DNA Analysis and the Shroud of Turin: Development of a Shroud CODIS

Kelly P. Kearse

Abstract

Since its development in the mid 1980s, DNA analysis has become a standard procedure utilized by law enforcement and legal systems in the forensic examination of human remains, and to help establish or exclude a connection to a crime scene. The recent progression of gene amplification and enrichment strategies, together with next generation sequencing techniques, have made the analysis of ancient and degraded DNA samples much more feasible than previously imagined. Human DNA has been isolated from the Shroud of Turin, although the results remain rather limited and controversial. Indeed, it is unknown if such DNA truly originates from blood cells present on the cloth or is the result of contamination from exogenous sources. Here, the potential and limitations of modern molecular biology techniques in the analysis of the Shroud of Turin are reviewed, including the evaluation of both nuclear and mitochondrial DNA.

Organization of the human genome: DNA structure and function

Deoxyribonucleic acid, or DNA, encodes the information important for the construction of all living things, from simple bacteria to human beings (1,2). Structurally, DNA consists of four bases, or nucleotides, which are symbolized by the letters A,T,C,G. These four bases are the building blocks of DNA. Similar to varying the arrangement of letters of an alphabet to create different words, which are then joined together to form sentences, variation in the number and sequence of the four bases (A,T,C,G) determines the identity and function of specific DNA regions. For example, the DNA sequence TTCGGCCAT would encode a different set of instructions than the sequence CTAGTGTCC or the sequence ATCCTTGCG, and so forth.

Essentially two types of DNA sequences exist: (i) genes, which encode specific products necessary for certain cellular functions; and (ii) noncoding segments, that are localized between genes and do not encode specific products. Together, genes and noncoding segments represent a set of instructions that is necessary for the creation and function of a particular organism. This genetic blueprint is referred to as a genome.

Sequencing of the entire human genome was completed in 2003, and showed that human DNA contains only 20,000-30,000 genes, a much smaller number than was originally estimated (3,4). Genes account for only about 2% of total human DNA, with the remainder being noncoding sequences. The purpose of noncoding DNA remains to be determined; certain noncoding sequences may function to regulate the expression of specific genes. Other noncoding sequences may be nonessential, i.e. “junk DNA” that serves no identifiable purpose (2-4).

Within a cell, both genes and noncoding sequences exist together on DNA strands that are wound around each other in the form of a double helix (Figure 1). Such strands are said to be complementary because the “A” base always pairs with the “T”

