Biostatistical approaches using minisatellite DNA patterns in paternity cases (mother-child-putative-father trios)

Prof. Dr. K. Hummel, Dipl.-Math. Nina Fukshansky, Institute for Blood-Group Serology, POB 880, D-7800 Freiburg (FRG), PD Dr. W. Bär, Institute for Forensic Medicine, University of CH-8028 Zürich (Switzerland) and Prof. Dr. K. Zang, Institute for Human Genetics of the University D-6650 Homburg (FRG)

The calculation of probabilities of paternity from fingerprints of a multilocus probe is relatively simple in two groups of cases: normal mother-child-putative-father (M-C-PF) trios, and cases where findings for the mother or for the PF are lacking but findings for their respective parents are available. Cases where there is no exclusion also lend themselves to biostatistical evaluation. A M-C-PF trio has 6 genetically compatible fingerprint patterns; one determines their frequencies from the pherograms by counting. If each visible band is regarded as a detectable property of a di-allele system (heterozygous or homozygous), than an invisible band represents the homozygous phenotype of the system's silent property.

In Essen-Möller's formula for the probability of paternity, $W = \frac{1}{1+Y/X}$,

X is the frequency of the trios with the real father,
Y with the false father.

Taking $\underline{a}$ is the mean band frequency, the following formulas are to use to calculate the likelihood ratios for the 6 unequivocal M-C-PF band constellations:

running no.	M C PF	Y/X	
1	- + +	a	$= f_1$
2	+ + +	$a(a+\sqrt{1-a})/(2a+\sqrt{1-a}\ -1)$	$= f_2$
3	+ + -	$a+\sqrt{1-a}$	$= f_3$
4	+ - -	$\sqrt{1-a}$	$= f_4$
5	- - +	$1+\sqrt{1-a}$	$= f_5$
6	+ - +	$1+\sqrt{1-a}$	$= f_6$

Furthermore it is possible that all three persons are homozygous in respect of the silent antithetic "allele" which means there is a 7th constellation which cannot be quantitatively determined from the pherogram:

7	- - -	$\sqrt{1-a}$	$= f_7$

To obtain it, one subtracts the number of obvious constellations ($n = n_1+n_2+n_3+n_4+n_5+n_6$) from the total of all the visible and invisible bands recorded by the probe (= N).

Advances in Forensic Haemogenetics 3
Edited by H. F. Polesky and W. R. Mayr

Essen-Möller's probability of paternity is then calculated as follows:

$$W_{EM} = 1/(1+(f_1^{n_1} \cdot f_2^{n_2} \cdot f_3^{n_3} \cdot f_4^{n_4} \cdot f_5^{n_5} \cdot f_6^{n_6} \cdot f_7^{n_7}))$$

where $n_1 \ldots n_7$ denote the respective "frequencies" of the different band constellations found.

To obtain a W_{EM} value in the individual case with the formula given above one must know

1. the typical mean band frequency for the probe, a; and
2. the total number N of visible and invisible bands recorded by the probe in question.

The values a and N characterize a multilocus probe sufficiently for the biostatistical evaluation of band patterns. These values are determined from the band constellations of, say, 100 M-C-PF cases ("families"). The method is as follows:

We proceed from the M-C-PF constellation + + +. The phenotype is based on genotype M-C-PF constellations with the overall frequency

$$p = x^2(3-2x),$$

where x = mean frequency of the "alleles" coding the bands.

We regard the appearance of the band constellation + + + for a M-C-PF trio as a "success" in a Bernoulli Test. The probability of this success is p. The number of times + + + constellation appears for a M-C-PF trio - null, one, two, three and more times - is given by the probabilities P(0), P(1), P(2), P(≧3).

