



Mixtures with relatives: A pedigree perspective



Thore Egeland^{a,b,*}, Guro Dørum^a, Magnus Dehli Vigeland^c, Nuala A. Sheehan^{d,e}

^aIKBM, Norwegian University of Life Sciences, Ås, Norway

^bNorwegian Institute of Public Health, Oslo, Norway

^cDepartment of Medical Genetics, Oslo University Hospital, Oslo, Norway

^dDepartment of Health Sciences, University of Leicester, UK

^eDepartment of Genetics, University of Leicester, UK

ARTICLE INFO

Article history:

Received 22 June 2013

Received in revised form 13 January 2014

Accepted 22 January 2014

Keywords:

Likelihoods

Pedigrees

Forensics

DNA mixtures

Related contributors

ABSTRACT

DNA mixture evidence pertains to cases where several individuals may have contributed to a biological stain. Statistical methods and software for such problems are available and a large number of cases can be handled adequately. However, one class of mixture problems remains untreated in full generality in the literature, namely when the contributors may be related. Disregarding a plausible close relative of the perpetrator as an alternative contributor (identical twin is the most extreme case) may lead to overestimating the evidence against a suspect. Existing methods only accommodate pairwise relationships such as the case where the suspect and the victim are siblings, for example. In this paper we consider relationships in full generality, conveniently represented by pedigrees. In particular, these pedigrees may involve inbreeding, for instance when the parents of an individual of interest are first cousins. Furthermore our framework handles situations where the opposing parties in a court case (prosecution and defence) propose different family relationships. Consequently, our approach combines classical mixture and kinship problems.

The basic idea of this paper is to formulate the problem in a way that allows for the exploitation of currently available methods and software designed originally for linkage applications. We have developed a freely available R package, *euroMix* based on another package, *paramlink*, and we illustrate the ideas and methods on real and simulated data.

© 2014 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

The broad motivation for this paper is the evaluation of DNA evidence in criminal cases. We deal specifically with *mixture evidence* which refers to cases where several individuals may have contributed to a biological stain recovered from a crime scene. Rape cases present an important example and occur frequently in forensic case work. Typically, the DNA profile based on a vaginal swab will indicate the presence of the victim and one or more men. The forensic scientist can evaluate the likelihood of the DNA data assuming the prosecutor and defence hypotheses, respectively, and report a *likelihood ratio* (LR).

Our calculations are based on qualitative data obtained from discretising the continuous measurements delivered in electrophorograms prior to calculation. This is in line with traditional

approaches and implementations. However, the continuous measurements have been used directly in recent publications [1–3]. Our reasons for taking the qualitative data as a starting point are mainly practical. It seems reasonable to first provide the extensions to standard implementations described below in the conventional setting before considering more ambitious models that utilise more of the information in the data. However, it should be noted that modelling the continuous data comes at a price in that more assumptions are required and implementation becomes less straightforward.

Statistical methods for mixtures have been developed when all contributors are assumed to be unrelated [4]. Most cases can therefore be handled adequately by existing software [5]. The motivation for this paper is cases involving related contributors. Typically, the evidence against a suspect with matching DNA profile will be strong if the alternative is that somebody unrelated to the suspect contributed: intuitively it is very unlikely that a randomly selected person will fit the evidence. However, it is much more likely that a brother of the suspect matches and this must be accounted for as explained in [6]. Fung and Hu have written widely

* Corresponding author at: Norwegian University of Life Sciences, Norway.

Tel.: +47 64966391.

E-mail address: Thore.Egeland@nmbu.no (T. Egeland).

on the situation where some of the contributors to the biological stain could be related (references include [7–10]). However, their approach is limited to specific relationships between two contributors. They argue that consideration of more than two relatives is not very relevant practically, and becomes difficult computationally as kinship coefficients involving more than two individuals would have to be derived. However, there are important cases involving more complicated family relationships and there are also, as we will see, alternatives to calculations based on kinship coefficients.

In this paper we propose a general method for handling mixture problems with related contributors. The relationships may involve any number of individuals and families with inbreeding. Furthermore, contrary to previous methods [9,4] the opposing hypotheses need not agree on the relationships. Hence, the approach presented here is also applicable to forensic identification problems where both mixture data and reference data for certain individuals are available. Combining genotype data at the individual level from reference samples with a trace DNA profile assumed to originate from a single individual was discussed in [11]. Here we discuss the much more general case where the trace could be a mixture from several individuals, some of whom may be related.

We formulate the problem in a way that allows for methods and software designed originally for linkage applications, to be applicable. Thus, well established implementations that are freely available can be used for these applications. General overviews of algorithms for pedigree likelihood calculations such as the Elston–Stewart algorithm (or more generally *peeling* algorithms) [12,13] used in our implementation together with further references, can be found in [14,15]. However, it should be noted that the traditional approach based on the general formulae presented in [16] is likely to be more efficient when contributors to the mixture are unrelated or simply pairwise related as summarised in [4,9].

