Bruce Budowle,[1] *Ph.D., Keith L. Monson,*[1] *Ph.D., Alan M. Giusti,*[1] *B.S., and Barry L. Brown,*[1] *Ph.D.*

Evaluation of Hinf I-Generated VNTR Profile Frequencies Determined Using Various Ethnic Databases

REFERENCE: Budowle, B., Monson, K. L., Giusti, A. M., and Brown, B. L., "**Evaluation of Hinf I-Generated VNTR Profile Frequencies Determined Using Various Ethnic Databases,**" *Journal of Forensic Sciences,* JFSCA, Vol. 39, No. 4, July 1994, pp. 988–1008.

ABSTRACT: Concerns have been raised about hypothetical problems arising from the use of statistics for determining the likelihood of occurrence of DNA profiles for forensic purposes. A major contention is that reference databases based on subgroups of a major population category rather than on general (or major) population groups, might yield large differences in the estimated likelihood of occurrence of DNA profiles. This hypothetical issue is based on the assertion by some people that the differences among subgroups within a race would be greater than between races (at least for forensic purposes). To evaluate the effects of the above concern the likelihood of occurrence of 615 Hinf I-generated target DNA profiles was estimated using fixed bin frequencies from various ethnic databases and the multiplication rule. Based on the data in this study, differences in allele frequencies at a particular locus do not have substantial effects on VNTR profile frequency estimates when subgroup reference databases from within a major population group are compared. In contrast, the greatest variation in statistical estimates occurs across-major population groups. Therefore, the assertion, by some critics that the differences among subgroups within a race would be greater than between races (at least for forensic purposes), is unfounded. The data in the study support that comparisons across major population groups provide valid estimates of DNA profile frequencies without forensically significant consequences. The data do not support the need for alternate procedures, such as the ceiling principle approach, for deriving statistical estimates of DNA profile frequencies.

KEYWORDS: pathology and biology, VNTR, allele frequency, population databases

Concerns have been raised about hypothetical problems arising from the use of statistics for determining the likelihood of occurrence of DNA profiles for forensic purposes [*1–4*]. The contention is that reference databases based on subgroups that comprise the United States, rather than on general (or major) population groups, might yield large differences in the estimated likelihood of occurrence of DNA profiles. Therefore, it would be necessary to assess the frequencies of DNA profiles in a variety of subpopulation groups before

Received for publication 22 Jan. 1993; revised manuscript received 31 March 1993 and 25 Jan. 1994; accepted for publication 27 Jan. 1994.

[1] Program Manager for DNA Research, Research Chemist, Chemist, and Research Chemist, respectively, Forensic Science Research and Training Center, FBI Academy, Quantico, VA.

This is publication number 93-07 of the Laboratory Division of the Federal Bureau of Investigation. Names of commercial manufacturers are provided for identification only and inclusion does not imply endorsement by the Federal Bureau of Investigation.

Copyright © 1994 by ASTM International

TABLE 1—*Reference databases and loci analyzed for Hinf I-based data.*[a]

Laboratory	D1S7	D7S21	D12S11	D2S44
Danish[b]	465	460	471	463
English Black[c]	...	...	...	224
English[c]	...	...	...	271
English Asian Indian[c]	...	...	...	237
English[d]	1598	1613	1603	883
English Black[d]	824	873	865	475
English Asian Indian[d]	106	105	106	66
Finnish[e]	...	...	...	115
Alsatian[f]	...	...	100	87
German[g]	652	648	650	194
German[h]	...	328	375	512
Italian[i]	...	...	...	334
Maoris[j]	...	...	77	75
New Zealander[j]	...	...	105	97
Norwegian[k]	...	166	166	166
Pacific Islanders[j]	...	...	132	121
Spanish[l]	...	...	...	113
Spanish[m]	...	176	198	191
Swedish[n]	280	281	280	279
Swiss[o]	...	460	479	484
Turkish[b]	...	103	101	102
United States Black[p]	313	289	310	306
United States Caucasian[p]	319	318	317	311

[a] The numbers in the locus columns represent the number of individuals typed.

[b] Databases provided by B. Eriksen, H. Hansen, N. Morling, and O. Svensmark, Institute of Forensic Genetics, University of Copenhagen, Denmark.

[c] Databases provided by P. Gill, The Forensic Science Service, Central Research and Support Establishment, Aldermaston, England.

[d] Databases provided by C. Buffery, F. Burridge, M. Greenhalgh, S. Jones, and G. Willot, Metropolitan Police Forensic Science Laboratory, London, England.

[e] Database provided by A. Sajantila, Department of Human Genetics, National Public Health Institute, Helsinki, Finland.

[f] Database provided by B. Ludes and H. Pfitzinger, Institut de Medicine Legale, Strasbourg, France.

[g] Database provided by J. Henke, Institut fur Blutgruppenforschung, Koln, Germany and L. Henke, Institut fur Blutgruppenforschung, Dusseldorf, Germany.

[h] Database provided by B. Brinkmann, Institut fur Rechtsmedizin, Westfalische Wilhelms-Universitat Munster, Germany.

[i] Database provided by V. Pascali, Immunohematology Laboratory, Department of Forensic Medicine, Universita Cattolica del Sacro Cuore, Rome, Italy.

[j] Databases provided by S. Cordiner, Institute of Environmental Health and Forensic Sciences, Wellington; F. Hamilton and J. Chambers, Victoria University, Wellington; and P. Stapleton, DNA Diagnostics, Auckland, New Zealand.

[k] Database provided by B. Olaisen, Institute of Forensic Medicine, University of Oslo, Norway.

[l] Database provided by C. Cabrero, Pharma Gen. S. A., Madrid, Spain.

[m] Database provided by A. Carracedo, Institute of Legal Medicine, University of Santiago de Compostela, Spain.

[n] Database provided by S. Holgersson, National Laboratory of Forensic Science (SKL), Linkoping, Sweden.

[o] Database provided by W. Bar, Institute of Legal Medicine, University of Zurich-Irchel, Zurich, Switzerland.

[p] Databases provided by L. Forman, Cellmark Diagnostics, Germantown, Maryland, USA.

assigning statistical estimates. In the preceding companion paper, Budowle et al. [5], using variable number of tandem repeat (VNTR) loci data, generated by Hae III digestion and restriction fragment length polymorphism (RFLP) analysis, demonstrated that these concerns were unfounded. Using worst-case scenario comparisons, they found that there were very

TABLE 2—*The number of loci carried by each target profile per population group after adjusting for operational constraints for estimating DNA profile frequencies.*[a]

615 Hinf I Target Profiles				
Number of Loci	Blacks	Caucasians	Asian Indians	Total
0	0	0	0	0
1	1	1	0	2
2	19	32	7	58
3	115	135	32	282
4	115	131	27	273

[a] Operational constraints are described in reference [5].

few forensically significant differences in the estimates of multiple locus VNTR profile frequencies in various subgroups within a major population category. The greatest range of frequencies in statistical estimates was observed between major population groups, not between their constituent subgroups. Differences in statistical estimates were deemed forensically significant when the likelihood of occurrence of the DNA profile would be meaningfully different [5–7]. To appreciate the effects on the differences of target profile estimates the reader should also refer to volume IV of VNTR Population Data: A Worldwide Study [8].

In the companion paper, Budowle et al. [5] investigated as many as nine regional U.S. Caucasian databases, but for non-U.S. Caucasian databases, four Canadian, a French, an Israeli, and a Swiss database were available with Hae III-generated RFLP data. This paper presents a comparison analysis of frequency estimates of Hinf I-generated VNTR profiles determined using various reference databases, with special attention to European Caucasian data.

Materials and Methods

RFLP population data for several VNTR loci were kindly provided by the contributors listed in Table 1. The data consisted of fragment lengths generated by digestion of genomic DNA with the restriction endonuclease Hinf I.

When using various Hinf I population reference databases, target profiles of 615 individuals from the Metropolitan Police Laboratory (London, England) Caucasian (N = 299), Asian Indian (N = 66), and Black (N = 250) databases were used. The likelihood of occurrence of each profile, using the loci D1S7, D2S44, D7S21, and D12S11, or a subset of these loci, was calculated in the databases contained within this study. The analyses were performed as described in the preceeding companion paper [5].

Averaged rebinned data of an English Caucasian database (from the Metropolitan Police Laboratory) were compared with the rebinned Spanish, Turkish, New Zealand, and Alsatian databases by scatter plot analysis.

Table 2 provides a breakdown of the target profiles and the number of loci per population group (after adjusting for operational constraints [5]) that were used for estimating DNA profile frequencies.

Finally, due to the large volume of data, only representative examples of all the binned data and/or cross-group scatter plot comparisons that were performed are provided in this report. More data are compiled in a separate compendium [8].

