Forensic Science International: Genetics 42 (2019) 1–7

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsigen

Research paper

Resolving a 150-year-old paternity case in Mormon history using DTC autosomal DNA testing of distant relatives

Ugo A. Perego[a,*], Martin Bodner[b], Alessandro Raveane[a], Scott R. Woodward[c], Francesco Montinaro[d,e], Walther Parson[b,f], Alessandro Achilli[a,*]

[a] Department of Biology and Biotechnology "L. Spallanzani", University of Pavia, Pavia, Italy
[b] Institute of Legal Medicine, Medical University of Innsbruck, Innsbruck, Austria
[c] Department of Biology, Utah Valley University, Orem, UT, United States
[d] Department of Zoology, Oxford University, South Parks Road, Oxford, United Kingdom
[e] Institute of Genomics, University of Tartu, Tartu, Estonia
[f] Forensic Science Program, Pennsylvania State University, University Park, PA, United States

ARTICLE INFO

Keywords:
Alleged paternity
Joseph Smith Jr.
Latter-day Saint (Mormon) movement
Polygamy
DNA testing
Autosomal SNPs
Mitochondrial DNA

ABSTRACT

Although autosomal DNA testing has been available for a number of years, its use to reconstruct genetic profiles of people that lived centuries in the past is relatively recent and there are no published cases where it was employed to verify a kinship relation, likely to be an alleged paternity, that occurred one and a half century ago.

DNA testing has already been employed to study the ancestry and posterity of Joseph Smith Jr., founder of the Latter-day Saint (Mormon) movement. Thanks to information found on the paternally inherited Y chromosome, a number of alleged paternities have been disproved, but obviously this analysis is not effective for alleged daughters. Likewise, his reconstructed mitogenome sequence, reported here for the first time, provides information about his maternal ancestry, but is useless in any paternity questions due to the strict maternal inheritance. Among all the children attributed to Joseph Smith Jr., Josephine Lyon, born in 1844, is perhaps the most frequently mentioned.

In the current study, 56 individuals, mostly direct descendants of Joseph Smith Jr. and Josephine Lyon, had their autosomal DNA tested to verify Josephine's biological paternity. Nearly 600,000 autosomal SNPs from each subject were typed and detailed genealogical data were compiled. The absence of shared DNA between Josephine's grandson and Joseph Smith Jr.'s five great-grandchildren together with various amounts of autosomal DNA shared by the same individual with four other relatives of Windsor Lyon is a clear indication that Josephine was not related to the Smith, but to the Lyon's family. These inferences were also verified using kinship analyses and likelihood ratio calculations.

1. Introduction

Kinship testing in forensic casework is commonly performed analyzing autosomal STR data of the involved individuals (over few generations) and/or haploid lineage markers, mitochondrial (mt)DNA and Y-chromosome, when direct maternal or paternal lines, even spanning multiple generations, are investigated. Kinship investigations on "historic" individuals have successfully been performed using (a combination of) autosomal STRs and haploid markers [1–7]. However, such analyses depend on the availability and quality of the individuals' DNA and are often restricted by the informativeness of STRs only in pedigrees spanning few generations [8], and of haploid markers only in direct maternal or paternal lines (to extant individuals) [9]. Alternative and additional marker sets are used in cases when those classical markers cannot be analyzed or do not suffice [10–12]. Autosomal SNPs, which are routinely typed in forensic genetics for individualization, phenotyping and investigating biogeographic ancestry [13] have shown to add value in kinship testing and to establish also distant relations [8,14–18]. Moreover, the validity and usefulness of combined genealogical and genetic data to reconstruct the history of populations and individuals is continually expanding, as demonstrated by recent studies [19–23].

In the past few years, DNA testing has also been employed to study the ancestry and posterity of Joseph Smith Jr. (1805–1844), founder of the Latter-day Saint (Mormon) movement. Among other teachings, Smith introduced the practice of plural marriage as part of his newly

* Corresponding authors.
E-mail addresses: ugo.jenna@gmail.com (U.A. Perego), alessandro.achilli@unipv.it (A. Achilli).

https://doi.org/10.1016/j.fsigen.2019.05.007
Received 6 February 2019; Received in revised form 31 May 2019; Accepted 31 May 2019
Available online 01 June 2019

founded faith [24,25]. However, many of his initial unions were not publicly known and it is possible that some of them were actual full-scale marriages, while others might have had a more spiritual or eternal nature void of physical intimacy. Consequently, before the advent of genetics, it has been difficult to determine whether some of the children born of women with ties to Joseph Smith Jr. were biologically his.

Joseph Smith Jr.'s Y-chromosome haplotype has been already analyzed [26,27] and was helpful in falsifying a number of paternities involving sons ascribed to Joseph Smith Jr. other than those born to his first recorded wife, Emma Hale (1804–1879) [28,29].

An unresolved genealogical question involving Joseph Smith Jr. deals with the paternity of Josephine Lyon, born on February 8, 1844. Josephine's mother, Sylvia Sessions (1818–1892), has been recorded as being married to Joseph Smith Jr. on 8 February 1842 in Nauvoo, Hancock County, Illinois, albeit she was already legally married to, but likely separated from, Windsor Lyon (1809–1849). Details about Sylvia's unions to both men (polyandry), particularly to Smith, are highly debated among Mormon historians [30]. However, among all the children attributed to Joseph Smith Jr. other than those born to Emma Hale, Josephine is perhaps the most frequently mentioned. This high interest is caused by an affidavit she signed in 1915 in which she stated:

Just prior to my mother's (Sylvia Sessions) death in 1882 she called me to her bedside and told me that her days on earth were about numbered and before she passed away from mortality she desired to tell me something which she had kept as an entire secret from me and from others, but which she now desired to communicate to me. She then told me that I was the daughter of the Prophet Joseph Smith Jr. [30].

For more than a century, scholars have been debating on the accuracy and real meaning of these words and whether Josephine was literally the biological daughter of Joseph Smith Jr.

