TECHNICAL NOTE

Major Nu En Huang[1] *and Bruce Budowle, Ph.D.*[2]

Fixed Bin Population Data for the VNTR Loci D1S7, D2S44, D4S139, D5S110, and D17S79 in Chinese from Taiwan

REFERENCE: Huang, N. E. and Budowle, B., **"Fixed Bin Population Data for the VNTR Loci D1S7, D2S44, D4S139, D5S110, and D17S79 in Chinese from Taiwan,"** *Journal of Forensic Sciences,* JFSCA, Vol. 40, No. 2, March 1995, pp. 287–290.

ABSTRACT: Fixed bin frequencies for the VNTR loci D1S7, D2S44, D4S139, D5S110, and D17S79 were determined in a Chinese sample population. The data were generated by RFLP analysis of Hae III-digested genomic DNA and chemiluminescence detection. The five VNTR loci meet Hardy-Weinberg expectations in the Chinese sample population, and there is little evidence for association of alleles between the VNTR loci. The frequency data can be used in forensic analyses and paternity tests to estimate the frequency of a DNA profile in Chinese.

KEYWORDS: Chinese, Population Databases, VNTR, Hardy-Weinberg Expectations, Linkage Equilibrium, RFLP

The highly polymorphic variable number of tandem repeat (VNTR) loci [*1–4*] can be typed by restriction fragment length polymorphism (RFLP) analysis. This technology affords reliable typing results, with a high degree of discrimination, from DNA derived from biological evidence [*5*]. In order to estimate the rarity of a DNA profile some population data are required. To date there have been two population studies on VNTR data, comprising D1S7, D2S44, D4S139, and D10S28 loci, in Chinese [*6,7*]. However, neither of these Chinese VNTR population studies assessed their data for Hardy-Weinberg equilibrium (HWE) expectations and linkage disequilibrium (that is, associations between loci). Moreover, no population data are available for Chinese on two other VNTR loci that have gained acceptance in the forensic arena (that is, D5S110 and D17S79) [*8,9*]. This paper describes RFLP population data on the VNTR loci D1S7, D2S44, D4S139, D5S110, and D17S79 in Chinese from Taiwan using the restriction enzyme Hae III and chemiluminescence detection.

Received for publication 15 May 1994; accepted for publication 25 July 1995.

[1]Criminalist, Criminal Investigation Bureau DNA Laboratory, Taipei, Taiwan, Republic of China.

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

This is publication number 94-11 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

Materials and Methods

Sample Preparation

Whole blood was obtained in EDTA Vacutainer tubes by venipuncture from 126 unrelated Chinese individuals collected by the Criminal Investigation Bureau DNA Laboratory in Taipei, Taiwan. The DNA was extracted by the phenol-chloroform method according to Maniatis et al. [*10*]. The quantity of extracted DNA was estimated using the slot-blot procedure described by Waye et al. [*11*] and/or UV_{260} absorbance. Approximately 300–400 ng of DNA were used for RFLP analysis.

Typing

The DNA was typed by RFLP analysis according to the method of Budowle and Baechtel [*12*] with the following modifications. The Southern transfer solution was 10× SSC (87.7 g NaCl and 44.1 g sodium citrate/liter, pH 7.0) and the DNA was immobilized on the neutral nylon membrane Nylon-1 (Life Technologies, GibcoBRL, Gaithersburg, MD). After transfer, the DNA was fixed to the damp membrane by exposure to 20,000/cm^2 microjoules (total output) of UV light and subsequently the membrane was baked at 80°C for 30 minutes. Hybridization and detection were according to the protocol supplied with the ACES 2.0 Kit (Life Technologies, GibcoBRL, Gaithersburg, MD). The chemiluminescent substrate was lumiphos+ and was kindly provided by L. Klevan (Life Technologies, GibcoBRL, Gaithersburg, MD). The probes MS1 (for D1S7), LH1 (for D5S110) pH30 (for D4S139) were purchased from Life Technologies (GibcoBRL, Gaithersburg, MD), and YNH24 (for D2S44) and V1 (for D17S79) were purchased from the Promega Corporation (Madison, WI). These probes for forensically validated VNTR loci were the only ones available at the time of our study that were labeled with alkaline phosphatase and yielded DNA profiles with minimal background staining. Size measurements of RFLP bands were made using an interactive image analysis system [*13*].

