



BSI Standards Publication

Information technology — Biometric data interchange formats

Part 14: DNA data

National foreword

This British Standard is the UK implementation of ISO/IEC 19794-14:2022+A1:2025. It supersedes BS ISO/IEC 19794-14:2022, which is withdrawn.

The start and finish of text introduced or altered by amendment is indicated in the text by tags. Tags indicating changes to ISO/IEC text carry the number of the ISO/IEC amendment. For example, text altered by ISO/IEC amendment 1 is indicated by A1 A1.

The UK participation in its preparation was entrusted to Technical Committee IST/44/-/3, Biometric data interchange formats.

A list of organizations represented on this committee can be obtained on request to its committee manager.

Contractual and legal considerations

This publication has been prepared in good faith, however no representation, warranty, assurance or undertaking (express or implied) is or will be made, and no responsibility or liability is or will be accepted by BSI in relation to the adequacy, accuracy, completeness or reasonableness of this publication. All and any such responsibility and liability is expressly disclaimed to the full extent permitted by the law.

This publication is provided as is, and is to be used at the recipient’s own risk.

The recipient is advised to consider seeking professional guidance with respect to its use of this publication.

This publication is not intended to constitute a contract. Users are responsible for its correct application.

© The British Standards Institution 2025
Published by BSI Standards Limited 2025

ISBN 978 0 539 26043 4

ICS 35.240.15

Compliance with a British Standard cannot confer immunity from legal obligations.

This British Standard was published under the authority of the Standards Policy and Strategy Committee on 31 October 2022.

Amendments/corrigenda issued since publication

Date	Text affected
30 June 2025	Implementation of ISO/IEC amendment 1:2025

INTERNATIONAL
STANDARD

ISO/IEC
19794-14
+A1

Second edition
2022-10

Amendment 1
2025-03

Information technology — Biometric data interchange formats —

Part 14: DNA data

*Technologies de l'information — Formats d'échange de données
biométriques —*

Partie 14: Données ADN



Reference number
ISO/IEC 19794-14:2022(E)

AVA Documents
Document Sharing Hub
© ISO/IEC 2022



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2022

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword	iv
Introduction	v
1 Scope	1
2 Normative references	1
3 Terms and definitions	2
3.1 Terms related to basic DNA concepts	2
3.2 Terms related to DNA profiling	4
3.3 Terms related to DNA databases	5
3.4 Terms related to DNA profile comparison and interpretation of results	6
4 Abbreviated terms	8
5 Conformance	9
6 DNA data format specification	9
6.1 Overview	9
6.2 Data conventions	10
6.2.1 Unknown field value	10
6.2.2 DNA data handlings	11
6.3 Content of the DNA XML schema	15
6.3.1 Overview	15
6.3.2 General header	15
6.3.3 Representations	18
6.3.4 Pedigrees	35
7 Registered format type identifier	39
Annex A (normative) DNA XML schema definition and sample encoding	40
Annex B (normative) Conformance testing methodology	59
Annex C (informative) DNA kit identifiers	67
Annex D (informative) DNA loci	73
Annex E (informative) Kinship interoperability tests — Pedigree test cases	79
Annex F (informative) Additional interoperability tests	110
Annex G (informative) DNA loci for identification purposes	135
Bibliography	137

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives or www.iec.ch/members_experts/refdocs).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents) or the IEC list of patent declarations received (see <https://patents.iec.ch>).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html. In the IEC, see www.iec.ch/understanding-standards.

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 37, *Biometrics*.

This second edition cancels and replaces the first edition (ISO/IEC 19794-14:2013), which has been technically revised. It also incorporates the Amendment ISO/IEC 19794-14:2013/Amd. 1:2016.

The main changes are as follows:

- [Clause 6](#) and [Annex A](#) have been technically revised to enable the standardized interchange of DNA profile search results;
- [Annex B](#) has been technically revised to reflect the revised data interchange format;
- New [Annexes E](#), [F](#) and [G](#) have been added.

A list of all parts in the ISO/IEC 19794 series can be found on the ISO and IEC websites.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html and www.iec.ch/national-committees.

Introduction

Forensic molecular genetics has evolved from a rapidly developing field with changing technologies into a highly recognized and generally accepted forensic science. Forensic genetics using deoxyribonucleic acid (DNA) profiling comprises a number of important applications. Examples are the investigation of biological stains to obtain evidence for the presence of an alleged perpetrator at a crime scene by comparing the genetic profiles from crime scene samples of human origin, to those available at DNA databases administered by law enforcement agencies. These also include the identification of unknown corpses in the context of both natural death and crime, immigration, paternity testing and disaster victim identification (DVI).

This document is based on DNA data from forensic DNA typing techniques that are commonly used, namely short tandem repeat (STR) profiling and other DNA typing techniques that are standardized by scientific bodies for the purpose of discriminating between individuals.

The purpose of this data interchange format is to enable the exchange of DNA data from different systems, not to impose any constraints on the specific DNA typing system/technique to be used. Where existing DNA data exchange formats have been referenced in the preparation of this document, these formats are listed as references.

Standard profiling systems exploit the non-coding parts of DNA that are referred to as “junk DNA”. The coding regions, which are richer in information pertaining to specific genetic traits of an individual, are deliberately avoided in order to maintain the privacy and civil rights of the donor. In addition, national data protection and privacy legislation can impose special security safeguards, such as (but not limited to) encryption of data transfers and/or storage.

This document supports XML (Extensible Markup Language) encoding, to support a spectrum of user requirements. [Annex A](#) specifies the schema against which XML-encoded DNA data XML documents are required to validate. It also contains a sample DNA data XML document. [Annex B](#) addresses the conformance testing methodology. [Annex C](#) lists some examples of DNA analysis kits. [Annex D](#) lists the names of DNA loci. [Annex E](#) lists interoperability test data for kinship searching in the form of pedigrees. In [Annex F](#), there is a description of interoperability tests at Level 3 (semantics). By means of the sample inclusion and comparison rules listed in [Annex G](#), a target can be identified among a number of candidates.

Information technology — Biometric data interchange formats —

Part 14: DNA data

1 Scope

This document specifies a data interchange format for the exchange of deoxyribonucleic acid (DNA) data for person identification or verification technologies that utilize human DNA. Consideration of laboratory procedures is out of scope of this document.

This document provides the ability for DNA profile data to be exchanged and used for comparison (subject to privacy regulations) with DNA profile data produced by any other system that is based on a compatible DNA profiling technique and where the data format conforms to this document.

This document is intended to cover current forensic DNA profiling or typing techniques that are based on short tandem repeats (STRs), including STRs on the X chromosome (X-STRs) the Y chromosome (Y-STRs), as well as mitochondrial DNA. A single DNA profile for a subject can contain data resulting from more than one of these different DNA techniques. This document enables data from multiple DNA techniques to be presented in a single DNA profile for a given subject.

This document has been prepared in light of ongoing efforts to reduce human involvement in the processing (enrolment and comparison) of DNA. In anticipation of the data format requirements for automated DNA techniques, this document describes a format for both processed and raw (electrophoretic) DNA data. A normative XML schema definition (XSD) is provided in [Clause A.1](#) for the syntax of DNA data XML documents. In [Clause A.2](#), there is a sample DNA data XML document.

This document is not intended for any other purposes than exchange of DNA for biometric verification and identification of individuals. In particular, it is not intended for the exchange of medical and other health-related information.

This document also specifies elements of conformance testing methodology, test assertions and test procedures as applicable to this document. It establishes test assertions pertaining to the structure of the DNA data format (Type A Level 1 as defined in ISO/IEC 19794-1:2011/Amd. 1:2013) and test assertions pertaining to internal consistency of the values contained within each field (Type A, ind Level 2 as defined in ISO/IEC 19794-1:2011/Amd. 1:2013). This document also specifies test assertions pertaining to the content of DNA data XML documents (Level 3 as defined in ISO/IEC 19794-1:2011/Amd. 1:2013). The successful completion of Level 1 and Level 2 is a prerequisite for carrying out the tests at Level 3.

The conformance testing methodology specified in this document does not establish:

- tests of other characteristics of biometric products or other types of testing of biometric products (e.g. acceptance, performance, robustness, security);
- tests of systems not claimed to conform to the requirements of this document.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 2382-37, *Information technology — Vocabulary — Part 37: Biometrics*

ISO 3166-1, *Codes for the representation of names of countries and their subdivisions — Part 1: Country code*

ISO 3166-2, *Codes for the representation of names of countries and their subdivisions — Part 2: Country subdivision code*

ISO/IEC 19794-1:2011, *Information technology — Biometric data interchange formats — Part 1: Framework*

ISO/IEC 19794-1:2011/Amd. 1:2013, *Information technology — Biometric data interchange formats — Part 1: Framework — Amendment 1: Conformance testing methodology*

ISO/IEC 19794-1:2011/Amd. 2:2015, *Information technology — Biometric data interchange formats — Part 1: Framework — Amendment 2: Framework for XML encoding*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 2382-37 and ISO/IEC 19794-1 and the following apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

- ISO Online browsing platform: available at <https://www.iso.org/obp>
- IEC Electropedia: available at <https://www.electropedia.org/>

3.1 Terms related to basic DNA concepts

3.1.1

deoxyribonucleic acid

DNA

complex molecule found in virtually every cell in the body that carries the genetic information from one generation to another

3.1.2

chromosome

structure within the cell that bears the genetic material as a linear strand of DNA

Note 1 to entry: In humans, each cell normally contains 23 pairs of chromosomes, for a total of 46. 22 of these pairs, called autosomes, look the same in both males and females. The 23rd pair, the sex chromosomes, differs between males and females. Sex chromosomes in males are different in size and are called X and Y. Sex chromosomes in females are identical in size and both are called X.

3.1.3

Y chromosome

organized structure of the DNA molecule containing male-specific DNA only

3.1.4

non-coding part of DNA

chromosome regions not genetically expressed, i.e. not known to provide for any functional properties of an organism

3.1.5

locus

unique physical location on the DNA molecule

Note 1 to entry: The plural of locus is loci.

3.1.6**allele**

member of two or more alternative forms of a DNA sequence found at a particular locus

3.1.7**tri-allelic pattern**

locus that shows an occasional detection of three alleles in single-source samples

Note 1 to entry: Tri-allelic patterns can show unbalanced peak heights (Type I: The sum of heights of two of the peaks is equal to the third) or balanced peak heights (Type II: The peaks of the three alleles are of a similar height).

3.1.8**chimera**

<genetic> individual having two different sets of DNAs with the code to make two separate individuals

Note 1 to entry: Otherwise said, this is a single individual composed of cells with more than one distinct genotype.

3.1.9**homozygote**

individual having the same (or indistinguishable) alleles at a particular locus due to the inheritance of the same allele from each parent

Note 1 to entry: A heterozygote is an individual having two different alleles at a particular locus.

