

## **Metrology Needs and NIST Resources for the Forensic Community**

Michael D. Coble, Margaret C. Kline, and John M. Butler

National Institute of Standards and Technology, Applied Genetics Group, 100 Bureau Drive MS  
8312, Gaithersburg, MD, 20899-8312

### **Abstract**

With the advent of Forensic DNA profiling in the mid-1980s, this technology has had a positive impact on the criminal justice system, helping to convict the guilty and exonerate the innocent. The field has evolved from focusing on multi-locus markers throughout the nuclear DNA genome to the use of autosomal Short Tandem Repeat (STR) markers. Other marker systems such as mitochondrial DNA and Y-chromosomal STR testing have also found an important niche for the identification of missing persons and historical investigations. Given the importance of forensic DNA testing, it is critical that laboratories include proper controls and validated procedures for making quality measurements. In the US, the National Institute of Standards and Technology (NIST) has developed several Standard Reference Materials (SRMs) to meet the needs of the forensic DNA community. Here we will discuss a brief history of forensic DNA testing and the development of NIST SRMs and educational resources for the field over the last twenty years.

**Keywords:** National Institute of Standards and Technology (USA), Standard Reference Materials, Short Tandem Repeat, DNA Quantification, STRBase, Forensic DNA testing

MD Coble · MC Kline · JM Butler (\*)

National Institute of Standards and Technology  
Applied Genetics Group  
100 Bureau Drive MS 8312  
Gaithersburg, MD 20899-8312

e-mail: [john.butler@nist.gov](mailto:john.butler@nist.gov)

## **Introduction**

With the discovery of the ABO blood groups in 1901, Dr. Karl Landsteiner essentially created the field of “Serological Identification.” For the next 80 years or so, scientific investigations that utilized biological evidence compared to a suspect were made using multiple serological protein systems to detect variations in isoenzyme systems such as PGM (Phosphoglucomutase) in blood and semen [1]. In the mid-1980s, the “father” of forensic DNA profiling, Dr. Alec Jeffreys, discovered the usefulness of genetic markers he characterized in genes from the muscle cells of grey seals for individualization of humans [2]. These markers, Variable Number Tandem Repeats (VNTRs) are characterized as having a long core of nucleotide bases (generally 15+ nucleotides) that are repeated in tandem. The particular VNTRs that Jeffreys discovered were multi-locus – found throughout the nuclear DNA genome – and effectively looked like a genetic “bar-code” when separated on a gel.

Using DNA from a technician in his lab (and her two parents) Jeffreys noted that all of the genetic bands from the technician could be explained by the banding pattern of her two parents. In fact, the first use of this new discovery was for an immigration case [3]. A woman returning from Ghana to the UK was accompanied by her son, who looked somewhat different from his passport photo. Immigration officials were unsure if the boy was in fact someone unrelated to the woman. Jeffreys was able to conclusively show that the boy was in fact the woman’s son.

The success of the first paternity test led to the first application of this technology to a criminal case. In 1983 and 1986 two young girls were found sexually assaulted and murdered in the

village of Narborough, Leicestershire. The *modus operandi* of the two cases was similar leading investigators to believe that the two murders were committed by the same individual. After the arrest of a suspect and his confession of the second murder, investigators approached Jeffreys to give them an unquestionable conviction. Jeffreys discovered that in fact the crime scene evidence matched both cases; however, the main suspect was not the person who committed the crime. After additional detective work, the true perpetrator, Colin Pitchfork, became the first man to be convicted with the help of DNA evidence.

Forensic labs eventually moved away from the multi-locus VNTR probes developed by Jeffreys to analyze single locus VNTR markers. However, this increased the overall time required to develop a DNA profile since it would take up to 1-2 weeks to develop each of the 6-8 single locus VNTR markers necessary to find a unique set of alleles to give a strong statistical association between a match of the perpetrator and the evidence. In addition to the increased time requirement of the assay, RFLP DNA profiling required both a large quantity (tens to hundreds of nanograms) and relatively high quality (i.e. non-degraded) DNA for testing.

### **Forensic DNA Profiling**

With the development of Polymerase Chain Reaction (PCR) and the need to test smaller quantities of evidentiary material (often degraded) forensic DNA testing moved away from the single locus VNTR markers toward the analysis of Short Tandem Repeat (STR) markers. STRs are similar to VNTRs, but have fewer repeating elements (from 2-7bp core repeating units in STRs rather than 15+bp repeating units with VNTRs). By decreasing the size of the core

repeating units, smaller fragments could be amplified (an advantage for degraded samples).

