IDENTIFICATION OF THE REMAINS OF THE ROMANOV FAMILY BY DNA ANALYSIS

Gill, Ivanov, Kimpton, Piercy, Benson, Tully, Evett, Hagelberg & Sullivan

Authors

- **Peter Gill**
  - Began his research into DNA in 1985
  - In 1993-1994 he lead the team that confirmed the identity of the remains of the Romanov family
  - Has published more than 180 papers
  - Leads the International Society of Forensic Genetics DNA commission
  - Currently a professor of Forensic Genetics at University of Oslo, Norway

Authors

- **Pavel Ivanov**
  - Molecular biologist
  - Assistant professor of medicine at Harvard Medical School
  - Director of Forensic DNA Unit of the Federal Center of Forensic Medical Expertise (FCFME)
- **Colin Kimpton**
  - PhD in Molecular Biology
  - Won the 1997 International Society for Forensic Haemogenetics Scientific Prize
  - Works on improving the quality, sensitivity and interpretation of DNA profiling

Tsar Nicholas

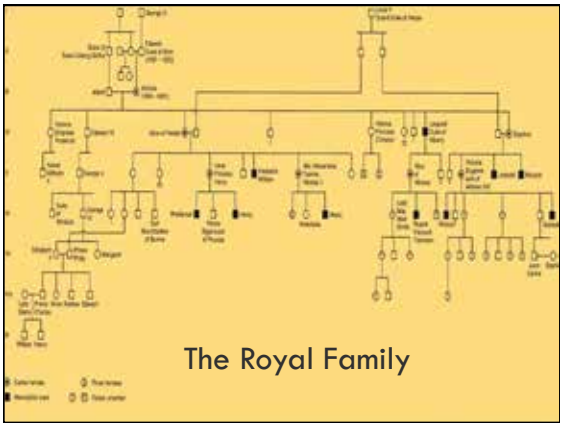
- Tsar Nicholas II (Nikolay Alexandrovich Romanov)
- Last Emperor of Russia
- Ruled from November 1st 1894 to March 2nd 1917
- Forced to abdicate following the February Revolution
- After which, he and his family were imprisoned and eventually executed

Romanov Family

- Tsar Nicholas married Alexandra Fedorova (Tsarina) in 1894
- Had five children
  - A son
    - Alexi
  - Four daughters
    - Olga
    - Tatiana
    - Maria
    - Anastasia

Hemophilia

- An inherited x-linked recessive disease
- Causes improper blood clotting or coagulation
- Is a result of a deficiency in the coagulation factor VIII
- This trait was introduced into the royal family due to a sporadic mutation in Queen Victoria
- It was then passed on to her granddaughter Tsarina Alexandra and her son Alexi



The Royal Family

Execution

- Romanov's were held prisoner at Ipatiev House in central Russia
- Sometime after July 16th, 1918, the Tsar, Tsarina, their four daughters and one son were put into a cellar with three servants and their family doctor, Eugeny Botkin
- The Bolshevik firing squad shot everyone in the cellar and those that were not killed were supposedly stabbed to death

Disposal of the Bodies

- The plan was to then dump the bodies down a mine shaft but, due to mechanical difficulties with the truck, the bodies were put in a quickly dug grave
- The bodies were covered with sulfuric acid to make them unidentifiable and the truck was driven over the grave to flatten it

History

- This information was provided by Nikolai Sokolov, an investigator that collected evidence about that family
- It is not known for certain if this series of events are completely accurate

Discovery of bodies

- Two historian investigators, Gely Ryabov and Alexander Avdonin, used the information from Sokolov to find the grave where the Romanov's were buried
- The Russian government took over the investigation and found nine skeletons within the grave that all had evidence of abuse

Determining Identity

- Some of the bodies had gold, platinum and porcelain dental work which was only seen in aristocrats at that time
- Through age estimation, sexing, odontology and computer facial reconstruction they were able to determine that the remains were from the Tsar, Tsarina and three of their five children
- In 1992 the authors of this paper were approached by the Russian government to use DNA testing to determine without a doubt the identity of the remains

DNA Extraction and Quantitation

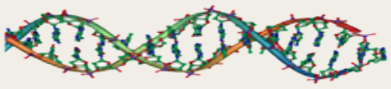
- The DNA was taken from the bones by a method called Hagelberg & Clegg
- The outer surfaces of the bone were removed by a flap wheel sander attached to a high speed electric drill
- The remaining bone was frozen with liquid nitrogen and then ground into a fine powder
- The powder was incubated overnight with EDTA, Proteinase K, and Tween
- The mixture was extracted with ethanol and chloroform and was then centrifuged for an hour
- The samples were washed with water and centrifuged for another hour
- The DNA extracts were quantitated by hybridization with a human-specific DNA probe kit

