

Email: [becky.hill@nist.gov](mailto:becky.hill@nist.gov)  
Phone: 301-975-4275

Carolyn R. Hill, Michael D. Coble, and John M. Butler

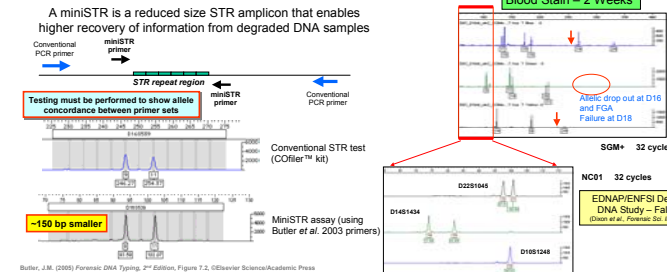
National Institute of Standards and Technology, Biochemical Science Division, 100 Bureau Drive MS 8311, Gaithersburg, MD 20899-8311

The development of new mini Short Tandem Repeat (miniSTR) loci for their use in forensic DNA typing is valuable in forensic casework involving DNA samples that are highly degraded due to exposure to environmental elements or contaminants (e.g. mass disasters such as the September 11, 2001 terrorist attack, the Indian Ocean Tsunami of December 2004, and Hurricane Katrina in August 2005). A number of studies have demonstrated that successful analysis of degraded DNA samples for human identification improve with smaller sized polymerase chain reaction (PCR) products (1). In addition, by moving PCR primers closer to the STR region, we have established that it is possible to decrease the incidence of allele or locus-droptout that may occur in degraded DNA samples.

The 27 new miniSTR loci were chosen based on their size and location on certain chromosomes. The candidate loci are all either located on chromosomes that differ from the 13 CODIS (COmbined DNA Index System) core loci or are at least ~50 Mb apart from an existing CODIS locus on the same chromosome, and therefore unlinked from that particular marker. New autosomal STR loci are being examined because many of the CODIS core loci have large allele ranges (e.g., D21S11 and FGA) that make it impossible to create small PCR products. The first 6 loci (grouped into 2 multiplexes of 3 loci, NC01 and NC02) were previously characterized (2), and the process described for their development and characterization form the basis of the current work involving the remaining 21 new miniSTR loci. Each locus was evaluated across more than 600 samples representing the three major populations in the U.S.: Caucasian, African American, and Hispanic. All of the data was genotyped and allele frequencies and population statistics were determined.

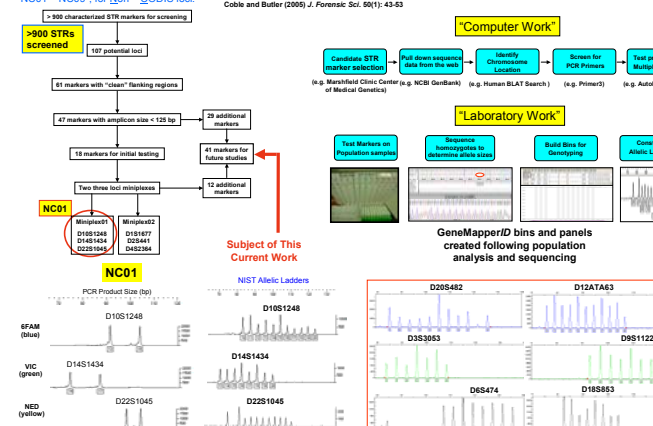
## Use of miniSTRs to Improve Analysis of Samples Containing Degraded DNA

In forensic casework, it is possible and quite likely that the DNA samples will be highly degraded due to exposure to environmental elements or natural contaminants that can result in loss of information at higher molecular weight STR loci. This loss of signal may be the result of the presence of PCR inhibitors or a DNA template that has been fragmented into smaller sizes, causing a partial DNA profile with allele or even complete locus dropout (1). Many of the commercial multiplex STR kits used in forensic DNA typing generate profiles up to 450 bp and can lead to loss of information at the higher molecular weight STR loci. In addition, many of the CODIS 13 core loci have large allele ranges that make it impossible to create small PCR products (1). Thus, by examining new STR loci that allow for the reduction of the size of the PCR products by moving the primers as close as possible to the STR repeat region, we have been able to observe a higher recovery of information from degraded DNA samples (2).



