Stochastic Models for Low Level DNA Mixtures

Dalibor Slovák[1,2], **Jana Zvárová**[1,2]

[1] EuroMISE Centre, Institute of Computer Science AS CR, Prague, Czech Republic

[2] Institute of Hygiene and Epidemiology, First Faculty of Medicine, Charles University, Prague, Czech Republic

Abstract

Objectives: The increasing sensitivity of forensic analysis methods allows to investigate less and less amount of biological samples. For samples of low quality or quantity, there are stochastic events that require intensive statistical analysis.
Methods: There are several models how to calculate the probability of a given set of alleles. We have described three of them and compared them to verify their accuracy.
Results: The two models proposed in [1] extend so far the most widely used model by the possibility of dropout and peak areas of individual alleles. The first one is incorrect, while the second model highly improves the possibility of DNA mixture analysis.
Conclusions: We have shown the inaccuracy of one of the recently proposed models. We have added the possibility of determining the dropout probability into the second model, otherwise this model overestimates the probabilities calculated.

Keywords

Forensic DNA interpretation, low level samples, allele peak areas, dropout probability

Correspondence to:

Dalibor Slovák
EuroMISE Centre, Institute of Computer Science, AS CR, v.v.i.
Address: Pod Vodarenskou vezi 2, Prague 8, Czech Republic
E–mail: slovak@cs.cas.cz

EJBI 2012; 8(5):25–30
recieved: August 16, 2012
accepted: September 4, 2012
published: November 22, 2012

1 Introduction

With the increasing sensitivity of methods used for forensic DNA analysis, collection of forensic traces can be accomplished from a very small amount of biological material. Therefore, the increasing number of incomplete or contaminated profiles and profiles originating from more contributors are appearing. The samples containing only a small amount of DNA (approximately up to 100 pg / ml) are called low level samples and various stochastic effects occur increasingly for these samples.

Some laboratories perform the analysis of samples twice or more. Curran et al. [2] introduced the set theory in order to enable the calculations to be made in these cases. However, we do not attempt to explain their theory in this paper.

The result of laboratory processing of DNA samples is electropherogram (epg), which displays the alleles present at particular loci and peak heights measured in relative fluorescence units (RFU). Currently the most common laboratory sets process sixteen loci.

Two main approaches to DNA mixture interpretation are currently discussed in forensic practice. The Random Man Not Excluded method (RMNE) calculates the probability of observing the DNA profile needed for evidence, given that the DNA profile comes from a random individual, unrelated to the suspect. In other words, it is the probability that the DNA profile from a random person is the same as the evidence DNA profile, and that this person therefore, due to the evidence, cannot be excluded from suspicion.

The Likelihood Ratio approach (LR) compares the probabilities of observing the evidence under two rival hypotheses: typically the prosecution hypothesis H_p, the probability that the suspect is one of the contributors to the mixture, and the defense hypothesis H_d, the probability that the suspect does not contribute to the mixture.

The advantage of the LR framework is that dropout can be assessed probabilistically and it is the only way to provide a meaningful calculation based on the probability of the evidence under H_p and H_d. A likelihood ratio approach is therefore preferred [3]. For a more detailed comparison of both methods, see [4].

If the allele which is present in the sample is not displayed on the epg we call such an event an (allelic) dropout. If no allele is displayed at the locus, we talk about locus dropout.

If n persons is assumed to contribute to the mixture, maximum of $2n$ alleles can appear at the locus. However, some alleles may be represented several times, others may be missing due to the dropout. The observed mixed profile is therefore usually made up of fewer alleles. Under such conditions, there are more possibilities how to reconstruct individual DNA profiles from observed mixed profile.

Kelly et al. [1] suggested two stochastic models to compute the probability of observing the mixed profile. They compare them with most commonly used model, designated there as the unconstrained combinatorial (UC) method. In this article, the comparison of the three models will be discussed.

Although this theory is easily extended to multiple loci, in the present article, we consider only one locus in the profile and some realities are omitted for simplification, e.g. contamination and drop-in possibility or population structure. The number of contributors to the mixture will be assumed to be known.

