Bayesian Analysis (2007) **2**, Number 2, pp. 333–348

A gamma model for DNA mixture analyses

R. G. Cowell,* S. L. Lauritzen† and J. Mortera‡

Abstract. We present a new methodology for analysing forensic identification problems involving DNA mixture traces where several individuals may have contributed to the trace. The model used for identification and separation of DNA mixtures is based on a gamma distribution for peak area values. In this paper we illustrate the gamma model and apply it on several real examples from forensic casework.

Keywords: DNA mixture, forensic, identification, mixture separation, peak area.

1 Introduction

We present a methodology for analysing the important problem of identification and separation or deconvolution of *mixed DNA traces*, where several individuals may have contributed to a DNA sample. Here we illustrate its use for analysing DNA mixtures arising in forensic casework. However, the problem of identifying components of mixtures has potential applications outside this area.

A *mixed DNA profile* is typically obtained from an unidentified biological stain or other trace found at a scene of crime. This occurs in rape cases, in burglaries where objects might have been handled by more than one individual, and also in a scuffle or brawl.

The analysis of DNA mixtures, for calculating the likelihoods of all hypotheses involving a specified set of known and unknown contributors to the mixture, based solely on the qualitative information on which alleles were present in the mixture, is illustrated in Evett et al. (1991) and Weir et al. (1997). In Mortera et al. (2003) Bayesian networks have also been constructed to address some of the challenging problems that arise in the interpretation of mixed trace evidence based on qualitative allele information.

However, the analysis becomes relatively complex when we want to use quantitative peak area values, which contain important additional information about the composition of the mixture. To handle such cases more sophisticated probabilistic modelling is required. This paper is concerned with the quantitative modelling of the peak areas to analyse DNA mixed profiles in order, for example, to infer the genotypes of individuals contributing to the mixture.

A model of peak areas based on conditional Gaussian (CG) distributions was presented in Cowell et al. (2006a), and an implementation of this model in a prototype software tool called MAIES is described in Cowell et al. (2006b). The CG model in

*Sir John Cass Business School, City University, London, http://www.staff.city.ac.uk/~rgc/
†Department of Statistics, University of Oxford, Oxford, http://www.stats.ox.ac.uk/~steffen/
‡Dipartimento di Economia, Università Roma Tre, mailto:mortera@uniroma3.it

these papers is a moment approximation to a more appropriate model based on gamma distributions. Here we present details of the gamma which is also amenable to implementation and exact inference in Bayesian networks (Cowell et al. 2006c).

The outline of the paper is as follows. In § 2 we give some genetic background material and in § 3 we illustrate the model for peak areas. In § 4 we illustrate the model on real forensic casework examples and show how well it predicts the genotypes of the individuals who have contributed to the mixture. Finally, in § 5 we discuss further work required to bring the methodology to a point where it could be applied to the routine analysis of casework.