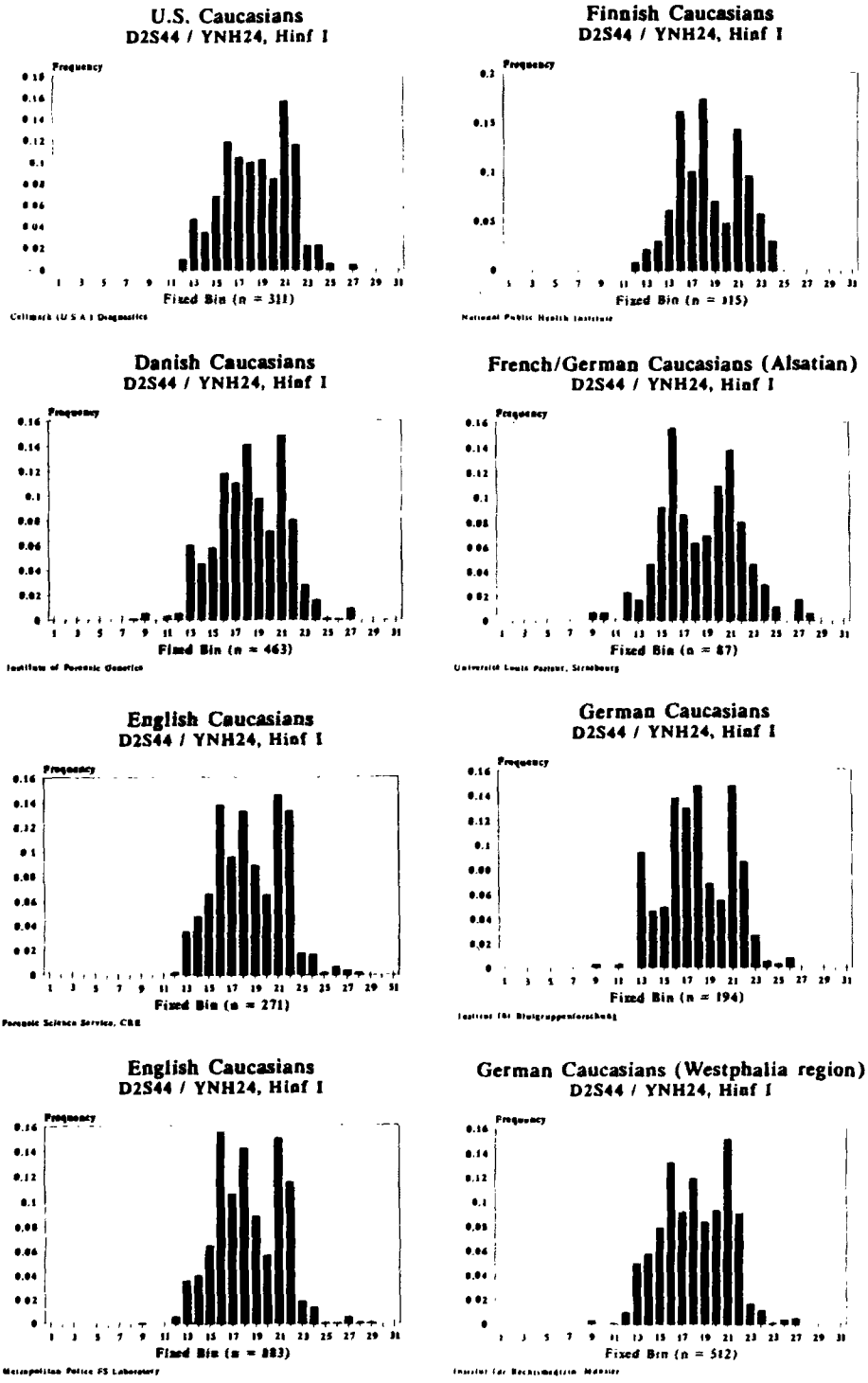
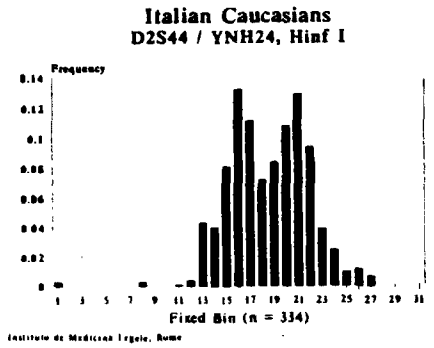
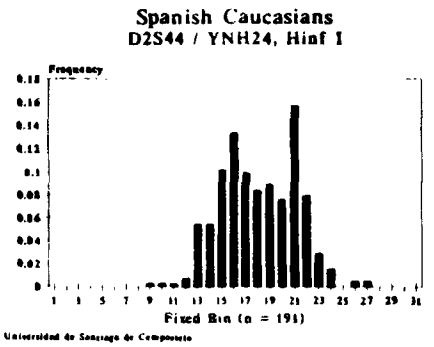
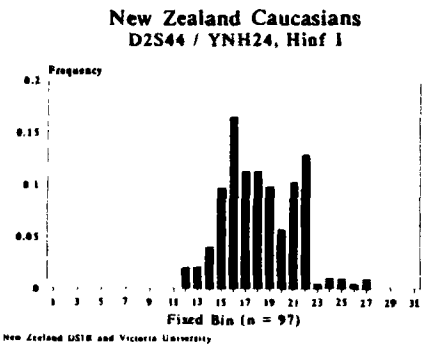
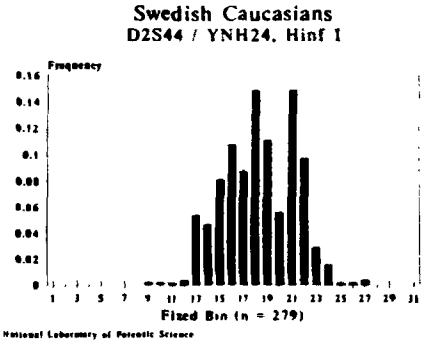


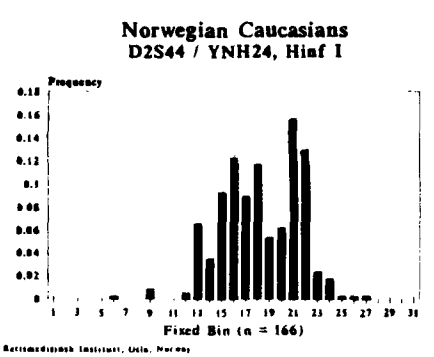
FIG. 1—*Histograms of 31 bin sorted frequency data for D2S44 from various reference population. The x axis defines the bin and the y axis defines the frequency of each bin. It should be noted that the frequency scale, or y axis, may vary among the histograms; this should be considered when evaluating the data.*