Because STR markers are typically less variable than VNTR markers, a larger number of STRs are required for finding a unique DNA profile – typically around 10-15 markers. Fortunately commercial companies produce STRs kits under stringent quality control protocols to prevent each laboratory from making their own STR multiplexes. The selection of a standard set of STR markers, rapidly amplified by PCR and separated by capillary electrophoresis, improved the speed and ease of developing a DNA profile from several weeks with single locus VNTRs to several hours with STRs.

In addition to the testing of autosomal STRs for DNA profiling of crime science evidence and human identification, testing of the haploid markers – the paternally inherited Y-chromosome and the maternally inherited mitochondrial DNA (mtDNA) genome – have provided the forensic scientist with additional tools for forensic investigations [4]. STRs on the Y-chromosome (Y-STRs) can be useful for analyzing evidence from sexual assaults where DNA from both the male perpetrator and female victim are mixed together. With autosomal STRs, the DNA profile from the evidence may be difficult to interpret – especially if the male component of the mixture is at a low level. Amplification of the Y-STRs examines only the male portion of the mixture and not the female component. In highly degraded DNA samples – where the total amount of nuclear DNA is at very low levels, testing of the mtDNA genome can be useful since the mtDNA content within a cell is typically found in hundreds to thousands of copies (whereas the nuclear genome is present in only two copies). For mtDNA testing, the sequence of the four nucleotide bases are determined and compared to a reference mtDNA sequence, the Cambridge Reference Sequence [5]. Although neither haploid marker system is unique – all males in a paternal lineage and

females in a maternal lineage will share the same type – sometimes having a limited result is better than no result at all.

### **Metrology Needs for Forensic DNA testing**

With the proven power and growth of DNA testing around the world, the need for a Standard Reference Material (SRM) became apparent as a way to assure consistency among labs and to measure the uncertainty of very large DNA fragments migrating through an agarose gel. In the US, the National Institute of Standards and Technology (NIST) developed the first forensic DNA SRM for *HaeIII* RFLP testing (SRM 2390; [6]). Two major technological advances during the 1980s and 1990s transformed the field of forensic DNA typing: PCR amplification and the development of fluorescence based genetic analyzers initially using denaturing polyacrylamide gels and later capillary electrophoresis instruments. PCR amplification allowed the analysis of lower quantities of often degraded DNA (ranging from ~0.5 to 2.0 nanograms of DNA) and the new instrumentation gave rise to automated detection without the need to pour gels with capillary electrophoresis. With these advances in technology, and the movement away from RFLP testing to STRs, SRM 2390 is now obsolete. However the new instrumentation and interlaboratory exercises conducted by NIST lead to the development of other SRMs to meet the needs of forensic testing.

Similar to RFLP testing, the genotyping of STR alleles is a qualitative rather than a quantitative measurement. For STR testing two measurements are used to genotype the alleles at each marker. First, an internal size standard is incorporated into each sample run to determine the size

of the fragments in base pairs (Figure 1, bottom). These size standard fragments contain a set of markers of known base pair length and are used to develop a standard curve to measure the size of each allele. Second, an allelic ladder is used to determine the number of nucleotide repeats (genotype) the alleles from each sample (Figure 1, top). The allelic ladder is constructed by the kit manufacturer to contain a set of common alleles in the population. Genotype designations are essentially size bins and when any measured allele falls within  $<0.5$  bp of an allele from the allelic ladder, the peak is designated to have the same number of core STR repeats as the ladder (Figure 1, middle). The use of allelic ladders or sequenced samples with a well defined number of repeat units was supported by a 1995 interlaboratory study conducted by NIST. The results of this study indicated that the exchange of data between laboratories was best accomplished using an allelic name rather than base pair size. [7]

The qualitative measure made in forensic DNA typing is the number of tandem repeating elements that vary among the chromosomes in a population. These repeats are discrete units that differ in a quantum, step-wise fashion. For example, in a tetranucleotide (4 bp core motif) STR marker system an allele with “13” repeating units differs from a “16” allele by twelve nucleotide bases (Figure 1). Therefore, the *count* of the repeat number (rather than the sizes of each PCR amplicon measured in base pairs) is the preferred measurement unit for STR typing. This assures that minor fluctuations in the electrophoretic mobility of an allele from one lab’s capillary electrophoresis instrument to another lab’s instrument can be accounted for. For example, if an allele from one sample runs at 150 bp in lab A, but in lab B, this sample runs at 150.8 bp – one could argue that these are two differing alleles that vary by  $\sim 1$  bp. By incorporating an internal size standard and using an allelic ladder, both of these alleles would be genotyped as having the

same repeat number. This is not to say that slight variations among alleles are not possible. Typically, electrophoretic sizing uncertainty is less than  $\pm 0.5$  bp (so that any single base variants can be distinguished), which on a tetranucleotide (4 bp) repeat would be equivalent to  $\pm 0.125$  repeats. If an allele doesn't match exactly to the allelic ladder (termed as an "off-ladder" allele) – replicate analysis of the sample can be performed to confirm the presence of these variants.

