogical record

Tommy Harding[1,2] | Emmanuel Milot[2,3] | Claudia Moreau[1,4] | Jean-Francois Lefebvre[1] | Jean-Sébastien Bournival[5] | Hélène Vézina[4,5,6] | Catherine Laprise[4,7] | Carles Lalueza-Fox[8] | Roger Anglada[9] | Brad Loewen[10] | Ferran Casals[9] | Isabelle Ribot[10] | Damian Labuda[1,11]

[1]Centre de recherche du CHU Sainte-Justine, Université de Montréal, Montréal, Canada

[2]Laboratoire de recherche en criminalistique, Département de chimie, biochimie et sciences de l'énergie, Université du Québec à Trois-Rivières, Trois-Rivières, Canada

[3]Centre international de criminologie comparée and Centre interuniversitaire d'études québécoises, Université du Québec à Trois-Rivières, Trois-Rivières, Canada

[4]Centre intersectoriel en santé durable, Université du Québec à Chicoutimi, Chicoutimi, Canada

[5]Projet BALSAC, Université du Québec à Chicoutimi, Chicoutimi, Canada

[6]Département des sciences humaines et sociales, Université du Québec à Chicoutimi, Chicoutimi, Canada

[7]Département des sciences fondamentales, Université du Québec à Chicoutimi, Chicoutimi, Canada

[8]Institut de Biologia Evolutiva, CSIC-Universitat Pompeu Fabra, Barcelona, Spain

[9]Genomics Core Facility, Universitat Pompeu Fabra, Barcelona, Spain

[10]Département d'anthropologie, Université de Montréal, Montréal, Canada

[11]Département de pédiatrie, Université de Montréal, Montréal, Canada

Correspondence
Tommy Harding and Damian Labuda, Centre de recherche du CHU Sainte-Justine, Université de Montréal, Montréal, Canada.
Email: tommy.harding@umontreal.ca (T. H.) and damian.labuda@umontreal.ca (D. L.)

Funding information
European Regional Development Fund, Grant/Award Number: PGC2018-095931-B-100; Fonds de Recherche du Québec - Santé; Fonds de Recherche du Québec-Société et Culture; Natural Sciences and Engineering Research Council of Canada

Abstract

Objectives: We describe a method to identify human remains excavated from unmarked graves in historical Québec cemeteries by combining parental-lineage genetic markers with the whole-population genealogy of Québec contained in the BALSAC database.

Materials and methods: The remains of six men were exhumed from four historical cemeteries in the province of Québec, Canada. DNA was extracted from the remains and genotyped to reveal their mitochondrial and Y-chromosome haplotypes, which were compared to a collection of haplotypes of genealogically-anchored modern volunteers. Maternal and paternal genealogies were searched in the BALSAC genealogical record for parental couples matching the mitochondrial and the Y-chromosome haplotypic signatures, to identify candidate sons from whom the remains could have originated.

Results: Analysis of the matching genealogies identified the parents of one man inhumed in the cemetery of the investigated parish during its operating time. The candidate individual died in 1833 at the age of 58, a plausible age at death in light of osteological analysis of the remains.

Discussion: This study demonstrates the promising potential of coupling genetic information from living individuals to genealogical data in BALSAC to identify historical human remains. If genetic coverage is increased, the genealogical information in BALSAC could enable the identification of 87% of the men (n = 178,435) married in Québec before 1850, with high discriminatory power in most cases since >75% of the parental couples have unique biparental signatures in most regions. Genotyping and identifying Québec's historical human remains are a key to reconstructing the genomes of the founders of Québec and reinhuming archeological remains with a marked grave.

KEYWORDS

ancient DNA, buried's identification, uniparental DNA markers, whole-population genealogy

1 | INTRODUCTION

A century ago, molecular markers (ABO blood groups) were used for the first time to study human diversity and to infer population histories (Hirschfeld & Hirschfeld, 1919). It took another 50 years to start exploring this diversity through studies of isoenzymes and protein polymorphisms (Cavalli-Sforza, Menozzi, & Piazza, 1994). The characterization of populations by their DNA began with the analysis of maternally-inherited mitochondrial DNA (mtDNA) followed by investigations of paternally-inherited Y-chromosome variation (Cann, Stoneking, & Wilson, 1987; Casanova et al., 1985; Giles, Blanc, Cann, & Wallace, 1980; Jobling & Tyler-Smith, 1995). Mitochondrial DNA and the Y chromosome, because of their strict inheritance rules and lack of recombination (with the exception of the Y-chromosome pseudoautosomal regions), remain central tools in the population genetics of paternal and maternal lineages. Notably, they are instrumental in the characterization of population stratification and past migrations (Calafell & Larmuseau, 2017; Underhill & Kivisild, 2007).

In a familial search context, forensic science and bioarchaeology regularly compare DNA samples isolated from human remains and living presumed relatives to verify their familial relationship. Analysis of ancient DNA (we use this expression for DNA recovered from archeological contexts, regardless of its age) has confirmed the identity of remains of eminent historical figures (Bogdanowicz et al., 2009, Charlier et al., 2013, but see Larmuseau et al., 2014). The bones of the Romanovs, the Russian royal family buried in 1918 in a mass grave after their execution by Bolshevik revolutionaries, have been identified through genetic analysis (Coble et al., 2009; Gill et al., 1994; Rogaev et al., 2009). Ancient DNA and genealogical links were used to identify the archeological remains of the 15th-century English king Richard III (King et al., 2014). These studies succeeded by comparing archeological samples from a hypothetical historical figure, with living relatives willing to provide their DNA for analysis. In Canada, ancient DNA recovered from 80 to 200-year-old cemeteries has been studied to infer past societal kinship systems and identify one individual, through comparison with living relatives (Dudar, Waye, & Saunders, 2003; Katzenberg et al., 2005). However, this approach does not apply to archeological finds that preclude any hypothesis as to the individual's familial identity, as is almost universally the case in mitigation archeology that uncovers many human remains every year.

Here, we explore the potential of identifying anonymous remains, excavated from four historical cemeteries in Québec, through the combined analysis of genetic data and whole-population genealogical data. The proposed approach may be conceived as a triangulation of ancient DNA from a random archeological individual, contemporary DNA with wide coverage within today's Québec population, and whole-population Québec genealogies. The anonymous archeological individuals were tentatively identified by comparing their lineage-specific Y-chromosome and mtDNA haplotypes with a collection of contemporary haplotypes of Québec individuals with known genealogies. This is feasible thanks to the high genealogical coverage of the historical Québec population since the first French settlement in 1608. Catholic parish registers include birth, marriage, and burial acts, allowing the genealogical reconstruction of the quasi-entire population over 350 years that is now compiled in the databases known as BALSAC (http://balsac.uqac.ca) and IMPQ (Infrastructure intégrée des microdonnées historiques de la population du Québec, https://impq.cieq.ca).

The method presented here relies on three essential components. First, the mtDNA and Y-chromosome haplotypes from the specimens are compared with the haplotypes of hundreds of living individuals who are genealogically anchored in the BALSAC database. Second, the matching haplotypes of the contemporary individuals are imputed to specific population founders by climbing maternal and paternal lines and subsequently attributed down to all the founders' descendants (Heyer et al., 2001; Heyer, Puymirat, Dieltjes, Bakker, & de Knijff, 1997; Moreau et al., 2011). Third, we search for male individuals in the genealogy sharing the same combination of Y-chromosome and mtDNA haplotypes as the historical specimens (Figure 1).

In this study, we test this approach empirically by analyzing the bones of six male individuals from unmarked graves. We also examine the whole-population genealogical records to determine the number of historical men that could potentially be identified using this approach and calculate the probability of erroneous matches. Despite

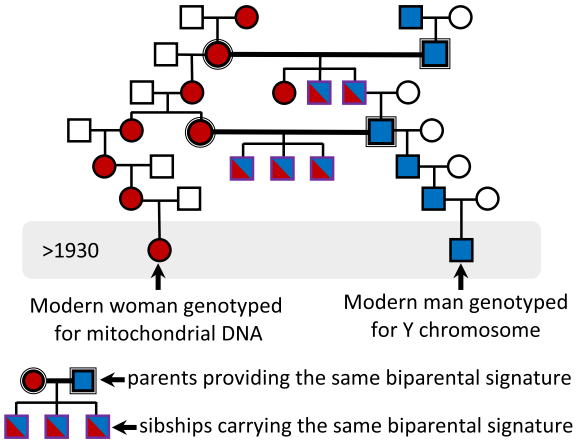

FIGURE 1 The methodological framework of this study. We search for parental couples in the uniparental genealogies of modern individuals carrying haplotypes matching the ancient DNA. Candidates (outlined in purple and containing red and blue triangles) to whom the historical remains could have belonged are the sons of parents with the matching biparental signature, that is, the sons of the mothers in the maternal genealogies of modern individuals with matching mitochondrial haplotype (in red) married to the fathers in the paternal genealogies of modern men with matching Y-chromosome profile (in blue)

a limited success due to the quality of the ancient DNA, the results show that successful identification, and thus anchoring archeological remains to the Québec genealogy, should be readily possible by increasing the genetic coverage of the contemporary population. Identified remains of ancestors may convey genetic information useful for many purposes, including the study of evolutionary factors shaping genetic diversity of extant populations, the study of the temporal evolution of genetic variants involved in diseases, and the calibration of paleogenome inference methods (Nelson et al., 2018).