Statistical Analysis

The base pair size data were binned according to the method of Budowle et al. [*14*]. Possible divergence from Hardy-Weinberg expectations (HWE) was determined by the likelihood ratio test [*15–17*] and the exact test [*18*]. An interclass correlation criterion

[19] for two-locus associations was used for detecting disequilibrium between the VNTR loci—D1S7, D2S44, D4S139, D5S110, and D17S79. Independence across the five VNTR loci also was determined by examining whether the observed variance of the number of heterozygous loci in the population sample is outside its confidence interval under the assumption of independence [20,21]. When appropriate, the Bonferroni procedure [22] was used to correct for multiple analyses to determine whether HWE or equilibrium between loci holds in a population.

Results and Discussion

The 31 fixed bin frequency distributions for the loci D1S7, D2S44, D4S139, D5S110, and D17S79 for Chinese are shown in Table 1. The loci are highly polymorphic, and there is no evidence for departure from HWE for any of the five VNTR loci based on the likelihood ratio test (15–17) and the exact test [18] (Table 1).

This is the first report on Oriental population data for the loci D17S79 and D5S110. As observed previously for Caucasians and African Americans [23], D17S79 in Chinese also is the least polymorphic of the forensically validated VNTR loci. In fact, only nine out of 31 bins contained any DNA fragments. Four of these bins had frequencies ranging from 0.127 to 0.270 (Table 1). Also similar to Caucasian and African American population data [23], the D5S110 locus in Chinese is highly polymorphic with most allele sizes ranging from 1000 to 10 000 base pairs (Table 1).

Analyses were performed to determine whether there were any detectable deviations from independence between D1S7, D2S44, D4S139, D5S110, or D17S79 loci. An interclass correlation test [19] demonstrated that there is little evidence for correlation between the alleles at any of the pairs of loci (Table 2). There was

TABLE 2—*Two locus inter-class correlation test for the VNTR loci in unrelated Chinese.*

Loci	Two-sided probability
D1S7/D2S44	0.533
D1S7/D4S139	0.511
D1S7/D5S110	0.023[a]
D1S7/D17S79	0.335
D2S44/D4S139	0.793
D2S44/D5S110	0.831
D2S44/D17S79	0.069
D4S139/D5S110	0.469
D4S139/D17S79	0.559
D5S110/D17S79	0.577

[a] = deviation at $P = 0.05$ level; with Bonferroni procedure level of rejection is $P = 0.005$.

TABLE 1—*Fixed bin frequencies for several VNTR loci in Chinese.*

BIN (bps)	D1S7[a] (n = 250)	D2S44[b] (n = 250)	D4S139[c] (n = 246)	D5S110[d] (n = 248)	D17S79[e] (n = 252)
0–639	.000	.000	.000	.000	.000
640–772	.000	.000	.000	.000	.000
773–871	.000	.016	.000	.000	.000
872–963	.000	.028	.000	.000	.000
964–1077	.000	.028	.000	.004	.000
1078–1196	.004	.020	.000	.004	.016
1197–1352	.004	.032	.000	.024	.270
1353–1507	.004	.140	.000	.012	.222
1508–1637	.012	.120	.000	.032	.202
1638–1788	.012	.204	.000	.040	.063
1789–1924	.016	.080	.000	.032	.060
1925–2088	.032	.100	.000	.056	.127
2089–2351	.040	.072	.000	.081	.032
2352–2522	.036	.044	.000	.060	.008
2523–2692	.032	.056	.000	.073	.000
2693–2862	.032	.028	.004	.044	.000
2863–3033	.032	.004	.016	.048	.000
3034–3329	.028	.008	.020	.097	.000
3330–3674	.052	.004	.037	.052	.000
3675–3979	.060	.008	.037	.056	.000
3980–4323	.076	.004	.069	.085	.000
4324–4821	.080	.004	.094	.036	.000
4822–5219	.052	.000	.098	.032	.000
5220–5685	.044	.000	.134	.040	.000
5686–6368	.056	.000	.122	.016	.000
6369–7241	.080	.000	.171	.004	.000
7242–8452	.064	.000	.098	.040	.000
8453–10093	.048	.000	.061	.020	.000
10094–11368	.048	.000	.016	.008	.000
11369–12829	.004	.000	.016	.000	.000
12830–	.052	.000	.008	.000	.000