3.1.10**short tandem repeat****STR**

short sequence of DNA that is repeated numerous times in direct succession

Note 1 to entry: The number of repeated units can vary widely between individuals and this high level of variation makes STRs particularly useful for discriminating between individuals.

Note 2 to entry: STR analysis is one of the most useful methods in forensic genetics for comparing specific loci on DNA from two or more samples.

3.1.11**autosomal STR****aSTR**

STR region found only in autosomal chromosomes in the nucleus of the cell

3.1.12**X-STR**

STR region found in female-specific DNA on the X chromosome only

3.1.13**Y-STR**

STR region found in male-specific DNA on the Y chromosome only

Note 1 to entry: Y-STR can be used to trace paternal lineages as it is male specific and only inherited from fathers to their sons.

3.1.14**mitochondrial DNA****mtDNA**

small circular DNA molecules located in structures used to provide energy to the cell (mitochondria)

Note 1 to entry: Mitochondria often are called the powerhouse of the cell. Their small size and abundant nature make them particularly useful when examining small or much-damaged biological material.

Note 2 to entry: The mitochondria, and thus mitochondrial DNA, are passed only from mother to offspring through the egg cell. It can be used to trace maternal lineages as it is only inherited from one's mother.

3.2 Terms related to DNA profiling

3.2.1

DNA profiling

DNA typing

technique used by scientists to discriminate between individuals by examining variations in their DNA

3.2.2

allelic ladder

artificial mixture of the common alleles present in the human population for a particular STR marker that is used during a DNA profiling process (capillary electrophoresis) in parallel with the sample of interest for accurate allele call determination

3.2.3

electropherogram

graphic representation of results of a DNA profiling process (capillary electrophoresis) with the X axis displaying the observed alleles and the Y axis recording the relative amount of DNA detected based on the relative fluorescent unit collected during analysis

Note 1 to entry: Electropherograms can be transmitted as image files if this is needed from partner DNA laboratories for validation of DNA profiles.

3.2.4

DNA profile

set of alphanumeric values describing the molecular structure at a group of loci identified in an individual's DNA

Note 1 to entry: A DNA profile is referred to as DNA fingerprint, DNA type or genetic fingerprint in other documents.

3.2.5

forensic DNA profile

DNA profile that represents a set of identification characteristics from non-coding parts of an analysed human DNA sample

3.2.6

mixed stain

biological stain that contains body fluids or tissues from more than one individual

EXAMPLE Contaminated sample, DNA sample taken from a swabbing of a surface of a drinking vessel or cigarette that has been shared

3.2.7

mixed DNA profile

DNA profile generated from a mixed stain

Note 1 to entry: In many cases where a sample consists of a stain or body fluid deposits of multiple individuals, the mixed DNA cannot be isolated when the sample is acquired.

Note 2 to entry: Where the profile of one or more of the mixed DNA sample contributors is known, the mixture can be separated into its contributing DNA profiles. One of the processes is called mixture deconvolution. This involves analysing the mixture DNA profile and exploiting the probabilistic and genetic hereditary properties of DNA to separate the profiles.

3.2.8

fully designated locus

locus of which all positions are reliably typed

Note 1 to entry: The locus status of a fully designated locus is "Normal".

3.2.9**partial locus**

locus at which not all the alleles show up

3.2.10**partial DNA profile**

DNA profile with partial loci or in which not all the loci targeted show up

EXAMPLE If 13 loci were targeted and only 9 could be reported, that would be termed a partial DNA profile.

Note 1 to entry: A DNA profile can be partial at the profile level, partial at the locus level or both.

3.2.11**DNA mobile processing unit**

fully-functional DNA laboratory that is mobile

3.2.12**rapid DNA instrument**

self-contained device that carries out a fully-automated DNA analysis of a DNA sample

3.3 Terms related to DNA databases**3.3.1****Interpol DNA Database**

central forensic DNA database to which all Interpol member states can submit forensic DNA profiles of unsolved crimes, criminals, missing persons or unknown human remains through their National Interpol Bureaus, both with classic DNA profile storage or search requests and through online DNA profile data transfers from their national DNA databases for automated searching

Note 1 to entry: Interpol runs also a separate Missing Persons DNA database (I-Familia) using family DNA comparison to identify unknown human remains.

3.3.2**Interpol Standard Set of Loci**

ISSOL

set of STR loci defined by the Interpol DNA Monitoring Expert Group, which recommends for use as common DNA loci for forensic DNA analyses in all forensic DNA kits and with minimum loading criteria to input a profile in the Interpol DNA Database to enable worldwide comparability of STR profiles and thus uniform crime fighting worldwide by usage of forensic DNA technology

3.3.3**Interpol DNA Monitoring Expert Group**

advisory board with senior experts from Interpol member states for creation of recommendations on the use of DNA in criminal and missing person investigations including creation of Interpol DNA profile interchange standards and forms as well as rules for the Interpol DNA Database

3.3.4**Prüm DNA Database Network**

decentralized database network system originally developed by some EU member states, in which biometric data, such as forensic DNA profiles, can be compared online and in real time with DNA profile search queries between the Prüm partner states

Note 1 to entry: The Prüm network has not only been implemented in a legally binding manner by all EU member states through EU legal acts but has also been extended through bilateral and multilateral state agreements to become a globally functioning Prüm data network system for biometric online data exchange (e.g. Western Balkan states).

3.3.5

European Standard Set of loci

ESS

set of STR loci defined by the ENFSI DNA Working Group which is recommended for use as minimum and common DNA loci for forensic DNA analyses in all forensic DNA kits and with minimum loading criteria to input a profile in the Prüm DNA Database Network to enable European comparability of STR profiles

3.3.6

ENFSI DNA Working Group

working group that supports the aims and objectives of ENFSI in the area of DNA casework analysis including definition of quality and STR loci standards for possible international forensic DNA cooperation

3.3.7

request

message containing one or more DNA profiles to be searched or stored or updated in or removed from a DNA profile database

3.3.8

response

message containing one or more answers depending on request message

Note 1 to entry: Match results, non-match results, error messages, notification of storage or deletion or update.

3.4 Terms related to DNA profile comparison and interpretation of results

3.4.1

power of discrimination

potential power of a genetic marker or set of markers to differentiate between any two individuals chosen at random

3.4.2

reference DNA profile

DNA profile of an identified person

3.4.3

target DNA profile

DNA profile contained in a request for comparison against a DNA profile database

3.4.4

exact match

outcome of a DNA search engine when all allele values of the compared loci are the same in two DNA profiles

3.4.5

rare allele value

allele value present in low frequency at a specific population and, therefore, much more significant than other alleles for identification purposes

3.4.6

wildcard

symbol substituting a rare allele value at a locus and matching any value at the corresponding locus in a DNA profile

Note 1 to entry: An asterisk is commonly used as a wildcard.

Note 2 to entry: Two different patterns can match in a wildcard search.

3.4.7**microvariant**

allele containing an incomplete repeat unit or appearing to have values beyond a specified range

Note 1 to entry: Many STR markers are composed of a specific sequence of four nucleotides (called nucleus, core or repeat unit). The sequence of nucleotides is repeated in tandem a number of times which varies. When one of the repeat units is incomplete (e.g. shows three nucleotides instead of four), the allele is called a microvariant.

3.4.8**mismatch**

outcome of a DNA search engine when only one difference, which involves a wildcard or a microvariant, is found in a comparison of two DNA profiles

3.4.9**near match**

outcome of a DNA search engine when only one of all allele values of the compared loci is different in two DNA profiles

3.4.10**match**

outcome of a DNA search engine that is either an exact match, a near match or a mismatch

3.4.11**non-match**

outcome of a DNA search engine other than exact match, near match or mismatch

3.4.12**match quality**

level of agreement between two DNA profiles

EXAMPLE The following match quality levels can be distinguished:

- Q1: exact match
- Q2: near match (only one potential difference involving a wildcard)
- Q3: near match (only one difference, which involves a microvariant)
- Q4: mismatch (only one difference other than wildcards or microvariants)

Note 1 to entry: Some DNA search engines use a likelihood ratio to quantify the match quality.

3.4.13**match count**

number of identical loci found in comparison of two DNA profiles

3.4.14**adventitious match**

match that happens by chance instead of having the same source or being linked by kinship

Note 1 to entry: In the case of DNA testing, not having enough distinguished characteristics (e.g. due to a partial DNA profile) can lead to adventitious matches. DNA search engine matches therefore always need forensic verification/validation for possible detection of adventitious matches.

3.4.15**candidate**

DNA profile found in a DNA profile database satisfying the defined matching criteria against the target DNA profile

3.4.16

hit

candidate confirmed by a DNA examiner

Note 1 to entry: A "no-hit" is a candidate rebutted by a DNA examiner, for example, detected adventitious match.

Note 2 to entry: Validation is required to be carried out in line with forensic quality management requirements (e.g. accreditation standards).

4 Abbreviated terms

AABB	American Association of Blood Banks
BDB	biometric data block
BIR	biometric information record
CBEFF	Common Biometric Exchange Formats Framework
CE	capillary electrophoresis
CODIS	Combined DNA Index System
CRS	Cambridge Reference Sequence
DLR	DNA loci reference
DVI	disaster victim identification
ENFSI	European Network of Forensic Science Institutes
FSA	fragment sequence analysis
GLP	Good Laboratory Practice
GPS	global positioning system
HV	hypervariable regions of mitochondrial DNA
ILAC	International Laboratory Accreditation Cooperation
ISFG	International Society of Forensic Genetics
IUPAC	International Union of Pure and Applied Chemistry
IUT	implementation under test
ICS	implementation conformance statement
NA	not available
NGS	next-generation sequencing
NIST	National Institute of Standards and Technology
ORI	originating agency identifier
PCR	polymerase chain reaction
POC	point of contact

QA	quality assurance
rCRS	revised Cambridge Reference Sequence
RDBMS	relational database management system
SNP	single-nucleotide polymorphism
SQL	Structured Query Language
UTC	Coordinated Universal Time
WGS	World Geodetic System
XML	Extensible Markup Language
XSD	XML schema definition

5 Conformance

An XML document conforms to this document if it satisfies the format requirements with respect to its structure, relations among its fields and relations between its fields and the underlying input that are specified within [Clause 6](#) and [Clause A.1](#).

Biometric data interchange format conformance tests conform to this document if they satisfy all the normative requirements set forth in [Annex B](#).

Implementations are not required to conform to all possible aspects of this document, but only to those that are claimed to be supported by the implementation in an implementation conformance statement (ICS), filled out in accordance with ISO/IEC 19794-1:2011/Amd. 1:2013 and [Table B.1](#) of this document.

6 DNA data format specification

6.1 Overview

XML documents encoding DNA data shall validate against the XML schema definition in [Clause A.1](#). In conformance to ISO/IEC 19794-1, a DNA data XML document may or may not be embedded in an appropriate CBEFF (Common Biometric Exchange Formats Framework) compliant biometric information record (BIR).