2 Methods

From the epg, not only alleles present may be found out but also the peak heights. This information can help us to distinguish e.g. component belonging to the dominant contributor, but even if it is not possible to divide precisely individual components of the mixture, peak heights can inform us about the presence of multiple copies of several alleles. However, the decision on whether the allele is present in multiple copies strongly depends on the assessment of forensic expert and his experience.

The calculation of a LR may proceed by either a binary, a semi-continuous, or a fully continuous method. The binary and semi-continuous methods treat alleles as present or absent, moreover the semi-continuous method assigns a probability to the events of dropout or non-dropout. Fully continuous method deals with the probability of drop-out and other stochastic events based on the heights of the peaks visualised at a locus. Only binary methods are compared here.

Software processing epgs usually shows two thresholds for more simple interpretation. If the signal is below the limit of detection (LOD), we consider it as a noise. The detection limit is usually determined as 25 or 50 RFU or is calculated as the average noise signal plus three its standard deviations.

The stochastic threshold T is a value above which the dropout is excluded. In case that there is only one signal above the stochastic threshold, it may be assumed that it is a homozygous profile [5]. T is usually in the range of 150-300 RFU or may be calculated as the average noise signal plus ten its standard deviations.

Now let us consider two examples with the limit of detection $LOD = 50$ RFU and the stochastic threshold $T = 300$ RFU. The observed profile will be denoted by X and the set of all occurring alleles (allelic vector) will be denoted by A.

Example 1

The alleles 13, 14 and 15 with values of 180, 195, and 212 RFU, respectively, are observed at the locus. The mixture is assumed to originate from two contributors. Thus it is the profile $X = [13, 14, 15]$ for which the peak heights on the epg are approximately the same for all alleles. Under these assumptions, one allele is missing in the allelic vector A - either there was a dropout, or some of the contributors is homozygote, or both contributors have an allele of the same type.

Example 2

The alleles 13, 14 and 15 with values of 150, 470 and 420 RFU, respectively, are observed at the locus. From the analysis of other loci in the same sample, the mixture is assumed to originate from three contributors. Thus it is the profile $X = [13, 14, 15]$ again but now there are three missing alleles to complete the allelic vector. The observed alleles also have quite different peak heights which encourage to the inclusion of multiple copies of some alleles into the allelic vector, but for now we let this opportunity unused. We will return to it later in the section 3.

Now we describe proposed models and show their application to both the examples mentioned above.

2.1 UC Model

The unconstrained combinatorial method does not allow for possibility of dropout nor include peak heights to the calculation. The allelic vector can be completed only by copies of alleles observed.

Example 1:

$$\begin{aligned}
&\mathsf{P}(X = [13, 14, 15]) = \\
&= \mathsf{P}\left(A \in \left\{[13^2, 14, 15], [13, 14^2, 15], [13, 14, 15^2]\right\}\right) = \\
&= \frac{4!}{2!1!1!}p_{13}^2 p_{14} p_{15} + \frac{4!}{1!2!1!}p_{13} p_{14}^2 p_{15} + \frac{4!}{1!1!2!}p_{13} p_{14} p_{15}^2 = \\
&= 12 p_{13} p_{14} p_{15}(p_{13} + p_{14} + p_{15}). \qquad (1)
\end{aligned}$$

Example 2:

$$\begin{aligned}
&\mathsf{P}(X = [13, 14, 15]) = \\
&= \mathsf{P}\left(A \in \left\{[13^4, 14, 15], [13^3, 14^2, 15], [13^3, 14, 15^2],\right.\right. \\
&\quad \left.\left.[13^2, 14^3, 15], [13^2, 14^2, 15^2], [13^2, 14, 15^3], [13, 14^4, 15],\right.\right. \\
&\quad \left.\left.[13, 14^3, 15^2], [13, 14^2, 15^3], [13, 14, 15^4]\right\}\right) = \\
&= \frac{6!}{4!1!1!}p_{13}^4 p_{14} p_{15} + \frac{6!}{3!2!1!}p_{13}^3 p_{14}^2 p_{15} + \frac{6!}{3!1!2!}p_{13}^3 p_{14} p_{15}^2 + \\
&+ \frac{6!}{2!3!1!}p_{13}^2 p_{14}^3 p_{15} + \frac{6!}{2!2!2!}p_{13}^2 p_{14}^2 p_{15}^2 + \frac{6!}{2!1!3!}p_{13}^2 p_{14} p_{15}^3 +
\end{aligned}$$