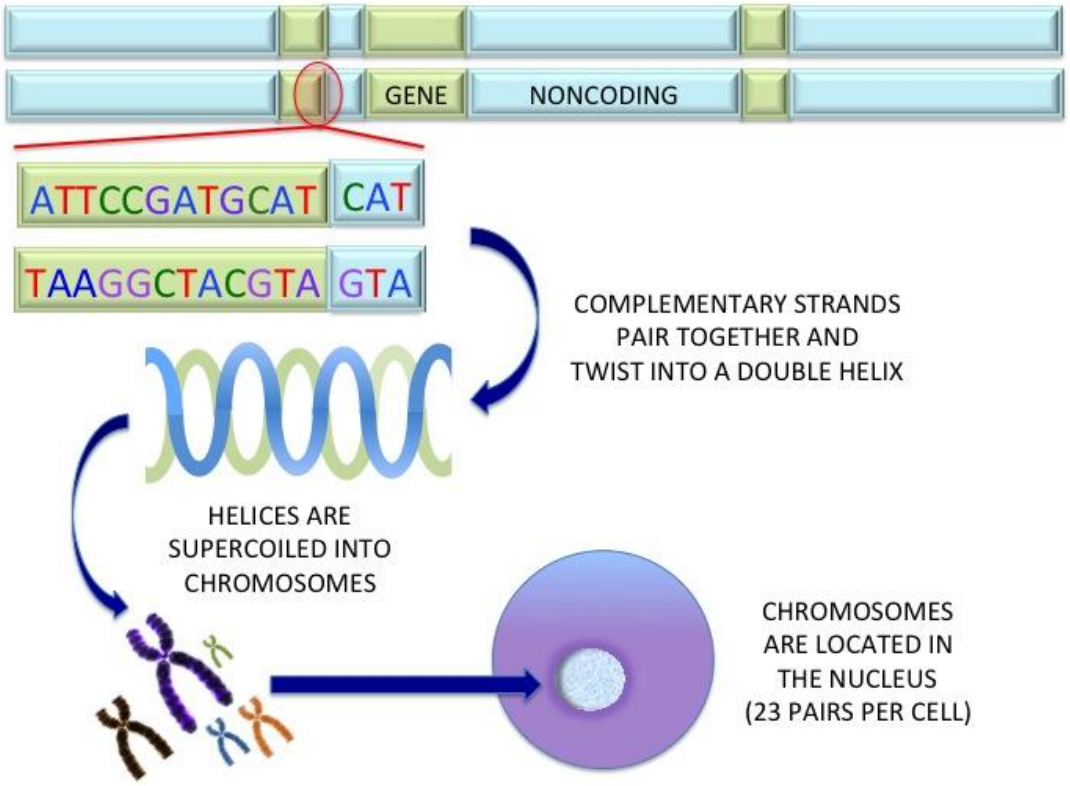

Figure 1. Organization of DNA within the cell. *Coding segments (genes) and noncoding segments exist on complementary strands of DNA. Complementary strands twist together to form double helices, which, in turn, are supercoiled together into chromosomes, localized in the nucleus (see text for details).*

base on the opposite strand. Similarly, "G" and "C" bases are always paired together. Thus, if one knows the sequence of one strand, the order of the complementary strand may be inferred. Helices are supercoiled around each other, similar to wires being twisted together, and packaged into chromosomes (Figure 1). Twenty three unique pairs of chromosomes exist in the nucleus of all cells in the body, except for two cell types: sex cells (sperm or egg), which have only half of this amount (23 chromosomes total, but no pairs); and red blood cells, which lose their nucleus as they mature and therefore are devoid of DNA (1,2,5).

DNA profiling and evaluation of specific DNA regions of interest

The vast majority of DNA (99.9%) is the same in all human beings, with only one-tenth of 1% of total DNA being unique for each individual (with the exception of identical twins). However, because of the vast size of the human genome (approximately 3 billion base pairs in length), such a small difference is sufficient to distinguish with a high level of confidence if the DNA in question originates from the same person or a different individual (1,2). Elucidation of the DNA sequence of a

single region, in of itself, is not that discriminatory; however, when many such regions are examined together, such analysis can establish the identity and relationship of DNA samples with a relatively high power of discrimination. This is the basis of DNA testing, or DNA profiling (1,2,6).

Standard (nuclear) DNA profiling involves the examination of highly variable regions found among the 23 pairs of chromosomes that are known as short tandem repeats, or STRs. STRs are short nucleotide sequences (~4 base pairs in length) that are repeated numerous times in a row in noncoding DNA regions scattered throughout the genome. A core of 13 STRs, plus the amelogenin gene to distinguish X and Y chromosomes, is currently utilized in the combined DNA index system (CODIS), a national DNA database that is kept on file by the Federal Bureau of Investigation (FBI), (6-8). By comparing a battery of specific STR sequences between one DNA sample and another, and considering their frequency among various populations, their relatedness can be determined with a relatively high probability. The greater the number of STR sequences that are examined, the higher the discriminatory power in determining the relationship of the samples in question. STR CODIS analysis is one of the most recognized standards by which an individual may be excluded or included in a particular case of interest.

Sequences other than STRs may also be evaluated to determine which form(s) of a particular gene is (are) present, for example, the ABO blood group genes (blood typing), and the human leukocyte antigens (tissue typing). Numerous companies offer genetic analysis services to the public (for a fee) to determine ancestry relationships among family members; such DNA tests as similar in principle, but are outside of the typical STR/CODIS analysis (6-8).

