European DNA Profiling (EDNAP) Meeting

Athens, Greece. 27 May 1995

Chairman: Peter D Martin

List of members attending is attached.

REPORT OF THE MEETING

The meeting was unusual in that Ian Evett and David Balding were specially invited to make presentations on the use of statistics applied to the various aspects of DNA analysis as used in crime case investigation and the presentation of evidence to the courts. The following are synopses of the talks.

1. A Comparison of Bayesian and Frequentist Approaches

Ian Evett, FSS Birmingham

The frequentist view of probability is that it is objective. It does not depend on human perception and is defined as the extension of a long run relative frequency. The Bayesian view of probability is that it is subjective. It is a measure of a person's degree of belief and different people will have different probabilities for the same proposition.

The frequentist view does not enable us to address the probability of a proposition; we have to use indirect methods of inference based only on data. The Bayesian view requires us to address the probability of a proposition. We use other information to form a prior probability and this is modified by the data to give a posterior probability.

The frequentist view provides no means for combining the evidence whereas the Bayesian view provides a framework for combining all kinds of evidence. Many frequentist statisticians agree that Bayesian inference provides a useful means for understanding evidence in legal proceedings. There are, however, several confusing issues of DNA statistics which have arisen from frequentist views.

Although it is usually assumed that the frequencies of separate alleles can only be multiplied if they are independent, in reality it only matters if the dependence effects have a serious effect on the answer. There is a major difficulty in that it is impossible to prove independence. Just because the test has failed to detect dependence effects it does not mean that they do not exist. In carrying out experiments to find whether it is possible to ignore the consequences of making an assumption of independence it is usual to be told that the estimation procedures must be conservative. In using such an approach the consequences are that evidence will be diluted in every case where the police have arrested the right man. In

99.99% of cases where the police have arrested the wrong man, the STR analysis will eliminate the suspect and the conservative approach has no effect. The forensic scientist has no mandate to distort the evidence in either direction and debate cannot be eliminated by being conservative.

2. Population Genetic Effects on DNA Evidence

David Balding, QMW College London

When assessing DNA evidence for court purposes the uncertainty can come from:

a small database

a non-random database

shared ancestry

It is the shared ancestry which has the greatest effect. In a crime case the chance of finding two people with the same DNA profile will be affected by the population in question. The frequency of particular alleles will be affected by relatives and ethnicity. It has been demonstrated how the frequencies of blood groups within the UK can vary due to isolation and migration patterns. Shared ancestry could have the effect of one or two orders of magnitude.

The uncertainty in shared ancestry becomes important because, in a casework situation we are considering the hypothesis that the suspect (whose DNA profile matches that from the crime scene) was not the donor of that particular stain. Therefore there are two people who have the same DNA profile.

It is necessary to know how much shared ancestry there is in Europe where pockets of in-breeding have been demonstrated. Excess homozygosity and subpopulation frequencies are sensitive to:

preferential mating

nature of the database*

null alleles

*The database is, typically, national but the suspect and the true perpetrator might come from the same small community.

3. Analysis of Quadruplex data Ian Evett

(The following is a short account of the talk as the full account will be published in the American Journal of Human Genetics)

The analysis of the data obtained using the quadruplex of loci: THO1, VWA, F13 and FES.

Caucasian samples were obtained from four different sources:

FSS	423
MPFSL	257
Derbyshire police	582
Strathclyde	140 (later increased to 400)

For pragmatic reasons most of the analysis was based on traditional significance testing. The analysis of the second generation multiplex will be via Bayesian methods.

The tests used were: Homozygosity, Exact test (Zaykin et al), Number of matches and Likelihood Ratio distribution.

The exact test (Zaykin et al. Genetica 1995) is a very powerful test and was used to test within and between locus independence (failure to show independence does not indicate serious dependence effects as absolute independence does not exist). In the first exact test using separate databases, low p-values were obtained from the MPFSL data in five of the higher order tests and they all involved F13A1. A low p-value was also obtained for the VWA/THO1 data from the FSS. All other p-values were greater than 0.05. In the second exact test using a combined database, there were some apparent dependence effects from VWA/THO1 and F13A1/FES but on further testing it has been demonstrated that 'errors' have either no practical significance or are conservative.

The report concludes that:

> The variations observed between the four Caucasian frequency distributions have no practical significance.
>
> Large scale application of independence tests inevitably leads to failures.
>
> The failures observed in this analysis have no practical forensic significance.
>
> A casework procedure involving multiplication within and between loci has been shown to be robust.