$$+\frac{6!}{1!4!1!}p_{13}p_{14}^4p_{15}+\frac{6!}{1!3!2!}p_{13}p_{14}^3p_{15}^2+\frac{6!}{1!2!3!}p_{13}p_{14}^2p_{15}^3+$$
$$+\frac{6!}{1!1!4!}p_{13}p_{14}p_{15}^4=$$
$$=30p_{13}p_{14}p_{15}\left(p_{13}^3+2p_{13}^2p_{14}+2p_{13}^2p_{15}+2p_{13}p_{14}^2+\right.$$
$$\left.+3p_{13}p_{14}p_{15}+2p_{13}p_{15}^2+p_{14}^3+2p_{14}^2p_{15}+\right.$$
$$\left.+2p_{14}p_{15}^2+p_{15}^3\right). \quad (2)$$

2.2 F and Q Models

F and Q models were suggested by Kelly et al. [1] as an extension of UC model. Compared to this model, they allow to calculate with the possibility of dropout and to use the information about peak heights.

In F model, any allele completing the observed profile to the allelic vector is denoted by F. For example, under conditions of Example 1 Kelly et al. state

$$\mathsf{P}(X=[13,14,15])=\mathsf{P}(A=[13,14,15,F])=$$
$$=\frac{4!}{1!1!1!1!}p_{13}p_{14}p_{15}=24p_{13}p_{14}p_{15}. \quad (3)$$

However, F model is incorrect due to the non-differentiation between observed and unobserved alleles. If the allele designated as F is of the same type as an allele already observed, the number of possible combinations is less than if we assume that all alleles are different. Thus, F model overestimates computed probabilities. In the case of Example 2, we get $120p_{13}p_{14}p_{15}$ which gives the senseless probability 1.875 for values $p_{13}=p_{14}=p_{15}=0.25$. Therefore, we will continue to consider only model Q.

In Q model, any allele which does not appear on the epg (e.g. due to the dropout) is denoted by Q. The probability of allele marked Q is equal to one minus the sum of the probabilities of observed alleles.

Example 1:

$$\mathsf{P}(X=[13,14,15])=\mathsf{P}\left(A\in\left\{[13^2,14,15],\right.\right.$$
$$\left.\left.[13,14^2,15],[13,14,15^2],[13,14,15,Q]\right\}\right)=$$
$$=\frac{4!}{2!1!1!}p_{13}^2p_{14}p_{15}+\frac{4!}{1!2!1!}p_{13}p_{14}^2p_{15}+\frac{4!}{1!1!2!}p_{13}p_{14}p_{15}^2+$$
$$+\frac{4!}{1!1!1!1!}p_{13}p_{14}p_{15}(1-p_{13}-p_{14}-p_{15})=$$
$$=12p_{13}p_{14}p_{15}(2-p_{13}-p_{14}-p_{15}). \quad (4)$$

Example 2:

$$\mathsf{P}(X=[13,14,15])=$$
$$=\mathsf{P}\left(A\in\left\{[13^4,14,15],[13^3,14^2,15],[13^3,14,15^2],\right.\right.$$
$$[13^3,14,15,Q],[13^2,14^3,15],[13^2,14,15^3],$$
$$[13^2,14,15,Q^2],[13^2,14^2,15^2],[13^2,14^2,15,Q],$$
$$[13^2,14,15^2,Q],[13,14^4,15],[13,14^3,15^2],$$
$$[13,14^3,15,Q],[13,14^2,15^3],[13,14^2,15^2,Q],$$
$$[13,14^2,15,Q^2],[13,14,15^4],[13,14,15^3,Q],$$
$$\left.\left.[13,14,15^2,Q^2],[13,14,15,Q^3]\right\}\right)=\ldots=$$
$$=30p_{13}p_{14}p_{15}\left(2-p_{13}-p_{14}-p_{15}\right)\times$$
$$\times\left(p_{13}^2+p_{14}^2+p_{15}^2+p_{13}p_{14}+p_{13}p_{15}+\right.$$
$$\left.+p_{14}p_{15}-2p_{13}-2p_{14}-2p_{15}\right). \quad (5)$$