Mitochondrial DNA versus nuclear DNA

Typically, when DNA analysis is discussed in the media, it is nuclear DNA which is being referenced as this constitutes the bulk of the human genome. However, in certain cases additional DNA sequences may be evaluated that are located outside of the nucleus in a cellular organelle termed the mitochondrion (Figure 2). Mitochondria perform numerous cellular functions, most notably the generation of cellular energy. The mitochondrial genome is much smaller than that of the nuclear genome, containing only 37 genes total, and unlike nuclear DNA, which is wound into chromosomes, mitochondrial DNA exists in the form of a small, circular closed loop (Figure 2), (1, 9-13). Mitochondrial DNA is composed of the same building blocks (the bases A,T,C,G) that are used to create nuclear DNA. While only one nucleus exists within a cell, hundreds to thousands of mitochondria may be present, each of which can contain several copies of mitochondrial DNA. Thus, unlike nuclear DNA, which exists as only 1 set (2 copies) per cell, 100-10,000 copies of mitochondrial DNA may be present (Figure 2), (1, 9-12). This property makes mitochondrial DNA analysis particularly useful in situations where sufficient nuclear DNA may not be available for evaluation, for example in aged samples where DNA degradation is a concern (9-12).

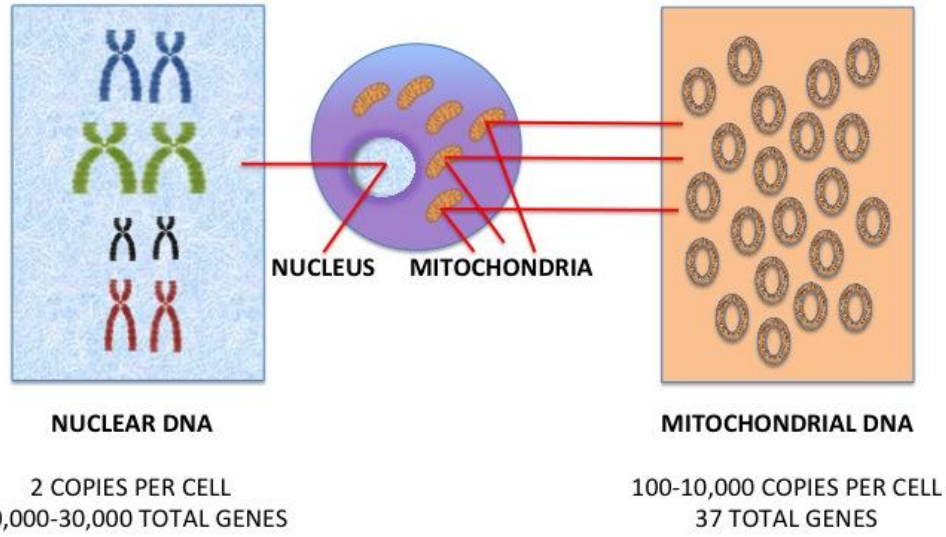

Figure 2. Nuclear DNA and Mitochondrial DNA. *Nuclear DNA exists in the form of 23 pairs of chromosomes localized in a (singular) nucleus within a cell. One cell may contain hundreds to thousands of mitochondria, each of which may contain several copies of mitochondrial DNA, present in the form of a closed, circular loop. Together, nuclear DNA and mitochondrial DNA comprise the human genome (see text for details).*

Another major difference that exists between nuclear DNA and mitochondrial DNA is their mode of inheritance. Unlike nuclear DNA, which is received from both the mother and the father, mitochondrial DNA is inherited strictly from the mother. During fertilization, only nuclear DNA from the sperm is effectively transferred to the egg cell. The nucleus in a sperm cell is located in the head region, which binds and penetrates the egg cell surface; mitochondria are located further down in the midpiece section, which transfer into the egg cell, but are subsequently destroyed (13,14). Thus, all surviving mitochondrial DNA in the fertilized egg cell and developing embryo are of maternal origin. Both male and female offspring born of the same mother carry identical mitochondrial DNA, which is the same as that of the mother, and the mother's maternal relatives (1,10,11). Males carry such mitochondrial DNA sequences but do not transmit them to subsequent offspring (Figure 3). Because mitochondrial DNA has a relatively high mutation rate, mitochondrial sequences typically differ from one unrelated person to another. Thus, it can be said with a relatively high confidence that no two individuals will express identical mitochondrial DNA sequences unless they share the same mother,

grandmother, or great grandmother, and so forth. Depending on the extent of similarity present within specific mitochondrial DNA segments,

