Manfred Kayser,[1] *Ph.D.; Silke Brauer;*[1] *Sascha Willuweit;*[2]
Hiltrud Schädlich,[1] *Ph.D.; Mark A. Batzer,*[3] *Ph.D.; Jennifer Zawacki;*[4]
Mechthild Prinz,[4] *M.D.; Lutz Roewer,*[2] *Ph.D.; and Mark Stoneking,*[1] *Ph.D.*

Online Y-chromosomal Short Tandem Repeat Haplotype Reference Database (YHRD) for U.S. Populations*

REFERENCE: Kayser M, Brauer S, Willuweit S, Schädlich H, Batzer MA, Zawacki J, Prinz M, Roewer L, Stoneking M. Online Y-chromosomal short tandem repeat haplotype reference database (YHRD) for U.S. Populations. J Forensic Sci 2002;47(3):513–519.

ABSTRACT: We describe here an online Y-chromosomal short tandem repeat haplotype reference database (YHRD) for U.S. populations, which represents 9-locus Y-STR haplotypes for 1705 African-Americans, European-Americans and Hispanics as of October 2001. This database is available online (http://www.ystr.org/usa/), free to access and was generated in order to supply the U.S. forensic DNA community with a valuable resource for frequencies of complete or incomplete 9-locus Y-STR haplotypes, as well as information about typing protocols and population genetic analyses. Pairwise R_{ST}-statistics derived from the Y-STR haplotypes indicate no significant substructure among African-American populations from different regions of the U.S., nor (usually) among European-American and Hispanic populations. Thus, pooling of Y-STR haplotype data from regional populations within these three major groups is appropriate in order to obtain larger sample sizes. However, pooling of different major populations is generally not recommended due to statistically significant differences between African-American populations and all European-American / Hispanic populations, as well as between some European-American and Hispanic populations.

KEYWORDS: forensic science, Y chromosome DNA, short tandem repeats, STRs, database, YHRD, African-American, European-American, Hispanic

Haplotypes based on Y-chromosomal short tandem repeat (STR) or microsatellite markers are becoming a powerful tool to identify and characterize male DNA in forensic analysis and paternity testing, and have already been used in court cases in various countries (1–8). Y-STR haplotypes are particularly useful in characterizing male culprit DNA in material from sexual assault/forcible rape cases (4,8–10). This is of special importance in countries like the United States, where the rate of rape cases is both high and increasing. In 1998, 1,531,044 cases of violent crime were recorded of which 93,103 (6.1%) were cases of forcible rape (11). This equals an average rate (or crime index) of about 34 cases of forcible rape per every 100,000 U.S. inhabitants, which is about 3.5 times the rate observed in 1960. Since 99.6% of offenders in cases of forcible rape are males (12), DNA identification of the culprit necessitates identifying the male component in a sample that typically consists of a mixture of male and female cellular material.

Conventionally, autosomal DNA markers are analyzed in sexual assault cases by separation of the female (epithelial cell) component of the evidence material (e.g., vaginal swabs of a rape victim) from the male (sperm cell) component by a differential lysis procedure (13). However, it can be difficult to achieve a complete separation of the male and the female components; in many cases the resulting DNA mixtures still contain a high amount of female DNA, leading to preferential amplification of the female victim DNA when the sample is analyzed with autosomal DNA markers (10). Also, in cases where small numbers of semen, or even male epithelial cells (e.g., from azoospermic males) are mixed with high numbers of female victim epithelial cells, the differential lysis procedure will not result in any enrichment of male cells.

An alternative strategy is to focus exclusively on Y chromosome markers, e.g., Y-STR haplotypes, which allows the male contribution to total genomic DNA isolated from the evidence to be analyzed directly, even in the presence of large amounts of female DNA, thereby avoiding the necessity of separation of the male and female components (4; 8–10). Moreover, the haploid nature of the Y chromosome simplifies the analysis of samples containing DNA from more than one male, and the hypervariability of Y-STR haplotypes potentially permits ready identification of multiple male culprits in rape cases, as has indeed already been shown in routine case work (8).

An additional application of Y-STR haplotypes involves deficiency cases of disputed paternity of a male offspring, where the alleged father is not available for DNA analysis (1,2). Since the Y chromosome is inherited unchanged from fathers to sons, unless rare mutations occur (14,15), any male relative of the deceased alleged father can potentially be used to replace the alleged father in Y-STR haplotype analysis, in order to test for paternity of a male

[1] Department for Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany.
[2] Institute for Legal Medicine, Humboldt-University Berlin, Berlin, Germany.
[3] Department of Biological Sciences, Louisiana State University, Baton Rouge, LA.
[4] Department of Forensic Biology, Office of Chief Medical Examiner, New York City, NY.

* This project was supported by grants from the National Institute of Justice: 98–IJ–CX–0014 (M.S.), 99–IJ–CX–K009 (M.A.B.), the Louisiana Board of Regents Millennium Trust Health Excellence Fund HEF (2000–05)–05 and HEF (2000–05)–01 (M.A.B.), and by the Max Planck Society and was presented at the American Academy of Forensic Sciences meeting in Seattle, February 19–24, 2001.