2 | MATERIALS AND METHODS

2.1 | Samples of human remains selected for ancient DNA analysis

Human remains (teeth and petrous bones) were selected from six historical unmarked graves (Table 1). They originate from four cemeteries located in the province of Québec (Figure 2). As all these interment sites date to a period ranging from the end of the 17th to the 19th centuries, they can potentially inform us about the genetic diversity of early colonial inhabitants compared to later phases. However, while the location of the burials is known within the cemetery, it is not possible to precisely date the individuals under study. In fact, the internal organization of these cemeteries that were used for more than a century, without grave markers, is complex and not fully understood, as cemetery areas were often reused at a later date, especially in urban contexts (Arkéos, 2008). Note that, since these specimens were more recent than typical ancient DNA research samples that date back to several tens of thousands of years, we refer to them as "historical" specimens but use the expression "ancient DNA" for their genetic material, considering it was subject to the same degradation patterns as thousand-year-old DNA and processed through state-of-the-art ancient DNA analysis.

These six individuals were selected from a larger sample (N = 30) on the basis of the following conditions: (a) the individual had to be a male (sex determination was first provided via the osteological analysis and then confirmed or disproved via paleogenetic evidence) and (b) the skeletal material had to provide enough DNA for replicated PCR amplifications (see *Genotyping of historical human remains* below). Out of the 30 individuals, only 6 met both conditions. Identification of the male sex (as well as the age at death) was done osteologically for all the individuals, although, because of variable skeletal preservation, the most reliable anatomical area such as the pelvis was not always available (see Table 1 for all the methods used).

Three individuals (12D-S1, 4K-S2, and 11D-S1) originated from the Notre-Dame cemetery, a Catholic parish cemetery in Montréal that remained active from 1691 to 1796 (Arkéos, 2008). Individual 20A-S12 was discovered in the Saint-Antoine cemetery (1799–1855), a Montréal burial ground that was opened after the closure of Notre-Dame (Ethnoscop, 2016b). Individual 7A9-S16 originated from Pointe-aux-Trembles (1709–1843), a village cemetery located near Montréal (Ethnoscop, 2016a). Finally, individual 2B7 was discovered in Sainte-Marie-de-Beauce (1748–1878), a rural cemetery near Québec City (Ethnoscop, 2006).

All corresponded to mature adults although the age estimation for 12D-S1 and 20A-S12 was less reliable because of poor skeletal preservation (when the pelvis was missing, methods based on the cranium were attempted, see Table 1). Interestingly, for the Notre-Dame individuals, isotopic analysis on skeletal tissues provided some biographical information (e.g., diet, possible place of residence), which confirmed that these individuals were settlers of European descent (Vigeant, 2012; Vigeant, Ribot, & Hélie, 2017). For example, individual 11D-S1 was probably a first generation migrant originating from warmer and/or more coastal areas, possibly from Western Europe given the historical context (Vigeant, 2012).

2.2 | Genotyping of historical human remains

DNA was extracted first from multiple samples of specimen 2B7: two teeth (in 2013, 2016, and 2018) and a temporal bone (in 2016). Teeth were prepared for DNA extraction by cutting the root in half, sanding the channel and the cementum to remove sources of exogenous DNA, and finally crushed. The sample from the petrous bone was obtained by drilling with a Dremel 400-series digital machine as described by Gamba et al. (2014). The analyst wore a complete bodysuit, a mask and two layers of gloves while working in a clean DNA room dedicated to work on ancient samples at the Genomics Core Facility, Universitat Pompeu Fabra (Barcelona, Spain). DNA was extracted, including negative controls, with the PrepFiler BTA forensic DNA extraction kit following the manufacturer's instructions (Thermo

TABLE 1 Samples of human remains selected for ancient DNA analysis

Site (code)	Code of skeleton	Sex estimation	Age at death	Sample for ancient DNA
Notre-Dame (BjFj-112)	12D-S1	♂ according to skull morphology (Buikstra & Ubelaker, 1994)	>*25 years* according to dental wear (Lovejoy, 1985)	Right superior permanent canine
	4K-S2	♂ according to pelvic morphometrics (Murail, Brůžek, Houët, & Cunha, 2005) and skull morphology (Buikstra & Ubelaker, 1994)	>*40 years* according to auricular surface (Schmitt, 2005) ~23–57 years according to pubicmorphology (Brooks &Suchey, 1990)	Right superior permanent canine
	11D-S1	♂ according to skull morphology (Buikstra & Ubelaker, 1994)	*~45–59 years* according to pubic morphology (Brooks & Suchey, 1990) *~35–55 years* according to rib end morphology (Lovejoy, 1985)	Left superior permanent third molar
Saint-Antoine (BjFj-37)	20A-S12	♂ according to skull morphology (Buikstra & Ubelaker, 1994)	>*50 years* according to cranial suture closure (Meindl & Lovejoy, 1985) & dental wear (Lovejoy, 1985)	Left inferior permanent canine
Pointe-aux-Trembles (BjFi-17)	7A9-S16	♂ according to skull morphology (Buikstra & Ubelaker, 1994)	>*30 years* according to vertebral changes (Albert & Maples, 1995) *~25–49 years* according to cranial sutures closure (Meindl & Lovejoy, 1985)	Left superior second permanent premolar
Sainte-Marie-de-Beauce (CcEs-1)	2B7	♂ according to pelvic morphometrics (Murail *et al*., 2005)	*~20–49 years* according to auricular surface (Schmitt, 2005) ~19–34 years according to pubic morphology (Brooks & Suchey, 1990) ~27–37 years according to rib end morphology (Hartnett, 2010)	Left superior second permanent molar, left superior second permanent premolar, petrous bone

Fisher Scientific). Human DNA in the extracts was quantified using the Quantifiler™ Human DNA quantification kit (Thermo Fisher Scientific). DNA was then extracted from the other specimens in 2018 (Table 1), as well as from buccal swabs from the technicians and scientists at the Genomics Core Facility to control for potential contamination. Results were not replicated by another independent ancient DNA laboratory. This project aimed to be a proof of concept and future analyses for formal remain identification should include a validation by another DNA laboratory, whenever possible.

Y-chromosome STRs were amplified with the AmpFLSTR™ Yfiler™ PCR amplification kit (Thermo Fisher Scientific). Amplicons were separated by capillary electrophoresis on a 3130×L instrument (injection time increased to 40 s relative to the manufacturer's instructions) and visualized using the GeneMapper™ software version 5 (Thermo Fisher Scientific). For each historical specimen, between four and 13 PCRs were done using varying amounts of DNA to recover at least three typed alleles for each locus, of which at least two alleles had to be identical. This was shown by Butler and Hill (2010) to be a reliable criterion to assign allele IDs in cases of low-template DNA. For match searches, we considered 12 loci already amplified from 434 modern volunteers (see below) and we retained loci for which the allele was consistently amplified at least twice and showed a peak of >50 relative fluorescence units (Benschop et al., 2011; Butler & Hill, 2010).

The mitochondrial hypervariable regions I and II were amplified in four segments using the PCR primers from the Human mtDNA D-Loop Hypervariable Region kit (Illumina). Next, adaptors were ligated using the Rubicon Genomics ThruPLEX DNA-seq kit (Takara Bio USA), and the library sequenced in a 2 × 150 MiSeq run with a nano flow cell (Illumina). Adapter and PCR primer sequences were trimmed with cutadapt version 1.17 (Martin, 2011) and the remaining sequences were further trimmed based on quality with Trimmomatic version 0.38 (Bolger, Lohse, & Usadel, 2014) using a sliding window of ten bases and a Phred quality score threshold of 35, keeping reads longer than 50 base pairs. Each read of a pair was then assembled with PEAR version 0.9.10 (Zhang, Kobert, Flouri, & Stamatakis, 2014). Resulting sequences were mapped on the revised Cambridge Reference Sequence of the human mitochondrial DNA (rCRS, GenBank accession number NC_012920.1, Andrews et al., 1999) with Bowtie2 version 2.3.0 (Langmead, Trapnell, Pop, & Salzberg, 2009) and analyzed for variant calls using Freebayes version 1.1.0 (Garrison & Marth, 2012) in the pooled-continuous mode (minimum alternate fraction: 0.01, ploidy: 1). As for the mtDNA Tree Build 17, positions located in homopolymeric nucleotide tracts (309, 315, 16,182, 16,183, and 16,193) were not considered.