[a]The number of individuals carrying a single band pattern = 3. HWE - Likelihood Ratio Test ($P = 0.704$) and Exact Test ($P = 0.627$).
[b]The number of individuals carrying a single band pattern = 5. HWE - Likelihood Ratio Test ($P = 0.217$) and Exact Test ($P = 0.258$).
[c]The number of individuals carrying a single band pattern = 7. HWE - Likelihood Ratio Test ($P = 0.159$) and Exact Test ($P = 0.166$).
[d]The number of individuals carrying a single band pattern = 5. HWE - Likelihood Ratio Test ($P = 0.653$) and Exact Test ($P = 0.584$).
[e]The number of individuals carrying a single band pattern = 21. HWE - Likelihood Ratio Test ($P = 0.118$) and Exact Test ($P = 0.200$).
[f]n equals the numbers of chromosomes.

one example of deviation between D1S7 and D4S139 ($P = 0.023$) out of a total of 10 interclass correlation tests. The amount of deviation observed is consistent with expectations. A Bonferroni procedure [22] was used as a correction when multiple tests are performed on a population sample. After correction, the data support that for the five VNTR loci the populations meet expectations of independence ($P = 0.005$ is the rejection level).

The observed variance (s_k^2) test [20,21] also was used to assess whether or not there is any evidence of substantial deviation from independence across the loci. There was no evidence of association for the five loci described in our Chinese sample population ($s_k^2 = 0.444$, 95% confidence interval of variance is 0.363 to 0.686).

There are population data in common for the loci D1S7, D2S44, and D4S139 among our Chinese database and two other Chinese databases [6,7]. It would not be meaningful to compare statistically these databases with a test for homogeneity, because the bin frequencies are subject to sampling variances and there are measurement biases among the laboratories. Budowle et al. [24,25] previously demonstrated that multiple VNTR loci are highly polymorphic in a variety of population groups and that subdivision, either by ethnic group or by U.S. geographic region, within a major population group did not substantially affect forensic estimates of the likelihood of occurrence of a DNA profile. However, to support the observation that estimates for multiple single-locus DNA profiles would not be substantially different among the three Chinese databases the bin frequencies for D1S7, D2S44, and D4S139 were compared in a similar manner to that described by Chow et al. [6] (Table 3). Our Taiwan Chinese 31-bin format frequency data were compared with data from California Chinese and Singapore Chinese [7]. Bins with a minimum of five events in at least one of the three Chinese databases for each of the VNTR loci were chosen. Based on a ratio of bin frequencies there are very few differences, most ratios were less than two-fold, between Taiwan and California or Taiwan and Singapore Chinese (Table 3). When the data are rebinned the few examples of ratios that exceeded two-fold were reduced (Table 3).

This is the first RFLP database generated by our laboratories using chemiluminescent detection. The profiles were typeable and consistent with those that would be expected with ^{32}P detection (data not shown, manuscript in preparation). Based on the statistical evaluation of the population data there is no more evidence of the presence of null alleles than would be expected with radioisotopic detection. We are currently evaluating the potential of replacing ^{32}P detection with chemiluminescence for RFLP typing of DNA from forensic biological evidence.

In conclusion, this report provides fixed bin frequencies for five VNTR loci in Chinese. The distribution of the genotype frequencies for all five VNTR loci meet HWE, and there is little evidence for association of alleles across the loci for our Chinese sample population. The data demonstrate that estimates of a multiple-locus profile frequency can be derived from our Chinese database for identity testing purposes using the product rule under the assumption of independence. In addition, based on this study and others [26,27], chemiluminescence can be used as a method of detection for RFLP analysis in lieu of ^{32}P for population database studies without loss of information.