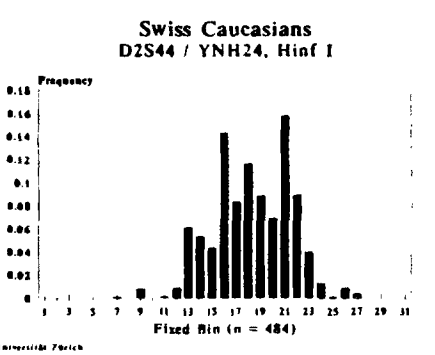


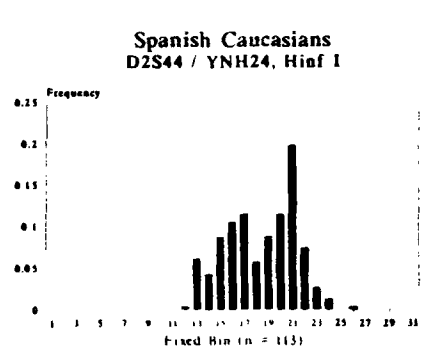












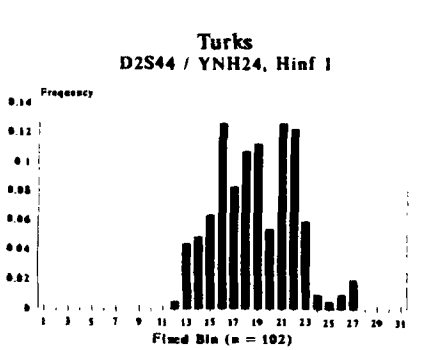


FIG. 1—*Continued.*

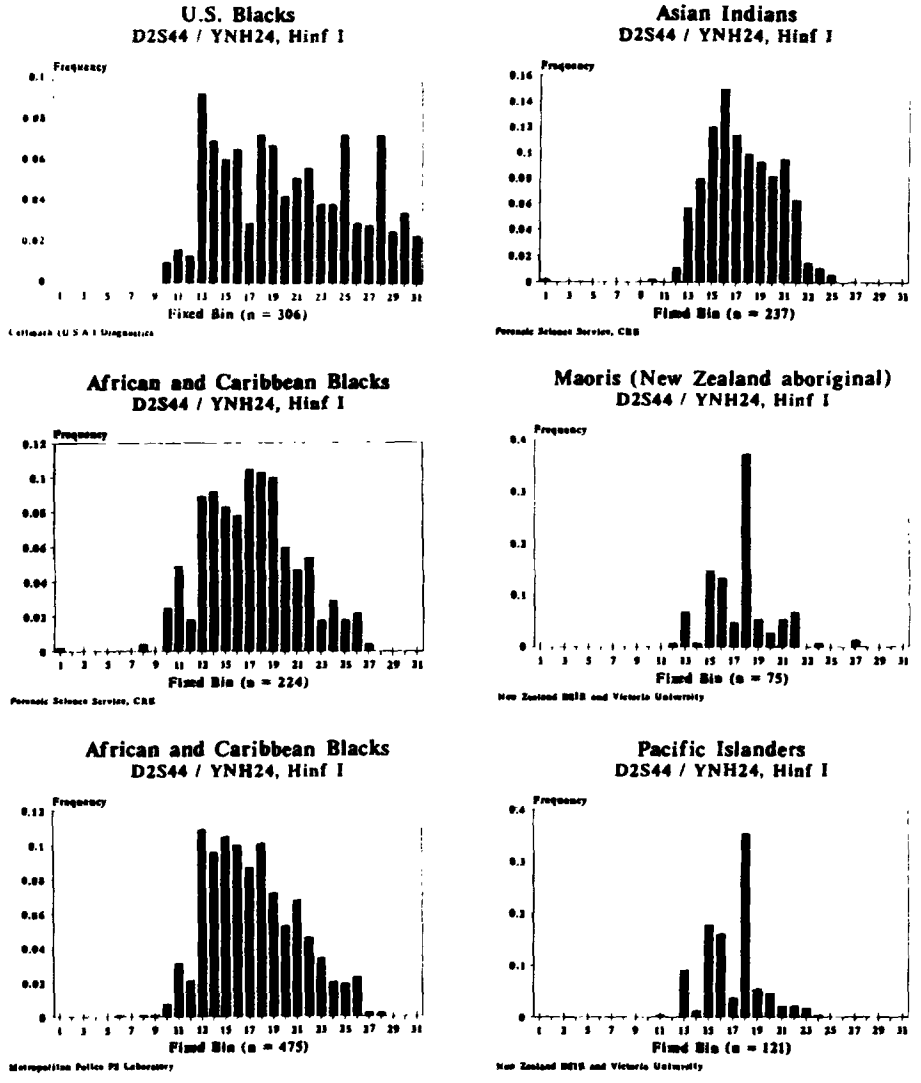


FIG. 1—*Continued.*

Results and Discussion

In accordance with Hae III-generated data [5], the VNTR loci studied were highly polymorphic in all databases described in Table 1. As an example of a locus that is polymorphic in all databases studied, Fig. 1 displays frequency histograms sorted into 31 bins for the locus D2S44 (Hinf I-based data) in various reference populations. Figure 1 shows that there are differences in binned allele frequencies among the various databases. Such differences are exemplified in Fig. 1 by bins 29–31 for D2S44 in US Caucasians, US Blacks, and Afro-Caribbeans. The bin frequencies are zero for these population samples except for US Blacks, where bins 29–31 have frequencies of 2.5%, 3.4%, and 2.3%, respectively. The explanation for this observation could be a founder effect, sampling

TABLE 3—*Distribution (percentages) of ratios of frequency estimates in various pairs of within-group reference populations using the loci D1S7, D2S44, D7S21, and D12S11 (Hinf I-based data).*

Reference Population Comparison	Bin Fmt[b]	Ratio Interval[a]					
		1	>1–2	>2–5	>5–10	>10–100	>100
English[c] v. Danish[d]	R	0.3	78.7	19.3	1.5	0.2	0
	31	0.3	82.6	15.9	1.1	0	0
English[c] v. German[e]	R	0	74.0	20.0	2.3	3.6	0.2
	31	0.3	73.7	22.1	2.9	1.0	0
English[c] v. Swedish[f]	R	0.5	56.3	23.9	8.5	10.6	0.3
	31	0.3	64.4	28.6	4.9	1.8	0
Danish[d] v. German[e]	R	0.3	75.0	20.0	2.1	2.6	0
	31	0.3	75.4	20.8	2.3	1.1	0
Danish[d] v. Swedish[f]	R	0	58.7	25.4	9.3	6.7	0
	31	0.3	70.7	26.8	2.0	0.2	0
German[e] v. Swedish[f]	R	0.3	58.7	25.4	8.3	7.2	0.2
	31	0.5	67.2	27.5	2.8	2.1	0
US Caucasian[g] v. English[c]	R	0.5	69.3	21.5	6.3	2.4	0
	31	0.5	72.5	23.4	2.8	0.8	0
US Caucasian[g] v. Danish[d]	R	0.3	72.2	24.2	2.9	0.3	0
	31	0.3	70.4	23.1	4.9	1.3	0
US Caucasian[g] v. German[e]	R	0.2	66.2	27.3	4.4	2.0	0
	31	0	67.3	28.9	3.4	0.3	0
US Caucasian[g] v. Swedish[f]	R	0.2	67.8	26.5	4.4	1.1	0
	31	0.2	73.7	17.7	4.2	4.2	0
US Black[g] v. English Black[c]	R	0.2	71.7	26.8	1.3	0	0
	31	0.2	65.2	29.4	4.2	1.0	0

[a] Ratios were determined by dividing the more common frequency by the less common frequency in the designated reference databases for each of the 615 target profiles. Each ratio interval represents the magnitude of the ratio for each target profile. The percentages are the portion of target profiles falling within each ratio interval.

[b] Bin Formats (Fmt) are: R = rebinned data; 31 = 31 bin data.

[c] Databases provided by Metropolitan Police Forensic Science Laboratory, London, England.

[d] Database provided by the Institute of Forensic Genetics, Copenhagen, Denmark.

[e] Database provided by the Institut fur Blutgruppenforschung in Koln and Dusseldorf, Germany.

[f] Database provided by National Laboratory of Forensic Science (SKL), Linkoping, Sweden.

[g] Databases provided by Cellmark Diagnostics, Germantown, Maryland, USA.

variance, and/or measurement biases [*9–12*]. There have been suggestions made that such allele frequency differences can lead to great variation (of at least two orders of magnitude) in the estimate of a multiple locus DNA profile when different ethnic databases are used [*3*]. Although it has been shown that the differences among allele frequencies in different population samples are diminished when the alleles from the entire set of loci comprising a DNA profile are used [*5,7,13–15*], it would be desirable to consider whether substantial differences occur for DNA profile frequency estimates between subgroups within a major population group. Therefore, the likelihood of occurrence of DNA profiles in various subgroup reference populations were evaluated for differences in DNA statistical estimates. In accordance with previous studies [*6,7,13,15*] the data strongly support that multiple locus VNTR DNA profiles are rare events in any relevant database.

TABLE 4—*Distribution (percentages) of ratios of frequency estimates in various pairs of Caucasian reference populations using the loci D2S44, D7S21, and D12S11 (Hin f I-based data).*

Reference Population Comparison	Bin Fmt[b]	Ratio Interval[a]					
		1	>1–2	>2–5	>5–10	>10–100	>100
English[c] v. German[d]	R	0.3	73.0	22.1	3.7	0.8	0
	31	0.3	75.3	20.5	3.6	0.3	0
English[c] v. Norwegian[e]	R	0.2	57.4	19.5	9.3	12.5	1.1
	31	0.3	67.5	24.9	5.0	2.3	0
English[c] v. Swiss[f]	R	0.5	76.4	18.0	3.3	1.8	0
	31	0.5	77.7	19.7	1.3	0.8	0
English[c] v. Spanish[g]	R	0.2	59.7	22.6	8.0	9.3	0.3
	31	0.3	62.1	28.1	7.5	2.0	0
	RS/R	0.2	73.0	20.2	5.5	1.1	0
English[c] v. Turkish[h]	R	0.5	59.7	18.2	6.5	13.3	1.8
	31	0.3	59.8	29.3	7.3	3.1	0.2
	RS/R	0.3	84.7	14.8	0.2	0	0
German[d] v. Norwegian[e]	R	0	62.6	19.8	9.3	8.1	0.2
	31	0	78.4	19.8	1.6	0.2	0
German[d] v. Swiss[f]	R	0.7	87.0	12.4	0	0	0
	31	0.5	83.1	15.4	1.0	0	0
German[d] v. Spanish[g]	R	0.5	71.4	21.6	3.3	3.3	0
	31	0.3	78.4	16.6	3.6	1.1	0
German[d] v. Turkish[h]	R	0	56.1	26.2	7.2	9.8	0.8
	31	0	60.0	32.7	5.5	1.8	0
Norwegian[e] v. Swiss[f]	R	0.5	60.0	22.1	8.9	8.3	0.2
	31	0.3	71.4	23.1	4.2	1.0	0
Norwegian[e] v. Spanish[g]	R	0.3	76.6	17.6	4.4	1.1	0
	31	0.2	75.0	19.2	3.9	1.8	0
Norwegian[e] v. Turkish[h]	R	0.5	72.5	25.7	1.1	0.2	0
	31	0.2	58.5	34.3	5.5	1.5	0
Norwegian[e] v. Swedish[i]	R	0.5	80.2	16.1	2.1	1.1	0
	31	0	70.7	26.0	2.8	0.5	0
Spanish[g] v. Swiss[f]	R	0.3	69.4	21.6	6.0	2.6	0
	31	0.3	69.9	23.4	4.6	1.8	0
Spanish[g] v. Turkish[h]	R	0.2	67.2	25.7	5.0	2.0	0
	31	0	65.7	28.9	3.4	2.0	0
Swiss[f] v. Turkish[h]	R	0.5	55.6	27.0	7.5	9.1	0.3
	31	0.2	55.6	33.5	7.5	3.3	0
Swedish[i] v. Turkish[h]	R	0.2	74.5	19.8	3.7	1.8	0
	31	0.3	60.8	29.6	7.2	2.0	0.2
Swiss[f] v. Swedish[i]	R	0.2	67.2	20.5	7.2	5.0	0
	31	0.2	76.9	19.7	2.1	1.1	0
US Caucasian[j] v. English[c]	R	0	77.6	16.7	4.7	1.0	0
	31	0	79.7	18.9	1.0	0.5	0
US Caucasian[j] v. German[d]	R	0.2	78.7	19.7	1.5	0	0
	31	0.2	79.0	17.7	2.3	0.8	0

