

STRBase: 10 Years and Beyond

New Internet Resources for the Human Identity Testing Community

John M. Butler

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

For more than a decade, the U.S. National Institute of Standards and Technology (NIST) has maintained the **Short Tandem Repeat DNA Internet DataBase (STRBase)**, which is located at <http://www.cstl.nist.gov/biotech/strbase/>. The purpose of STRBase has been and continues to be an attempt to bring together the abundant literature and information in the human identity testing field in a cohesive fashion to make current and future work easier. STRBase recently became NIST Standard Reference Database (SRD) 130. New resources that are regularly added to STRBase can now be found quickly at <http://www.cstl.nist.gov/biotech/strbase/updates.htm>. A page was recently created to track null alleles detected through DNA testing with different primer sets (see <http://www.cstl.nist.gov/biotech/strbase/NullAlleles.htm>). In addition, a new software program enabling concordance checking with multiple data sets can be downloaded from <http://www.cstl.nist.gov/biotech/strbase/software.htm>. Hundreds of new pages of information have been added in the past few years and numerous PowerPoint slides, NIST publications and presentations, software programs, and other useful information is available for download.

NIJ-Funded Projects at NIST

<http://www.cstl.nist.gov/biotech/strbase/NIJprojects.htm>



Projects Performed by NIST Human Identity Project Team

- ABI 3100 performance with various STR typing systems
- ABI 3130xl upgrade evaluation
- AutoDimer: software to enable rapid multiplex PCR design
- Autosomal SNP loci
- Autosomal STR loci: beyond the CODIS markers
- Biometric dry storage device DNA stability studies
- Cat STR assay: development of the "MeowPlex"
- DNA quantitation methods: evaluation and comparison
- DNA recovery from aged bloodstains on untreated filter paper
- DNA recovery from bloodstains on various filter papers
- DNA recovery from various tubes
- DNA stability studies
- Dog STR assay development
- Expert Systems: FSS-3 software evaluation
- Genetic variation in U.S. populations
- Interlaboratory studies: NIST Mixed Stain Study #1 & #2
- Interlaboratory studies: NIST Mixed Stain Study #3
- Interlaboratory studies: DNA quantitation (QS04)
- Interlaboratory studies: STR mixture interpretation (MIX05)
- Literature collection on topics related to forensic DNA typing
- Locus-specific brackets (LSBs) for Y-STR typing
- Luminex system for Y-SNP typing
- Mass spectrometry SNP typing
- MiniFiler concordance study
- MiniFiler assays for recovery of information from degraded DNA
- miniSTR: new loci and assays/support to other labs
- Mitochondrial DNA: LINEAR ARRAY evaluation
- Mitochondrial DNA: coding region SNP assay development
- Mitochondrial DNA: Affymetrix Mixup evaluation
- mixSTR: a tool to aid mixture DNA cases
- Multiplex PCR assay development (2000-present)
- Multiplex_QA: development of data quality assessment software
- Nuclear DNA from hair shafts
- Quality control methods for PCR primers
- SNP typing: ancestry informative marker assays
- SNP typing: identification assays
- SNP website for possible forensic SNP loci
- SRM 2391, 2391a, 2391b: PCR-based DNA Profiling Standard
- SRM 2395: Human Y-Chromosome DNA Profiling Standard
- SRM 2372: Human X-DNA Quantitation Standard
- STRBase: a human identity testing community resource
- STR kit primer sequence analysis
- Supplying samples to collaborators and colleagues
- Training materials: STR typing using capillary electrophoresis
- Training materials: mixture interpretation and LCN DNA testing
- Validation resources, standardization, and training
- Variant allele sequencing to determine basis for allele dropout
- Y-chromosome: Duplication characterization
- Y-chromosome: Examination of the DYS464 multi-copy marker
- Y-chromosome: Y-SNP variation in U.S. populations
- Y-chromosome: Y-STR kit evaluations
- Y-chromosome: Y-STR multiplex assay development
- Y-chromosome: Y-STR mutation rate determination
- Y-chromosome: Y-STR value of additional loci

<http://www.cstl.nist.gov/biotech/strbase/>

Short Tandem Repeat DNA Internet Database

NIST Standard Reference Database SRD 130 [Recent Updates]

Serving the forensic DNA and human identity testing communities for over 10 years... These data are intended to benefit research and application of short tandem repeat DNA markers to human identity testing. The authors are solely responsible for the information herein.

This database has been accessed **258424** times since 10/02/97. (Counter courtesy www.digits.com - see disclaimer)

Created by **John M. Butler** and **Dennis J. Reeder** (NIST Biochemical Science Division), with invaluable help from **Jan Redman**, **Christian Ruitberg** and **Michael Tung**. Site creators' curriculum vitae available using links above.

Partial support for the design and maintenance of this website is being provided by The National Institute of Justice through the NIST Office of Law Enforcement Standards.

General Information

- Recent Updates**
- Purpose of STRBase
- Publications and Presentations from NIST Project Team
- NIJ-Funded Projects
- Training Materials
- Links to other web sites
- Glossary of commonly used terms

Forensic STR Information

- STRs101: Brief Introduction to STRs
- Core Loci:** FBI CODIS Core STR Loci and European Core Loci
- STR Fact Sheets (observed alleles and PCR product sizes)
- Multiplex STR kits**
- Sequence information (annotated)
- Variant Allele Reports
- Tri-Allelic Patterns
- Mutation Rates for Common Loci
- Published PCR primers
- Y-chromosome STRs
- miniSTRs (short amplicons) – including MiniFiler**
- Null Alleles - discordance observed between STR kits**
- STR Reference List - **now >2900 references**