Received 8 June 2001; and in revised form 18 Oct. 2001; accepted 18 Oct. 2001.

offspring. However, in general it has to be remembered that with Y-chromosome markers one is only able to characterize paternal lineages and thus differentiate between paternal lineages but not between individuals within a particular lineage. Thus, in cases of non-exclusions in forensic applications, every member of a particular paternal lineage has equal probability of being the biological father in a paternity case or the culprit in a crime case (2).

As in any DNA analysis, in the event of a match between the Y-STR haplotypes for a case sample and a suspect sample, it is desirable to have an estimate of the probability that a match would occur by chance. Multiplication of single locus allele frequencies to obtain estimated Y-STR haplotype frequencies is not appropriate since all STRs are from the non-recombining portion of the Y chromosome and hence are completely linked. As with mitochondrial DNA, databases of complete Y-STR haplotypes have to be generated as a source for estimating frequencies. Because of the much higher diversity of combined Y-STR haplotypes compared with single Y-STR loci, such a database has to be much larger in order to be able to serve as a reliable representation of the underlying population haplotype frequencies. Also, since the Y chromosome is more sensitive to genetic drift due to its haploid and paternal mode of inheritance, populations are more likely to show statistically significant differences in regard to their Y-STR haplotypes (16). The potential for population structure must therefore be considered when generating Y-STR haplotype databases.

Recently, a large and growing Y-STR haplotype reference database (YHRD) for European populations has been made available (7,17). As of October 2001, this database contained 3805 different 9-locus Y-STR haplotypes from 7784 individuals of 58 populations of European ancestry. Here, we report the creation of a Y-STR haplotype reference database (YHRD) for U.S. populations, based on the same standard of nine loci. This YHRD for U.S. populations is available online and free of access (18) and has two major objectives. The first is to supply the U.S. forensic DNA community with a valuable resource for obtaining Y-STR haplotype frequencies needed for calculating matching or paternity probabilities in cases of non-exclusions in forensic analysis and paternity testing. The second is to test for significant geographic stratification and genetic heterogeneity based on Y-STR haplotypes among African-American, European-American, and Hispanic populations in the U.S.

Material and Methods

The data in this study were obtained from a total of 1705 individuals (30 populations) from 11 geographic regions in the United States: 599 African-Americans (10 populations), 628 European-Americans (11 populations) and 478 Hispanics (9 populations). The population origin of each individual was self-reported. All samples came from U.S. crime laboratories, with the exception of the Louisiana and Acadian samples (provided by M.B.). Most of the samples were received as dried bloodstains and DNA was extracted using the IsoQuick (Orca Research) extraction kit (19). Nine Y-STR loci [DYS19 (or DYS394), DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS385a/b] were analyzed in a pentaplex and a quadruplex PCR or alternatively in a single nanoplex PCR and detected on ABI PRISM 373 or 377 DNA Sequencers (Applied Biosystems). Detailed protocols are available from the website of the YHRD for U.S. populations (18). The discriminatory capacity was calculated as the number of individuals divided by the number of different haplotypes. Haplotype diversity was calculated as described elsewhere (20). R_{ST} values and associated probability values, estimated from 10,000 permutations, were calculated using the software package ARLEQUIN (21). A neighbor-joining tree was produced from the pair wise R_{ST} values using the relevant programs in PHYLIP (22) and viewed using the program TREEVIEW (23).

Results and Discussion

Y-STR Loci Characteristics/Haplotype Format

Nine Y-STRs were chosen to construct haplotypes: DYS19 (or DYS394), DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, and DYS385a/b. PCR primers for DYS385 amplify two polymorphic Y-specific loci, most likely due to a duplication of the Y chromosomal region that includes DYS385. DYS19 and DYS394 are synonymous loci amplified with different PCR primer pairs as revealed by DNA sequence analysis (Kayser et al. in preparation). For two of the nine Y-STR loci amplification products from the human X chromosome have been reported (DYS391, DYS393), which can be explained by sequence homology between parts of the human X and Y chromosome (24,25). However, since those sequences are not completely identical, X-chromosomal amplification can be avoided by stringent PCR conditions that are optimal for Y-specific primers and/or alternative primer design (24).

These loci have been chosen for the following reasons: high discrimination capacity (1,5,7,26); well-established molecular characteristics (1,5,14,15); large amount of haplotype data available from worldwide populations (7,16) locus specific mutation rates estimated from studies of confirmed father/son pairs (14,15); forensic validation according to the guidelines of the DNA advisory board (DAB) performed successfully, including accuracy, precision, reproducibility, species specificity, sensitivity, stability in different stain substrates, and mixture performance (8, 27, Sudhir Sinha, in preparation); collaborative studies across different laboratories and detection platforms successfully completed (1,28,29); all nine loci can be analyzed in one or two mutiplex reactions (Fig. 1); recommended for and successfully applied to court use (3,4,6,9,10,30); and nomenclature is in concordance with the International Society of Forensic Genetics (ISFG) guidelines for forensic STR analysis (31,32). The 9-locus Y-STR haplotype used here was recently called the "minimal haplotype," indicating that these nine Y-STR loci are required as backbone haplotype for the characterization of Y chromosome lineages in forensic applications due to their sufficient molecular and forensic evaluation (5,7), which was also underlined recently in the recommendations on forensic analysis using Y-chromosome STRs formulated by the DNA commission of the ISFG (34). Already, several publications have appeared that include casework examples that demonstrate the reliability of the Y-STR approach based on these loci and emphasize the advantage of these Y-STR loci in cases where the male DNA is the minor contributor in a mixture (4,8–10,34,35).