In this study, we shed light on the biological paternity of Josephine Lyon through the analysis of autosomal DNA markers from distant relatives (descendants beyond the second degree) of Josephine Lyon, Joseph Smith Jr., and Windsor Lyon. This specific historic paternity case of obvious religious and historiographic interest also constitutes an unprecedented example of a forensic paternity investigation in a well-documented pedigree applying a genomic approach [31–34]. Since autosomal STRs and uniparental lineage markers would not be informative enough, the paternity event is clarified using SNP genetic information from second- to fifth-degree-relatives two centuries apart.

2. Methods and materials

2.1. Ethics statement

All experimental procedures and individual written informed consent, obtained from all donors, were reviewed and approved by the Western Institutional Review Board, Olympia, Washington (USA). Each participant was also informed about the purpose of the research and the use of GEDmatch for the analysis. Submission to GEDmatch was anonymous, with the use of alias, a generic e-mail address and with limited access from the public to their data (*i.e.* they were not included in the SNP SHARING POOL option and the samples were uploaded using the 'RESEARCH' option, which, according to their website, will prevent them from showing in comparison results with other kits).

2.2. Sample collection and genetic profiling

The majority of samples (n = 52) collected for this study were processed by 23andMe, a commercial company offering direct-to-consumer (DTC) autosomal genetic tests [35,36] with a customized chip produced by Illumina, covering 610,545 SNPs: 585,541 on autosomes, 2129 on the Y chromosome, 19,588 on the X chromosome, 3287 on the mitochondrial DNA. A small number of samples were processed at Family Tree DNA (N = 1), covering 716,007 SNPs (698,194 on autosomes, 17,813 on the X chromosome), and Ancestry.com (N = 3), covering 668,942 SNPs (666,531 on autosomes, 1691 on the Y chromosome, 525 on the X chromosome, 195 on mitochondrial DNA). We extrapolated these values from the raw data obtained from the three companies employed only for data production purposes.

2.3. Genealogical relationships and reference genetic distance

The autosomal DNA proportion shared with a given ancestor, in theory, halves with every interject generation. Thus, autosomal DNA testing for genealogical purposes is limited to investigate family relationships within the last five or six generations. Beyond that, genetic segments become too small and eventually disappear due to chance [15]. This means that although we can be genealogically related to all our ancestors, we might carry a clear genetic signature for relatively few of them. Centimorgan (cM) is a genetic unit for measuring genetic linkage (in terms of recombination frequency). In humans, one cM (*i.e.* a recombination frequency of 1% between two loci) roughly corresponds to about 1 million base pairs on average [37]. Because we inherit 50% of autosomal DNA from each of our parents and consequently we pass 50% of our DNA to our children, we can utilize these averages to calculate approximately how much DNA we would expect to observe between two closely related individuals [38]. For this study, a minimum threshold > 6 cMs was used as a reliable proof of biological relationships [33], because shorter segments could be the results of chance (identical by state or IBS) and not necessarily the result of a genealogical relationship (identical by descent or IBD).

The observed values of shared IBD in centimorgans were estimated using GEDmatch, an open-source database that provides DNA and genealogical analysis tools for researchers and genealogists, with the aim to compare the observed shared DNA values calculated on our dataset with the amount expected calculating a linear decrement of shared DNA through generations or observed in about 25,000 pairs of relatives in "The Shared cM Project 3.0 tool v4" [33,39]. For every SNP, it records half or full match among pairs of individuals and considers a true signal of IBD only if an uninterrupted match longer than a specific threshold is observed (> 6 cMs in our analyses).

The investigation presented in this study was quite timely, as the questioned paternity event occurred in 1844, which means that the future posterity of Joseph Smith Jr., Josephine and Windsor Lyon might not carry enough autosomal DNA to resolve the question of their alleged biological relationship.

A total of 56 participants agreed to take part in the current study and their relationships are depicted in Fig. 1. These individuals were selected based on their relationship to either Joseph Smith Jr. or to Josephine Lyon, with the final objective to obtain two balanced datasets for genetic comparison. Particular attention during the selection criteria was placed on the number of generations separating the living descendant to the ancestor of interest (a vertical pedigree approach, for example a great-grandchild was preferred over a great-great-grandchild), and on the spread or degree of separation that these descendants shared among themselves. A horizontal pedigree approach was also adopted, for example second- or third-degree cousins were chosen over first degree relationships. Naturally, these criteria could only be applied based on the availability of living descendants who consented to be tested. During the recruiting process, a few additional individuals related to either family volunteered to be included even if their genetic contribution was likely not as fundamental to the success of the study as those meeting the selection criteria. The main objective in following these principles was to build a dataset of individuals carrying as much autosomal DNA from either Joseph Smith Jr. or Josephine Lyon to ensure that any DNA segments that the two families might have in common was indeed from their two ancestors.

If the alleged paternity would be supported, the six candidates expected to carry the largest amount of autosomal DNA are those closer in number of generations to Joseph Smith Jr. and Josephine Lyon (Fig. 2).