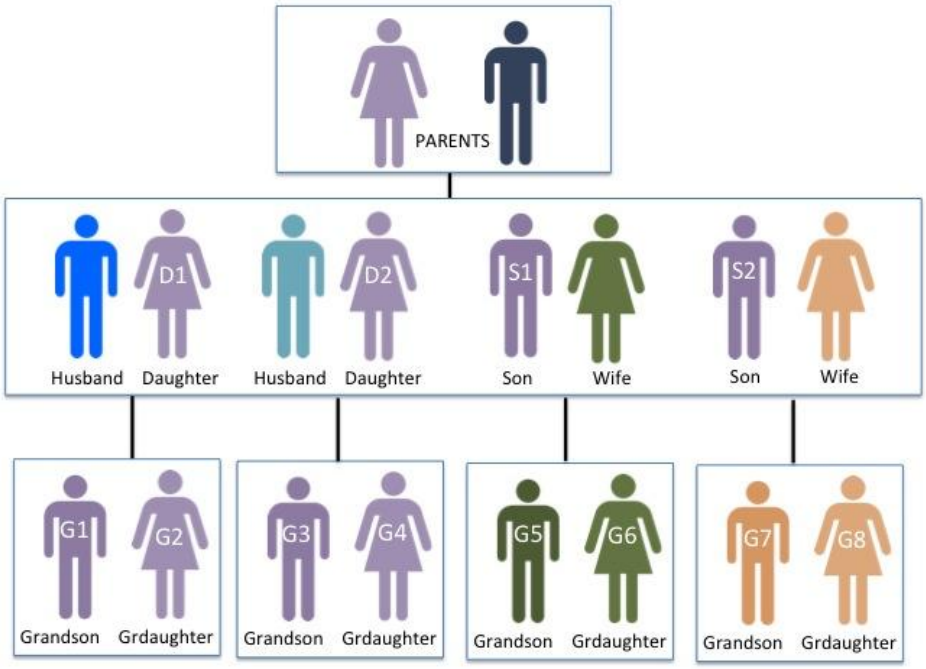

Figure 3. Maternal inheritance of mitochondrial DNA. *In this example, the parents give birth to two daughters (D1, D2) and two sons (S1,S2), all of which inherit the mother's mitochondrial DNA, as indicated by the purple shading. When the sons and daughters marry, only the daughters (D1, D2) will pass their mother's mitochondrial DNA on to their offspring (G1-G4). The offspring of the two sons, (G5-G8), will receive mitochondrial DNA from their respective mothers, i.e. the two women that each son married (see text for details).*

ancestry may be traced back to previous generations that existed hundreds of years ago or before (1, 10-12). Likewise, it can also be determined with a relatively high certainty that if two samples do not contain identical mitochondrial DNA sequences, they are not maternally related. Therefore, it is extremely unlikely that two unrelated people will share the same mitochondrial DNA. This property is especially useful when determining the relatedness of different DNA samples to one another.

Similar to nuclear DNA analysis, the discriminatory power of mitochondrial DNA analysis is proportional to the number of specific regions that are examined. Unlike nuclear DNA analysis, however, mitochondrial DNA testing cannot distinguish between a brother and a sister or even certain cousins that share the same maternal lineage (see Figure 3).

Analysis of mitochondrial DNA is based on a similar principle as nuclear DNA profiling, involving the examination of certain regions within the mitochondrial genome. Specifically, sequences located within so-called hypervariable regions,

termed HV1 and HV2, are the most studied. Additional regions within the mitochondrial genome may also be included for more definitive analysis (1,10-12). As mentioned above, due to the relative high copy number of mitochondrial DNA within the cell, such analysis is particularly useful when limited amounts of DNA may be present (10-12).

Previous DNA studies on the Shroud of Turin

DNA analysis of the Shroud has been reported; the most recent of these being over 15 years ago, when the sensitivity and capability of molecular biology techniques was far less than that which currently exists. In the late 1990s, Garza-Valdes reported in the book "The DNA of God" the sequencing of portions of three genes from threads taken from Shroud bloodstains: the beta-globin gene (a subunit of hemoglobin), and the amelogenin X and amelogenin Y genes, present on X and Y chromosomes, respectively. The threads examined were from the left-hand area and the occipital region (the back of the head). Garza-Valdes concluded "all three segments of human genes tested were positive, indicating the blood of the Man on the Shroud came from a human male" (15).

