Int J Legal Med (2002) 116 : 161–164

TECHNICAL NOTE

R. Gaytmenn · D. P. Hildebrand · D. Sweet
I. A. Pretty

Determination of the sensitivity and specificity of sibship calculations using AmpF/STR Profiler Plus

Received: 26 April 2001 / Accepted: 9 October 2001

Abstract In circumstances where a known DNA reference sample from the deceased's belongings or biological parents is not available, more complex kinship analyses are possible. The purpose of the work reported here is to determine the sensitivity and specificity of the sibship analysis utilising multiple STR loci. Using all nine Profiler Plus loci, likelihood ratios for biologically-related siblings ranged from slightly less than 1 to over 45,000. When allelic dropout was mimicked, likelihood ratios ranged from less than 1 to over 1,000. Thus, the results of this study have a direct application to forensic laboratories faced with identifications involving sibling comparisons.

Keywords Sibship · Siblings · STR loci · Human identification

Introduction

In situations where investigators are not able to make an identification of found human remains by conventional methods such as fingerprints and dental radiographs, DNA technology can be implemented to identify an individual. Amplification of trace DNA evidence using polymerase chain reaction (PCR) technology is now a well-established forensic technique. Since the target regions of the DNA molecule are microsatellite loci, or short tandem repeats (STR) in the range of 100–300 base pairs, it is possible to successfully amplify even partially degraded DNA [1]. The STR loci are amplified and analyzed simultaneously (multiplexed) producing a genotype at numerous polymorphic loci [2, 3].

DNA comparisons between a sample from the deceased individual and a known reference sample can be either direct or indirect. Known samples for direct comparisons are established to have come from the deceased, such as a hair from the deceased's hairbrush, a biopsy sample or a Pap-smear. Based on likelihood ratio calculations, the highest level of conclusions result from direct comparisons. Indirect comparisons are made between a known family member and the deceased using a kinship analysis, such as paternity, reverse paternity or sibship. Paternity or reverse paternity comparisons using genetic profiles of both putative parents typically result in the next highest level of conclusion. For parentage analysis, the child must have received one allele from the mother and the other from the father. A genotype set can be exclusionary if the alleged parents are not the biological parents of the individual, or as a result of mutation. Since mutational events are possible, many laboratories require that there are exclusions in a number of STR loci before calling an exclusion [4]. The average number of exclusionary loci in the AmpF*l*STR Profiler PCR amplification kit determined by one parentage study was five, with the minimum number of exclusionary loci being two [5].

In situations where neither direct comparison nor parentage analysis is feasible, sibship analyses by DNA comparison to an alleged sibling are possible. Sibship analyses are more complicated since there are no obligatory alleles between siblings that make it possible to exclude a biological relationship with absolute certainty. Although full siblings are more likely to share genetic components than not, a lack of shared alleles at any particular locus does not exclude two individuals from being related [6].

A previous study by Wenk et al. used three independent, polymorphic VNTR loci from 20 sibling pairs to develop and assess the validity of a sibling comparison test [7]. Results indicated that two pairs of siblings were not related. Given the inclusion of parental samples in the study to ensure true sibship, this is unlikely. This is sig-

R. Gaytmenn · D. Sweet
Bureau of Legal Dentistry,
146-2355 East Mall, Vancouver, BC, Canada V6T 1Z4

D. P. Hildebrand
British Columbia Institute of Technology,
3700 Willingdon, Burnaby, BC, Canada V5G 3H2

I. A. Pretty (✉)
University of Liverpool, Department of Clinical Dental Sciences,
Edwards Building, Daulby Street, Liverpool, L69 3GN, UK
e-mail: ipretty@liv.ac.uk,
Tel.: +44-151-7065288, Fax: +44-151-7065908

Table 1 Certainty thresholds for likelihood ratios

Likelihood ratio	Evidence strength
1– 33	Weak
33– 100	Fair
100– 330	Good
330–1000	Strong
> 1000	Very strong

nificant in that it could lead to incorrect conclusions in casework.