3 Inclusion of Peak Heights

As can be seen in equation (5), the number of possible allelic vectors and the complexity of their quantification increases very markedly with a higher number of unknown alleles. In fact, the possibility of the peak height inclusion was not employed to the calculation.

Since the peaks of alleles 14 and 15 (470 and 420 RFU) in Example 2 are above the stochastic threshold (300 RFU) and are significantly higher than the third observed value (150 RFU), alleles 14 and 15 can be assumed to be present in two copies. Taking the peak height into account, observed profile X may be adjusted to $X^*=[13,14^2,15^2]$. Quantification is thus considerably simplified:

$$\mathsf{P}(X^*=[13,14^2,15^2])=$$
$$=\mathsf{P}\left(A\in\left\{[13^2,14^2,15^2],[13,14^3,15^2],[13,14^2,15^3],\right.\right.$$
$$\left.\left.[13,14^2,15^2,Q]\right\}\right)=\frac{6!}{2!2!2!}p_{13}^2p_{14}^2p_{15}^2+$$
$$+\frac{6!}{1!3!2!}p_{13}p_{14}^3p_{15}^2+\frac{6!}{1!2!3!}p_{13}p_{14}^2p_{15}^3+$$
$$+\frac{6!}{1!2!2!1!}p_{13}p_{14}^2p_{15}^2(1-p_{13}-p_{14}-p_{15})=$$
$$=30p_{13}p_{14}^2p_{15}^2(6-3p_{13}-4p_{14}-4p_{15}). \quad (6)$$

The model Q is in this part an appropriate extension of the UC model.

4 Probability of Dropout

As was mentioned, the model Q enables to calculate also with possibility of dropout. Due to the small amount of DNA, allelic dropout of one or more alleles is very common in low level samples. Ignoring the possibility of dropout tends to the disfavour of defense [6] so there are some methods to inform about probabilities of dropout ([7], [8]).

However, the model Q includes dropout to the calculation without considering of its probability. We think that this approach is as incorrect as the exclusion of dropout itself and may results in a strong overestimation of calculated probabilities.

Let us suppose that the dropout probability is determined as $d \in (0,1)$. If the probability of allelic vector is calculated considering allelic dropout, this probability should be multiplied by d. For example, the fourth summand in equation (6) must be multiplied by a value of d:

$$\begin{aligned}\mathsf{P}(X^* = [13, 14^2, 15^2]) &= \frac{6!}{2!2!2!}p_{13}^2p_{14}^2p_{15}^2 + \\ &+ \frac{6!}{1!3!2!}p_{13}p_{14}^3p_{15}^2 + \frac{6!}{1!2!3!}p_{13}p_{14}^2p_{15}^3 + \\ &+ d\frac{6!}{1!2!2!1!}p_{13}p_{14}^2p_{15}^2(1-p_{13}-p_{14}-p_{15}) = \\ &= 30p_{13}p_{14}^2p_{15}^2 \times \\ &\times [6d + 3p_{13}(1-2d) + 2(p_{14}+p_{15})(1-3d)]. \quad (7)\end{aligned}$$

The original formula may be obtained by putting the value of $d = 1$ which means that the dropout occurred with the probability equal to 1. However, it would exclude the possibility that the allele is a copy of some of the observed alleles.