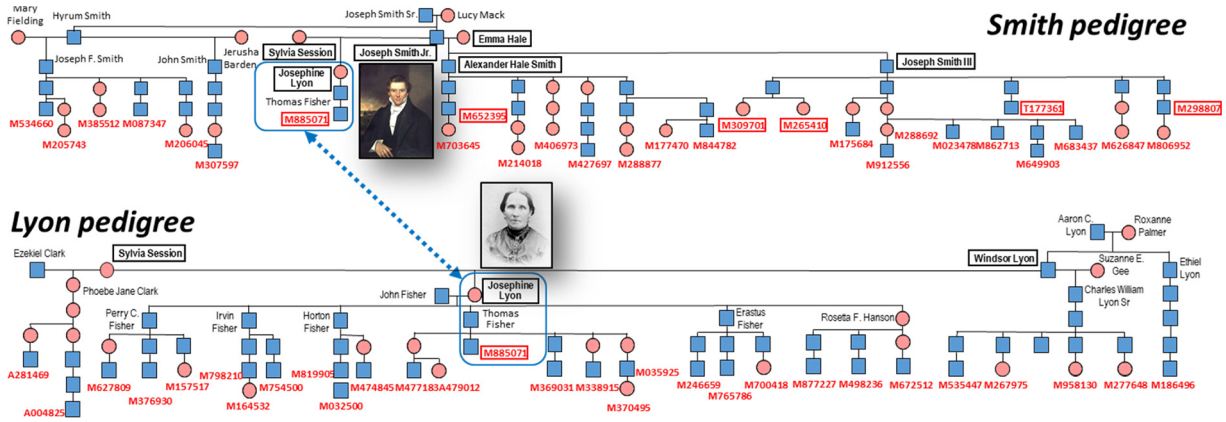

Fig. 1. Genealogical data. The two pedigrees encompass the 56 individuals that had their autosomal DNA tested to verify Josephine Lyon's paternity. Names of participants were replaced with anonymous IDs for privacy purposes. Samples with framed IDs are the most informative ones, as summarized in Fig. 2. See Text S1 for more genealogy details. (Photographs courtesy Community of Christ and Clark Layton).

Joseph Smith Jr. begat nine biological children with his first wife, Emma Hale [30]. Four sons lived to adulthood, but only two of them, Joseph Smith III and Alexander Hale Smith, have known living biological posterity. All children and grandchildren of Joseph Smith Jr. were deceased at time of sampling. Five great-grandchildren were still alive and agreed to contribute a DNA sample (M652395, T177361, M298807, M309701 and M265410, expectedly carrying about 12.5% of Joseph Smith Jr.'s autosomal DNA). Josephine Lyon gave birth to ten children, with seven surviving to adulthood. Descendants from six children donated DNA samples to the current study, including the only grandchild that was still alive at time of sampling (M885071, with expected ˜25% of Josephine autosomal DNA). Under the assumption that Joseph Smith Jr. was the biological father of Josephine Lyon, her grandson M885071 would share an expected 1.563% of autosomal DNA (or approximately 106.25 cMs, Table 1) with Joseph Smith Jr.'s great grandchildren included in this study. Given the number of generations separating these individuals, it is unlikely that we can confirm a biological paternity, but we should find strong evidence to either support or exclude it.

2.4. Kinship analysis

PLINKv1.9 [40] was used to merge and convert the data from the UCSC build37 reference genome. A total of 227,200 SNPs were retained to compare all 56 individuals, after filtering for missing data (threshold = 0.01). It is worth mentioning that similar results (data not shown) were obtained when considering only 23andMe data, 52 samples and 501,495 filtered SNPs.

The kinship coefficient, defined as the probability that two randomly chosen alleles are identical by descent (IBD), provides an estimate of the relatedness between pairs of individuals; this coefficient can be measured using different algorithms implemented in several software. Here, the kinship coefficient was estimated in KING v1.4 96 (*–kinship* and *–ibs* flags), which allows checking for pair-wise relationships by assuming the existence of a population structure [41]. Different ranges of the estimated kinship coefficients, *i.e.* [> 0.354], [0.177, 0.354], [0.0884, 0.177] and [0.0442, 0.0884], were used to infer different degrees of relationships, *i.e.* a duplicate sample, 1 st degree, 2nd degree and 3rd degree, respectively. Eventually, the kinship values were used first to build an UPGMA (Unweighted Pair Group Method with Arithmetic Mean) tree (based on Euclidean distances) using the *hclust()* function in R and then to create an heatmap with the *pheatmap* package [42].

The pairwise proportion of IBD was also investigated using the *–genome* flag in PLINKv1.9 and the results were plotted using *ggplot2* package in R. We took into consideration the IBD score (or PI_HAT) in the form $p(\text{IBD} = 2) + 0.5*p(\text{IBD} = 1)$.

2.5. Pairwise likelihood ratio calculations

Pairwise Likelihood Ratios (LRs) were computed using SNP analyzer, a software package that performs kinship calculations using large

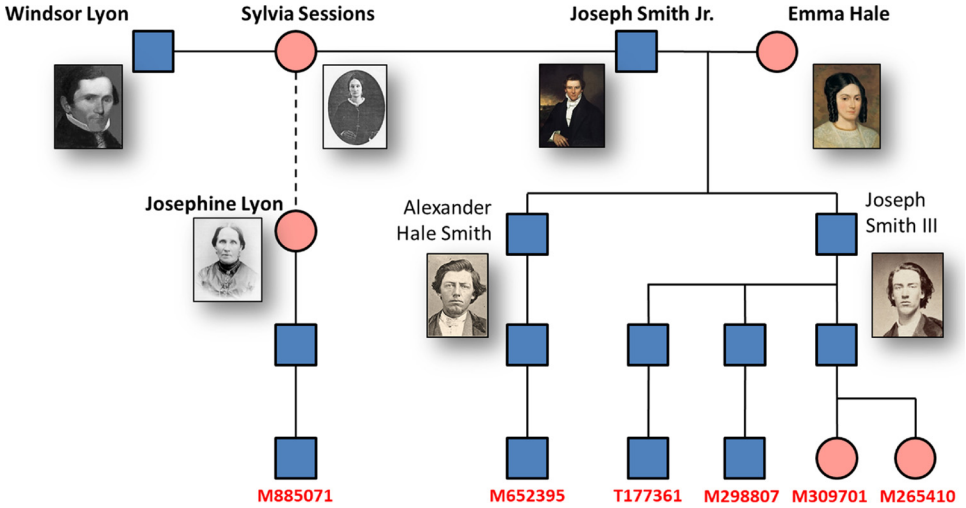

Fig. 2. A schematic genealogical tree including the six most informative samples. This pedigree includes the six candidates expected to carry the largest amount of autosomal DNA derived from Sylvia Sessions, Windsor Lyon (biological father) and Joseph Smith Jr. (alleged father). (Photographs courtesy Community of Christ and Clark Layton).

Table 1
Comparison between the Josephine Lyon's grandson and five great grand-children of Joseph Smith Jr. (alleged father).