Database Structure

The structure and the principles of the YHRD for U.S. populations have been adapted from the YHRD for European populations (7,17). The website of the YHRD for U.S. populations is free of access and consists of so far eight pages (18). The starting "About" page provides, in addition to the aims of the database, information about the current state of the database and the features of the Y-STR loci chosen for haplotype construction. The "Quality Control/Validation" page gives information about the quality assurance criteria under which the data were generated. This page also supplies de-

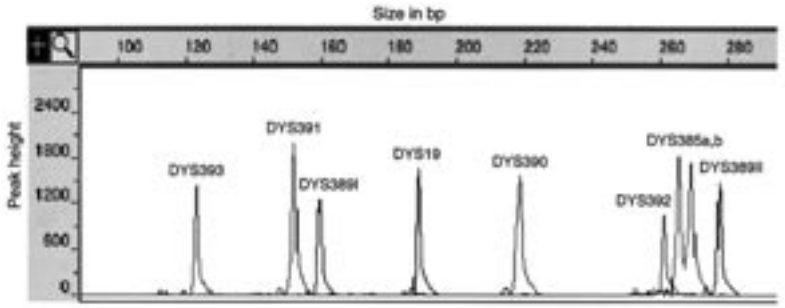

FIG. 1—*Electropherogram from simultaneous fragment–length analysis of nine Y–STRs (Nanoplex) using an ABI PRISM 377 DNA Sequencer and the Genescan software (Applied Biosystems).*

tailed information about the results of the Y-STR validation according to the guidelines of the DNA advisory board (DAB). The "Haplotype contribution" page gives detailed instructions on how to submit Y-STR haplotype data to the database. The "Primer & Protocols" page provides detailed protocols for the typing methods used to generate the data, including all primer sequences. The "Population Analysis" page gives essential results of population comparison analyses within and between African-American, European-American, and Hispanic populations, based on Y-STR haplotypes. A "Contact" and "Acknowledgment" page are also available. The "Start Search" page allows the user to enter and search for a Y-STR haplotype and provides all resulting query information.

Data Entry

Data entry into the YHRD for U.S. populations includes the complete 9-locus Y-STR haplotype according to the repeat nomenclature proposed previously (1), with the alteration of the allele nomenclature for the loci DYS389I and DYS389II, where in both cases three additional repeats have been included. Additional information entered in the database are the name of the population, including both a regional geographic characterization as well as the major U.S. population group (African-American, European-American, Hispanic) to which the typed individual belongs, and an individual identification number. The latter is not available for online users of the database, but allows for potential further inclusion of additional Y chromosome loci into the haplotype format of the database. No other individual information is stored in the database. As a prerequisite for data entry into the YHRD for U.S. population, participating laboratories must pass a quality control test that involves—consistent with the requirements of the YHRD for European populations—blind haplotyping of five DNA samples, which are available from the corresponding author. Once the quality test has been passed successfully, complete 9-locus haplotype data, together with the information about the geographic origin of the individuals typed, can be submitted electronically (*Microsoft Excel* table format). More details are available from the "Haplotype contribution" page of the website.

Database Query

The search for a particular Y-STR haplotype and its occurrence in the database only requires entering all single-locus alleles in the given mask at the "Start search" page. Although entering the complete 9-locus haplotype is recommended in order to obtain sufficient haplotype resolution, the database also allows entering of and searching for incomplete haplotypes, down to single loci. No information about the population origin of the individual of the entered haplotype is needed—the database program automatically compares the entered haplotype with all haplotypes in the database on the basis of a simple match search. The result of the Y-STR database query for the entered haplotype appears as the "Haplotype query summary," which provides the number of matches obtained for all African-American, European-American, and Hispanic populations. This enables the database user to calculate the haplotype frequency separately for each major population group. If the entered haplotype does not yet exist in the database, the response "no matches found" is given for the respective group. In the "Detailed information" section, the complete haplotype is shown including the matched population(s). In case an incomplete haplotype was entered, all matches appear also for those haplotypes that might differ by alleles at loci that were not included. This provides an indication of the further resolution that would be theoretically obtainable by typing more loci towards the complete minimal haplotype. As another feature of the query results, a map of the U.S. is shown with blue datapoints indicating all geographic regions that are covered by the database and red datapoints indicating regions with database populations where matches were observed. The legend to the map supplies the number of matches for each region. In the "Population query summary" all populations included in the database are listed in a table with the absolute number of matches per population and the total number of individuals typed per population. This enables the database user to calculate population-specific haplotype frequencies. Finally, a link to the YHRD for European populations (17) allows a further haplotype search against 3805 haplotypes from 7784 European individuals from 58 populations (as of October 2001).