There are two kinds of fields (in XML also known as elements): simple and combined. A simple field contains only one simple data object, and a combined field contains one or more fields that can be simple or combined. Simple and combined fields are implemented by the XML mechanisms “simple type” and “complex type”, respectively.

The structure of a DNA data XML document is depicted in [Figure 1](#).

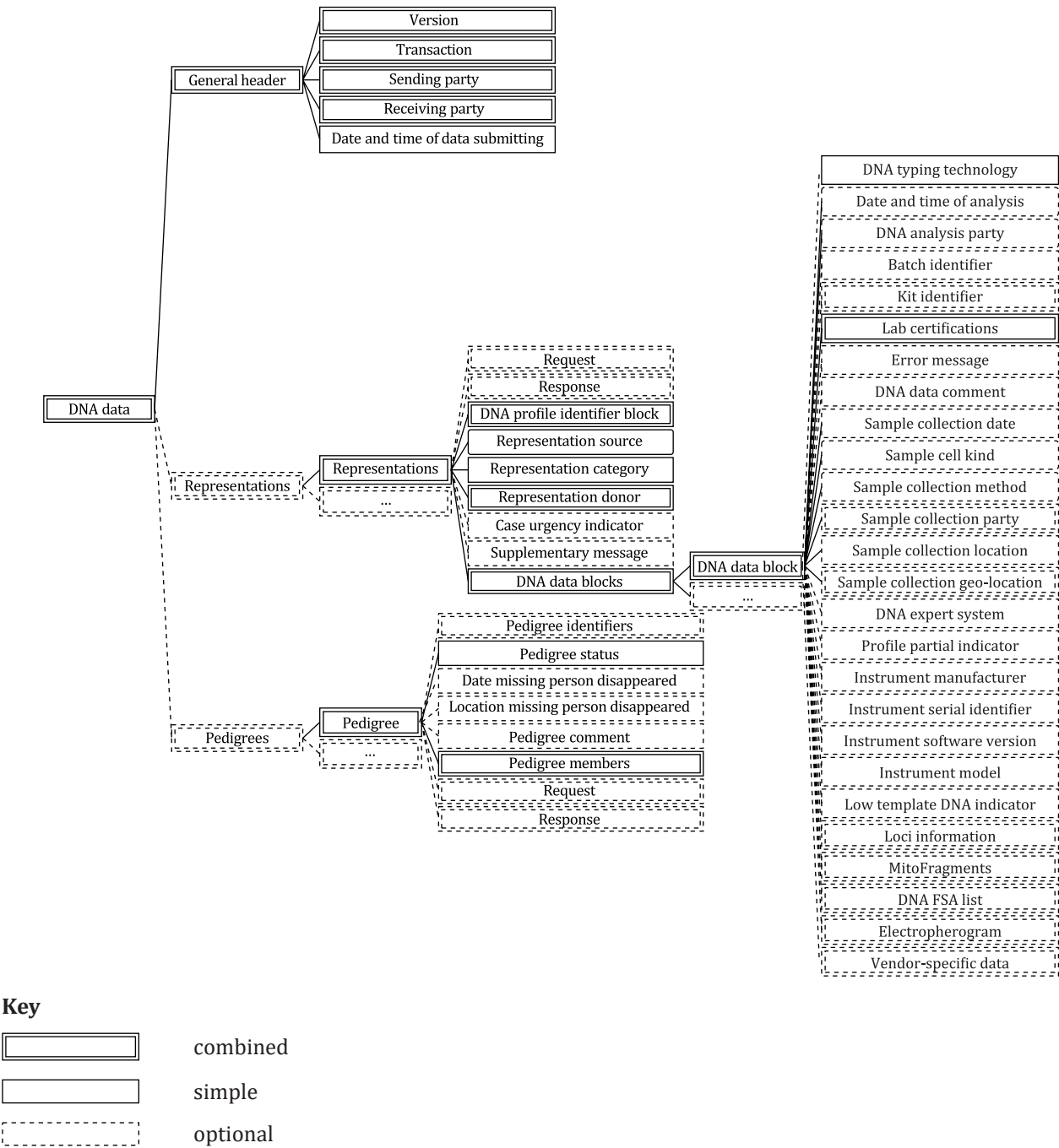


Figure 1 — DNA data format

6.2 Data conventions

6.2.1 Unknown field value

For mandatory fields of unrestricted string type, a value “Unknown” shall be used to denote that the information to be encoded in this field is not yet determined.

6.2.2 DNA data handlings

6.2.2.1 Mixed DNA profiles

For mixed profiles, the following data fields should be exchanged:

- Allele calls
- Electropherogram image
- Fragment sequence analysis (FSA) data, (e.g. fsa, hid files)
- Deconvoluted single source profiles with confidence intervals
- Algorithm/tool used to deconvolute

Deconvolution should not be carried out by hand. Currently there are sufficiently available software tools to infer contributing profiles to a DNA mixture (e.g. EuroForMix,^[1] CaseSolver,^[2] which are free). Free software exists to enable the development of searches for mixtures of more than two contributors (up to five), for specific cases, using the user's own database outside the Combined DNA Index System (CODIS).^[3]

6.2.2.2 Single source DNA profile with more than two allele values

More than two alleles can appear:

- in only one locus (tri-allelic patterns in single source samples);
- in several loci when the profile comes from DNA mixture (more than one contributor).

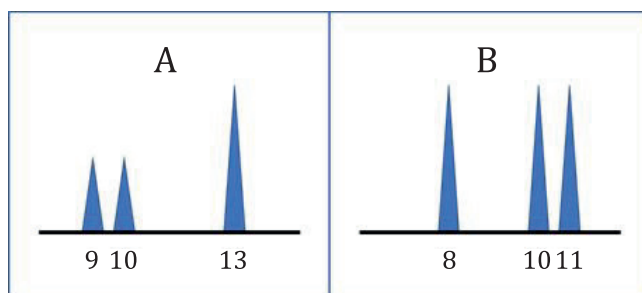
Two types of tri-allelic patterns exist (see [Figure 2](#)):

- type I (different peak heights) is due to somatic mutation of one allele that occurs during an individual's development. These patterns are characterized by uneven peak heights for the two variants of the affected allele that sum to the height of a non-mutated allele.
- type II (same peak heights) is mainly due to a localized duplication event or chromosomal aneuploidy (e.g. trisomy, three chromosomes instead of two chromosomes in a pair). These patterns are usually characterized by the appearance of peaks of equal height.

For the incidence of tri-allelic patterns, References [\[4\]](#) and [\[5\]](#) can be consulted.

NOTE The incidence of tri-allelic patterns type II is different in different loci.

EXAMPLE Loci D21S11 and D18S51 show more tri-allelic patterns than other loci due to trisomy in chromosomes 21 and 18 respectively; Penta_D is also located on chromosome 21; TPOX show tri-allelic patterns in African populations (2,4 %), with the extra allele almost always being allele 10 which is really located on X-chromosomes.^[6]

**Key**

- A type I
B type II

Figure 2 — Types of tri-allelic patterns

In addition to sending data in the most precise form known, there are several alternatives for exporting tri-allelic patterns.

- 1) First allele: If three alleles are provided for one locus, the first allele will be accepted and the remaining two alleles will be automatically converted for the export to a wildcard (*) and searched against all. Type I in [Figure 2](#) would be exported as “9; *” and Type II as “8; *”.
- 2) Any allele: One of the alleles in a locus, independent of its position, will be accepted in the case of non-mutated alleles. In this case, the other two allele values will be substituted by a wildcard. However, it is not always possible to verify which is the non-mutated allele. Type I in [Figure 2](#) would be exported as “13; *” and Type II as “8; *”.

While in Type I patterns the non-mutated allele can be inferred, this is not the case for Type II patterns. Therefore, option 1) proves more effective for use in practice.

If the DNA profile is a mixture, then use the recommendations for DNA data handling for mixed DNA profiles; see [6.2.2.1](#).

The parties involved in an automatic procedure on DNA data exchange should follow the following recommendations:

- generally, DNA profiles should be transmitted for comparison with their fully-designated allele values if they are fully available, as well as tri-allelic and/or rare off-ladder allele values.
- if DNA profiles in national data storage components store rare allele values as wild cards (*), they should be transmitted for comparison as stored in the national data storage components.

6.2.2.3 DNA profile with homozygosis

In case of a homozygote, both alleles should be given the same value in a DNA database according to the related genotype.

There are two cases to be differentiated:

- Only one allele available: a partial locus with an allelic dropout;
- Homozygote: same value for both alleles.

Allelic drop-out can result from various reasons. In evidential samples, degradation or very small amounts of DNA in the sample can lead to allelic dropout. Stochastic effects during the polymerase chain reaction (PCR) can cause the template DNA of some of the alleles to be insufficiently present for copying during the first cycles.

Reference samples of good quality can also show an apparent allelic drop-out. The reasons are very different from the ones described in the previous paragraph. In this case, a mutation in the primer binding zone can cause the lack of amplification of one of the two alleles of a locus. Usually, if the sample is amplified by using a different kit (i.e. with different primers), both alleles can be detected.

There should also be a designation that the profile is partial as partial DNA profiles can have bias, limited information or drop-in/out events present. These should be reviewed with caution and both of the parties (submitting and receiving) should know there is a partial DNA profile for review.

The combination of metadata fields and allele calls in this document allows exchanging of homozygous profiles and profiles with only one allele due to null allele or allele drop-out.

EXAMPLE 1 Only one allele available – locus with one reliably typed allele and a potential null/silent allele.

```
<LocusInformation>
  <LocusHeader>
    <LocusMarker>TPOX</LocusMarker>
    <LocusStatus>SilentAllele</LocusStatus>
    <AnalyzedBy>O.P. Wilson</AnalyzedBy>
    <LocusCategory>Autosomal</LocusCategory>
  </LocusHeader>
  <AlleleCalls>
    <AlleleCall>
      <Operator>Equal</Operator>
      <AlleleValue>8</AlleleValue>
    </AlleleCall>
  </AlleleCalls>
</LocusInformation>
```

EXAMPLE 2 Only one allele available – locus with one reliably typed allele and a suspected allele dropout.

```
<LocusInformation>
  <LocusHeader>
    <LocusMarker>D10S1248</LocusMarker>
    <LocusStatus>Partial</LocusStatus>
    <AnalyzedBy>O.P. Wilson</AnalyzedBy>
    <LocusCategory>Autosomal</LocusCategory>
  </LocusHeader>
  <AlleleCalls>
    <AlleleCall>
      <Operator>Equal</Operator>
      <AlleleValue>13</AlleleValue>
    </AlleleCall>
  </AlleleCalls>
</LocusInformation>
```

EXAMPLE 3 Only one allele available – locus with somewhat uncertain allele call.

```
<LocusInformation>
  <LocusHeader>
    <LocusMarker>D13S317</LocusMarker>
    <LocusStatus>NotDefinitive</LocusStatus>
    <AnalyzedBy>O.P. Wilson</AnalyzedBy>
    <LocusCategory>Autosomal</LocusCategory>
  </LocusHeader>
  <AlleleCalls>
    <AlleleCall>
      <Operator>Equal</Operator>
      <AlleleValue>9</AlleleValue>
    </AlleleCall>
  </AlleleCalls>
</LocusInformation>
```

EXAMPLE 4 Homozygote – same value for both alleles in the database.

```
<LocusInformation>
  <LocusHeader>
    <LocusMarker>CSF1PO</LocusMarker>
    <LocusStatus>Normal</LocusStatus>
    <AnalyzedBy>O.P. Wilson</AnalyzedBy>
    <LocusCategory>Autosomal</LocusCategory>
  </LocusHeader>
```

```

    <AlleleCalls>
      <AlleleCall>
        <Operator>Equal</Operator>
        <AlleleValue>10</AlleleValue>
      </AlleleCall>
    </AlleleCalls>
  </LocusInformation>

```

Moreover, the receiving party can view/examine the raw data (electropherogram image and .fsa file in the XML document) to perform additional analysis as needed.