In 1995, Canale and coworkers performed DNA analysis of samples taken from both the Turin Shroud (from the soles of the left and right feet) and the Sudarium of Oviedo. Several sequences were examined, including the amelogenin X and Y genes, THO1 (tyrosine hydroxylase), vWA (von Willebrand factor), FES/FPS (tyrosine kinase), and F13A1 (coagulation factor XIII), (16). Two of these regions, THO1 and vWA, are part of the CODIS standard 13 STRs (6,7). The authors reported that contamination between male and female DNA, and all of the other sequences examined was evident (16, 17). It was stated that the presence of male DNA on the Shroud was more noticeable that female DNA, and contamination by persons who touched the cloths was considered as a contributing factor (16,17). Consequently, such results were considered essentially null and void. Contamination among DNA samples becomes evident when the presence of additional, distinct forms of a gene (termed alleles) is detected (1,2,6,8). While numerous forms of gene may exist within population (from several to hundreds, to even thousands depending on the particular gene), a single individual can express at most two forms: one inherited from the mother and one from the father (see Figure 4). Thus, when greater than two forms of a gene (alleles) are present, it can be concluded that the sample is composed of DNA from more than one individual (Figure 4). Amplification of mitochondrial DNA, but not nuclear DNA, has been reported from the Sudarium (18), although no sequence information was made available. As far as the author is aware, evaluation of mitochondrial DNA from the Shroud has not been reported.

In hindsight, there are two main concerns associated with the previous DNA studies on the Shroud: first, no evidence exists to show that any of the reported DNA sequences actually came from blood cells. All DNA sequences that have been examined to date are shared between blood cells and other cell types in the body, including skin cells (see below). Second, as previously suggested (16), any confirmatory DNA analysis of the Shroud should include evaluation of multiple sequences from numerous bloodstains to provide an accurate comparison of the

reproducibility of these findings (see below). Such studies are important in establishing that any DNA present originates from a single source. In previous

