

Genomic identification in the historical case of the Nicholas II royal family

Evgeny I. Rogae^{a,b,c,d,1}, Anastasia P. Grigorenko^{b,d}, Yuri K. Moliaka^b, Gulnaz Faskhutdinova^b, Andrey Goltsov^d, Arlene Lahti^e, Curtis Hildebrandt^e, Ellen L. W. Kittler^f, and Irina Morozova^a

^aDepartment of Genomics and Laboratory of Evolutionary Genomics, Vavilov Institute of General Genetics, Russian Academy of Science, Gubkina Street, 3, Moscow, 119991, Russian Federation; ^bBrudnick Neuropsychiatric Research Institute, University of Massachusetts Medical School, 303 Belmont Street, Worcester, MA 01604; ^cFaculty of Bioinformatics and Bioengineering, Lomonosov Moscow State University, Moscow, 119991, Russian Federation; ^dResearch Center of Mental Health, Russian Academy of Medical Science, Zagorodnoe Shosse 2/2, Moscow, 113152, Russia; ^eMolecular World, Inc., Thunder Bay, ON, Canada P7B 2T1; and ^fUniversity of Massachusetts Medical School, Center for AIDS Research, Worcester, MA 01605

Communicated by James D. Watson, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, November 14, 2008 (received for review October 8, 2008)

Accurate unambiguous identification of ancient or historical specimens can potentially be achieved by DNA analysis. The controversy surrounding the fate of the last Russian Emperor, Nicholas II, and his family has persisted, in part, because the bodies of 2 children, Prince Alexei and 1 of his sisters, have not been found. A grave discovered in 1991 contained remains putatively identified as those of the Russian Royal family. However, not all family members were represented. Here, we report the results of genomic analyses of new specimens, the human remains of 2 burned skeletons exhumed from a grave discovered in July 2007, and the results of a comprehensive genomic analysis of remains from the 1991 discovery. Additionally, ≈117 years old archival blood specimens from Nicholas II were obtained and genotyped, which provided critical material for the specific determination of individual identities and kinship identifications. Results of genotypic analyses of damaged historical specimens were evaluated alongside samples from descendants of both paternal and maternal lineages of the European Royal families, and the results conclusively demonstrate that the recently found remains belong to children of Nicholas II: Prince Alexei and his sister. The results of our studies provide unequivocal evidence that the remains of Nicholas II and his entire family, including all 5 children, have been identified. We demonstrate that convergent analysis of complete mitochondrial genome sequences combined with nuclear DNA profiles is an efficient and conclusive method for individual and kinship identification of specimens obtained from old historic relics.

complete mitochondrial genome | genetic identification | historical relics | human | STR profile

The genetic identification of old biological specimens, e.g., human remains from burial sites, is often limited to analysis of short degraded DNA fragments. Ancient DNA analysis techniques have been used in evolutionary analysis of long mitochondrial (mt) DNA sequences from paleontological specimens, including pre-Holocene mammalian remains (1–4). The development and application of comprehensive DNA-tests for identification of very old forensic or historical samples is of significant interest. In the early 1990s, a first grave containing human remains was found in the Ural region of Russia. Forensic investigation suggested that those remains belonged to the Romanov Imperial family of Russia, specifically Emperor Nicholas II, his wife Empress Alexandra, their children, 3 servants and a court physician, who were murdered during the Russian Civil War in 1918 (5–8). However, the remains of 2 children of the Romanov family were not identified in this investigation, and the controversy of their whereabouts remained. Legends persisted for a century that Alexei and Anastasia, the 2 youngest children of the Romanov family, survived those turbulent times. In July 2007, bone fragments of 2 burned skeletons were discovered in a second grave at a bonfire site in the same Ural region near Yekaterinburg. Forty-four bone fragments and teeth specimens were exhumed from the second grave. The bone

fragments were badly damaged by fire and presumably by sulfuric acid. Preliminary anthropological examinations of the semiburied bone fragments excavated from the shallow pit suggest that they are from a boy who was approximately between 10 and 14 years of age and of a young woman who was approximately between 18 and 23 years of age at the time of their deaths. We demonstrate here that convergent analysis of complete mt genome sequences coupled with nuclear (especially Y chromosome) profiles can be efficiently applied to historical relics for complex human kinship and individual identification.

Results and Discussion

For genetic analysis we selected 3 relatively well-preserved bone specimens (N141, N146, and N147) from the second grave. The bone samples from the putative skeletal remains of Nicholas II (N4), his wife Alexandra (N7) and 3 of their daughters (N3, N5, N6) from the first grave were also available. DNA was extracted from these specimens for this study. The maternal reference samples were collected from descendants of 2 maternal lineages, Queen Victoria (1819–1901) and Empress Maria Feodorovna (1847–1928) (known also as Princess Dagmar, daughter of Louise of Hesse-Cassel and Christian IX, King of Denmark). The paternal reference samples were obtained from male lineage descendants of Emperor Nicholas I (1796–1855) (*SI Materials and Methods, Reference Samples*). Because de novo mutations may theoretically occur in subjects separated by several generations, at least 2 genealogical paternal lineages were recruited in comparative analysis (Figs. 1–3).