Current State of the Database and Population Analyses

As of October 2001 the YHRD for U.S. populations contained 1705 individual entries, consisting of complete 9-locus Y-STR haplotypes, of 30 populations from 11 geographic regions of the United States of America: 599 African-Americans (10 populations), 628 European-Americans (11 populations), and 478 Hispanics (9 populations), (Fig 2). There were 1116 different 9-locus Y-STR haplotypes observed, i.e., 65.5% of all U.S. Americans analyzed are distinguishable by their individual 9-locus Y-STR haplotype (Table 1). The discriminatory capacity is even higher when the total population is divided into the three major groups (Table 1). The haplotype diversity was highest in the African-Americans (0.9982), next highest in the European-Americans (0.9957), and lowest in the Hispanics (0.9948); the overall haplotype diversity was 0.9974 (Table 1). Haplotype diversity can also be used to ad-

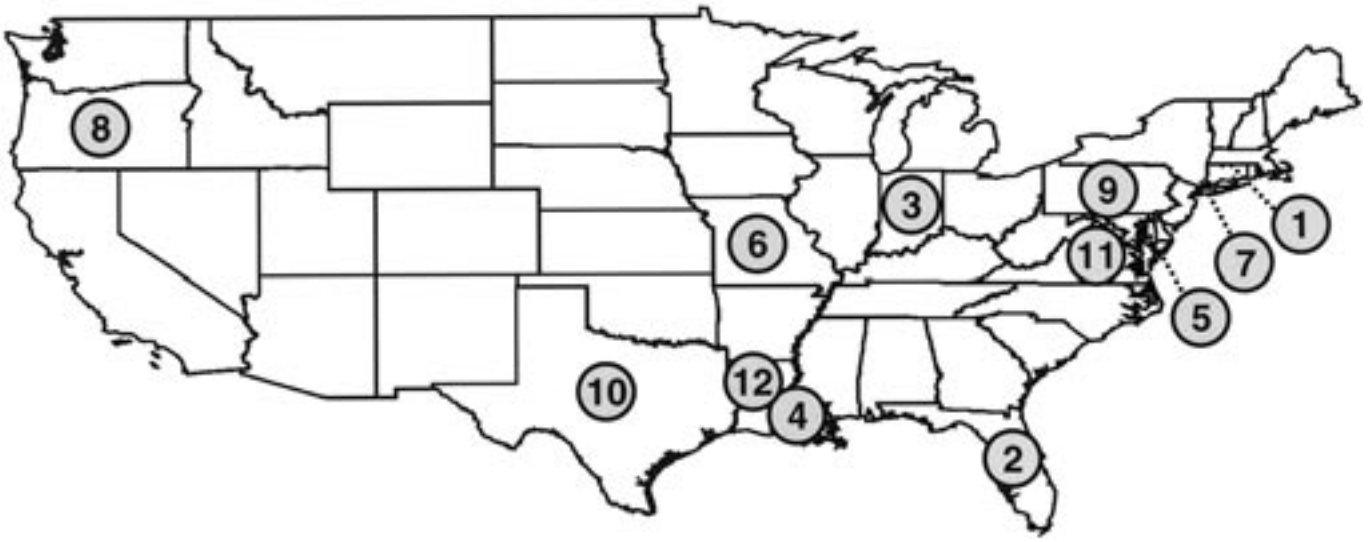

FIG. 2—*Map of the U.S. with 11 geographic regions covered by the YHRD for U.S. populations as of October 2001. 1: Connecticut; 2: Florida; 3: Indiana; 4: Lousiana; 5: Maryland; 6: Missouri; 7: New York City; 8: Oregon; 9: Pennsylvania; 10: Texas; 11: Virginia; 12: Acadian. From most of the regions all three major U.S. population groups (African–Americans, European–Americans and Hispanics) were sampled.*

TABLE 1—*Current state of the Y–chromosomal haplotype reference database (YHRD) for U.S. populations as of October 2001.*

	African–American	European–American	Hispanic	Total
No. of Individuals	599	628	478	1705
No. of Haplotypes	454	437	354	1116
Discriminatory Capacity	75.8%	69.6%	74.1%	65.5%
Haplotype Diversity	0.9982	0.9957	0.9948	0.9974
(± Standard Deviation)	(± 0.0003)	(± 0.0007)	(± 0.0011)	(± 0.0003)
No. of Haplotypes Observed Only Once (%)	377 (83.0)	361 (82.6)	301 (85.0)	896 (80.3)
Occurrence of Most Frequent Haplotype (%)	12 (2.0)	25 (3.98)	19 (3.97)	53 (3.1)

dress a question of practical interest: if multiple unrelated males have contributed to a sample, what is the probability that Y-STR haplotype profiling will reveal a mixture? In the simplest case, involving two males, the probability of detecting a mixture is the same as the haplotype diversity, about 99.7%. However, a complication arises in that multiple alleles are occasionally observed at single Y-STR loci (14). Although the frequency of multiple alleles is rare, about 0.12% (14), it is possible that such multiple alleles, which presumably reflect multiplication of at least the entire Y-STR locus, will be mistaken for mixtures (or vice versa) (15).

