Best Practice Manual for DNA Pattern Recognition and Comparison

ENFSI-BPM-DNA-01
Version 01 - November 2015

With the financial support of the Prevention of and Fight against Crime Programme
European Commission - Directorate -General Home Affairs
Background

This Best Practice Manual (BPM) belongs to a series of 10 BPMs issued by the European Network of Forensic Science Institutes (ENFSI) in November 2015. The series covers the following forensic disciplines:

1. Forensic Examination of Digital Technology
2. Forensic Examination of Handwriting
3. Chemographic Methods in Gunshot Residue Analysis
4. Road Accident Reconstruction
5. Microscopic Examination and Comparison of Human and Animal Hair
6. Fingerprint Examination
7. DNA Pattern Recognition and Comparison
8. Application of Molecular Methods for the Forensic Examination of Non-Human Biological Traces
9. Forensic Recovery, Identification and Analysis of Explosives Traces
10. Forensic Investigation of Fire Scenes which have resulted in Fatalities*
11. Forensic Investigation of Fire Scenes which involve the Clandestine Manufacture of Improvised or Homemade Explosive Devices*
12. Forensic Investigation of Fire Scenes which Involve the Clandestine Manufacture of Illicit Synthetic Drugs*

* The three specific areas on Forensic Investigation of Fire Scenes (numbers 10 -12) were combined into one BPM ‘Investigation of Fire Scenes’.

In the years 2014 and 2015, so-called Activity Teams have - in parallel - developed the 10 BPMs. The activities were performed within the project ‘Towards European Forensic Standardisation through Best Practice Manuals (TEFSBPM)’ and co-ordinated by the ENFSI Quality and Competence Committee. The realisation of the BPMs was supported by the Prevention of and Fight against Crime Programme of the European Commission – Directorate General Home Affairs (code: PROJECT HOME/2012/ISEC/MO/4000004278). The core project concept was that the BPMs will enhance the quality of the forensic services available to law enforcement and justice across Europe and thereby encourage forensic standardisation and cross-border cooperation between countries.

ENFSI expects that the issuing of this series will stimulate the improvement of already existing BPMs as well as the creation of new BPMs on disciplines that are not covered yet.

Official language
The text may be translated into other languages as required. The English language version remains the definitive version.

Copyright
The copyright of this text is held by ENFSI. The text may not be copied for resale.

Further information
For further information about this publication, contact the ENFSI Secretariat. Please check the website of ENFSI (www.enfsi.eu) for update information.
Best Practice Manual for DNA Pattern Recognition and Comparison

CONTENTS

1. AIMS .......................................................... 4
2 SCOPE .......................................................... 4
3 TERMS AND DEFINITIONS ............................................ 5
4 RESOURCES .......................................................... 6
  4.1 Personnel ......................................................... 6
  4.2 Equipment ....................................................... 7
  4.3 Reference materials ........................................... 7
  4.4 Accommodation and environmental conditions ............ 7
  4.5 Materials and reagents ......................................... 7
5 METHODS .......................................................... 7
  5.1 Acquisition of data ............................................. 7
  5.2 Generation of composite and/or consensus profiles in case of replicates .................. 8
  5.3 Quality assessment ............................................ 8
  5.4 Comparison of genotype data .................................. 9
5.5 Peer review ..................................................... 10
6 VALIDATION AND ESTIMATION OF UNCERTAINTY OF MEASUREMENT .......... 10
  6.1 Validation ....................................................... 10
  6.2 Estimation of uncertainty of measurement .................. 11
7 PROFICIENCY TESTING ........................................... 11
8 INITIAL ASSESSMENT .............................................. 11
9 PRIORITISATION AND SEQUENCE OF EXAMINATIONS .......................... 11
10 RECONSTRUCTION OF EVENTS ................................... 11
11 HANDLING ITEMS .................................................. 11
12 EVALUATION AND INTERPRETATION .............................. 12
13 PRESENTATION OF EVIDENCE .................................. 12
14 HEALTH AND SAFETY ............................................. 13
15 REFERENCES .......................................................... 13
16 AMENDMENTS AGAINST PREVIOUS VERSION .......................... 14

