Uses of the NIST 26plex STR Assay for Human Identity Testing

Peter M. Vallone[1], Carolyn R. Hill[1], Kristen E. Lewis[2], Toni M. Diegoli[3], Michael D. Coble[3], and John M. Butler[1]

Email: petev@nist.gov

[1] U.S. National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, MD 20899-8311
[2] Department of Genome Sciences, University of Washington, Seattle, WA 98195-5065
[3] U.S. Armed Forces DNA Identification Laboratory, 1413 Research Drive, Rockville, MD 20850

P-022

Ongoing work at the U.S. National Institute of Standards and Technology has focused on the characterization of 26 autosomal STR loci for human identity testing. These 26 loci are in addition to the existing 13 U.S. core loci and those found in PowerPlex16 and Identifiler commercial STR typing kits. The amplification of the 26 loci has been optimized for degraded extracts in unique 3plex panels and also for reference samples as a large single reaction 26plex assay[1,2]. A study has been performed comparing genotypes obtained with the 26plex primers to those with miniSTR 3plex panels for allele drop out and concordance[3]. The forensic utility of the 26plex assay was evaluated for situations where additional loci are beneficial. These assays were used to type samples in a complex kinship scenario and also to test the benefits of additional loci when simulating familial searching. The utility of this large multiplex was also tested on DNA extracted from degraded bone samples. The 26plex can serve as a low cost assay (compared to commercially available kits) useful for both sorting highly comingled remains and providing additional markers for increased statistical support for samples that use "non-trio" family references for human identification.

[1] Hill, C.R., Kline, M.C., Coble, M.D., Butler, J.M. (2008) Characterization of 26 miniSTR loci for improved analysis of degraded DNA samples. *J. Forensic Sci.* 53:73-80.
[2] Hill, C.R., Butler, J.M., and Vallone, P.M. (2009) A new 26plex assay for use in human identity testing. *J. Forensic Sci. 54:* 1008-1015
[3] Butler, J.M., Hill, C.R., Decker, A.E., Kline, M.C., Reid, T.M., Vallone, P.M. (2007) New autosomal and Y-chromosome STR loci: characterization and potential uses. *Proceedings of the Eighteenth International Symposium on Human Identification.*

Timeline

- 2004: EDNAP degraded DNA study shows value of NIST miniSTR assays — Dixon *et al. Forensic Sci. Int.* 164: 33-44
- 2004: Coble, M.D. and Butler, J.M. (2005) Characterization of new miniSTR loci to aid analysis of degraded DNA. *J. Forensic Sci.* 50: 43-53.
- 2005: Europe adopts new loci D10S1248, D14S1434 and D22S1045 — Gill *et al. Forensic Sci. Int.* 2006;156:242-244
- 2005: D2S441 replaces D14S1434 — Gill *et al. Forensic Sci. Int.* 2006;163:155-157
- 2005: Mutation rate studies & initial kinship data collected
- 2006: Design of a single tube reference multiplex (26 loci + Amelogenin) PCR begins
- 2007: Applied Biosystems MiniFiler released
- 2008: Completion of a 25 STR locus + Amelogenin multiplex = 26plex
- 2008: Hill, C.R., Kline, M.C., Coble, M.D., Butler, J.M. (2008) Characterization of 26 miniSTR loci for improved analysis of degraded DNA samples. *J. Forensic Sci.* 53(1):73-80.
- 2008: 26 loci are certified for the US NIST SRM 2391b (September, 2008)
- 2009: Hill, C.R., Butler, J.M., and Vallone, P.M. (2009) A new 26plex assay for use in human identity testing. *J. Forensic Sci. 54:* 1008-1015
- 2009: Promega releases commercial kit containing D10S1248, D22S1045, and D2S441

Assay Development

Additional loci were originally selected as candidates for miniSTR assays

- Certain CODIS and existing kit loci are not amenable to miniSTR assay design
- Large allele range (FGA)
- STR flanking region sequence that results in larger amplicons (D7S820 and D21S11)

In 2004 - 2005 **Dr. Michael D. Coble** performed a survey of STRs to find loci that would make good candidates

Loci Criteria:
- Heterozygosity > 0.7
- Moderate allele range (= low mutation rates)
- Tri & Tetra nucleotide repeat motifs
- **Not linked to CODIS/kit loci**