Josephine Lyon's grandson	Joseph Smith Jr.'s great-grandchildren	Alleged Relationship[a]	Average Expected cMs	Avgs and ranges observed in Shared cM Project (v. 3.0)	Actual cMs
M885071	T177361	Half 2C[a]	106.25	117 (9-397)	0
M885071	M298807	Half 2C	106.25	117 (9-397)	0
M885071	M265410	Half 2C	106.25	117 (9-397)	0
M885071	M309701	Half 2C	106.25	117 (9-397)	0
M885071	M652395	Half 2C	106.25	117 (9-397)	0

[a] Half2C: half second cousin.

Table 2
Comparison between the Josephine Lyon's grandson and five relatives of Windsor Lyon (biological father).

Josephine Lyon's grandson	Windsor Lyon's relative	Alleged Relationship[a]	Average Expected cMs	Avgs and ranges observed in Shared cM Project (v. 3.0)	Actual cMs
M885071	M277648	Half 2C1R	53.13	73 (0-341)	117.4
M885071	M958130	Half 2C1R	53.13	73 (0-341)	84.2
M885071	M186496	3C1R	26.56	48 (0-173)	30.1
M885071	M267975	Half 2C1R	53.13	73 (0-341)	19.7
M885071	M535447	Half 2C1R	53.13	73 (0-341)	0

[a] Half 2C1R: half second cousin once removed; 3C1R: third cousin once removed.

sets of SNPs [43]. LRs assess the probability of observing the genetic data under two competing hypotheses about the relatedness between two individuals and are calculated from IBD patterns within pedigrees. They are provided in several degrees of thinning (or pruning), choosing a minimum distance of 0.1–8 cM s between the SNPs considered for calculations. This naïve method aims to select enough markers but mitigating the possible effect of linkage disequilibrium (LD) [18]. Summary statistics were calculated from the ˜300,000 overlapping markers between the person of interest and the reference population. Population frequency data from Utah residents, with Northern and Western European ancestry, from the 1000 Genomes project [44] were chosen as assumedly closest. Default settings were used (no genotyping error; minimum allele frequency = 0.4). A total of 162 calculations were performed on the 27 individuals considered to be the most informative to solve the paternity case.

3. Results and discussion

3.1. Confirming the genetic relationships within the two families

We started by comparing Joseph Smith Jr.'s five grandchildren (M652395, T177361, M298807, M309701 and M265410) and Josephine's grandson (M885071) to all other individuals in their respective tree (Figs. 1 and 2 and Table S1). The degrees of relationship range from parent/child all the way to third cousins twice removed (3C2R). A positive linear correlation is observed as closer Smiths and Lyons relatives show higher amounts of cMs shared respectively and vice versa. Naturally, in a few of the more distant familial relations, the observed amount of shared cMs was less than the minimum threshold of 6 cM. These results support the correctness of the genealogical information, thus demonstrating that these six samples are indeed related to everyone that participated in the study within each respective family group (Smith and Lyon, Fig. 1).

3.2. Verifying the biological paternity of Josephine Lyon

Once the genealogical information on the two families was confirmed, the next step was a comparison between Josephine Lyon's only surviving grandson with each of Joseph Smith Jr.'s great-grandchildren (Table 1). None of the five Smith's shared any amount of autosomal DNA with Josephine's grandchild, while at the half second-degree cousin relationship, which the five descendants of Joseph Smith Jr. would allegedly share with Josephine's grandchild, it is expected to observe an average of 106.25 cMs (based on a linear decrement) or 117 (9–397) cMs (based on The Shared cM Project, v. 3.0). As a positive control, the observed range for a similar relationship within each family was 27.7–177.5 cM s (twenty occurrences in the Smith family). Therefore, the observed absence of shared autosomal DNA between Josephine's grandson and Joseph Smith Jr.'s five great-grandchildren indicates that the five Smiths are probably not biologically related within the alleged relationship degree.

In order to test the paternity of Windsor Lyon, we checked the autosomal DNA from Josephine Lyon's grandson shared with five individuals related to Josephine through her mother Sylvia Sessions or through Windsor Lyon's family, but bearing no apparent close relationship to the Smith family (Table 2). Even if one of them (M535447) did not share any cMs, which is a likely possibility at those distant degrees of relationship (Half 2C1R), four individuals (M277648, M958130, M186496, and M267975) shared autosomal DNA with Josephine's grandson, ranging from 19.8 to 117.5 cMs. The closeness of Josephine's descendant M885071 with the Lyon family rather than the Smith pedigree was also confirmed by the average proportion of IBD fragments shared with Windsor Lyon's that is significantly higher (Wilcoxon *p-value* < 0.05) than with the Joseph Smith Jr.'s descendants (Fig. S1).

The overall results are well summarized by the UPGMA genetic tree based on kinship values showing a well-defined separation between Smiths and Lyons (Fig. 3). Notably, all the Josephine's great and grandsons fall in the group of Lyon family among the previously mentioned descendants of Sylvia Sessions (A281469, A004825) and Windsor Lyon (M535447, M267975, M958130, M277648 and M186496). The location of these samples in the UPGMA clustering tree, together with the relationship highlighted in the heatmap, further supports the relationship between Josephine and the Lyons. Moreover, genetic data are consistent with the genealogical relationships reported in Fig. 1 for both families, with one notable exception. In fact, sample M307597, which belongs to the Smith family, clusters with the Lyon family (the only pink square among the blue ones). A possible explanation is the distant relationship M307597 shares with Joseph Smith Jr., as he is a descendant of Joseph's brother Hyrum through his first wife Jerusha Barden. According to the genealogical data provided, his closest Joseph Smith Jr.'s descendants tested in the current study would be his third cousin twice removed (3C2R). Table S1 shows that M307597 has extremely little or no DNA in common with the five Joseph Smith Sr.'s great grandchildren included in the current study. As both Hyrum Smith and Josephine Lyon's descendants relocated and settled in the