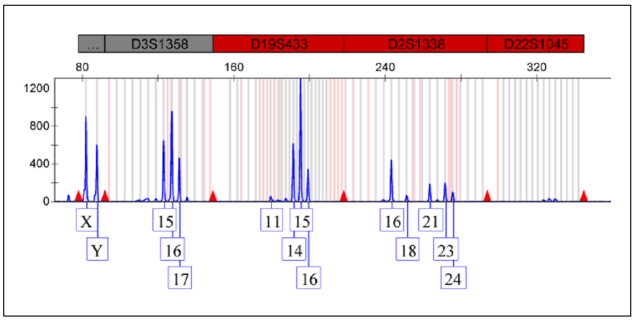

Figure 1: Part of the mixed profile.

If the possibility of two dropouts is assumed, the parameter d must also be considered in the square; if three dropouts are assumed, third power of d is necessary etc. In equation (5), the parameter d should appear in the first, second, and third power. In practice, summands with second and third power have usually an order of magnitude too small to affect the overall probability and could be neglected. See [9] for more complex discussion.

5 Comparison of Models

Figure 1 shows epg of DNA mixture for which three persons are assumed to be contributors. At locus D19S433, four peaks are displayed. Table 1 shows peak heights and allele frequencies in Czech population [10]. There are two suspects with alleles $14, 15$ and $15, 16$. Both calculations are performed independently.

Table 1: Locus D19S433: present alleles and their frequencies in the Czech population.

Allele	Value (RFU)	Frequency
11	55	0.0035
14	610	0.3617
15	1385	0.172
16	391	0.0408

The likelihood ratio is equal to the proportion of probabilities of evidence under prosecution and defense hypotheses:

$$LR = \frac{\mathsf{P}(E|H_p)}{\mathsf{P}(E|H_d)},$$

where H_p means "suspect and two unknown persons contributed to the mixture" and H_d means "three unknown persons contributed to the mixture".

In the following examples we calculate LRs first for the suspect's profile $S_1 = [14, 15]$. Since peak of allele 11 is small, it will be considered later.

5.1 UC Model

Let us evaluate UC model with crime scene profile $X = [14, 15, 16]$ and suspect's profile $S_1 = [14, 15]$.

Hypothesis H_p assumes two persons having together at least one allele 16 and no other than 14, 15 and 16.

$$\begin{aligned}\mathsf{P}(E|H_p) &= \mathsf{P}(X = [14,15,16]|S_1 = [14,15]) = \\ &= 12p_{14}^2p_{15}p_{16} + 12p_{14}p_{15}^2p_{16} + 12p_{14}p_{15}p_{16}^2 + 6p_{14}^2p_{16}^2 + \\ &+ 6p_{15}^2p_{16}^2 + 4p_{14}p_{16}^3 + 4p_{15}p_{16}^3 + 4p_{14}^3p_{16} + \\ &+ 4p_{15}^3p_{16} + p_{16}^4 = 0.0278018\end{aligned}$$

Hypothesis H_d assumes three persons having together alleles 14, 15 and 16 only.

$$\begin{aligned}\mathsf{P}(E|H_d) &= \mathsf{P}(X = [14,15,16]) = \\ &= 30p_{14}p_{15}p_{16}\left(p_{14}^3 + 2p_{14}^2p_{15} + 2p_{14}^2p_{16} + 2p_{14}p_{15}^2 + \right. \\ &+ 3p_{14}p_{15}p_{16} + 2p_{14}p_{16}^2 + p_{15}^3 + 2p_{15}^2p_{16} + \\ &+ \left. 2p_{15}p_{16}^2 + p_{16}^3\right) = 0.01076452\end{aligned}$$

Thus LR for UC model is

$$LR_1 = \frac{\mathsf{P}(E|H_p)}{\mathsf{P}(E|H_d)} = 2.582726. \quad (8)$$

5.2 Original Q Model

If Q model is considered, it may be assumed from analysis of peak heights that allele 15 occurs twice at least. Then the crime scene profile is $X = [14, 15^2, 16]$. The possibility of dropout may be included and let put $p_Q = 1 - p_{14} - p_{15} - p_{16}$.