Since each PCR reaction contained a complex mixture of sequences, including the original historical mitochondrial sequences and their derived forms due to postmortem damage, we used a series of in-house Python scripts to characterize the pools of sequences. For each

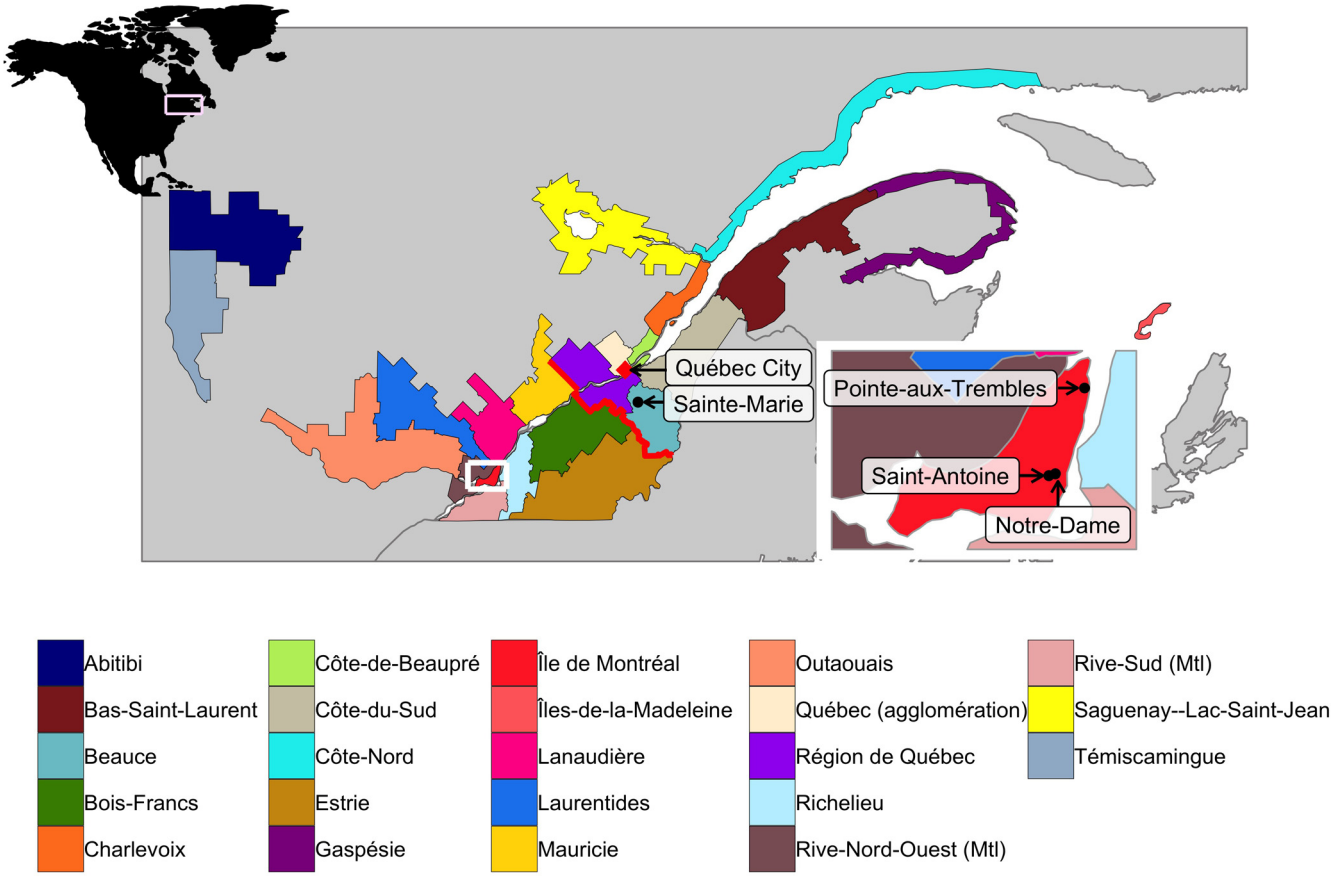

FIGURE 2 Location of the study cemeteries (black dots: Sainte-Marie, Pointe-aux-Trembles, Saint-Antoine, and Notre-Dame) in Québec, and the 23 regions as defined by the BALSAC team. The location of the study area in North America is displayed by the light purple rectangle in the top-left corner. The island of Montréal (white rectangle on the main map) is shown in greater detail on the bottom-right corner. A red diamond shows the location of Québec city. The division between western regions and eastern regions is indicated by the red line

individual, we assumed that the most abundant amplicon types originated from the historical individual. We then confirmed the historical nature of these sequences by identifying other sequences in the pool that were identical except at positions where we found misincorporated bases (thymine instead of cytosine, C → T, and adenine instead of guanine, G → A) that are typical of postmortem degradation in ancient DNA sequences. Furthermore, since each hypervariable region was amplified in two pieces, we confirmed that the overlapping parts (positions 16,159–16,236 for HVRI and 172–285 for HVRII) of most abundant amplicons were identical to ensure their authenticity. We confirmed the excess of C → T and G → A misincorporations in the data originating from the historical samples by comparing them to a set of 79 modern mitochondrial samples. The complete mitochondrial genome for these modern samples was sequenced on the MiSeq platform (2 × 250 bp) following the Human mtDNA genome kit protocol for Illumina platforms.

2.3 | Genealogies and genetic signatures of maternal and paternal lineages

The genealogy of virtually all French Canadian individuals married in Québec between the beginning of European settlement in the early 17th century and 1960 was compiled in the BALSAC database (N = 4,364,381 individuals, as of March 2015, http://balsac.uqac.ca). In prior studies, the coverage was extended beyond 1960 to include lines of a subset of contemporary individuals from varying regions of Québec, that were genotyped at 12 short tandem repeats (STRs; DYS391, DYS389I, DYS389II, DYS439, DYS438, DYS437, DYS19, DYS392, DYS393, DYS390, DYS385a, and DYS385b) on the Y chromosome (n = 434 men) and sequenced at the mtDNA D-loop for hypervariable regions I and II (HVRI positions 16,069 to 16,383 and HVRII positions 58 to 370; n = 530 women and 433 men; Moreau, Vézina et al., 2011; Moreau et al., 2009; Roy-Gagnon et al., 2011; and unpublished data). The men for whom the Y chromosome was typed belonged to 314 distinct paternal genealogies and the individuals for which the mtDNA was typed belonged to 423 distinct maternal genealogies. Among the 290 unique Y-chromosome haplotypes, 24 (8%) were found in more than one genealogy, and among the 297 unique mitochondrial haplotypes, 79 (27%) were found in more than one genealogy. In addition to saliva or peripheral blood samples, consenting individuals provided the information required to reconstruct their ascending genealogies. This research was approved by the Research Ethics Board at the Centre hospitalier universitaire Sainte-Justine in Montréal.

We complemented the genealogical data contained in BALSAC with data on burial acts contained in the IMPQ register (https://impq.cieq.ca). Notably, the IMPQ register contains the locations and dates of inhumation of most individuals buried between 1627 and 1849 in Québec (N = 585,621). The cemeteries from which the studied remains were exhumed were administered by the parishes of Notre-Dame de Montréal (Saint-Antoine and Notre-Dame cemeteries), Saint-Enfant-Jésus-de-la-Pointe-aux-Trembles (Pointe-aux-Trembles cemetery), and Sainte-Marie-de-Beauce (Sainte-Marie cemetery). Since the Saint-Antoine and Notre-Dame cemeteries were administered by the same parish, we cannot distinguish between them when analyzing burial acts compiled in the IMPQ database.

2.4 | DNA match searches

We searched for DNA matches between a given historical specimen and individuals in BALSAC with known haplotypes. First, we identified the modern individuals that carried the same mtDNA haplotype as the historical specimen. We then ascended their maternal lineages until the lineage founders and descended the genealogies to identify all the maternal ancestors that bore the matching haplotype. Second, we repeated the same strategy for the modern men that carried the same Y-chromosome haplotype as the historical specimen by ascending and descending their paternal genealogies to identify all the men that potentially carried this Y-chromosome haplotype as well. For each specimen, we only compared the alleles at the Y-STR loci that passed the quality filters described in the section *Genotyping of historical human remains*. Third, we listed the parental couples that were constituted of women in the matching maternal genealogies and of men in the matching paternal genealogies. To be selected, these parents needed to have had children married in Québec at least 75 years (the maximum interval between marriage and burial years observed in the IMPQ register) before the opening of the historical specimen cemetery (minimum = 1621) and at most at the closure year of the cemetery (Notre-Dame = 1621–1796, Saint-Antoine = 1724–1855, Pointe-aux-Trembles = 1634–1843, and Sainte-Marie = 1673–1878). We considered parents that had only married daughters as well since they might have had unmarried sons, absent from BALSAC but listed in the IMPQ register.

Finally, we examined candidates starting with those buried in the investigated cemetery based on the IMPQ register and excluding those that were buried elsewhere. We constructed pedigrees with the R package Kinship to investigate the genealogical context of candidates and to identify genealogical errors. Given the current modern sample, genealogical error rates resulting from false paternity or undeclared adoption was estimated to be 0.0038 for the mtDNA and 0.0080 for the Y chromosome (Doyon et al., in preparation). Previous studies of the BALSAC genealogies by Heyer et al. (1997) and Jomphe (2011) reported similar values (<0.75% rate), among the lowest reported worldwide (Anderson, 2006). The life history of candidates was examined when the information was available in the IMPQ register (e.g., dates and places of birth, marriage, and death of candidates and of their parents and children). The final stage involved repeating the above procedure by considering modern haplotypes differing by one or two mutations relative to the historical haplotypes.