This is publication number 94-11 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.

TABLE 3—*Number of bins[a] between Taiwan Chinese[b] and either California or Singapore Chinese[c] with particular frequency ratios (maximum frequency: minimum frequency).*

Ratio	D1S7[d]		D2S44[e]		D4S139[f]	
	Cal	Sing	Cal	Sing	Cal	Sing
$X \leq 2.0$	18	16	13	12	12	12
$2.0 < X \leq 3.0$	2	2	1	2	0	1
$3.0 < X \leq 4.0$	1	2	1	0	0	0
$X > 4.0$	0	1	1	2	1	0

[a]The bins for comparison were selected from 31-bin format data and had to contain a minimum of five events in at least one of the three Chinese population samples.

[b]Data from current study.

[c]California and Singapore bin data from *VNTR Population Data: A Worldwide Study* [7].

[d]After rebinning the data according to the method of Budowle et al. [14], none of the D1S7 comparisons exceeded a three-fold ratio.

[e]After rebinning the data, all D2S44 comparisons fell within the $X \leq 2.0$ category except for one which was less than three-fold.

[f]After rebinning the data, all D4S139 comparisons fell within the $X \leq 2.0$ category.

References

[1] Wyman, A. R. and White R., "A Highly Polymorphic Locus in Human DNA," *Proceedings of the National Academy of Sciences USA,* Vol. 77, 1980, pp. 6754–6758.

[2] Jeffreys, A. J., Wilson, V., and Thein, S. L., "Individual-Specific Fingerprints of Human DNA," *Nature,* Vol. 316, 1985, pp. 76–79.

[3] Jeffreys, A. J., Royle, N. J., Wilson, V., and Wong, Z., "Spontaneous Mutation Rates to New Length Alleles at Tandem-Repetitive Hypervariable Loci in Human DNA," *Nature,* Vol. 323, 1988, pp. 278–281.

[4] Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M., Kumlin, E., and White, R., "Variable Number of Tandem Repeat Markers for Human Gene Mapping," *Science,* Vol. 235, 1987, pp. 1616–1622.

[5] Adams, D. E., Presley, L. A., Baumstark, A. L., Hensley, K. W., Hill, A. L., Anoe, K. S., Campbell, P. A., McLaughlin, C. M., Budowle, B., Giusti, A. M., Smerick, J. B., and Baechtel, F. S., "DNA Analysis by Restriction Fragment Length Polymorphisms of Blood and Other Body Fluid Stains Subjected to Contamination and Environmental Insults," *Journal of Forensic Sciences,* Vol. 36, 1991, pp. 1284–1298.

[6] Chow, S. T., Tan, W. F., Yap, K. H., and Ng, T. L., "The Development of DNA Profiling Database in an Hae III Based RFLP System in Chinese, Malays, and Indians in Singapore," *Journal of Forensic Sciences,* Vol. 38, 1993, pp. 874–884.

[7] Hae III Bin Tables, In *VNTR Population Data: A Worldwide Study,* Volume I-B, Federal Bureau of Investigation, Quantico, VA, 1993, pp. 753–774.

[8] Balazs, I., Baird, M., Clyne, M., and Meade, E., "Human Population Genetic Studies of Five Hypervariable DNA Loci," *American Journal of Human Genetics,* Vol. 44, 1989, pp. 182–190.

[9] Amour, J. A., Povey, S., Jeremiah, S., and Jeffreys, A. J., "Systematic Cloning of Human Minisatellites from Ordered Charomid Libraries," *Genomics,* Vol. 8, 1990, pp. 501–512.

[10] Maniatis, T., Fritsch, E. F., and Sambrook, J., *Molecular Cloning. A Laboratory Manual,* Cold Spring Harbor Laboratory Publication, New York, 1982, pp. 458–459.

[11] Waye, J. S., Presley, L., Budowle, B., Shutler, G. G., and Fourney, R. M., "A Simple Method for Quantifying Human Genomic DNA in Forensic Specimen Extracts," *Biotechniques,* Vol. 7, 1989, pp. 852–855.