6.2.2.4 DNA profile with rare allele values

Rare allele values can be very effective for identification purposes, if available. As indicated in [6.2.2.2](#), DNA profiles with their fully-designated allele values should be transmitted for comparison if they are available. However, different national DNA databases differ in the way rare allele values are recorded. A wildcard, which substitutes a rare allele value and denotes any value, may be used when data is exchanged for searching.^[7]

6.2.2.5 DNA profile with loci drop-out

Blank spaces exist in a partial DNA profile (a DNA profile for which complete typing results are not obtained at all tested loci). This can be due to limited DNA template, DNA degradation, inhibition, preferential amplification and/or stochastic effects.

A partial DNA profile should be left as it is, not supplemented by wildcards. Loci dropouts should be handled in XML by using the profile partial indicator.

EXAMPLE The missing values in a profile from using GlobalFiler Express kit with D7S820 and D2S1338 missing can be exchanged in the following XML code example:

```

<Representation>
  <Request>
    <RequestCategory>DataSubmissionAndSearch</RequestCategory>
  </Request>
  <DnaProfileIdBlock>
    <DnaProfileId>U654321</DnaProfileId>
    <CountryCode>BI</CountryCode>
    <FederalStateCode>Unknown</FederalStateCode>
    <OrganizationCode>Unknown</OrganizationCode>
  </DnaProfileIdBlock>
  <RepresentationSource>SingleSourceStain</RepresentationSource>
  <RepresentationCategory>UnidentifiedPerson</RepresentationCategory>
  <RepresentationDonor>
    <RepresentationDonorIndicator>Unknown</RepresentationDonorIndicator>
    <DonorGender>Male</DonorGender>
    <DonorVitalStatus>Dead</DonorVitalStatus>
  </RepresentationDonor>
  <CaseUrgencyIndicator>false</CaseUrgencyIndicator>
  <SupplementaryMessage>D7S820 and D2S1338 yield no data.</SupplementaryMessage>
  <DnaDataBlocks>
    <DnaDataBlock>
      <DnaTypingTechnology>STR</DnaTypingTechnology>
      <KitId>GlobalFiler Express</KitId>
      <LabCertifications>
        <LabCertification>
          <LabCertificationValue>IlacGuild19Accreditation</
LabCertificationValue>
          <ScopeOfAccreditations>
            <ScopeOfAccreditation>Nuclear</ScopeOfAccreditation>
          </ScopeOfAccreditations>
        </LabCertification>
      </LabCertifications>
      <ProfilePartialIndicator>true</ProfilePartialIndicator>
      <LociInformation>
        <LocusInformation>
          <LocusHeader>
            <LocusMarker>D3S1358</LocusMarker>

```

```

        <LocusStatus>Normal</LocusStatus>
        <AnalyzedBy>Unknown</AnalyzedBy>
        <LocusCategory>Autosomal</LocusCategory>
    </LocusHeader>
    <AlleleCalls>
        <AlleleCall>
            <Operator>Equal</Operator>
            <AlleleValue>14</AlleleValue>
        </AlleleCall>
        <AlleleCall>
            <Operator>Equal</Operator>
            <AlleleValue>15</AlleleValue>
        </AlleleCall>
    </AlleleCalls>
</LocusInformation>
</LociInformation>
</DnaDataBlock>
</DnaDataBlocks>
</Representation>

```

6.2.2.6 Specific DNA profiles

Specific DNA profiles can include those from staff, control/positive controls, known contamination, vendor profiles and/or high-profile casework.

- Control – positive controls
- Elimination – known contaminations, vendor profiles
- Population – DNA profiles for population frequency or study purpose
- Training – DNA profiles for DNA operation training, proficiency and/or competency
- Staff – DNA profiles from DNA operation staff for contamination monitoring

6.2.2.7 Chimera

In terms of handling a chimera, all alleles should be entered as well as notes on the origin and fact that it is a chimera in the SupplementaryMessage field (6.3.3.9). If there are additional renditions of the DNA profile, those should be linked to each other.

If differentiation between two separate chimera DNA profiles is possible, it should be carried out with a specific note to indicate this case.

6.3 Content of the DNA XML schema

6.3.1 Overview

A DNA data XML document consists of the fields shown in [Table 1](#).

Table 1 — Structure of a DNA data XML document

Field	Item type	Description	Mandatory / optional
GeneralHeader	GeneralHeaderType	See 6.3.2	Mandatory
Representations	Sequence of elements of RepresentationType	See 6.3.3	Optional
Pedigrees	Sequence of elements of PedigreeType	See 6.3.4	Optional

6.3.2 General header

6.3.2.1 Overview

The general header consists of the fields shown in [Table 2](#).

Table 2 — Structure of GeneralHeaderType

Field	Item type	Description	Mandatory / optional
Version	VersionType	See 6.3.2.2	Mandatory
Transaction	TransactionType	See 6.3.2.3	Mandatory
SendingParty	PartyType	See 6.3.2.4	Mandatory
ReceivingParty	PartyType	See 6.3.2.5	Mandatory
DateAndTimeOfDataSubmitting	dateTime	See 6.3.2.6	Mandatory

6.3.2.2 Version

The version field shall consist of a major version number and a minor revision number. The format is given in ISO/IEC 19794-1:2011/Amd. 2:2015. In an XML document following this document, the version field shall contain major version 4 and minor revision 0.

NOTE In an XML document following the first edition (ISO/IEC 19794-14:2013), the version field contains major version 3 and minor revision 0.

6.3.2.3 Transaction

A transaction field consists of the fields shown in [Table 3](#).

Table 3 — Structure of TransactionType

Field	Item type	Description	Mandatory / optional
TransactionID	string	Unique identifier of the transaction	Mandatory
CommunicationDirection	string	See Table 4	Mandatory
TransactionProcessingStatus	string	See Table 5	Mandatory only when communication direction is "Response"
TransactionProcessingMessage	string	If communication direction is "Response", transaction-level message (e.g. error or warning) from processing the corresponding request transaction	Mandatory only when communication direction is "Response"
RespondingToRequestID	string	If communication direction is "Response", TransactionID of the corresponding request transaction	Mandatory only when communication direction is "Response"

[Table 4](#) shows valid values for the communication direction.

Table 4 — Valid values for communication direction

Valid value	Description
Request	Request to a DNA profile database
Response	Response to a request transaction

[Table 5](#) shows valid values for the transaction processing status.

Table 5 — Valid values for transaction processing status

Valid value	Description
Success	Success of processing the request transaction
Fail	Failure of processing the request transaction

6.3.2.4 Sending party

The SendingParty field shall describe the sending party. It consists of the fields shown in [Table 6](#).

Table 6 — Structure of PartyType

Field	Item type	Description	Mandatory / optional
CountryCode	string	ISO 3166-1 alpha-2 country code	Optional
OrganizationName	string	Name of the party	Mandatory
OrganizationCode	string	Short identifier of the party EXAMPLE Originating agency identifier (ORI) for a law enforcement agency in the USA or a component identifier such as "DHS57/DHS/CBP/CSIS" within the Department of Homeland Security in the USA	Optional
PartyCategory	PartyCategoryType	See Table 7	Mandatory only for sending party
OrganizationPOCName	string	Name of the person in charge	Optional

The PartyCategory field consists of the fields shown in [Table 7](#).

Table 7 — Structure of PartyCategoryType

Field	Item type	Description	Mandatory / optional
OrganizationCategory	string	See Table 8	Mandatory
UnitCategory	string	See Table 9	Optional
UnitLocation	string	—	Optional

With rapid DNA instruments, an organization is not required to be a laboratory to produce DNA profiles.

EXAMPLE A police booking station can produce high-quality arrestee profiles using a rapid DNA instrument. In this case, the organization category is "G", and the DNA unit category value is "R". [Table 8](#) shows the valid values for the organization category.

Table 8 — Valid values for organization category

Valid value	Description
G	Government
I	Industry
O	Other (none of the specified values)
U	Unknown

[Table 9](#) shows the valid values for the unit category.

Table 9 — Valid values for unit category

Valid value	Description
L	Laboratory (traditional DNA processing)
R	Rapid DNA instrument
M	DNA mobile processing unit
U	Unknown

6.3.2.5 Receiving party

The ReceivingParty field shall describe the receiving party. It consists of the fields shown in [Table 6](#).

6.3.2.6 Date and time of data submitting

The DateAndTimeOfDataSubmitting field shall encode date and time of data submission, not data capture. This field shall be stated in Coordinated Universal Time (UTC).

6.3.3 Representations

6.3.3.1 Overview

For each DNA profile, there is a representation field. The structure of a representation is described in [Table 10](#).

Table 10 — Structure of a representation

Field	Item type	Description	Mandatory / optional
Request	RequestType	See 6.3.3.2	Mandatory only when communication direction is “Request”
Response	ResponseType	See 6.3.3.3	Mandatory only when communication direction is “Response”
DnaProfileIdBlock	DnaProfileIdBlockType	See 6.3.3.4	Mandatory
RepresentationSource	string	See 6.3.3.5	Mandatory
RepresentationCategory	string	See 6.3.3.6	Mandatory
RepresentationDonor	RepresentationDonorType	See 6.3.3.7	Mandatory
CaseUrgencyIndicator	boolean	See 6.3.3.8	Optional
SupplementaryMessage	string	See 6.3.3.9	Optional
DnaDataBlocks	Sequence of elements of DnaDataBlockType	See 6.3.3.10	Mandatory

6.3.3.2 Request

A request field shall be present when the communication direction is “Request”. The structure of a request field is described in [Table 11](#).

Table 11 — Structure of RequestType

Field	Item type	Description	Mandatory / optional
RequestCategory	string	See Table 12	Mandatory
UserDefined	UserDefinedType	Supplementary data, see Table 13	Optional
Description	string	Explanation of the request	Optional

[Table 12](#) shows the valid values for the request category.