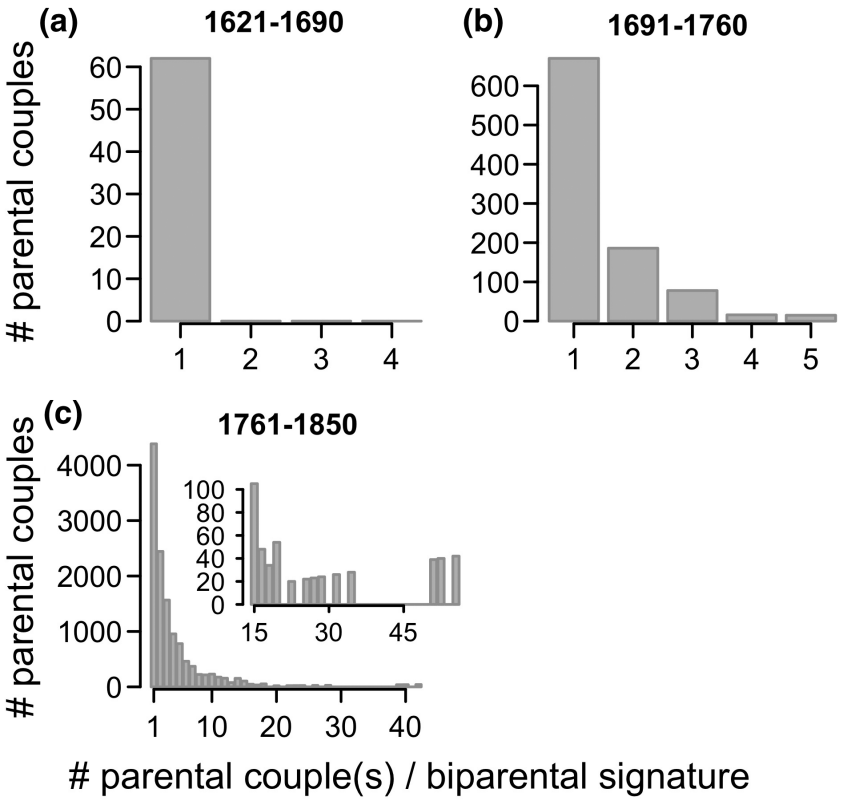

FIGURE 3 The number of parental couples who shared the same biparental signature (combined paternal-maternal genealogies) and for whom the first children to marry did so in (a) 1621–1690, (b) 1691–1760, and (c) 1761–1850. Parents belonged to the uniparental genealogies of genotyped modern individuals, that is, their sons are currently identifiable. Bars too small to be seen in c are enlarged in the inset

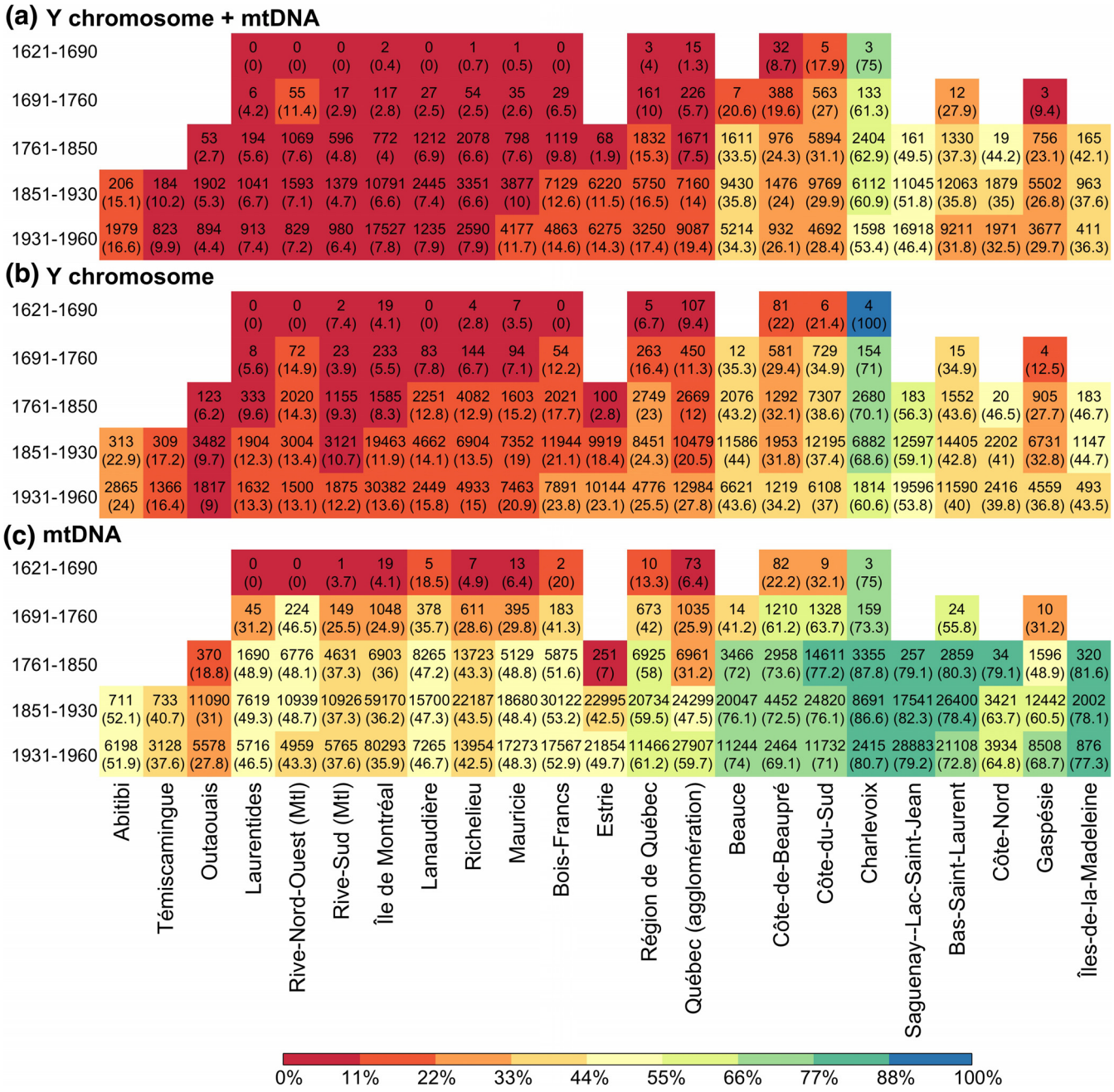

FIGURE 4 Spatiotemporal genetic coverage of currently identifiable men in BALSAC genealogies. The number (color-coded percent in parentheses) of men related to genotyped modern individuals through (a) both their maternal and paternal lineages, (b) their paternal lineages, and (c) their maternal lineages are indicated as a function of their marriage year and region. Blank cells correspond to years before the foundation of a region. From left to right, regions are ordered from west to east

2.5 | Coverage and diversity analyses

We assessed what periods and what regions, given our contemporary mtDNA and Y-chromosome database, provided a maximum of men that could be identified to guide future decision making in the choice of sampling sites. We divided the data into spatiotemporal slices: five temporal periods (e.g., Figure 4) and 23 geographical regions (see Figure 2). We used 1931–1960 as a proxy of the modern population. We divided the previous years into four periods centered on the passage in 1760 from the French to the British Regime. The fourth period starts at 1850 since graves in historical Québec cemeteries were typically anonymous until the mid-19th century (Oliver-Lloyd, 2008). We refer to Québec City and Beauce regions as "eastern" regions, along with those farther east, and the remaining Québec regions as "western" regions (red line on Figure 2).

We also studied the diversity of biparental signatures, that is, the combinations of maternal and paternal genealogies created by the marital union of two parents. We generated frequency spectra to

evaluate the discriminatory power of Y-chromosome and mtDNA markers for the identification of men in the historical population of Québec. We especially focused on identifiable men, that is, men whose paternal and maternal genealogies persisted up until the present. They include *currently identifiable historical men*, that is, men that are related to the genotyped modern individuals through both of their parental genealogies.

3 | RESULTS

3.1 | The integrity of ancient DNA

The amount of human DNA extracted from teeth ranged from 0.006 to 9.1 ng (0.12–182.73 pg/μl; Table 2). We authenticated these samples by examining the postmortem DNA damage due to the deamination of cytosine residues that leads to the incorporation of thymines at cytosine sites (Briggs et al., 2007). We detected increased C → T misincorporation rates in mitochondrial sequencing reads relative to modern samples (Figures S1 and S2). Likewise, a decrease in the amplification success of longer Y-chromosome STRs amplicons was consistent with postmortem fragmentation of Y-chromosome DNA (Figure S3).