In 1994, the US Federal Bureau of Investigation's DNA Advisory Board (DAB) established Quality Assurance Standards (QAS) for laboratories conducting forensic DNA testing [8]. Under Standard 9 (Analytical Procedures) the QAS state (section 9.5): "The laboratory shall check its DNA procedures annually or whenever substantial changes are made to the protocol(s) against an appropriate and available NIST standard reference material or standard traceable to a NIST standard." A revision of the QAS standards effective July 1, 2009 continued this policy (now section 9.5.5) and only changed "protocol(s)" to "procedure" in the new standards (<http://www.cstl.nist.gov/strbase/QAS/Final-FBI-Director-Forensic-Standards.pdf>).

To ensure accurate and comparable measurements between laboratories, NIST developed SRM 2391 [9] to meet the QAS requirement for the forensic DNA community. Currently, SRM 2391b (the third generation of the SRM) contains 12 components of DNA for genotyping [10]. Over 40 different STR markers used by forensic and paternity testing laboratories have been certified by NIST for the number of repeats at each allele by Sanger sequencing. There are also a number of considerations for developing the SRM including (a) Homogeneity (each component is consistent to minimize container to container variability), (b) Purity (single source samples are



used unless otherwise stated), and (c) Stability (certification of the SRM is 5-6 years under appropriate storage conditions (refrigerated or frozen, out of sunlight)).

Additional SRMs developed by NIST for the forensic DNA community are outlined in Table 1. These include SRMs for the haploid Y-chromosomal STR loci [10] and mtDNA sequence confirmation [11, 12]. Given the strict requirements for the amount of target DNA to be amplified with commercial STR kits – insufficient quantities of DNA can suffer from stochastic effects during PCR where some alleles are not amplified (drop-out); excessive quantities of DNA can generate artifact peaks in multiple fluorescent dye channels, confounding interpretation – NIST produced a human DNA quantification standard: SRM 2372 [13].

Beyond the requirement in the US for a yearly verification of laboratory procedures using the NIST SRM, there are a number of aspects of measurement quality that can be realized with the forensic SRM tools. For example, the NIST SRM 2372 quantification standard can be used to calibrate an in-house “standard.” A dilution of the SRM standard can be used to develop a standard curve to then calibrate either an in-house or a manufacturer’s standard. This “NIST-traceable” lot can now be stored and used accordingly. SRMs can also be used in the validation of a new STR kit as part of the concordance between the previous STR kit and the newly adopted kit. The SRMs can also be used to train new employees or in external Quality Assurance (e.g. proficiency testing) purposes. Finally, SRMs can also be used as an internal Quality Control check for new laboratory instrumentation.

### **Standard Reference Database (SRD)**

STRBase (<http://www.cstl.nist.gov/biotech/strbase/>) is a freely available website created by Dr. John Butler in 1997 to, “benefit research and application of short tandem repeat DNA markers to human identity testing.” The website is broken into four major parts: General Information, Forensic STR Information, Other DNA Marker Information and Non-Human DNA Resources, and Lab Resources and Tools.

The General Information section contains an introduction to STRBase, including training materials for laboratories to use for teaching newly hired technicians and analysts. Publications and presentations from the Human Identity Project Team for the forensic community are available. In the Forensic STR Information section, one can find an extensive amount of information on STRs including fact sheets, STR sequence information, the commercially available multiplex STR kits, rare alleles observed in DNA testing, and a comprehensive STR reference list of over 3400 published articles on STR testing. In the “Other DNA Marker Information and Non-Human DNA Resources” section one can find information on forensic Single Nucleotide Polymorphism (SNP) testing, mtDNA testing, and non-human STR testing (e.g. canine and feline STR testing). Finally, the Lab Resources and Tools section contains additional information on population data with STR testing (including over 600 population samples tested on multiple genetic markers at NIST), NIST-developed software (freely available) for the forensic community, information on the SRMs, validation information, and inter-laboratory studies conducted at NIST.