Mitochondrial DNA Analysis. Despite the severe damage to the bone specimens we successfully extracted DNA from the bone fragments and sequenced the standard short hypervariable regions (HVR1 and HVR2) of mtDNA. We found that the HVR sequences from 3 bone specimens excavated from the second grave were identical to one another (N141, N146, and N147). The mtDNA sequences from the N146 and N147 samples were obtained in replications in 2 different laboratories and confirmed to be identical. We also determined the HVR sequences from skeletons N7 (putative remains of Empress Alexandra) and N3, N5, and N6 (putative remains of Romanov children) from the first grave. We found that the mtDNA sequences from the second grave match to sequences from the remains in the first grave.

For subsequent analysis, we used 2 specimens from the second grave (N146 and N147), which yielded a quantity of DNA sufficient for nuclear and complete mitochondrial DNA analysis.

Author contributions: E.I.R. designed research; E.I.R., A.P.G., Y.K.M., G.F., A.G., A.L., C.H., E.L.W.K., and I.M. performed research; E.I.R., A.P.G., Y.K.M., and G.F. analyzed data; and E.I.R., A.P.G., and Y.K.M. wrote the paper.

The authors declare no conflict of interest.

¹To whom correspondence should be addressed. E-mail: rogaev@vigg.ru.

This article contains supporting information online at www.pnas.org/cgi/content/full/0811190106/DCSupplemental.

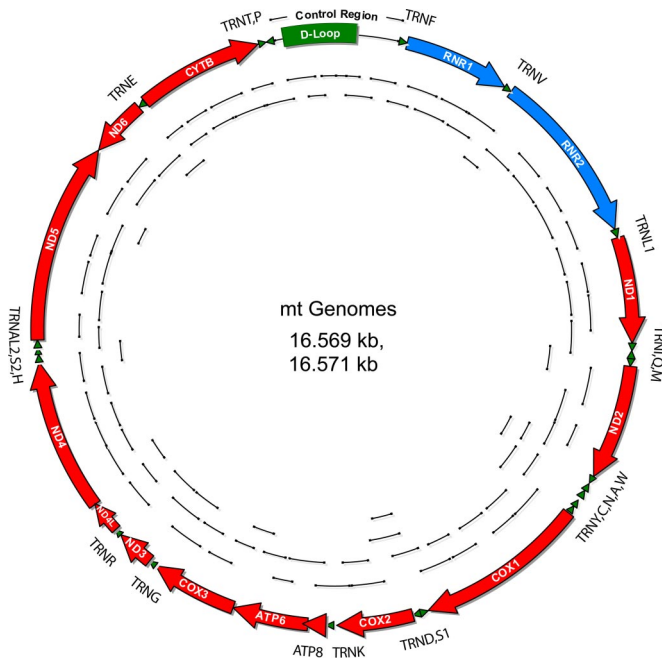


Fig. 1. Complete mt genome sequences retrieved from the putative remains of Prince Alexei (N146), his sister (N147) (second grave) and their parents, Empress Alexandra (N7) and Emperor Nicholas II (N4) (first grave). Short amplification products are shown inside of the mt genome circle. The complete mtDNA sequences from N7, N146 and N147 specimens were identical and each is 2 nt longer than the mtDNA sequence from the N4 skeleton specimen because of insertion of 2 nt in positions 524.1 (A) and 524.2 (C).

To increase sensitivity and discrimination, we developed amplification strategies for short amplicons in an attempt to retrieve complete mitochondrial (cmt) DNA sequences from the highly degraded DNA (Fig. 1, *Material and Methods*, Figs. S1 and S2 and *SI Materials and Methods*). Employing 3 different amplification approaches (*Material and Methods*), we successfully determined genomic cmtDNA sequences for N146 and N147, and the N7 specimen. Next, we determined cmtDNA sequences for living relatives from 2 different branches of the maternal lineage of Queen Victoria who is grandmother of Empress Alexandra (*Material and Methods*, Fig. 2, and Fig. S2). The resulting alignment showed a perfect match between the 16,571-bp cmtDNA sequences of N146, N147 and N7 samples and sequences of the great²-granddaughter and great⁴-granddaughter of Queen Victoria (Table 1). We designated this mitotype as “Queen Victoria” type.

In large population databases for HVR1 sequences, which include Russian, East and West European populations (Table S1), we found that this “Queen Victoria” mtDNA type (identical in putative remains of Empress Alexandra, her children and Queen Victoria’s living descendants) is very rare in human populations. In a collection of >70,000 individuals with available HVR1 data we found only 1 individual with an identical profile (Table S1). We requested and obtained DNA from this individual (M.T.), who is of Canadian origin, but with presumable ancestral origins from Germany, for a more comprehensive analysis. Additional analysis of D-loop showed mismatches for position 524.1 and 524.2 (no insertion in the tested individual, and 524.1A and 524.2C in “Queen Victoria” type mt DNA). The entire mtDNA sequences retrieved from the putative remains of Empress Alexandra and her children show no match to any individual from the population database for human cmtDNA sequences (Table S1). Thus, the mtDNA profiles found in skeletal remains N146; N147; and N7, N3, N5, and N6 (“Queen Victoria” mtDNA type) are unique (not found in available databases).

