## TECHNICAL NOTE

*Mei Yun Wu,[1] M.D.; Lang Pin Zou,[2] M.D.; Bin Shen,[3] M.D.; Guang Yun Sun,[1] Ph.D; Rong Hua Li,[1] Ph.D.; and Guo Di Chen,[1] M.D.*

# STR HUMARA Locus Gene and Genotype Frequencies in Han and Bei Populations in China*

**REFERENCE:** Wu MY, Zou LP, Shen B, Sun GY, Li RH, Chen GD. STR HUMARA locus gene and genotype frequencies in Han and Bei populations in China. J Forensic Sci 1999;44(5):1039–1041.

**ABSTRACT:** For the purpose of the population genetics study of the HUMARA locus, the allele, and genotype frequencies were determined in two Chinese population samples (Han-101, Bei-113) using PCR, PAGE, and silver staining. Fourteen alleles were found. The size of amplified fragments were 258 bp-315 bp. The observed heterozygosities were 0.83 in the Han population and 0.73 in the Bei population respectively. The expected heterozygosities were 0.91 in the Han population and 0.97 in the Bei population respectively. Both populations meet Hardy-Weinberg expectation, Han population $x^2 = 17.7206$, $df = 11$, $p > 0.05$; Bei population $x^2 = 7.4268$, $df = 10$, $p > 0.05$. The discrimination power were 0.95 in females and 0.89 in males in the Han population, 0.94 in females and 0.88 in males in the Bei population. Thus, the allelic frequency data can be used in the personal identification and parentage testing in the forensic science practice. The PCR test established in this study is robust and reproducible.

**KEYWORDS:** forensic science, DNA typing, short tandem repeats HUMARA, population genetics, Han and Bei populations, China

The gene and genotype frequencies for the STR HUMARA locus were determined in a Han population of 101 unrelated individuals living in Chengdu, Sichuan Province and a Bei population of 113 unrelated individuals living in Yunnan Province in PR China.

### Materials and Methods

*DNA Samples and Extraction*

DNA was extracted from EDTA blood samples collected from 101 and 113 unrelated blood donors of both Han and Bei populations in Sichuan and Yunnan Provinces respectively in China using a modification of the proteinase K (Sigma) digestion, phenol/chloroform extraction and ethanol precipitation procedure (1). The extracted DNA was dissolved in the TE buffer for PCR analysis.

*PCR Condition*

Primer sequences described by Hammond et al. (2) were used.

5′ tcc aga atc tgt tcc aga gcg tgc 3′

5′ get gtg aag gtt gct gtt cct cat 3′

20 ng of genomic DNA was amplified in a total volume of 25 μL consisting of 0.1 μm of each primer, 0.2 mM of each dNTP (Pharmacia), 0.5 μ of Taq polymerase (PE), 50 mM of KCI, 10 mM of Tris-HCI (pH 8.5), 1.5 mM of $MgCI_2$, and 0.01 mg/mL gelatin. The PCR was carried out in a Gene Amp PCR system (PC-Cetus Model 9600) for 30 cycles consisting of 95°C for 30 s, 60°C for 20 s, and 72°C for 30 s. After each cycle, an additional extension at 72°C for 30 s was performed. The amplified products were then separated on a vertical polyacrylamide gel (T = 8%, C = 5%) electrophoresis followed by silver staining (3). PBR 322/Msp1 was used as an external size standard. The internal size standard was prepared by mixing amplifying specific alleles from individuals of known genotypes. The smallest allele was arbitrarily designated 1 and others numbered consequently taking into consideration of the repeat size of 3 bp since the STR HUMARA locus is trimeric.

*Calculations*

The frequency of each allele was calculated from the members of each type in sample tested. Unbiased estimate of the expected heterozygosity was obtained from the method described by Chakraborty et al. (3). The Hardy-Weinberg test was carried out according to the method described by Hou Y et al. (4). Individualization potentials were calculated using a standard formula of Sensabaugh (5), which is expressed as the sum of the squares of all possible genotype frequencies in a polymorphic system. The DP value was calculated from the individualization potentials. DP = 1-individualization potentials.

[1] Professor, associate professor, associate professor, and associate professor, respectively. Department of Forensic Biology, West China University of Medical Sciences, The People's Republic of China.

[2] Lecturer. Court of Justice of Yunnan Province. The People's Republic of China.

[3] Professor. School of Forensic Medicine, Kuming Medical College, The People's Republic of China.