Sex Testing

- Sex determination was carried out by using an amplified portion of the X-Y homologous gene amelogenin, which is responsible for the development of the tooth enamel
- This method was used as it is effective even on highly degraded genetic material
- They took 20 picograms of DNA from each skeleton which was then amplified through 39 cycles of polymerase chain reactions
- The PCR products were put into an agrose gel and electrophoresis was performed to separate the products
- X products with 106 base pairs and Y products with 112 base pairs were generated
- With the use of DNA from the nine samples along with analysis of the bones they were able to determine that four males and five females were present in the grave

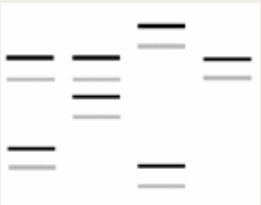
DNA

- Most of our DNA is identical to that of others
- Inherited regions vary from person to person
- Variations are called polymorphisms
- Polymorphisms are due to a different number of copies of the repeat element that can occur in a population



Short Tandem Repeats (STRs)

- Short Tandem Repeats are usually considered "junk DNA"
- The number of times a DNA sequence is repeated for a given STR is different in every individual
- For this reason STR's are often used for forensic studies



Short Tandem Repeat Analysis

- A molecular biology method
- Compares specific loci on DNA from two or more samples
- Measures the number of repeating units
- Uses a polymerase chain reaction (PCR)
  - Generates thousands of copies of the DNA sequence
  - Probes are attached to regions on the DNA and PCR determines the length of the short tandem repeats

Short Tandem Repeat Analysis

- In this experiment 5 loci were analyzed:
  - HUMTH01
  - HUMVWA31
  - HUMF13A1
  - HUMFES/FPS
  - HUMACTBP2
- These loci were amplified from each of the nine skeletons

Short Tandem Repeat Analysis

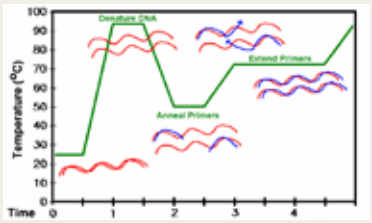
- 20-40 picograms of template DNA were added to each amplification reaction
- Each locus was amplified individually in a 50 μl reaction volume that contained 1x PARR buffer, 1.25 U Taq polymerase, 200 μM dNTPs and 0.25 μM of each primer

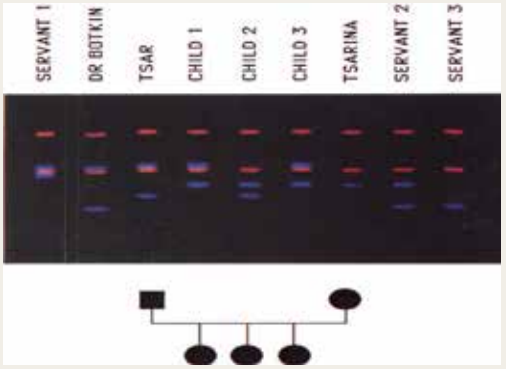
Short Tandem Repeat Analysis

- The DNA was heated at 94°C for 45s (denaturing), cooled to 54°C for 60s (annealing) and heated to 72°C for 60s (extending)
  - HUMVWA31: 38 cycles
  - HUMTH01: 35 cycles
  - HUMF13A1: 43 cycles
  - HUMFES/FPS: 40 cycles
  - HUMACTBP2: 38 cycles
- The fragment sizes were then determined

How DNA is copied in a PCR

http://www.youtube.com/watch?v=2KoLnIwoZKU&feature=c4-overview-vl&list=PL854F3A5D92AFC121





Short Tandem Repeat Analysis

- Amplification of genomic DNA at levels of 50 picograms or less of heterozygotes may completely fail
  - For this reason each sample was analyzed a minimum of four times.
  - Homozygotes were analyzed a minimum of six times
- Error: occasionally the amplification of HUMF13A1 and HUMFES/FPS resulted in additional artificial bands
  - The artificial bands were of constant size and occurred outside of the expected allele range
  - True allele bands were not affected

Table 1 STR genotypes[a] for the nine skeletons

Skeleton	HUMVWA/31	HUMTH01	HUMF13A1	HUMFES/FPS	HUMACTBP2
1 (servant)	14,20	9,10	6,16	10,11	ND
2 (doctor)	17,17	6,10	5,7	10,11	11,30
3 (child)	15,16	8,10	5,7	12,13	11,32
4 (Tsar)	15,16	7,10	7,7	12,12	11,32
5 (child)	15,16	7,8	5,7	12,13	11,36
6 (child)	15,16	8,10	3,7	12,13	32,36
7 (Tsarina)	15,16	8,8	3,5	12,13	32,36
8 (servant)	15,17	6,9	5,7	8,10	ND
9 (servant)	16,17	6,6	6,7	11,12	ND

[a]Allele designation for all loci except HUMACTBP2 is based on the number of repeat units (determined by sequencing of specific alleles — data not shown). The allele designation for HUMACTBP2 is based on an arbitrary scale identical to that of Kimpton *et al.*[2].