26 candidates were selected and termed 'NC' for non-CODIS/Core loci

26 New STR Loci for Human Identity Testing

Initial miniSTR work
- Low level multiplex assays developed (10 miniplexes)
- Intended for use on degraded samples
- Sensitivity down to 100 pg (with 28 cycles)

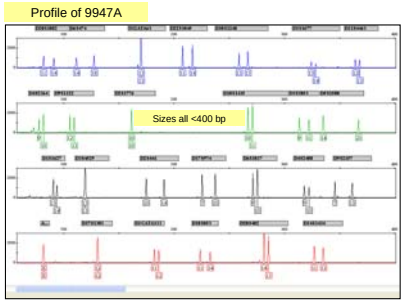
Utility of miniplexes
- Degraded DNA
- Low copy number analysis

Standard Reference Material
- The 26 loci are certified for NIST SRM 2391b

Reference multiplex
- Goal: to type all 26 STR loci in a single reaction
- 65 to 400 base pair amplicons
- Majority of PCR primers redesigned - **no longer miniSTRs**
- D8S1115 was omitted from the final reference multiplex
- **26plex = 25 STRs + Amelogenin**

Utility of a large multiplex
- Kinship testing
- Reference samples
- Population studies
- Mutation rate studies
- Genotype concordance with miniSTR primer sets

26plex Profile (25 STRs + Amelogenin)



The PCR primers and conditions for the 26plex are described in: Hill, C.R., Butler, J.M., and Vallone, P.M. (2009) A new 26plex assay for use in human identity testing. *J. Forensic Sci. 54:* 1008-1015

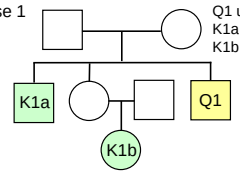
Target DNA template amount ~1 ng of genomic DNA (30 cycles)
Sensitivity as low as 100 pg (with 30 cycles)
GeneMapper IDv3.2 bins and panels developed for POP4 and POP6 (http://www.cstl.nist.gov/biotech/strbase/str26plex.htm#Bins-and-Panels)

These projects were and are supported by Grant Number 2003-IJ-R-029 and 2008-DN-R-121, which is an interagency agreement between NIJ and the NIST Office of Law Enforcement Standards, awarded by the National Institute of Justice, Office of Justice Programs, US Department of Justice. Points of view in this document are those of the authors and do not necessarily represent the official position or policies of the US Department of Justice and US Department of Defense. Certain 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 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.

Kinship Casework

Using the miniSTR assays in combination with the 26plex

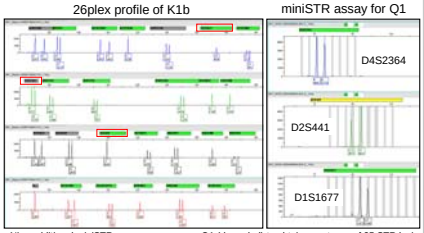
Case 1



Q1 unknown (degraded)
K1a brother
K1b maternal niece

K1a and K1b were typed with the NIST 26plex assay and Identifiler (Applied Biosystems)

Q1 was typed using 10 NIST miniplexes and MiniFiler (Applied Biosystems)



Nine additional miniSTR assays were run on Q1 (degraded) to obtain genotypes of 25 STR loci that overlapped with the 26plex loci

Likelihood Ratios (LR) were Calculated

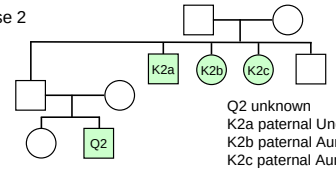
Allele frequency data for Identifiler and 26plex loci were obtained from: http://www.cstl.nist.gov/biotech/strbase/NISTpop.htm

System	LR (Brother K1a)	LR (Niece K1b)
26plex	3.75E+06	3.60
MiniFiler	295.62	1.92
Combined	1.11E+09	6.91

Full Sibling Calculation
The additional 25 STRs **significantly increased** the LR from ~300 to ~3.7 million
Uncle-Niece Calculation
However, the additional loci did **not** have a significant effect on the LR (due to lack of allele sharing).