TABLE 2 Mass and concentration of ancient DNA extracts

	Human DNA mass (ng)		Human DNA conc. (pg/μl)	
Specimen	Extract 1	Extract 2	Extract 1	Extract 2
12D-S1	0.070	0.250	1.40	5.00
4K-S2	0.006	0.650	0.12	13.00
11D-S1	0.746	1.600	14.92	32.00
20A-S12	0.197	Undetectable	3.95	Undetectable
7A9-S16	9.136	NA[a]	182.73	NA[a]
2B7	NA[b]	NA[b]	NA[b]	NA[b]

[a]DNA was not extracted a second time from specimen 7A9-S16.
[b]DNA from specimen 2B7 was extracted from multiple sources several times before the other samples and was either not detected using the Quantifiler™ method (possibly due to the presence of PCR inhitors) or quantified by fluorometry (i.e., quantification includes microbial DNA) and ranged between 0.3 and 0.4 ng/μl.

3.2 | Historical genetic profiles and genealogical candidates

The number of successfully typed Y-STR loci per historical specimen varied between 6 (specimen 20A-S12) and 12 (specimen 7A9-S16; Table 3). The haplotypes of specimens 12D-S1, 4K-S2, 11D-S1, 7A-S16, and 20A-S12 matched the haplotypes of 1, 2, 8, 14, and 62 individuals in our modern sample, respectively (Table 3). Four modern haplotypes were closest to 2B7's haplotype and each differed by one mutation at either DYS437, DYS390, or DYS385a.

The mitochondrial profile of specimen 12D-S1 matched 24 genotyped modern individuals. Specimens 4K-S2, 20A-S12, and 2B7, sharing the same mitochondrial haplotype, matched 85 genotyped modern individuals (Table 4). The mitochondrial sequences from specimens 11D-S1 and 7A9-S16 differed by two mutations from the haplotypes in our collection.

For historical specimens with both mitochondrial and Y-chromosome profiles matching modern haplotypes (12D-S1, 4K-S2, 20A-S12), we searched for parental couples who shared these haplotypes and that had children married at periods when the respective cemeteries were operational. Sibships (sibling groups) of such parents carry a biparental genetic signature allowing their unambiguous identification as sibships, and further individualization can be achieved through the analysis of archeological and complementary data. Following this route, we found candidate parents for specimen 20A-S12, exhumed from the Saint-Antoine cemetery located in the Notre-Dame de Montréal parish. For this specimen, a total of 201 parents belonged to matching genealogies and had children married between 1724 and 1855 (Table 5). Of these, 159 parents had 379 sons and 42 parental couples had daughters only. These parents represented 122 distinct combinations of paternal-maternal genealogies.

One of the matching couples had one of their sons inhumed in Notre-Dame parish according to the IMPQ register. This son was born in Pierrefonds on Montréal Island in 1775 and was married in 1803 in Saint-Eustache (Northwest of Montréal). He died in 1833 at the age of 58, a plausible age at death given the available bone material of 20A-S12. His parents had another son, but the date and place of death of this second son is unknown. The man born in 1775 is, therefore, the strongest candidate in our data.

A second matching couple had three sons buried in this parish. However, these sons all died under the age of 3, too young to be

TABLE 3 Y-STR alleles of historical human remains

Specimen	DYS 391	DYS 389_I	DYS 389_II	DYS 439	DYS 438	DYS 437	DYS 19	DYS 392	DYS 393	DYS 390	DYS 385a	DYS 385b	# Matching modern individuals
12D-S1	11	13	29	?	?	?	?	12	14	24	11	?	1
4K-S2	10	14	?	?	12	?	14	13	13	24	11	?	2
11D-S1	11	13	?	12	?	14	14	?	13	?	11	14	8
20A-S12	11	13	?	12	?	15	?	?	13	?	11	?	62
7A9-S16	11	13	29	12	12	15	14	13	13	24	11	14	14
2B7	?	13	29	12	12	14	14	?	13	24	12	?	0

TABLE 4 Mitochondrial data from historical human remains

Specimen	Mapped reads	% Mapped reads	% Reads assignable to historical sample[a]	Genetic profile (mutations relative to rCRS)	# Matching modern individuals
12D-S1	54,260	98.42%	95.9%	A73G, A263G, T16224C, T16311C	24
4K-S2	73,516	99.82%	94.8%	A263G	85
11D-S1	45,171	99.69%	88.6%	A73G, A263G, T16092C, A16162G	0
20A-S12	36,895	99.16%	90.1%	A263G	85
7A9-S16	18,625	98.73%	100%	T72C, T195C, T16298C	0
2B7	23,066	99.74%	100%	A263G	85

[a]Percentage of reads with coverage >100× that encoded mutations expected from the historical specimen with and without misincorporations of C → T or G → A attributable to postmortem damage.

TABLE 5 Numbers of uniparental genealogies in BALSAC, parents, sons of these parents and daughter-only sibships with haplotypes perfectly matching the historical ones (Y + mtDNA, both Y and mtDNA haplotypes; Y, Y haplotypes only; mtDNA, mtDNA haplotypes only)

	# Genealogies[a]	# Parents	# Married sons	# Daughter-only sibships
Specimen 12D-S1				
Y + mtDNA	0	0	0	0
Y	1	26	45	9
mtDNA	5	680	1,200	153
Specimen 4K-S2				
Y + mtDNA	0	0	0	0
Y	2	3	5	0
mtDNA	39	1,806	3,012	495
Specimen 11D-S1				
Y + mtDNA	0	0	0	0
Y	6	192	359	35
mtDNA	0	0	0	0
Specimen 20A-S12				
Y + mtDNA	122	201	379	42
Y	49	3,881	7,373	841
mtDNA	39	7,619	13,626	1,708
Specimen 7A9-S16				
Y + mtDNA	0	0	0	0
Y	12	542	1,011	115
mtDNA	0	0	0	0
Specimen 2B7				
Y + mtDNA	0	0	0	0
Y	0	0	0	0
mtDNA	39	11,372	20,439	2521

[a]Numbers of paternal genealogies with matching Y haplotypes, number of maternal genealogies with matching mtDNA haplotypes, and number of distinct combinations of paternal + maternal genealogies with matching Y + mtDNA haplotypes.

retained as candidates (Table 1). Among the other parents belonging to genealogies matching the profiles of 20A-S12, 54 had 91 sons that were buried elsewhere than in the Notre-Dame parish. These sons can thus be excluded. However, the burial place being unknown for the remaining 286 sons (children of 135 parents including 32 parents that had sons buried elsewhere), they are all plausible candidates. That so many candidates matched the haplotypes of specimen 20A-S12 is due to the low number of typed loci in his Y-chromosome profile. A total of 22 distinct modern Y haplotypes (found in genealogies established in Québec before 1850) matched the alleles at the six typed loci in 20A-S12's profile. Additional genetic data from this historical specimen are required to identify which one of the modern Y haplotypes, if any, the individual 20A-S12 has carried.

TABLE 6 Number of people buried in studied parishes (ND, Notre-Dame de Montréal; PaT, Saint-Enfant-Jésus-de-la-Pointe-aux-Trembles; SMdB, SainteMarie-de-Beauce) based on the IMPQ database

Parish	Total # deaths[a]	# Men[b] (% deaths)	# Y[c] (% men)	# mtDNA[c] (% men)	# Y + mtDNA[c] (% men)
ND	47,924	4,092 (8.5)	281 (6.9)	1,283 (31.4)	136 (3.3)
PaT	2,926	311 (10.6)	24 (7.7)	88 (28.3)	10 (3.2)
SMdB	4,696	499 (10.6)	222 (44.5)	319 (63.9)	145 (29.1)

[a]Number of people buried in studied parishes based on the IMPQ database which covers the years 1691-1849 for ND, 1709-1843 for PaT, and 1748-1849 for SMdB.
[b]Number of married men in BALSAC with burial acts in the IMPQ database. Percentage values in parentheses relate to the proportion of individuals buried in the parish.
[c]Number of married men in BALSAC, buried in studied parishes and related to genotyped individuals through their paternal lineage only (# Y), their maternal lineage only (# mtDNA) and both their parental lineages (# Y + mtDNA). Percentage values in parentheses relate to the proportion of men buried in the parish.

3.3 | Current population coverage and discriminatory power

Before 1850, 311 married men were buried in Saint-Enfant-Jésus-de-la-Pointe-aux-Trembles parish, 499 in Sainte-Marie-de-Beauce and 4,092 in Notre-Dame de Montréal (Table 6). Between 3 and 29% of these men are currently identifiable since they were related through both their paternal Y-chromosome and their maternal mtDNA lines to modern genotyped individuals (Figure 1).

In Québec, 13,730 parental couples with children married before 1850 were genealogically linked through their paternal and maternal lines to the genotyped modern individuals. These couples were the parents of 12% of the men married before 1850 (n = 26,672 men). They represented 7,091 distinct biparental signatures, that is, distinct combinations of maternal and paternal genealogies. Most of the parental couples that had children married before 1760 carried distinct biparental signatures: 100% of them in 1621–1690 and 69% of them in 1691–1760 (Figure 3). Between 1761 and 1850, 35% of parental couples had unique biparental signatures and up to 43 couples shared the same signature (Figure 3). However, geographical partitioning of the data, based on the place of marriage of the first married children, indicated that some regions (e.g., Charlevoix, Côte-du-Sud, Bas-Saint-Laurent, and Région de Québec) were enriched in couples sharing the same biparental signature (Figure S4). Outside these regions, 56% of couples had unique biparental signatures, and no more than 10 shared the same biparental signature.