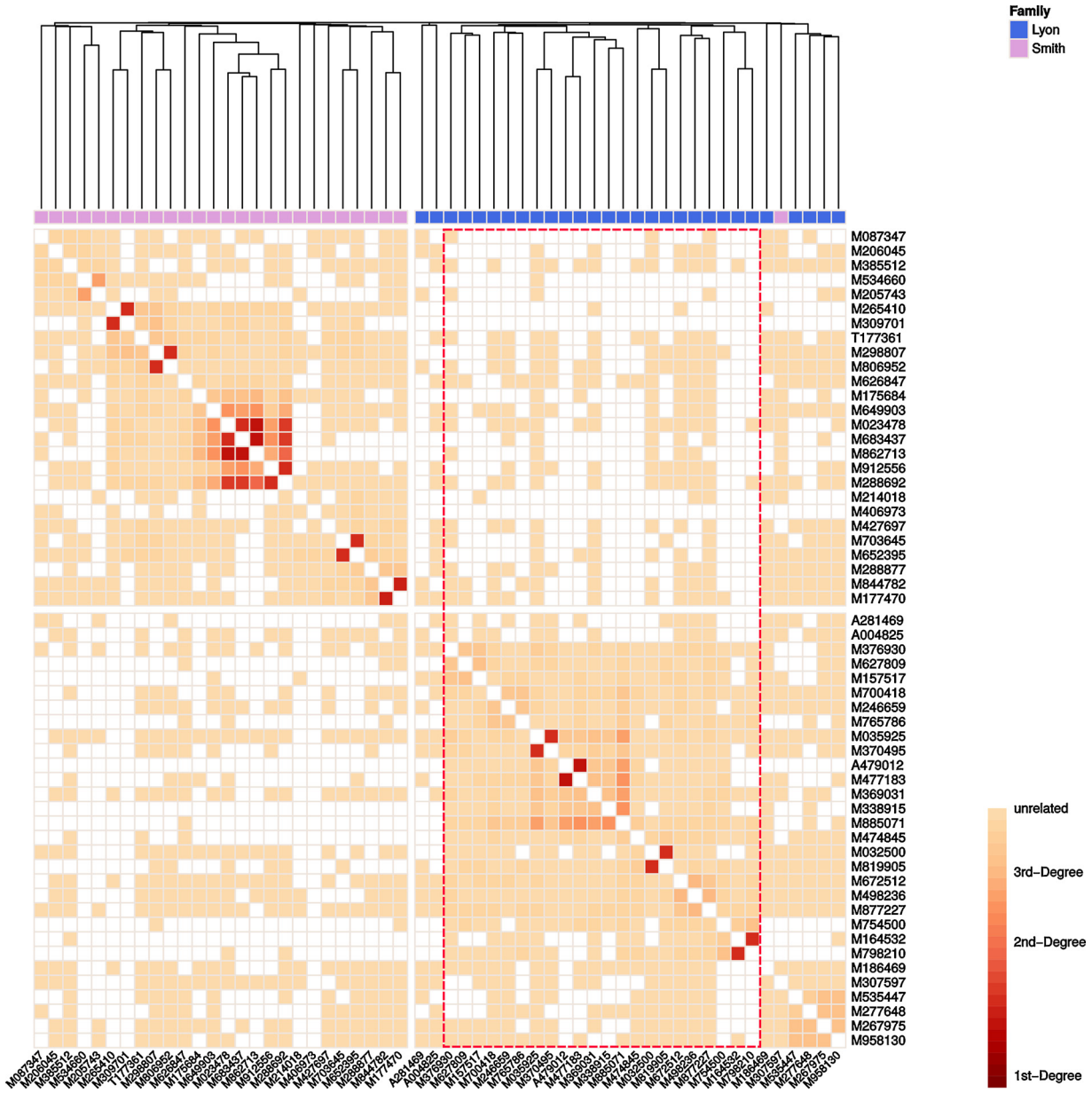

Fig. 3. UPGMA clustering tree inferred from the heatmap of the 56 samples. Darker colors indicate closer relationship and the red box encloses the columns with the kinship values of Josephine Lyon's direct descendants. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

state of Utah, it is likely (although it has not been verified yet) that M307597 could have some common, more recent ancestors with the Lyons. This hypothesis would explain why M307597 clusters with the Lyon and not with the Smith family in Fig. 3.

3.3. Supporting our results with Pairwise LR

To corroborate our inferences, a total of 162 Pairwise LR comparisons were also performed to test the genetic relationships between the most informative offspring pairs.

We started by comparing the 18 most informative descendants of Josephine Lyon (*i.e.* her grandson M885071, and her 17 great grandchildren processed with 23andMe: M477183, M369031, M338915, M035925, M798210, M754500, M819905, M474845, M627809, M376930, M157517, M246659, M765786, M700418, M877227, M498236, M672512; Fig. 1) and the five most informative descendants of Joseph Smith Jr. and Emma Hale (*i.e.* their five great grandchildren M652395, M309701, M265410, T177361, M298807; Fig. 2). None of the 90 pairwise comparisons supported the hypothesized genetic relationships of half second cousin (half 2C, for M885071) or half second cousin once removed (half 2C1R, for all others, respectively). Results were largely inconclusive and provided limited support (LR < < 10) [45] for either hypothesis with all marker selections, which may indicate that even more markers would be necessary to resolve this case [18]. Thus, there is no evidence for the paternity of Joseph Smith Jr.

Exceedance probability simulations have shown that high LR results are possible in very distant true relations of individuals [18]. This is demonstrated in the second batch of analyses. In fact, the LR results for the denser marker set favor genetic relations of the 18 descendants of Josephine Lyon (see above) to the four available descendants of Windsor Lyon and Suzanne E. Gee (*i.e.* their four great great-grandchildren M535447, M267975, M958130, M277648) (Fig. 1). Among the 72 pairwise comparisons performed, assuming half second cousin once removed (half 2C1R) relations for M885071 and half third cousin

(half 3C) relations for all others, respectively, 40 yielded LR results > 1,000, which is considered very strong support [45]. In general, high LRs were yielded in comparisons including all the individuals, except for only three individuals related to Josephine Lyon through three of her children that showed no indication for the investigated relationship with the chosen markers. The results indicate strong evidence for relatedness between individuals through Windsor Lyon, thus also for the paternity hypothesis.