We also generated and sequenced the multiple short amplicons to reconstruct the complete mt genome sequence for putative remains of Nicholas II (N4–46 femur specimen from skeleton N4). The reference cmtDNAs were determined by sequencing of mtDNA from a maternal relative of Nicholas II. The 16,569-bp cmtDNA sequence of the N4 skeleton specimen matched to the reference sample from a maternal descendant of Empress Maria Feodorovna (Princess Dagmar), the mother of Nicholas II. We designated this mitotype as a “Dagmar” mtDNA type. There was an exception at one position for heteroplasmic mutation 16169C/T found in the N4 subject (which is also described in refs. 6–8) showing apparent rapid segregation to 2 homoplasmic nucleotide sites in this maternal lineage. The descendant of one sister of Nicholas II bears homoplasmic 16169 T (Table 2 and Fig. 2B), whereas a descendant of other sister of Nicholas II, as reported in ref. 8, showed homoplasmic 16169 C. These observations indicate that the heteroplasmic mutation apparently inherited from Empress Maria in only 1 or a few generations (Fig. 2 and Table 2). The predominant HVR1 type (with 16169C) occurs quite frequently in populations, however, the cmtDNA sequences from the putative Nicholas II remains (“Dagmar” mtDNA type either with 16169C or 16169T position) have never been observed in the available database for cmtDNA (Table S1). The data further demonstrate the biological relationship of individual N4 to the Royal maternal lineage of Empress Maria F. and provide evidence that the dual 16169C/T observed in skeleton N4 is indeed a heteroplasmic mutation (rather than contamination), which was inherited from Nicholas II’s mother and segregated rapidly to homoplasmic positions in her descendants via presumable mechanisms of biological bottleneck or reduction of mtDNA content in germ line (9, 10).

Nuclear DNA Analysis. Although the quantity of DNA from N146 and N147 samples was quite limited the nuclear DNA genotyping was also efficient in analysis of the bone specimens. The gender-identification DNA tests revealed that specimen N147 is from a female and N146 is from a male (*Material and Methods*, Table 3, and Figs. S3–S5) implying that the bone fragments could be from Alexei and 1 of his sisters. Gender-analysis was initially performed with different sets of PCR primers for a standard amelogenin gene region (*Material and Methods*). Previous reports revealed failure of the amelogenin sex determination tests in some cases. Deletions or mutations in primer binding sites for X and Y chromosome amelogenin homologs occur in human populations and may lead to gender misidentification (11–13). To avoid this potential problem, we have developed a gender-identification system adapted for the analysis of degraded DNA. The PCR primers in this system detect short DNA fragments for homologous X and Y regions located a long distance ($\approx 6-8$ Mb) from the amelogenin gene (AMELY and AMELX loci) (*Material and Methods*). Thus, the application of 2 independent gender-identification tests in multiple replications provides reliable, conclusive data. The analysis demonstrates that samples from the N147 specimen (second grave) and skeletons N7, N3, N5, N6 (first grave) belong to females and samples from specimen N146 (second grave) and skeleton N4 (first grave) belong to males (Figs. S3–S5).

There are several living descendants from unbroken paternal lineages of Russian Emperor Nicholas I (1796–1855) who must inherit from him the same Y haplotype as Nicholas II or Prince Alexei. Analysis of a nonrecombinant part of Y chromosome markers was performed with multiple replications for N146 sample and for sample from N4 skeleton in separate experiments. Y chromosome genotyping for multiple short tandem repeats (STR) loci demonstrated a match between Y-STR profiles of N146 (putative Alexei) and N4 (putative Nicholas II) individuals and the second cousins of Nicholas II derived from unbroken paternal lineages of Emperor Nicholas I (Fig. 3 and Table 3). This haplotype is unique. It was not found in the large population databases for multilocus Y-STRs (Table S2), and it is reported here. To evaluate