## Selection and Characterization of New miniSTR Loci

The 27 new miniSTR loci were chosen based on their size and location on certain chromosomes. The candidate loci are all either located on chromosomes that differ from the 13 CODIS core loci or are at least ~50 Mb apart from an existing CODIS locus on the same chromosome, and therefore unlinked from that particular marker. The 27 new loci are separated into 9 groups of 3 loci, and are named "NC01 - NC09" for Non-CODIS loci.



## Miniplex Sets Used in This Study

(Chosen to avoid linkage with CODIS 13 STRs to enable use of the product rule)

- NC01: D10S1248, D14S1434, D2S11045
- NC02: D4S2364, D2S441, D1S1677
- NC03: D3S3053, D6S474, D20S482
- NC04: D1GATA13, D2S1776, D4S2408
- NC05: D1S1627, D8S2500, D8S1115
- NC06: D3S4529, D6S1017, D9S2147
- NC07: D9S1122, D10S1436, D12AT6A3
- NC08: D17S1301, D18S853, D20S1082
- NC09: D6S1027, D17S974, D11S4443

**Rejected Loci and Why**  
D10S1430: complex repeats  
D10S2327: tri-, quad-allelic profiles  
D15S817: poor heterozygosity  
D15S818: tri-, quad-allelic profiles

\*32 loci x 663 samples = 21,216 total data points in this study

For more information, see STRBase:  
<http://www.cstl.nist.gov/biotech/strbase/nist.htm>  
<http://www.cstl.nist.gov/biotech/strbase/newSTRs.htm>  
<http://www.cstl.nist.gov/biotech/strbase/miniSTR/timeline.htm>

## Population Data

A set of ~650 U.S. population samples (<http://www.cstl.nist.gov/biotech/strbase/NISTpop.htm>) were used to assess variation in African Americans, Caucasians, and Hispanics. The 3 loci boxed in blue are the new core European loci.

| Marker      | Allele   | B-FAM: BLUE |    |    | VIC: GREEN |    |    | NED: YELLOW |    |    |
|-------------|----------|-------------|----|----|------------|----|----|-------------|----|----|
|             |          | AF          | AA | H  | AF         | AA | H  | AF          | AA | H  |
| <b>NC01</b> | D10S1248 | 11          | 11 | 11 | 11         | 11 | 11 | 11          | 11 | 11 |
| <b>NC02</b> | D4S2364  | 11          | 11 | 11 | 11         | 11 | 11 | 11          | 11 | 11 |
| <b>NC03</b> | D2S11045 | 11          | 11 | 11 | 11         | 11 | 11 | 11          | 11 | 11 |
| <b>NC04</b> | D17S1301 | 11          | 11 | 11 | 11         | 11 | 11 | 11          | 11 | 11 |
| <b>NC05</b> | D1S1627  | 11          | 11 | 11 | 11         | 11 | 11 | 11          | 11 | 11 |
| <b>NC06</b> | D3S4529  | 11          | 11 | 11 | 11         | 11 | 11 | 11          | 11 | 11 |
| <b>NC07</b> | D9S1122  | 11          | 11 | 11 | 11         | 11 | 11 | 11          | 11 | 11 |
| <b>NC08</b> | D17S1301 | 11          | 11 | 11 | 11         | 11 | 11 | 11          | 11 | 11 |
| <b>NC09</b> | D6S1027  | 11          | 11 | 11 | 11         | 11 | 11 | 11          | 11 | 11 |