Figure 3 shows the Y-STR haplotype distribution for the three major population groups. Within all three groups the vast majority (83–85%) of haplotypes were observed only once, reflecting the large amount of Y-STR haplotype diversity (Table 1). The most frequent haplotype (MFH) in African-Americans (DYS19-DYS389I-DYS389II-DYS390-DYS391-DYS392-DYS393-DYS385a,b: 15-13-31-21-10-11-13-16,17) occurred with a frequency of 2%, was observed in 8 of the 10 African-American populations analyzed, and elsewhere only found in one Hispanic male. The MFH in European-Americans (14-13-29-24-11-13-13-11,14) was found with a frequency of 4%, occurred in 8 of the 11 European-American populations, was also one of two MFHs in Hispanics with a frequency of 4% (8 out of 9 populations), and occurred in 9 African-Americans (6 out of 10 populations). The other Hispanic MFH (13-14-30-24-9-11-13-13,14) was found with a frequency of 4% and occurred in 7 out of 9 Hispanic populations, but not elsewhere. Graphs showing the frequency of the 20 most frequent haplotypes per major population with their allelic designation are available from the "Population Analysis" page of the U.S. population YHRD website.

A comparison of the U.S. YHRD with the European YHRD (7784 individuals with 3805 different 9-locus haplotypes from 58 populations as of October 2001 (17)) indicates that haplotype diversity is lower in European-Americans (0.9957) and Hispanics (0.9948) than in the pooled European populations (0.9974), but that the discrimination capacity value is higher in European-Americans (69.6%) and Hispanics (74.1%) than in pooled Europeans (48.9%). Further expansion of the YHRD for U.S. populations may reveal whether this is an effect of the smaller sample size for the U.S. database. The MFH among European-Americans (14-13-29-24-11-13-13-11,14) was with 3.1% frequency also the MFH for general Europeans (240 out of 7784 individuals from 48 out of 58 populations). The MFH in Hispanics (13-14-30-24-9-11-13-13,14) was found in 20 out of 7784 European males (0.26%) from 12 populations, whereas the MFH in African-Americans (15-13-31-21-10-11-13-16,17) was not found among the entire European dataset.

Population differentiation was assessed by computing R_{ST} values between each pair of populations and also among the three major groups of populations. R_{ST} is an analogue of F_{ST} (36) that takes into account mutational differences between Y-STR haplotypes. Although R_{ST} between the pooled European-Americans and Hispanics

is much lower than between African-Americans and European-Americans/Hispanics, all pair wise comparisons between the pooled major groups show highly significant differences (Table 2). Within African-Americans all populations revealed non-significant differences. Also the differences between the majority of European-Americans (except four pairs including Texas and two pairs including Virginia) on the one hand and the majority of Hispanic populations (except four pairs including Texas) on the other hand were not statistically significant. This reflects the close relationship between populations of the same group origin. When comparing populations between the three major groups, all African-American populations

TABLE 2—*R_{ST} values (above diagonal) and their P–values estimated from 10,000 permutations (below diagonal) for three major U.S. population groups.*

	African–American	European–American	Hispanic
African–American	×	0.36518	0.23202
European–American	< 0.0001	×	0.05821
Hispanic	< 0.0001	< 0.0001	×