The likelihood ratio (LR) is the ratio of two probabilities; the probability of the genetic evidence given event one divided by the probability of the genetic evidence given event two [8]. Guidelines for establishing a degree of certainty based on the calculated LR value (Table 1) have been utilised in this study as a means of determining a certainty threshold [9]. While the scale is arbitrary, it is used to determine the strength of the evidence when interpreting LRs. Additionally, by calculating known error rates this study will help to satisfy the increased judicial regulations that require experts to define error rates for quoted tests.

STRs have been shown to be highly discriminating in determining both direct [3, 10] and parent/child relationships [5, 11, 12]. The intent of the work reported here is to extend the study performed by Wenk et al. [7] to STRs and to meet requirements imposed by elevated court scrutiny. Using 19 known related sibling pairs, as well as random cross-pairing of unrelated individuals, the sensitivity and specificity of the sibship analysis using AmpF*l*STR Profiler Plus (Applied Biosystems, Foster City, Calif.) were calculated at a range of certainty thresholds as well as a variety of loci. Degraded samples are often submitted as questioned samples in cases involving the identification of human remains. Often these degraded samples produce partial profiles when the largest loci do not amplify during PCR (allelic drop-out). The loss of these loci generally results in a lower level of certainty in the conclusion. Partial profiles were artificially generated by excluding the largest loci to mimic allelic drop-out. Thus, the results of this study have a direct application to forensic laboratories faced with identifications involving sibling comparisons.

Materials and methods

Sample preparation and amplification

The biological samples used in this study consisted of buccal swabs stored on FTA cards (Fitzco, Maple Plain, MN) from 21 unrelated volunteer families of Canadian Caucasian decent, consisting of a mother, father and two siblings. Permission to conduct this testing was granted by the Clinical Research Ethics Board of the University of British Columbia. Each volunteer gave their informed consent prior to inclusion in this study. The sample collection process was conducted under blind testing conditions to maintain anonymity and reduce bias.

Samples were removed from the cards by punching 1.2 mm circles using a Harris Micro-Punch (Shunderson Communications, Ottawa, ON) and a self-healing mat. The circles were placed directly into 0.2 ml PCR tubes and washed three times for 5 min at RT with FTA purification reagent (200 µl) and twice for 5 min at RT with TE buffer (10 mM Tris HCl pH 8.0, 0.1 mM EDTA pH 8.0; 200 µl). The buffer was removed from the last TE wash and the circles were dried (60 °C for 20 min).

When the samples were dry, 5 µl of sterile distilled water was added followed by 7.5 µl of PCR master mix (5.25 µl reaction mix, 2.75 µl Profiler Plus primer set, 0.25 µl AmpliTaq Gold). The PCR samples were amplified according to the manufacturer's protocol (95 °C for 11 min, 28 cycles of 94 °C for 1 min, 59 °C for 1 min, 72 °C for 1 min, 60 °C for 45 min, 25 °C infinite).

Following PCR 1 µl of each sample was added to 12 µl of deionized formamide, heated to 95 °C (3 min) and cooled in an ice bath (3 min). Separation and detection of the amplified product was carried out on an ABI 310 Genetic Analyser (Applied Biosystems, Foster City, Calif.) with GeneScan 2.1 and Genotyper 2.0 software. Any samples that failed to amplify using the above procedure were re-punched with a 2 mm Harris Micro-Punch and re-amplified. In order to accommodate the larger sample size, the PCR reaction volume was doubled (25 µl).

Parentage and sibling analysis

Although samples were collected from known family units, the parentage of each mother-father-child trio was confirmed for both siblings by paternity analysis. The likelihood ratio for these calculations took the form $LR = Pr(G_C/G_M, G_F, H_1)/Pr(G_C/H_2)$, where GC is the genotype of the child, GM is the genotype of the putative mother, GF is the genotype of the putative father, H1 is the probability of observing the child's genotype given the genotypes of the mother and father and the proposition that the child is the biological child of these parents and H2 is the probability of observing the child's genotype given the alternative proposition that a random member of the population is the putative child. Likelihood ratios greater than 1 indicate that the genetic evidence is more probable if the putative parents are the biological parents. Likelihood ratios less than 1 indicate that the genetic evidence is more probable if the putative parents are not the biological parents. For every trio tested, likelihood ratios were greater than 1, indicating that common parentage was very likely for each sibling pair. Once it was corroborated that each child was in fact the biological offspring of each set of parents, the parents were removed from the study to mimic a true sibship analysis. Negative controls were generated by random pairing of unrelated individuals from the nuclear families to produce 19 non-sibling matches.

