Supporting Information

Rogaev et al. 10.1073/pnas.0811190106

SI Material and Methods

History of the Case. The last Russian Emperor (Tsar) Nicholas II (1868–1918), his wife Empress (Tsarina) Alexandra Feodorovna (1872–1918), their 4 daughters (Olga, Tatiana, Maria and Anastasia), their son, heir of the dynasty, Prince Alexei, along with their court physician and 3 servants, are believed to have been murdered in 1918 during the Civil War in the Ural region of central Russia.

In 1991 a shallow grave was found near Yekaterinburg with remains of several skeletons. DNA and anthropological studies suggested that the remains were those of Nicholas II and the Romanov family and their attendants (1–4). The connection of the individuals found in the grave to members of the European Royal family was based mainly on analysis of short hypervariable region (HVR) mitochondrial (mt) DNA fragments. The data on mtDNA sequences in human populations were insufficient at that time. Paternal lineages had not been studied. Presumable controversies over the authenticity of the remains have been discussed (5). The remains of 2 of the children still had not been found. In 2 primary anthropological studies performed in 1992 there was disagreement on the individual identity of skeleton fragments belonging to Tatiana (21-year-old), Maria (19 years old) and Anastasia (17 years old). The anthropological testing implied that the bodies of one of the youngest Romanov daughters (Anastasia or Maria) and Alexei were missing from the grave. Speculations that 2 of the Emperor's children (Anastasia and Alexei) escaped the murders have persisted to the present day (6, 7). Anthropological evaluations of age, proportion of the skeleton bones, skulls and facial reconstructions performed by Russian forensic anthropologists strongly suggested that skeletons N3, N5 and N6 correspond to characteristics of Olga, Tatiana and Anastasia and, thus, Maria remained a missing person (1). There are some archival documents with the recollections of this crime left by murderers asserting that 2 or more bodies were burned and buried in a separate place (1).

The Orthodox Church canonized the Nicholas II and his family as "passion bearers," but acknowledged that more comprehensive evidence is required to verify the authenticity of these relics.

On August 24, 2007, Russian archaeologists announced the discovery of damaged and partially burned bone fragments found at a second burial site near Yekaterinburg. Along with the remains of the 2 bodies, archaeologists found shards from a container of sulfuric acid and bullets of various caliber.

The criminal forensic and historical investigation into the deaths of the Nicholas II family was reopened. The evidentiary items (bone samples) were obtained for official investigation in this forensic casework in accordance to protocols and regulations required for chain of custody. E.I.R. was assigned by the Prosecutor Office of Russian Federation as an expert to lead the genetic analysis of the bone specimens under criminal casework investigation.

Reference Samples. The revised Cambridge Reference Sequence (rCRS) for complete mitochondrial genome sequence (AC_000021) corrected for errors (identified in the original "Cambridge" sequence) was used for pairwise comparison with mitochondrial genome sequences determined in this study. The Emperor Nicholas II mtDNA sequence (Princess "Dagmar" type) showed substantially more mismatches with the rCRS than mtDNA sequence of Empress Alexandra ("Queen's Victoria" type) (Tables 1 and 2).

The modern biological samples were obtained by collection of buccal swabs, blood specimens and hairs from members of Royal families of Europe and Romanov relatives.

For mtDNA analysis the samples from relatives of the Queen Victoria maternal lineage were collected from Princess O. [who is a granddaughter of King Alfonso XIII of Spain (1886–1941) and a great-great-granddaughter of Queen Victoria (1819–1901)] and her daughter D.; Princess K. [who is a great-great-granddaughter of Princess Victoria of Hesse (1863–1950), the sister of Empress Alexandra Feodorovna] and her daughter V.

The samples for the maternal lineage of Nicholas II were collected from great-granddaughter of Grand Duchess Xenia Alexandrovna (1875–1960), sister of Nicholas II (Fig. 2).

For comparative analysis of the non-recombinant region of the Y-chromosome we collected and genotyped the samples from Prince N. and from his first degree brother Prince D., who are great-grandsons of Grand Duke Nicholas (1831–1891) [son of Emperor Nicholas I (1796–1855)] and from Prince A., who is a great-grandson of Grand Duke Michael (1832–1909) (son of Emperor Nicholas I) and his sons (Fig. 3).

The samples from all living descendants were collected with their informed consent. The archival bloodstain specimens of Nicholas II were also available and used as described below. Photographs of the museum item are shown in Fig. S6 *D* and *E*. We do not demonstrate the photographs of the bone specimens due to ethical issues.

Analysis of DNA from the Bone Specimens. DNA was extracted in 2 primary laboratories that were not previously used for work with human DNA: in a newly equipped laboratory for ancient DNA study (VIGG) and in an especially designed isolated laboratory facility to work with low copy number DNA (UMASS MS). In addition, the extractions for a few selected samples were also replicated in a third DNA forensic laboratory (MWI). All of the procedures were performed in a sterile PCR hood. All staff wore coveralls, powder-free latex gloves, hairnets, respirators, facemasks and shoe covers. Regular decontamination of all surface areas included cleaning with bleach (1:10 diluted commercial bleach) and UV-treatment.

The procedures for DNA extractions varied slightly for different samples and in different laboratories. Generally, DNA was isolated from ≈170–750 mg of a bone sample. Bones were purified mechanically, e.g., by Micromot 50/E drill (Proxxon) followed by chemical decontamination in a series of 50-mL tubes containing diluted 1:3 commercial bleach, MilliQ H_2O, 0.5% SDS, 20% Ethanol and HPLC grade water (Chromasolv, Sigma). Cleaned bones were ground in a liquid nitrogen grinding mill (SPEX Freezer/Mill), in a RETSCH Mixer Mills MM 200, or in a blender with a minicontainer.

Pulverized bone material was collected into sterile plastic tubes containing solution 5 mL of 0.5M EDTA and 50 μL of 20–100 mg/mL proteinase K (Sigma) with or without SDS and incubated overnight at 56 °C under rotation in Isotemp* Hybridization Incubators (FisherScientific) or in other thermoincubators with shakers. After overnight incubation, lysis supernatant was concentrated on Amicon Ultra 4–30kD columns (Millipore) and 250–500 μL of concentrated sample was finally purified using QIAquick PCR purification kit (Qiagen). For some extractions, the Amicon concentration procedure was omitted. DNA was eluted from the QIAquick columns in H_2O or TE buffer. In parallel, extraction negative/blank controls were

prepared following identical extraction procedures to detect possible contamination.

A quantitative DNA analysis performed by Quant-iT PicoGreen Assay (Invitrogen) detecting the amount of total DNA and by a human specific DNA quantification kit (Plexor HY assay) (Promega) suggested that non-human microbial DNA is a significant component of DNA samples extracted from the bone specimens, in particular, from bone samples from the second grave. Nevertheless, successful genotyping was performed by several mt and nuclear-DNA systems insensitive to microbial DNA contaminations. The bone specimens identified only by numbers were used for DNA extractions in 2 physically separated primary laboratories; the identifying code numbers for independent DNA extraction were assigned to specimens in a third laboratory; and the code numbers or marks were assigned to all samples prepared for sequencing or to file chromatograms keeping personal identification of each sample confidential. Thus, the technical design of the experiments was as "blind" as possible for these studies.

Despite the fact that DNA in N146, N147 and N141 specimens was degraded, the PCR fragments of mtDNA shorter than 350–400 bp were successfully generated after primary PCR from small aliquots of several DNA extracts. Based on the assumption that hundred or a thousand mt genomes occur per somatic cell, we generally used <10 pg of human DNA from the alleged specimens for multiplex or individuals mt DNA PCR analysis. In initial experiments the extracts were tested for quantity and potential contamination by external human DNA. The selected extracts were used further for a more detailed mtDNA and nuclear DNA analysis.

The amplifications of HVR1 16009 bp–16400 bp or 16009bp-16365bp and HVR2 35bp-375bp regions were performed using PCR primers producing overlapping PCR fragments. This procedure generated the 392 bp or 357 bp HVR1 and 342 bp HVR2 sequences (excluding the primer nucleotide sequences), which were used for primary analysis. The accuracy was achieved by multiple replications. Direct sequencing of PCR products for HVR1 and HVR2 obtained from different extracts of N146 and N147 samples and in different laboratories (VIGG, Moscow and UMASS MS, Worcester) produced clean chromatogram reads in most cases. The sequences generated in replicate experiments from the same specimens were identical. The PCR products were also cloned and individual clones were resequenced. The cloned sequences contained the same SNPs, which were found in total PCR product sequences. As anticipated, the cloned sequences occasionally contained rare mutations. These random mutations are expected for single molecules cloned from amplification products due to Taq-polymerase errors or postmortem DNA modifications described, e.g., for ancient DNA. For reconstruction of complete mitochondrial genomes the multiple PCR products were generated from the bone specimens and sequenced with careful control for potential contamination by external DNA or nuclear mitochondrial pseudogenes (numts). Replications and analysis of overlapped regions from independent PCR products demonstrated no nucleotide mismatches. Importantly, the complete mtDNA sequences from N146 and N147 specimens matched perfectly to complete mtDNA sequences from living maternal relatives determined in later experiments.