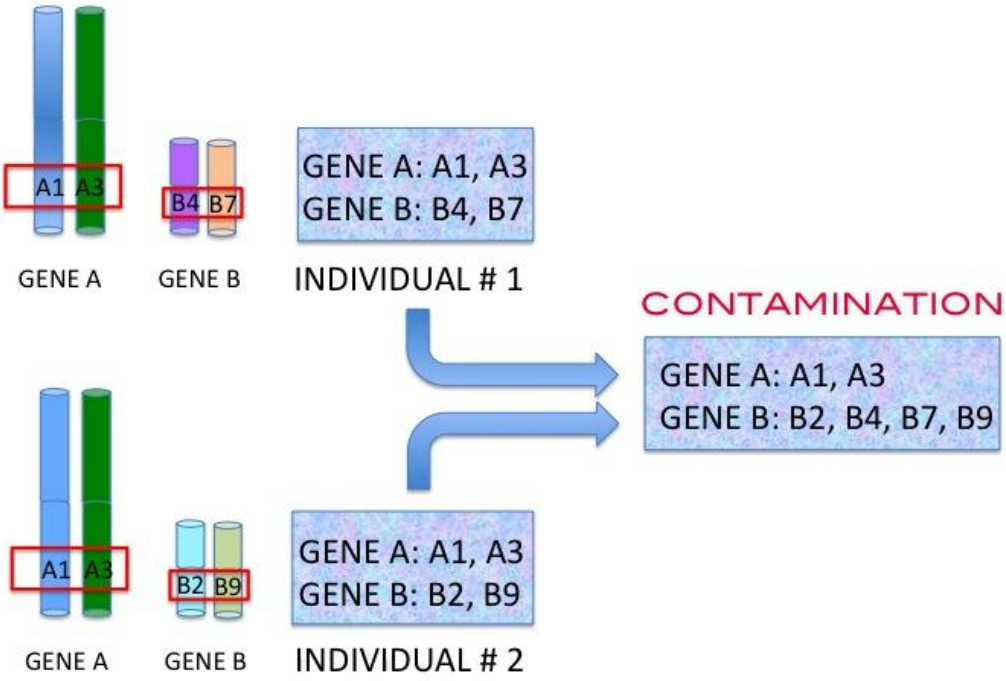

Figure 4. A generalized analysis of DNA samples. *Here, the analysis of two hypothetical genes (GENE A and GENE B) in two individuals is shown. A person may express up to two forms (maximum) of a gene, one that is inherited from the mother and one from the father. In this example, analysis of GENE A would not distinguish between Individual #1 and Individual #2 as both express identical forms (A1 and A3). In contrast, analysis of GENE B would distinguish between these two individuals as Individual #1 and Individual #2 express multiple distinct forms (B4, B7 and B2, B9, respectively). If a sample contains >2 forms of any gene (GENE B, in this example), this is an indication that contamination has occurred, i.e. that DNA from more than one individual is present (see text for details).*

studies (16), it was suggested that further investigation carried out on a larger number of samples taken from protected areas may make it possible to ascertain the original traces of DNA present on the Shroud (and Sudarium). This is especially relevant given the technological advances in DNA isolation and sequencing that have occurred within the past two decades. Such rapid progress has resulted from a combination of: (i) being able to amplify even minute quantities of DNA, over a billion copies of a single DNA sequence can be generated in four hours time using the polymerase chain reaction (PCR) method (19,20); and (ii) the development of computers with the capacity to direct many automated functions in the sequencing steps, and the ability to store and analyze large amounts of data (21). So-called next

generation DNA sequencing methods have become so sensitive within the past few years that well volumes in the zeptoliter (10^{-21} L) range, containing a single DNA replication enzyme, may now be evaluated (21). To put this into perspective, such analysis is able to effectively examine a volume as small as one millionth of one billionth of a single drop. It is now feasible to obtain sequence data from a single DNA molecule (even without amplification), something that was unprecedented as recently as 10-15 years ago.

Finally, it is noteworthy to mention that all previous Shroud DNA studies were performed prior to the development of more sensitive sequencing techniques, relying primarily on PCR amplification for DNA detection. PCR is most efficient at copying DNA fragments that are at least 80 bases in length, which could be problematic for certain aged DNA that might exist in smaller pieces. Next generation sequencers are able to read each base separately, which has made analysis of many ancient DNA genomes (often isolated in small fragments), now possible (21-28). In addition, modified DNA sequencing methods are currently available that aid in study of highly fragmented and degraded DNA, based on analysis of mini-STRs combined with single nucleotide polymorphisms (SNPs). Modern multiplex analysis allows the examination of multiple sequences within a single sample, minimizing the amount of material that is required for study (21-28).

Development of a Shroud CODIS: Potential and limitations of DNA studies

Development of a Shroud CODIS (combined DNA index system) would involve evaluating the presence or absence of a defined set of DNA sequences isolated from various regions of the Shroud, to establish a collective profile of the DNA that exists on the cloth. First and foremost, it would be important to determine the extent of DNA heterogeneity that may be present by evaluating samples taken from multiple bloodstains, including those on the reverse side. Conceivably, such distally located threads may be relatively more protected from contamination as compared to those on the regularly exposed side. Analysis of adjacent non-bloodstained fibers could also prove useful in evaluating the extent of DNA contamination, in general, that may be present within the region. For serious study, it would be worth considering that a DNA sample (simple cheek swab) be submitted by all available and willing persons that come in contact with the Shroud. Such samples could prove extremely useful in helping to index any background DNA sequences that may exist. Similar to the approach used in Human Genome Project, confidentiality could be maintained to ensure anonymity of results (3,4). Essentially, there are three possibilities regarding the origin(s) of any DNA present in Shroud bloodstains: i) The DNA is exclusively representative of endogenous DNA from white blood cells; ii) The DNA is exclusively representative of contaminating exogenous DNA (skin, sweat, saliva, tears); or iii) The DNA is a mixture of both endogenous and exogenous sequences. Development of a Shroud CODIS would help to discriminate between these various possibilities and place in context any previous sequence data that has been obtained.