Then, for P(i):

$$P(0) = (1-p)^N$$
$$P(1) = N \cdot p(1-p)^{N-1}$$
$$P(2) = 1/2\,(N(N-1)) \cdot p^2(1-p)^{N-2}$$
$$P(\geq 3) = 1-(P(0)+P(1)+P(2)).$$

Among m real M-C-PF trios the + + + constellation may not appear, m_0, appear once, m_1, twice, m_2, or three or more times, m_3;

$$m = m_0+m_1+m_2+m_3.$$

The probability of this series is proportional to

$$L = P(0)^{m_0} \cdot P(1)^{m_1} \cdot P(2)^{m_2} \cdot P(\geq 3)^{m_3}.$$

If M denotes the mean number of visible bands per person in the pherograms of n M-C-PF trios then

$$M = a \cdot N.$$

M can be determined from readings of the pherograms without difficulty; if p is known, the equation $p = x^2(3-2x)$ can be solved for a:

$$a = x(2-x).$$

One uses the maximum-likelihood method to obtain parameter p, solving for that value that will maximize the likelihood function:

$$\ln L = m_0 \cdot \ln P(0)+m_1 \cdot \ln P(1)+m_2 \cdot \ln P(2)+m_3 \cdot \ln P(\geq 3).$$

Using probe 33.15 on 94 M-C-PF trios gave the following results:

The mean number of the visible bands per person was 7.78 (M = 7.78). The frequencies of the + + + constellation were

$m_0 = 57$
$m_1 = 23$
$m_2 = 11$
$m_3 = 3$
$m = 94$

Using the maximum-likelihood method, the total number of visible and invisible bands recorded by the probe 33.15 is $N = 71$

with a mean band frequency of $a = 0.11$.

One could similarly run the other 5 unequivocal M-C-PF constellations. Of these, the only really significant one is the constellation + - + because in it, as in constellation + + +, one band occurs in two non-related individuals.

Once the maximum number of bands recorded by the probe is known by 71 it is not difficult to calculate the frequency of the constellation - - - for the individual case: $n_7 = N-n$.

Using a = 0.11, the formula for W_{EM} will give Essen-Möller's probability of paternity for a M-C-PF constellation. W_{EM}% tells one how accurate one could be if in 100 serologically analogous cases (= cases with the same fingerprints) one would categorically decide that in every case the putative father was the real F of the C. 100-W% would give the expectation of error for such a categorical decision.

Obviously, the calculation of W_{EM} presupposes formal genetics of band heredity. Regardless of features, some experts regard all formal genetics for multilocus band patterns as arbitrary. This explains why they hesitate to calculate Essen-Möller probabilities from multilocus probes.

Instead of the "general" solution (using W_{EM}), one can choose a "limited" solution. This is based on the M-C-PF constellation - + + alone. The likelihood ratio Y/X of a single - + + constellations has the value a (= mean band frequency). If n such constellations appear in the pherograms, then $Y/X(n) = a^n$.

Substituting in a formula based on the Essen-Möller principle gives a probability of paternity

$$W = \frac{1}{1+Y/X} = \frac{1}{1+a^n}.$$

The subscript A (in W_A) signifies that W_A is not based on all band constellations (like W_{EM}) but only on those which could produce an exclusion in the event of non-paternity (that is, on the M-C constellation - +). W_A(%) tells one how often one would be accurate if in 100 "analogous" cases - that is, cases with the same number of - + M-C constellations - one would regard all non-excludable men as real fathers; 100-W_A(%) would be the quota of error for such a categorical decision.

In most cases W_A is less than W_{EM} because W_A omits the M-C-PF constellation - - - that also strengthens the positive indication. The "advantage" of W_A is that it is based only on the principle that the child must have inherited from the PF a certain number of bands. There are no other, least of all formal genetic, prerequisites for W_A. The disadvantage of W_A is that its application is restricted to trios (normal as well as incest trios) and grandparental cases. If the exclusion expectation for non-kinship is very low or even zero - which is almost always the case in deficiency cases - the value for W_A will not be very helpful. In such cases only the probability value W_{EM} can provide useful information. - Our computer print-outs for trios and grandparental cases include both W_{EM} and W_A.