The notation and formulae described by Wenk et al. [7] were used in this study. Siblings were arbitrarily labelled as S1 or S2. The likelihood ratio for the sibling analysis, called a sibling index (SI), took the form $SI = Pr(G_{S2}/G_{S1}, H_1)/Pr(G_{S2}/H_2)$, where G_{S2} is the genotype of sibling 2 and G_{S1} is the genotype of sibling 1. The numerator denotes the probability of observing the genotype of S2 given the genotype of S1 and the proposition that the two individuals are truly related. The evaluation of the numerator depends on the genotype band/peak pattern observed for each locus and the use of the kinship coefficients for full siblings. The denominator denotes the probability of observing the S2 genotype under the proposition that the two individuals are unrelated, which is simply the genotype frequency of S2. Combined SI's were determined by multiplying the individual values of the chosen loci. The allele frequencies used in both the parentage and sibling calculations were obtained from the Canadian Caucasian database.

Sensitivity and specificity of the sibship analysis

The sensitivity of the test is the probability of the test being positive given that the two people being tested are in fact siblings. The specificity of the test is the probability of the test being negative given that the two people are not siblings. Sensitivity and specificity of these tests are stable properties that remain constant when various proportions of siblings and non-siblings are tested [13]. The positive predictive value (PPV) tells what proportion of subjects have been correctly identified as siblings. The negative pre-

dictive value (NPV) tells what proportion of subjects have been correctly identified as non-siblings. Accuracy is the proportion of sibship results that agree with the parentage results. Predictive values are not constant because they are dependent on the prevalence of siblings in the studied population. In this study, prevalence, the proportion of siblings to non-siblings, was arbitrarily set to 50% so there were an equivalent number of sibling and non-sibling comparisons.

Results and discussion

Of the 21 unrelated families originally used in this study, only 19 were actually used for the sibling analysis as 2 families were excluded from the study due to different mutation events at a single locus. The SI results for the known sibling pairs are shown in Table 2.

SI calculations using all nine Profiler Plus loci

Due to the variation of relative allele frequencies in the population, a wide range of combined SI values was observed. When using all nine loci (D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317, D7S820) the values ranged from slightly less than 1 to over 45,000. Under these circumstances, only one known sibling pair resulted in a SI less than 1, indicating the evidence is more probable if the siblings were unrelated. Given the confirmed parentage of the siblings, this is not likely. Random pairing of non-siblings was created by matching unrelated individuals from the 19 nuclear families and used to calculate combined SI's. The 19 comparisons resulted in no SI values greater than 1 when all nine loci were tested (data not shown). Sensitivity and specificity of the results are shown in Table 3. When adopting a certainty threshold with a lower limit of 1 (i.e. all SI values greater than 1 are considered positive), the sibship calculation has 94.7% sensitivity, 100.0% specificity, a PPV of 100.0%, a NPV of 95.0% and an accuracy of 97.4%. This means that 94.7% of siblings will be considered as siblings, 100% of non-siblings will be considered non-siblings, 100% of the subjects are correctly identified as siblings, 95.0% of the subjects are correctly identified as non-siblings and 97.4% of the subjects are correctly identified as either siblings or non-siblings. When adopting a certainty threshold greater than 100 (i.e. all SI values greater than 100 are considered positive), the sibship calculation has 52.6% sensitivity, 100.0% specificity, a PPV of 100.0%, a NPV of 67.8% and an accuracy of 76.3%.