### **Future Needs for Forensic DNA Metrology**

One of the core values of NIST is, “Customer focus: We anticipate the needs of our customers and are committed to meeting or exceeding their expectations” ([http://www.nist.gov/public\\_affairs/mission.cfm](http://www.nist.gov/public_affairs/mission.cfm)). It can be difficult to predict what new genetic marker system will be used in the future of forensic DNA testing. In the relatively brief history of forensic DNA testing, we have already observed one methodology (RFLP testing) become obsolete and for now, STRs are the preferred method of DNA profiling. Regardless of the markers used for forensic testing, one thing is certain: given the importance of forensic DNA to the criminal justice system, there will be a need for traceable materials to make quality measurements for the forensic DNA community. We also realize the need to provide training and education for the forensic community and having STRBase as a tool to provide relevant information to our customers is necessary in a growing field like DNA testing.

### **Acknowledgements**

The authors would like to thank David Duewer, Becky Hill, Jan Redman, Erica Butts, Kristen Lewis O’Connor, Peter Vallone, and all former members, collaborators, interns and associates of the NIST human identity efforts over the past two decades. The authors would also like to acknowledge the National Institute of Justice for their funding to our NIST Group through the NIST Office of Law Enforcement Standards interagency agreement 2008-DN-R-121. The opinions and assertions contained herein are solely those of the author and are not to be construed as official or as views of the U.S. Department of Commerce. Commercial equipment, instruments and materials are identified in order to specify experimental procedures as completely as possible. In no case does such identification imply a recommendation or endorsement by the U.S. Department of Commerce nor does it imply that any of the materials, instruments or equipment identified are necessarily the best available for the purpose.

## References

1. Rees B, Rothwell TJ, Bonnar J (1974) Correlation of phosphoglucomutase isoenzymes in blood and semen. *Lancet* 303:783
2. Jeffreys AJ, Wilson V, Thein SL (1985) Individual-specific 'fingerprints' of human DNA. *Nature* 316:76-79
3. Jeffreys AJ, Brookfield JF, Semeonoff R (1985) Positive identification of an immigration test-case using human DNA fingerprints. *Nature* 317:818-819
4. Butler, J. M. (2009). *Fundamentals of Forensic DNA Typing*. Burlington, MA: Elsevier Academic Press Inc., U.S.
5. Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, Young IG (1981) Sequence and organization of the human mitochondrial genome. *Nature* 290:457-465
6. Duewer DL, Richie KL, Reeder DJ (2000) RFLP band size standards: NIST standard reference material 2390. *J Forensic Sci* 45:1093-1105
7. Kline MC, Duewer DL, Newall P, Redman JW, Reeder DJ, Richard M (1997) Interlaboratory Evaluation of Short Tandem Repeat Triplex CTT. *J Forensic Sci* 42:897-906
8. DNA Advisory Board (2000) Quality Assurance Standards for Forensic DNA Testing Laboratories. *Forensic Science Communications*, Volume 2. Available online: <http://www2.fbi.gov/hq/lab/fsc/backissu/july2000/codis2a.htm>
9. Levin BC, Cheng H, Kline MC, Redman JW, Richie KL (2001) A review of the DNA standard reference materials developed by the National Institute of Standards and Technology. *Fresenius J Anal Chem* 370:213-219

10. Kline MC, Decker AE, Hill CR, Butler JM (2006) NIST SRM Updates: Value-added to the Current Materials in SRM 2391b and SRM 2395. Poster presentation at the International Symposium on Human Identification, Nashville, TN. Available online: [http://www.cstl.nist.gov/biotech/strbase/pub\\_pres/Promega2006\\_Kline.pdf](http://www.cstl.nist.gov/biotech/strbase/pub_pres/Promega2006_Kline.pdf)
11. Levin BC, Cheng H, Reeder DJ (1999) A human mitochondrial DNA standard reference material for quality control in forensic identification, medical diagnosis, and mutation detection. *Genomics* 55:135-146
12. Levin BC, Holland KA, Hancock DK, Coble M, Parsons TJ, Kienker LJ, Williams DW, Jones M, Richie KL (2003) Comparison of the complete mtDNA genome sequences of human cell lines--HL-60 and GM10742A--from individuals with pro-myelocytic leukemia and leber hereditary optic neuropathy, respectively, and the inclusion of HL-60 in the NIST human mitochondrial DNA standard reference material--SRM 2392-I. *Mitochondrion* 2:387-400
13. Kline MC, Duewer DL, Travis JC, Smith MV, Redman JW, Vallone PM, Decker AE, Butler JM (2009) Production and certification of NIST Standard Reference Material 2372 Human DNA Quantitation Standard. *Anal Bioanal Chem* 394:1183-1192

