

STRBase: A Short Tandem Repeat DNA Database

By John M. Butler* and Dennis J. Reeder
National Institute of Standards and Technology,
Biotechnology Division, Gaithersburg, MD

To make short tandem repeat (STR) DNA data more readily available to the DNA typing community, we have gathered over 450 references from 35 sources (including conference proceedings and book chapters) and have compiled them into an internet accessible database which we have named "STRBase". This web site brings together the abundant literature on STRs in a cohesive fashion which we hope will make future work in this field easier.

This database includes STR Fact Sheets that contain information on commonly used markers. These fact sheets show observed alleles (and reported microvariants) together with their repeat structure and PCR product sizes using published primers. Hyperlinks are included to GenBank® (U.S. Department of Health and Human Services), which makes the entire sequence for each STR locus readily

accessible. Additionally, a list of population studies and references specific to each STR locus can be reached via hyperlinks. Multiplex sets of STRs and commercial sources for allelic ladders are also hyperlinked. In addition, STRBase includes a chromosomal index of identity testing markers.

We located over 750 population studies reported in the literature and documented them in STRBase by listing the STR system, the population examined, the number of unrelated individuals tested, and the reference. This index of population studies provides a valuable tool for locating references that contain STR allele frequencies to aid in calculating matching probabilities for DNA typing cases.

A brief review of techniques that have been successfully applied to resolving and detecting STR alleles is included. The established tech-

niques of polyacrylamide gel electrophoresis with silver staining, fluorescent scanning and automated fluorescent detection systems are discussed along with emerging methods such as capillary electrophoresis, capillary array electrophoresis, microchip CE analysis and time-of-flight mass spectrometry. Hyperlinks to groups and individuals working in STR research are also included. Organizations involved in DNA typing, commercial sources of instrumentation or DNA typing systems, paternity testing laboratories, electronic journals containing STR-related publications and other useful DNA databases (e.g., GenBank® and the Genome Data Base) are listed and hyperlinked. If your laboratory or company has a web page which is not included in the database, please contact us and your link will be added.

STRBase will be available to the general public later this year through the world wide web at www.str.nist.gov/home.html. It is currently available at the temporary location <http://ibm4.carb.nist.gov:8800/dna/home.htm>.

The literature will be searched regularly for new publications, and updates to STRBase will be made periodically. Comments on the database, suggestions for further improvements or submissions should be sent to Dr. Dennis J. Reeder, (Attn: STRBase), National Institute of Standards and Technology, Biotechnology Division, Building 222, Room A353, Gaithersburg, MD 20899 or by e-mail: dennis.reeder@nist.gov. We hope this database will serve as a valuable tool for the continued development and application of STR systems to the field of human identity testing.

*Current address: GeneTrace Systems, Inc., 333 Ravenswood Avenue, Menlo Park, CA 94025
GenBank is a registered trademark of the U.S. Dept. of Health and Human Services.