Table 2 Sibship indices for known sibling pairs

Family	Complete[a]	Partial (6 loci)[b]	Partial (3 loci)[c]
1	262.2	36.6	3.5
2	65.7	1.1	4.7
3	0.02	1.3	1.1
4	50.9	11.6	0.5
5	18.8	5.5	2.2
6	1389.8	74.6	38.2
7	155.6	682.1	8.2
8	7715.6	1290.3	7.9
9	1.8	0.1	0.2
10	45699.0	4020.9	1111.5
11	6473.1	177.2	9.9
12	11.2	3.6	1.4
13	11836.7	140.6	35.7
14	7474.6	7.1	0.1
15	4414.2	179.4	0.9
16	2443.4	296.6	64.8
17	2.1	131.0	1.7
18	2.2	73.0	0.2
19	7.1	3.2	0.01

[a] 9 loci of AmpF*l*STR Profiler Plus.
[b] 6 loci: D3S1358, vWA, D8S1179, D21S11, D5S818, D13S317.
[c] 3 loci: D3S1358, D8S1179, D5S818.

SI calculations using six Profiler Plus loci

Amplification of DNA from skeletonized or highly degraded remains often results in allelic drop-out of the largest loci of Profiler Plus due to stochastic events or degradation. In order to mimic this scenario, we repeated the SI calculations but omitted three loci (FGA, D18S51 and D7S820). As in the previous example, the SI values took a wide range (from less than 1 to over 4,000). One known sibling pair resulted in a SI of less than 1. As above, random pairing of siblings known to be unrelated were also used to calculate combined SI's and 19 comparisons of non-siblings resulted in 2 SI values greater than 1 (data not shown). Sensitivity and specificity of the results are shown in Table 4.

SI calculations using three Profiler Plus loci

Of the 19 families tested, 6 resulted in SI's less than 1. Sibling index calculations using 19 unrelated sibling pairs resulted in 5 SI's with values greater than 1 (data not shown). Sensitivity and specificity of the results are shown in Table 5.

Table 3 Validity of sibship indices using 9 loci

Evidence strength	Sensitivity (%) (95% C.I.)	Specificity (%) (95% C.I.)	PPV (%)	NPV (%)	Accuracy (%)
> 1 (weak)	94.7 (76.7–99.7)	100.0 (85.4–100.0)	100.0	95.0	97.4
> 33 (fair)	63.2 (40.3–82.3)	100.0 (85.4–100.0)	100.0	73.1	81.6
> 100 (good)	52.6 (30.6–73.9)	100.0 (85.4–100.0)	100.0	67.8	76.3
> 330 (strong)	42.1 (21.8–64.6)	100.0 (85.4–100.0)	100.0	63.3	71.1
> 1000 (very strong)	42.1 (21.8–64.6)	100.0 (85.4–100.0)	100.0	63.3	71.1

Table 4 Validity of sibship indices using 6 loci

Evidence strength	Sensitivity (%) (95% C.I.)	Specificity (%) (95% C.I.)	PPV (%)	NPV (%)	Accuracy (%)
> 1 (weak)	84.2 (62.8–95.8)	89.5 (69.4– 98.2)	88.9	85.0	86.8
> 33 (fair)	57.9 (35.3–78.2)	100.0 (85.4–100.0)	100.0	70.4	79.0
> 100 (good)	42.1 (21.8–64.6)	100.0 (85.4–100.0)	100.0	63.3	71.1
> 330 (strong)	15.8 (4.2–37.2)	100.0 (85.4–100.0)	100.0	54.3	57.9
> 1000 (very strong)	10.5 (1.8–30.6)	100.0 (85.4–100.0)	100.0	52.8	55.3

Table 5 Validity of sibship indices using 3 loci

Evidence strength	Sensitivity (%) (95% C.I.)	Specificity (%) (95% C.I.)	PPV (%)	NPV (%)	Accuracy (%)
> 1 (weak)	66.7 (43.1–85.3)	73.7 (50.9– 89.7)	71.7	68.9	68.4
> 33 (fair)	21.1 (7.1–43.3)	100.0 (85.4–100.0)	100.0	55.9	60.5
> 100 (good)	5.3 (0.3–23.3)	100.0 (85.4–100.0)	100.0	51.4	47.4
> 330 (strong)	5.3 (0.3–23.3)	100.0 (85.4–100.0)	100.0	51.4	47.4
> 1000 (very strong)	5.3 (0.3–23.3)	100.0 (85.4–100.0)	100.0	51.4	47.4

The specificity of tests for both six and nine loci were high. However, the sensitivity varied depending on the certainty threshold. The higher the certainty threshold, the less sensitive the test. A high value for sensitivity is obtained when the number of false negatives is low, while a high value for specificity is obtained when the number of false positives is low. The thresholds of sensitivity and specificity a laboratory decides to adopt may depend on whether they are evaluating a criminal case or making an identification and in the case of an identification, whether circumstantial evidence is also available. In this study accuracy is a reflection of the number of true siblings detected plus the number of non-siblings detected divided by the total number of possible sibling pairs. Accuracy values ranged from 97.4% to 68.4% at the lowest certainty threshold to 71.0%–47.4% at the highest level.