One of the most fundamental issues that remains to be determined regarding DNA and the Shroud is if there is any DNA present that can be connected exclusively to blood cells (20). To date, none of the reported DNA sequence analysis addresses this

primary question. These experiments would require evaluation of immune receptor gene rearrangement, which is unique to white blood cells (5,20). Contaminating DNA from skin cells or any other source than blood cells would test negative in such a system. Another fundamental issue that has yet to be firmly resolved is the sex typing of the bloodstains. As previously mentioned, testing of the amelogenin gene is a standard addition in the 13 STR multiplex evaluation procedure. Sampling from several bloodstains, together with specific evaluation to the degree of heterogeneity that may be present (including additional STRs located on X, Y chromosomes) could decisively establish the male or female genotype of the bloodstains. Another area of interest that DNA analysis could potentially advance is to determine if the Shroud bloodstains originate from a single individual. These data might also be useful in addressing the suggestion that certain bloodstains have been "touched up" with the successive addition of blood throughout the Shroud's history. Finally, through probing of specific target sequences, the ABO and Rh factor genes, it might be possible to confirm the shroud blood type genetically, which would extend previous studies using serological techniques (29,30).

Taken together, such studies could establish the collective profile of the DNA sequences that are present on the Shroud. By extension, this Shroud CODIS could potentially be used to determine if a similar (identical) DNA profile was present on related artifacts, for example the Sudarium of Oviedo or the Tunic of Argentuli.

The two main limitations that most likely exist in the DNA analysis of the Shroud are: (i) the limited amounts of endogenous DNA that may be present and (ii) the amount of heterogeneity (contamination) that may exist. The discriminatory power of STR (and mitochondrial DNA) analysis is a function of the number of sequences for which data is collected. Definitive results from all 13 STR loci, which may not be achievable, would provide the highest discriminatory power to define the DNA profile representative of a particular Shroud bloodstain. Comparison of multiple STRs (and/or mitochondrial DNA) among bloodstains would provide an indication if DNA is present from a single or multiple source(s). How many different types of sequences might be required for a confident identification of the Shroud DNA profile is difficult to estimate: exclusion based on disparity relies on much simpler statistical probabilities than inclusion due to multiple matching (2,6,8).

It is unknown if degradation issues would preclude data collection and analysis even with more modern sequencing methods that are now available. Previous reports have indicated that DNA is, in fact, present on the Shroud, albeit in the pre-restoration time period (15,16). The restoration procedures may have introduced additional, exogenous DNA samples on the cloth, or perhaps even removed a degree of superficial, contaminating DNA from certain sites. Contamination could be a major issue; comparative evaluation of threads taken from multiple bloodstains should provide some insight as to the degree of heterogeneity that may exist. Parallel examination of sequencing data from a singular bloodstain sampled from both the front and reverse sides of the cloth might prove particularly informative.

Finally, DNA studies are by nature comparative. However, it should not be misunderstood that evaluation of DNA sequencing data would allow scientists to determine the identity of the man on the Shroud. No historic DNA profile of specific individuals, i.e. Jesus, exists for comparison. Provided sufficient DNA was intact, it

is conceivable that a more detailed evaluation of sequences (outside of standard CODIS STRs) could potentially provide certain information related to ancestry. It should be emphasized, however, that inclusion or exclusion to a particular heritage or region is not based on the evaluation of single, or even several, gene(s), but requires the analysis and consideration of multiple factors. At this stage, any suggestion of putative expected results would be unjustified speculation. Indeed, it remains to be determined if any endogenous DNA is even present and if standard sequence analyses of multiple regions are tenable. The generation of a Shroud CODIS would be most useful for internal comparison, among bloodstains, to conclusively establish what the DNA profile of the bloodstains consists of. DNA sequence data could provide insight into many important fundamental questions about the nature of the bloodstains, which at present remain inconclusive. Depending on the fidelity of such results, such data could possibly be extended to other, related artifacts to support or exclude their potential relationship with each other.

Acknowledgements

Thank you to John LaForest and Barrie Schwortz for critical reading of the manuscript. As always, thank you to my wife, Kathy, for everything.

References

1. Watson, J.D., et al., *Molecular Biology of the Gene*, The Benjamin/Cummings Publishing Company, Inc., Menlo Park, CA (2013).

