Available online at www.sciencedirect.com
ScienceDirect
Journal of Forensic and Legal Medicine 15 (2008) 329–334
JOURNAL OF FORENSIC AND LEGAL MEDICINE
www.elsevier.com/jflm

Original Communication

# Systematic evaluation of sensitivity and specificity of sibship determination by using 15 STR loci

Chang En Pu MS (Scientist Specialist and Section Chief) [a,*], Adrian Linacre PhD (Senior Lecturer) [b]

[a] *Forensic DNA Section, Scientific and Technical Research Center, Ministry Justice Investigation Bureau, No. 74 Chun_hya Road, Hsin-tien, Taipei County 231, Taiwan*
[b] *Centre for Forensic Science, Department of Pure and Applied Chemistry, University of Strathclyde, Glasgow, UK*

Received 17 June 2007; received in revised form 30 August 2007; accepted 31 December 2007
Available online 11 April 2008

**Abstract**

Paternity disputes and other forms of kinship testing are routinely resolved using short tandem repeat (STR) DNA loci. Sibship determination is encountered in instances where the DNA profiles of two individuals are compared to determine if they are siblings. If either parent is available for testing then the situation is simplified but if neither parent of the two individuals is available for DNA testing, a combined sibling indices (CSI) for the determination of sibship between two people can be determined. Support for kinship is also based upon the sharing of alleles, particularly when both alleles are shared at the same locus, termed two-allele-sharing-loci (TASL). We report on the combination of CSI and TASL to enhance the determination of sibship. The 15 STR loci that comprise the Identifiler® loci were applied to three populations using pairs of full siblings or unrelated pairs. Based upon the data obtained, the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) can be applied to determining whether two DNA profiles come from full or non-sibling pairs. This report highlights the problems inherent in this form of kinship testing and recommends a combination use of CSI and TASL for sibship determination.


*Keywords:* Short tandem repeat (STR); Kinship; Sibship; Combined sibship index (CSI); Sensitivity; Specificity

## 1. Introduction

Sibship determination is encountered in instances such as linking human remains to a relative and when neither of the biological parents of the two individuals is available for testing. Theoretically at any one locus there is a 0.5 chance that two siblings will share one allele, a 0.25 chance that they will share neither allele and a 0.25 chance that they will share both alleles. The chance that two unrelated individuals share either one or both alleles at any one locus is dependent upon the frequency of the alleles.[1,2] A probability of sibship can be determined based upon the frequency of the matching alleles in the population and will increase when a high number of high discriminating loci are examined.[3] Uncertainty using DNA testing to resolve sibship increases if the parents were heterozygous rather than homozygous.[1,4] There is no evidence of sibship if there is no two-allele-sharing-locus (TASL) between the two profiles,[5] however confidence is increased if a number of TASL existed.

Based upon the degree of sharing of alleles between two DNA profiles it is possible to determine a combined sibship indices (CSI).[1] When the index is less than 1 the two individuals might not be related as siblings. If the index is over 1 then the data supports the existence of a sibling relationship. Other cut-off point have been recommended such a CSI $\geqslant 3$.[2] These figures are guides as it was found that 1.6% of random pairs of DNA profiles had CSI greater than 1[2] when using 15 STR loci. In a study using 16 STR

* Corresponding author. Tel.: +886 2 29112241x3740; fax: +886 2 291 38599.
*E-mail address:* pu_macros@yahoo.com.tw (C.E. Pu).

1752-928X/$ - see front matter © 2008 Elsevier Ltd and FFLM. All rights reserved.
doi:10.1016/j.jflm.2007.12.018

loci 0.1% of unrelated pairs of DNA profiles were found to have CSI > 100 and 0.01% for CSI > 1000.[6] Using the 15 STR loci used in the Identifiler® kit, none of the non-sibling pairs were found with CSI ⩾ 1, while all sibling pairs have CSI > 10.[7] In a different study using the same loci, 6.06% of sibling pairs exhibited CSI < 1 and 9.1% of random unrelated pairs had CSI > 1.[8] The variation of percentage of random pairs with CSI > 1 found above maybe owing to the different levels of consanguinity that might be present in those populations.

We have extended the studies using the 15 STR comprising the Identifiler® loci to study three populations using 357,630 full sibling pairs and 178,815 non-sibling pairs generated from DNA profiles of random population. Using this high number of sample pairs, it is possible to evaluate the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the 15 STR loci for discriminating between full and non-siblings. The relationship between the CSI values and TASL is reported.