[12] Budowle, B. and Baechtel, F. S., "Modifications to Improve the Effectiveness of Restriction Fragment Length Polymorphism Typing," *Applied and Theoretical Electrophoresis,* Vol. 1, 1990, pp. 181–187.

[13] Monson, K. L. and Budowle, B., "A System for Semi-Automated Analysis of DNA Autoradiograms," In *Proceedings of the International Symposium on the Forensic Aspects of DNA Analysis,* Government Printing Office, Washington, D.C., 1989, pp. 127–132.

[14] 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.

[15] Edwards, A., Hammond, H., Jin, L., Caskey, C. T., and Chakraborty, R., "Genetic Variation at Five Trimeric and Tetrameric Repeat Loci in Four Human Population Groups," *Genomics,* Vol. 12, 1992, pp. 241–253.

[16] Chakraborty, R., Fornage, M., Guegue, R., and Boerwinkle, E., "Population Genetics of Hypervariable Loci: Analysis of PCR Based VNTR Polymorphism Within a Population," In *DNA Fingerprinting: Approaches and Applications,* T., Burke, G., Dolf, A. J., Jeffreys, and R. Wolff, Eds., Birkhauser Verlag, Berlin, 1991, pp 127–143.

[17] Weir, B. S., "Independence of VNTR Alleles Defined by Fixed Bins," *Genetics,* Vol. 130, 1992, pp. 873–887.

[18] Guo, S. W. and Thompson, E. A., "Performing the Exact Test of Hardy-Weinberg Proportion for Multiple Alleles," *Biometrics,* Vol. 48, 1992, pp. 361–372.

[19] Karlin, S., Cameron, E. C., and Williams, P. T., "Sibling and Parent-Offspring Correlation Estimation with Variable Family Size," *Proceedings of the National Academy of Sciences U.S.A.,* Vol. 78, 1981, pp. 2664–2668.

[20] Brown, A. H. D., Feldman, M. W., and Nevo, E., "Multilocus Structure of Natural Populations of Hordeum Spontaneum," *Genetics,* Vol. 96, 1980, pp. 523–536.

[21] Chakraborty, R., "Detection of Nonrandom Association of Alleles from the Distribution of the Number of Heterozygous Loci in a Sample," *Genetics,* Vol. 108, 1984, pp. 719–731.

[22] Weir, B. S., "Multiple Tests," In, *Genetic Data Analysis,* Sinauer Associates, Inc., Sunderland, Massachusetts, 1990, pp. 109–110.

[23] Budowle, B., Monson, K. L., Anoe, K., Baechtel, F. S., Bergman, D., et al., "A Preliminary Report on Binned General Population Data on Six VNTR Loci in Caucasians, Blacks, and Hispanics from the United States," *Crime Laboratory Digest,* Vol. 18, No. 1, 1991, pp. 9–26.

[24] Budowle, B., Monson, K. L., Giusti, A. M., and Brown, B., "The Assessment of Frequency Estimates of Hae III-Generated VNTR Profiles in Various Reference Databases," *Journal of Forensic Sciences,* Vol. 39, 1994, pp. 319–352.

[25] Budowle, B., Monson, K. L., Giusti, A. M., and Brown, B., "Evaluation of Hinf I-Generated VNTR Profile Frequencies Determined Using Various Ethnic Databases," *Journal of Forensic Sciences,* Vol. 39, 1994, pp. 988–1008.

[26] Giles, A. F., Booth, K. J., Parker, J. R., Garman, A. J., Carrick, D. T., Akhaven, H., and Schaap, A. P., "Rapid, Simple, Non-Isotopic Probing of Southern Blots for DNA Fingerprinting," In, *Advances in Forensic Haemogenetics 3,* H. F. Polesky and W. R., Mayr, Eds., Springer-Verlag, Berlin, 1989, pp. 40–42.

[27] Budowle, B., Klevan, L., and Eisenberg, A. J., "RFLP Typing: A New Highly Polymorphic VNTR Locus and Chemiluminescent Detection," In, *Advances in Forensic Haemogenetics 5,* W. Bar, A. Fiori, and U. Rossi, Eds., Springer-Verlag, Berlin, 1994, pp. 245–252.

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