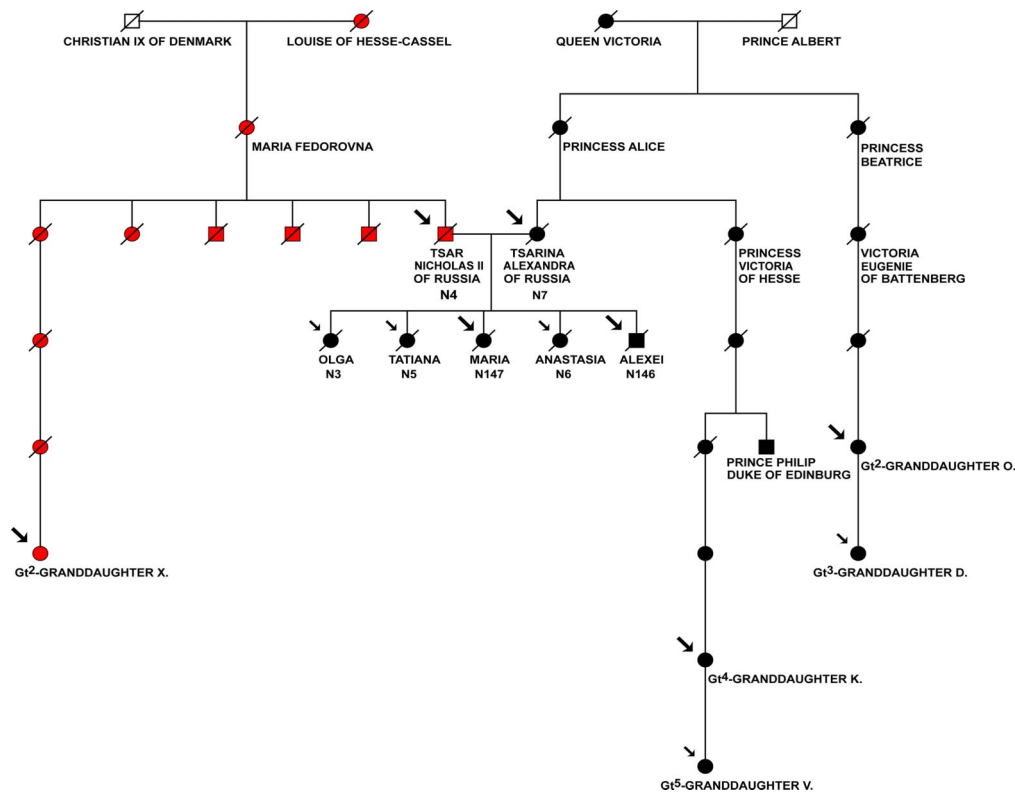


Fig. 2. Analysis of maternal lineages of Romanov family. Pedigree represents 2 maternal lineages of Romanov family. The mt haplotype of Nicholas II Romanov is inherited from Empress Maria Feodorovna (Princess Dagmar, the daughter of Louise of Hesse-Cassel and Christian IX, King of Denmark). The haplotype of Empress Alexandra Feodorovna is inherited from Princess Alice, the daughter of Queen Victoria. The complete mitochondrial genome sequences were retrieved from the bone fragments from the second grave: N146, putative Prince (Tsarevich) Alexei, and N147, putative Grand Duchess Maria; and from the first grave: skeleton N7, putative Empress (Tsarina) Alexandra Feodorovna, and skeleton N4, putative Emperor (Tsar) Nicholas II. Complete mt genome sequences or HVR sequences were also determined for the maternal relatives. Large arrows denote individuals with complete mtDNA sequences and small arrows show individuals with HVR sequences determined in this study.

the population frequency of “Nicholas I” Y-STR haplotype we used the U.S. Consolidated Y-STR Database where genotypes for 15–17 STR loci are available (www.usystrdatabase.org). This database also contains multiple European populations. No single match was observed in the Y-STR database for the 17 or 15 loci haplotypes determined in the putative remains of Nicholas II (N4) and Alexei (N146). Additionally, no identical Y-STR haplotypes were found in cohorts of individuals of Russian origin (Table 3 and Table S2).

Further gender and autosomal chromosome genotyping with STR-multiplex systems developed for degraded DNA (*SI Materials*

and Methods) showed that the male (N146) and female (N147) individuals from the second grave have genotype characteristics nonidentical to any genotype determined for Romanov family remains found in the first grave, but consistent with a biological kinship connection (*SI Materials and Methods*, Fig. S5). These data clearly indicated that the newly found bone specimens may belong to Prince Alexei and 1 of the daughters of the Imperial family.

We conclude that this nuclear DNA analysis (supported also by anthropological data) provides the following evidence: There are remains of a female (N147) and male (N146) from a second grave; the second grave specimens are not from skeleton N7 (putative mother, Empress Alexandra) or from skeleton N4 (putative father, Emperor Nicholas II); however, they are related through paternal and maternal lineages.

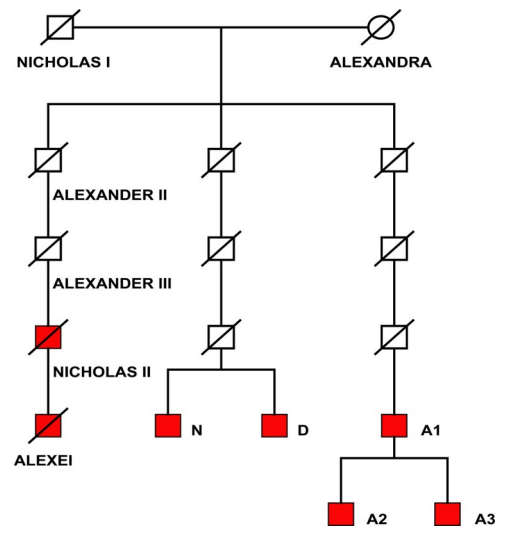


Fig. 3. Analysis of paternal lineages of Romanov family. Y profiles were determined for putative remains of Alexei and Nicholas II and their living cousins of unbroken male lineages of Nicholas I (shown in red) (*SI Materials and Methods*).

DNA Analysis of Archival Nicholas II Bloodstain Samples. An excellently preserved shirt that archival records indicate belonged to Emperor Nicholas II with presumable traces of his blood was located in the archives of the Hermitage Museum in St. Petersburg (*SI Materials and Methods*). The shirt was kept as a family historical relic after Nicholas II survived an assassination attempt in Japan in 1891. Thus, there was an intriguing opportunity to increase the discriminating capacity of DNA testing by a direct comparison of DNA profiles determined for skeleton N4 and the Nicholas II archival bloodstains. However, the contamination by foreign DNA and DNA degradation are common problems highly anticipated for an ≈117-year-old archival specimen.

We extracted DNA in 2 laboratories (Vavilov Institute of General Genetics and University of Massachusetts Medical School) from samples collected from the archival bloodstain items. The samples were taken from at least 4 different locations on the shirt (*SI Materials and Methods*).