Acknowledgements
U. Neuhaus-Steinmetz1, S. Dornseifer1, B. Haldemann1, T. Heylen2, C. Aelbrecht2, O. Bleka3, H.J. Larsen4, S. Willuweit4, A. Alonso5, J. M. Teodoridis7, J. Morzfeld3, L. Zatkalikova8, M. Krupsky9, B. Berger9, W. Parson9, N. Morling4 and P. Gill3,10, are thanked for their contributions to the realisation of this BPM.

1Institute of Forensic Science, Landeskriminalamt, Berlin, Germany;
2National Institute of Criminalistics and Criminology, Brussels, Belgium;
3Norwegian Institute of Public Health, Oslo, Norway;
4Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark;
5Department of Forensic Genetics, Institute of Legal Medicine and Forensic Sciences, Charité - Universitätsmedizin, Berlin, Germany;
6National Institute of Toxicology and Forensic Science, Madrid, Spain;
7Institute of Forensic Science, Bundeskriminalamt, Wiesbaden, Germany;
8Institute of Forensic Science, Bratislava, Slovakia;
9Institute of Legal Medicine, Medical University of Innsbruck, Austria;
10University of Oslo, Norway
1 AIMS

This Best Practice Manual (BPM) aims at providing a framework of procedures, quality principles, training processes and approaches to the expert evaluation of forensic DNA data. This BPM can be used by member laboratories of ENFSI and other forensic science laboratories to establish and maintain working practices in the field of forensic DNA examination that will deliver reliable results, maximise the quality of the information obtained and produce robust DNA evidence. The use of consistent methodology and the production of more comparable results will facilitate interchange of data between laboratories.

The term Best Practice Manual is used to reflect the scientifically accepted practices at the time of creating. The term BPM does not imply that the practices laid out in this manual are the only good practices used in the forensic field. In this series of ENFSI Practice Manuals the term BPM has been maintained for reasons of continuity and recognition.

2 SCOPE

This BPM is aimed at experts in the field and assumes prior knowledge in the discipline. It is not a standard operating procedure (SOP) and addresses the requirements of the judicial systems in general terms only.

The BPM provides optimised workflows that can be followed either manually or can serve as specification for expert analysis software. It covers parts of the analytical process downstream of the work at crime scenes and in ‘wet labs’. The scope of this BPM comprises the examination of nuclear DNA profiles based on quality parameters, DNA pattern recognition, comparison of DNA profiles and the evaluation of results (Compare Fig. 1).

This BPM neither covers evaluation of electropherograms (i.e. peak detection and allele designation), nor report writing, nor searches in DNA databases.
Fig. 1: Scope of the BPM for DNA Pattern Recognition and Comparison in relation to the forensic DNA typing workflow.

3 TERMS AND DEFINITIONS

For the purposes of this Best Practice Manual (BPM), the relevant terms and definitions are given in ENFSI documents, the ILAC G19 "Modules in Forensic Science Process", as in standards like ISO 9000, ISO 17000 and 17020 apply. For forensic DNA analysis, the relevant terms and definitions are given in “Appendix 1: Glossary of Terms” of [1], as well as in [2-6]. Further field-specific definitions are provided below to assure precise, unambiguous language throughout the BPM to guarantee a unified understanding.

**Allele match**: defined as each allele that is present in one profile and in a compared profile.

**Allele mismatch**: defined as each allele that is present in one profile but is absent in a compared profile.
Complete profile: for known genotypes: all alleles of the known genotype(s) were detected at all investigated loci; for unknown genotypes: alleles are present for all loci analysed.

Composite profile: containing all detected alleles of replicates, i.e. present in at least one replication.

Consensus profile: containing only reproduced alleles from two or more replicate analyses

DNA profile: the observed alleles at an investigated set of chromosomal loci. It ideally reflects the genotype(s) of the DNA contributor(s).