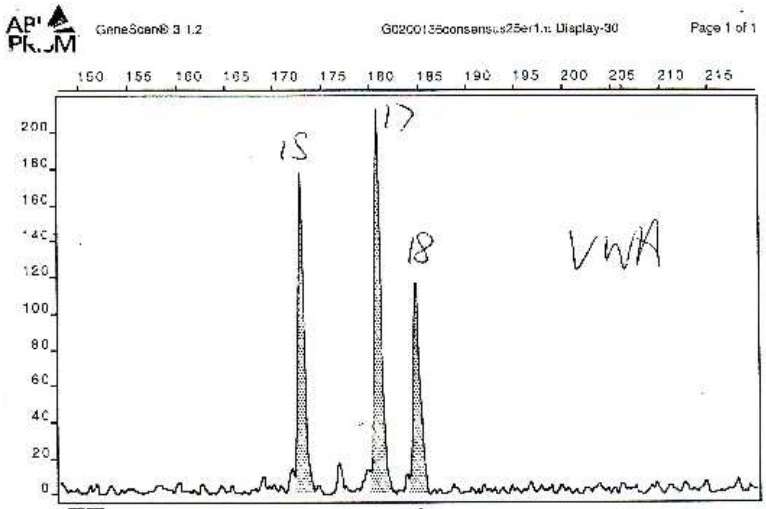
2 Genetic background

We now introduce some basic genetic facts about DNA profiles, for a more detailed explanation see Butler (2005). A gene is a particular sequence of the four *bases*, represented by the letters A, C, G and T. A specific position on a chromosome is called a *locus* (hence there are two genes at any locus of a chromosome pair). A *DNA profile* consists of measurements on a number of *forensic markers*, which are specially selected *loci* on different chromosomes. Current technology uses around 12 – 20 *short tandem repeat* (STR) markers. Each such marker has a finite number (up to around 20) of possible values, or *alleles*, generally positive integers. For example, an allele value of 5 indicates that a certain word (e.g. *CAGGTG*) in the four letter alphabet is repeated exactly 5 times in the DNA sequence at that locus. If a partial repeat sequence is present then the size of the partial repeat is given in bases after a decimal point. For example, the allele named 9.3 consists of 9 repeats and a partial repeat of 3 base pairs.

An individual's *DNA profile* comprises a collection of *genotypes*, one for each marker. Each genotype consists of an unordered pair of alleles, one inherited from the father and one from the mother (though one cannot distinguish which is which). When both alleles are identical the individual is *homozygous* at that marker, and only a single allele value is observed; else the individual is *heterozygous*.

Databases have been gathered from which allele frequency distributions, for various populations, can be estimated for each forensic marker. Here we refer to the gene frequencies reported in Butler *et al.* (2003).

The results of a DNA analysis are usually represented as an *electropherogram* (EPG) measuring responses in *relative fluorescence units* (RFU), see Figure 1. When we are in the presence of a DNA trace where three or more alleles are present on some marker, the trace must have been left by more than one individual. The allele values (repeat numbers) correspond to peaks with a given height and area and are determined by size values along the horizontal axis (see the third column in Figure 1). The band intensity around each allele in the relative fluorescence units represented, for example, through their *peak areas* (column 5 of Figure 1), contains important information about the composition of the mixture. Peak area is measured automatically by the apparatus essentially corresponding to the shaded areas in Figure 1.



Dye/Sample Peak	Minutes	Size	Peak Height		Peak Area	Data Point
18B, 6	55.39	172.83	15	181	1116	2077
18B, 7	57.33	180.75	17	219	1503	2150
18B, 8	58.32	184.76	18	120	869	2187

Figure 1: An electropherogram (EPG) of marker VWA from a mixture. Peaks represent alleles at 15, 17 and 18 and the areas and height of the peaks express the quantities of each. Since the peak around allelic position 17 is the highest this indicates that the 17 allele is likely to be a homozygote or a shared allele between two heterozygotes. This image is supplied through the courtesy of LGC Limited, 2004.

3 The Model

3.1 Notation and setup

We consider I potential contributors to a DNA mixture, assume that M markers are to be used in the analysis of the mixture, and that marker m has A_m allelic types, $m = 1, \ldots, M$.

The model is idealized in that it ignores complicating artifacts such as stutter, drop-out alleles and so on. Now prior to amplification, and provided the mixture sample has not been degraded to the point of breaking up tissue cells, the sample put into the amplification apparatus will consist of an unknown number of cells from each of the I contributors to the mixture. Then, with every cell containing exactly two alleles from each marker, the fraction or *proportion* of cells θ_i, from individual i, is also a common measure across the markers of the amount of DNA from individual i. Let $\theta = (\theta_i, i = 1, \ldots, I)$ denote the vector of these proportions of DNA so that $\theta_i \geq 0$ and $\sum_i \theta_i = 1$. If γ denotes the total amount of DNA in the sample, individual i then

contributes $\gamma\theta_i$.

In an ideal amplification apparatus, the proportion of alleles of each allelic type would be preserved without error. We model departures from this ideal through random variation. The post-amplification proportions of alleles for each marker are represented in the peak area information. For a specific marker m, we define the *peak weight* w_a of an allele a with *repeat number* λ_a by scaling the peak area as

$$w_a = \lambda_a \alpha_a,$$

where α_a is the peak area around allele a. Multiplying the area with the repeat number is a crude way of correcting for the fact that alleles with a high repeat number tend to be less amplified than alleles with a low repeat number. For issues concerning heterozygous imbalance see Clayton and Buckleton (2004).

We further assume that

- The pre-amplification mixture proportion θ is constant across markers, for the reasons outlined above;
- The peak weight for an allele, W_a is approximately proportional to the amount of DNA of allelic type a;
- W_a is the sum of the allele a weights of all contributors.