TABLE 4—*Continued.*

Reference Population Comparison	Bin Fmt[b]	Ratio Interval[a]					
		1	>1–2	>2–5	>5–10	>10–100	>100
US Caucasian[j] v. Norwegian[e]	R	0.2	55.9	29.6	10.4	3.9	0
	31	0.2	68.9	26.5	2.6	1.8	0
US Caucasian[j] v. Swiss[f]	R	0.2	76.7	22.3	0.8	0	0
	31	0.3	69.8	26.3	3.3	0.3	0
US Caucasian[j] v. Spanish[g]	R	0.5	66.2	27.5	5.0	0.8	0
	31	0.8	74.1	22.9	2.1	0	0
US Caucasian[j] v. Turkish[h]	R	0	60.8	25.4	8.3	5.4	0
	31	0.3	69.3	26.5	3.3	0.7	0
US Caucasian[j] v. Swedish[i]	R	0.3	67.5	25.7	5.4	1.1	0
	31	0.3	75.4	16.4	4.2	3.6	0

[a]Ratios were determined by dividing the more common frequency by the less common frequency in the designated reference databases for each of the 615 target profiles. Each ratio interval represents the magnitude of the ratio for each target profile. The percentages are the portion of target profiles falling within each ratio interval.

[b]Bin Formats (Fmt) are: R = rebinned data; 31 = 31 bin data; and RS/R = rebinned random sampling data of the larger-sized reference population compared with rebinned data of the smaller-sized reference population.

[c]Database provided by Metropolitan Police Forensic Science Laboratory, London, England.

[d]Database provided by Institut fur Rechtsmedizin, Munster, Germany.

[e]Database provided by Institute of Forensic Medicine, Oslo, Norway.

[f]Database provided by Institute of Legal Medicine, Zurich, Switzerland.

[g]Database provided by Institute of Legal Medicine, Santiago, Spain.

[h]Database provided from Institute of Forensic Genetics, Copenhagen, Denmark.

[i]Database provided by National Laboratory of Forensic Science (SKL), Linkoping, Sweden.

[j]Database provided by Cellmark Diagnostics, Germantown, Maryland, USA.

The observations for Hinf I-generated data are similar to those for Hae III-generated data. Therefore, only a representative portion of the data are presented and will not be explained as explicitly as the previous Hae III-based data [*5*]. The reader again should refer to VNTR Population Data: A Worldwide Study [*8*] and Budowle et al. [*5*] for additional data.

Tables 3 to 7 show the distribution of the ratios of frequency estimates calculated in various pairs of reference populations using various combinations of VNTR loci. As was demonstrated for Hae III-generated VNTR data, there are very few differences in within-population group (that is, subgroups of a major population category) frequency estimates. Ratios greater than one order of magnitude, with frequencies more common than 1/1,000,000 (or for that matter 1/100,000), were unlikely occurrences.

It was shown for the Hae III-generated data that the frequencies of occurrence of non-Caucasian target profiles being estimated using Caucasian databases contribute most of the ratios that exceed one order of magnitude [*5*]. The same trend holds for Hinf I-generated VNTR data. Tables 8 to 10 show the number of Caucasian target profiles composed of four, three, or two loci, respectively, that were more common than 1/1,000,000 and had a ratio greater than one order of magnitude. The frequencies of occurrence of Caucasian target profiles do not vary significantly when different Caucasian databases are employed. The same trends hold for US and English Black sample populations. There were no Black target profiles that were more common than 1/1,000,000 and differed by more than one order of magnitude when using rebinned Black population data (Table 3). With 31 bin sorted Black data, there was only one target profile that was more common than 1/1,000,000

TABLE 5—*Distribution (percentages) of ratios of frequency estimates in various reference within group population comparisons. Based on number of loci in each target profile (Hinf I-based data).*[a]

Reference Populations	Bin Fmt[c]	No. of Loci[d]	Ratio Interval[b]					
			1	>1–2	>2–5	>5–10	>10–100	>100
US Caucasian v. Norwegian	R	1	0	100	0	0	0	0
	31	1	0	100	0	0	0	0
	R	2	0.4	71.6	21.5	5.0	1.5	0
	31	2	0.4	81.6	17.2	0.8	0	0
	R	3	0	43.4	36.2	14.7	5.8	0
	31	3	0	58.9	33.9	4.0	3.2	0
US Caucasian v. Spanish	R	1	0	100	0	0	0	0
	31	1	0	100	0	0	0	0
	R	2	0.8	71.6	23.8	2.7	1.2	0
	31	2	1.5	83.1	14.6	0.8	0	0
	R	3	0.3	61.5	30.7	6.9	0.6	0
	31	3	0.3	67.0	29.6	3.2	0	0
US Caucasian v. Turkish	R	1	0	100	0	0	0	0
	31	1	0	100	0	0	0	0
	R	2	0	67.4	22.2	6.9	3.5	0
	31	2	0.4	73.6	21.8	3.8	0.4	0
	R	3	0	55.2	28.2	9.5	6.9	0.3
	31	3	0.3	65.5	30.5	2.9	0.9	0

[a] The reference population comparisons were selected from Table 11. Because of space limitations, only three examples of US v. European comparisons are displayed.

[b] Ratios were determined by dividing the largest frequency by the smallest frequency for each of the 615 target profiles observed across the designated set of reference databases. Each ratio interval represents the magnitude of the ratio for each target profile. The percentages in each interval are the portion of target profiles within each ratio interval.

[c] Bin Formats (Fmt) are: R = rebinned data; 31 = 31 bin data.

[d] The column has been broken down into the number of loci each target profile carries. There were no zero locus target profiles, 2 one locus target profiles, 58 two locus target profiles, and 282 three locus profiles.

(that is, 1/35,200) and had a ratio greater than one order of magnitude (that is, 54.8). Again, because deviations based on ratios will show a large variance due to sampling, the very few observed differences are extreme examples.

Because of space limitations, it was not possible to subdivide all the data into one, two, three, and four locus target profile frequency estimate comparisons. Table 5 displays three examples of such from Table 4. The databases were chosen to illustrate comparisons of a US Caucasian with different European databases.

To assess the differences across subgroups for target DNA profiles with more common frequencies (that is, generally two or one locus target profiles), the number of profiles with frequencies more common than 1/1,000 and ratios greater than two-fold were determined (Tables 11 to 13). As anticipated, target profiles with fewer loci had more observations that fit the criteria of being more common than 1/1,000 and showing greater than two-fold ratios than did target profiles with more loci. Even for estimates with more common frequencies, there appears to be little evidence for substantial differences between within-group comparisons. This is particularly so for US versus various European databases.