* This project was supported by the grant from the National Natural Fund (39370754).

Received 27 Aug. 1998; and in revised form 31 Dec 1998; accepted 5 Jan. 1999.

## Results and Discussion

All samples from these two population groups were typed successfully (photo). The STR HUMARA genes are X-linked. Two amplified product bands could be seen in a female's blood sample, whereas only one band could be found in male's blood. In case of the female, one band indicates the homozygote, whereas two bands indicate the heterozygote.

### *Allele Frequency Distribution*

The size of the HUMARA locus amplified products were 255 bp-315 bp as reported by Fregeau CJ et al. (6). In the present study, 14 alleles were found in both 101 unrelated and 113 unrelated individuals of the Han and the Bei populations respectively (Table 1) (Fig. 1). The allele frequency distribution exhibited unimodal pattern (Fig. 2).

### *Genotype Frequency Distribution*

Thirty-one different genotypes (14 in males and 20 in females) were detected in the Han population (Table 2) and 32 (14 in males and 24 in females) in the Bei population. In the Han population, the expected heterozygosity was 0.91 and the observed heterozygosity was 0.83 in females. The discriminating power (DP) was 0.94 in females and 0.88 in males. In the Bei population, the expected heterozygosity was 0.97 and the observed heterozygosity was 0.73 in females. The DP value was 0.95 in females and 0.89 in males. Because of the population sample size and the large number of different genotypes, any genotype class with less than seven observations was pooled for determining the Hardy-Weinberg equilibrium. Both populations meet H-W expectation.

The STR HUMARA [AGA] locus is located on the X chromosome within the protein coding region of the exon 1 of the human androgen receptor gene and within the Drosophia Notch gene. ARA is the abbreviation of the androgen receptor gene. It maps to XCenq13.

The results of the present study showed that this locus was polymorphic. It can be used for parentage testing and personal identification in the forensic science practice. Meanwhile, this locus is X-linked. So that, it can be used for sex determination of blood, hair, or tissues of human origin.

A population genetics study of STR HUMARA locus has been performed by Edwards et al. (7) by using the fluorescent multiplex PCR. Fourteen alleles were found in US Black, 10 in White, 13 in Hispanic, and 12 in Asians in a sample of 40 unrelated individuals. One year later, another paper was published by Edwards et al. (8). They had found 15 HUMARA alleles in US White (240). Eighteen in Black (250), 16 in Mexicans-Americans (205), and 14 in Asians (97) by PCR. The number in the parenthesis represents the number of individuals studied. Our results were the same with allele number of US black and Asians reported by Edwards et al. (7,8).

The allele frequency profile of these two populations were quite similar. But minor difference was found in that allele seven was clearly variable across these two populations. Strict further comparison of allele frequencies between different populations was not

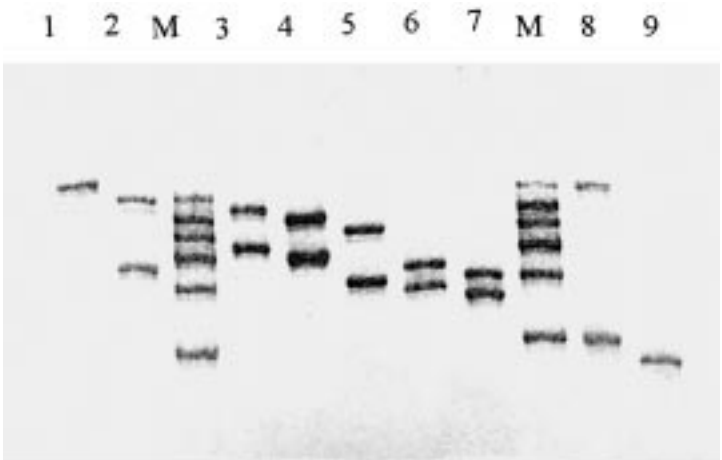


FIG. 1—*Photo. Genotypes of STR HUMARA locus. From 1 to 9: 15–15, 8–14, 10–13, 9–12, 7–11, 6–8, 5–7, 2–14, 1–1. M: Allelic ladder (alleles 2,7,9,11,12, and 14 from bottom to top).*