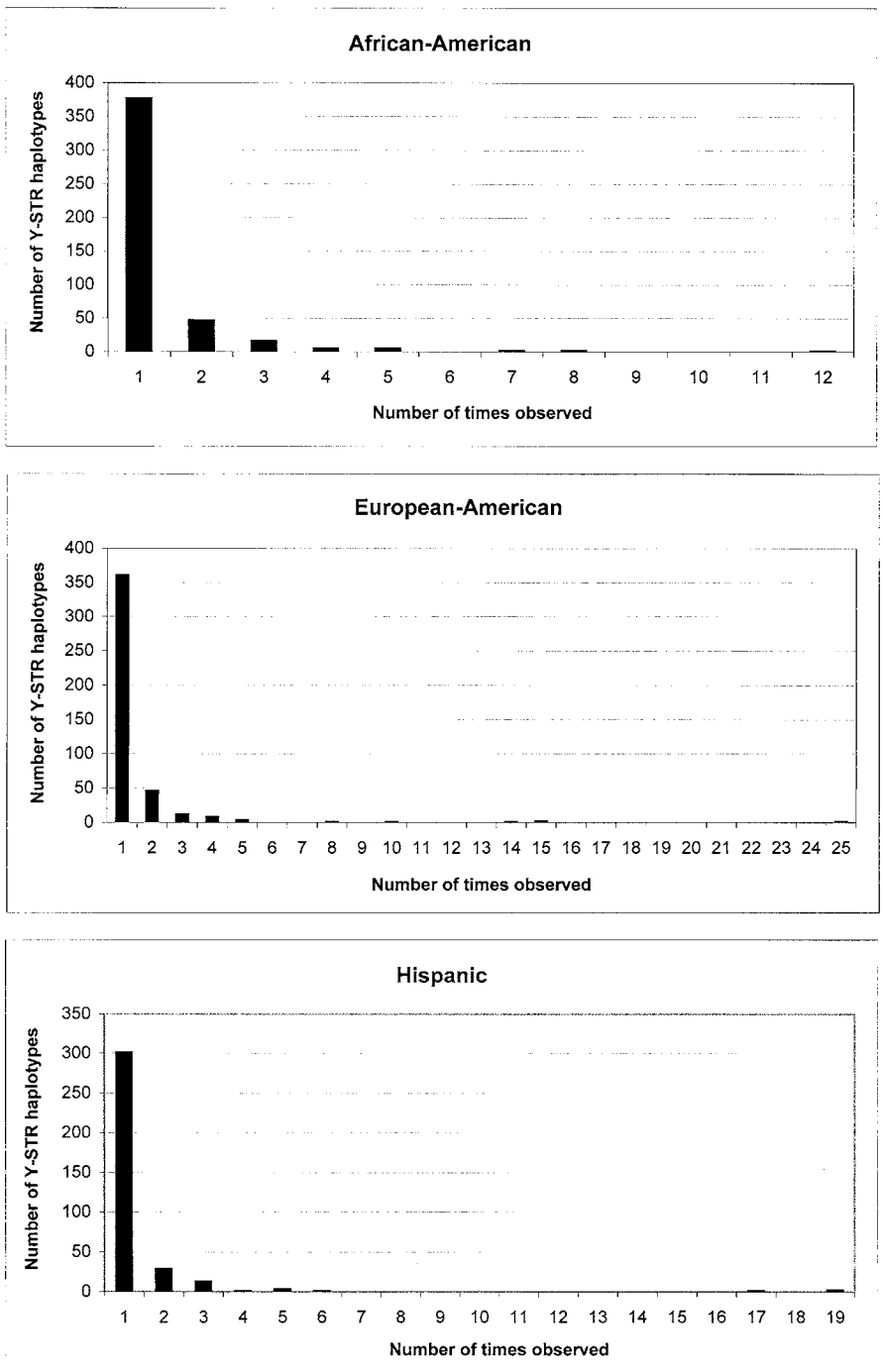

FIG. 3—*Distribution of Y–STR haplotypes in the three major U.S. population groups.*

were significantly different from all European-American and Hispanic populations, reflecting strong differences between African and non-African Americans with respect to their Y-STR haplotypes. However, comparing European and Hispanic populations to each other revealed non-significant differences in 62 of the 99 pair wise comparisons, with most of the significant differences between pairs including either the samples from Texas or Virginia. This reflects the closer relationship between populations of general European and Hispanic origin. These results are also evident from a NJ tree of the pairwise R_{ST} values where all European-American population (except Florida European-Americans) cluster together, separated from but still close to all Hispanic populations. European-American and Hispanic populations are strongly separated from all African-American populations, forming a distinct group (Fig. 4).

To summarize, based on their 9-locus Y-STR haplotypes, populations from the three major groups of the United States (African-Americans, European-Americans and Hispanics) are most similar to populations from the same major group. However, all African-American populations are significantly differentiated from European-Americans and Hispanics. European-Americans and Hispanic populations are different from each other in aggregate, but the differences are often not statistically significant on a population pair wise level.

Conclusions

The YHRD for U.S. populations is a dynamically programmed, publicly available database (18) and was generated to provide a valuable resource for obtaining Y-STR haplotype frequencies of African-Americans, European-Americans, and Hispanics needed for the calculation of matching probabilities in cases of non-exclusions in forensic DNA analysis. Based on Y-STR haplotypes and under the principle of R_{ST}-statistics, population substructure was not detectable between African-American populations from different geographic regions of the U.S., and usually not between different European-American and different Hispanic populations, respectively, so that pooling of regional populations within these three major groups is appropriate in order to obtain sufficiently large sample sizes. However, pooling major populations, e.g. European-American and Hispanic populations, is not recommended due to significant differences between populations of different major groups with respect to Y-STR haplotypes. At this stage the YHRD for U.S. populations should be seen as an initial attempt to generate a national U.S. database for Y-STR haplotypes, with the ultimate goal of including all major populations from as many geographic regions of the U.S. as possible. In order to enlarge the database we encourage the U.S. forensic DNA community to obtain Y-STR haplotype data from additional regional U.S. populations and submit them to the YHRD for U.S. populations.

Acknowledgments

We thank the following colleagues for providing blood or DNA samples: Bruce Budowle, Tamyra Moretti, Cecilia H. von Beroldingen, Teresa M. Long, Thomas Grant, Barbara Llewellyn, Chris Tomsey, Joanne B. Sgueglia, Mohammad A. Tahir, and Keith