In addition to the amelogenin assay, the gender identification for degraded DNA extracted from the bone specimens was performed by PCR in a reduced (12.5 μL) volume with AmpliTaq Gold DNA polymerase (Applied Biosystems) for 35 cycles using the following primers: FEM4shDIR 5′-6FAM-AACAAGAAAATCTGCCTTTGTCA-3′ (6FAM fluorophore from IDT, Inc., spectral emission 520 nm) and FEM4shREV 5′-ACAGTGGGGTCTCAGCAGTAA-3′ to amplify Y-chromosome (117 bp) and X-chromosome (119 bp) DNA fragments. In our preliminary experiments the oligonucleotide primers for the FEM4 locus, along with other markers detecting X–Y loci designed in our laboratory, were validated using serial dilutions of degraded and non-degraded human DNA templates. The FEM4 marker proved to be the most efficient and reliable especially for low copy number (20–100 pg) degraded DNA templates.

The X–Y loci genotype (male) or X locus genotype (female) were confirmed by replications in multiple independent PCR amplifications for both the amelogenin and FEM4 locus; this was the standard procedure for all bone specimens reported in this study.

Y-STR profiles were obtained using AmpFlSTR Yfiler kit (Applied Biosystems) according to the manufacturer's protocol with minor modifications for degraded DNA samples. The AmpFlSTR MiniFiler PCR Amplification Kit (Applied Biosystems) and the PowerPlex S5 System (Promega), designed for analysis of degraded DNA, were used for autosomal STR profiling of the bone specimens from the first and second grave.

In addition, a more informative AmpFlSTR Identifiler PCR Amplification Kit (Applied Biosystems) was used to generate an STR profile for 15 STR markers for the bone sample 4–46 (putative Nicholas II).

As a positive control, we used 007 DNA sample (for Applied Biosystems kits AmpFlSTR Yfiler and AmpFlSTR MiniFiler) and 9947A DNA sample (for Promega PowerPlexS5 and Applied Biosystems AmpFlSTR Identifiler).

To improve the efficacy of PCR amplification for a low amount of highly degraded DNA the number of cycles was increased and the volume of reaction was reduced. PCR reactions for Y-STR or autosomal STR analyses were performed in a total volume of 12.5 or 25 μL, at 33–37 cycles, with ≈60–500 pg of input human genomic DNA. Blank DNA extraction was used as a negative control for all STR amplifications. In a primary analysis of mtDNA or nuclear DNA the extracts demonstrating a mixture of individual profiles were excluded from further analysis. The level of sensitivity for recovery of full STR profiles by the testing commercial multiloci STR kits was determined as 60–100 pg (the human diploid nuclear genome is approximately 6 pg). The lower quantitative limit of template would be difficult to assess a priori for degraded DNA. The minimal amount of human genomic DNA used for STR analysis in our analysis of the bone specimens was ≈60–500 pg. The established threshold for determination of contamination in the analysis of STR profiles was extremely stringent. Only samples with no detectable allele amplification signal in the blank DNA extraction negative control in any replicate STR amplification for that DNA extraction sample were used for this study.

Dropout and drop in alleles are potential problems in analyses of low copy number and degraded DNA. Thus, minimally, duplicates of all alleles were required to be included in a profile (15). Serial replications were made for each sample from different extracts. Homozygous loci were assumed as authentic if multiple replications of the allele for autosomal STR marker were observed in independent amplifications. Full or partial profiles were detected in replicate amplifications for the bone samples. At least 3–4 independent observations after replicate PCR amplifications were required for each autosomal STR allele to be confirmed as authentic and for autosomal STR marker to be included in analysis. Only the genotypes for STR markers, which were detected in all bone samples that meet this criteria (as shown in Fig. S5) were used for statistical calculations.

The number of independent replicate amplifications with STR kits for degraded DNA (AmpFlSTR MiniFiler, Applied Biosystems and the PowerPlex S5 System, Promega) varied for each sample. The analysis of apparently heterozygous autosomal STR loci demonstrates that the dropout of alleles occur in 20–30% cases in some bone samples or even in a higher rate for Y-STR

loci (some Y-STR alleles are relatively large). The dropout of alleles was apparent for the specimens from the second grave (particularly, for N146 sample) whereas efficacy of amplification of autosomal STR alleles for samples with relatively high DNA quantity and quality (samples from skeletons N4 and N6) was near 100%. Thus, in total, the STR genotypes were determined using data from at least 9 or more (up to 20) separate amplifications for the specimens from different skeletons.

The consistency of Y-STR and autosomal STR typing was assessed by analysis of multiple replications of allelic results from different extracts of the same bone or bloodstain (from Nicholas II shirt item) specimens.

Electrophoretic analysis was performed, in most cases, using DNA and Genetic Analyzers (Applied Biosystems). The electrophoresis data were analyzed with GeneMapper ID software v3.2 (Applied Biosystems).

To enhance a signal intensity and to reduce background for STR profiles, in some cases, we purified the multiplex STR amplification products as suggested for genotyping of low copy number DNA templates (12). For example, ≈15 μL of Yfiler PCR product for N146 specimens was purified using Qiagen's MiniElute PCR purification kit with a final elution volume 13 μL in TE buffer, then 1 μL of purified PCR product was taken for further electrophoresis using a 96-capillary 3730xl DNA Analyzer (Applied Biosystems).

The HVR sequences for the skeletons from the first grave determined in this study are consistent with the data obtained in previous studies (2, 4). The HVR sequences for Queen Victoria maternal descendants determined in this study matched to the HVR sequences reported previously for Prince Philip (2). One STR marker (TH01) tested in our study was also used by other group in the previous analysis of remains from the first grave (2). Apparently the TH01 allele, defined as 9.3 (consisting of 9 tetranucleotide repeats and a partial repeat of 3 nt) in our study, was designated as allele 10 in data reported by Gill et al. (2). This discrepancy is most likely explained by subtle differences in size of PCR fragments for alleles 9.3 and 10, thus allele 9.3 and 10 could not be resolved under the electrophoresis conditions applied in the previous study (2).

Archival Blood Specimens of Emperor Nicholas II. In 1890–1891 Nicholas II, then-heir to the throne was on an around-the-world voyage. On 11 May 1891, during his visit to Osaka, Japan, he was attacked and injured in an apparent assassination attempt. The escort policemen swung at Nicholas II's head with a saber; however the following blow was parried by Prince George of Greece and Denmark who was accompanying Nicholas II. Although the wound was not life-threatening, Nicholas II was severely bleeding and a long scar remained on the right side of his forehead.

Surprisingly, the shirt worn by Nicholas II, with traces of his blood, has been stored as a historic relic since that event, and it was recently found in archives of the State Hermitage Museum in St. Petersburg. The experts (EIR) were invited to evaluate the possibility of extracting Nicholas II's DNA from the relic.

DNA Isolation from the ≈117-Year-Old Bloodstains of Nicholas II. The usefulness of old archival dried blood specimens for DNA typing, such as those found on Nicholas II's shirt, was questionable. The relic apparently was exposed to different environmental conditions and handled by many individuals dealing with the museum items. Although it has been stored for many years in Hermitage Museum, the shirt was apparently transported and exhibited and the storage conditions during that time are uncertain. It was not ruled out that the relic had been subjected to chemical preservation, which is often the procedure for museum items, but which would destroy DNA. Environmental exposure to light (UV), heat or humidity are also conditions that rapidly destroy DNA in bloodstains. Although discrete specimens (e.g., bone) can be decontaminated by removing the outermost layer, these procedures cannot be applied to the non-discrete bloodstains mingled with fabric and other materials. The visual inspection of the shirt identified several spots with possible traces of the blood, particularly abundant on the internal part of the shirt sleeve cuff and collar. The biological material was recovered from 4 different blood spots. The same blood spot was swabbed at least 3 times, and the first swab was discarded as potentially containing external DNA contaminants.

To minimize any potential contamination problem, the DNA was isolated only from the second and the third swabs for each stain using a QIAamp DNA Mini Kit, Qiagen. The extraction was performed according to a DNA Purification from Dried Blood Spots protocol, provided by the manufacturer, except that the volume of all buffers used before washing steps and incubation times were increased. Briefly, swabs were placed in 2-mL tubes containing 360 μL of buffer ATL and incubated at 85 °C for 15 min. Then 40 μL of proteinase K was added, and after vortexing, samples were incubated at 56 °C for 20 min. Then 400 μL of buffer AL was added and samples were incubated at 70 °C for 15 min, mixed with 400 μL of ethanol and supernatants were applied to the spin columns. After standard washing steps with AW1 and AW2 buffers DNA was eluted in 35 μL of buffer AE, first elution, and 50 μL of water, second elution.