Considering all the men married before 1930, 326,642 (15%) are currently identifiable, given our limited genetic database. Their distribution varied geographically and with time (Figure 4). The best coverage is observed in regions east of Québec City, where up to 75% of men are currently identifiable. In western regions, the coverage never exceeds 17% (due to sampling bias in our database).

4 | DISCUSSION

We developed a framework to identify historical remains within a family tree of Québec individuals. This was made possible given the quasi-complete genealogical tree of the individuals married in Québec from the population foundation in the 17th century down to the present.

One of the main factors restricting the identification of human remains in this study was the limited coverage of the available genetic modern data. Three of the specimens (11D-S1, 7A9-S12, and 2B7) had only one of their uniparental haplotypes matching genealogies present when their respective cemeteries were operational (Table 5). Two specimens (12D-S1, 4K-S2) had both their Y-chromosome and mitochondrial haplotypes matching, but their combinations were absent from currently genotyped genealogies. This shows that not all the genealogies of the historical specimens were represented in our modern sample. In particular, we had no biparental genetic signatures for 71–97% of men buried in the studied cemeteries, especially for those located on Montréal Island (Table 6). The coverage of the regions west of Québec city had lower proportions of identifiable men based on the current haplotype collection (Figure 4). Future sampling efforts should thus target western regions, especially those located on and around Montréal Island. In addition, since the same mitochondrial HVI-II haplotypes were carried by more than one founder, their differentiation into distinct ones would require more extended sequencing. This also applies to the Y-chromosome haplotypes, especially given the low number of loci considered (e.g., using the same set of Y-STRs loci, 19% of volunteers shared the same haplotype in the United States; Hanson & Ballantyne, 2007) and that not all markers were successfully typed in the historical samples. Greater discriminatory power is to be achieved by sequencing the entire mitochondrial genome, and by genotyping additional Y-STR loci. The ongoing sequencing projects of large Québec cohorts (e.g., CARTaGENE, https://www.cartagene.qc.ca/; Awadalla et al., 2013), including the participants' genealogical information, are currently filling the gaps in genetic coverage.

In principle, immigrant men cannot be identified since most often their maternal signatures are unknown. Due to the maternal transmission of mtDNA, human remains identification using this genetic marker requires that a matrilineal relative of the anonymous subject has left matrilineal descendants in the current population. As most men usually immigrated alone, during the first decades of the colony, the fraction of identifiable men was lower (14% before 1690), notably because the matrilineal ancestors of these men did not establish

themselves in Québec (80% of men before 1690; Figure S5). Between 1691 and 1760, while no matrilineal relatives are known in BALSAC for 23% of men, 70% of them are identifiable. After 1761, the percentage of men without known maternal relatives drops under 14%, and more than 82% are identifiable. Interestingly, isotope analysis ($\delta^{18}O$) of specimen 11D-S1 indicated that this individual might have immigrated from warmer possibly coastal regions (Vigeant, 2012). While his mitochondrial haplotype was absent from our modern sample, it is known in the Basque population (Behar et al., 2012), information that can be used to identify these remains. Hypothetically, individual 11D-S1 immigrated to New France alone carrying a maternal signature unknown to BALSAC, precluding his identification.

Nevertheless, provided an increase in the genetic coverage (ongoing sequencing effort—see below), this method holds great promise. Considering the individuals in BALSAC that are related through both their uniparental lineages to individuals in the contemporary population, 178,435 (87%) men married before 1850 are potentially identifiable. When all the men married in Québec between 1621 and 1930 are considered, this number increases to 1,632,877 (76%) men. Sorting them according to the time and place of their marriage, we determined that >90% of men are identifiable in most regions after 1760 (Figure S6). Before 1760, the percentage of parental couples with unique biparental signature ranges from 75% to 100%, and none of the regions have more than five parental couples sharing the same biparental signature (Figures S7). Between 1761 and 1850, 77% to 100% of parental couples in western regions have unique biparental signatures, whereas in eastern regions 37–100% of couples have unique biparental signatures (Figure S8). The lower proportions in eastern regions, such as Charlevoix and Côte-du-Sud, correlate with their lower genetic diversity. Outside these regions, no more than 15 couples share the same biparental signature (Figure S8). The discriminatory power increases further at the parish level. According to the IMPQ register, 100%, 85% and 63% of fraternal sibships buried in 1621–1690, 1691–1760, and 1761–1850, respectively, have biparental signatures that are unique in their parish. None of the fraternal sibships buried in 1621–1690, a maximum of two fraternal sibships in 1691–1760 and a maximum of eight fraternal sibships in 1761–1850 share the same signature as other fraternal sibships in their parish. In other words, from several thousands of individuals buried in the most populous cemeteries, the number of fraternal sibships to investigate through further characterizations (e.g., archeological or historical searches) could be reduced to a unique sibship for the majority of men, and to a dozen sibships in worst-case scenarios.

Several applications emerge from the method developed in this study. For example, it may serve to identify anonymous human remains excavated from historical cemeteries to reinhume them in a signed grave. It may also help the forensic identification of French Canadian soldiers who died overseas during the First World War. The Casualty Identification Program has the mandate to repatriate the bodies of missing Canadians dead on the war fields prior to 1970 (estimated to more than 27,000), to identify and bury them with a proper grave. The method presented here could help to target plausible candidates.

A more fundamental objective would be to aid in the reconstruction of the historical genome of Québec, starting from contemporary sequences, back in time along the corresponding genealogical lines. Without ancient DNA data, this task depends upon the number of related modern individuals that can be sampled to maximize the number of shared ancestral segments (Bryca et al., 2010; Gravel et al., 2013; Moreno-Estrada et al., 2013). A pair of related individuals are expected to share only a fraction of $2^{-(m-1)}$ of their genomes, where m denotes the number of meiosis separating them (Browning & Browning, 2012). With 12 meioses separating fifth cousins, they are expected to share 0.05% of their genomes. In addition, due to recombination, the expected mean length of a shared segment is 100 cM/m, or ~8.3 cM in this case (~0.25% of the genome). Thus, for one of the fifth cousins sharing such a segment, four will carry none on average (Browning & Browning, 2012). Therefore, accessing genomic information of well-identified genealogical ancestors through ancient DNA sequencing would greatly help in calibrating such reconstruction by providing exact ancestral genomic information a few generations in the past.

While only applicable at a historical depth in populations with known genealogies, the identification of ancestral individuals and the sequencing of their genomes would provide a source of invaluable information to study human genetics on the micro-evolutionary scale (Larmuseau, Van Geystelen, van Oven, & Decorte, 2013; Milot & Stearns, in press; Peischl et al., 2018). This is important because we need to confront real data with theoretical predictions that are often restricted to formally tractable simple population models (Gravel & Steel, 2015). The sequencing of historical Québec genomes would further shed light on the early distinctions between regional populations, the selection factors acting on the front of migration waves and the genetic circumstances that advantaged certain regions rather than others (Bherer et al., 2011; Moreau, Bhérer, et al., 2011).

In conclusion, using genealogical information, we identified several plausible male candidates matching one historical specimen, although the available molecular data limited our capacity to narrow down the list to a single person. Greater genetic coverage of paternal and maternal lineages and more genetic markers are required to increase the chance to meet the right mitochondrial/Y-chromosome combination and to increase confidence in the match. However, our study demonstrates that identifying historical remains by coupling genealogical and genetic information is achievable and that it can reduce the pool of candidates from several hundreds of thousands of individuals to a more manageable set of dozens of candidates. The method presented here might at last systematically connect bio-archeological evidence of a founder population to its extensive genealogical record.

ACKNOWLEDGMENTS

We thank the community of Sainte-Marie-de-Beauce (Conseil de la Fabrique) and Ville de Montréal (François Bélanger, Réserves des Collections Archéologiques de la Ville de Montréal), who allowed us to analyze and sample the burials from the historical cemeteries before their future reburial. We also thank the participants for providing

samples for genetic testing and family information for genealogical reconstruction. The participants were recruited in the *Groupe de référence en épidémiologie génétique et en génétique des populations humaines au Québec*, as well as in the *Saguenay–Lac-Saint-Jean cohort*. Computations were made on the supercomputers Cedar and Mammouth from Simon Fraser University and Université de Sherbrooke, respectively, managed by Compute Canada and Calcul Québec. The operation of these supercomputers is funded by the Canada Foundation for Innovation, the ministère de l'Économie et de l'Innovation du Québec and the Fonds de recherche du Québec-Nature et technologies. The shapefiles used to draw the map of Québec (Figure 2) was kindly provided by the Centre interuniversitaire d'études québécoises. This research was supported by the Fonds de recherche du Québec—Santé via the Réseau de médecine génétique appliquée (D.L. and H.V.), the National Sciences and Engineering Research Council of Canada (D.L.), the Fonds de recherche du Québec—Société et Culture (B.L., D.L., I.R., research group ArchéoScience-ArchéoSociale directed by Adrian Burke), and the Fondo Europeo de Desarrollo Regional—Ministerio De Economia Y Competitividad of Spain (CLF; grant number PGC2018-095931-B-100).