Table 12 — Valid values for request category

Valid value	Description
DataSubmission	Add the DNA profile
DataEdit	Edit (i.e. replace) the DNA profile
DataRemove	Remove the DNA profile
DataVerify	Verify whether the DNA profile exists in the data storage/database

Table 12 (continued)

Valid value	Description
DataSubmissionAndSearch	Add the DNA profile and search
Search	Search the DNA profile but do not add it to the data storage/database
KinshipSearch	Perform kinship search but do not add to the storage database
KinshipVerify	Perform kinship verification but do not add it to the storage/database
Other	None of the specified values. Details shall be included in the Description field.

The structure of UserDefinedType is described in [Table 13](#).

Table 13 — Structure of UserDefinedType

Field	Item type	Description	Mandatory / optional
TypeCode	string	Type identifier	Optional
Data	base64Binary	Binary data	Mandatory

6.3.3.3 Response

The response field contains the comparison result from processing a request transaction. A response field shall be present when the value of communication direction is “Response”. The structure of a response field is described in [Table 14](#).

Table 14 — Structure of ResponseType

Field	Item type	Description	Mandatory / optional
ResponseCategory	string	See Table 15	Mandatory
RespondingToProfileId	DnaProfileIdBlockType	Link a DNA match in the response to the profile in the request. This element allows multiple matches of a DNA profile to be included in a response message.	Mandatory only when response category is “MatchCandidate”
RespondingToPedigreeId	IdType	Link a pedigree match in the response to the pedigree in the request.	Mandatory only when the response contains a pedigree
UserDefined	UserDefinedType	Supplementary data, see Table 13	Optional
MatchQuality	string	Match quality, see Table 16	Optional
Description	string	Explanation of the response	Optional

[Table 15](#) shows valid values for the response category field.

Table 15 — Valid values for response category field

Valid value	Description
UnableToProcess	In response to requests of any category: Unable to carry out the request. The reason/error message should be included in the Description field.
DataAdded	In response to requests of the categories “DataSubmission” or “DataSubmissionAndSearch”: The DNA profile is successfully added. In response to requests of the category “DataEdit”: The DNA profile did not exist and is successfully added.

Table 15 (continued)

Valid value	Description
DataEdited	In response to requests of the category “DataEdit”: The DNA profile is successfully edited/replaced. In response to requests of the categories “DataSubmission” or “DataSubmissionAndSearch”: The DNA profile already existed and is successfully edited/replaced.
DataRemoved	In response to requests of the category “DataRemove”: The DNA profile is successfully removed.
DataVerifiedExist	In response to requests of the category “DataVerify”: The DNA profile exists in the data storage/database.
DataVerifiedNotExist	In response to requests of the category “DataVerify”: The DNA profile does not exist in the data storage/database.
NonMatch	In response to requests of the categories “Search” or “DataSubmissionAndSearch”: No matching candidate found by a DNA search engine.
Match	In response to requests of the categories “Search” or “DataSubmissionAndSearch”: Matching candidate found by a DNA search engine.
NoHit	In response to requests of the categories “Search” or “DataSubmissionAndSearch”: Matching candidate found by a DNA search engine but rebutted by a DNA examiner.
Hit	In response to requests of the categories “Search” or “DataSubmissionAndSearch”: Matching candidate found by a DNA search engine and confirmed by a DNA examiner.
MatchCandidate	In response to requests of the categories “Search” or “DataSubmissionAndSearch”: DNA profile of matching candidate contained in response.
KinshipVerified	In response to requests of the category “KinshipVerify”: Kinship verified for the profiles/pedigree.
KinshipNegated	In response to requests of the category “KinshipVerify”: Kinship negated for the profiles/pedigree.
KinshipCandidate	In response to requests of the category “KinshipSearch”: Candidate pedigree contained in response.
Other	In response to requests of any category: None of the specified values. Details shall be included in the Description field.

[Table 16](#) shows valid values for the match quality field.

Table 16 — Valid values for match quality field

Valid value	Description
1-ExactMatch	Exact match (all identical match).
2-OneWildcardDifference	Only one difference, which involves a wildcard.
3-OneMicrovariantDifference	Only one difference, which involves a microvariant.
4-OneOtherDifference	The target profile and candidate profile have only one difference between them; and the difference involves neither wildcards nor microvariants.

6.3.3.4 DNA profile identifier block

The DNA profile identifier block shall encode a unique identifier for the DNA profile. The structure of a DNA profile identifier block is described in [Table 17](#).

Table 17 — Structure of DnaProfileIdBlockType

Field	Item type	Description	Mandatory / optional
DnaProfileId	string	Identifier of the DNA profile	Mandatory
CountryCode	string	ISO 3166-1 alpha-2 country code	Optional

Table 17 (continued)

Field	Item type	Description	Mandatory / optional
FederalStateCode	string	ISO 3166-2 alpha-2 country subdivision code	Optional
OrganizationCode	string	Short identifier of the organization maintaining the DNA profile EXAMPLE Originating agency identifier (ORI) for a law enforcement agency in the USA or a component identifier such as "DHS57/DHS/CBP/CSIS" within the Department of Homeland Security in the USA	Optional

6.3.3.5 Representation source

[Table 18](#) lists valid values for the representation source.

Table 18 — Valid values for representation source

Valid value	Description
Composite	The DNA profile is a composite profile from multiple samples from the same individual. The purpose of exchanging a composite DNA profile is to share the most complete and accurate DNA data of an individual that the sending party has, so that searching and/or verification can be optimized.
Person	The DNA profile comes from a reference sample of a person.
MixedStain	The DNA profile comes from an evidentiary or latent sample that has multiple contributors. For example, a mixture sample from a rape kit.
SingleSourceStain	The DNA profile comes from an evidentiary or latent sample of a single individual. For example, a sample obtained from a bloody fingerprint on the glass window of a burgled house. This sample has been verified to come from a single individual.
MixedOrSingleSourceStain	The DNA profile comes from an evidentiary or latent sample, and whether its source is a single individual or multiple individuals cannot be determined.
DeconvolutedFromMixedStain	The DNA profile comes from a mixture deconvolution analysis of a mixture sample.

6.3.3.6 Representation category

The valid values for the representation category are shown in [Table 19](#).

Table 19 — Valid values for representation category

Valid value	Description
Arrestee	DNA profile from a person who has been arrested
Asylee	DNA profile from an asylum seeker
BiologicalChild	DNA profile from a reference sample voluntarily provided by an adult child or provided with the parental/guardian consent for a minor child of a reported missing person
BiologicalFather	DNA profile from a known reference sample voluntarily provided by the biological father of a reported missing person
BiologicalMother	DNA profile a known reference sample voluntarily provided by the biological mother of a reported missing person
BiologicalSibling	DNA profile from a known reference sample voluntarily provided by the full or half biological adult sibling or provided with parental/guardian consent of the full or half biological minor sibling of a reported missing person
Control	DNA profile from a known sample to be used as a positive control in quality assurance (QA) processes

Table 19 (continued)

Valid value	Description
ConvictedOffender	DNA profile from an allowable offender, which includes convicted offenders (persons convicted or adjudicated of specified offenses), sex offenders and juveniles adjudicated delinquent, as defined by statute
Deportee	DNA profile from a person who has been deported
Detainee	DNA profile from a person who has been detained
Elimination	DNA profile from a known individual obtained to eliminate an individual as a possible source of DNA left by the perpetrator of a crime EXAMPLE DNA profile from a consensual sex partner
FamilyRelative	DNA profile from a known reference sample voluntarily provided by a relative of a missing person
Forensic	DNA profile from a biological sample from and associated with a crime scene
ImmigrationApplicant	DNA profile from an immigration applicant
ImmigrationSponsor	DNA profile from an immigration sponsor
Insurgent	DNA profile from a violent, armed rebel who is against authority
KnownOrSuspectedTerrorist	DNA profile from a person who is a known or suspected terrorist
MaternalRelative	DNA profile from a known reference sample voluntarily provided by a maternal biological relative who is not a mother, child or sibling of a reported missing person
MissingPerson	DNA profile from a known reference sample or a DNA sample from personal items purported to belong to the missing person (e.g. toothbrush) and compared to close biological relatives, if possible
Other	None of the specified values Details shall be included in the SupplementaryMessage field
PaternalRelative	DNA profile from a known reference sample voluntarily provided by a paternal biological relative who is not a father, child or sibling of a reported missing person
Population	DNA profile to be used for population frequency or study purpose
Refugee	DNA profile from a person who has been recognized under the 1951 Convention relating to the status of refugees to be a refugee
Spouse	DNA profile from a known reference sample voluntarily provided by a presumptive parent of a common child, used in kinship analysis (parentage trio)
Staff	DNA profile from laboratory personnel, law enforcement personnel, individuals that can come into contact with DNA samples, used to detect contamination in the DNA profiles
Suspect	DNA profile from a sample collected by law enforcement that is assumed to have come from a known individual who is believed to be a suspect in a crime EXAMPLE DNA profile from a sample from a cup used by the suspect during an interview or from a cigarette butt discarded by a suspect
Training	DNA profile to be used in DNA operation training, proficiency and/or competency
UnaccompaniedMinor	DNA profile from an unaccompanied minor that can be used to search for the minor's parents/relatives for family reunification
UnidentifiedPerson	DNA profile from the recovered deceased (body parts, tissue) or an individual who is unidentified (i.e. children who cannot and others who cannot or refuse to identify themselves)
Unspecified	Not stated in detail
Victim	DNA profile from an individual whose identity is known to law enforcement and who has been a subject of a crime
Volunteer	DNA profile from a person who volunteered a DNA sample

Some of the entries in [Table 19](#) (e.g. for detainee, arrestee, convicted offender) can have different meaning in different jurisdictions.

The procedure “deconvolute” is used to analyse mixture data in the data exchange by setting the representation source to “DeconvolutedFromMixedStain” and “DNA Expert System” to the mixture deconvolution software/tool name used to perform the mixture deconvolution.

6.3.3.7 Representation donor

6.3.3.7.1 Overview

The RepresentationDonor field contains information about the person whose DNA profile is exchanged. The fields of the representation donor field are listed in [Table 20](#).

Table 20 — Structure of RepresentationDonorType

Field	Item type	Description	Mandatory / optional
RepresentationDonorIndicator	string	See Table 21	Mandatory
DonorGender	string	See Table 22	Mandatory
DonorVitalStatus	string	See Table 23	Optional

6.3.3.7.2 Representation donor indicator

The representation donor indicator shows whether the identity of the donor of the DNA profile is known or not. The valid values for the representation donor indicator are shown in [Table 21](#).

NOTE The representation donor indicator can be useful in:

- a criminal investigation case where the DNA profile is a single source profile from the perpetrator. The indicator shows whether the perpetrator’s identity is known or not.
- a missing person case where the DNA profile comes from an unidentified person. The indicator shows whether the missing person has been positively identified or not.