The proportion of pairwise comparisons between Josephine Lyon's and Windsor Lyon's descendants that did not yield high LR in favor of the putative relation were dispersed among individuals that yielded extremely high LRs with other individuals in the same tree, which likely pinpoints the variation in the small amount of shared cMs between very distantly related individuals. False rates, indicating the percentage of unrelated pairs that is falsely concluded as related (LR > 1000) [43] were very low (< 0.02) (data not shown). In this context, it appears highly unlikely that the negative kinship results from all 90 pairwise comparisons between Josephine Lyon's and Joseph Smith Jr.'s descendants (that, in addition, were putatively closer related than Windsor Lyon's descendants) would be caused by this effect.

3.4. Completing the genetic profile of Joseph Smith Jr. with his mitogenome

In order to completely investigate the genetic profile of Joseph Smith Jr. we have also reconstructed his maternal lineage by sequencing the entire mitochondrial DNA (using the previously described protocol [46]) from a descendant of Katherine Smith, one of Joseph Smith Jr.'s sisters (Fig. S2). His maternal line is characterized by the haplotype 152C 263 G 315.1C 750 G 1438 G 2706 G 4769 G 7028 T 8860 G 9039A 14180C 15,326 G 16519C relative to the rCRS [47]. It belongs to haplogroup HV18 [48], a rare lineage (five samples in EMPOP v4/R12, https://empop.online/) [49,50] that has been defined for the first time by analyzing modern Iranians [51] and that is nowadays present in Western Europe and the Middle East. It has been recently identified also in an ancient Iron Age from the South Baltic region [52].

4. Conclusion

In this project, a total of 56 individuals representing both Joseph Smith Jr.'s and Josephine Lyon's families, including controls from other relatives, contributed samples for autosomal DNA testing to solve the highly debated question of the Joseph Smith Jr.'s alleged paternity of Josephine Lyon. Using pedigrees may be a pitfall, because they are necessarily assumptions from historical/genealogical data, which might lead to a plethora of possible errors, ranging from typos during pedigree transcriptions to possible extra-pair paternity or maternity events [9,23]. Actually, this was an issue when preliminary genetic data from a small number of descendants (three Smiths and six Lyons) were published [28]. DNA sharing between the Smith/Lyon family member pairs was observed, but it was quickly determined that it could have been the result of additional intermarriages and not necessarily from a Joseph and Sylvia alleged offspring. Here, a much larger number of individuals were tested and the reconstructed pedigrees were concordant with the results from autosomal SNPs, thus strengthening the genetic outcome.

Although a reconstruction of Joseph Smith Jr. and Josephine Lyon's DNA through their descendants separated by three or more generations will never provide the same level of accuracy as a paternity test with DNA obtained directly from Joseph and Josephine, data presented in this study was consistent and offered the strongest evidence to date toward clarifying their alleged father/daughter relationship. Based on this analysis, it appears that Josephine did not share a biological tie with the founder of Mormonism and that perhaps what Sylvia Sessions told her daughter would mean something other than the biological relationship many historians have perceived. This is certainly an additional piece of genetic evidence in future researches on the subject of Joseph Smith Jr.'s practice of polygamy.

This study confirms that autosomal DNA testing is providing the opportunity to address many genealogical questions that could not be answered by analysis of autosomal STRs or the uniparental Y-chromosome and mitochondrial molecules. The weak signal of genetic inheritance of shared cM segments is here strengthened by the high number of individuals available from the well documented pedigrees, the high number of markers and the use of three methods independently. Errors could always be introduced in these types of studies by incorrectly transcribing genealogical records, lab errors or even illegitimacies. However, when a considerable number of individuals are carefully selected and tested for the same family history case, the strength of both the results and the study conclusions increase. This study proves that a considerable number of individuals carefully selected and tested for the same family history case might enable a result even when the "actual" individuals are not available.

Data availability

Mitogenome sequence has been deposited in GenBank (xxxxx)

To access the genotyped data, researchers should send a signed letter to A.A. containing the following text: (a) I will not distribute the data outside my collaboration; (b) I will not post the data publicly;(c) I will make no attempt to connect the genetic data to personal identifiers for the samples; (d) I will not use the data for any selection studies; (e) I will not use the data for medical or disease-related analyses; (f) I will not use the data for commercial purposes.'

Acknowledgments

This study would have not been possible without the support of key individuals in both Joseph Smith Jr.'s and Josephine Lyon's families. We are especially grateful to all the fifty-six DNA donors. We are grateful to the generous individuals that made financial contributions through the crowdfunding website https://www.gofundme.com/JosephSmithDNA. This funding was used for the DNA testing and tools needed for this project.

This study received support from: i) the University of Pavia strategic theme "Towards a governance model for international migration: an interdisciplinary and diachronic perspective" (MIGRAT-IN-G) (to A.A.); ii) the Italian Ministry of Education, University and Research (MIUR), *i.e.* Dipartimenti di Eccellenza Program (2018–2022) to the Dept. of Biology and Biotechnology "L. Spallanzani" at the University of Pavia (to A.A. and A.R.) and PRIN2017 project No. 20174BTC4R (to A.A.); iii) the European Research Council (ERC), *i.e.* Horizon 2020 Consolidator Grant CoG-2014 No. 648535 (to A.A.).

We are also grateful to Ann Turner and Blaine Bettinger for providing valuable feedback and suggestions to early drafts of this paper, and to the PhD programs of the University of Pavia and to Cristian Capelli at the Department of Zoology to the University of Oxford for facilitating the research visit of A.R..

We thank Daniel Kling (Dep. of Forensic Sciences, University Hospital Oslo, Oslo, Norway) for providing SNP analyzer and friendly support with using the software.

We thank the two anonymous reviewers for useful comments.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.fsigen.2019.05.007.

References

[1] E.A. Foster, M.A. Jobling, P.G. Taylor, P. Donnelly, P. De Knijff, R. Mieremet, T. Zerjal, C. Tyler-Smith, Jefferson fathered slave's last child, Nature 396 (6706) (1998) 27.