DATA AVAILABILITY STATEMENT

The data on the genealogically anchored samples can be found in the Québec Reference Sample database (http://www.quebecgenpop.ca/). Sequence reads were deposited in the NCBI Sequence Read Archive under the accession number PRJNA578852. Genealogical and genetic data were handled in R version 3.5.1 (R Core Team, 2013). R and Python scripts used are available on demand.

ORCID

Tommy Harding https://orcid.org/0000-0002-7823-2755

REFERENCES

Albert, A. M., & Maples, W. R. (1995). Stages of epiphyseal union for thoracic and lumbar vertebral centra as a method of age determination for teenage and young adult skeletons. *Journal of Forensic Sciences*, *40*(4), 623–633.

Anderson, K. G. (2006). How well does paternity confidence match actual paternity? Evidence from worldwide nonpaternity rates. *Current Anthropology*, *47*(3), 513–520. https://doi.org/10.1086/504167

Andrews, R. M., Kubacka, I., Chinnery, P. F., Lightowlers, R. N., Turnbull, D. M., & Howell, N. (1999). Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nature Genetics*, *23*(2), 147. https://doi.org/10.1038/13779

Arkéos. (2008). *Interventions archéologiques sur le site de la première église Notre-Dame (BjFj-112), Vieux-Montréal*. Internal archaeological report. Service de la mise en valeur du territoire et du patrimoine.

Awadalla, P., Boileau, C., Payette, Y., Idaghdour, Y., Goulet, J.-P., Knoppers, B., … Laberge, C. (2013). Cohort profile of the CARTaGENE study: Quebec's population-based biobank for public health and personalized genomics. *International Journal of Epidemiology*, *42*(5), 1285–1299. https://doi.org/10.1093/ije/dys160

Behar, D., Harmant, C., Manry, J., van Oven, M., Haak, W., Martinez-Cruz, B., … The Genographic Consortium. (2012). The Basque paradigm: Genetic evidence of a maternal continuity in the Franco-Cantabrian region since pre-Neolithic times. *American Journal of Human Genetics*, *90*(3), 486–493. https://doi.org/10.1016/j.ajhg.2012.01.002

Benschop, C. C., van der Beek, C. P., Meiland, H. C., van Gorp, A. G., Westen, A. A., & Sijen, T. (2011). Low template STR typing: Effect of replicate number and consensus method on genotyping reliability and DNA database search results. *Forensic Science International: Genetics*, *5*(4), 316–328. https://doi.org/10.1016/j.fsigen.2010.06.006

Bherer, C., Labuda, D., Roy-Gagnon, M. H., Houde, L., Tremblay, M., & Vézina, H. (2011). Admixed ancestry and stratification of Quebec regional populations. *American Journal of Physical Anthropology*, *144*(3), 432–441. https://doi.org/10.1002/ajpa.21424

Bogdanowicz, W., Allen, M., Branicki, W., Lembring, M., Gajewska, M., & Kupiec, T. (2009). Genetic identification of putative remains of the famous astronomer Nicolaus Copernicus. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(30), 12279–12282. https://doi.org/10.1073/pnas.0901848106

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Briggs, A. W., Stenzel, U., Johnson, P. L. F., Green, R. E., Kelso, J., Prüfer, K., … Pääbo, S. (2007). Patterns of damage in genomic DNA sequences from a Neandertal. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(37), 14616–14621. https://doi.org/10.1073/pnas.0704665104

Brooks, S., & Suchey, J. M. (1990). Skeletal age determination based on the os pubis: A comparison of the Ascadi-Nemeskéri and Suchey-Brooks methods. *Human Evolution*, *5*(3), 227–238. https://doi.org/10.1007/bf02437238

Browning, S. R., & Browning, B. L. (2012). Identity by descent between distant relatives: Detection and applications. *Annual Review of Genetics*, *46*, 617–633. https://doi.org/10.1146/annurev-genet-110711-155534

Bryca, K., Velezb, C., Karafetc, T., Moreno-Estradaa, A., Reynoldsa, A., Autona, A., … Ostrer, H. (2010). Genome-wide patterns of population structure andadmixture among Hispanic/Latino populations. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 8964–8961. https://doi.org/10.1073/pnas.0914618107

Buikstra, J. E., & Ubelaker, D. H. (1994). *Standards for data collection from human skeletal remains: Proceedings of a seminar at the Field Museum of natural history*. Fayetteville, USA: Arkansas Archaeological Survey.

Butler, J. M., & Hill, C. R. (2010). Scientific issues with analysis of low amounts of DNA. *Profiles in DNA*, *13*(1). Accessed on January 15, 2019. https://www.promega.in/resources/profiles-in-dna/2010/scientific-issues-with-analysis-of-low-amounts-of-dna

Calafell, F., & Larmuseau, M. H. D. (2017). The Y chromosome as the most popular marker in genetic genealogy benefits interdisciplinary research. *Human Genetics*, *136*(5), 559–573. https://doi.org/10.1007/s00439-016-1740-0

Cann, R. L., Stoneking, M., & Wilson, A. C. (1987). Mitochondrial DNA and human evolution. *Nature*, *325*(6,099), 31–36. https://doi.org/10.1038/325031a0

Casanova, M., Leroy, P., Boucekkine, C., Weissenbach, J., Bishop, C., Fellous, M., … Siniscalco, M. (1985). A human Y-linked DNA polymorphism and its potential for estimating genetic and evolutionary distance. *Science*, *230*(4,732), 1403–1406. https://doi.org/10.1126/science.2999986

Cavalli-Sforza, L. L., Menozzi, P., & Piazza, A. (1994). *The history and geography of human genes*. Princeton, NJ: Princeton University Press.

Charlier, P., Olalde, I., Sole, N., Ramirez, O., Babelon, J. P., Galland, B., … Lalueza-Fox, C. (2013). Genetic comparison of the head of Henri IV and the presumptive blood from Louis XVI (both kings of France). *Forensic Science International*, *226*(1–3), 38–40. https://doi.org/10.1016/j.forsciint.2012.11.018

Coble, M. D., Loreille, O. M., Wadhams, M. J., Edson, S. M., Maynard, K., Meyer, C. E., … Finelli, L. N. (2009). Mystery solved: The identification of the two missing Romanov children using DNA analysis. *PLoS One*, *4* (3), e4838. https://doi.org/10.1371/journal.pone.0004838

R Core Team. (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Doyon, A., Labuda, D., Moreau, C., Larmuseau, M., Vézina, H., Noël, Y., … Milot, E. (in preparation). Population-wide assessment of spatiotemporal variation in DNA lineage marker frequencies and implications for their use in forensic science.

Dudar, J. C., Waye, J. S., & Saunders, S. R. (2003). Determination of a kinship system using ancient DNA, mortuary practice, and historic records in an upper Canadian pioneer cemetery. *International Journal of Osteoarchaeology*, *13*(4), 232–246. https://doi.org/10.1002/oa.680

Ethnoscop. (2006). *Site du premier cimetière de Sainte-Marie, CcEs-1. Inventaire et fouilles archéologiques* 2003-2004. Internal archaeological report. Ministère des transports du Québec, Direction de Chaudière-Appalaches.

Ethnoscop. (2016a). *Interventions archéologiques dans le cadre du projet de construction de la Maison du citoyen à Pointe-aux-Trembles, 2014, BjFi-17*. Internal archaeological report. Rivière-des-Prairies, Pointe-aux-Trembles et Ville de Montréal.

Ethnoscop. (2016b). *Projet de réaménagement de la place du Canada. Site archéologique du cimetière Saint-Antoine (1799–1854), BiFj-37, Montréal. Fouille et supervision archéologique (2014)*. Internal archaeological report. Ville de Montréal.

Gamba, C., Jones, E. R., Teasdale, M. D., McLaughlin, R. L., Gonzalez-Fortes, G., Mattiangeli, V., … Pinhasi, R. (2014). Genome flux and stasis in a five millennium transect of European prehistory. *Nature Communications*, *5*, 5257. https://doi.org/10.1038/ncomms6257

Garrison, E., & Marth, G. (2012). Haplotype-Based Variant Detection from Short-Read Sequencing. *arXiv*, 1207.3907.