PCR amplification of mtDNA sequences and sequencing reactions were performed as described for DNA isolated from the bone specimens. However, since the quality of DNA from the bloodstains was unknown, a different set of PCR primer oligonucleotides was designed (available upon request) to amplify short 64–109 bp DNA fragments harboring positions for very rare mtDNA SNPs, which we identified in the specimen (4–46 femur) from the N4 skeleton.

The short amplicons for HVR1 and HVR2 and for the sequences harboring rare SNPs across the mitochondrial genome were efficiently generated after the primary PCR. DNA sequencing determined rare variants at positions: 315.1, 1842, 1888, 2850, 4216, 4917, 6257, 7022, 8697, 10463, 11812, 13368, 13965, 14233, 14905, 15452, 15607, 15928, 16126, 16169 in the DNA specimens from 3 bloodstain extracts in replications (up to 5 or 7 replicate PCRs for some SNPs). The sequences covering all other polymorphic sites were also determined in extracts from the bloodstains with the set of primer oligonucleotides applied for the bone sample analysis (Table 2). The mtDNA SNPs in Nicholas II's archival blood specimens were identical to those found in the bone samples from skeleton N4 and in the Nicholas II's maternal relatives (as described in Main Text). This unique haplotype was not found in the available database for complete mitochondrial genome sequences (Table S1).

The mitochondrial and nuclear STR typing showed that 1 of the DNA extracts (N 2M) contained a mixture of DNA from Nicholas II and other individuals. Three other extracts (N 1U, N 2U and N 3U) from bloodstains on sleeve cuff and collar showed no detectable contamination by external human DNA in STR analysis.

Full Y-STR and autosomal STR profiles were obtained using the AmpFlSTR Yfiler kit (Applied Biosystems) and the AmpFlSTR Identifiler PCR Amplification Kit (Applied Biosystems), the AmpFlSTR MiniFiler PCR Amplification Kit (Applied Biosystems) and the PowerPlex S5 System (Promega) according to protocols provided by the manufacturers. The dropout of STR alleles was rare in replicate amplifications for the blood stain extracts. We were impressed that the quantity and quality of the extracted DNA were sufficient for the mtDNA and nuclear DNA analysis considering the condition and age of the specimen. The data demonstrate that despite partial DNA degradation high quality uncontaminated DNA profiles can be obtained from >100-year-old archival bloodstains.

Statistical Calculations and Interpretation of the Data. Full consideration of mathematical assumptions and multiple scenarios in this complex casework is beyond the scope of this report and will be presented elsewhere. In brief, we adopted standard mathematical calculations (13–15) assuming the haploid nature of the mtDNA and Y-STR markers (13–15). We found no single mismatch in mt genome sequences or multilocus Y-profiles between samples from the human remains (putative Alexei and his sister) and the reference relatives. Thus, the interpretation of the case is developed under consideration that no mutation events occurred across generations for the indicated individuals and their maternal or paternal relatives. The calculation in the identification test was based on a likelihood ratio (LR) of two mutually exclusive hypotheses that the remains are from relatives of the Romanov family (H_1 hypothesis) or they are from unrelated randomly chosen individuals (H_0 hypothesis). For Y-chromosome and mtDNA data likelihoods were based on Y-chromosome and mt haplotype frequencies, and likelihood ratios were calculated as

$$LR = \frac{P(E|H_1)}{P(E|H_0)}$$

where $P(E|H_1) = 1$ and $P(E|H_0)$ is population frequency of the haplotype.

The probability of observing an allele or haplotype frequency (mitotype or Y-STR markers) were estimated in accordance to reported recommendations (13, 17–19).

To make a correction on database size and frequency of the haplotype, the 95% C.I. (confidence interval) upper limit of haplotype frequencies were also calculated with formulas (14) for cases when no or 1 and more haplotypes in database matched to the casework haplotype.

All population databases for mtDNA or Y- and autosomal STRs used in this study included populations (e.g., German, British, Slavic, Russian) relevant to this casework.

For comparison we used a variety of mt databases, e.g., a large Mitosearch database and our internal EUROS database, which includes many Russian, Slavic, East European and West European populations (Table S1). The SWGDAM forensic database has also been used for calculations (data not shown) but the samples in this database, collected in U.S. populations, are irrelevant to this study.

Y-STR profile frequencies in world populations were estimated using the US consolidate (www.usystrdatabase.org) and the Y Chromosome Haplotype Reference Database (YHRD, www.yhrd.org/index.html) that include many European populations and Russian cohorts. To determine average population frequencies of autosomal STR alleles we pooled non-overlapping allele frequency data for European populations from 2 large databases - ALFRED (http://alfred.med.yale.edu/alfred/) and "The Distribution of the Human DNA-PCR Polymorphisms database" (www.uni-duesseldorf.de/WWW/MedFak/Serology/database.html) and a recently published dataset for Russian populations (16). Genotype frequencies for autosomal STR loci were calculated using Hardy-Weinberg formulas. Likelihood ratio of observed match of autosomal STR profiles of the sample from N4 skeleton from the first grave and DNA isolated from the Nicholas II's bloodstains was calculated applying product rule to likelihood ratios for individual STR loci defined as $1/f$, where f is a genotype frequency. The resulting estimate was combined with likelihood ratios obtained for observed mtDNA and Y-STR haplotypes (13, 17, 18). To test the hypothesis that N146, N147 samples and specimens from N3, N5 and N6 skeletons belong to the children of the persons whose remains are designated as N4 and N7, likelihood ratios for autosomal STR loci were calculated using formulas defined for the missing person identification case when genotypes of both parents are known (13, 17, 18). For each alleged child the likelihood ratio for autosomal STR profile was combined with likelihood ratios for mtDNA or for mtDNA and Y-STR haplotypes in case of the sample N146.

Likelihood ratio (LR) test for N7 (putative Empress Alexandra) subject based on mtDNA frequency shows that it is ≈14,5 thousand (or 4,8 thousand, 95% C.I. upper limit) times (using EUROS database) or ≈35,8 thousand (or 24,2 thousand, 95% C.I. upper limit) times (using Mitosearch database) more likely that the remains belong to relative of maternal living descendants of Queen Victoria (Table S1) than to unrelated random individual.

Although analysis of complete mtDNA from the N4 skeleton (putative Nicholas II) shows 1 ambiguous nucleotide position (heteroplasmy at 16169 C/T position (Table 2) all the positions for 22 rare SNPs in mt genomes from the N4 bone sample and Nicholas II bloodstain specimens were identical.

Based upon comparison with DNA profiles from Nicholas II's bloodstain calculated likelihood ratios for individual identification of the specimen N4–46 as belonging to Nicholas II rather than to an unrelated random individual is more than hundred sextillion for autosomal STR plus Y-STR or autosomal STR plus mtDNA profiles or more than 5 hundred septillion for all 3 (mtDNA, autosomal- and Y-STR profiles) identification systems (Table S5).

Although multiple scenarios (e.g., occurrence of relatives) theoretically can be considered further for statistical interpretation the extremely big LR numbers for individual identification coupled with family relationship analysis provide sufficient evidence that the remains belong to Nicholas II, but not to any other individual including siblings (e.g., brothers). The data and conclusions obtained here are applied, of course, only to the evidentiary items (selected bone specimens) tested in this study. For example, it cannot be ruled out that some bone fragments attributed to 1 skeleton by anthropological analysis alone may, in fact, belong to other morphologically similar skeletons found in the same grave.

When calculating only the Y-STR profile, the likelihood ratio that subject N146 is a paternal lineage relative of Nicholas II rather than a random unrelated individual is 4.2×10^3 (or 1.4×10^3, 95% C.I. upper limit) (using US Consolidated Y-STR database). Calculating only mtDNA profile the likelihood ratios that either subject N146 or subject N147 are maternal relatives of Empress Alexandra are 1.4×10^4 (or 4.8×10^3, 95% C.I. upper limit) (using EUROS database) or 3.6×10^4 (or 2.4×10^4, 95% C.I. upper limit) (using Mitosearch database). Assuming both Y-haplotype and mt haplotype data and multiplying LRs, the evidence is at least 6.0×10^7 (or 6.7×10^6, 95% C.I. upper limit) (using EUROS and US Consolidated Y-STR databases) more likely that remains belong to Alexei; and 8.7×10^{11} (or $\approx 3.3 \times 10^{10}$, 95% C.I. upper limit) more likely that remains belong to 2 Romanov children (son and daughter) than to individuals unrelated to the Romanovs. From a historical perspective, no relative of the Romanov family, other than Alexei, with these genotype characteristics (Y-STR and mtDNA) can be imagined to be at that location at that time.