Table 21 — Valid values for representation donor indicator

Valid value	Description
Known	The identity of the donor of the DNA profile is known
Unknown	The identity of the donor of the DNA profile is not known

6.3.3.7.3 Gender

[Table 22](#) lists valid values for gender.

Table 22 — Valid values for gender

Valid value	Description
Female	—
Male	—
Other	<p>None of the specified values</p> <p>Details may be included in the nearest comment field on a higher level</p> <p>NOTE 1 Details about the gender of a representation donor would be included in the SupplementaryMessage field of the representation</p> <p>NOTE 2 Details about the gender of a pedigree member would be included in the PedigreeComment field of the pedigree</p>

Table 22 (continued)

Valid value	Description
Unknown	The gender is not known

6.3.3.7.4 Donor vital status

[Table 23](#) lists valid values for donor vital status.

Table 23 — Valid values for donor vital status

Valid value	Description
Alive	—
Dead	—

6.3.3.8 Case urgency indicator

The valid values for the case urgency indicator are shown in [Table 24](#).

Table 24 — Valid values for case urgency indicator

Valid value	Description
True	The request/response is urgent The reason for urgency should be described in the SupplementaryMessage field
False	The request/response is not urgent

6.3.3.9 Supplementary message

If present, the SupplementaryMessage field shall contain additional information or comments about the entire DNA profile. Information or comments related to a specific DNA typing technology should be recorded in the DnaDataComment field of the technology-specific DNA data block.

6.3.3.10 DNA data blocks

6.3.3.10.1 Overview

For each DNA typing technology (see [Table 26](#)), there may be a DNA data block. The structure of a DNA data block is described in [Table 25](#).

Table 25 — Structure of DnaDataBlockType

Field	Item type	Description	Mandatory / optional
DnaTypingTechnology	string	See 6.3.3.10.2	Mandatory
DateAndTimeOfAnalysis	dateTime	See 6.3.3.10.3	Optional
DnaAnalysisParty	PartyType	See 6.3.3.10.4	Optional
BatchId	string	See 6.3.3.10.5	Optional
KitId	string	See 6.3.3.10.6	Optional
LabCertifications	Sequence of elements of LabCertType	See 6.3.3.10.7	Mandatory
ErrorMessage	string	See 6.3.3.10.8	Optional
DnaDataComment	string	See 6.3.3.10.9	Optional
SampleCollectionDate	dateTime	See 6.3.3.10.10	Optional
SampleCellKind	string	See 6.3.3.10.11	Optional

Table 25 (continued)

Field	Item type	Description	Mandatory / optional
SampleCollectionMethod	string	See 6.3.3.10.12	Optional
SampleCollectionParty	PartyType	See 6.3.3.10.13	Optional
SampleCollectionLocation	string	See 6.3.3.10.14	Optional
SampleCollectionGeoLocation	GeoLocationType	See 6.3.3.10.15	Optional
DnaExpertSystem	string	See 6.3.3.10.16	Optional
ProfilePartialIndicator	boolean	See 6.3.3.10.17	Optional
InstrumentManufacturer	string	See 6.3.3.10.18	Optional
InstrumentSerialId	string	See 6.3.3.10.19	Optional
InstrumentSoftwareVersion	string	See 6.3.3.10.20	Optional
InstrumentModel	string	See 6.3.3.10.21	Optional
LowTemplateDnaIndicator	boolean	See 6.3.3.10.22	Optional
LociInformation	Sequence of elements of LocusInfoType	See 6.3.3.10.23	Mandatory when DNA typing technology is STR; otherwise, absent
MitoFragments	Sequence of elements of MitoFragmentType	See 6.3.3.10.24	Mandatory when DNA typing technology is mtDNA; otherwise, absent
DnaFsaList	Sequence of elements of DnaFsaType	“Raw” DNA data or binary DNA FSA data See 6.3.3.10.25	Optional Most DNA FSA software/tools require FSA data from both the sample and the allelic ladder. So, if DNA FSA data is present, then two DNA FSA fields may be included: one for the sample and one for the ladder.
Electropherogram	EpgType	See 6.3.3.10.26	Optional
VendorSpecificData	VendorSpecificDataType	See 6.3.3.10.27	Optional

6.3.3.10.2 DNA typing technology

The DnaTypingTechnology field shall represent the technology utilized to profile the DNA sample. [Table 26](#) lists valid values for DNA typing technology.

Table 26 — Valid values for DNA typing technology

Valid value	Description
STR	STR analysis on nuclear DNA (including autosomal STR, X-STR and Y-STR)
mtDNA	Mitochondria first generation sequencing (FGS) technology, also called the mitochondria chain termination method or Sanger sequencing method
Other	None of the specified values Details shall be included in the DnaDataComment field.

6.3.3.10.3 Date and time of analysis

If present, the DateAndTimeOfAnalysis field shall encode the date and time of data analysis in UTC.

6.3.3.10.4 DNA analysis party

If present, the DnaAnalysisParty field shall capture the DNA profile audit trail. It consists of the fields shown in [Table 6](#).

6.3.3.10.5 Batch identifier

If present, the BatchIdentifier field shall encode the batch or lot number from the DNA analysis.

6.3.3.10.6 DNA kit identifier

If present, the DnaKitIdentifier field shall encode an identifier for the DNA analysis kit used. Examples of DNA kit identifiers are listed in the informative [Annex C](#).

6.3.3.10.7 Laboratory certifications

The LabCertifications field shall provide the certification status of the laboratory that processed the DNA. More than one certification is possible. The structure of a lab certification field is described in [Table 27](#).

Table 27 — Structure of LabCertType

Field	Item type	Description	Mandatory / optional
LabCertificationValue	string	Certifications of the laboratory that processed the DNA See Table 28	Mandatory
ScopeOfAccreditation	Sequence of strings	Scopes of accreditation of the laboratory that processed the DNA See Table 29	Mandatory More than one scope of accreditation is possible.

The valid laboratory certification values are listed in [Table 28](#).

Table 28 — Valid laboratory certification values

Valid value	Description
NoCertification	No certification
GlpValidation	GLP (Good Laboratory Practice) validation
AabbCertification	AABB (American Association of Blood Banks) certification
IsoIec17025	ISO/IEC 17025
IlacGuild19Accreditation	ILAC (International Laboratory Accreditation Cooperation) G19:08/2014
Other	None of the specified values Details shall be included in the DnaDataComment field.
Unknown	Not known

The valid values for scope of accreditation are listed in [Table 29](#).

Table 29 — Valid values for scope of accreditation

Valid value	Description
Nuclear	DNA profile determination with autosomal STR markers, DNA profile determination with Y-STR markers, DNA profile determination with X-STR markers, DNA profile determination with single-nucleotide polymorphism (SNP)
Mitochondrial	Haplotype determination with Sanger sequencing, haplotype determination with massively parallel sequencing
Database	National or state DNA profile database, other named DNA profile database, international DNA profile database
Other	None of the specified values Details shall be included in the DnaDataComment field.

6.3.3.10.8 Error message

If present, the ErrorMessage field shall contain any error or warning pertaining to processing the DNA data.

EXAMPLE Inconsistent DNA data (donor gender and amelogenin data do not match or invalid locus marker) or insufficient DNA data (insufficient number of loci).

6.3.3.10.9 DNA data comment

If present, the DnaDataComment field shall contain additional information or comments about the technology-specific DNA data.

6.3.3.10.10 Sample collection date

If present, the SampleCollectionDate field shall encode the date and time the sample was collected. This field shall be stated in UTC.

6.3.3.10.11 Sample cell kind

If present, the SampleCellKind field shall represent the cell type from where the sample was extracted. This field is only needed (and useful) when the representation is about a single sample. The valid values for the sample cell kind are shown in [Table 30](#).

Table 30 — Valid values for sample cell kind

Valid value	Description
Blood	Blood
Bone	Bone
BuccalCell	Buccal cell
CommingledBiologicalMaterial	Commingled biological material
Hair	Hair
Saliva	Saliva
Semen	Semen
Skin	Skin
SweatFingerprint	Sweat/fingerprint
Tissue	Tissue
ToothPulp	Tooth (including pulp)
Other	None of the specified values Details shall be included in the DnaDataComment field.

6.3.3.10.12 Sample collection method

If present, the SampleCollectionMethod field shall describe the method used to collect the sample.

EXAMPLE If the sample cell kind is "Hair", the sample can be gathered from the donor directly or picked up from a crime scene.

6.3.3.10.13 Sample collection party

If present, the SampleCollectionParty field shall indicate the agency that collected the sample. The fields of the SampleCollectionParty field are listed in [Table 6](#).

6.3.3.10.14 Sample collection location

If present, the SampleCollectionLocation field shall denote the location where the sample was collected.

EXAMPLE Houses, buildings and/or postal addresses, etc.

6.3.3.10.15 Sample collection geo-location

If present, this field shall contain the GPS (global positioning system) coordinates of the location where the sample was collected. This supports processing of large-scale disasters. [Table 31](#) shows the structure of the sample collection geo-location field.

Table 31 — Structure of SampleCollectionGeoLocation

Field	Item type	Description	Mandatory / optional
Latitude	float	Latitude in WGS (World Geodetic System) 1984 + for North, - for South	Mandatory
Longitude	float	Longitude in WGS 1984 + for East, - for West	Mandatory

6.3.3.10.16 DNA expert system

If present, the DnaExpertSystem field shall give any expert system(s) used to obtain the allele calls and/or deconvoluted profiles from a mixture.

EXAMPLE ANDE Expert System v.2.0.5, GeneMapperID-X v1.4 and OSIRIS v.x.

6.3.3.10.17 Profile partial indicator

The valid values for the profile partial indicator are shown in [Table 32](#).

Table 32 — Valid values for profile partial indicator

Valid value	Description
True	The DNA profile is a partial DNA profile
False	The DNA profile is complete

6.3.3.10.18 Rapid DNA instrument manufacturer

If a rapid DNA instrument was used, the InstrumentManufacturer field shall give the name of the manufacturer. The valid values for rapid DNA instrument manufacturer are listed in [Table 33](#).

Table 33 — Valid values for rapid DNA instrument manufacturer

Valid value	Description
ThermoFisher	Thermo Fisher
ANDE	ANDE
Other	None of the specified values Details shall be included in the DnaDataComment field.
Unknown	Not known

6.3.3.10.19 Rapid DNA instrument serial identifier

If a rapid DNA instrument was used, the InstrumentSerialId shall give a unique identifier of the rapid DNA instrument series.

6.3.3.10.20 Rapid DNA instrument software version

If a rapid DNA instrument was used, the InstrumentSoftwareVersion shall give the version number of the software that was running on the instrument when the DNA data was produced.