4. Variance between subpopulations David Balding

Artificially generated databases were used to determine the variance which might exist between the component subpopulation databases.

The conclusions reached were:

> For the European populations considered (Northern Ireland, Dundee, general UK Caucasian, Italy, Greece and Greek Cypriot), there was variation in the F value from one locus to another, but values were often under 1% and usually under 2%.
>
> All of the required data is still not available. There is a need to determine the level of differentiation in more isolated European populations.
>
> If a gross miscalculation is made e.g. using a Pakistani database to calculate a frequency for Chinese, there could be a serious miscalculation of the frequency (unless a very large F value is used).

5. Computerised calculation of Essen-Möller Index applied to identification cases.

Thore Egeland, Oslo.

A synopsis of this presentation is attached at Annex 1

6. Standardisation of the STR system ACTBP2 (SE33) using a sequenced ladder.

Anke Möller, Münster.

So far over 50 different ACTBP2 (SE33) alleles have been identified. Three different sequence variants, depending on fragment lengths, have been found as follows:

202-259 bp = 4bp repeat unit
265-309 bp = 4bp repeat unit + hexanucleotide unit
311-323 bp = 4bp repeat unit + 2 hexanucloetide units

The Münster laboratory has prepared an allelic ladder which consists of 23 alleles. The final selection was made from 90 alleles, which had been sequenced, to provide a uniform ladder for use in a denaturing gel. Detection of the alleles is by fluorescence tagging.

An inter-laboratory exercise was organised for 8 participating laboratories who were asked to specify their practical conditions. These included the type of sequencer, the internal standard, the electrophoretic and amplification conditions. All participants used denaturing gels and, with the exception of one laboratory, fluorescent detection was via an ABI sequencer. The other laboratory used an ALF sequencer. The organisers specified which of the two primers was dye-labelled.

Two unknown samples were provided together with the cell line K562. These were typed with reference to:

1. Internal standard (relative bp)

2. SE33 ladder (allelic designation)

In order to study intra-gel reproducibility participants were asked to type the ladder five times on the same plate.

For the 8 laboratories 5 different internal standards were used:

3 laboratories used GS2500 (ROX)
2 laboratories used GS 350 (ROX)
1 laboratory used GS 500 (ROX)
1 laboratory used GS2500 (TAMRA)
1 laboratory used an ALF internal standard

Results

Sample 1 = 17:27 relative to the allelic ladder
Sample 2 = 15:30 relative to the allelic ladder
K562 = 26:28 relative to the allelic ladder

All laboratories correctly typed all samples.

If samples are assessed relative to the internal standards, the results are highly variable with a SD between 1% and 2% (the value of the SD depends on the allele: the larger the allele, the larger the SD).

Inter-laboratory SD is 1% to 2% but the intra-gel SD values were below 0.5%. The maximum differences between laboratories (7bp with mean = 4.5bp) show that it is necessary to estimate the allele size relative to the allelic ladder and not to the internal standard.

Results from both the ABI and ALF sequencers showed the correct assignment of alleles.

Conclusions

1. There is considerable inter-laboratory variation of results when alleles are assigned with reference to an internal standard.

2. The ladder alleles were correctly assigned by all laboratories using the nomenclature proposed by the Münster laboratory.

3. All laboratories obtained the correct results.

4. It is essential to use a uniform ladder and an agreed nomenclature in combination with the relevant gel system.

7. **Population data - independent inheritance of STR loci**

Hermann Schmitter, BKA Wiesbaden

A synopsis of this presentation is attached at Annex 2.

8. **UK National DNA Database**

David Werrett, FSS Birmingham

A synopsis of this report is attached at Annex 3.

9. As there is still a desire to find further linking systems which could be used for inter-laboratory co-operation, Peter Gill suggested that members should inform others of any STRs which they consider could be used. A questionnaire has n been compiled and is attached at Annex 4.

Could you please complete this questionnaire and return it to Peter Gill at the FSS, Birmingham. This will be an agenda item at the EDNAP session of the ISFH meeting in September.

10. Sincere thanks were given to Chris Konialis for providing such an agreeable venue, and to David Balding and Ian Evett for providing such worthwhile and informative presentations.

11. Participants agreed that Dr Scheitauer at the Institut für Rechtsmedizin, Innsbruck should be invited to join EDNAP as the representative for Austria. Peter Martin was asked to communicate the wishes of the meeting.

12. A list of participants is attached at Annex 5.