3.2 The network model

The diagram in Figure 2 indicates the model structure for each marker in the case of two contributors where four distinct alleles are present in the mixture. The central layer of the diagram represents genotypes of the two contributors each of which could originate from a specific individual, here named *suspect* and *victim* respectively, or from an unknown contributor. The central layer also includes the joint genotype of the two contributors. The genotypes of suspect, victim and unknown contributors are represented in the top layer. The nodes in the bottom layer represent expected values of each of four possible peak areas as determined by the genotypes of the contributors and the fraction θ contributed by the the first of these. Similar diagrams represent other markers and these all form elements in an object-oriented Bayesian network used to perform an exact calculation of relevant quantities using efficient local computation algorithms (Lauritzen and Spiegelhalter 1998), thus avoiding the use MCMC methods. For details on this implementation see Cowell et al. (2006c).

3.3 The conditional Gamma Model

Let W_{ia}^m denote the contribution of individual i to the peak weight at allele a of marker m, and let n_{ia}^m denote the number of alleles of type a possessed by individual i. The key distributional premise of our model is to assume that

$$W_{ia}^m \sim \Gamma(\rho_m \gamma \theta_i n_{ia}^m, \eta_m) \tag{1}$$

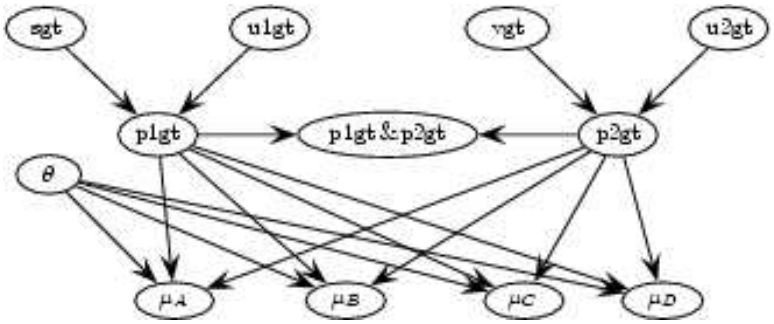


Figure 2: Network model for a specific marker describing a mixture with two contributors.

where $\Gamma(\alpha, \beta)$ denotes the gamma distribution with density

$$f(w) = \frac{\beta^\alpha}{\Gamma(\alpha)} w^{\alpha-1} e^{-\beta w},$$

ρ_m is an amplification factor, and η_m a scale parameter.

The individual contributions W_{ia}^m are unobservable, whereas the measurement process yields observations of the total weights W_{+a}^m, being the sum of the corresponding weights for the contributors

$$W_{+a}^m = \sum_i W_{ia}^m \sim \Gamma\left(\rho_m \sum_i \gamma \theta_i n_{ia}^m, \eta_m\right), \quad m = 1, \ldots, M; a = 1, \ldots, A_m. \tag{2}$$

We further introduce the *weighted allele number*

$$B_a^m = \sum_i \gamma \theta_i n_{ia}^m,$$

which has the property that

$$B_+ = \sum_a B_a^m = \sum_a \sum_i \gamma \theta_i n_{ia}^m = \sum_i \gamma \theta_i \left(\sum_a n_{ia}^m\right) = \sum_i 2\gamma\theta_i = 2\gamma$$

is twice the total amount of DNA and is independent of m. Letting

$$\mu_a^m = \frac{B_a^m}{B_+} = \frac{\sum_i \theta_i n_{ia}^m}{2}$$

we have

$$W_{+a}^m \sim \Gamma(2\rho_m \mu_a^m \gamma, \eta_m), \tag{3}$$

and

$$W^m_{++} = \sum_a W^m_{+a} \sim \Gamma(2\rho_m\gamma, \eta_m), \tag{4}$$

the latter being independent of $\mu^m = (\mu^m_a, a = 1, \ldots, A_m)$. To avoid arbitrariness in scaling, we consider the *relative peak weight* R^m_a, obtained by scaling with the total peak weight as

$$R^m_a = \frac{W^m_{+a}}{W^m_{++}},$$

so that the set of relative peak weights on each marker follows a Dirichlet distribution that does not depend on the scale parameter η_m:

$$R^m = \{R^m_a\} \sim \mathrm{Dir}(2\rho_m\mu^m_a\gamma), \tag{5}$$

having

$$E[R^m_a] = \mu^m_a \tag{6}$$

and

$$V[R^m_a] = \frac{\mu^m_a(1-\mu^m_a)}{2\rho_m\gamma+1} = \sigma^2_m\mu^m_a(1-\mu^m_a), \tag{7}$$