We initially tested mtDNA sequences, specifically, HVR1 and HVR2, and highly informative regions harboring very rare mt SNPs across mt genome, which we previously identified from the putative

Table 1. Analysis of complete mtDNA sequences: Single-nucleotide polymorphisms (SNPs) found throughout the entire mitochondrial genome sequence retrieved from the remains N146 (putative Alexei), N147 (putative Maria), N7 (putative Empress Alexandra Feodorovna), and distant Queen Victoria maternal line relatives

NN	SNP position	CS	N 146	N146 em [†]	N 147	N 7	V1	V2	Frequency	Location
1	263	A	G	G	G	G	G	G	0.997	D-loop
2	315.1	—	C	C	C	C	C	C	0.961	D-loop
3	524.1	—	A	A	A	A	A	A	<i>0.014</i>	D-loop
4	524.2	—	C	C	C	C	C	C	<i>0.014</i>	D-loop
5	750	A	G	G	G	G	G	G	0.992	12S rRNA
6	1438	A	G	G	G	G	G	G	0.969	12S rRNA
7	3010	G	A	A	A	A	A	A	0.203	16S rRNA
8	4137	C	T	T	T	T	T	T	<i>0.001</i>	ND1
9	4769	A	G	G	G	G	G	G	0.989	ND2
10	8860	A	G	G	G	G	G	G	0.998	ATPase6*
11	15326	A	G	G	G	G	G	G	0.994	Cytb*
12	16111	C	T	T	T	T	T	T	<i>0.019</i>	D-loop
13	16357	T	C	C	C	C	C	C	<i>0.013</i>	D-loop
14	16519	T	C	C	C	C	C	C	0.597	D-loop

Rare SNPs (<10%) are indicated in italics. The frequencies of single nucleotide polymorphisms were obtained from mtDB-Human Mitochondrial Genome Database (www.genpat.uu.se/mtDB) or, for insertion/deletion polymorphisms in HVR regions, from population database (www.fbi.gov/hq/lab/fsc/backissu/april2002/miller1.htm). V1 - Great² granddaughter of Queen Victoria; V2 - Great⁴ granddaughter of Queen Victoria; CS—Cambridge sequence, AC_000021. Positions different from revised Cambridge sequence (rCRS) are shown.

*Nonsynonymous Thr -> Ala substitutions.

[†]Whole-genome amplification in emulsion PCR.

Nicholas II bone samples (as described above) (*SI Materials and Methods*).

Remarkably, we have found the same rare mt SNPs in these blood samples (Table 2). One DNA sample extraction (2M) contained the mtDNA “Dagmar” type SNPs anticipated for Nicholas II, but also other mtDNA variants. In addition to Nicholas II DNA, external DNA from other individuals was detected contaminating the stain (the ratio of “Dagmar” type to contaminated DNA types was estimated as at least 50%). Three other independent DNA extractions (1U, 2U, 3U) from blood-stains on shirt collar and sleeve cuff showed no signs of contamination and contained only 1 unique mitotype that matched to the mtDNA profile of skeleton N4 (4–46 sample) and a maternal relative of Nicholas II (Fig. 4, Fig. S6A, and Table 2). Interestingly, the heteroplasmy 16169 C/T was found in the blood specimens identical to the bone specimens illuminating that the heteroplasmy is not a tissue specific phenomenon and persists in a relatively similar proportion (predominant C and minor T variants) in somatic cells in different tissues of Nicholas II.

Next, given the relatively high quality of the extracted DNA, we attempted to determine autosomal and Y-STR profiles for the blood samples. We determined full profiles for 17 Y chromosome markers and 15 autosomal STR marker profiles in multiple replications in 3 different extracts (Table 3, Tables S3 and S4, Fig. S6 B and C, and *SI Materials and Methods*). The genetic profiles perfectly aligned to the STR profiles in samples from N4-46 specimen from skeleton N4 (putative Nicholas II). Thus, we were able to perform a highly discriminating calculation of matches between the reference sample (blood specimen) and the alleged sample (the bone specimen). If one combines the autosomal- and Y-STR profiles, likelihood ratio that the bone sample belongs to Nicholas II rather than a random unrelated individual is on the order of sextillions (>10²³). If one were to combine all 3 (including also mtDNA) genotyping systems the likelihood ratio is on the order of septillions (*SI Materials and Methods*, Table S5).

Interpretation of the Data for the Samples N146 and N147. The sufficient evidence is presented for the authenticity of the Emperor Nicholas II remains (skeleton N4) and for his spouse Empress Alexandra Feodorovna (skeleton N7) (as described above). The bipaternal connection of both the female and male from the second grave to 2 individuals from the first grave (father Nicholas II and

mother Alexandra) is confirmed by STR analysis (Fig. S5). The data support the hypothesis that previously unaccounted for remains of the daughter and son of Nicholas II and Alexandra have been found. The mtDNA sequences are identical in specimens for N3, N5, N6 and N147 and N146 (putative children) and N7 (putative mother) and this mtDNA type is extremely rare in populations (Table S1). The autosomal STR profiles were different for each individual in this specimen group. The STR profiles for male N4 individual and female N7 individual are consistent with having parent-child relationship with each N146, N147 and N3, N5, N6 individuals. Consideration of the autosomal STR, Y-STR and mtDNA profiles resulting from the genotyping demonstrates the biological kinship connection as one of immediate family: mother (N7), father (N4), 4 daughters (N3, N5, N6 and N147), and son (N146) (Fig. S5).

Likelihood estimations show that it is >10⁸ or even 10⁹ more likely that newly found remains and remains from the first grave belong to the Romanov’s children than to random individuals unrelated to the Romanov family (*SI Materials and Methods*). Taken together, our genotyping data establish beyond reasonable doubt that the remains of the last Russian Emperor, Nicholas II Romanov, his wife Empress Alexandra, their 4 daughters (Grand Duchesses Olga, Tatiana, Maria and Anastasia), and their son (Crown Prince Alexei) (Figs. S7 and S8) have been identified. Thus, none of the Nicholas II Romanov family members survived the massacre.

