ACCURACY, PRECISION, AND SITE-TO-SITE REPRODUCIBILITY IN ANALYSIS OF DNA POLYMORPHISMS FOR IDENTITY TESTING

I. Balazs, J.Chimera,* A. Eisenberg, D. Endean,** R. Giles,*** J. Kateley,**** S. Maguire, B. Schall, A. Turck, R. McKee

Lifecodes Corporation, Saw Mill River Road, Valhalla, NY 10595, USA
* Roche Biomedical Laboratories, Burlington, NC
** The Blood Center of Southeastern Wisconsin, Milwaukee, WI
*** GeneScreen, Dallas, TX
**** Sparrow Hospital, Lansing, MI

INTRODUCTION

Interest in using RFLP analysis of DNA for identity testing is growing rapidly. As an increasing number of diverse laboratories begin practicing the technique, the concerns over the reproducibility of the methodology and over the utility of the data generated will continue to grow. Since the procedure for generating an RFLP analysis is composed of a large number of steps, there are potentially a large number of opportunities to introduce variances from lab to lab. The concerns can be categorized into two broad areas. First, it must be shown that the RFLP result (ie. allele sizes) obtained from any given sample does not vary significantly when determined in various laboratories. Second, we must assure ourselves that data bases used for determining probabilities of identity or paternity are valid. This paper is intended to begin compiling data to address these concerns.

Five different laboratories have used identical protocols and sets of reagents to process blood samples for paternity testing. The data generated have been used to estimate the magnitude of site-to-site variations in RFLP analyses and the degree of coincidence of databases constructed with data from the various laboratories.

ALLELE SIZE

DNA was extracted with a simplified, non-organic method which gives a high yield of intact DNA which is suitable for RFLP analysis (Turck 1989). Samples were restricted with PstI and electrophoresed on a 28 cm 0.6% agarose gel. One lane of each gel contained 500 ng of K562 DNA which had been digested with PstI. DNA was transferred from the gel to a nylon membrane according to a modified Southern procedure (Southern 1975), and the membranes were then hybridized with ^{32}P-labelled DNA probes for the loci D2S44, D14S13, and D17S79. After autoradiographs were prepared, all alleles from the samples and from the K562 control DNA were sized using a digitizing tablet along with sizing software provided by Lifecodes.

The procedures employed included various standards and controls for two reasons. The first is to allow the monitoring of each stage of the analysis (extraction, restriction, electrophoresis,

Advances in Forensic Haemogenetics 3
Edited by H. F. Polesky and W. R. Mayr

Southern transfer and hybridization). The second reason is to provide the means for proper measurement of the allele sizes from the resulting autoradiographs. Accordingly, the sizing of alleles on each membrane includes the measurement of the alleles from the K562 control DNA. Obtaining consistent values in the same laboratory over a period of time for the K562 alleles provides an important quality assurance parameter.

Similarly, consistent sizes of the K562 alleles obtained at different laboratories ameliorates concerns of site-to-site variations in RFLP analysis of DNA for identity and paternity testing. An analysis of the values obtained from a multitude of sizings of K562 alleles in three laboratories is shown in Table 1.

Table 1. Allele Sizes for Three Loci of K562 DNA

Locus	N	Mean (kb)	S.D. (kb)	% Err	Range (kb)
D2S44 (1)	203	11.963	0.058	0.49	0.303
D2S44 (2)	203	10.859	0.068	0.62	0.471
D13S14	112	5.011	0.022	0.45	0.170
D17S79 (1)	230	4.007	0.023	0.57	0.171
D17S79 (2)	203	3.552	0.021	0.60	0.162