where

$$\sigma^2_m = \frac{1}{2\rho_m\gamma+1}.$$

Conditional on μ^m, the total peak weight W^m_{++} and relative peak weights are independent and independent of everything else in the network model. Thus information about mixture composition from the peak areas about μ^m is contained in the relative weights R^m_a and the corresponding likelihood factorizes as

$$\begin{aligned} L(\mu \,|\, W) &= f(W \,|\, \mu) = L(\mu \,|\, R, W_{++}) \\ &\propto L(\mu \,|\, R) = \prod_a \frac{{r_a}^{2\rho\gamma\mu_a - 1}}{\Gamma(2\rho\gamma\mu_a)} = \prod_a \frac{{r_a}^{\mu_a(\frac{1}{\sigma^2}-1)}}{\Gamma\left(\mu_a(\frac{1}{\sigma^2}-1)\right)}, \end{aligned} \tag{8}$$

where we have suppressed the dependence on the marker m. This factorization now enables efficient computation of quantities of interest using a probabilistic expert system (Cowell et al. 2006c).

4 Application of our model to forensic casework

An analysis of a mixed trace can have different purposes, several of which can be relevant simultaneously, making a unified approach particularly suitable. However, for the sake of exposition we consider the issues separately. The first focus of our analysis will be that of *evidential calculation*, detailed in § 4.1. Here a *suspect* with known genotype is held and we want to determine the likelihood ratio for the hypothesis that the suspect has contributed to the mixture vs. the hypothesis that the contributor is a randomly chosen individual. We distinguish two cases: the other contributor could be (i) a *victim*

with a known genotype; or (ii) a *contaminator* with an unknown genotype possibly without a direct relation to the crime. This could be a laboratory contamination or any other source of contamination from an unknown contributor.

Another use of our network is the *separation of profiles*, i.e. identifying the genotype of each of the possibly unknown contributors to the mixture, the evidential calculation playing a secondary role. This use is illustrated in § 4.2.

In all of the examples considered in this paper we shall use the value $\sigma_m^2 = 0.01$ for all markers. This is the value used in Cowell et al. (2006a) and Cowell et al. (2006b) for conditional Gaussian models, reflecting the variability in peak areas reported in the literature (Gill et al. 1997).

Table 1: *Evett* data showing mixture composition, peak areas and relative weights from a 10:1 mixture of two individuals, with suspect's genotype specified.

Marker	Alleles	Peak area	Relative weight	Suspect
D8	10	6416	0.4347	10
	11	383	0.0285	
	14	5659	0.5368	14
D18	13	38985	0.8871	13
	16	1914	0.0536	
	17	1991	0.0592	
D21	59	1226	0.0525	
	65	1434	0.0676	
	67	8816	0.4284	67
	70	8894	0.4515	70
FGA	21	16099	0.5699	21
	22	10538	0.3908	22
	23	1014	0.0393	
THO1	8	17441	0.4015	8
	9.3	22368	0.5985	9.3
vWA	16	4669	0.4170	16
	17	931	0.0884	
	18	4724	0.4747	18
	19	188	0.0199	

4.1 Evidential calculations

This section illustrates how to calculate the weight of the evidence—in the form of a likelihood ratio, $LR = \Pr(\mathcal{E} \mid H_0)/Pr(\mathcal{E} \mid H_1)$—for a given *suspect* to have contributed to a trace under different circumstances.

The evidence $\mathcal{E}$, could consist of DNA profiles extracted from a *suspect*, s, a *victim*, v, and the mixed trace. In this case we compute the likelihood ratio in favour of the hypothesis that the victim and suspect contributed to the mixture: $H_0 : v\&s$, vs. the

hypothesis that the victim and an unknown individual, u_1 contributed to the mixture: $H_1 : v\&u_1$. A variant has an *unknown contaminator*, u_2 instead of a victim, in which case the hypotheses are $H_0 : u_2\&s$ versus $H_1 : u_1\&u_2$.

Our example is taken from Evett et al. (1998) and has only information of the genotype from one potential contributor, here named the *suspect*, whereas the other unknown contributor is termed *contaminator*. The data refers to a 10:1 mixture of two individuals. The data is displayed in Table 1 and is henceforth referred to as the *Evett* data. Table 2 displays the logarithm of this likelihood ratio together with the corresponding ratio when peak weights are ignored, and the ratios when the mixture proportion θ is assumed known at given values.