Drop-in: may explain alleles in the DNA profile that are additional to the assumed main contributor(s), i.e. artefacts or contaminations.

Drop-out: failure to detect certain alleles (false negatives). Failure to detect any alleles present at a particular locus can be referred to as locus drop-out.

Forensic DNA expert: a person trained and experienced in forensic DNA typing and may function as expert witness in a court of law. Qualification requirements and performed tasks may differ among laboratories and/or legislations.

Genotype: an individual’s genetic mark-up (compare DNA profile).

Major component: in multiple source profiles that contain alleles of sufficient different peak heights, alleles with the largest peak heights per locus are assigned to a major component. They can be extracted and, e.g. treated as a distinct profile representing one or more persons. Compare P. Gill et al. [7].

Multiple source profile: a DNA profile that consists of several genotypes, alternatively referred to as “mixture”.

Partial profile: not all alleles of the known genotype(s) were detected at the investigated loci.

Reference profile: a DNA profile of a known individual. The donor might be a suspect, a(n) (assumed) victim, or any other potential contributor (e.g. crime scene investigators, witnesses, other attendees). See also stain DNA profile as counterpart.

Single source profile: a DNA profile that allegedly consists of one genotype only.

Stain DNA profile: a DNA profile which originates from detectable biological stains or from touch (i.e. trace) DNA. See also reference profile as counterpart.

Target: profile of interest used during comparison of DNA profiles. The alleles present in the target define the maximum number of possible allele matches.

4 RESOURCES

4.1 Personnel
This BPM is aimed at experts in the field of forensic DNA analysis and assumes prior knowledge in the discipline. It has to be assured that the experts’ education is in accordance with national rules and regulations. The process of forensic DNA analysis may be divided into steps performed by different staff categories. A forensic DNA expert performing evaluation and interpretation should at least possess the following competences in terms of scientific knowledge in the fields:

- Interpretation of preliminary tests (for blood, semen, saliva etc.)
- DNA purification, extraction, quantification and amplification
- STR fragment analysis
- Analysis using automated genotyping software
- Evaluation of DNA profiles and their comparison, see [8]

For further information on requirements, see Chapter 5, “General Competencies” of [9].

Forensic DNA experts expected to present written and oral testimony should have received instruction and / or mentoring in the procedural requirements of the particular criminal justice system in which the evidence is to be presented. Compare [9].
4.1.1 Training and assessment
For training and assessment of personnel please see [9].

4.2 Equipment
The following equipment is recommended:
- Computer system including appropriate screen(s) and colour printer(s)
- Software for table and text processing, e.g. office software
- Software for generation, evaluation, and visualisation of electropherograms (EPGs)
- Software for mixture interpretation
- Statistical software tools, e.g. for likelihood calculations

4.3 Reference materials
- Allele frequency data sets of case-relevant populations (autosomal see e.g. ENFSI DNA WG STR Population Database [10], mtDNA see EMPOP [11] and Y-chromosomal see YHRD [12])
- Elimination database (containing DNA profiles of potential contaminators)
- DNA profiles from standard reference materials, proficiency testing exercises, etc.

4.4 Accommodation and environmental conditions
Does not apply, since this BPM does not cover crime-scene-related or laboratory activities.

4.5 Materials and reagents
- Stain characterisation data, e.g. from tests for blood, semen, saliva, etc. (optional)
- Quantification data, i.e. real-time PCR quantitative and qualitative DNA data (optional)
- EPGs or electronically stored data files as provided by a genotyping software

5 METHODS

DNA pattern recognition and comparison workflow can be structured into the following steps:
1. Acquisition of data
2. Generation of composite and/or consensus profiles in case of replicates
3. Quality assessment, i.e. classification of profiles according to completeness, peak heights, amount of DNA input, presence of PCR inhibitors, DNA degradation index and minimum number of contributors
4. Comparison of genotype data
   - With respect to single source profiles
   - With respect to multiple source profiles
5. Evaluation and verification of results, covered in Section 12

In the following, the beforementioned steps 1 to 4 are described in more detail:

5.1 Acquisition of data
DNA data can be obtained from one or more replicates and acquired from various sources, e.g.
- Genotype analysis software (i.e. tables containing assigned alleles and peak heights)
- Quantitative real-time PCR data
- Handwritten or electronic data sheets
- LIMS
- National DNA databases
and can include analytical (experimental) values and results (see Section 5.1.1) as well as additional data (i.e. metadata as listed in Section 5.1.2).
5.1.1 Analytically-oriented data
- Assigned alleles
- Peak heights or areas
- DNA fragment lengths
- DNA quantification data, i.e. concentration and volume information of total DNA and male DNA
- DNA quality data, e.g. regarding degradation, presence of inhibitors
- PCR kit(s) and parameters, e.g. cycle number, amount of template DNA
- Electrophoresis parameters, e.g. injection time, voltage
- Parameters for fragment analyses, e.g. analytical threshold and stutter filters

5.1.2 Additional data
In order to distinguish profiles and differentiate between cases, metadata have to be joined with DNA data. Below, some classes of metadata are suggested:
- Identifier(s), unambiguously linking a profile to a criminal case, a contributor, lab batch processes etc.
- Specimen category, e.g. “reference”, “stain”, “touch / trace DNA” (classifications have to be defined by the laboratory and can therefore vary in detail between different laboratories)
- DNA extraction method, e.g. differential lysis
- Source-level data, i.e. tissue, fluid or cell type of origin
- Additional data for report generation, e.g. description of DNA sample, name of analyst(s), dates and times of individual work steps etc.
- Case circumstances may be needed for evaluation of DNA comparison results
Profiles submitted by external laboratories or acquired through national DNA databases normally lack these details. They are usually provided as an allele table together with a unique identifier but without any experimental or evaluative details.

5.2 Generation of composite and/or consensus profiles in case of replicates
Replicates are commonly displayed in allele tables, either as composite profiles or as consensus profiles. For definition of the terms see Section 2.

5.3 Quality assessment
Quality assessment criteria are established by each laboratory. The following scheme is provided as a guideline but is by no means compulsory:

5.3.1 DNA quantity and quality
Real-time PCR data evaluation allows classification of DNA profiles according to DNA input, DNA degradation and presence of inhibitors.

5.3.2 Locus drop-out
There are circumstances in which a profile is not “complete” (occurrence of locus drop-out, i.e. investigated loci without any detected alleles present). Reasons for locus drop-outs could be for instance DNA degradation, PCR inhibition, primer site mutations and/or low template DNA. Therefore, a profile with alleles at all investigated loci along with appropriate peak heights (above lab threshold) is an indicator for a sufficient amount of non-degraded DNA for PCR analysis. For the evaluation of matches it has to be taken into account that in an incomplete profile the number alleles matching a reference is reduced.
The number of locus drop-outs could be used as a qualitative description in report writing.
5.3.3 Contributors
A multiple source profile originates from at least two contributors. An indicator for the detection of multiple source profiles is the presence of more than two alleles at one or more loci. The minimal number of potential contributors to the profiles might be assessed according to current recommendations (see publication list below). However, the occurrence of multiple alleles at one locus (the number should be designated by the lab), as typically observed in complex mixtures, will hinder this assessment. Further reading: Gill et al. [7].
For a collection of current and recommended publications, see Butler [13], Chapter 6, “Reading List and Internet Resources”, “Estimating the Number of Contributors”.

5.3.4 Contamination
Profiles may be checked (according to national rules and regulations) for contamination which may arise via:
1. Cross contamination between samples
2. DNA from lab personal or other persons involved in the investigation.
For details please see [3, 14].

5.3.5 Reference profiles
The quality requirements of reference profiles have to be defined and monitored by the lab. External reference profiles are generally assumed to be qualified. The origin has to be stated.

5.4 Comparison of genotype data
The objective is to compare case-related stain DNA profiles with each other, with profiles from other cases and with persons, i.e. reference profiles.