Other DNA Marker Information and Non-Human Resources

- Amelogenin information**
- Sex-typing markers
- Sexual SNP information
- Mitochondrial DNA (mtDNA)
- Cat STR Information
- Dog STR information

Lab Resources and Tools

- Addresses for scientists working with STRs
- Training Materials
- STR Allele Sequencing
- Population data – **downloadable OmniPop program**
- Data from NIST U.S. Population Samples
- NIST-Developed Software including mixSTR and Multiplex_QA
- NIST Standard Reference Material for PCR-Based Testing
- New STR Markers under Development at NIST
- Chromosomal Locations
- DNA Advisory Board Quality Assurance Standards
- Interlaboratory Studies
- NIST Mixture 2005 Interlab Study MIX05 Data
- Validation information**
- DNA Quantitation – including information on SRM 2372
- Technology for resolving STR alleles

Planned Future Additions

- Mixture Interpretation – with worked examples
- Statistical Analysis Resources - with worked examples
- Known STR Flanking Region Variation
- Slides showing applications for STR typing

Acknowledgments: Jan Redman for entry of variant alleles, references, and address information. Funding support from the National Institute of Justice through interagency agreement 2003-IJ-R-029 to the NIST Office of Law Enforcement Standards.

References

- Ruitberg, C.M., Reeder, D.J., Butler, J.M. (2001) STRBase: a short tandem repeat DNA database for the human identity testing community. *Nucleic Acids Res.* 29: 320-322.
- Butler, J.M. and Reeder, D.J. (1997) STRBase: a short tandem repeat DNA database for the DNA typing community. *Profiles in DNA, Promega Corporation*, Volume 1, No. 2, p. 10. http://www.promega.com/profiles/102/ProfilesDNA_102_10.pdf
- Butler, J.M., Ruitberg, C.M., Reeder, D.J. (1998) STRBase: a short tandem repeat DNA internet-accessible database. *Proceedings of the Eighth International Symposium on Human Identification 1997*, Promega Corporation, pp. 38-47. <http://www.promega.com/geneticidproc/us8symproc13.html>
- Butler, J.M. (2007) New resources for the forensic genetics community available on the NIST STRBase website. *Progress in Forensic Genetics* 12 (in press).

New Null Allele Section for Tracking Discordance Between Primer Sets

Results from Concordance Studies

Link to Publication (NIST only)

Schematic of Kit Loci Size Ranges and Dye Labels

D16S539 with SRM 2391b Component 8

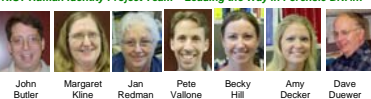
STR Fact Sheet with Locus Chromosomal Location and All Known Alleles and PCR Primers

http://www.cstl.nist.gov/biotech/strbase/str_fact.htm

| Other Names | Chromosomal Location | GenBank | Accession |
|-------------|------------------------------------|------------|----------------|
| D16 | 16q24.1 | G09253 | has 11 repeats |
| 16A10 | Chr 16: 84,944 kb (May 2004, NCBI) | AC024913.3 | has 11 repeats |
| 16A10 | 16q24.1 | | |

| Allele (Repeat #) | Seq 1 | Seq 2 | Seq 3 | Seq 4 | Repeat Structures | Ref. |
|-------------------|-------|-------|-------|-------|-------------------|------|
| 4 | 260Tg | 129Tg | 229Tg | 248Tg | variant allele | |
| 5 | 264Tg | 133Tg | 233Tg | 252Tg | IGATAL | 721 |
| 6 | 268Tg | 137Tg | 237Tg | 256Tg | variant allele | |
| 7 | 272Tg | 141Tg | 241Tg | 260Tg | variant allele | |
| 8 | 276Tg | 145Tg | 245Tg | 264Tg | IGATAL | 721 |
| 9 | 280Tg | 149Tg | 249Tg | 268Tg | IGATAL | 721 |
| 9.5 | 283Tg | 152Tg | 252Tg | 271Tg | variant allele | |
| 10 | 284Tg | 153Tg | 253Tg | 272Tg | IGATAL | 721 |
| 11 | 286Tg | 157Tg | 257Tg | 276Tg | IGATAL | 721 |
| 11.5 | 291Tg | 160Tg | 260Tg | 279Tg | variant allele | |
| 12 | 292Tg | 161Tg | 261Tg | 280Tg | IGATAL | 721 |
| 12.1 | 293Tg | 162Tg | 262Tg | 281Tg | variant allele | |
| 12.2 | 294Tg | 163Tg | 263Tg | 282Tg | variant allele | |
| 13 | 296Tg | 165Tg | 265Tg | 284Tg | IGATAL | 721 |
| 13.1 | 297Tg | 166Tg | 266Tg | 285Tg | variant allele | |
| 13.5 | 299Tg | 168Tg | 268Tg | 287Tg | variant allele | |
| 14 | 300Tg | 169Tg | 269Tg | 288Tg | IGATAL | 721 |
| 14.3 | 303Tg | 172Tg | 272Tg | 291Tg | variant allele | |
| 15 | 304Tg | 173Tg | 273Tg | 292Tg | IGATAL | 721 |
| 16 | 308Tg | 177Tg | 277Tg | 296Tg | variant allele | |

NIST Human Identity Project Team – Leading the Way in Forensic DNA...



- NIST publications and presentations:** <http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm>
- NIST-developed software:** <http://www.cstl.nist.gov/biotech/strbase/software.htm>
- NIST-collected population data:** <http://www.cstl.nist.gov/biotech/strbase/NISTpop.htm>

Poster available for download from STRBase: http://www.cstl.nist.gov/biotech/strbase/pub_pres/Butler_Promega2007poster.pdf

Disclaimer

This project was funded by the National Institute of Justice through interagency agreement 2003-IJ-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 US Department of Justice. 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.