Also, as seen with Hae III-generated data [5], the variation in the breadth of the scatter plots is smaller for within-major population group comparisons (for example, Norwegian vs. Turkish) than for between major-group comparisons (for example, English Black vs. Asian Indian) (Fig. 2), (Table 7). The data support the premise that the range of estimates of the likelihood of occurrence based on general major population group reference databases

TABLE 6—*Distribution (percentages) of ratios of frequency estimates in various pairs of reference populations using the loci D2S44 and D12S11 (Hinf I-based data).*

Reference Population Comparison	Bin Fmt[b]	Ratio Interval[a]					
		1	>1–2	>2–5	>5–10	>10–100	>100
English[c] v. New Zealanders[d]	R	0	80.8	14.6	3.6	1.0	0
	31	0	78.5	17.4	3.1	1.0	0
	RS/R	0	88.8	10.9	0.3	0	0
English[c] v. Alsatian[e]	R	2.6	64.6	27.0	2.9	2.9	0
	31	2.6	65.2	26.3	3.4	2.4	0
	RS/R	0	75.3	23.3	0.8	0.7	0
English[c] v. Pacific Islanders[d]	R	0	26.2	34.8	20.5	16.9	1.6
	31	0	28.0	34.1	18.0	17.6	2.3
English[c] v. Maoris[d]	R	0	36.9	36.4	13.8	12.0	0.8
	31	0	38.4	34.1	17.6	9.8	0.2
Maoris[d] v. Pacific Islanders[d]	R	0	55.9	32.0	7.6	4.4	0
	31	0	61.5	30.7	4.7	3.1	0

[a] Ratios were determined by dividing the more common frequency by the less common frequency in the designated reference databases for each of the 615 target profiles. Each ratio interval represents the magnitude of the ratio for each target profile. The percentages are the portion of target profiles falling within each ratio interval.

[b] Bin Formats (Fmt) are: R = rebinned data; 31 = 31 bin data; and RS/R = rebinned random sampling data of the larger-sized reference population compared with rebinned data of the smaller-sized reference population.

[c] Database provided by the Metropolitan Police Forensic Science Laboratory, London, England.

[d] Database provided by S. Cordiner, Institute of Environmental Health and Forensic Sciences, Wellington; F. Hamilton and J. Chambers, Victoria University, Wellington; and P. Stapleton, DNA Diagnostics, Auckland, New Zealand.

[e] Database provided by the Institut de Medecine Legale, Strasbourg, France.

TABLE 7—*Distribution (percentages) of ratios of frequency estimates in various pairs of cross major-group reference populations using the loci D1S7, D2S44, D7S21, and D12S11 (Hinf I-based data).*

Reference Population Comparison	Bin Fmt[b]	Ratio Interval[a]					
		1	>1–2	>2–5	>5–10	>10–100	>100
English Blacks[c] v. English Caucasians[c]	R	0.2	17.9	24.4	15.9	35.9	5.7
	31	0.2	16.3	23.3	15.4	34.8	10.1
English Caucasians[c] v. English Indians[c]	R	0.2	35.8	28.0	10.4	19.5	6.2
	31	0	48.3	34.1	10.4	6.5	0.7
English Blacks[c] v. English Indians[c]	R	0	16.4	26.5	16.9	36.1	4.1
	31	0	18.2	26.5	21.3	30.4	3.6

[a] Ratios were determined by dividing the more common frequency by the less common frequency in the designated reference databases for each of the 615 target profiles. Each ratio interval represents the magnitude of the ratio for each target profile. The percentages are the portion of target profiles falling within each ratio interval.

[b] Bin Formats (Fmt) are: R = rebinned data; 31 = 31 bin data.

[c] Databases provided by Metropolitan Police Forensic Science Laboratory, London, England.

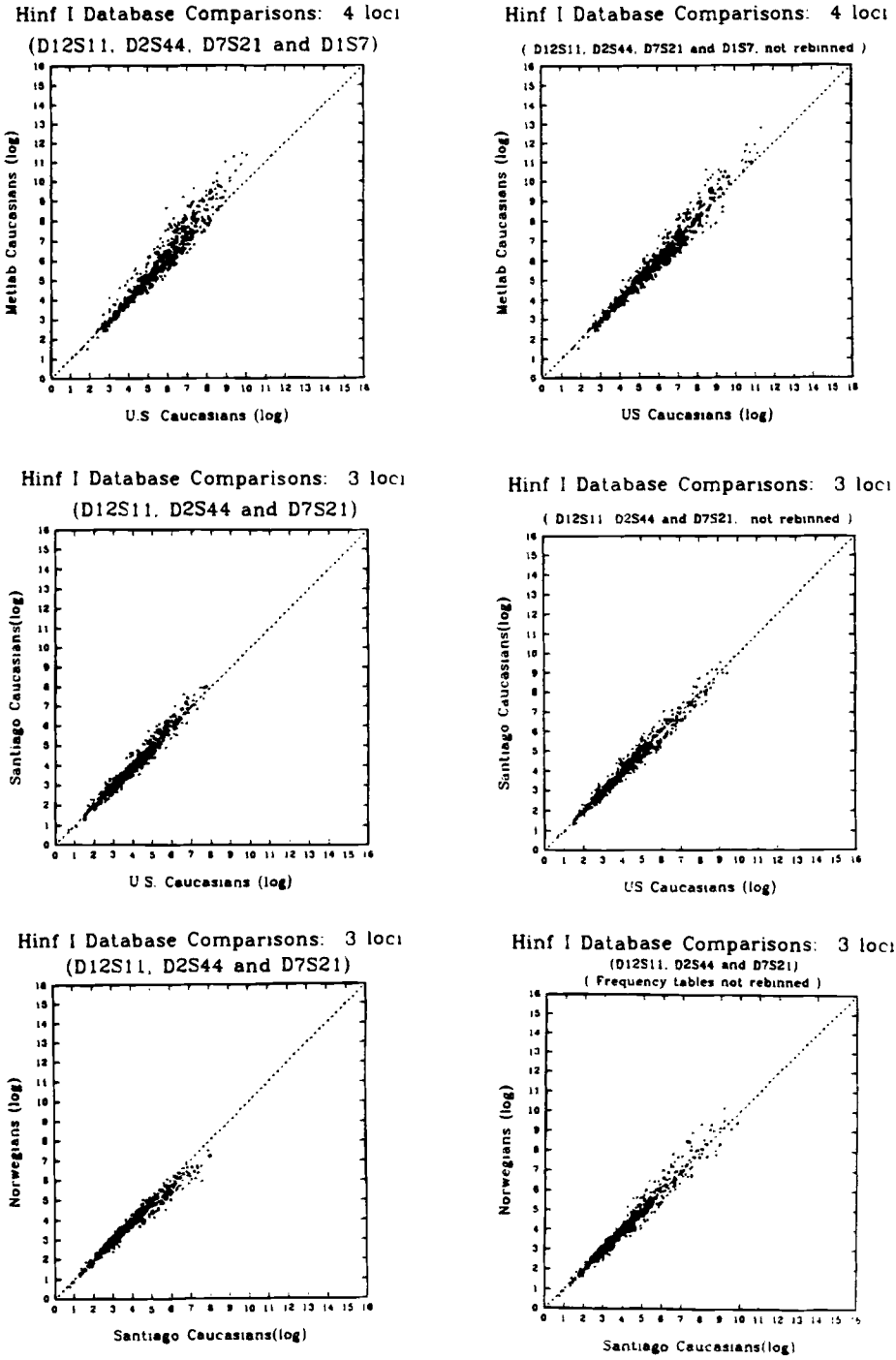


FIG. 2—*Examples of scatter plot comparisons of various reference populations databases. The likelihood of occurrence of 615 Hinf I-generated target RFLP profiles was estimated by the fixed bin method using various reference populations. The x and y axes of each scatter plot are labeled with each reference population used in a comparison. The first column of scatter plots displays comparisons using rebinned data and the second column of scatter plots displays comparisons using 31 bin data.*

TABLE 8—*Caucasian target DNA profiles (out of a total of* N = 299) *where: (1) frequency estimate ratios*[a] *are greater than ten-fold, and (2) at least one of the frequency estimates is more common than 1/1,000,000, in various pairs of within-group reference populations using the loci D1S7, D2S44, D7S21, and D12S11 (Hinf I-based data).*

Reference Population Comparison	Bin Fmt[a]	No. of Profiles	More Common Frequency	Ratio[b]
English[c] v. Danish[d]	R	0	NA	NA
	31	0	NA	NA
English[c] v. German[e]	R	1	1/82,600	13.0
	31	0	NA	NA
English[c] v. Swedish[f]	R	1	1/476,000	12.9
	31	0	NA	NA
Danish[d] v. German[e]	R	1	1/82,600	34.1
	31	0	NA	NA
Danish[d] v. Swedish[f]	R	0	NA	NA
	31	0	NA	NA
German[e] v. Swedish[f]	R	2	1/430,000	11.1
			1/82,600	19.0
	31	0	NA	NA
US[g] v. English[c]	R	0	NA	NA
	31	0	NA	NA
US[g] v. Danish[d]	R	0	NA	NA
	31	0	NA	NA
US[g] v. German[e]	R	1	1/82,600	11.9
	31	0	NA	NA
US[g] v. Swedish[f]	R	0	NA	NA
	31	0	NA	NA

[a]Bin Formats (Fmt) are: R = rebinned data; 31 = 31 bin data.

[b]Ratios were determined by dividing the more common frequency by the less common frequency in the designated reference databases for each of the 615 target profiles. Each ratio interval represents the magnitude of the ratio for each target profile. The percentages are the portion of target profiles falling within each ratio interval.