5.4.1 Types of DNA profile comparison
1. Comparison of reference profiles to single source profiles
2. Comparison of single source profiles
3. Comparison of reference profiles to multiple source profiles
4. Comparison of single source profiles to multiple source profiles

In general, one profile (in the following called the "target") is compared to several profiles (the compared profiles). The succeeding steps should be determined by the lab, a typical example is given:
In the first step, reference profiles of known individuals can serve as the target. In the second step, targets are recruited from single source profiles that have not yet been assigned to a reference profile. These single source profiles are generally referred to as “unknown individuals”. For the last step, remaining multiple source profiles with major or dominant contributors serve as the target, if not assigned to an individual (known or unknown) during the preceding steps.

5.4.2 Types of comparison
a) High stringency ("perfect match"): all the alleles of the target profile are present in the compared profile and vice versa. High stringency is used in comparison types 1 and 2 (as defined in the section above).
b) Moderate stringency: if the target profile is a reference sample and the compared profile represents a mixture, then all the alleles of the target profile should be present in the mixture. Alleles in the compared profile that are not present in the target profile (unexplained alleles) are allowed.
c) Low stringency: At least one allele is shared between the target and the compared profiles. This stringency is used to find parent-child-relationships in criminal investigations (where allowed according to national rules and regulations). Only loci with alleles in both profiles, i.e. in the target and compared profile, are considered for comparison. This may occur, for instance, in case the comparison involves DNA profiles obtained with PCR kits which have different numbers of loci.

Ideally the number of matching alleles is identical to the number of alleles present in the target and the number of mismatches is zero. However, to avoid false negatives, a defined number of mismatches might be allowed.

The reasons for false negatives can be allele drop-outs (e.g. degradation or problems with primer concordance in different STR PCR kits). Potential profile matches should be evaluated according to current guidelines (Compare [7]).

5.4.3 Outcome of DNA profile comparison

The result of a comparison between a target profile and one or more compared profiles is a list of candidate matches according to the stringency level(s) applied. The list should be ranked according to the statistical weight of the match (as described in [8]; see also Section 12). The number of compared loci and alleles (total, matching & mismatching) should be given. Where applicable, these numbers should be provided for both the composite and the consensus profile of the compared DNA sample. For visualisation, all alleles within one locus can be assigned a colour according to match stringency at the locus level.

5.5 Peer review

Peer review is the evaluation of work by one or more persons of similar competence to the producers of the work (peers). It constitutes a form of self-regulation by qualified members of a profession within the relevant field. Peer review methods are employed to maintain standards of quality, improve performance, and provide credibility.

To ensure high quality and to avoid errors, transfer of data (e.g. DNA profiles and identifiers), evaluation and interpretation should be peer reviewed. Usually, this is accomplished by ‘internal peer review’ via qualified colleagues within the reporting scientists’ institution. Therefore, appropriate documentation and peer review is mandatory for quality control. Furthermore, statistical evaluation has to be reviewed. Special emphasis should be put on critical findings that have an impact on the case.

6 VALIDATION AND ESTIMATION OF UNCERTAINTY OF MEASUREMENT

6.1 Validation

6.1.1 Validation of new methods

The laboratory should only use properly validated methods for the expert evaluation of forensic DNA data, e.g. following ISO 17025 [15] or equivalent standards. For more details on ENFSI validation requirements see [16].

Requirements on the validation of software used in forensic context may be taken from the following sources: [17-21]. The Scientific Working Group on Digital Evidence (SWGDE) provide recommendations on purpose and scope of testing, requirements to be tested, test scenario selection, test data and documentation of test data [22].
6.2 Estimation of uncertainty of measurement
Uncertainty cannot be estimated but minimised by paying attention to key issues, such as:
- Competence level of staff
- Allele frequencies/population data
- Transfer of data
- Profile compiling (e.g. false assignment of alleles to composite or consensus profile)
- Comparison of profiles: false or missing matches
- Formulation of useful hypotheses
- STR analysis
- Source and integrity of the sample, including the possibility of mixing or inverting the samples
- Potential relationships between contributors
- Evaluation of potential contamination

Software that assists with profile comparison and other error prone tasks can reduce error rates, and is therefore recommended.