TABLE 1—*The allele frequencies of STR HUMARA locus in both Han and Bei populations.*

| | | Han Population | | | | Bei Population | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Number | | | | Number | | | |
| Allele (bp) | Size | 72X,Chr. (72 M) | 58X,Chr. (29 F) | Total | Pro. | 68X,Chr. (68 M) | 90X,Chr. (45 F) | Total | Pro. |
| 1 | 255 | | | | | | | | |
| 2 | 258 | 1 | | 1 | 0.008 | 1 | | 1 | 0.006 |
| 3 | 261 | 1 | 2 | 3 | 0.023 | 1 | 1 | 2 | 0.013 |
| 6 | 270 | 5 | 4 | 9 | 0.069 | 3 | 1 | 4 | 0.025 |
| 7 | 273 | 11 | 8 | 19 | 0.146 | 4 | 3 | 7 | 0.044 |
| 8 | 276 | 12 | 9 | 21 | 0.162 | 9 | 16 | 25 | 0.158 |
| 9 | 279 | 10 | 12 | 22 | 0.169 | 13 | 17 | 30 | 0.190 |
| 10 | 282 | 14 | 8 | 22 | 0.169 | 8 | 20 | 28 | 0.177 |
| 11 | 285 | 4 | 8 | 12 | 0.092 | 8 | 15 | 23 | 0.146 |
| 12 | 288 | 4 | 3 | 7 | 0.054 | 4 | 8 | 12 | 0.076 |
| 13 | 291 | 3 | 3 | 6 | 0.046 | 7 | 4 | 11 | 0.070 |
| 14 | 294 | 4 | | 4 | 0.031 | 4 | 3 | 7 | 0.044 |
| 15 | 297 | 1 | 1 | 2 | 0.015 | 2 | 2 | 4 | 0.025 |
| 16 | 300 | 1 | | 1 | 0.008 | 3 | | 3 | 0.019 |
| 21 | 315 | 1 | | 1 | 0.008 | 1 | | 1 | 0.006 |

NOTE: Chr.= chromosome; M= male; F= female; Pro.=proportion.

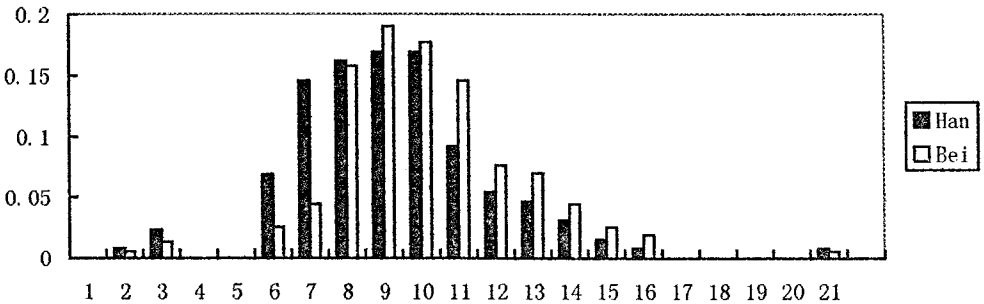


FIG. 2—*HUMARA allele frequency distribution from 101 unrelated individuals of Han population and 113 unrelated individuals of Bei population.*

TABLE 2—*The genotype frequencies of STR HUMARA locus in both Han and Bei populations.*