The profiles for STR genotypes for N4, N7, N3, N5, N6 (first grave) and N146, N147 (second grave) subjects are different demonstrating that the newfound specimens from the second grave (N146 and N147) cannot belong to individuals from the first grave. DNA profiles for N4 and N7 are not consistent with a parent-child relationship. DNA profiles of N4 and N7 are consistent with having a parent-child relationship with each profile N3, N5, N6, N146, N147. In family trio comparison the 3–6 allele mismatches were observed between each of the putative offspring and putative father (N4) or 2–6 mismatches were observed between each of the putative offspring and putative mother (N7), but no allele was detected in the putative

offspring, which was not observed in at least 1 of the putative parents (N4 and N7).

Assuming both parents are known, the likelihood ratios for identification of children based on missing person tests were calculated for N146 (Alexei): $\approx 1.4 \times 10^9$ (or $\approx 7 \times 10^8$, 95% C.I. upper limit) for mtDNA combining with autosomal STRs, and 1.6×10^8 (or 4.0×10^7, 95% C.I. upper limit) for Y-STR combining with autosomal STRs, and 5.8×10^{12} (or 9.7×10^{11}, 95% C.I. upper limit) for combination of mtDNA, autosomal STRs and Y-STRs; N147 (Maria): 1.6×10^9 (or 7.9×10^8, 95% C.I. upper limit) for mtDNA combining with autosomal STRs. The LR for other children N3 (Olga), N5 (Tatiana), N6 (Anastasia) was also extremely high varying from 8.6×10^8 (or 4.4×10^8, 95% C.I. upper limit) to 1.3×10^{10} (6.1×10^9, 95% C.I. upper limit) for mtDNA combining with autosomal STR profiles.

The family relationships showing that the data are consistent with the hypothesis that the remains belong to 5 offsprings and 2 parents in a nuclear family were also supported by estimation made by the M-FISys software version 9.09 (2008) [Gene Codes Forensics, Inc., access and assistance was kindly provided by H. Cash (Gene Codes Corporation, Ann Arbor, MI)].

For example, the LR that the autosomal STR profile for N7 (putative Alexandra, spouse of Nichols II) would be observed given the hypothesis that N4 (putative Nicholas II) is the spouse of N7 with offspring N3, N4, N5, N6, N146, N147 is $\approx 3 \times 10^7$.

Thus, in conjunction with available historical and anthropological data, the comprehensive genomic analyses presented here provide sufficient evidence for identification of remains of Emperor Nichols II (N4) and members of his family, Empress Alexandra Feodorovna (N7), Grand Duchesses Olga (N3), Tatiana (N5), Anastasia (likely N6), Maria (likely N147) and Prince Alexei (N146).

From a methodological perspective, we would like to note that while this manuscript was in preparation and under review, 2 studies reported that FLX sequencing can be used to determine the complete human mitochondrial DNA for paleo-human specimens found in cold temperature glacial or permafrost conditions (20, 21). Here, we demonstrate that accurate genomic analysis can be applied for the precise identification of archival or historical human samples. The multiplex PCR assay for amplification of degraded human DNA, as we demonstrate in this study, is particularly applicable when only very limited historical or museum archival biological materials are available. The >90–100-year-old specimens (e.g., some semiburned bone samples) were obtained from wet soil in shallow graves from continental climate conditions or as museum archive items. Importantly, the complete mitochondrial genome sequences and Y-chromosome and autosomal STR profiles can be efficiently analyzed for old, historical human specimens stored in common environmental conditions.

1. Comp Aksuchits V (1998) *Pokayanie, Materials of Government Commission* (in Russian) (Vybor, Moscow), pp 1–285.
2. Gill P, et al. (1994) Identification of the remains of the Romanov family by DNA analysis. *Nat Genet* 6:130–135.
3. Ivanov PL, et al. (1996) Mitochondrial DNA sequence heteroplasmy in the Grand Duke of Russia Georgij Romanov establishes the authenticity of the remains of Tsar Nicholas II. *Nat Genet* 12:417–420.
4. Rogaev EI (1998) Analysis of mitochondrial DNA of putative bone remains of Nicholas II and his nephew. *Pokayanie, Materials of Government Commission* (in Russian) (Vybor, Moscow) pp 171–182.
5. Hofreiter M et al.; Gill P, Hagelberg E; Knight A, et al. (2004) Ongoing controversy over Romanov remains. *Science* 306:407–410 (Letters).
6. Kurth P (1985) *Anastasia: The Riddle of Anna Anderson* (Back Bay Books, Boston).
7. Petrov V, Egorov G, Lysenko IV (1998) The escape of Alexei, Son of Tsar Nicholas II. What happened the Night the Romanov family was executed (Abrams, New York).
8. Blow MJ, et al. (2008) Identification of ancient remains through genomic sequencing. *Genome Res* 18:1347–1353.
9. Rieder MJ, Taylor SL, Tobe VO, Nickerson DA (1998) Automating the identification of DNA variations using quality-based fluorescence re-sequencing: Analysis of the human mitochondrial genome. *Nucleic Acids Res* 26:967–973.
10. Rogaev EI, Grigorenko AP (2005) Russian Federation Patent no. 2258531.
11. Gill P, Whitaker J, Flaxman C, Brown N, Buckleton J (2000) An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA. *Forensic Sci Int* 112:17–40.
12. Smith PJ, Ballantyne J (2007) Simplified low-copy-number DNA analysis by post-PCR purification. *J Forensic Sci* 52:820–829.
13. Gjertson DW, et al. (2007) ISFG: Recommendations on biostatistics in paternity testing. *Forensic Sci Int Genetics* 1:223–231.
14. Holland MM, Parsons TJ (1999) Mitochondrial DNA sequence analysis. Validation and use in forensic casework. *Forensic Sci Rev* 11:21–50.
15. Rolf B, Keil W, Brinkmann B, Roewer L, Fimmers R (2001) Paternity testing using Y-STR haplotypes: Assigning a probability for paternity in cases of mutations. *Int J Legal Med* 115:12–15.
16. Malyarchuk BA (2007) Variation of 15 autosomal microsatellite DNA loci in the Russian population. *Molecular Biology (Rus)* 41:1–4.
17. Buckleton JS, Triggs CM, Walsh SJ (2005) Forensic DNA evidence interpretation (CRC, Boca Raton, FL).
18. Forensic Biology Section Procedure Manual, Section III. (2008) *Virginia Department of Forensic Science Document 210-D400*, pp 97–105.
19. Gusmao L, et al. (2006) DNA Commission of the International Society of Forensic Genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis. *Forensic Sci Int* 157:187–197.
20. Gilbert MT, et al. (2008) Paleo-Eskimo mtDNA genome reveals matrilineal discontinuity in Greenland. *Science* 320:1787–1789.
21. Ermini L, et al. (2008) Complete mitochondrial genome sequence of the Tyrolean Iceman. *Curr Biol* 18:1687–1693.