6.3.3.10.21 Rapid DNA instrument model

If a rapid DNA instrument was used, the InstrumentModel field shall give the name of the rapid DNA instrument model. The valid values for rapid DNA instrument model are listed in [Table 34](#).

Table 34 — Valid values for rapid DNA instrument model

Valid value	Description
RapidHitId	RapidHIT ID
ANDE6C	ANDE 6C
Other	None of the specified values Details shall be included in the DnaDataComment field
Unknown	Not known
NOTE These are examples of suitable products available commercially. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO or IEC of these products.	

6.3.3.10.22 Low template DNA indicator

The LowTemplateDnaIndicator field is used to indicate that the DNA profile is obtained from a sample with only a small amount of DNA. Stochastic effects inherent to analysis of low amounts of DNA yield allele or locus drop-outs. Additionally, increasing detection sensitivity can result in a greater potential for contamination or allele drop-ins. This indicator allows the receiver to use the appropriate inclusion and match rules to effectively handle the DNA profile. The valid values for the low template DNA indicator are shown in [Table 35](#).

Table 35 — Valid values for the low template DNA indicator

Valid value	Description
True	The DNA profiling has not produced good quantization values
False	The DNA profiling has produced good quantization values

6.3.3.10.23 STR DNA profile

A STR DNA profile is encoded in a sequence of locus information fields. The structure of a locus information field is described in [Table 36](#).

Table 36 — Structure of LocusInfoType

Field	Item type	Description	Mandatory / optional
LocusHeader	LocusHeaderType	See Table 37	Mandatory
AlleleCalls	Sequence of elements of AlleleCallType	Sequence of two or more allele calls (repeat numbers) determined by comparison of peak size in an electropherogram to allelic ladder allele peak sizes See Table 40	Mandatory

The structure of a locus header field is described in [Table 37](#).

Table 37 — Structure of LocusHeaderType

Field	Item type	Description	Mandatory / optional
LocusMarker	string	Name of the locus For identification/verification, see D.2 and D.3 . For other purposes, see D.1	Mandatory
LocusStatus	string	See Table 38	Mandatory
AnalyzedBy	string	Name of the person who analysed or reviewed the locus genotype For a DNA profile generated by a rapid DNA instrument, AnalyzedBy may contain the name of the administrator of the instrument	Mandatory
BatchId	string	Identifier for the batch within which the DNA profile was analysed	Optional This field should be used only in case the DNA profile is a composite of multiple DNA test runs (e.g. data from Profiler Plus and Cofiler).
KitId	string	Identifier for the DNA analysis kit used Examples are listed in the informative Annex C .	Optional This field should be used only in case the DNA profile is a composite of multiple DNA test runs (e.g. data from Profiler Plus and Cofiler)
LocusCategory	string	See Table 39	Mandatory

Valid values for locus status are listed in [Table 38](#).

Table 38 — Valid values for locus status

Valid value	Description
Normal	Homozygous or heterozygous locus in which the allele(s) is/are reliably genotyped
SilentAllele	Locus with a single reliably genotyped silent or null allele
NotDefinitive	Locus at which the genotype contains ambiguity. EXAMPLE Alleles derived from a mixture deconvolution NOTE DeconvolutedFromMixedStain is a defined value in the representation source field.
Partial	Partial locus, i.e. locus with allele drop-out(s)

The valid values for locus category are listed in [Table 39](#).

Table 39 — Valid values for locus category

Valid value	Description
Autosomal	Autosomal STR
X-STR	X-STR
Y-STR	Y-STR
Other	None of the specified values Details shall be included in the DnaDataComment field

The structure of an AlleleCall field is described in [Table 40](#).

Table 40 — Structure of AlleleCallType

Field	Item type	Description	Mandatory / optional
Operator	string	See Table 41	Mandatory
AlleleValue	string	Value determined by comparison of peak size in an electropherogram to allelic ladder allele peak sizes Possible values: repeat number as integer or two integers separated by “.”, X or Y if locus marker is “Amelogenin” or wildcard “*”	Mandatory

The valid values for operator are listed in [Table 41](#).

Table 41 — Valid values for operator

Valid value	Description
Equal	The allele call is equal to the value in the AlleleValue field
BelowLowerLimit	The allele call is below the lower limit of the allelic ladder. The value of the lower limit of the allelic ladder is specified in the AlleleValue field
AboveUpperLimit	The allele call is above the upper limit of the allelic ladder. The value of the upper limit of the allelic ladder is specified in the AlleleValue field

Alleles with operator “AboveUpperLimit” and “BelowLowerLimit” represent off-ladder alleles. The allele value of an off-ladder allele may be substituted by a wildcard “*” representing a rare allele value significant for identity.

6.3.3.10.24 Mitochondrial DNA data

Mitochondrial DNA (mtDNA) is a molecule containing around 16 569 units called nucleotides. There are four types of nucleotides: A (adenine), C (cytosine), T (thymine) and G (guanine). The sequence of nucleotides that make up the human mitochondrial genome was first analyzed by Anderson et al. in 1981;^[8] this mtDNA sequence is called the Anderson or Cambridge Reference Sequence (CRS). Subsequently, Andrews et al. reproduced the same analysis on the same biological material but using more advanced technology.^[9] They found some errors in the CRS, then gave a revised version of the sequence (revised Cambridge Reference Sequence or rCRS). In both versions, each of the nucleotides of the mitochondrial genome was assigned a consecutive number, starting at 1 and ending at 16569.

The mtDNA control region is a part of the mtDNA molecule that extends from nucleotide position 16024 to 16569 and from 1 to 576. However, not all laboratories analyze the whole mtDNA control region but rather only a part of it. Forensic laboratories analyze the so-called hypervariable regions of mitochondrial DNA HV1 and HV2, which refer to the following mtDNA ranges (nucleotide positions):

- HV1: from 16024 to 16365
- HV2: from 73 to 340.

When exchanging data, the range that is analyzed by the laboratory shall be precisely defined since this range has the same function as the name of an autosomal STR (aSTR) marker.

EXAMPLE The following are examples of mtDNA analysis ranges:

- 16024 to 16365
- 16024 to 16569
- 73 to 340
- 1 to 576
- 16024 to 576

Figure 3 illustrates the structure of mtDNA data. A mitochondrial DNA profile is encoded in a sequence of MitoFragment fields.

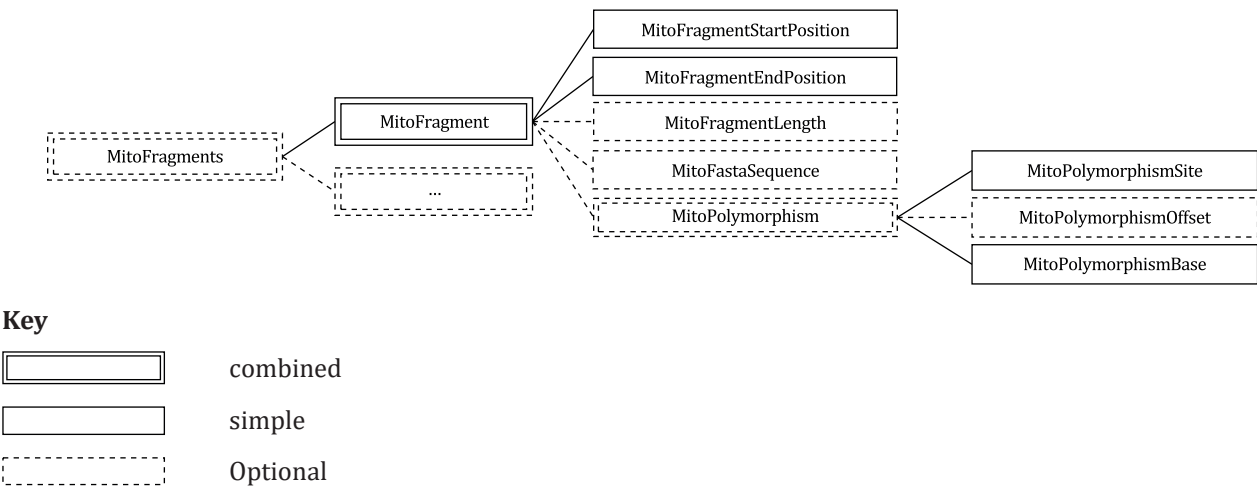


Figure 3 — Structure of mtDNA data

The structure of a MitoFragment field is described in Table 42.

Table 42 — Structure of MitoFragmentType

Field	Item type	Description	Mandatory / optional
MitoFragmentStartPosition	integer	Start position of the mtDNA analysis range	Mandatory
MitoFragmentEndPosition	integer	End position of the mtDNA analysis range	Mandatory
MitoFragmentLength	integer	Length of the mtDNA analysis range	Optional
MitoFastaSequence	string	Sequence in FASTA format	Optional
MitoPolymorphism	MitoPolymorphismType	See Table 43	Optional

The structure of a mitochondrial polymorphism field is described in Table 43.

Table 43 — Structure of MitoPolymorphismType

Field	Item type	Description	Mandatory / optional
MitoPolymorphismSite	integer	Location of the polymorphism	Mandatory
MitoPolymorphismOffset	integer	Location of an insertion	Optional
MitoPolymorphismBase	string	Nucleotide at the specified MitoPolymorphismSite See Table 44	Mandatory

EXAMPLE The following XML snippet expresses an mtDNA fragment 1 to 576 with a polymorphism 309.1C:

```
<MitoFragment>
  <MitoFragmentStartPosition>1</MitoFragmentStartPosition>
  <MitoFragmentEndPosition>576</MitoFragmentEndPosition>
  <MitoPolymorphism>
    <MitoPolymorphismSite>309</MitoPolymorphismSite>
    <MitoPolymorphismOffset>1</MitoPolymorphismOffset>
    <MitoPolymorphismBase>C</MitoPolymorphismBase>
  </MitoPolymorphism>
</MitoFragment>
```

```
</MitoPolymorphism>
</MitoFragment>
```

mtDNA data can be very useful in the identification of bodies through their relatives. In fact, mtDNA data is often used to distinguish adventitious aSTR matches that result after kinship searches (low power of discrimination). Any mtDNA range (e.g. only HV1) can be used in this way after an aSTR search has been done.

A distinction can be made between mass comparisons and one-to-one comparisons. Mass international comparisons based only on mtDNA data (without aSTR) may result in a high number of adventitious matches since the power of discrimination of mtDNA is not as high as that of aSTR data. For instance, some haplotypes show a frequency of 5 % in some populations. Therefore, mass international comparisons are optional and subject to agreement of the countries involved. In the case of mass international comparisons, the exchange of information for at least the 16024 to 16365 and 73 to 340 ranges is recommended if only mtDNA data is exchanged. If the DNA information to be exchanged also includes aSTR data, then any mtDNA range can be exchanged. Also, if the type of comparison is one to one, no restrictions regarding the mtDNA range are necessary.