[2] C. Haas, N. Shved, F.J. Rühli, C. Papageorgopoulou, J. Purps, M. Geppert, S. Willuweit, L. Roewer, M. Krawczak, Y-chromosomal analysis identifies the skeletal remains of Swiss national hero Jörg Jenatsch (1596-1639), Forensic Sci. Int. Genet. 7 (6) (2013) 610–617.

[3] C.M. Bauer, M. Bodner, H. Niederstätter, D. Niederwieser, G. Huber, P. Hatzer-Grubwieser, K. Holubar, W. Parson, Molecular genetic investigations on Austria's patron saint Leopold III, Forensic Sci. Int. Genet. 7 (2) (2013) 313–315.

[4] T.E. King, G.G. Fortes, P. Balaresque, M.G. Thomas, D. Balding, P.M. Delser, R. Neumann, W. Parson, M. Knapp, S. Walsh, Identification of the remains of King Richard III, Nat. Commun. 5 (2014) 5631.

[5] P. Gill, P.L. Ivanov, C. Kimpton, R. Piercy, N. Benson, G. Tully, I. Evett, E. Hagelberg, K. Sullivan, Identification of the remains of the Romanov family by DNA analysis, Nat. Genet. 6 (2) (1994) 130.

[6] J.A. Irwin, S.M. Edson, O. Loreille, R.S. Just, S.M. Barritt, D.A. Lee, T.D. Holland, T.J. Parsons, M.D. Leney, DNA identification of "Earthquake McGoon" 50 years postmortem, J. Forensic Sci. 52 (5) (2007) 1115–1118.

[7] M.D. Coble, O.M. Loreille, M.J. Wadhams, S.M. Edson, K. Maynard, C.E. Meyer, H. Niederstätter, C. Berger, B. Berger, A.B. Falsetti, Mystery solved: the identification of the two missing Romanov children using DNA analysis, PloS ONE 4 (3) (2009) e4838.

[8] J.M. Butler, Fundamentals of Forensic DNA Typing, Academic Press, 2009.

[9] M.H. Larmuseau, M. Bodner, The biological relevance of a medieval king's DNA, Biochem. Soc. Trans. 46 (4) (2018) 1013–1020.

[10] P. Grubwieser, B. Zimmermann, H. Niederstätter, M. Pavlic, M. Steinlechner, W. Parson, Evaluation of an extended set of 15 candidate STR loci for paternity and kinship analysis in an Austrian population sample, Int. J. Legal Med. 121 (2) (2007) 85–89.

[11] P.M. Schneider, Beyond STRs: the role of diallelic markers in forensic genetics, Transfus. Med. Hemother. 39 (3) (2012) 176–180.

[12] A.O. Tillmar, P. Mostad, Choosing supplementary markers in forensic casework, Forensic Sci. Int. Genet. 13 (2014) 128–133.

[13] C. Phillips, L. Prieto, M. Fondevila, A. Salas, A. Gómez-Tato, J. Álvarez-Dios, A. Alonso, A. Blanco-Verea, M. Brión, M. Montesino, Ancestry analysis in the 11-M Madrid bomb attack investigation, PLoS ONE 4 (8) (2009) e6583.

[14] C. Phillips, M. Fondevila, M. García-Magariños, A. Rodriguez, A. Salas, A. Carracedo, M.V. Lareu, Resolving relationship tests that show ambiguous STR results using autosomal SNPs as supplementary markers, Forensic Sci. Int. Genet. 2 (3) (2008) 198–204.

[15] C. Phillips, M. García-Magariños, A. Salas, Á. Carracedo, M.V. Lareu, SNPs as supplements in simple kinship analysis or as core markers in distant pairwise relationship tests: when do SNPs add value or replace well-established and powerful STR tests? Transfus. Med. Hemother. 39 (3) (2012) 202–210.

[16] D. Kling, J. Welander, A. Tillmar, Ø. Skare, T. Egeland, G. Holmlund, DNA microarray as a tool in establishing genetic relatedness—Current status and future prospects, Forensic Sci. Int. Genet. 6 (3) (2012) 322–329.

[17] R. Li, H. Li, D. Peng, B. Hao, Z. Wang, E. Huang, R. Wu, H. Sun, Improved pairwise kinship analysis using massively parallel sequencing, Forensic Sci. Int. Genet. 38 (2019) 77–85.

[18] D. Kling, On the use of dense sets of SNP markers and their potential in relationship inference, Forensic Sci. Int. Genet. 39 (2019) 19–31.

[19] N. Zaitlen, P. Kraft, N. Patterson, B. Pasaniuc, G. Bhatia, S. Pollack, A.L. Price, Using extended genealogy to estimate components of heritability for 23 quantitative and dichotomous traits, PLoS Genet. 9 (5) (2013) e1003520.

[20] A. Jagadeesan, E.D. Gunnarsdóttir, S.S. Ebenesersdóttir, V.B. Guðmundsdóttir, E.L. Thordardottir, M.S. Einarsdóttir, H. Jónsson, J.-M. Dugoujon, C. Fortes-Lima, F. Migot-Nabias, Reconstructing an African haploid genome from the 18th century, Nat. Genet. 50 (2) (2018) 199.

[21] B.M. Henn, L. Hon, J.M. Macpherson, N. Eriksson, S. Saxonov, I. Pe'er, J.L. Mountain, Cryptic distant relatives are common in both isolated and cosmopolitan genetic samples, PLoS ONE 7 (4) (2012) e34267.

[22] J. Kaplanis, A. Gordon, T. Shor, O. Weissbrod, D. Geiger, M. Wahl, M. Gershovits, B. Markus, M. Sheikh, M. Gymrek, Quantitative analysis of population-scale family trees with millions of relatives, Science 360 (6385) (2018) 171–175.

[23] M.H.D. Larmuseau, A. Van Geystelen, M. Van Oven, R. Decorte, Genetic genealogy comes of age: perspectives on the use of deep-rooted pedigrees in human population genetics, Am. J. Phys. Anthropol. 150 (4) (2013) 505–511.