We will commence with a motivating example to provide the essential background and notation after which we will formulate the problem more generally. Three examples are presented in Section 3, the last one based on a real case. In Section 4 we discuss the assumptions underlying our approach, along with possible extensions and suggestions for future work. Some further examples are presented in the supplementary material and the online documentation.

2. Methods

2.1. Motivating example

It is convenient to introduce the context and essential concepts via an example. The example is specifically constructed to allow for simple calculations. A DNA mixture from three individuals is available and consists of *alleles* denoted 1 and 2 from a *genetic marker*. In addition, reference samples are available from two

undisputed contributors to the mixture: the typed individuals labelled 3 and 4 in Fig. 1.

The genotype and identity of the third contributor is unknown and disputed. For this paper we are assuming *unlinked markers* and *linkage equilibrium* [14] and therefore we only need one marker to explain the concepts since likelihood ratios are multiplied across independent markers.

We have included in our `euroMix` package (described below) a database of frequencies for the 17 markers in the PowerPlex ESX 17 System, plus 6 additional markers. As we will see, this set of markers gives sufficient power to discriminate between the opposing hypotheses in realistic examples. Data like this could occur in a crime context where individuals 3 and 4 could be victims in a murder investigation, for instance, and where the third contributor could be the presumed perpetrator, or murderer. The identity of the perpetrator is disputed, and the suspects are individuals 5 and 6 in the two suggested pedigrees, respectively, shown in Fig. 1. Neither of them has been genotyped. On the other hand, this could be an identification problem as noted in the introduction where the task is to determine the pedigree connecting the typed individuals.

The assessment of the evidence starts by formulating the prosecution (P) and defence (D) hypotheses which in this case are “The contributors are the individuals 3, 4 and 5 related as shown in the left part of Fig. 1” (H_P) and “The contributors are the individuals 3, 4 and 6 related as shown in the right part of Fig. 1” (H_D). The likelihood ratio is

$$LR = \frac{P(R, T|H_P)}{P(R, T|H_D)} = \frac{P(R|T, H_P) P(T|H_P)}{P(R|T, H_D) P(T|H_D)}, \quad (1)$$

where $R = \{1, 2\}$ denotes the set of alleles found in the mixture evidence and $T = \{(3, g_3), (4, g_4)\}$, with $g_3 = 1/1$ and $g_4 = 2/2$ being the known genotypes of individuals 3 and 4. We have tried to stay close to the notation of [17], but some modification is required as genotypes are assigned to individual members of a pedigree. Thus, each element of T above is a pair (i, g_i) , where g_i is the genotype of a contributing individual i . It is commonly assumed [9,17] that $P(T|H_P) = P(T|H_D)$ using the argument that information on typed individuals will not vary between the prosecution and defense hypotheses. While this seems reasonable, we note that it is only true if the two hypotheses dictate the same family relationships between the typed individuals. In this case the likelihood ratio in (1) simplifies to

$$LR = \frac{P(R|T, H_P)}{P(R|T, H_D)}. \quad (2)$$

In our simple example, it can be seen from the two pedigrees in Fig. 1 that the mentioned assumption does not hold. This means that the simplification in (2) does not apply and we must work with the general expression for the likelihood ratio in (1). In [9]

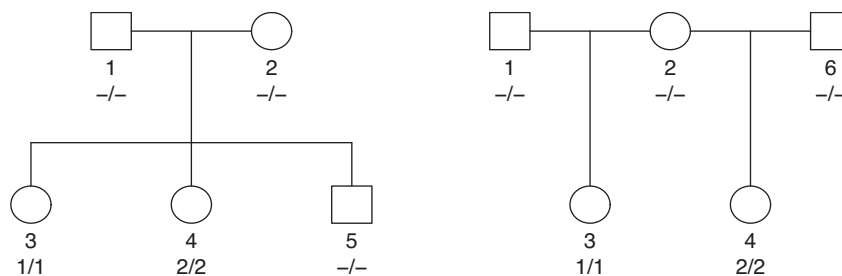


Fig. 1. The pedigrees representing the prosecution (left) and defence (right) hypotheses of the motivating example. The victims 3 and 4 are known contributors to the mixture $R = \{1, 2\}$. The opposing hypotheses disagree on whether the victims are sisters or half sisters, and whether the third, untyped, contributor is their brother (5) or the father (6) of one of them.

Download English Version:

<https://daneshyari.com/en/article/6554133>

Download Persian Version:

<https://daneshyari.com/article/6554133>

[Daneshyari.com](https://daneshyari.com)