Table 2: Logarithm of the likelihood ratios in favour of $H_0 : u_2\&s$ vs. $H_1 : u_1\&u_2$ for the *Evett* data.

			For fixed $\theta =$								
	Alleles	Areas	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\log_{10}$ LR	4.40	8.23	-299.23	-222.68	-145.75	-68.78	5.06	8.26	8.52	8.53	8.53

Note that the strengthening of evidence against the suspect is more dramatic when peak area information is used: the logarithm of the likelihood ratio changes from 4.4 to 8.23, corresponding to an additional factor around 6000. Also the likelihood ratio is essentially constant over a region which completely covers the posterior plausible range $0.85 < \theta < 0.95$.

The posterior distribution of the mixture proportion θ is displayed in Figure 3. The maximum occurs around the value 0.90, which is a little off the true 10:1 mixture proportion.

4.2 Separation calculations

Deconvolution of mixtures or separating a mixed DNA profile into its components has been studied by Perlin and Szabady (2001); Wang et al. (2002); Bill et al. (2005), among others. A mixed DNA profile has been collected and the genotypes of one or more unknown individuals who have contributed to the mixture are desired, for example with the purpose of searching for a potential perpetrator among an existing database of DNA profiles.

For a two-person mixture, the easiest case to consider is clearly that of separation of a single unknown profile, i.e. when the genotype of one of the contributors to the mixture is known. The case when both contributors are unknown is more difficult. In the latter situation this is only possible to a reasonable accuracy when the contributions to the DNA mixture have taken place in quite different proportions.

We concentrate on the problem of separating a mixture into two components, using peak area and allele repeat number information but no information regarding the two contributors to the mixture. We begin by using the *Evett* data again, but this time ig-

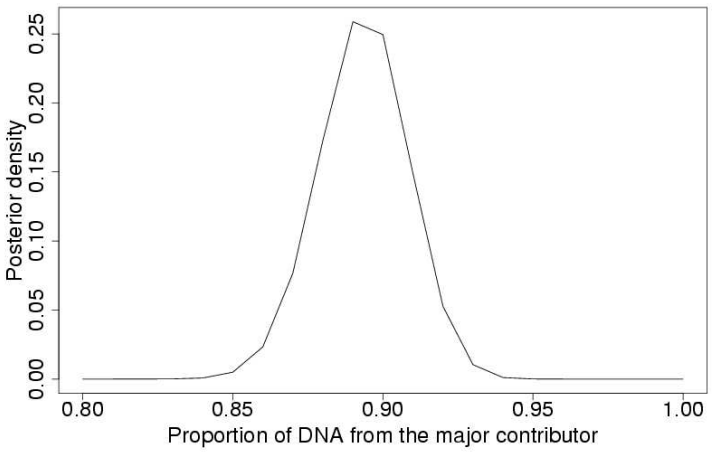


Figure 3: Posterior distribution of the mixture proportion for the *Evett* data using the suspect's genotype.

noring the information on the suspect. The predicted genotypes of the two contributors are shown in Table 3, with the suspect's profile being predicted correctly.

Table 3: Predicted genotypes of both contributors for *Evett* data, with suspect (p1) correct on every marker. The number in brackets is the product of individual marker probabilities.

Marker	Genotype p1	Genotype p2	Probability
D8	10 14	11 14	0.8693
D18	13 13	16 17	1
D21	67 70	59 65	1
FGA	21 22	21 23	0.8819
THO1	8 9.3	9.3 9.3	0.8640
vWA	16 18	17 19	1
joint			0.6720 (0.6705)

Our next example is taken from Appendix B of Clayton et al. (1998) and illustrates use of the *amelogenin* marker in the analysis of DNA mixtures when the individual contributors are of opposite sex. Peak area analysis of the amelogenin marker in DNA indicated an approximate 2:1 ratio for the amount of female to male DNA contributing to the mixture. Peak area information was available on six other markers, the information is shown in Table 4; we shall refer to this as the *Clayton* data.

Table 4: *Clayton* data showing mixture composition, peak areas and relative weights together with the DNA profiles of both victim and suspect. For the marker D21 the allele designation in brackets is as given in Clayton *et al.* (1998) using the Urquhart *et al.* (1994) labelling convention.