Short Tandem Repeat Analysis

- If these remains were of the Romanov family then the STR analysis indicates that one princess and Tsarevitch Alexei were missing from the grave
  - This aligns with historical accounts of two bodies being burned, buried separately or the possibility that two individuals survived the massacre

Laboratory Organization

- To obtain the best results possible all extractions were performed in a laminar flow cabinet
  - A carefully enclosed bench is used to avoid contamination
- This cabinet was set up in a dedicated lab with equipment used solely for this experiment
- Amplification was performed in yet another lab so the amplified material never entered the extraction lab

Laboratory Organization

- Negative controls where bone powder was not added to the extraction mixture was used in each experiment
  - If a negative control gave a signal the experiment was subsequently rejected
  - This rarely occurred
- Sequenced data results were confirmed by sequencing the data in both directions

mtDNA

- mtDNA is "a small, circular molecule"
- Can range from 16-18,000 base pairs in most vertebrates
- mtDNA genome is self replicating
- Passed on from mother to offspring through the maternal egg cytoplasm
- The sperm mitochondria does not contribute
- Has found much use in evolutionary as well as population genetics

mtDNA Analysis

- In mitochondrial DNA there are two hyper variable regions - these are used in mtDNA testing
- HVR1 is a low resolution region while HVR2 is high resolution
- In this experiment both hyper variable regions were sequenced in all samples
- The exception to this was skeleton nine
  - Nucleotides 16216 to 16360 were sequenced manually from a single strand
- Duplicate extractions were performed on the 'family group' and no differences were shown to be present between samples of a single individual

Samples Taken

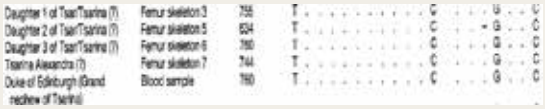
- All samples were extracted from bone DNA
- HVR1 was determined from 380 nucleotides in the first region
- HVR2 from 360 nucleotides
- Quality of these samples were deemed to be comparable to that of a fresh blood sample
- Comparisons of sequenced data showed six different mtDNA sequences were present in the group of nine skeletons
- These sequences varied on average by six nucleotides

Amplification

- The amplification of mtDNA extracted from the bones was as follows:
  - Two rounds of Polymerase Chain Reaction from 50 picograms of extracted DNA
  - The first round of amplification occurred over 30 cycles of PCR
  - Next 4 pairs of internal primers were used in second amplification

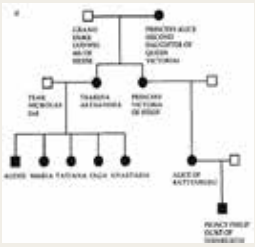
Tsarina

- The sequences of the Tsarina and three children were identical
- HRH Prince Phillip the Duke of Edinburgh, the great nephew of Tsarina, provided a sample of blood
- Gave an identical sequence to that of the Tsarina and children
- Confirms the identity of Tsarina and her three children



Tsarina

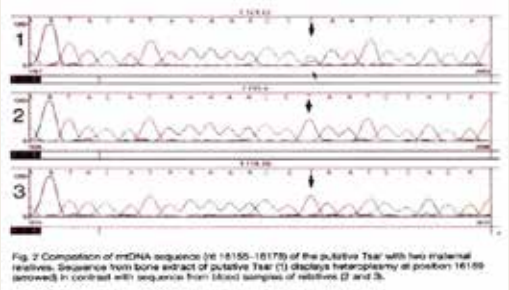
- Pedigree depicting the relation between Prince Phillip and the Tsarina
- This relation is of unbroken maternal lineage



Tsar

- mtDNA was extracted from the skeleton
- Compared with mtDNA of two relatives of unbroken maternal descent of his grandmother, Louise of Hesse-Cassel
- Gave the same sequence with the exception of one nucleotide
  - Situated at 16169
- Tsar determined to have heteroplasmy
  - When more then one exact sequence of DNA is present
  - Some cells show a mutation while others do not
- C:T at a ration of 3.4 :1 which was estimated from peak areas