## 2. Materials and methods

DNA profiles from a Chinese population were obtained using the ABI AmpFISTR Identifiler® PCR Amplification Kit (Applied Biosystem, Foster City, CA, USA). The STR products were analyzed with an ABI Prism 3100 XL Genetic Analyzer.

STR profiles (450) of random members of the Taiwanese Chinese population were processed by Microsoft Excel Macros controlled by Visual Basic program written by authors of this study. Every member of the population was paired with every other member to form random pairs, e.g. for Chinese population in this study (450 × 449)/2, equaling 101,025 pairs, were made. Every pair was set to have two children, resulting in 202,050 sibling pairs being generated. DNA profiles from the 15 STR loci from a Caucasian ($n = 301$) and American African ($n = 256$) population were obtained from short tandem repeat DNA internet database.[9] These data were processed in the same way as those of the Taiwanese population to generate sibling pairs and random pairs. Inevitably for computer based populations no account of substructure is made and mating occurs randomly. The combined sibship indices (CSI) were calculated for each simulated sibling pairs or random pairs by using standard formulae,[1] and allele frequency tables used for calculation of CSI were adjusted by using 5/2N rule.[10]

The rate of false negatives equaled the percentage of real sibship testing cases (in this study the simulated sibling pairs) that would be excluded based upon any given cut-off point of CSI or TASL. The rate of false positives equaled the percentage of random pairs of DNA profiles where their CSI or TASL was greater than any recommended cut-off value. The sensitivity of the test is based upon 1 – the % of false negatives, the specificity of the test is based upon 1 – the % of false positives, the positive predictive value (PPV) = the proportion of subjects correctly identified as siblings and the negative predictive value

Table 1
Maximum and minimum CSI for three populations using 15 STR systems

| Population | Profiles | Type | Pairs | Maximum | Minimum |
|---|---|---|---|---|---|
| Chinese | 450 | Simulated siblings | 202,050 | 4.06E+16 | 1.68E−05 |
| | | Random pairs | 101,025 | 4.79E+04 | 2.33E−09 |
| Caucasians | 301 | Simulated siblings | 90,300 | 1.23E+14 | 3.42E−04 |
| | | Random pairs | 45,150 | 1.52E+04 | 6.38E−09 |
| African Americans | 256 | Simulated siblings | 65,280 | 3.54E+14 | 9.52E−04 |
| | | Random pairs | 32,640 | 6.24E+08 | 3.43E−09 |

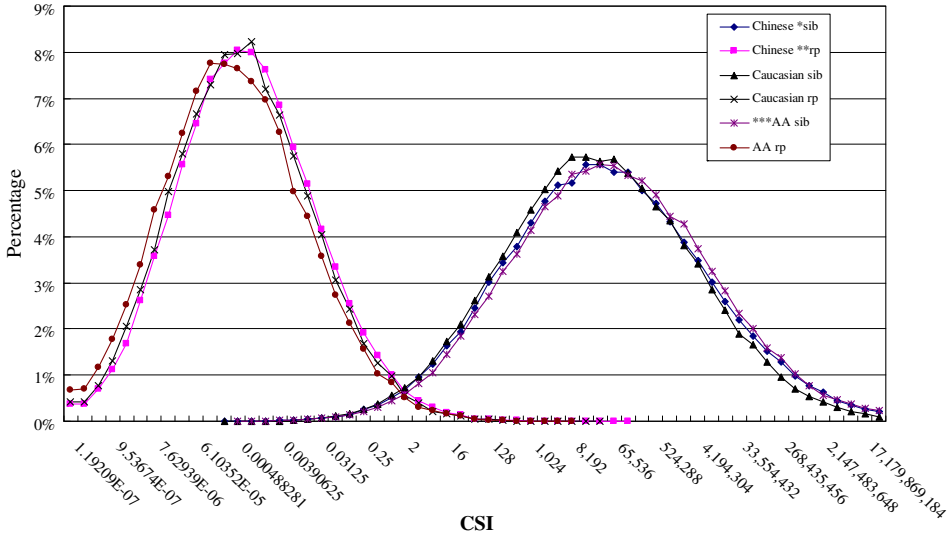


Fig. 1. Ratio distribution of CSI for three populations (*sib: siblings; **rp: random pairs; ***AA: African Americans).