## Population Data for NC01 and NC02 Loci is Being Collected by Other Labs

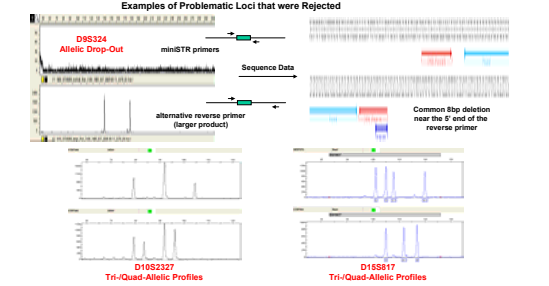


| Source                               | Number of Samples | Population        | Country  | Lab  | Year |
|--------------------------------------|-------------------|-------------------|----------|------|------|
| Coble and Butler (2005)              | 154               | African-Americans | USA      | NIST | 2005 |
| Coble and Butler (2005)              | 278               | U.S. Caucasians   | USA      | NIST | 2005 |
| Coble and Butler (2005)              | 140               | U.S. Hispanics    | USA      | NIST | 2005 |
| Asamura et al. (2005)                | 142               | Japanese          | Japan    | NIST | 2005 |
| Yong et al. (2005) in press          | 135               | Chinese           | China    | NIST | 2005 |
| Yong et al. (2005) in press          | 132               | Malaysian         | Malaysia | NIST | 2005 |
| Yong et al. (2005) in press          | 178               | East Indian       | India    | NIST | 2005 |
| Prof. Ramon Domenech (2005) in press | 100               | Italian           | Italy    | NIST | 2005 |
| GenetiX (2005) in preparation        | 200               | Romanian          | Romania  | NIST | 2005 |
| GenetiX (2005) in preparation        | 200               | Romanian          | Romania  | NIST | 2005 |
| Simon Lim, in preparation            | 226               | Chinese           | China    | NIST | 2005 |
| Simon Lim, in preparation            | 238               | Malaysian         | Malaysia | NIST | 2005 |
| Simon Lim, in preparation            | 238               | Malaysian         | Malaysia | NIST | 2005 |
| Simon Lim, in preparation            | 238               | Malaysian         | Malaysia | NIST | 2005 |
| <b>Total</b>                         | <b>2363</b>       |                   |          |      |      |

## Materials and Methods

PCR primer sequences and full assay conditions will be included in a forthcoming publication

General protocols are available on the NIST STRBase website:  
[http://www.cstl.nist.gov/biotech/strbase/miniSTR/Updated\\_NC01\\_protocol.pdf](http://www.cstl.nist.gov/biotech/strbase/miniSTR/Updated_NC01_protocol.pdf)  
[http://www.cstl.nist.gov/biotech/strbase/miniSTR/Protocol\\_AllelicLadderReamp.pdf](http://www.cstl.nist.gov/biotech/strbase/miniSTR/Protocol_AllelicLadderReamp.pdf)



## Characteristics of New miniSTR Loci

Comparison of heterozygosity values for 27 non-CODIS miniSTRs (black) and the 15 markers from the ABI Identifier kit (blue) across the U.S. samples examined in this study. The 3 loci boxed in blue are the new core European loci. The loci are arranged according to their Overall Rank, which is determined by the Heterozygosity Rank added to the Size Range Rank.