LR calculations were performed with either DNA-View (Charles Brenner) or an Excel spreadsheet program (Steven Myers, California DOJ)

Case 2



Q2 unknown
K2a paternal Uncle
K2b paternal Aunt
K2c paternal Aunt

All samples were typed with the NIST 26plex assay and Identifiler (Applied Biosystems)

System	LR (Uncle K2a)	LR (Aunt K2b)	LR (Aunt K2c)
26plex	10.44	2.12	2.52
Identifiler	0.34	0.59	2.16
Combined	3.55	1.25	5.44

The additional 25 STRs **marginally increased** the LR for the Nephew-Uncle & Nephew-Aunt relationships. This is not unexpected given the amount of allele sharing expected from relatives outside of the traditional paternity trios. The use of lineage markers (Y and mitochondrial) might provide additional information in such a case.

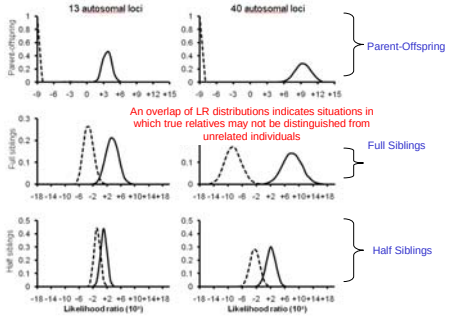
Simulated Kinship Testing

- One application of Kinship Testing is to find a close relative of an unknown individual using databases of 10-15 autosomal STR profiles. An association is used as an investigative lead to identify the unknown.
- However, the number of false positive leads is large and increases with the size of the database and the distance of the relationship to be evaluated.
- Can additional 25 autosomal STR loci from the NIST 26plex provide robust, sensitive, and specific kinship associations?

Methods

- NIST US Population samples were previously genotyped for 15 Identifiler loci (n=572). A majority of this database (n=551) was also genotyped for the 26plex loci for a total of 40 autosomal STRs.
- An unknown profile was simulated with a known relationship to a database profile by creating the appropriate family structures. Only one true kinship relationship was present in each of the 10,000 independent kinship tests. Parent-offspring, full sibling, and half sibling (or equivalent) relationships were tested.
- For each pairwise comparison between the evidence and database profiles, a likelihood ratio (LR) was calculated. After each search, LRs were ordered from highest to lowest. The ordered rank of the true pair was used to calculate the cumulative frequency of the ranks of the true relationships. These frequencies can be thought of as the probability of finding the true relative after investigating a given proportion of individuals in the database.

Results



Distributions of LRs for true relatives (solid lines) and for unrelated persons (dashed lines). The additional 25 STR loci improved discrimination of true relatives.

With 40 loci, the profile with the highest LR was the true relative in all 10,000 parent-offspring simulations, in 9,998 of 10,000 full sibling simulations, and in 6,574 of 10,000 half sibling simulations.

LR Threshold	Parent-Offspring PT	Parent-Offspring FP	Full Siblings PT	Full Siblings FP	Half Siblings PT	Half Siblings FP
13 Loci						
100	0.95	18	0.70	5	0.03	2
10	1.00	45	0.86	60	0.27	184
1	1.00	79	0.95	440	0.73	4,570
40 Loci						
100	1.00	0	0.97	0	0.27	3
10	1.00	0	0.99	1	0.57	50
1	1.00	0	1.00	5	0.83	607

The probability of finding a true relative (PT), if one exists in the database, and the number of false positives (FP) that would have to be investigated for different LR thresholds are set after kinship testing. Based on a database of 100,000 individuals with allele frequencies of the NIST samples.

There is a balance between setting a LR threshold high enough to minimize the number of false positives while maximizing the probability of finding true relatives.

For the example above (LR = 1), if 40 autosomal STRs were included in the profile, the true full sibling would be found with a probability of 1.00 and only five uninvolved persons would appear at least equally likely to be related.

Conclusions
- Searching profiles based on 40 loci increases the probability of finding a true parent-offspring, full sibling, or half sibling relative in the filtered list of likelihood ratio values, while decreasing the number of false positives.
- The additional 25 autosomal STR loci from the 26plex help provide robust, sensitive, and specific kinship testing.
- The use of additional genetic information protects the privacy of unrelated individuals (false positives) while providing more specific information for the true relative.

Acknowledgements
AFDIL Suzanne Barritt, Layla Shahmirzadi, Chris Johnson, Kimberly Sturk, Brion Smith and Lou Finelli
University of Washington Bruce Weir and Mary-Claire King

Poster available for download from STRBase http://www.cstl.nist.gov/biotech/strbase/pub_pres/ValloneISFG2009poster.pdf