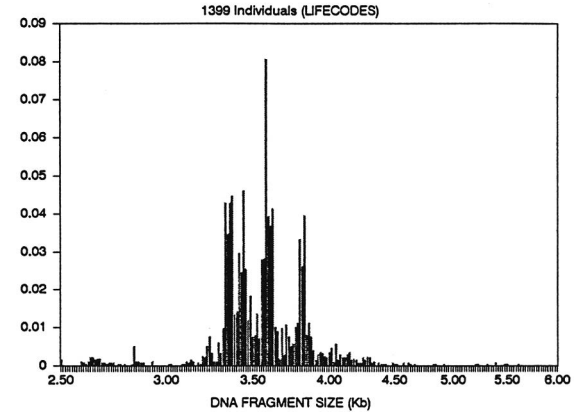
The maximum standard deviation corresponds to 0.62% of the DNA fragment size, which is consistent with a previous report which claimed a 0.6% error in the measurement of allele sizes by RFLP analysis (Balazs 1989). The same error is observed over a size range of at least 12 kb to 3.5 kb. Thus, by applying the three standard deviation rule, one is 99% certain that a DNA fragment measured to be 12 kb is between 11.784 kb and 12.216 kb, or a fragment measured to be 3.5 kb is between 3.446 kb and 3.554 kb. Furthermore, the data are highly consistent among the laboratories. The maximum difference between the mean of an allele size determined in one laboratory and the combined mean for all laboratories is 0.4%.

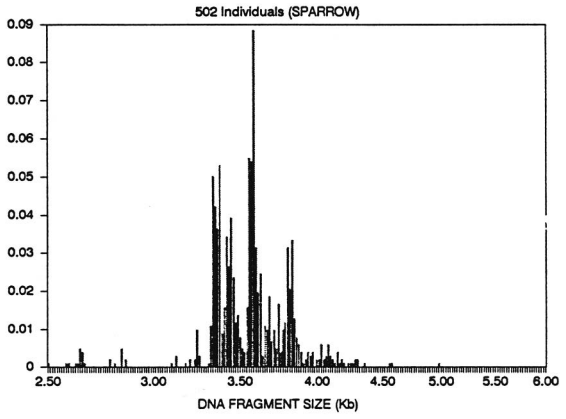
A second approach to determining the level of reproducibility of this technology was to split samples for analysis in two laboratories. Five laboratories participated in the study. Samples submitted for paternity testing to four of the laboratories where split. One aliquot was analyzed in the receiving laboratory and the other was analyzed in our laboratory. Allele sizes for all samples were determined, and the data were subjected to a pairwise analysis. The percent difference in the measured values of each allele was calculated as $[(A_L-A_B)/((A_L+A_B)/2]*100$, where A_L was the size measured in our laboratory and A_B was the size measured at the receiving laboratory. Table 2 compiles the data for the 877 pairs of sizes (two sizings for each allele) that were analyzed, including 59 complete paternity trios. In every case, both laboratories arrived at the same determination of inclusion or exclusion of the alleged father. The average standard deviation of the per cent differences of the data from the four beta sites participating in the study was 0.85.

Table 2. Variations in Two Independent Determinations of 877 Allele Sizes

Beta Site	N	Range of % Difference	Standard Deviation
1	261	(-1.95) - (2.40)	0.86
2	386	(-2.46) - (2.73)	0.74
3	159	(-4.80) - (3.01)	0.95
4	71	(-3.36) - (2.12)	0.85

ALLELE FREQUENCY DISTRIBUTION





In order to validate a database of VNTR alleles, it must be shown that the frequency distribution is not biased by sampling error. Allele frequency distributions for PstI derived alleles at the loci D2S44, D17S79, and D14S13 were constructed by independent laboratories (Lifecodes, The Blood Center of Southeastern Wisconsin, Roche Biomedical Laboratories, and Sparrow Hospital). The distributions do not show statistically significant differences. As an example, Fig. 1 shows the data obtained for the D17S79 locus in American Caucasians.

Fig. 1 Allele frequency distributions of D17S79 alleles in PstI-digested DNA from American Caucasians as determined by two independent laboratories.

References

Balazs I, Baird M, Clyne M, Meade E (1989) Human population genetic studies of five hypervariable DNA loci. Am J Hum Genet 44:182-190

Southern E (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503-517

Turck A, Schall B, Maguire S, Belluscio L, Nawoschik S, McKee R, Grimberg J, Eisenberg A (1989) A simple and efficient non-organic procedure for the extraction of DNA from evidentiary samples. These proceedings