| Locus           | Repeat Type | Repeat Position | Chromosome        | N (Overall) | Heterozygosity Rank | African | Caucasian | Hispanic | Observed Size Range | Size Range (bp) | Overall Rank |
|-----------------|-------------|-----------------|-------------------|-------------|---------------------|---------|-----------|----------|---------------------|-----------------|--------------|
| D2S1248         | Ta          | ATA             | Chr 2: 103,077 Mb | 651         | 0.844               | 4       | 0.854     | 0.840    | 0.779               | 71              | 101          |
| A146            | CTG         | TAA             | Chr 2: 36,78 Mb   | 663         | 0.784               | 11      | 0.817     | 0.785    | 0.717               | 16              | 119          |
| D2S441 (NC02)   | Ta          | TCTA            | Chr 2: 68,21 Mb   | 660         | 0.774               | 12      | 0.798     | 0.760    | 0.721               | 70              | 119          |
| D10S2324 (NC02) | Ta          | GGAA            | Chr 10: 100,27 Mb | 653         | 0.762               | 15      | 0.825     | 0.785    | 0.743               | 83              | 122          |
| D19S1435        | Ta          | GATA            | Chr 10: 22,33 Mb  | 653         | 0.766               | 15      | 0.798     | 0.770    | 0.720               | 82              | 123          |
| D19S1437        | Ta          | ATT             | Chr 11: 100,6 Mb  | 656         | 0.766               | 12      | 0.793     | 0.727    | 0.683               | 81              | 126          |
| D19S1438        | Ta          | AAGG            | Chr 11: 101,11 Mb | 659         | 0.803               | 9       | 0.875     | 0.794    | 0.744               | 104             | 126          |
| D15S817 (NC02)  | Ta          | GGAA            | Chr 11: 160,73 Mb | 655         | 0.746               | 23      | 0.743     | 0.740    | 0.743               | 81              | 117          |
| FGA             | Ta          | CTTT            | Chr 4: 179,48 Mb  | 664         | 0.740               | 26      | 0.807     | 0.806    | 0.803               | 81              | 121          |
| D6S1017         | Complex     | Ta              | Chr 2: 18,48 Mb   | 654         | 0.744               | 1       | 0.829     | 0.844    | 0.871               | 106             | 244          |
| D6S1018         | Complex     | Ta              | Chr 2: 19,83 Mb   | 659         | 0.826               | 7       | 0.802     | 0.836    | 0.800               | 109             | 242          |
| D6S200          | Complex     | Ta              | Chr 2: 58,73 Mb   | 654         | 0.747               | 20      | 0.757     | 0.747    | 0.729               | 85              | 125          |
| D14S1044 (NC01) | Complex     | Ta              | Chr 12: 50,23 Mb  | 653         | 0.686               | 26      | 0.688     | 0.671    | 0.679               | 79              | 108          |
| D3S1598         | Complex     | Ta              | Chr 4: 45,4 Mb    | 659         | 0.703               | 16      | 0.757     | 0.753    | 0.717               | 87              | 145          |
| D6S1111         | Complex     | Ta              | Chr 6: 42,63 Mb   | 664         | 0.683               | 30      | 0.620     | 0.683    | 0.729               | 83              | 106          |
| D19S1437        | Ta          | ATA             | Chr 11: 100,6 Mb  | 656         | 0.699               | 15      | 0.792     | 0.663    | 0.600               | 73              | 100          |
| D6S1179         | Complex     | Ta              | Chr 12: 50,23 Mb  | 659         | 0.774               | 13      | 0.763     | 0.779    | 0.786               | 123             | 211          |
| D19S1438        | Ta          | ATCA            | Chr 11: 100,73 Mb | 659         | 0.878               | 1       | 0.893     | 0.914    | 0.944               | 264             | 344          |
| D4S2364 (NC02)  | Complex     | Ta              | Chr 4: 59,58 Mb   | 650         | 0.511               | 41      | 0.385     | 0.501    | 0.684               | 67              | 83           |
| D3S1338         | Complex     | Ta              | Chr 2: 218,11 Mb  | 659         | 0.882               | 3       | 0.903     | 0.882    | 0.843               | 28              | 33           |
| D3S1339         | Ta          | GATA            | Chr 17: 17,23 Mb  | 648         | 0.729               | 27      | 0.712     | 0.724    | 0.814               | 84              | 168          |
| D3S4529         | Ta          | GATA            | Chr 6: 42,63 Mb   | 660         | 0.761               | 18      | 0.762     | 0.723    | 0.629               | 111             | 139          |
| D6S474          | Complex     | Ta              | Chr 1: 12,99 Mb   | 648         | 0.761               | 19      | 0.765     | 0.802    | 0.679               | 107             | 215          |
| D6S1017         | Ta          | GGAA            | Chr 11: 160,73 Mb | 655         | 0.808               | 8       | 0.793     | 0.817    | 0.814               | 293             | 293          |