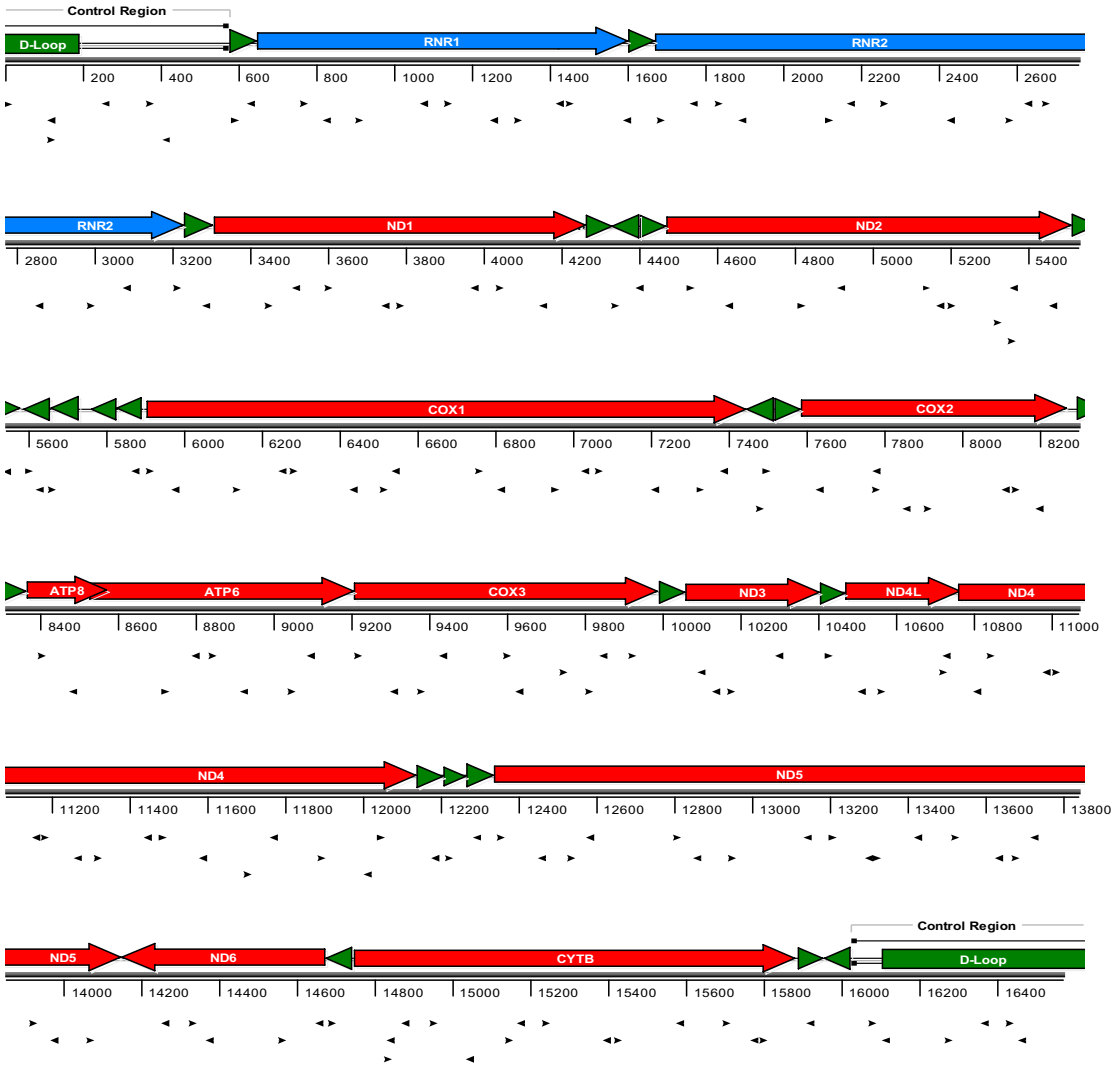

Fig. S1. Primer design for multiplex PCR to retrieve complete human mitochondrial genome sequences from degraded human DNA by small overlapping amplicons.

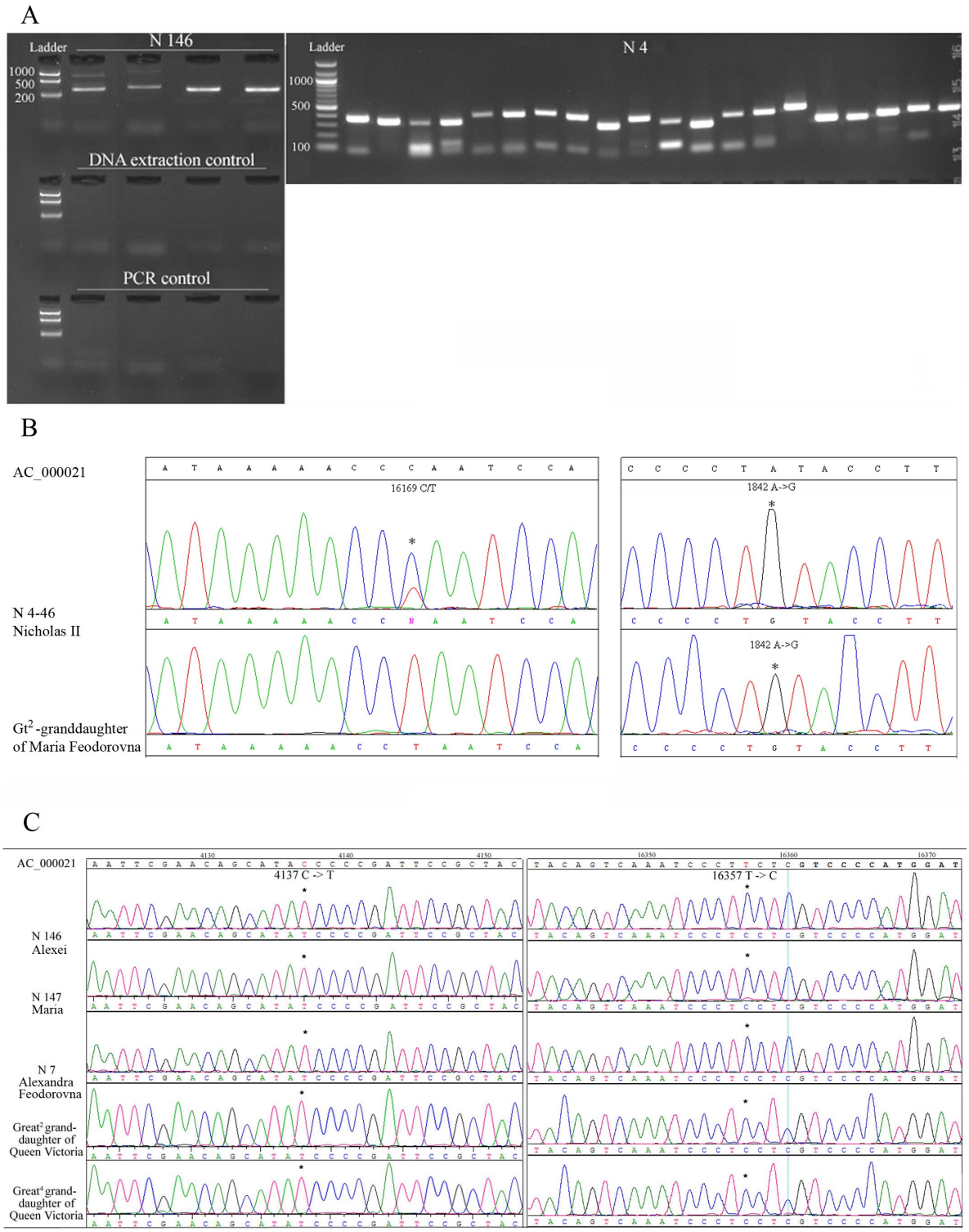

Fig. S2. Mitochondrial DNA analysis. (*A*) Examples of agarose gel electrophoresis of small amplicons in complete mtDNA amplification assay. (*Left*) Primary PCR products (sample N146) and no product for blank extraction and PCR negative controls. (*Right*) After multiplex PCR the amplification products were generated with ''individual'' pair of primers (sample N4–46 from N4, putative Nicholas II skeleton). Low molecular weight fragments represent primer dimmers. (*B*) Comparison of sequence chromatograms of heteroplasmic site in HVR1 of N 4–46 and homoplasmic site in living maternal relative of Nicholas II (left side). (*B* and *C*) Informative rare polymorphic sites in mtDNA retrieved from the tested bone specimens identical to the living relatives from the European Royal families.

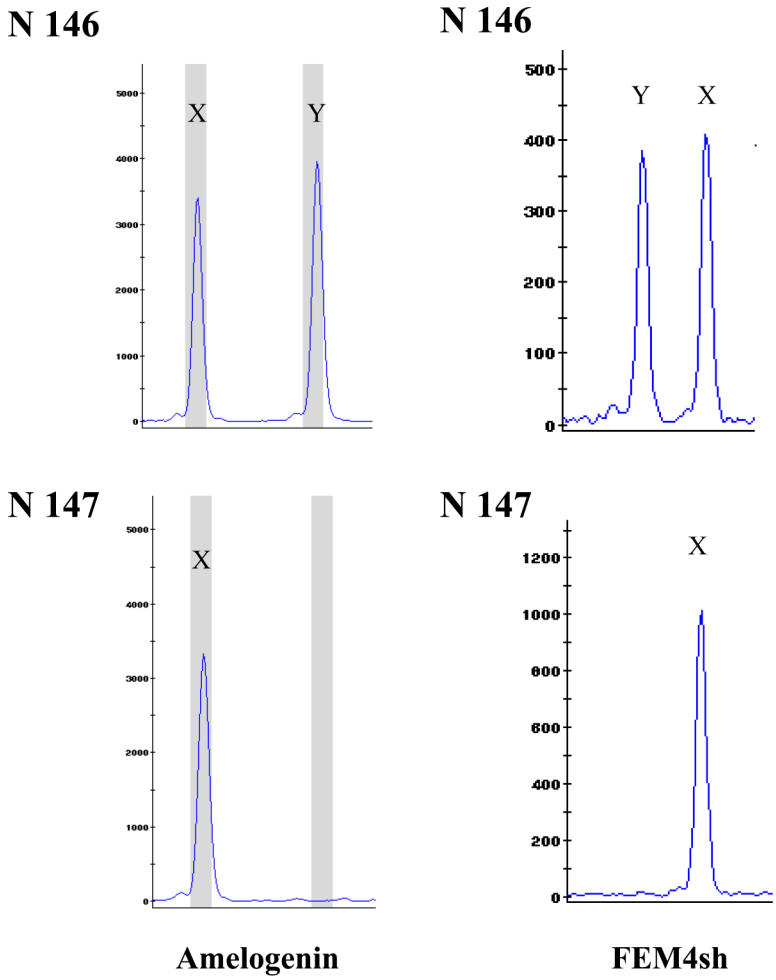

Fig. S3. Example of gender identification of the remains from the second grave with amelogenin and FEM4sh X–Y PCR primers.