7 PROFICIENCY TESTING

There is no special proficiency test covering the whole process addressed with this BPM. Existing proficiency tests partially cover the process, e.g. GEDNAP and the GEP proficiency testing program [23,24].

Proficiency tests should be used to test and assure the quality of DNA pattern recognition and comparison. A list of currently available PT/CE schemes as put together by the QCC is available at the ENFSI Secretariat. "Guidance on the conduct of proficiency tests and collaborative exercises within ENFSI" [25] provides information for the ENFSI Expert Working Groups (EWGs) on how to organise effective proficiency tests (PTs) and collaborative exercises (CEs) for their members.

8 INITIAL ASSESSMENT

Source and integrity of a sample have to be assessed and mentioned in the report.

If there is an evaluation to be made (i.e. a comparison between a DNA sample profile and a reference profile), further information on the case is needed. For details, see [8].

9 PRIORITISATION AND SEQUENCE OF EXAMINATIONS

This sub-section does not apply, since this BPM does not cover crime-scene-related or laboratory activities.

10 RECONSTRUCTION OF EVENTS

This section does not apply since it is beyond the scope of this BPM.

11 HANDLING ITEMS

This sub-section does not apply, since this BPM does not cover crime-scene-related or laboratory activities.
12 EVALUATION AND INTERPRETATION

In the evaluation phase, the expert assesses the weight of a match and a non-match, considering the possibility of drop-ins and drop-outs. The expert might rank and sort out matches for a comprehensive presentation in reports or in court.

According to national quality standards, sample DNA profiles have to be evaluated whether they are qualified for entry into the national DNA database. Considerations regarding evaluation, especially statistical approaches that can and should be used during decision making, are described in the [8]. This guideline provides basic guidance on evaluation and interpretation where, among other resources, a detailed example of a DNA case is given.

There are statistical frameworks for further assignment of probabilities to drop-ins and drop-outs (see ISFG website [26] for listing of software).

Additionally, it is recommended that each lab sets up a framework for evaluation and interpretation considering the following points:

- **Classification**
  - Samples (reference vs case-related DNA samples)
  - DNA profiles (single source vs mixtures, major vs minor components, full vs partial profiles etc.)

- **Rules for allele designation**

- **Pattern recognition in mixtures and/or low template DNA**

- **Statistical inference**

- **Available relevant databases (DBs, see also Section 4.3)**
  - Contamination/Elimination DBs
  - Estimated allelic proportions (also commonly termed allele frequencies)
  - Quality cut-off for storage in national DNA DBs

The risks described in Section 6.2 (the estimation of uncertainty of measurement) have to be taken into account especially during evaluation and interpretation.

13 PRESENTATION OF EVIDENCE

The overriding duty of those providing expert testimony is to the court and to the administration of justice. As such, evidence should be provided with honesty, integrity, objectivity and impartiality. Evidence can be presented to the court either orally or in writing. Only information which is supported by the examinations carried out should be presented. Presentation of evidence should clearly state the results of any evaluation and interpretation of the examination.

Written reports should include all the relevant information in a clear, concise, structured and unambiguous manner as required by the relevant legal process. Written reports should be reviewed. Expert witnesses should resist responding to questions that take them outside their field of expertise unless specifically directed by the court, and even then a declaration as to the limitations of their expertise should be made.

DNA examination results could be presented as tables and/or explained in text form. For inclusion of statistical weight of a match please see [8].
14 HEALTH AND SAFETY

National and institutional regulations for occupational safety and health in the office apply.

15 REFERENCES

25. QCC-PT-001. Guidance on the conduct of proficiency tests and collaborative exercises within ENFSI", version 001, 27/06/2014

16 AMENDMENTS AGAINST PREVIOUS VERSION

Not applicable (first version)
Best Practice Manual for DNA Pattern Recognition and Comparison
ENFSI-BPM-DNA-01
Version 01 - November 2015