Marker	Alleles	Peak area	Relative weight	Suspect	Victim
Amelogenin	X	1277	0.8298	X	XX
	Y	262	0.1702	Y	
D8	13	3234	0.6372		13
	14	752	0.1596	14	
	15	894	0.2032	15	
D18	14	1339	0.1462	14	
	15	1465	0.1714	15	
	16	2895	0.3612		16
	18	2288	0.3212		18
D21	28 (61)	373	0.1719	28	
	30 (65)	590	0.2913		30
	32.2 (70)	615	0.3259		32.2
	36 (77)	356	0.2109	36	
FGA	22	534	0.1547	22	
	23	2792	0.8453	23	23
THO	5	5735	0.2756		5
	7	10769	0.7244	7	7
vWA	15	1247	0.1633	15	
	16	1193	0.1667	16	
	17	2279	0.3383		17
	19	2000	0.3318		19

Table 5: Predicted genotypes of both contributors for *Clayton* data, with victim (p1) and male suspect (p2) correct on every marker. The number in brackets is the product of individual marker probabilities.

Marker	Genotype p1	Genotype p2	Probability
Amelogenin	X X	X Y	0.9979
D8	13 13	14 15	0.9661
D18	16 18	14 15	0.9996
D21	30 32.2	28 36	0.9844
FGA	23 23	22 23	0.9994
THO	5 7	7 7	0.9019
vWA	17 19	15 16	0.9995
joint			0.8797 (0.8785)

In Table 5 we show the results of separating the mixture using peak area information only, without using information on either the suspect or victim. All markers are correctly identified. Figure 4 shows the posterior distribution of the mixture proportion; the peak at around 0.65 corresponds to a mixture ratio of 1.86:1, in line with the approximate 2:1 estimated in Clayton et al. (1998).

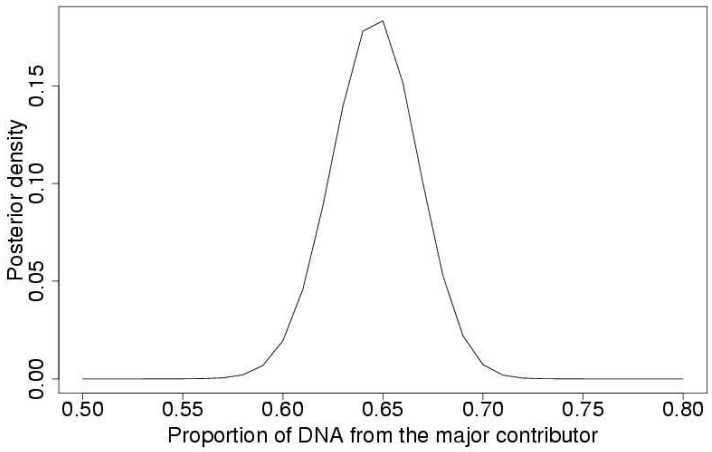


Figure 4: Posterior distribution of mixture proportion from *Clayton* data using no genotypic information.

Our final example uses data from Perlin and Szabady (2001), henceforth referred to as the *Perlin* data, displayed in Table 6. Ignoring information of the known contributors to the mixture, we obtain the posterior distribution for θ shown in Figure 5, with the mode at 0.69 very close to the value reported of 0.7.

The predicted genotypes of the two contributors are shown in Table 7. There is a genotype classification error in marker vWA, however the low marginal predictive probability for the genotype on this marker would indicate that the classification is not reliable; to a lesser extent this is also true for marker D19. The correct allelic combination is the second most probable genotype combination with a posterior probability of 0.3040, which is close to the probability of the most likely combination, 0.3602.

Table 6: *Perlin* data showing mixture composition, peak areas, relative weights, suspect's and victim's genotypes from a 7:3 mixture of two individuals.

Marker	Alleles	Peak area	Relative Weight	Suspect	Victim
D2	16	0.3190	0.1339		16
	18	0.6339	0.2992	18	
	20	0.3713	0.1947		20
	21	0.6758	0.3722	21	
D3	14	1.0365	0.5010	14	14
	15	0.9635	0.4990	15	15
D8	9	0.7279	0.2832	9	
	12	0.2749	0.1426		12
	13	0.6813	0.3829	13	
	14	0.3160	0.1913		14
D16	11	1.4452	0.6801	11	
	13	0.2889	0.1607		13
	14	0.2660	0.1593		14
D18	12	0.3443	0.1504		12
	13	0.6952	0.3290	13	
	14	0.6755	0.3443	14	
	17	0.2850	0.1764		17
D19	12.2	0.6991	0.3109	12.2	
	14	0.6060	0.3092		14
	15	0.6949	0.3799	15	
D21	27	0.2787	0.1289		27
	29	0.7876	0.3913	29	
	30	0.9337	0.4798	30	30
FGA	19	1.0580	0.4621	19	19
	24	0.2830	0.1561		24
	25.2	0.6589	0.3817	25.2	
THO1	6	0.3178	0.1268		6
	7	1.0074	0.4691	7	7
	9	0.6749	0.4041	9	
vWA	17	1.4755	0.7265	17	
	18	0.5245	0.2735		18