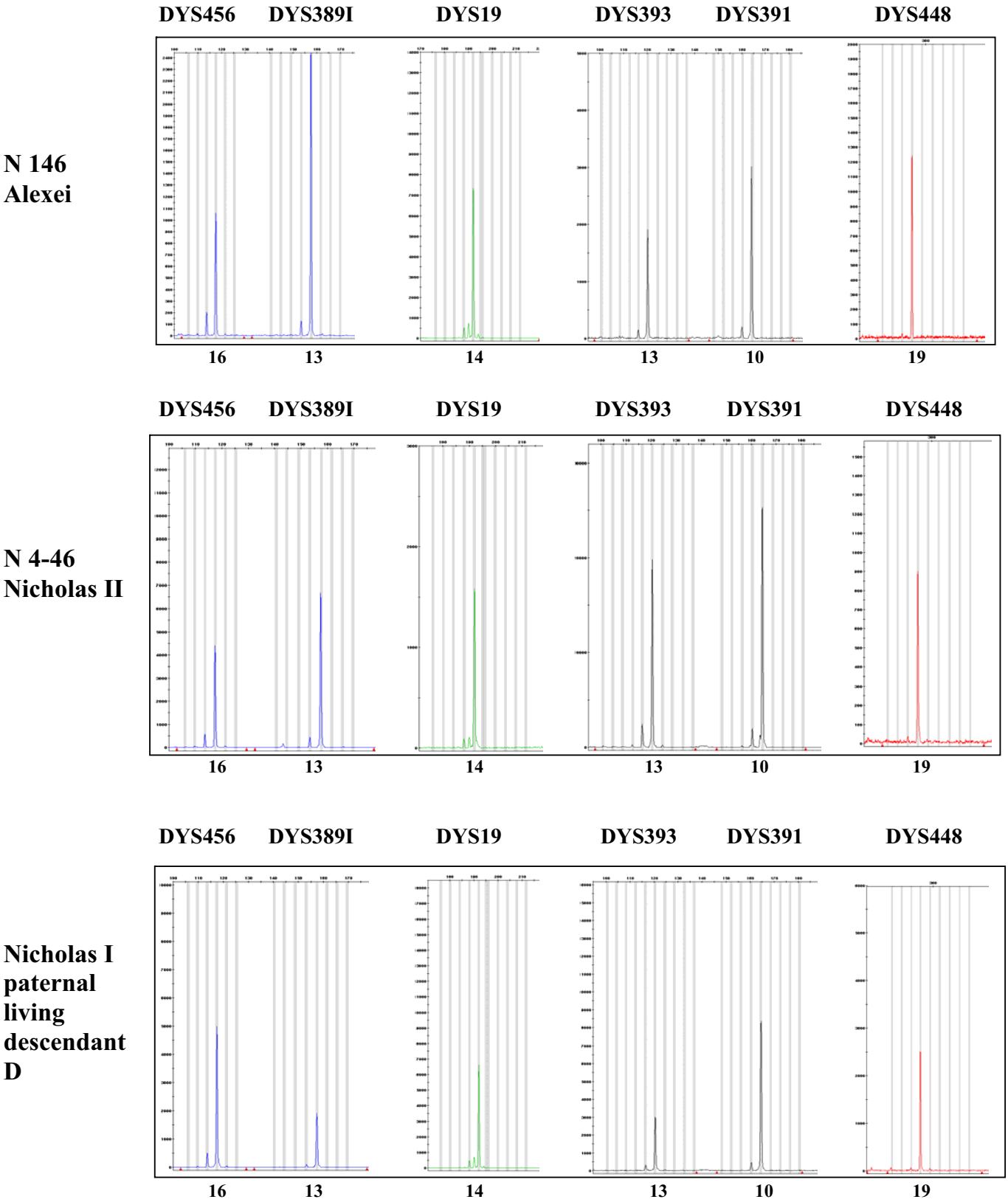

Fig. S4. Examples of Y-STR alleles matched between samples from N146 (putative Alexei) bone specimen from the second grave and N4 skeleton specimen (putative Nicholas II) from the first grave and from Romanov male descendants.

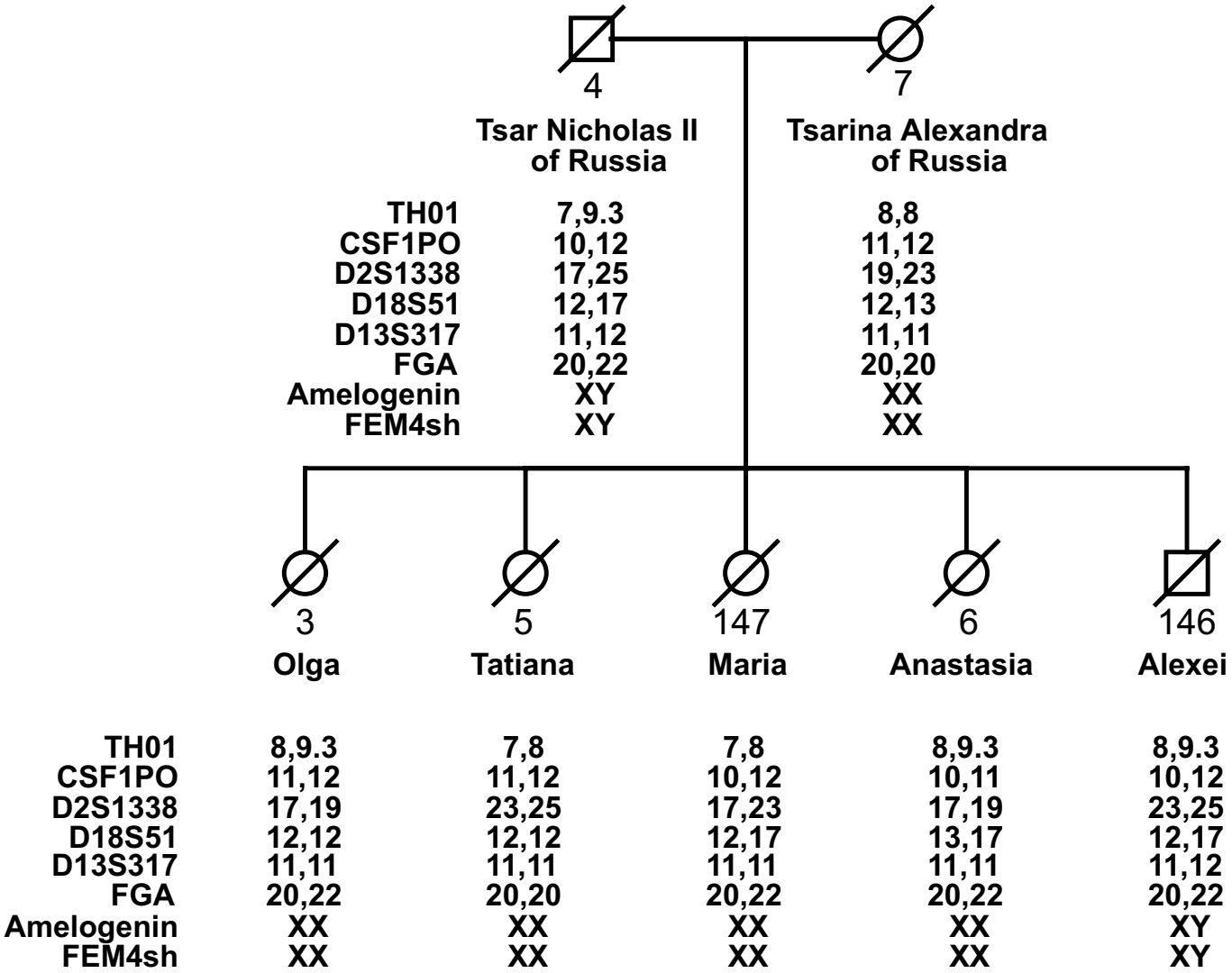

Fig. S5. Sex and autosomal STR analysis show that the remains (N146 and N147 specimens) have genotypes that are non-identical to any of the genotypes from skeletons in the first grave. The composition of genotypes of all individuals is consistent with the hypothesis of immediate family relationships consisting of 2 parents and 5 children. Independent amplifications were performed and the loci for STR alleles are shown, which were observed in several replications for all these individuals.