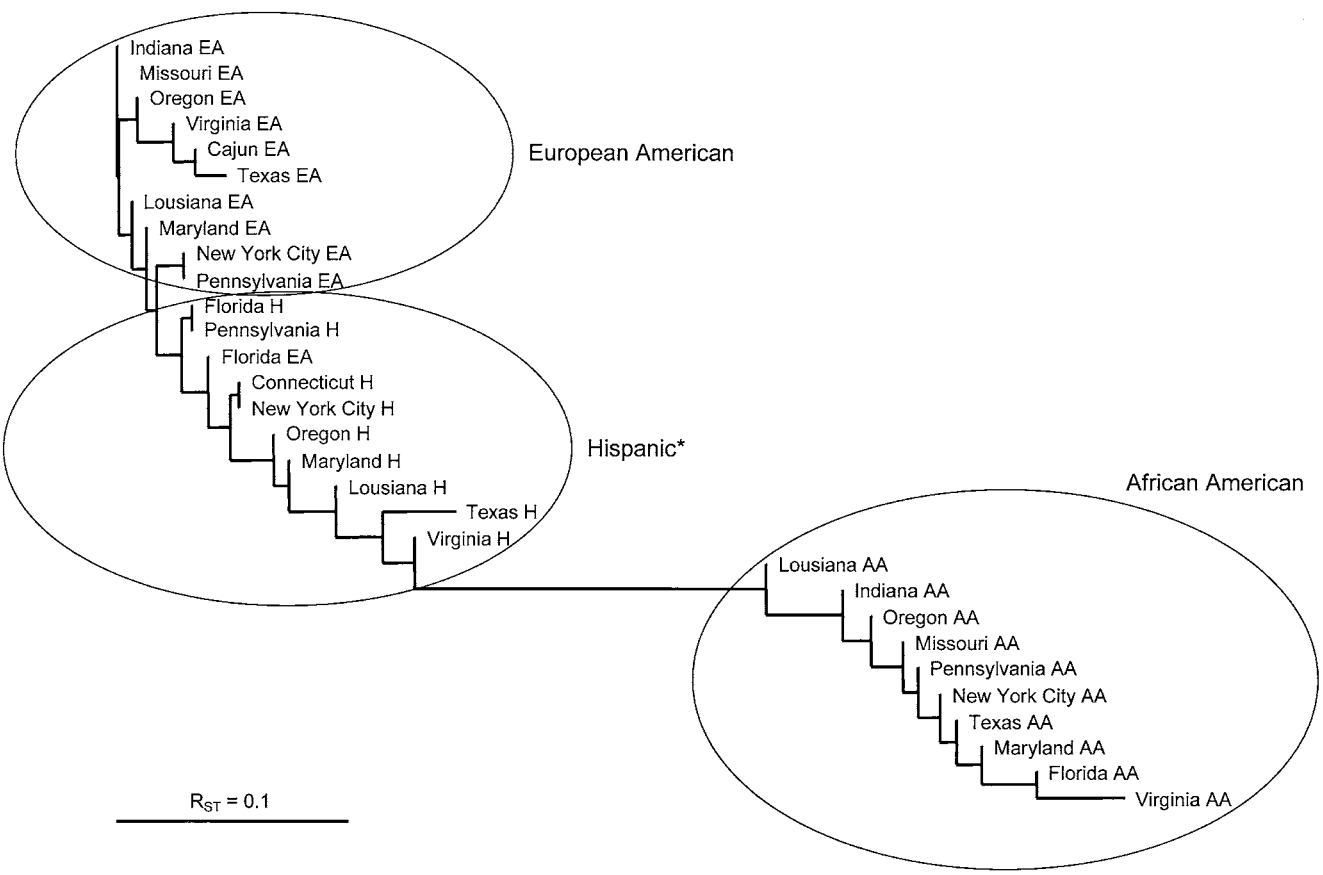

FIG. 4—*Neighbor–Joining tree of pairwise R_{ST} values with major U.S. population groups highlighted. * includes all Hispanic populations and Florida European–Americans.*

McKenney. In addition, we are grateful to Kevin Hiester and Stephanie Clifford for technical assistance in an early phase of the project.