Conclusions

This study, which represents initial findings of a small sample size, illustrates that results based on sibship calculations should be interpreted with caution. Comparison of the results for nine loci, six loci and three loci give some indication of the recommended number of loci to include when determining sibship. As expected, the poor sensitivity and specificity of the triplex for sibship determination does not warrant its use in forensic casework. Due to the lack of obligatory alleles, one cannot determine with absolute certainty whether or not two individuals are siblings. The more loci used, however, the stronger the statistical argument. As the sensitivity and specificity results indicate, sibship analysis is better used as an inclusion of sibship than an exclusion. Even the use of nine loci can result in false negatives or SI values that are only slightly in favour of kinship (i.e. Families 9, 17, 18 and 19 in Table 2). Testing of kits that utilise more discretionary loci would further this study by denoting loci that would be more advantageous for sibship analysis.

Acknowledgements The authors wish to thank Dr. George Carmody and Dr. Ron Fourney for supplying STRquest II software and data from the Canadian Caucasian DNA database.

References

1. Fregeau CJ, Fourney RM (1993) DNA typing with fluorescently tagged short tandem repeats: a sensitive and accurate approach to human identification. Biotechniques 15:100–119
2. Kimpton CP, Gill P, Walton A, Urquhart A, Millican ES, Adams M (1993) Automated DNA profiling employing multiplex amplification of short tandem repeat loci. PCR Methods Appl 3: 13–22
3. Kimpton CP, Oldroyd NJ, Watson SK, Frazier RR, Johnson PE, Millican ES, Urquhart A, Sparkes BL, Gill P (1996) Validation of highly discriminating multiplex short tandem repeat amplification systems for individual identification. Electrophoresis 17:1283–1293
4. Gunn P, Trueman K, Stapleton P, Klarkowski DB (1997) DNA analysis in disputed parentage: the occurrence of two apparently false exclusions of paternity, both at short tandem repeat (STR) loci, in the one child. Electrophoresis 18:1650–1652
5. Brinkmann B, Pfeiffer H, Schürenkamp M, Hohoff C (2001) The evidential value of STRs. An analysis of exclusion cases. Int J Legal Med 114:173–177
6. Bourke MT, Ladd C, Yang NCS, McGilvray MK, Lee HC, Bieber FR (2000) Sib pair identity at multiple STR loci: implications for interpretation of forensic DNA casework. In: Proceedings of the 52nd Annual Meeting of the American Academy of Forensic Sciences. Reno, NV, pp 30–31
7. Wenk RE, Traver M, Chiafari FA (1996) Determination of sibship in any two persons. Transfusion 36:259–262
8. Evett IW, Weir BS (1998) Interpreting DNA evidence: statistical genetics for forensic scientists. Sinauer Associates, Sunderland, MA
9. Evett IW (1991) Strength of evidence in likelihood ratios. The use of statistics in forensic science. Ellis Horwood, New York
10. Hochmeister MN, Budowle B, Borer UV, Rudin O, Bohnert M, Dirnhofer R (1995) Confirmation of the identity of human skeletal remains using multiplex PCR amplification and typing kits. J Forensic Sci 40:701–705
11. Thomson JA, Pilotti V, Stevens P, Ayres KL, Debenham PG (1999) Validation of short tandem repeat analysis for the investigation of cases of disputed paternity. Forensic Sci Int 100:1–16
12. Zupaniè Pajniè I, Šterlinko H, Balaic J, Komel R (2001) Parentage testing with 14 STR loci and population data for 5 STRs in the Slovenian population. Int J Legal Med 114:178-180
13. Brunette DM (1996) Critical thinking. Quintessence, London