A

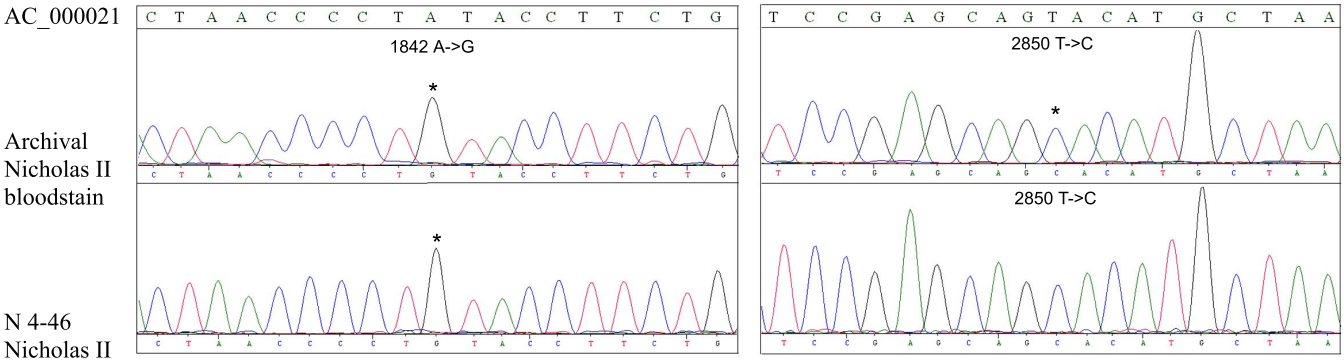

Fig. S6. (*A*–*C*) DNA typing in archival blood specimens of Nicholas II. Examples of mt-DNA rare variations (*A*), Y-STR (*B*) and autosomal STRs (*C*). (*D* and *E*) Front view of shirt (*D*) and shirt sleeve cuff (*E*) of shirt of Nicholas II, archival specimen from State Hermitage Museum in St. Petersburg.

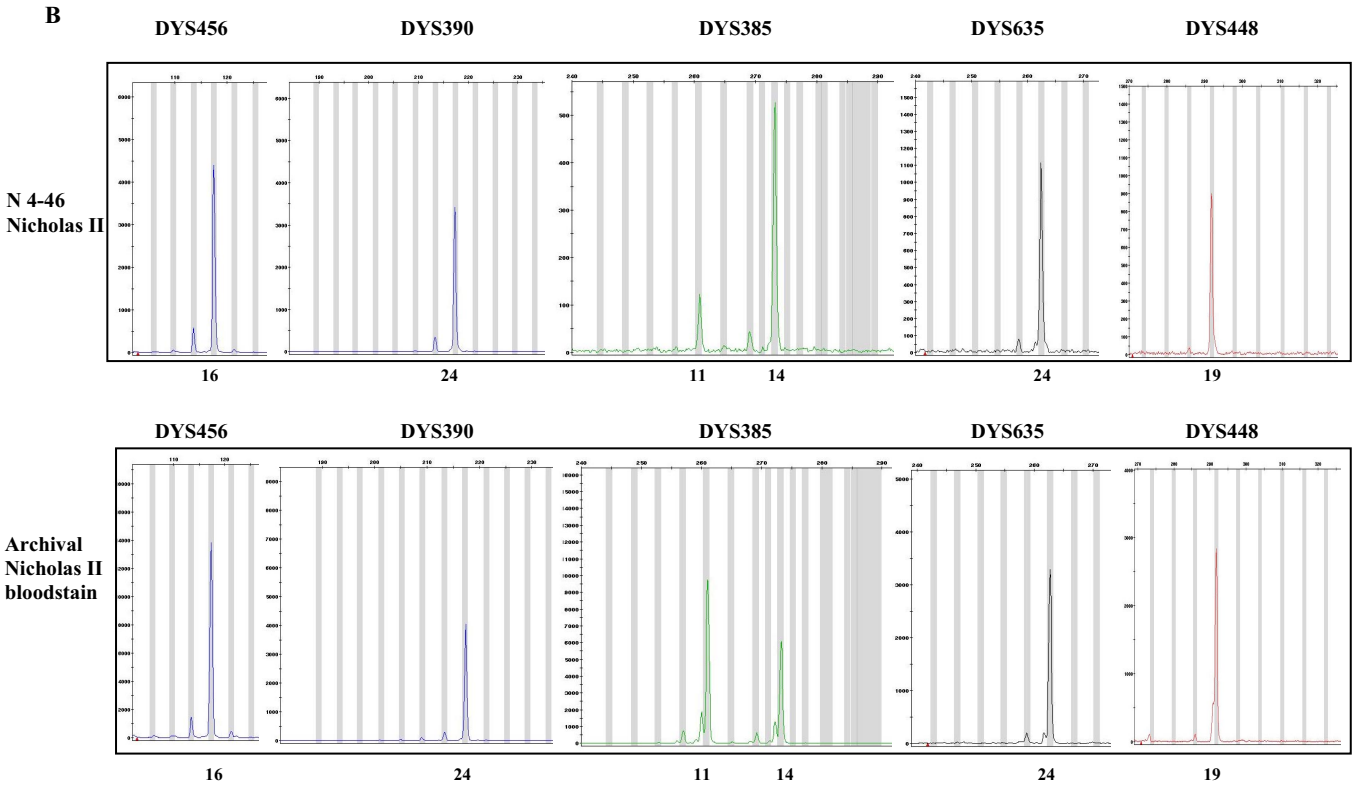

Fig. S6. (continued)

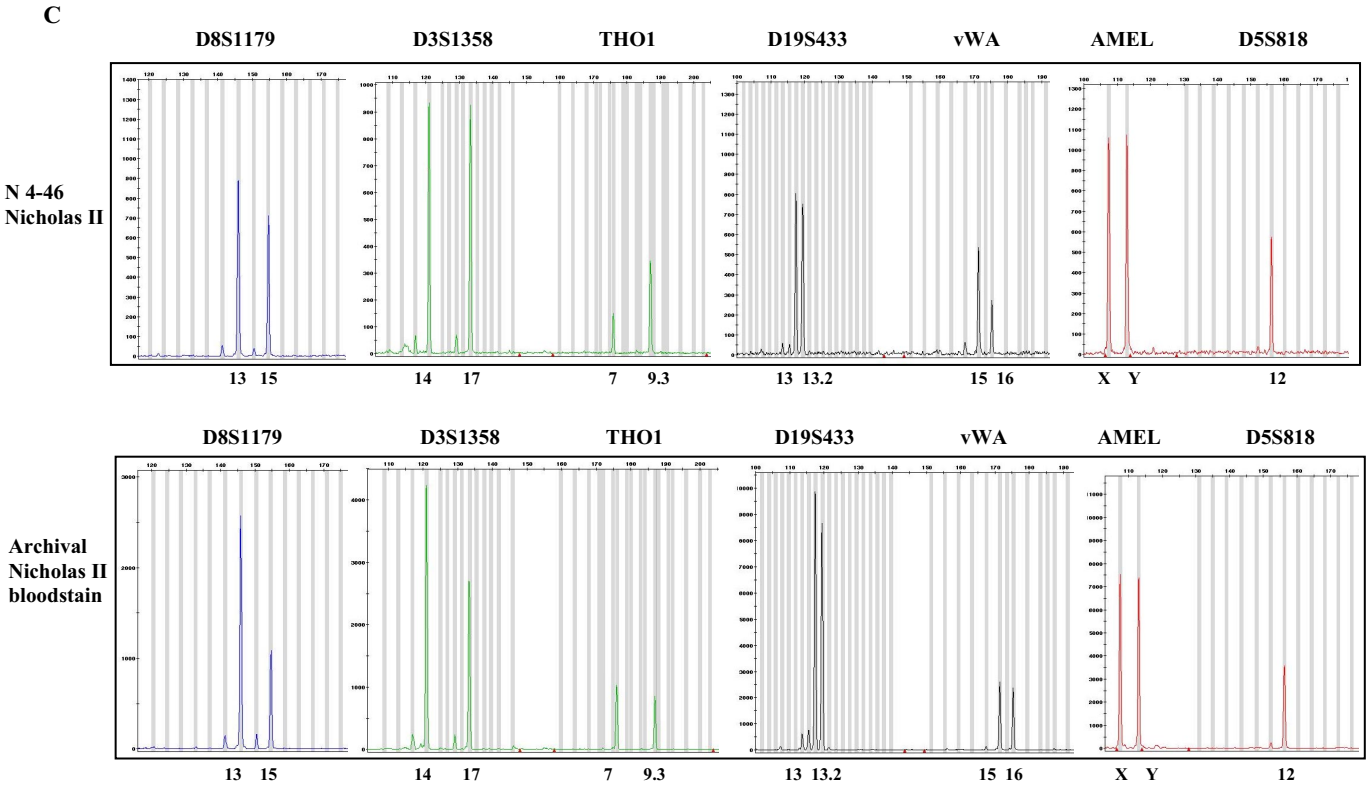

Fig. S6. (continued)

Fig. S6. (continued)

Fig. S7. Romanov family. (*A*) Romanov children: Grand Duchesses Maria, Grand Duchess Olga, Grand Duchess Tatiana, Grand Duchess Anastasia and Prince Alexei (1912). (*B*) Emperor Nicholas II (right), his wife Alexandra Feodorovna and their children with German Kaiser Wilhelm II (left) (1909). (*C*) Grand Duchess Maria Romanova (1910). (*D*) Prince Alexei Romanov (1916). The photographs from Romanov family archives were published with permission of the State Archives of the Russian Federation.