| D6S1018         | Ta          | GATA            | Chr 2: 18,48 Mb   | 654         | 0.763               | 17      | 0.740     | 0.801    | 0.774               | 124             | 181          |
| D6S1019         | Ta          | TAA             | Chr 11: 100,73 Mb | 654         | 0.721               | 21      | 0.772     | 0.746    | 0.691               | 81              | 116          |
| D6S1020         | Ta          | AAT             | Chr 11: 100,73 Mb | 653         | 0.733               | 23      | 0.744     | 0.713    | 0.700               | 72              | 102          |
| D6S1021         | Ta          | GATA            | Chr 6: 42,63 Mb   | 654         | 0.722               | 31      | 0.726     | 0.746    | 0.681               | 81              | 116          |
| D6S1022         | Ta          | GATA            | Chr 11: 17,4 Mb   | 654         | 0.688               | 38      | 0.673     | 0.652    | 0.727               | 77              | 111          |
| GATA113 (D1)    | Ta          | GATA            | Chr 11: 17,4 Mb   | 654         | 0.722               | 31      | 0.726     | 0.746    | 0.681               | 81              | 116          |
| D6S1122         | Ta          | GATA            | Chr 9: 78,9 Mb    | 659         | 0.734               | 26      | 0.753     | 0.742    | 0.686               | 93              | 125          |
| D19S1438        | Ta          | GATA            | Chr 11: 100,73 Mb | 654         | 0.730               | 30      | 0.780     | 0.878    | 0.743               | 106             | 156          |
| D19S1439        | Ta          | GATA            | Chr 11: 100,73 Mb | 659         | 0.786               | 14      | 0.798     | 0.733    | 0.733               | 237             | 348          |
| D19S1440        | Ta          | GATA            | Chr 11: 100,73 Mb | 659         | 0.881               | 2       | 0.877     | 0.874    | 0.829               | 106             | 156          |
| D19S1441        | Ta          | TATC            | Chr 11: 100,73 Mb | 658         | 0.747               | 21      | 0.693     | 0.748    | 0.843               | 193             | 237          |
| D19S1442        | Ta          | GAT             | Chr 11: 100,73 Mb | 658         | 0.881               | 3       | 0.877     | 0.874    | 0.829               | 106             | 156          |
| T601            | Ta          | TGAT            | Chr 11: 21,24 Mb  | 659         | 0.745               | 25      | 0.759     | 0.721    | 0.744               | 100             | 204          |
| D6S1018         | Ta          | AGAT            | Chr 2: 123,19 Mb  | 659         | 0.721               | 32      | 0.726     | 0.720    | 0.744               | 110             | 211          |
| D6S1019         | Ta          | AGAT            | Chr 2: 123,19 Mb  | 659         | 0.749               | 24      | 0.739     | 0.733    | 0.693               | 109             | 202          |
| D17S1301        | Ta          | TCTA            | Chr 17: 70,18 Mb  | 664         | 0.649               | 40      | 0.626     | 0.671    | 0.644               | 114             | 218          |
| TPCA            | Ta          | GATA            | Chr 11: 14,4 Mb   | 659         | 0.707               | 34      | 0.703     | 0.688    | 0.679               | 139             | 241          |

## Summary

- New miniSTR markers are being characterized and information will be made available on STRBase (<http://www.cstl.nist.gov/biotech/strbase/newSTRs.htm>).
- Several of these miniSTR loci have been recommended for adoption by the European DNA community as new core loci. (Gill et al. 2006)
- In addition to increasing the successful typing of degraded materials, these loci can also provide additional discrimination in complex paternity cases or missing persons cases.

## Acknowledgments

Funding from the National Institute of Justice through interagency agreement 2003-JJ-R-029 to the NIST Office of Law Enforcement Standards. Points of view are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice. Commercial equipment, instruments and materials are identified in order to specify experimental procedures completely as possible. In no case does such identification imply a recommendation or endorsement by the National Institute of Standards and Technology nor does it imply that any of the materials, instruments, or equipment identified are necessarily the best available for the purpose.

## Disclaimer

This project was funded by the National Institute of Justice through interagency agreement 2003-JJ-R-029 to the NIST Office of Law Enforcement Standards. Points of view are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice. Commercial equipment, instruments and materials are identified in order to specify experimental procedures completely as possible. In no case does such identification imply a recommendation or endorsement by the National Institute of Standards and Technology nor does it imply that any of the materials, instruments, or equipment identified are necessarily the best available for the purpose.