2. Micklos, D.A., et al., *DNA Science: A first course*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (2012).

3. "All about the Human Genome Project", http://www.genome.gov/10001772, (2013).

4. Venter, J.C., et al., "The Sequence of the Human Genome", *Science* 291: 1304 (2001.)

5. Janeway, C., et al. *Immunobiology*, Garland Science Publishing, New York, NY (2005).

6. "Combined DNA Index System (CODIS)", http://www.fbi.gov/about-us/lab/biometric-analysis/codis, (2013).

7. "CODIS core loci", http://www.nfstc.org/pdi/Subject04/pdi_s04_m02_02.htm, (2013).

8. Norrgard, K., "Forensics, DNA Fingerprinting, and CODIS", *Nature Education* 1:1 (2008).

9. Budowle, B., et al., "Forensics and Mitochondrial DNA: Applications, Debates, and Foundations", *Annu. Rev. Hum. Genet.* 4: 199 (2003).

10. Isenberg, A.E. and Moore, J.M., "Mitochondrial DNA analysis at the FBI laboratory", Forensic Science Communications, 1:1 (1999).

11. Melton, T., et al., "Forensic mitochondrial DNA analysis: Current practice and future potential", Forensic Science Review, 24: 102 (2012).

12. Coble, et al., "Effective strategies for forensic analysis in the mitochondrial DNA coding region", Int. J. Leg. Med., 120:27 (2006).

13. Sutovsky, P., et al., “Ubiquitinated sperm mitochondria, selective proteolysis, and the regulation of mitochondrial inheritance in mammalian embryos”, Biol. Reprod. 63: 582 (2000).

14. Sutovsky, P., et al., “Degradation of paternal mitochondria after fertilization: implications for heteroplasmy, assisted reproductive technologies and mtDNA inheritance”, Reprod. Biomed. Online 8: 24 (2004).

15. Garza-Valdes, L., *The DNA of God?,* Doubleday, New York, USA (1999).

16. Casarino, et al., “Ricerca dei polimorfismi del DNA sulla sindone e sul Sudario di Oviedo”, Sindon N.S. Quad. 8:39 (1995).

17. Petrosillo, O. and Marinelli, E., “The Enigma of the Shroud”, Publishers Enterprises Group, San Gwann, Malta (1996).

18. “The Second International Conference on the Sudarium of Oviedo”, http://www.shroud.com/pdfs/n65part6.pdf, (2007).

19. Mullis, K.B., *The Polymerase Chain Reaction*, Birkhauser, Boston, MA (1994).

20. Kearse, K.P., “DNA on the Shroud of Turin: Distinguishing endogenous from exogenous DNA. http://www.shroud.com/pdfs/kearse2.pdf, (2012).

21. Kircher, M., and Kelso, J. “High-throughput DNA sequencing-concepts and limitations”, Bioesssays 32: 524 (2010).

22. Green, R.E., “The Neandertal genome and ancient DNA authenticity”, EMBO J. 28: 2494 (2009).

23. Green, R.E., et al., “A draft sequence of the Neantertal genome”, Science 328: 710 (2010).

24. Fu, Q. et al., “DNA analysis of an early modern human from Tianyuan Cave, China”, PNAS 110: 1 (2012).

25. Meyer, M. “A high-coverage genome sequence from an archaic Denisovan individual” Science 338: 222 (2012).

26. Grigorenko, A.P., et al. “Achievements and peculiarities in studies of ancient DNA and DNA from complicated specimens”, Acta Naturae 1: 58 (2009).

27. Rizzi, E., et al., “Ancient DNA studies: new perspectives on old samples”, Genetics Selection Evolution 44: 21 (2012).

28. Brotherton, P., et al., “Preferential access to genetic information from endogenous hominin ancient DNA and accurate quantitative SNP-typing via SPEX” Nucleic Acids, Res. 38: 7 (2010).

29. Kearse, K.P., “Empirical evidence that the blood on the Shroud of Turin is of human origin: Is the current data sufficient?”, http://www.shroud.com/pdfs/kearse1.pdf, (2012).

30. Kearse, K.P. “Blood on the Shroud of Turin: an Immunogical review”, http://www.shroud.com/pdfs/kearse.pdf, (2012).