[24] R.P. Howard, The changing RLDS response to mormon polygamy: a preliminary analysis, John Whitmer Hist. Assoc. J. 3 (1983) 14–29.

[25] B.C. Hales, Joseph Smith's personal polygamy, J. Mormon Hist. 38 (2) (2012) 163–228.

[26] J. Gitschier, Inferential genotyping of Y chromosomes in latter-day saints founders and comparison to Utah samples in the HapMap project, Am. J. Hum. Genet. 84 (2) (2009) 251–258.

[27] U.A. Perego, N.M. Myres, S.R. Woodward, Reconstructing the Y-chromosome of Joseph Smith: genealogical applications, J. Mormon Hist. 31 (2) (2005) 42–60.

[28] U.A. Perego, Joseph Smith Jr., The question of polygamous offspring, and DNA analysis, in: N.G. Bringhurst, C.L. Foster (Eds.), The Persistence of Polygamy: Joseph Smith and the Origins of Mormon Polygamy, John Whitmer Books, 2010, pp. 233–256.

[29] M. Kennedy, Joseph and Emma's Family, Ensign, (2008) https://www.lds.org/ensign/2008/02/joseph-and-emmas-family?lang=eng#.

[30] B.C. Hales, The Joseph Smith — Sylvia Sessions Plural Sealing: Polyandry or Polygyny? Mormon Hist. Studies 9 (1) (2008) 41–57.

[31] S.R. Browning, B.L. Browning, Identity by descent between distant relatives: detection and applications, Annu. Rev. Genet. 46 (2012) 617–633.

[32] S.C. Nelson, S.M. Fullerton, Bridge to the literature"? Third-party genetic interpretation tools and the views of tool developers, J. Genet. Couns. 27 (4) (2018) 770–781.

[33] Y. Erlich, T. Shor, I. Pe'er, S. Carmi, Identity inference of genomic data using long-range familial searches, Science 362 (6415) (2018) 690–694.

[34] M.D. Ramstetter, S.A. Shenoy, T.D. Dyer, D.M. Lehman, J.E. Curran, R. Duggirala, J. Blangero, J.G. Mezey, A.L. Williams, Inferring identical-by-descent sharing of sample ancestors promotes High-Resolution relative detection, Am. J. Hum. Genet. 103 (1) (2018) 30–44.

[35] R. Khan, D. Mittelman, Consumer genomics will change your life, whether you get tested or not, Genome Biol. 19 (1) (2018) 120.

[36] E. Vayena, Direct-to-consumer genomics on the scales of autonomy, J. Med. Ethics 41 (4) (2015) 310–314.

[37] J.E. Richards, R.S. Hawley, The Human Genome, Academic Press, 2010.

[38] W. Hill, B. Weir, Variation in actual relationship as a consequence of mendelian sampling and linkage, Genet. Res. 93 (1) (2011) 47–64.

[39] B.T. Bettinger, The Shared cM Project – Version 3.0, (2017) https://thegeneticgenealogist.com/wp-content/uploads/2017/08/Shared_cM_Project_2017.pdf.

[40] C.C. Chang, C.C. Chow, L.C. Tellier, S. Vattikuti, S.M. Purcell, J.J. Lee, Second-generation PLINK: rising to the challenge of larger and richer datasets, Gigascience 4 (1) (2015) 7.

[41] A. Manichaikul, J.C. Mychaleckyj, S.S. Rich, K. Daly, M. Sale, W.-M. Chen, Robust relationship inference in genome-wide association studies, Bioinformatics 26 (22) (2010) 2867–2873.

[42] R. Kolde, Pheatmap: pretty heatmaps, R package version, (2012).

[43] D. Kling, SNP analyzer – a tool to analyze large sets of genetic markers accounting for linkage, forensic sci, Int. Genet. Suppl. Ser. 6 (2017) e587–e588.

[44] The Genomes Project Consortium, The Genomes Project Consortium, A global reference for human genetic variation, Nature 526 (2015) 68.

[45] I. Evett, B.S. Weir, Interpreting DNA Evidence: Statistical Genetics for Forensic Scientists, Oxford University Press, Incorporated, 1998.

[46] A. Achilli, C. Rengo, C. Magri, V. Battaglia, A. Olivieri, R. Scozzari, F. Cruciani, M. Zeviani, E. Briem, V. Carelli, P. Moral, J.M. Dugoujon, U. Roostalu, E.L. Loogväli, T. Kivisild, H.J. Bandelt, M. Richards, R. Villems, A.S. Santachiara-Benerecetti, O. Semino, A. Torroni, The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool, Am. J. Hum. Genet. 75 (5) (2004) 910–918.

[47] R.M. Andrews, I. Kubacka, P.F. Chinnery, R.N. Lightowlers, D.M. Turnbull, N. Howell, Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA, Nat. Genet. 23 (2) (1999) 147.

[48] M. Van Oven, M. Kayser, Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation, Hum. Mutat. 30 (2) (2009) E386–E394.

[49] N. Huber, W. Parson, A. Dür, Next generation database search algorithm for forensic mitogenome analyses, Forensic Sci. Int. Genet. 37 (2018) 204–214.

[50] W. Parson, A. Dür, EMPOP–a forensic mtDNA database, Forensic Sci. Int. Genet. 1 (2) (2007) 88–92.

[51] M. Derenko, B. Malyarchuk, A. Bahmanimehr, G. Denisova, M. Perkova, S. Farjadian, L. Yepiskoposyan, Complete mitochondrial DNA diversity in Iranians, PloS ONE 8 (11) (2013) e80673.

[52] I. Stolarek, A. Juras, L. Handschuh, M. Marcinkowska-Swojak, A. Philips, M. Zenczak, A. Dębski, H. Kóčka-Krenz, J. Piontek, P. Kozlowski, A mosaic genetic structure of the human population living in the South Baltic region during the Iron Age, Sci. Rep. 8 (1) (2018) 2455.