Hypothesis H_p assumes two persons having together alleles 15 and 16.

$$\begin{aligned}\mathsf{P}(E|H_p) &= \mathsf{P}\left(X=[14,15^2,16]|S_1=[14,15]\right)= \\ &= p_{15}p_{16}\left(4p_{16}^2+6p_{15}p_{16}+12p_{14}p_{16}+12p_{16}p_Q+\right. \\ &\quad +4p_{15}^2+12p_{14}^2+12p_{14}p_{15}+12p_{15}p_Q+ \\ &\quad \left.+\ 12p_Q^2+24p_{14}p_Q\right)=0.0674637\end{aligned}$$

Hypothesis H_d assumes three persons with alleles 14, 15, 15 a 16.

$$\begin{aligned}\mathsf{P}(E|H_d) &= \mathsf{P}\left(X=[14,15^2,16]\right)= \\ &= 30p_{14}p_{15}^2p_{16}\left(2p_{14}^2+2p_{14}p_{15}+3p_{14}p_{16}+6p_{14}p_Q+\right. \\ &\quad +p_{15}^2+2p_{15}p_{16}+4p_{15}p_Q+2p_{16}^2+ \\ &\quad \left.+\ 6p_{16}p_Q+6p_Q^2\right)=0.0377721\end{aligned}$$

LR for original Q model is

$$LR_2=\frac{\mathsf{P}(E|H_p)}{\mathsf{P}(E|H_d)}=1.786072. \tag{9}$$

5.3 Modified Q Model

The process from section 4 is applied. The crime scene profile is $X=[14,15^2,16]$ again and dropout probability is $d=0.45$.

Hypotheses H_p and H_d are the same as in the original Q model, the only change is inclusion of parameter d.

$$\begin{aligned}\mathsf{P}(E|H_p) &= \mathsf{P}\left(X=[14,15^2,16]|S_1=[14,15]\right)= \\ &= p_{15}p_{16}\left(4p_{16}^2+6p_{15}p_{16}+12p_{14}p_{16}+12dp_{16}p_Q+\right. \\ &\quad +4p_{15}^2+12p_{14}^2+12p_{14}p_{15}+12dp_{15}p_Q+ \\ &\quad \left.+\ 12d^2p_Q^2+24dp_{14}p_Q\right)=0.03685446\end{aligned}$$

$$\begin{aligned}\mathsf{P}(E|H_d) &= \mathsf{P}\left(X=[14,15^2,16]\right)= \\ &= 30p_{14}p_{15}^2p_{16}\left(2p_{14}^2+2p_{14}p_{15}+3p_{14}p_{16}+6dp_{14}p_Q+\right. \\ &\quad +p_{15}^2+2p_{15}p_{16}+4dp_{15}p_Q+2p_{16}^2+ \\ &\quad \left.+\ 6dp_{16}p_Q+6d^2p_Q^2\right)=0.01691434\end{aligned}$$

LR for modified Q model is

$$LR_3=\frac{\mathsf{P}(E|H_p)}{\mathsf{P}(E|H_d)}=2.178889. \tag{10}$$

5.4 Modified Q Model with Allele 11

Now, allele 11 is also included to the calculation using modified Q model; it means crime scene profile $X=[11,14,15^2,16]$. Dropout probability is $d=0.45$ again.

Hypothesis H_p assumes two persons with alleles 11, 15 and 16.

$$\begin{aligned}\mathsf{P}(E|H_p) &= \mathsf{P}\left(X=[11,14,15^2,16]|S_1=[14,15]\right)= \\ &= 12p_{11}p_{15}p_{16}\left(p_{11}+2p_{14}+p_{15}+p_{16}+2dp_Q\right)= \\ &= 0.0003889084\end{aligned}$$

Hypothesis H_d assumes three persons with alleles 11, 14, 15, 15 and 16.

$$\begin{aligned}\mathsf{P}(E|H_d) &= \mathsf{P}\left(X=[11,14,15^2,16]\right)= \\ &= 180p_{11}p_{14}p_{15}^2p_{16}\left(p_{11}+p_{14}+p_{15}+p_{16}+2dp_Q\right)= \\ &= 0.0002634395\end{aligned}$$