Materials and Methods

DNA Extraction from Old Biological Specimens. DNA was initially extracted in 2 independent, isolated, newly equipped laboratories (in Vavilov Institute of General Genetics and University of Massachusetts Medical School) that were not previously used for work with human DNA and dedicated especially to work with ancient DNA. For replication, the DNA extractions were repeated in a third forensic DNA laboratory (Molecular World, Inc.). The experimental procedures were performed in a sterile PCR hoods under standards for ancient DNA work with regular decontamination measures and all precautions to avoid a risk of contamination by modern DNA molecules.

Initially, DNA extractions and primary mitochondrial DNA profiles were obtained from the bone samples from the second grave [N141, N146 (originally labeled 146.1), N147 specimens] and later from the skeleton specimens from the first grave [N4, N7, N3, N5, N6 (putative remains of Nicholas II, his wife Alexandra and their daughters Olga, Tatiana and Anastasia correspondingly)] and reference samples. The blank DNA extraction negative controls were used in all extraction and genotyping experiments. At least 3–4 independent DNA extractions generated from each of the N146 (left femoral bone fragment) and N147 (right femoral

Table 2. Analysis of complete mtDNA sequences: Single-nucleotide polymorphisms (SNPs) found throughout the entire mitochondrial genome sequence retrieved from the N4 skeleton samples (putative Nicholas II) and descendant of Empress Maria (Princess Dagmar) of Hesse-Cassel maternal line

NN	Position	CS	N 4	MF1	MF2	Frequency	Location
1	73	A	G	G	G	0.834	D-loop
2	263	A	G	G	G	0.997	D-Loop
3	315.1	—	C	C	C	0.961	D-Loop
4	709	G	A	A	A	0.164	12S rRNA
5	750	A	G	G	G	0.992	12S rRNA
6	1438	A	G	G	G	0.969	12S rRNA
7	1842	A	G	G	G	0.002	16S rRNA
8	1888	G	A	A	A	0.053	16S rRNA
9	2706	A	G	G	G	0.805	16S rRNA
10	2850	T	C	C	C	0.004	16S rRNA
11	4216	T	C	C	C	0.090	ND1*
12	4769	A	G	G	G	0.989	ND2
13	4917	A	G	G	G	0.048	ND2*
14	6257	G	A	A	A	0.006	COI
15	7022	T	C	C	C	0.004	COI
16	7028	C	T	T	T	0.813	COI
17	8697	G	A	A	A	0.047	ATPase6
18	8860	A	G	G	G	0.998	ATPase6*
19	10463	T	C	C	C	0.047	tRNA Arg
20	11251	A	G	G	G	0.087	ND4
21	11719	G	A	A	A	0.777	ND4
22	11812	A	G	G	G	0.033	ND4
23	13368	G	A	A	A	0.049	ND5
24	13965	T	C	C	C	0.006	ND5
25	14233	A	G	G	G	0.034	ND6
26	14687	A	G	G	G	0.008	tRNA Glu
27	14766	C	T	T	T	0.774	Cytb*
28	14905	G	A	A	A	0.051	Cytb
29	15326	A	G	G	G	0.994	Cytb*
30	15452	C	A	A	A	0.087	Cytb*
31	15607	A	G	G	G	0.055	Cytb
32	15928	G	A	A	A	0.049	tRNA Thr
33	16126	T	C	C	C	0.089	D-loop
34	16169	C	C/T†	T	C/T†	0.990/0.005	D-loop
35	16294	C	T	T	T	0.057	D-loop
36	16296	C	T	T	T	0.024	D-loop
37	16519	T	C	C	C	0.597	D-Loop

Rare SNPs (<10%) are indicated in italics. These multiple rare SNPs and HVR1 SNPs along with 16169 C/T heteroplasmy were also detected in the archival Nicholas II bloodstain specimens. The frequencies of single nucleotide polymorphisms were obtained from mtDB-Human Mitochondrial Genome Database (www.genpat.uu.se/mtDB) or, for insertion/deletion polymorphisms in HVR regions, from population database (www.fbi.gov/hq/lab/fsc/backissu/april2002/miller1.htm). MF1, great² granddaughter of Empress Maria F; MF2, blood stain specimen from Nicholas II, son of Empress Maria F; CS, Cambridge sequence, AC.000021. Positions different from rCRS are shown. *Nonsynonymous substitutions.

†N 4 and MF2 samples display C/T heteroplasmy at position 16169 (C is a predominant form).

bone fragment) samples were used for each genotyping system. At least 2–3 independent extracts have been used for each bone specimen for genotyping remains from the first grave. Small bone fragments were cleaned with rigorous physical and chemical decontamination procedures and were ground or drilled to powder. In most cases DNA was isolated from only ~170–750 mg of cleaned bone specimens using extraction procedures with 0.5M EDTA and proteinase K and purification with silica-spin QIAquick columns (QIAquick PCR purification kit, Qiagen) with some modifications. In our previous work, this method has provided efficient DNA purification from historical bone specimens (8). DNA extraction from museum bloodstain specimens was performed as described in *SI Materials and Methods*. Quantification of human-specific DNA was performed using the Plexor HY assay (Promega) with the 7500 Real-Time PCR System (Applied Biosystems).