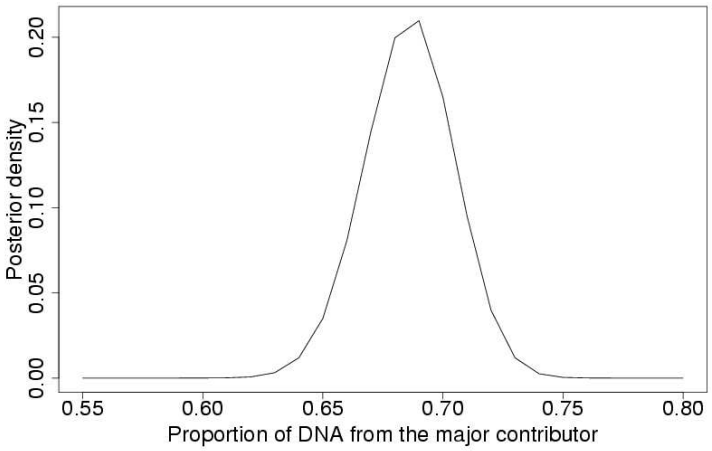


Figure 5: Posterior distribution of mixture proportion from *Perlin* data using no genotypic information.

Table 7: Predicted genotypes of both contributors for *Perlin* data. The number in brackets is the product of individual marker probabilities. There is a classification error in marker vWA (italicized).

Marker	Genotype p1	Genotype p2	Probability
D2	18 21	16 20	0.9995
D3	14 15	14 15	0.9898
D8	9 13	12 14	0.9990
D16	11 11	13 14	0.9998
D18	13 14	12 17	1
D19	12.2 15	14 14	0.7750
D21	29 30	27 30	0.9688
FGA	19 25.2	19 24	0.9970
THO1	7 9	6 7	0.9297
vWA	*17 18*	*17 17*	0.5378
joint			0.3602 (0.3697)

5 Discussion

In the previous sections we have demonstrated how a probabilistic expert system based on a gamma model for the peak areas can be used for analysing DNA mixtures using peak area information, yielding a coherent way of predicting genotypes of unknown contributors and assessing evidence for particular individuals having contributed to the mixture. The gamma model appears to perform well, as can be seen in the examples presented in § 4.1 and § 4.2. The model developed by Cowell et al. (2006a) based on conditional Gaussian distributions gives similar results to those obtained here and this approximative model thus seems satisfactory to use in such relatively simple situations. However, the Gaussian model has some logical difficulties, for example associated with the fact that the relative peak areas necessarily vary in the unit interval and in general may be difficult to approximate with normal distributions. The gamma model we have described does not need any additional *ad hoc* approximations and deals well both with additive composition of peak areas and renormalization. This could be of particular importance when the network model is elaborated to include other artifacts such as silent alleles, dropout and stutter phenomena.

There are further issues which need to be considered which we highlight here. As was the case for the Gaussian model, the variance factors may depend on the marker and on the amount of DNA analysed, but for simplicity we used marker independent values in our analyses. Our model is robust to small changes in these parameter estimates. Our model shows these variance factors to depend critically on the total *amount* of DNA available for analysis, and as this necessarily is varying from case to case, a calibration study should be performed to take this properly into account. However, we find it comforting that the system itself would warn against trusting an uncertain prediction, by yielding an associated low classification probability, as shown for the marker vWA in Table 7.

Methods for diagnostic checking and validation of the model should be developed based upon comparing observed weights to those predicted when genotypes are assumed correct. Such methods could also be useful for calibrating the variance parameter σ^2. Cowell et al. (2006a) use a model which also includes measurement error. This is particularly simple to incorporate in the conditional Gaussian model, but appears to have only a minor influence on the results.