LR for modified Q model with allele 11 is

$$LR_4=\frac{\mathsf{P}(E|H_p)}{\mathsf{P}(E|H_d)}=1.476272. \tag{11}$$

5.5 Suspect S_2

Calculations for the second suspect $S_2=[15,16]$ are similar. $\mathsf{P}(E|H_d)$ are the same as for first suspect but $\mathsf{P}(E|H_p)$ and hence LRs are much higher:

- LR = 9.929154 for UC model.
- LR = 10.88783 for original Q model.
- LR = 11.58568 for modified Q model.
- LR = 9.904598 for modified Q model with allele 11.

6 Conclusion

Suppose the number of contributors is known and let us briefly summarize the possible statistical processing of epg.

If the number of observed alleles is twice the number of contributors, then all necessary alleles are known and the probability of the profile may be directly calculated. If any alleles are missing in the allelic vector, the procedure from the section 3 is used. The stochastic threshold T is set and the alleles whose peak is above threshold are counted twice. Thereby the set of present alleles is determined more precisely.

If the allelic vector is still incomplete (i.e. the number of alleles $\neq 2n$), all the possibilities of adding any number of alleles present may be calculated. If the possibility of

dropout is also assumed, its probability is predicted and the modified Q model is used as was shown in section 4.

As shown in section 5, substantially different results can be obtained according to the used model and investigated profiles. Generally speaking, the rare alleles present in the profile of the suspect, the higher the likelihood ratio and thus the posterior probability of guilt of the suspect.

When comparing UC and Q model, higher LR was received first and then smaller. On the other hand, it appears that adding of parameter d increases LR because it reduces the denominator more than the numerator.

Acknowledgements

The paper has been supported by the SVV-2012-264 513 project of Charles University in Prague. The authors are grateful to Vlastimil Stenzl for providing the data.

References

[1] Kelly H, Bright J-A, Curran J, Buckleton J. The interpretation of low level DNA mixtures. Forensic Sci Int Genet. 2012; 6: 191–197

[2] Curran JM, Gill P, Bill MR. Interpretation of repeat measurement DNA evidence allowing for multiple contributors and population substructure. Forensic Sci Int. 2005; 148: 47-–53.

[3] Gill P, Brenner CH, Buckleton JS, Carracedo A, Krawczak M, Mayr WR, Morling N, Prinz M, Schneider PM, Weir BS. DNA Commission of the International Society of Forensic Genetics: recommendations on the interpretation of mixtures. Forensic Sci Int. 2006; 160: 90-–101.

[4] Buckleton J, Curran J. A discussion of the merits of random man not excluded and likelihood ratios. Forensic Sci Int Genet. 2008; 2: 343–348

[5] Gill P, Puch-Solis R, Curran J. The low-template-DNA (stochastic) threshold - Its determination relative to risk analysis for national DNA databases. Forensic Sci Int Genet. 2009; 3: 104–111

[6] Balding DJ, Buckleton J. Interpreting low template DNA profiles. Forensic Sci Int Genet. 2009; 4: 1–10

[7] Tvedebrink T, Eriksen PS, Mogensen HS, Morling N. Estimating the probability of allelic drop-out of STR alleles in forensic genetics. Forensic Sci Int Genet. 2009; 3: 222–226

[8] Haned H, Egeland T, Pontier D, Pène L, Gill P. Estimating drop-out probabilities in forensic DNA samples: A simulation approach to evaluate different models. Forensic Sci Int Genet. 2011; 5: 525–531

[9] Gill P, et al. DNA commission of the International Society of Forensic Genetics: Recommendations on the evaluation of STR typing results that may include drop-out and/or drop-in using probabilistic methods. Forensic Sci Int Genet. 2012; in press

[10] Šimková H, Faltus V, Marvan R, Pexa T, Stenzl V, Brouček J, Hořínek A, Mazura I, Zvárová J. Allele frequency data for 17 short tandem repeats in a Czech population sample. Forensic Sci Int Genet. 2009; 4: e15–e17