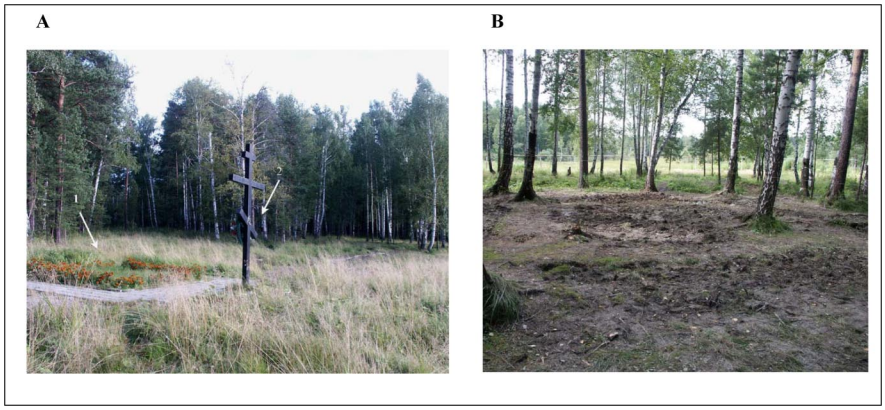

Fig. S8. Burial sites of Romanovs and their attendants near Yekaterinburg. (*A*) The grave 1 and grave 2 (located on a distance of ≈60 meters from the first grave) are designated by arrows 1 and 2 correspondingly. (*B*) The excavation site for the grave 2.

Table S1. Analysis of complete mtDNA sequences for N7, N4 skeletons and N146, N147 bone specimens

Data base	Mitotype	Number of matching mitotypes	Number of samples in Database	Mitotype frequency estimate	95% C.I. Upper Limit
		N7, N146, N147 (Tsarina Alexandra and children)			
EUROS*	16111T 16357C	0	14486	0.000069	0.000207
Mitosearch†	16111T 16357C 16519C	1	71664	0.000028	0.000041
mtDB‡	mt genome	0	2704	0.000370	0.001107
		N4, variant 16169C (Tsar Nicholas II)			
EUROS*	16126C 16294T 16296T	86	14486	0.006005	0.007188
Mitosearch†	16126C 16294T 16296T 16519C	86	71664	0.001214	0.001454
mtDB‡	mt genome	0	2704	0.000370	0.001107
		N4, variant 16169T (Tsar Nicholas II)			
EUROS*	16126C 16169T 16294T 16296T	0	14486	0.000069	0.000207
Mitosearch†	16126C 16169T 16294T 16296T 16519C	0	71664	0.000014	0.000042
mtDB‡	mt genome	0	2704	0.000370	0.001107

Frequency of mitochondrial haplotypes in population databases. We also identified the 16111T, 16357C and 16519C rare haplotypes in mtDNA from the putative remains of the children (N3, N5, and N6) of Tsarina Alexandra (N7) from the first grave.

*EUROS database includes literature data and our own data for mitochondrial HVR1 from ethnically determined populations, particularly relevant to this study: Western and Eastern Europe, Slavic populations, Russians from different geographic areas and other Eurasian ethnic groups occupying territory of Former Russian Empire (within borders of 1917).

†Mitosearch (www.mitosearch.org).

‡mtDB, Human Mitochondrial Genome Database (www.genpat.uu.se/mtDB), comprises 1,865 complete sequences and 839 coding region sequences.

Table S2. Y-chromosomal STR analysis. Frequency of "Nicholas I" Y-STR type (identical in N146 (putative Alexei), N4 (putative Nicholas II), Nicholas II archival blood specimens, and samples from Nicholas II living cousins) in population databases

Database	Compared loci	Population/ Region	Number of samples in database	Number of matches	Y haplotype frequency estimate	95% C.I. upper limit
US Cons.	A	World	4163	0	0.000240	0.000719
YHRD	B	Russia	1261	0	0.000792	0.002373

Y-STR profile frequencies in world populations were estimated using US consolidate database (www.usystrdatabase.org/) and data for Russians were extracted from Y Chromosome Haplotype Reference Database (YHRD, www.yhrd.org/index.html). A, 17 loci DYS19, DYS385ab, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS438, DYS439, DYS437, DYS448, DYS456, DYS458, DYS635 (YGATA C4), YGATA H4 or 15 loci without DYS385ab. B, 11 loci DYS19, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS385ab, DYS438, DYS439 or 9 loci without DYS385ab.

Table S3. Y-chromosomal STR profile from skeleton N4 and bloodstain specimens of Nicholas II

	N 4–46	Archival Nicholas II bloodstain (extract N 1U)	Archival Nicholas II bloodstain (extract N 2U)	Archival Nicholas II bloodstain (extract N 3U)
DYS456	16	16	16	16
DYS389I	13	13	13	13
DYS390	24	24	24	24
DYS389II	29	29	29	29
DYS458	17	17	17	17
DYS19	14	14	14	14
DYS385	11, 14	11, 14	11, 14	11, 14
DYS393	13	13	13	13
DYS391	10	10	10	10
DYS439	11	11	11	11
DYS635	24	24	24	24
DYS392	13	13	13	13
Y-GATA-H4	12	12	12	12
DYS437	15	15	15	15
DYS438	12	12	12	12
DYS448	19	19	19	19

The Y-STR genotype was confirmed in multiple replications for two extracts from the 4–46 femur. Complete Y-STR profile was determined for all three DNA extractions from bloodstain samples of Nicholas II using AmpFlSTR Yfiler system. DNA extraction negative controls (showing no amplification products) were included in all STR genotyping experiments. The positive controls (ABI control DNA or DNA from the investigators) showed different STR profiles.

Table S4. Autosomal STR profile of skeleton N4 and blood specimens of Nicholas II

	N 4–46	Archival Nicholas II bloodstain (extract N 1U)	Archival Nicholas II bloodstain (extract N 2U)	Archival Nicholas II bloodstain (extract N 3U)
CSF1PO	10, 12	10, 12	10, 12	10, 12
D2S1338	17, 25	17, 25	17, 25	17, 25
D3S1358	14,17	14,17	14,17	14,17
D5S818	12, 12	12, 12	12, 12	12, 12
D7S820	12, 12	12, 12	12, 12	12, 12
D8S1179	13, 15	13, 15	13, 15	13, 15
D13S317	11, 12	11, 12	11, 12	11, 12
D16S539	11, 14	11, 14	11, 14	11, 14
D18S51	12, 17	12, 17	12, 17	12, 17
D19S433	13, 13.2	13, 13.2	13, 13.2	13, 13.2
D21S11	32.2, 33.2	32.2, 33.2	32.2, 33.2	32.2, 33.2
FGA	20, 22	20, 22	20, 22	20, 22
TH01	7, 9.3	7, 9.3	7, 9.3	7, 9.3
TROX	8,8	8,8	8,8	8,8
vWA	15, 16	15, 16	15, 16	15, 16
AMEL	XY	XY	XY	XY

The DNA typing was performed for three independent DNA samples extracted from the bloodstains and two different DNA samples extracted from 4–46 femur. Multiplex STR kits including AmpFℓSTR Identifiler (Applied Biosystems), PowerPlex S5 (Promega) and AmpFlSTR MiniFiler (Applied Biosystems) were used in serial replications. The genotypes were confirmed in replications using all three multiplex systems.

Table S5. Likelihood ratios (LR) for individual identification of putative remains of Nicholas II

Autosomal STR*	5.03×10^{19}	(2.25×10^{19})
Y-STR	4.16×10^{3}	(1.39×10^{3})
mtDNA	2.71×10^{3}	(9.03×10^{2})
Autosomal STR* and Y-STR	2.10×10^{23}	(3.13×10^{22})
Autosomal STR and mtDNA	1.36×10^{23}	(2.03×10^{22})
Autosomal STR*, Y-STR and mtDNA	5.67×10^{26}	(2.83×10^{25})

Comparison of DNA profiles from skeleton N4 samples and Nicholas II archival blood specimens. *, LRs for 15 autosomal STR loci are presented. Likelihood ratios calculated using 95% upper limit estimates of allele and haplotype frequencies are shown in parenthesis.