Once the analysis range has been defined, the polymorphisms of the particular sample have to be defined (haplotype). The nomenclature for mitochondrial haplotypes has been defined by the ISFG (International Society for Forensic Genetics).^[10] Currently, the revised Cambridge Reference Sequence (rCRS) is the sequence used to name the different mitochondrial polymorphisms. Thus, the nomenclature of each haplotype only reflects the differences between the haplotype of interest and the rCRS. When a difference between an individual's sequence and the rCRS is observed, only the nucleotide position and the nucleotide differing from the reference standard are recorded. For example, at site 73, the rCRS sequence has an A, but a large portion of the population carries a G at site 73. Such an individual's mtDNA sequence is described as 73G.

[Table 44](#) shows the character values established by the International Union of Pure and Applied Chemistry (IUPAC) for DNA base type measurements. The valid sequence alphabet is “A”, “T”, “C”, “G” and the IUPAC ambiguity codes. Any location that does not have a value shall be set to N (i.e. A, G, T or C).

Table 44 — Valid values for mitochondrial polymorphism base, established by IUPAC

IUPAC value	Description
G	Guanine
A	Adenine
T	Thymine
C	Cytosine
R	G, A
Y	T, C
M	A, C
K	G, T
S	G, C
W	A, T
H	A, C, T
B	G, T, C
V	G, C, A
D	G, A, T
N	G, A, T, C
–	Deletion

Insertions are usually labelled with “.1X”, (where X = A, T, C or G, depending on the nucleotide inserted). If two nucleotides are inserted, they are labelled as .1X and .2X. For instance, if two C residues are inserted after nucleotide position 309, they will be named as: 309.1C 309.2C. Deletions are recorded

by listing the missing site followed by 'del' or '-' (i.e., 249del or 249-). Nomenclature of mtDNA in the forensic field is more complicated than these examples (e.g. heteroplasmies). For a detailed guide on mtDNA nomenclature, the reader can consult Reference [10].

Some problems can arise regarding mtDNA nomenclature when aligning the haplotype of the sample of interest with the rCRS. These problems affect the results of the mtDNA searches. A specific haplotype can be named in more than one way. In [Figure 4](#), the nucleotide sequence between positions 145 and 160 is shown for both rCRS and a specific sample. The sample can be aligned with the rCRS in two ways, resulting in two different nomenclatures. If one laboratory sends the haplotype using Alignment 1 in [Figure 4](#), but the receiver has stored their haplotypes using Alignment 2, the match will never be found even if the haplotypes are identical.

	145	146	147	148	149	150	151	152		153	154	155	156	157	158	159	160	Nomenclature
rCRS	C	T	C	A	T	C	C	T		A	T	T	A	T	T	T	A	
Sample	C	T	C	A	T	C	T	C		A	T	T	A	T	T	T	A	
Alignment 1	C	T	C	A	T	C	T	C		A	T	T	A	T	T	T	A	151T 152C
Alignment 2	C	T	C	A	T	C	-	T	C	A	T	T	A	T	T	T	A	151del 152.1C

Figure 4 — mtDNA alignment problems

ISFG recommends aligning haplotypes by following phylogenetic rules, but this is not always straightforward because this requires a deep knowledge of all mtDNA haplotypes described and the relation between them regarding their origin. Two strategies can be followed to solve this problem.

- Developing searching algorithms that do not consider the differences with the rCRS, such as that developed in the EMPOP database.^[11] Instead of searching differences to rCRS (e.g. 151T 152C or 151del 152.1C), the haplotype is converted into the FASTA format, and the search is performed with this string (in the example: CTCATCTCATTATTTA).
- Using the EMPOP alignment tool^[18] before entering a haplotype in the database. This avoids different alignments for the same DNA sequence since EMPOP always aligns the haplotypes in the same way following phylogenetic rules (see an example in [Figure 5](#)).

Query

Result

Details

Neighbors

Alignment

Haplogrouping

Phylogenetic alignment

Input profile	73G	263G	309.1C	309.2C	315.1C	498-	499A	16183C	16188.1C		16193.1C		16217C	16304C	16519C	
Phylogenetic alignment	73G	263G	309.1C	309.2C	315.1C	498-	499A	16183C		16189C	16190T	16193.1C	16193.2C	16217C	16304C	16519C

Figure 5 — EMPOP alignment tool

6.3.3.10.25 DNA FSA data

The structure of a DNA FSA field is described in [Table 45](#).

Table 45 — Structure of DnaFsaType

Field	Item type	Description	Mandatory / optional
Id	string	Identifier for the DNA FSA data	Optional
DnaFsaCategory	string	See Table 46	Mandatory
StorageFormat	string	See Table 47	Mandatory
BinaryObject	base64Binary	base64Binary encoding of the FSA data	Mandatory

The valid values for the category field are listed in [Table 46](#).

Table 46 — Valid values for DNA FSA data category

Valid value	Description
Sample	The FSA data is for the sample
Ladder	The FSA data is for the ladder

The valid values for the DNA FSA storage format are listed in [Table 47](#).

Table 47 — Valid values for DNA FSA storage format

Valid value	Description
ab1	The binary DNA FSA file is in Sanger sequence format and has the filename extension “ab1”. This DNA FSA storage format is used for mtDNA data.
fsa	The binary DNA FSA file has the filename extension “fsa”. This DNA FSA storage format is used for STR data
hid	The binary DNA FSA file has the filename extension “hid”. This DNA FSA storage format is used for STR data

6.3.3.10.26 Electropherogram

The structure of an electropherogram field is described in [Table 48](#).

Table 48 — Structure of EpgType

Field	Item type	Description	Mandatory / optional
ImageId	string	Identifier for the electropherogram	Optional
StorageFormat	string	Data format of the image, usually the image file extension EXAMPLE “png”, “jpg” or “jpeg”	Mandatory
BinaryObject	base64Binary	base64Binary encoding of the image	Mandatory

6.3.3.10.27 Vendor-specific data

The vendor-specific data field consists of a type code and a binary data block of that type as specified in ISO/IEC 19794-1:2011/Amd. 2:2015.

6.3.4 Pedigrees

6.3.4.1 Overview

Pedigrees may be included in the request of a kinship search for missing person or mass fatality incident use cases. In some cases, a pedigree may be included in the request of a familial search for a criminal investigation if the law permits. The response to a kinship search request may contain multiple pedigrees if the kinship search returns multiple candidate kinship matches. The structure of PedigreeType is described in [Table 49](#).

Table 49 — Structure of PedigreeType

Field	Item type	Description	Mandatory / optional
PedigreeIds	Sequence of elements of IdType	See 6.3.4.2	Mandatory when communication direction is “Request”; otherwise, optional
PedigreeStatus	string	See 6.3.4.3	Mandatory
DateMissingPersonDisappeared	date	See 6.3.4.4	Optional

Table 49 (continued)

Field	Item type	Description	Mandatory / optional
LocationMissingPersonDisappeared	string	See 6.3.4.5	Optional
PedigreeComment	string	See 6.3.4.6	Optional
PedigreeMembers	Sequence of elements of PedigreeMbrType	See 6.3.4.7	Mandatory At least two
Request	RequestType	See 6.3.3.2	Mandatory when communication direction is “Request”; otherwise, absent
Response	ResponseType	See 6.3.3.3	Mandatory when communication direction is “Response”; otherwise, absent

6.3.4.2 Pedigree identifier

In a missing person or mass fatality incident case, a pedigree is often associated with a pedigree identifier, a missing person case number, etc. Multiple pedigree identifiers may be included for tracking and cross-reference purposes. In a request, at least one pedigree identifier shall be included. [Table 50](#) shows the structure of the pedigree identifier field.

Table 50 — Structure of IdType

Field	Item type	Description	Mandatory / optional
IdName	string	Class of identifiers	Mandatory
IdValue	string	Actual identifier	Mandatory

EXAMPLE The following XML snippet shows a pedigree identifier “12345” and a reported missing person case “RM20200412120911001”:

```
<PedigreeIds>
  <PedigreeId>
    <IdName>Pedigree identifier</IdName>
    <IdValue>12345</IdValue>
  </PedigreeId>
  <PedigreeId>
    <IdName>Reported missing person case</IdName>
    <IdValue>RM20200412120911001</IdValue>
  </PedigreeId>
</PedigreeIds>
```

6.3.4.3 Pedigree status

The valid values for pedigree status are listed in [Table 51](#).

Table 51 — Valid values for pedigree status

Valid value	Description
Claimed	In a request: The kinship/relationship described by the pedigree has not been verified yet In a response: This pedigree was the original input
Verified	The kinship/relationship described by the pedigree has been verified by DNA analysis
Negated	The kinship/relationship described by the pedigree has been negated by DNA analysis
Candidate	A candidate for the target member of the pedigree has been identified

6.3.4.4 Date missing person disappeared

If present, the DateMissingPersonDisappeared field shall encode the date of disappearance of a missing person.

6.3.4.5 Location missing person disappeared

If present, the LocationMissingPersonDisappeared field shall denote the place where a missing person disappeared.

6.3.4.6 Pedigree comment

If present, the pedigree comment shall hold any comments or notes about the pedigree.

EXAMPLE If the family references in a missing person case have been verified to be true prior to being used in a kinship search request, then “The known part of the pedigree has been verified” may be included in the pedigree comment field.

6.3.4.7 Pedigree members

A pedigree shall have at least two members. Genetic data should be consistent with the pedigree information provided. Each pedigree may have one or more “ancillary” members representing pedigree members without a DNA profile. An ancillary pedigree member ties another pedigree member together to define the kinship. A pedigree member field may be annotated uniquely with an ID attribute of built-in XSD type ID, as in <PedigreeMember id="ID-1">.

[Table 52](#) shows the structure of a pedigree member field.

NOTE MotherId and FatherId and ancillary pedigree members constitute a common mechanism to define a pedigree of any complexity. The alternative is to specify the name of the relationship such as “mother”, “grandmother”, “aunt”, etc. The problem with using relationship names is that it is not scalable for complicated relationships where more than two pedigree members are involved; and it is possible that the relationship name is not precise enough to define a kinship search.

Table 52 — Structure of PedigreeMbrType

Field	Item type	Description	Mandatory / optional
PedigreeMemberIds	Sequence of elements of IdType	Identifiers of the DNA profile of the pedigree member See Table 50 If a DNA profile is associated with a pedigree member, then the identifier of the DNA profile should be specified in the pedigree member identifier field The associated DNA profile should be included in the representations field of the XML document (see 6.3.3)	Optional because a pedigree member can be “ancillary” without a known DNA profile There may be multiple pedigree member identifiers for linkage and/or cross-referencing
MotherId	Reference to the ID attribute of a pedigree member specified within the same XML document	Identifier of the pedigree member’s biological mother	Optional
FatherId	Reference to the ID attribute of a pedigree member specified within the same XML document	Identifier of the pedigree member’s biological father	Optional
PedigreeMemberStatus	string	See Table 53	Mandatory