[c]Databases provided by Metropolitan Police Forensic Science Laboratory, London, England.

[d]Databases provided by the Institute of Forensic Genetics, Copenhagen, Denmark.

[e]Database provided by the Institut fur Blutgruppenforschung in Koln and Dusseldorf, Germany.

[f]Database provided by National Laboratory of Forensic Science (SKL), Linkoping, Sweden.

[g]Databases provided by Cellmark Diagnostics, Germantown, Maryland, USA.

will be greater than the range of estimates derived from various within-group reference databases (Tables 6 and 7).

The data support the premise that any of the databases within a major population category could serve as a reliable reference database for estimating the DNA profile frequencies (for example, English for Caucasians) and wrongful bias effects would not likely be encountered. Furthermore, the comparisons of various European databases with the US database tended to show on average smaller differences in frequency estimates than did various European estimates (Tables 3, 4, 8, and 9). Even though English versus Danish comparisons had the smallest differences, using a US database in lieu of either an English or Danish database would result in few differences in DNA profile estimates. The observation that US data comparisons generally show smaller differences than the various within-European comparisons should be expected. A greater gene flow across these groups would be expected in

TABLE 9—*Caucasian target DNA profiles (out of a total of* N = 299) *where: (1) frequency estimate ratios[a] are greater than ten-fold, and (2) at least one of the frequency estimates is more common than 1/1,000,000, in various pairs of within-group reference populations using the loci D2S44, D7S21, and D12S11 (Hinf I-based data).*

Reference Population Comparison	Bin Fmt[a]	No. of Profiles	More Common Frequency	Ratio[b]
English[c] v. German[d]	R	0	NA	NA
	31	0	NA	NA
English[c] v. Norwegian[e]	R	1	1/3,620	25.9
	31	1	1/11,100	11.4
English[c] v. Swiss[f]	R	0	NA	NA
	31	0	NA	NA
English[c] v. Spanish[g]	R	1	1/4,840	19.4
	31	0	NA	NA
	RS/R	0	NA	NA
English[c] v. Turkish[h]	R	1	1/7,160	13.1
	31	2	1/6,750	11.4
			1/325,000	13.2
	RS/R	0	NA	NA
German[d] v. Norwegian[e]	R	2	1/8,680	12.8
			1/124,000	15.9
	31	0	NA	NA
German[d] v. Swiss[f]	R	0	NA	NA
	31	0	NA	NA
German[d] v. Spanish[g]	R	0	NA	NA
	31	0	NA	NA
German[d] v. Turkish[h]	R	2	1/3,920	12.1
			1/89,900	21.9
	31	2	1/3920	12.1
			1/163,000	26.3
Norwegian[e] v. Swiss[f]	R	1	1/10,200	10.7
	31	0	NA	NA
Norwegian[e] v. Spanish[g]	R	0	NA	NA
	31	0	NA	NA
Norwegian[e] v. Turkish[h]	R	1	1/3,500	13.5
	31	5	1/231,000	18.6
			1/10,400	13.4
			1/3,500	13.6
			1/11,200	16.4
			1/7020	10.7
Norwegian[e] v. Swedish[i]	R	0	NA	NA
	31	0	NA	NA
Spanish[g] v. Swiss[f]	R	0	NA	NA
	31	0	NA	NA
Spanish[g] v. Turkish[h]	R	1	1/3,760	12.6
	31	3	1/320,000	13.4
			1/3,760	12.6
			1/7,330	10.2
Swiss[f] v. Turkish[h]	R	2	1/4,430	10.7
			1/4,570	10.5
	31	2	1/138,000	31.1
			1/4,430	10.7

TABLE 9—*Continued.*

Reference Population Comparison	Bin Fmt[a]	No. of Profiles	More Common Frequency	Ratio[b]
Swedish[i] v.	R	0	NA	NA
Turkish[h]	31	0	NA	NA
Swiss[f] v.	R	0	NA	NA
Swedish[i]	31	0	NA	NA
US[j] v.	R	0	NA	NA
English[c]	31	0	NA	NA
US[j] v.	R	0	NA	NA
German[d]	31	0	NA	NA
US[j] v.	R	1	1/12,100	12.5
Norwegian[e]	31	0	NA	NA
US[j] v.	R	0	NA	NA
Swiss[f]	31	0	NA	NA
US[j] v.	R	1	1/13,000	11.6
Spanish[g]	31	0	NA	NA
US[j] v.	R	0	NA	NA
Turkish[h]	31	0	NA	NA
US[j] v.	R	0	NA	NA
Swedish[i]	31	0	NA	NA

[a]Bin Formats (Fmt) are: R = rebinned data; 31 = 31 bin data; and RS/R = rebinned random sampling data of the larger-sized reference population compared with rebinned data of the smaller-sized reference population.

[b]Ratios were determined by dividing the more common frequency by the less common frequency in the designated reference databases for each of the 615 target profiles. Each ratio interval represents the magnitude of the ratio for each target profile. The percentages are the portion of target profiles falling within each ratio interval.

[c]Database provided by Metropolitan Police Forensic Science Laboratory, London, England.

[d]Database provided by Institut fur Rechtsmedizin, Munster, Germany.

[e]Database provided by Institute of Forensic Medicine, Oslo, Norway.

[f]Database provided by Institute of Legal Medicine, Zurich, Switzerland.

[g]Database provided by Institute of Legal Medicine, Santiago, Spain.

[h]Database provided from Institute of Forensic Genetics, Copenhagen, Denmark.

[i]Database provided by National Laboratory of Forensic Science (SKL), Linkoping, Sweden.

[j]Database provided by Cellmark Diagnostics, Germantown, Maryland, USA.

the US [*6,16,17*]. Therefore, the US Caucasian database likely will be a weighted average of the European databases. Since very few differences were observed for US versus European comparisons (and these differences will wane, even with a small amount of gene flow among the ethnic groups) generally no wrongful bias should be encountered in the US when estimating DNA profile frequencies using general US reference databases.

Recently, Krane et al. [*18*] reported that when using different Caucasian databases the portion of target DNA profiles with ratios of the likelihood of occurrence that exceeded one order of magnitude was greater than that observed in the study here. They observed that 22 to 34% of their target DNA profiles had ratios greater than one order of magnitude. Additionally, Krane et al. [*19*] observed that approximately 80% of their target profiles were estimated as less common when a Caucasian target profile was estimated using a different Caucasian ethnic database. This study and the companion paper by Budowle et al. [*5*] found little evidence to support the magnitude of the findings of Krane, et al. [*18*], other than when using a general Caucasian database the magnitude of the difference in

TABLE 10—*Caucasian target DNA profiles (out of a total of* N = 299*) where: (1) frequency estimate ratios*[a] *are greater than ten-fold, and (2) at least one of the frequency estimates is more common than 1/1,000,000, in various pairs of within-group reference populations using the loci D2S44 and D12S11 (Hinf I-based data).*

Reference Population Comparison	Bin Fmt[a]	No. of Profiles	More Common Frequency	Ratio[b]
English[c] v.	R	0	NA	NA
New Zealanders[d]	31	0	NA	NA
	RS/R	0	NA	NA
English[c] v.	R	1	1/1,050	10.2
Alsatian[e]	31	0	NA	NA
	RS/R	0	NA	NA

[a] Bin Formats (Fmt) are: R = rebinned data; 31 = 31 bin data; and RS/R = rebinned random sampling data of the larger-sized reference population compared with rebinned data of the smaller-sized reference population.

[b] Ratios were determined by dividing the more common frequency by the less common frequency in the designated reference databases for each of the 615 target profiles. Each ratio interval represents the magnitude of the ratio for each target profile. The percentages are the portion of target profiles falling within each ratio interval.

[c] Database provided by the Metropolitan Police Forensic Science Laboratory, London, England.

[d] Database provided by S. Cordiner, Institute of Environmental Health and Forensic Sciences, Wellington; F. Hamilton and J. Chambers, Victoria University, Wellington; and P. Stapleton, DNA Diagnostics, Auckland, New Zealand.

[e] Database provided by Institut de Medicine Legale, Strasbourg, France.

estimates was reduced. Possibly, Krane et al. [*18*] have found some genetic differences between Finns and Italians that should be considered as potentially impacting on DNA profile frequency estimates for forensic purposes. However, inadequate sampling would be a more plausible explanation for their observations [*19*]. Krane et al. [*18*] had only 51, 41, and 56 Finnish individuals in their D2S44, D10S28, and D16S85 locus databases, respectively, and 78, 73, and 75 Italians in their D2S44, D10S28, and D16S85 locus databases, respectively. Perhaps, larger databases for Finns and Italians and an alternate method for estimating DNA frequencies, such as described in the present paper and elsewhere [*5,7,19,20*], would reduce some of the effects of sampling variance. Then, most, if not all, of the differences detected by Krane et al. [*18*] probably would not be considered as affecting the rarity of the estimates.