DNA Extraction from Modern Specimens. All manipulations with DNA from living Romanov relatives were performed in another buildings physically separated

Table 3. Y chromosome STR analysis: STR haplotypes

Marker	N 4	N 146	Romanov family members	Archival Nicholas II bloodstain	Control DNA ABI, 007
DYS456	16	16	16	16	15
DYS389I	13	13	13	13	13
DYS390	24	24	24	24	24
DYS389II	29	29	29	29	29
DYS458	17	17	17	17	17
DYS19	14	14	14	14	15
DYS385	11, 14	11, ND	11, 14	11, 14	11, 14
DYS393	13	13	13	13	13
DYS391	10	10	10	10	11
DYS439	11	11	11	11	12
DYS635	24	24	24	24	24
DYS392	13	13	13	13	13
Y-GATA-H4	12	12	12	12	13
DYS437	15	15	15	15	15
DYS438	12	12	12	12	12
DYS448	19	19	19	19	19

The Y-STR alleles were determined for bone 4–46 specimen from N4 skeleton (putative Nicholas II) from the first grave and N146 (putative Alexei) from the second grave and for males from unbroken paternal lines of Nicholas I. The STR alleles were identified in multiple independent PCR replications using at least 3 different DNA extractions. Only alleles which appeared in a minimum of two replicative assays were scored as authentic alleles. Under these conditions, a complete profile was generated from the bone sample of skeleton N4 and the archival blood specimens of Nicolas II. Low copy number highly-degraded DNA is prone to drop-out alleles for STR markers. One marker DYS385 detects two loci on the Y chromosome. High molecular weight allele (DYS385/ 14) was observed only once in replication experiments for DNA extracted from N146 sample, thus this allele for the N146 sample is formally indicated as nondetermined (ND).

from the ancient DNA laboratories. Buccal cell collection was performed using Whatman's sterile Omni swabs. Blood was spotted and dried on a sterile medical bandage in the doctor's office. DNA was purified from each sample using the



B

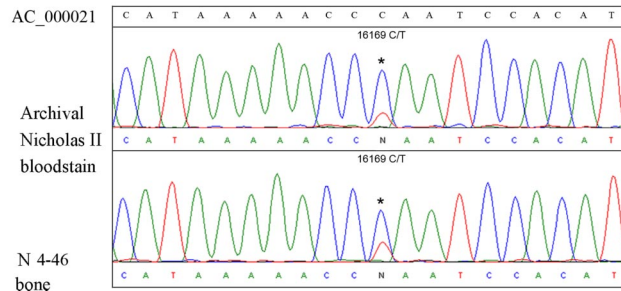


Fig. 4. MtDNA analysis of Nicholas II bloodstain specimen and specimen of skeleton N4. (A) the shirt of Nicholas II with visible traces of bloodstains used for DNA extraction. (B) Comparison of sequence chromatograms of Nicholas II archival sample and 4-46 femur; heteroplasmy was found in both bone and blood specimens.

QIAamp DNA Mini Kit (Qiagen), following the protocol supplied for DNA purification from swabs and dried blood spots. Informed Consent was obtained from all living relatives participating in the study (*SI Materials and Methods*).

Analysis of HVR1 and HVR2. MtDNA genomic sequence intervals for HVR1 from 16,009 bp to 16,365 bp or 16,400 bp, and for HVR2 from 34 bp to 375 bp were amplified as short overlapping fragments by 37 cycles of PCR in a 50 μ L of reaction volume using 1 μ L of 1:10–1:20 diluted original DNA extract from bone specimens and TaqGold DNA polymerase (Applied Biosystems) or PicoMaxx High Fidelity PCR Sytem (Stratagene), according to the manufacturer's protocol. To control for potential DNA contamination, the blank DNA extractions and blank PCR were used as negative controls throughout all tests. PCR products were purified from agarose gel by QIAquick Gel Extraction kit or MinElute Gel Extraction kit, depending on PCR efficiency. Purified PCR products were sequenced in both directions. For additional verification, the HVR1 and HVR2 products from the remains from the second grave were cloned into the pGEM-T or pGEM-T Easy vectors (Promega) in Moscow's and US laboratories according to the manufacturer's protocol. Clones containing inserts were sequenced using T7 and SP6 universal primers.

Analysis of Complete Mitochondrial Genome Sequences. Eighty-eight pairs of oligonucleotide PCR primers were designed to amplify short overlapping 164- to 383-bp mtDNA sequences covering the entire mitochondrial genome. The locations of all of the amplicons on the mitochondrial genome map are shown on the Fig. 1 and primer positions are shown on Fig. S1. These primers were used for either standard, as described above, or multiplex PCR.

To retrieve complete mtDNA from the very small amount of degraded DNA templates, we developed 3 strategies. In a first approach, a series of single PCR amplifications were performed using aliquots of very diluted human DNA extracted from the remains.

In a second approach, to increase the amount of DNA template whole-genome emulsion amplification was initially performed as described by Blow et al. (14) [technical support and protocols were provided by T. Woyke (U.S. Department of Energy Joint Genome Institute, Walnut Creek, CA)]. The GS adaptors (GS Adapter-1 and GS Adapter-2) were directly ligated to end-repaired degraded DNA and emulsion amplification of original DNA was performed with the adaptor primers. Then, aliquots of amplified total DNA were used in single PCR amplifications with specific mtDNA primers.

In another strategy, to reconstruct cmtDNA sequences multiplex PCR amplifications were used. For multiplex amplification the primers were combined in 3 sets of 25, 30, and 33 pairs for 3 multiplex PCR reactions. The reactions were performed in 50- μ L volumes for 35 cycles. After multiplex amplification, 1 μ L of purified PCR product was used as a template for a second round of PCR with each of the individual pairs of primers. PCR products were sequenced as described above. The PicoMaxx High Fidelity PCR System enzymes were used for mtDNA amplification. The sequences of the PCR primers and conditions are available upon request.