References

1. Kayser M, Caglià A, Corach D, Fretwell N, Gehrig C, Graziosi G, et al. Evaluation of Y chromosomal STRs: a multicenter study. Int J Legal Med 1997;110:125–133,141–9.
2. Kayser M, Krüger C, Nagy M, Geserick G, de Knijff P, Roewer L. Y-chromosomal DNA-analysis in paternity testing: experiences and recommendations. In: Olaisen B, Brinkmann B, Lincoln PJ, editors. Progress in forensic genetics. 7: Amsterdam: Elsevier, 1998;494–6.
3. Pascali VL, Dobosz M, Brinkmann B. Coordinating Y–chromosomal STR research for the courts. Int J Legal Med 1998;112:1.
4. Honda K, Roewer L, de Knijff P. Male DNA typing from 25–year–old vaginal swabs using Y chromosomal STR polymorphisms in a retrial request case. J Forensic Sci 1999;44:868–72.
5. Roewer L, Kayser M, de Knijff P, Anslinger K, Betz A, Caglia A, et al. A new method for the evaluation of matches in non–recombining genomes: application to Y–chromosomal short tandem repeat (STR) haplotypes in European males. Forensic Sci Int 2000;114:31–43.
6. Corach D, Filgueira Risso L, Marino M, Penacino G, Sala A. Routine Y–STR typing in forensic casework. Forensic Sci Int 2001;118:131–5.
7. Roewer L, Krawczak M, Willuweit S, Nagy M, Alves C, Amorim A, et al. Online reference database of European Y–chromosomal short tandem repeat (STR) haplotypes. Forensic Sci Int 2001;118:106–13.
8. Prinz M, Ishii A, Colemann A, Baum HJ, Shaler RC. Validation and casework application of a Y chromosome specific STR multiplex. Forensic Sci Int 2001;120:177–88.
9. Betz A, Bassler G, Dietl G, Steil X, Weyermann G, Pflug W. DYS STR analysis with epithelial cells in rape cases. Forensic Sci Int 2001; 118:126–30.
10. Dekairelle AF, Hoste B. Application of a Y–STR pentaplex PCR (DYS19, DYS389I and II, DYS390, and DYS393) to sexual assault cases. Forensic Sci Int 2001;118:122–5.
11. http://www.disastercenter.com/crime/uscrime.htm
12. Greenfeld LA. Sex offenses and offenders: an analysis of data and rape and sexual assault. Washington D.C.: US Department of Justice, Bureau of Justice Statistics, 1997.
13. Gill P, Jeffreys AJ, Werrett D. Forensic application of DNA fingerprints. Nature 1985;318:577–9.
14. Kayser M, Roewer L, Hedman M, Henke L, Henke J, Brauer S, et al. Characteristics and frequency of germline mutations at microsatellite loci from the human Y chromosome, as revealed by direct observation in father/son pairs. Am J Hum Genet 2000;66:1580–8.
15. Kayser M, Sajantila A. Mutations at Y–STR loci: implications for paternity testing and forensic analysis. Forensic Sci Int 2001;118:116–21.
16. Kayser M, Krawczak M, Excoffier L, Dieltjes P, Corach D, Cagliá A, et al. Extensive analysis of chromosome Y microsatellite haplotypes in globally dispersed human populations. Am J Hum Genet 2001;68: 990–1018.
17. http://www.ystr.org/europe
18. http://www.ystr.org/usa
19. Buchanan AV, Sherry S, Weiss KM, McGarvey S, Neel JV, Stoneking M. Extraction of DNA from frozen red–blood cells. Hum Biol 1993; 65:647–54.
20. Nei M. Molecular evolutionary genetics. New York: Columbia University Press, 1987.
21. http://anthropologie.unige.ch/arlequin.
22. http://evolution.genetics.washington.edu/phylip.html
23. http://taxonomy.zoology.gla.ac.uk/rod/rod.html
24. Gusmao L, Sanchez–Diz P, Gonzales–Neira A, Brion M, Lareu MV, Amorim A, et al. Homologous STRs in the X and Y chromosome: a contribution to understanding human sexual chromosome evolution. In: Sensabaugh GF, Lincoln P, Olaisen B, editors. Progress in Forensic Genetics. 8: Amsterdam: Elsevier Science B.V., 2000.
25. Dupuy BM, Gedde–Dahl T, Olaisen B. DXYS267: DYS393 and its X chromosome counterpart. Forensic Sci Int 2000;112:111–21.
26. Caglia A, Dobosz M, Boschi I, d' Aloja E, Pascali VL. Increased forensic efficiency of a STR–based Y–specific haplotype by addition of the highly polymorphic DYS385 locus. Int J Legal Med 1998;111:142–6.
27. DNA Advisory Board. Quality assurance standards for forensic DNA testing laboratories. Washington D.C.: US Dept. of Justice, FBI Publications, 1998.
28. Schneider PM, d'Aloja E, Dupuy BM, Eriksen B, Jangblad A, Kloosterman AD, et al. Results of collaborative study regarding the standardization of the Y–linked STR system DYS385 by the European DNA Profiling (EDNAP) group. Forensic Sci Int 1999;102:159–65.
29. Carracedo A, Beckmann A, Bengs A, Brinkmann B, Caglia A, Capelli C, et al. Results of a collaborative study of the EDNAP group regarding the reproducibility and robustness of the Y chromosome STRs DYS19, DYS389I, and II, DYS390 and DYS393 in a PCR pentaplex format. Forensic Sci Int 2001;119:28–41.
30. Prinz M, Ishii A, Sansone M, Baum HJ. Y chromosome specific STR testing and the US legal system. In: Sensabaugh GF, Lincoln PJ, Olaisen B, editors. Progress in Forensic Genetics. 8: Amsterdam: Elsevier Science B.V., 2000;591–4.
31. DNA recommendations–1994 report concerning further recommendations of the DNA commission of the ISFH regarding PCR based polymorphisms in STR (short tandem repeat) systems. Int J Legal Med 1994;107:159–60.
32. DNA recommendations–further report of the DNA commission of the ISFH regarding the use of tandem repeat systems. Forensic Sci Int 1997;87:179–84.
33. Gill P, Brenner C, Brinkmann B, Budowle B, Mayr E, Jobling M, et al. International Society of Forensic Genetics recommendations on forensic analysis using Y–chromosome STRs. Int J Legal Med 2001;114:305–9.
34. Prinz M, Boll K, Baum H, Shaler B. Multiplexing of Y chromosome specific STRs and performance for mixed samples. Forensic Sci Int 1997;85:209–18.
35. Gehrig C, Hochmeister M, Budowle B. Swiss allele frequencies and haplotypes of 7 Y–specific STRs. J Forensic Sci 2000;45:436–9.
36. Wright S. The interpretation of population structure by F–statistics with special regards to systems of mating. Evol 1965;19:395–420.

Additional information and reprint requests:
Dr. Manfred Kayser
Max Planck Institute for Evolutionary Anthropology
Department for Evolutionary Genetics
Inselstrasse 22
D–04103 Leipzig, Germany
Ph: +49–341–9952506
Fax: +49–341–9952555
E–mail: kayser@eva.mpg.de