Conclusions

The fixed bin method was used to assess the forensic significance of estimating the likelihood of occurrence of target DNA profiles in various reference databases. This is not to suggest that the only valid approach for statistical estimates is fixed binning; it was merely an approach for facilitating this study. Because comparisons using data sorted into 31 bins show very few differences in DNA profile frequency estimates, alternative methods, such as floating bin procedures, should lead to similar conclusions if alleles are defined according to the laboratory's empirically determined measurement error [*7*]. Based on the data described in this paper and the companion paper by Budowle et al. [*5*], differences in allele frequencies at a particular locus generally do not create substantial differences in frequency estimates of multiple locus VNTR profiles when different subgroup reference databases from within a major population group are compared. Using a Norwegian database in place of, for example, a Spanish database will not likely result in forensically significant differences in the estimates of DNA profile frequencies. The very few differences that were

TABLE 11—*Target DNA profiles (out of a total of* N = *615, consisting of 299 Caucasians, 250 Blacks and 66 Asian Indians) and the Caucasian subset of the total DNA target profiles* (N = *299) where: (1) frequency estimate ratios*[a] *are greater than two-fold, and (2) at least one of the frequency estimates is more common than 1/1,000, in various pairs of within-group reference populations (Hinf I-based data).*

Reference Population Comparison	Bin Fmt[a]	No. of Profiles	Ratio Range of Profiles[b]	No. of Caucasian Profiles	Ratio Range of Caucasian Profiles[b]
English[c] v. Danish[d]	R	1	2.0	0	NA
	31	1	2.0	0	NA
English[d] v. German[e]	R	3	2.5–2.7	1	2.5
	31	3	2.3–2.5	1	2.4
English[c] v. Swedish[f]	R	3	2.4–2.6	2	2.4–2.6
	31	3	2.4–2.6	1	2.4
Danish[d] v. German[e]	R	1	2.4	0	NA
	31	1	2.3	0	NA
Danish[d] v. Swedish[f]	R	0	NA	0	NA
	31	0	NA	0	NA
German[e] v. Swedish[f]	R	3	2.2–3.1	1	2.4
	31	3	2.2–2.9	1	2.4
US[g] v. English[c]	R	3	2.1–2.9	2	2.1–2.1
	31	3	2.1–2.9	2	2.1–2.1
US[g] v. Danish[d]	R	2	2.1–2.5	1	2.1
	31	2	2.1–2.5	1	2.1
US[g] v. German[e]	R	6	2.1–3.2	3	2.1–2.2
	31	4	2.1–3.2	2	2.1–2.2
US[g] v. Swedish[f]	R	4	2.1–3.7	3	2.1–2.8
	31	4	2.1–3.7	3	2.1–2.8

[a] Bin Formats (Fmt) are: R = rebinned data; 31 = 31 bin data.

[b] Ratios were determined by dividing the more common frequency by the less common frequency in the designated reference databases for each of the DNA target profiles. Each ratio interval represents the magnitude of the ratio for each target profile. The percentages are the portion of target profiles falling within each ratio interval.

[c] Databases provided by Metropolitan Police Forensic Science Laboratory, London, England.

[d] Database provided by the Institute of Forensic Genetics, Copenhagen, Denmark.

[e] Database provided by the Institut fur Blutgruppenforschung in Koln and Dusseldorf, Germany.

[f] Database provided by National Laboratory of Forensic Science (SKL), Linkoping, Sweden.

[g] Databases provided by Cellmark Diagnostics, Germantown, Maryland, USA.

observed will be diminished further when measurement error biases due to the analytical system are considered (that is, simply typing the sample in the same laboratory in which the database was generated will reduce the effects of measurement error and systematic bias); and in the case of US databases the result of ethnic admixture will reduce further the magnitude of differences between VNTR frequency profile estimates. Estimates of the likelihood of occurrence of a DNA profile using major population group databases (for example, Caucasian and Black) provide a greater range of frequencies than would estimates from subgroups of a major population category. Comparisons of major population groups provide valid estimates of DNA profile frequencies without generating a wrongful bias.

One could argue, that since the ethnic composition of the various databases in our study is not well-defined, the databases could roughly have the same ethnic composition; thus,

TABLE 12—*Target DNA profiles (out of a total of* N = *615, consisting of 299 Caucasians, 250 Blacks and 66 Asian Indians) and the Caucasian subset of the total DNA target profiles* (N = *299) where: (1) frequency estimate ratios*[a] *are greater than two-fold, and (2) at least one of the frequency estimates is more common than 1/1,000, in various pairs of within-group reference populations using the loci D2S44, D7S21, and D12S11 (Hinf I-based data).*

Reference Population Comparison	Bin Fmt[a]	No. of Profiles	Ratio Range of Profiles[b]	No. of Caucasian Profiles	Ratio Range of Caucasian Profiles[b]
English[c] v. German[d]	R	9	2.1–2.5	5	2.1–2.1
	31	9	2.1–2.5	5	2.1–2.1
English[c] v. Norwegian[e]	R	22	2.0–3.8	16	2.0–3.0
	31	23	2.0–3.3	19	2.0–3.1
English[c] v. Swiss[f]	R	9	2.2–2.7	7	2.2–2.7
	31	9	2.2–2.7	7	2.2–2.7
English[c] v. Spanish[g]	R	35	2.0–5.0	21	2.1–5.0
	31	33	2.0–5.0	21	2.0–5.0
	RS/R	29	2.0–4.5	19	2.0–4.5
English[c] v. Turkish[h]	R	15	2.1–17.2	8	2.1–3.0
	31	20	2.0–5.8	12	2.0–5.5
	RS/R	15	2.0–2.8	9	2.0–2.8
German[d] v. Norwegian[d]	R	13	2.0–6.7	9	2.0–6.7
	31	9	2.1–2.7	7	2.1–2.7
German[d] v. Swiss[f]	R	5	2.0–2.9	2	2.4–2.9
	31	5	2.0–2.7	2	2.4–2.7
German[d] v. Spanish[g]	R	4	2.3–6.3	2	2.3–6.3
	31	3	2.1–2.7	2	2.3–2.7
German[d] v. Turkish[h]	R	28	2.1–13.8	17	2.1–7.3
	31	31	2.1–5.6	20	2.1–4.4
Norwegian[e] v. Swiss[f]	R	20	2.0–7.2	11	2.0–2.8
	31	17	2.0–2.9	10	2.0–2.6
Norwegian[e] v. Spanish[g]	R	14	2.0–4.8	12	2.0–4.8
	31	14	2.0–4.4	12	2.0–4.4
Norwegian[e] v. Turkish[h]	R	44	2.0–4.6	38	2.0–4.6
	31	53	2.0–4.6	39	2.0–4.6
Norwegian[e] v. Swedish[i]	R	16	2.0–2.6	10	2.0–2.6
	31	16	2.0–3.9	9	2.1–2.6
Spanish[g] v. Swiss[f]	R	15	2.1–6.4	7	2.1–2.5
	31	12	2.2–4.1	7	2.2–2.6
Spanish[g] v. Turkish[h]	R	33	2.0–3.9	24	2.0–3.9
	31	32	2.0–4.9	21	2.0–3.6
Swiss[f] v. Turkish[h]	R	27	2.0–5.3	17	2.0–3.5
	31	31	2.0–5.9	20	2.0–4.6
Swedish[i] v. Turkish[h]	R	20	2.0–4.2	17	2.0–4.2
	31	28	2.0–6.2	18	2.0–4.7
Swiss[f] v. Swedish[i]	R	17	2.0–5.1	8	2.0–3.6
	31	15	2.0–3.6	8	2.0–3.6
US[j] v. English[c]	R	12	2.0–3.3	10	2.0–3.3
	31	12	2.0–3.3	10	2.0–3.3
US[j] v. German[d]	R	8	2.0–2.4	6	2.0–2.4
	31	8	2.0–2.4	6	2.0–2.4

TABLE 12—*Continued.*

Reference Population Comparison	Bin Fmt[a]	No. of Profiles	Ratio Range of Profiles[b]	No. of Caucasian Profiles	Ratio Range of Caucasian Profiles[b]
US[j] v. Norwegian[e]	R	28	2.0–4.9	22	2.0–3.5
	31	25	2.0–3.5	20	2.0–3.5
US[j] v. Swiss[f]	R	17	2.0–2.9	13	2.0–2.9
	31	18	2.0–3.3	13	2.0–2.9
US[j] v. Spanish[g]	R	23	2.0–4.3	14	2.0–3.6
	31	22	2.0–3.6	14	2.0–3.6
US[j] v. Turkish[h]	R	14	2.0–6.1	7	2.0–3.8
	31	11	2.0–2.9	5	2.0–2.6
US[j] v. Swedish[i]	R	19	2.0–3.6	11	2.1–3.6
	31	19	2.1–3.7	11	2.1–3.6

[a]Bin Formats (Fmt) are: R = rebinned data; 31 = 31 bin data; and RS/R = rebinned random sampling data of the larger-sized reference population compared with rebinned data of the smaller-sized reference population.