Heteroplasmy



Heteroplasmy

- Observed in the duplicate extractions performed on skeleton four presumed to be the Tsar
- Confirmed this observation by sequencing individual recombinant clones using PCR, followed by transformation to E. coli
- Of the sequenced clones, 28 showed cysteine at 16169 and seven showed tyrosine giving a 4:1 ratio
  - This ratio was similar to the estimated ratio from Figure 2

Heteroplasmy

- At the time this paper was written no data was available on the rate of heteroplasmic mutations becoming fixed in a maternal lineage
  - Or on the frequency of heteroplasmy in this region
- Background signals of many sequencing techniques mask all but the most pronounced heteroplasmy
- The mtDNA sequences of the Tsar's mother and grandmother were unavailable, thus they could not determine if either showed the heteroplasmy as well
- Conflicting evidence had been shown regarding rate of conversion of heteroplasmy to homoplasmy
- Following this, the authors planned to undertake an extensive study of maternal lineages in humans

Analysis

- Short Tandem Repeat analysis evidence showed that it was likely that skeletons three to seven were related (showed STR patterns that would be expected in a group of related individuals)
- It also gave evidence that the skeletons of individuals four and seven were most likely the parents, with skeletons three, five, and six being those of the children
- The STR analysis however cannot directly prove that the family was in fact the Romanov family
- In order to decipher whether or not this was the Romanov family mtDNA was compared with that of known relatives

Statistical Analysis

- This led to two fairly obvious possibilities:
  - The group was in fact the Romanov family
  - The group is an unrelated group of people
- It was also possible that this was a family group that was related to the Romanovs
  - This possibility was ignored due to historical evidence denying any possible related family with the same ages and sexes
- The possibility that this was a hoax was also ignored since there was no realistic scenario where this would happen

Statistical Analysis

- A final possibility that was disregarded was that there was some type of corruption or contamination, which would cause impure results.
- This was disregarded because:
  - Samples were taken from two different bones at different times, yet yielded the same result
  - Bone samples were amplified and extracted blindly, and in different labs
- Finally, the two chosen sequences differed by one base pair only while the average between two random people of the same race is eight base pairs

Statistical Analysis

- mtDNA from the Tsar was found to be heteroplasmic
- This complicated the experiment as they then had to calculate the probability that a mutation occurred in the Tsar himself
- Bayesian Inference was used for interpretation purposes
  - P(E/R)= probability of evidence if group is Romanov family
  - P(E/R')= probability of evidence if group is unrelated group

Statistical Analysis

- First calculation: Likelihood Ratio LR=p(E/R)/p(E/R')
  - Numerator represents the probability of there being a single mutation in the Tsar and no mutations in generations between the Tsar and the rest of the relatives
    p=(3.2x10^-3)
  - Denominator represents the probability of two random unrelated people varying by a single base pair (using white Caucasian data base of 100 individuals)
    p=(4.8x10^-5)
- This first calculation concluded that it was 70 times more likely for the Tsar to have had the mutation than it was for the skeletons to be unrelated yet only differ by a single base pair

Statistical Analysis

- The second calculation assumed that no mutation occurred in the Tsar or any relatives: LR=p(E/R)/p(E/R')
  - Numerator represents probability that no mutations occur in the Tsar or the family
    p=0.95
  - Denominator was calculated from the observation of chance matches seen In a database
    p=1.2x10^-5
- This calculation showed that it is 80,000 times more likely for R to occur!

Discussion

- When victims are discovered many years after murders or disasters identification can be problematic
- mtDNA analysis:
  - Identity can be determined from living descendants even when separated by several generations
- Short tandem repeat analysis:
  - Can be used to determine relationships between parents and siblings

Discussion

- Anthropological evidence gives information of the probability of the remains being those of the Tsar's family
  - There was evidence that the family was aristocratic
    - Fillings were gold, porcelain and platinum
  - All the bodies had wounds
  - The bodies were of the correct age and sex
  - Found in the correct location

Discussion

- This study was the first of its kind where both mtDNA and STR were used in a major historical investigation as identification tools
- This study shows it is almost certain that the remains were those of the Romanov royal family with the exception of two of the Romanov children

98.5%

Probability that the remains are those of the Romanov family

THE END!

Presented by

Kristie Boyd
Adam Collins
Stephanie Ellsworth
Alana Kavanagh

References

- http://www.mun.ca/biology.scarr/MGA2-05-17.html
- https://www.mun.ca/biology/scarr/mtDNA_genome.html
- http://www.biology.arizona.edu/human_bio/activities/blackett2/str_description.html
- http://www.forensicdnacenter.com/dna-str.html
- http://www.genebase.com/learning/article/63
- http://www.ncbi.nlm.nih.gov/genome/probe/doc/TechPCR.shtml
- http://www.mun.ca/biology/scarr/PCR_simplified.html