All 3 approaches worked well for the highly degraded DNA from the bone specimens. The mtDNA sequences retrieved from the bone samples by different methods (as described above) were identical in all replications. The multiple informative SNPs were confirmed for the bone specimens or Nicholas II blood-stain samples using other set of oligonucleotide primers amplifying very short sequences (64–109 bp) harboring the mt SNPs (*SI Materials and Methods*).

Complete mt genome sequences for modern specimens (collected as buccal swabs or dry blood samples) were retrieved with another set of primers for longer PCR products as described with minor modifications (15). DNA sequences were assembled into contigs (contiguous sequences) using Seqman software, DNAS-TAR, using rCRS sequence (accession number AC.000021) as a scaffold. All PCR amplification reactions were performed on GeneAmp PCR System 9700 (Applied Biosystems).

Sex Identification. Gender was primarily identified using amelogenin primers included in commercial STR multiplex assays [AmpFISTR MiniFiler (Applied Biosystems) and PowerPlex 55 (Promega)] and separate amelogenin primers kindly provided by Promega. We have also designed a pair of primers that are able to efficiently detect alternative genomic regions on the X and Y chromosomes on degraded low-copy DNA template. FEM4 locus is located at a distance of \approx 8 Mb from the AMELX gene on the X chromosome and at a distance of \approx 6 Mb from the AMELY gene on the Y chromosome (16). The primers were designed to amplify short DNA fragments on the Y (117 bp) and X (119 bp) chromosomes for this locus (*SI Materials and Methods*). All gender results for the bone specimens were independently confirmed using different kits and sex-determination systems.

Nuclear STR Analysis. Y chromosome and autosomal chromosome STR-profiles for the historical and archival specimens were obtained with commercial multiplex systems (Promega and Applied Biosystems) and modifications for analysis of degraded DNA (17) (*SI Materials and Methods*). The criteria for analysis and interpretation of low copy number DNA profiling were applied and, the potential problems for contamination, imbalanced STR allele picks and drop-out alleles (17) were carefully addressed in this study (*SI Materials and Methods*).

Statistical Interpretation. Statistical interpretation was based on standard mathematical calculations (18–20) of likelihood ratios (LR) assuming the haploid nature of the mtDNA and Y-STR markers (17–19) and that no mutation events, as observed from our genotyping data, occurred in generations for alleged individuals and their maternal or paternal relatives (*SI Materials and Methods*).

ACKNOWLEDGMENTS. The primary analysis and conclusions for mtDNA HVR profiles from the newfound remains were made by the authors in December 2007 to January 2008. We thank B. Malyarchuk for information on population data for mtDNA, T. Andreeva for technical support, V. Gromov and N. Nevolin for assistance with the specimens, G. V. Valenbachov and coworkers of Hermitage museum for access to archival items, S. Nikitin and S. V. Mironenko for assistance with photographs, Promega and Applied Biosystems for reagents, H. Cash for discussion and all participants and relatives who donated the samples for this study.

- Cooper A, et al. (2001) Complete mitochondrial genome sequences of two extinct moas clarify ratite evolution. *Nature* 409:704–707.
- Krause J, et al. (2006) Multiplex amplification of the mammoth mitochondrial genome and the evolution of Elephantidae. *Nature* 439:724–727.
- Rogaev EI, et al. (2006) Complete mitochondrial genome and phylogeny of pleistocene mammoth *Mammuthus primigenius*. *PLoS Biol* 4:403–410.
- Gilbert MT, et al. (2007) Whole-genome shotgun sequencing of mitochondria from ancient hair shafts. *Science* 317:1927–1930.
- Comp Axsuchits V (1998) *Pokayanie, Materials of Government Commission* (in Russian) (Vybor, Moscow), pp 1–285.
- Gill P, et al. (1994) Identification of the remains of the Romanov family by DNA analysis. *Nat Genet* 6:130–135.
- 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.
- 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.
- Cree LM (2008) A reduction of mitochondrial DNA molecules during embryogenesis explains the rapid segregation of genotypes. *Nat Genet* 40:249–254.
- Gocke CD, Benko FA, Rogan PK (1998) Transmission of mitochondrial DNA heteroplasmy in normal pedigrees. *Hum Genet* 102:182–186.
- Thangaraj K, Reddy AG, Singh L (2002) Is the amelogenin gene reliable for gender identification in forensic casework and prenatal diagnosis? *Int J Legal Med* 116:121–123.
- Chang YM, Burgoyne LA, Both K (2003) Higher failures of amelogenin sex test in an Indian population group. *J Forensic Sci* 48:1309–1313.
- Shadrach B, Commane M, Hren C, Warshawsky I (2004) A rare mutation in the primer binding region of the amelogenin gene can interfere with gender identification. *J Mol Diagn* 6:401–405.
- Blow MJ, et al. (2008) Identification of ancient remains through genomic sequencing. *Genome Res* 18:1347–1353.
- 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.
- Rogaev EI, Grigorenko AP (2005) Russian Federation Patent no. 2258531.
- 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.
- Gjertson DW, et al. (2007) ISFG: Recommendations on biostatistics in paternity testing. *Forensic Sci Int Genetics* 1:223–231.
- Holland MM, Parsons TJ (1999) Mitochondrial DNA sequence analysis. validation and use in forensic casework. *Forensic Sci Rev* 11:21–50.
- 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.