[b]Ratios were determined by dividing the more common frequency by the less common frequency in the designated reference databases for each of the 615 target profiles. Each ratio interval represents the magnitude of the ratio for each target profile. The percentages are the portion of target profiles falling within each ratio interval.

[c]Database provided by Metropolitan Police Forensic Science Laboratory, London, England.

[d]Database provided by Institut fur Rechtsmedizin, Munster, Germany.

[e]Database provided by Institute of Forensic Medicine, Oslo, Norway.

[f]Database provided by Institute of Legal Medicine, Zurich, Switzerland.

[g]Database provided by Institute of Legal Medicine, Santiago, Spain.

[h]Database provided from Institute of Forensic Genetics, Copenhagen, Denmark.

[i]Database provided by National Laboratory of Forensic Science (SKL), Linkoping, Sweden.

[j]Database provided by Cellmark Diagnostics, Germantown, Maryland, USA.

there is no way of determining whether or not the results in this study are meaningful. However, this argument would not be a likely scenario for the Hinf I-based data. Although a strict regimen for sample collection was not undertaken for the databases, the samples did derive from various countries and it would seem very unlikely that, for example, the Norwegian and Spanish databases would have the same ethnic composition. Since there are very few differences between comparisons of European databases, and since US populations of potential contributors of a forensic DNA sample are mixtures of ethnic groups, it would be anticipated that regionally derived US general population databases have roughly the same ethnic composition.

Acknowledgments

We would like to express our thanks and gratitude to those people who contributed their population data. Without their generosity this study could not have been undertaken. We also would like especially to thank Dr. Paul Debenham whose assistance facilitated the collection of the Hinf I population data. We also would like to thank Ranajit Chakraborty, Bernie Devlin, and Bruce Weir for their suggestions.

References

[1] Lander, E. S., "Population Genetic Considerations in the Forensic Use of DNA Typing," *Banbury Report 32: DNA Technology and Forensic Science*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1989, pp. 143–156.

TABLE 13—*Target DNA profiles (out of a total of* N = *615, consisting of 299 Caucasians, 250 Blacks and 66 Asian Indians) and the Caucasian subset of the total DNA target profiles* (N = *299) where: (1) frequency estimate ratios*[a] *are greater than two-fold, and (2) at least one of the frequency estimates is more common than 1/1,000, in various pairs of within-group reference populations using the loci D2S44 and D12S11 (Hinf I-based data).*

Reference population comparison	Bin Fmt[a]	No. of profiles	Ratio range of profiles[b]	No. of caucasian profiles	Ratio range of caucasian profiles[b]
English[c] v.	R	47	2.0–34.3	16	2.0–4.4
	31	37	2.0–10.7	13	2.0–4.8
New Zealanders[d]	RS/R	35	2.0–4.2	16	2.0–2.7
English[c] v.	R	110	2.0–55.1	55	2.0–4.2
	31	115	2.0–11.0	60	2.0–4.6
Alsatian[e]	RS/R	98	2.0–8.0	48	2.0–4.4

[a] Bin Formats (Fmt) are: R = rebinned data; 31 = 31 bin data; and RS/R = rebinned random sampling data of the larger-sized reference population compared with rebinned data of the smaller-sized reference population.

[b] Ratios were determined by dividing the more common frequency by the less common frequency in the designated reference databases for each of the 615 target profiles. Each ratio interval represents the magnitude of the ratio for each target profile. The percentages are the portion of target profiles falling within each ratio interval.

[c] Database provided by the Metropolitan Police Forensic Science Laboratory, London, England.

[d] Database provided by S. Cordiner, Institute of Environmental Health and Forensic Sciences, Wellington; F. Hamilton and J. Chambers, Victoria University, Wellington; and P. Stapleton, DNA Diagnostics, Auckland, New Zealand.

[e] Database provided by Institut de Medicine Legale, Strasbourg, France.

[2] Lander, E. S., "Invited Editorial: Research on DNA Catching Up With Courtroom Application," *American Journal of Human Genetics,* Vol. 48, 1991, pp. 819–823.

[3] Lewontin, R. C. and Hartl, D. L., "Population Genetics in Forensic DNA Typing," *Science,* Vol. 254, 1991, pp. 1745–1750.

[4] National Research Council, "DNA Typing: Statistical Bases for Interpretation," *DNA Technology in Forensic Science,* Chapter 3, Washington, D.C., National Academy Press, 1992, pp. 74–96.

[5] Budowle, B., Monson, K. L., Giusti, A. M., and Brown, B. L., "The Assessment of Frequency Estimates of Hae III-Generated VNTR Profiles in Various Reference Databases," *Journal of Forensic Sciences,* Vol. 39, No. 2, March 1994, pp. 319–352.

[6] Chakraborty, R. and Kidd, K. K., "The Utility of DNA Typing in Forensic Work," *Science,* Vol. 254, 1991, pp. 1735–1739.

[7] Monson, K. L. and Budowle, B., "VNTR Population Frequency Estimation for Forensics: Effect of Reference Population and Calculation Method," *Journal of Forensic Sciences,* Vol. 38, 1993, pp. 1037–1050.

[8] *VNTR Population Data: A Worldwide Study, Volumes I–IV,* Federal Bureau of Investigation, Washington, D.C., 1993.

[9] Devlin, B., Risch, N., and Roeder, K., "Statistical Comments on the NRC's Report on DNA Typing," *Journal of Forensic Sciences,* Vol. 39, 1994, pp. 28–40.

[10] Devlin, B., Risch, N., and Roeder, K., "Statistical Evaluation of DNA Fingerprinting: A Critique of the NRC's Report," *Science,* Vol. 259, 1993, pp. 748–750.

[11] Laber, T. L., O'Connor, J. M., Iverson, J. T., Liberty, J. A., and Bergman, D. L., "Evaluation of Four Deoxyribonucleic Acid (DNA) Extraction Protocols for DNA Yield and Variation in Restriction Fragment Length Polymorphism (RFLP) Sizes Under Varying Gel Conditions," *Journal of Forensic Sciences,* Vol. 37, 1992, pp. 404–424.

[12] Schneider, P. M., Fimmers, R., Woodroffe, S., Werrett, D. J., Bar, W., Brinkmann, B., Eriksen, B., Jones, S., Kloosterman, A. D., Mevag, B., Pascali, V. L., Rittner, C., Schmitter, H., Thomson, J. A., and Gill, P., "Report on a European Collaborative Exercise Comparing DNA Typing Results Using a Single Locus VNTR Probe," *Forensic Science International,* Vol. 49, 1991, pp. 1–15.

[13] Devlin, B., Risch, N., and Roeder, K., "Forensic Inference from DNA Fingerprints," *Journal of the American Statistics Association,* Vol. 87, 1992, pp. 337–350.

[14] Risch, N. and Devlin, B., "On the Probability of Matching DNA Fingerprints," *Science,* Vol. 255, 1992, pp. 717–720.
[15] Weir, B. S., "Independence of VNTR Alleles Defined by Fixed Bins," *Genetics,* Vol. 130, 1992, pp. 873–887.
[16] Kennedy, R. J. R., "Single or Triple Melting Pot? Intermarriage Trends in New Haven," *American Journal of Sociology,* Vol. 49, 1944, pp. 331–339.
[17] Spuhler, J. N. and Clark, P. J., "Migration into the Human Breeding Population of Ann Harbor, Michigan 1900–1950," *Human Biology,* Vol. 33, 1961, pp. 223–231.
[18] Krane, D. E., Allen, R. W., Sawyer, S. A., Petrov, D. A., and Hartl, D. L., "Genetic Differences at Four DNA Typing Loci in Finnish, Italian, and Mixed Caucasian Populations," *Proceedings of the National Academy of Sciences USA,* Vol. 89, 1992, pp. 10583–10587.
[19] Budowle, B., Monson, K. L., and Giusti, A. M., "A Reassessment of Frequency Estimates of Pvu II-Generated VNTR Profiles in a Finnish, an Italian, and a General United States Caucasian Database: No Evidence for Ethnic Subgroups Affecting Forensic Estimates," *American Journal of Human Genetics,* 1994, (in press).
[20] Budowle, B., Giusti, A. M., Waye, J. S., Baechtel, F. S., Fourney, R. M., Adams, D. E., Presley, L. A., Deadman, H. A., and Monson, K. L., "Fixed-Bin Analysis for Statistical Evaluation of Continuous Distributions of Allelic Data from VNTR Loci, for Use in Forensic Comparisons," *American Journal of Human Genetics,* Vol. 48, 1991, pp. 841–855.

Address requests for reprints or additional information to
Bruce Budowle, Ph.D.
FSRTC
FBI Academy
Quantico, VA 22135