We also need to explore how to extend the model to handle Y-chromosome and mitochondrial DNA haplotype data. As mentioned earlier, we emphasize that for the moment we have not dealt with incorporating artifacts such as stutter, pull-up, allelic dropout, etc., but we hope to pursue these and other aspects in the future. It may be that in incorporating such artifacts our networks will become too complex for exact inference based on evidence propagation in Bayesian networks, and a Monte-Carlo simulation approach may be required.

In a general review of the analysis of DNA evidence, Foreman et al. (2003) include several applications of PES and emphasize their potential by predicting that this methodology "will offer solutions to DNA mixtures and many more complex problems

in the future". We hope that the models presented in this paper shall indeed be part of the solutions anticipated by Foreman *et al.*

References

Bill, M., Gill, P., Curran, J., Clayton, T., Pinchin, R., Healy, M., and Buckleton, J. (2005). "PENDULUM — a guideline - based approach to the interpretation of STR Mixtures." *Forensic Science International*, 148: 181–189.

Butler, J. M. (2005). *Forensic DNA typing*. USA: Elsevier. (Second Edition).

Clayton, T. M. and Buckleton, J. S. (2004). "Mixtures." In Buckleton, J. S., Triggs, C. M., and Walsh, S. J. (eds.), *Forensic DNA Evidence Interpretation*, chapter 7, 217–274. CRC Press.

Clayton, T. M., Whitaker, J. P., Sparkes, R., and Gill, P. (1998). "Analysis and interpretation of mixed forensic stains using DNA STR profiling." *Forensic Science International*, 91: 55–70.

Cowell, R. G., Lauritzen, S. L., and Mortera, J. (2006a). "Identification and separation of DNA mixtures using peak area information." *Forensic Science International*. To appear.

— (2006b). "MAIES: A Tool for DNA Mixture Analysis." In Dechter, R. and Richardson, T. (eds.), *Proceedings of the 22nd Conference on Uncertainty in Artificial Intelligence (UAI)*, 90–97.

— (2006c). "Object-oriented Bayesian networks for DNA mixture analyses." (Manuscript available at `http://www.staff.city.ac.uk/~rgc`).

Evett, I., Gill, P., and Lambert, J. (1998). "Taking account of peak areas when interpreting mixed DNA profiles." *Journal of Forensic Sciences*, 43: 62–69.

Evett, I. W., Buffery, C., Wilcott, G., and Stoney, D. (1991). "A Guide to Interpreting Single Locus Profiles of DNA Mixtures in Forensic Cases." *Journal of the Forensic Science Society*, 31: 41–47.

Foreman, L., Champod, C., Evett, I., Lambert, J., and Pope, S. (2003). "Interpreting DNA Evidence: A Review." *International Statistical Review*, 71(3): 473–495.

Gill, P. D., Sparkes, R., and Kimpton, C. (1997). "Development of Guidelines to Designate Alleles using a STR multiplex system." *Forensic Science International*, 89: 185–197.

Lauritzen, S. L. and Spiegelhalter, D. J. (1998). "Local computations with probabilities on graphical structures and their application to expert systems (with discussion)." *Journal of the Royal Statistical Society, SeriesB*, 50: 157–224.

Mortera, J., Dawid, A. P., and Lauritzen, S. L. (2003). "Probabilistic Expert Systems for DNA Mixture Profiling." *Theoretical Population Biology*, 63: 191–205.

Perlin, M. and Szabady, B. (2001). "Linear mixture analysis: a mathematical approach to resolving mixed DNA samples." *Journal of Forensic Sciences*, 46: 1372–1378.

Urquhart, A., Kimpton, C. P., Downes, T. J., and Gill, P. (1994). "Variation in short tandem repeat sequences – a survey of twelve microsatellite loci for use as forensic identification markers." *International Journal of Legal Medicine*, 107(1): 13–20.

Wang, T., Xue, N., and Wickenheiser, R. (2002). "Least Square Deconvolution (LSD): A New Way of Resolving STR/DNA Mixture Samples." Presentation at the 13th International Symposium on Human Identification, October 7–10, 2002, Phoenix, AZ.

Weir, B. S., Triggs, C. M., Starling, L., Stowell, L. I., Walsh, K. A. J., and Buckleton, J. S. (1997). "Interpreting DNA Mixtures." *Journal of Forensic Sciences*, 42: 213–222.

Acknowledgments

This research was supported by a PRIN05 grant from MIUR. We thank Caryn Saunders for supplying the EPG image used in Figure 1. We also thank the Editor for constructive comments on a previous version of this paper.