| Genotypes | | Number Observed/Proportions | | | |
|---|---|---|---|---|---|
| | | Han Population | | Bei Population | |
| Males | Females | Males | Females | Males | Females |
| 2–y | 2–2 | 1/0.014 | | 1/0.015 | |
| 3–y | 3–3 | 1/0.014 | | 1/0.015 | |
| | 3–9 | | 1/0.034 | | |
| | 3–15 | | 1/0.034 | | 1/0.022 |
| 6–y | 6–6 | 5/0.069 | | 3/0.044 | |
| | 6–7 | | 1/0.034 | | 1/0.022 |
| | 6–8 | | 2/0.070 | | |
| | 6–9 | | 1/0.034 | | |
| 7–y | 7–7 | 11/0.153 | | 4/0.058 | |
| | 7–8 | | 2/0.070 | | |
| | 7–9 | | 4/0.138 | | 2/0.044 |
| | 7–11 | | 1/0.034 | | |
| 8–y | 8–8 | 12/0.166 | 1/0.034 | 9/0.132 | 1/0.022 |
| | 8–9 | | 2/0.070 | | 3/0.067 |
| | 8–10 | | 1/0.034 | | 4/0.089 |
| | 8–11 | | | | 3/0.067 |
| | 8–12 | | | | 2/0.044 |
| | 8–13 | | | | 1/0.022 |
| | 8–14 | | | | 1/0.022 |
| 9–y | 9–9 | 10/0.138 | | 13/0.191 | 2/0.044 |
| | 9–10 | | 1/0.034 | | 3/0.067 |
| | 9–11 | | 2/0.070 | | 2/0.044 |
| | 9–12 | | | | 1/0.022 |
| | 9–13 | | 1/0.034 | | |
| | 9–14 | | | | 2/0.044 |
| 10–y | 10–10 | 14/0.194 | 2/0.070 | 8/0.118 | 4/0.089 |
| | 10–11 | | | | 2/0.044 |
| | 10–12 | | 1/0.034 | | 2/0.044 |
| | 10–13 | | 1/0.034 | | 1/0.022 |
| 11–y | 11–11 | 4/0.056 | 2/0.070 | 8/0.118 | 3/0.067 |
| | 11–12 | | 1/0.034 | | 1/0.002 |
| | 11–13 | | 1/0.034 | | |
| | 11–14 | | | | 1/0.022 |
| 12–y | 12–12 | 4/0.056 | | 4/0.059 | 1/0.022 |
| 13–y | 13–13 | 3/0.042 | | 7/0.103 | 1/0.022 |
| 14–y | 14–14 | 4/0.056 | | 4/0.059 | |
| 15–y | 15–15 | 1/0.014 | | 2/0.029 | |
| 16–y | 16–16 | 1/0.014 | | 3/0.044 | |
| 21–y | 21–21 | 1/0.014 | | 1/0.015 | |
| Total | | 72 | 29 | 68 | 45 |
| | | | $\chi^2 = 17.7206$ | | $\chi^2 = 7.4268$ |
| | | | Df = 11 | | Df = 10 |
| | | | $p > 0.05$ | | $p > 0.05$ |

Chi-square test based on observed/expected homozygosity. Small classes, those with fewer than seven persons, were combined for statistical tests.

carried out for the reason of different designations of HUMARA alleles and different PCR conditions in different reports (8).

The allele frequency distributions appear uni-model and symmetrical in both Han and Bei populations. The mechanism by which the different distribution modes arose has not been ascertained.

In females, the expected heterozygosity of the Han population was 0.91 and of the Bei population 0.97 which tend to be similar with those of US Hispanic (0.91), Black (0.89), White (0.87), and Asians (0.89) (7).

Our results indicate that the proper resolution and size determination can be achieved by using the non-denaturing native polyacrylamide gel electrophoresis because of increased electrophoretic migration and using an allele ladder which will compensate the potential variation in allele identification due to inter-laboratory differences in the techniques and equipments.

In conclusion, two Chinese populations (Han, Bei) data base has been established for STR HUMARA locus. The data demonstrated that this locus profile frequency can be derived for identity testing purpose.

## References

1. Sambrook J, Fritsch EF, Maniatis T. Molecular clonning; A Laboratory Manual, Cold Spring Harbour Laboratories Press, Cold Spring Habour, New York 1989.
2. Hammond HA, Jin L, Zhong Y, Caskey CT, Chakraborty R. Evaluation of 13 short tandem repeat loci for use in personal identification application. Am J Hum Genet 1994;55:175–89.
3. Budowle B, Chakraborty R, Giusti AM, Eisenberg AJ, Allen RC. Analysis of the VNTR-locus D1S80 by the PCR followed by high-resolution PAGE. Am J Hum Genet 1991;48:137–44.
4. Hou Y, Prinz M, Staak M. Comparison of different test for deviation from Hardy-Weinberg equilibrium of AmpFLP population data. In: Bar W, Friori A, Rossi V. Advances in forensic haemogenetics 5. Springer-Verlag: Berlin, Heidelberg, New York 1994:511–14.
5. Sensabaugh GF. Biochemical markers of individuality. In: Saferstein R, editor. Forensic Science Handbook. Prentice- Hall, Englewood Cliffs, NJ, 1982; 338–415.
6. Fregeau CJ, Fourney RM. DNA typing with fluorescent tagged short tandem repeats: a sensitive and accurate approach to human identification. BioTechniques 1993;15(1):100–19.
7. Edwards A, Civitello A, Hammond HA, Caskey CT. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am J Hum Genet 1991;49:746–56.
8. Edwards A, Hammond HA, Jin L, Caskey T, Chakraborty R. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics 1992;12:241–53.

Additional information and reprint requests:
Wu Mei Yun, M.D., Professor
School of Forensic Medicine
West China University of Medical Sciences
No. 17, Section 3, South Remin Road
Chengdu (610044). P.R. China