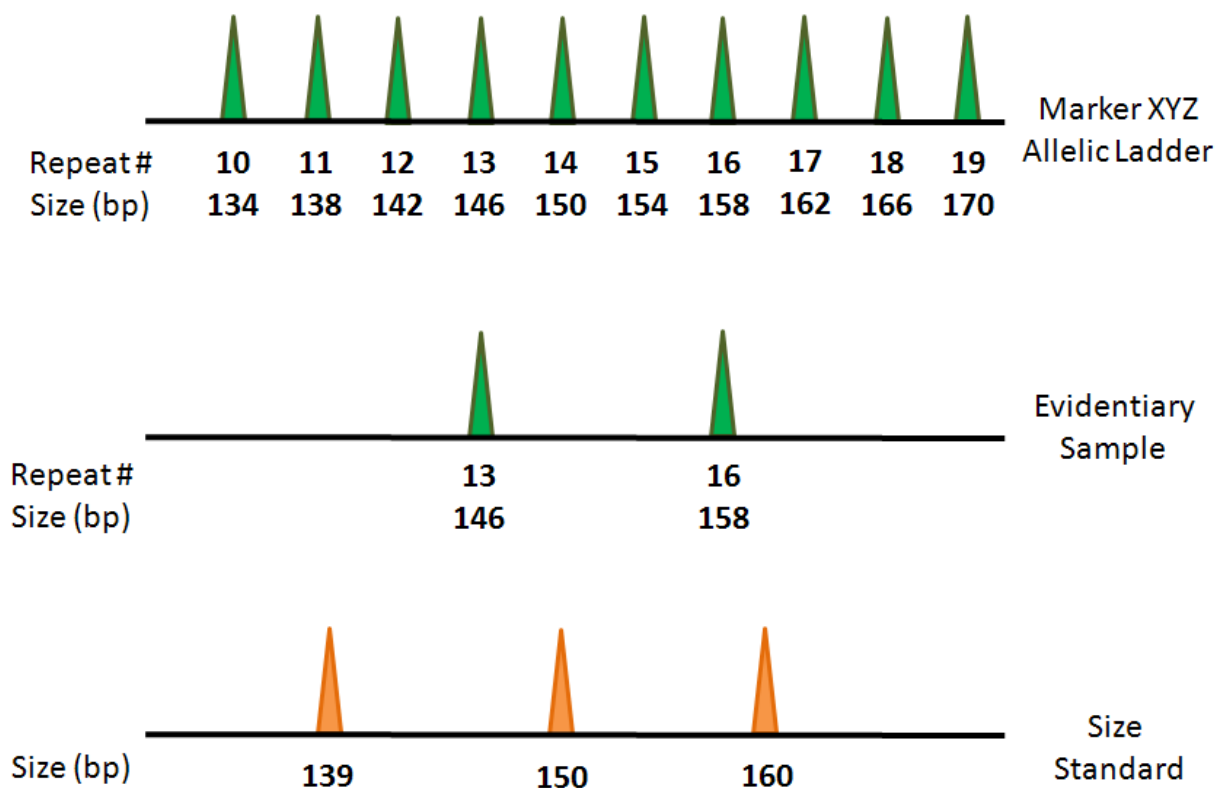


Figure 1. How Size Standards and Allelic Ladders are used to Genotype Forensic DNA Samples. A size standard (bottom) contains DNA fragments of known lengths (in bp) and is used to develop a standard curve for sizing the peaks in the evidentiary sample (middle) and the allelic ladder of the hypothetical marker XYZ (top). The number of repeating units in the sample is determined by an allelic ladder (top) which is provided in the commercial kit and covers the most commonly occurring alleles in a population (e.g. in this example, over 99% of the population would have alleles from 10 to 19 repeats). Based upon the size of the alleles in the evidentiary sample and the alignment of the allelic ladder, the genotype of the evidentiary sample for hypothetical marker (XYZ) is determined to be “13, 16.”

Table 1. NIST SRMs Developed for the Forensic DNA Community

SRM	Name	Year Released (r = year revised)	Comments	Reference
2390	RFLP DNA Profiling	1992, r2001	Now obsolete	[6]
2391	STR PCR DNA Profiling	1995-1998	Certified for the marker D1S80	[8, 9]
2391a	STR PCR DNA Profiling	1998-2002	Initially certified for only 4 STR markers	[8, 9]
2391b	STR PCR DNA Profiling	2003, r2008	Supply will be exhausted in 2011	[8, 9]
2391c	STR PCR DNA Profiling	2011	Is in preparation stage	--
2392, 2392-I	Mitochondrial DNA Sequencing	1999, 2003, r2009	2392-I contains the HL-60 cell line	[10, 11]
2395	Human Y-Chromosome DNA Profiling	2003, r2008	Certified for over 40 Y-Chr markers	[9]
2372	Human DNA Quantitation	2007	For platform testing	[12]