Giles, R. E., Blanc, H., Cann, H. M., & Wallace, D. C. (1980). Maternal inheritance of human mitochondrial DNA. *Proceedings of the National Academy of Sciences of the United States of America*, *77*(11), 6715–6719. https://doi.org/10.1073/pnas.77.11.6715

Gill, P., Ivanov, P. L., Kimpton, C., Piercy, R., Benson, N., Tully, G., … Sullivan, K. (1994). Identification of the remains of the Romanov family by DNA analysis. *Nature Genetics*, *6*(2), 130–135. https://doi.org/10.1038/ng0294-130

Gravel, S., & Steel, M. (2015). The existence and abundance of ghost ancestors in biparental populations. *Theoretical Population Biology*, *101*, 47–53. https://doi.org/10.1016/j.tpb.2015.02.002

Gravel, S., Zakharia, F., Moreno-Estrada, A., Byrnes, J. K., Muzzio, M., Rodriguez-Flores, J. L., … Bustamante, C. D. (2013). Reconstructing native american migrations from whole-genome and whole-exome data. *PLoS Genetics*, *9*(12), e1004023. https://doi.org/10.1371/journal.pgen.1004023

Hanson, E. K., & Ballantyne, J. (2007). An ultra-high discrimination Y chromosome short tandem repeat multiplex DNA typing system. *PLoS One*, *2*(8), e688. https://doi.org/10.1371/journal.pone.0000688

Hartnett, K. M. (2010). Analysis of age-at-death estimation using data from a new, modern autopsy sample— Part II: Sternal end of the fourth rib. *Journal of Forensic Sciences*, *55*(5), 1152–1156. https://doi.org/10.1111/j.1556-4029.2010.01415.x

Heyer, E., Puymirat, J., Dieltjes, P., Bakker, E., & de Knijff, P. (1997). Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees. *Human Molecular Genetics*, *6*(5), 799–803. https://doi.org/10.1093/hmg/6.5.799

Heyer, E., Zietkiewicz, E., Rochowski, A., Yotova, V., Puymirat, J., & Labuda, D. (2001). Phylogenetic and familial estimates of mitochondrial substitution rates: Study of control region mutations in deep-rooting pedigrees. *American Journal of Human Genetics*, *69*(5), 1113–1126. https://doi.org/10.1086/324024

Hirschfeld, L., & Hirschfeld, H. (1919). Serological differences between the blood of different races. *The Lancet*, *194*(5,016), 675–679. https://doi.org/10.1016/S0140-6736(01)48686-7

Jobling, M. A., & Tyler-Smith, C. (1995). Fathers and sons: The Y chromosome and human evolution. *Trends in Genetics*, *11*(11), 449–456. https://doi.org/10.1016/S0168-9525(00)89144-1

Jomphe, M. (2011). *Validation des généalogies reconstituées à BALSAC à partir de données génétiques (Document I-C-243)*. BALSAC, 13 pages.

Katzenberg, M. A., Oetelaar, G., Oetelaar, J., Fitzgerald, C., Yang, D., & Saunders, S. R. (2005). Identification of historical human skeletal remains: A case study using skeletal and dental age, history and DNA. *International Journal of Osteoarchaeology*, *15*(1), 61–72. https://doi.org/10.1002/oa.748

King, T. E., Fortes, G. G., Balaresque, P., Thomas, M. G., Balding, D., Delser, P. M., … Schurer, K. (2014). Identification of the remains of King Richard III. *Nature Communications*, *5*, 5631. https://doi.org/10.1038/ncomms6631

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, *10*, R25. https://doi.org/10.1186/gb-2009-10-3-r25

Larmuseau, M. H. D., Delorme, P., Germain, P., Vanderheyden, N., Gilissen, A., Van Geystelen, A., … Decorte, R. (2014). Genetic genealogy reveals true Y haplogroup of House of Bourbon contradicting recent identification of the presumed remains of two French Kings. *European Journal of Human Genetics*, *22*(5), 681–687. https://doi.org/10.1038/ejhg.2013.211

Larmuseau, M. H. D., Van Geystelen, A., van Oven, M., & Decorte, R. (2013). Genetic genealogy comes of age: Perspectives on the use of deep-rooted pedigrees in human population genetics. *American Journal of Physical Anthropology*, *150*(4), 505–511. https://doi.org/10.1002/ajpa.22233

Lovejoy, C. O. (1985). Dental wear in the Libben population: Its functional pattern and role in the determination of adult skeletal age at death. *American Journal of Physical Anthropology*, *68*(1), 47–56. https://doi.org/10.1002/ajpa.1330680105

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal*, *17*(1), 10–12. https://doi.org/10.14806/ej.17.1.200

Meindl, R. S., & Lovejoy, C. O. (1985). Ectocranial suture closure: A revised method for the determination of skeletal age at death based on the lateral-anterior sutures. *American Journal of Physical Anthropology*, *68* (1), 57–66. https://doi.org/10.1002/ajpa.1330680106

Milot, E., & Stearns, S. C. (in press). Selection on humans in cities. In M. Szulkin, J. Munshi-South, & A. Charmantier (Eds.), *2020, Urban evolutionary biology*. Oxford: Oxford University Press.

Moreau, C., Bhérer, C., Vézina, H., Jomphe, M., Labuda, D., & Excoffier, L. (2011). Deep human genealogies reveal a selective advantage to be on an expanding wave front. *Science*, *334*(6,059), 1148–1150. https://doi.org/10.1126/science.1212880

Moreau, C., Vezina, H., Jomphe, M., Lavoie, E. M., Roy-Gagnon, M. H., & Labuda, D. (2011). When genetics and genealogies tell different stories-maternal lineages in Gaspesia. *Annals of Human Genetics*, *75*(2), 247–254. https://doi.org/10.1111/j.1469-1809.2010.00617.x

Moreau, C., Vezina, H., Yotova, V., Hamon, R., de Knijff, P., Sinnett, D., & Labuda, D. (2009). Genetic heterogeneity in regional populations of Quebec—Parental lineages in the Gaspe Peninsula. *American Journal of Physical Anthropology*, *139*(4), 512–522. https://doi.org/10.1002/ajpa.21012

Moreno-Estrada, A., Gravel, S., Zakharia, F., McCauley, J. L., Byrnes, J. K., Gignoux, C. R., … Bustamante, C. D. (2013). Reconstructing the population genetic history of the Caribbean. *PLoS Genetics*, *9*(11), e1003925. https://doi.org/10.1371/journal.pgen.1003925

Murail, P., Brůžek, J., Houët, F., & Cunha, E. (2005). DSP: A tool for probabilistic sex diagnosis using worldwide variability in hip-bone

measurements. *Bulletins et Mémoires de la Société d'Anthropologie de Paris*, *17*, 167–176.

Nelson, D., Moreau, C., Vriendt, M. d., Zeng, Y., Preuss, C., Vézina, H., … Gravel, S. (2018). Inferring transmission histories of rare alleles in population-scale genealogies. *American Journal of Human Genetics*, *103* (6), 893–906. https://doi.org/10.1016/j.ajhg.2018.10.017

Oliver-Lloyd, V. (2008). *Le patrimoine archéologique des cimetières euroquébécois, étude produite dans le cadre de la participation du Québec au Répertoire canadien des lieux patrimoniaux, volet archéologique*. Québec, Ministère de la Culture, des Communications et de la Condition féminine, Études thématiques, Secteurs d'intervention: patrimoine–archéologie.

Peischl, S., Dupanloup, I., Foucal, A., Jomphe, M., Bruat, V., Grenier, J. C., … Excoffier, L. (2018). Relaxed selection during a recent human expansion. *Genetics*, *208*(2), 763–777. https://doi.org/10.1534/genetics.117.300551

Rogaev, E. I., Grigorenko, A. P., Moliaka, Y. K., Faskhutdinova, G., Goltsov, A., Lahti, A., … Morozova, I. (2009). Genomic identification in the historical case of the Nicholas II royal family. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(13), 5258–5263. https://doi.org/10.1073/pnas.0811190106

Roy-Gagnon, M. H., Moreau, C., Bherer, C., St-Onge, P., Sinnett, D., Laprise, C., … Labuda, D. (2011). Genomic and genealogical investigation of the French Canadian founder population structure. *Human Genetics*, *129*(5), 521–531. https://doi.org/10.1007/s00439-010-0945-x

Schmitt, A. (2005). Une nouvelle méthode pour estimer l'âge au décès des adultes à partir de la surface sacro-pelvienne iliaque. *Bulletins et Mémoires de la Société d'Anthropologie de Paris*, *17*(1–2), 51–73.

Underhill, P. A., & Kivisild, T. (2007). Use of Y chromosome and mitochondrial DNA population structure in tracing human migrations. *Annual Review of Genetics*, *41*, 539–564. https://doi.org/10.1146/annurev.genet.41.110306.130407

Vigeant, J. (2012). *Immigration et alimentation à Montréal aux XVIIe et XVIIIe siècles: Essai d'interprétation à partir d'analyses isotopiques sur des populations archéologiques* (Master's dissertation), Université de Montréal, Montréal, Canada.

Vigeant, J., Ribot, I., & Hélie, J.-F. (2017). Dietary habits in New France during the 17th and 18th centuries. *American Journal of Physical Anthropology*, *162*(3), 462–475. https://doi.org/10.1002/ajpa.23129

Zhang, J., Kobert, K., Flouri, T., & Stamatakis, A. (2014). PEAR: A fast and accurate Illumina paired-end reAd mergeR. *Bioinformatics*, *30*(5), 614–620. https://doi.org/10.1093/bioinformatics/btt593

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Harding T, Milot E, Moreau C, et al. Historical human remains identification through maternal and paternal genetic signatures in a founder population with extensive genealogical record. *Am J Phys Anthropol*. 2020;171: 645–658. https://doi.org/10.1002/ajpa.24024