Forensic Science International: Genetics 10 (2014) 64–72

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

GHEP-ISFG collaborative exercise on mixture profiles of autosomal STRs (GHEP-MIX01, GHEP-MIX02 and GHEP-MIX03): Results and evaluation

M. Crespillo [a,b,*], P.A. Barrio [a,b], J.A. Luque [a,b], C. Alves [c], M. Aler [d], F. Alessandrini [e], L. Andrade [f], R.M. Barretto [g], A. Bofarull [b], S. Costa [h], M.A. García [i], O. García [j], A. Gaviria [k], A. Gladys [l], A. Gorostiza [m], A. Hernández [n], M. Herrera [o], L. Hombreiro [p], A.A. Ibarra [q], M.J. Jiménez [r], G.M. Luque [s], P. Madero [t], B. Martínez-Jarreta [u], M.V. Masciovecchio [v], N.M. Modesti [w], F. Moreno [x], S. Pagano [y], S. Pedrosa [z], G. Plaza [A], E. Prat [B], J. Puente [C], F. Rendo [D], T. Ribeiro [E], A. Sala [F], E. Santamaría [G], V.G. Saragoni [H], M.R. Whittle [I]

[a] Mixture Commission of the GHEP-ISFG (The Spanish and Portuguese Speaking Working Group of the International Society for Forensic Genetics), Spain
[b] INTCF – National Institute of Toxicology and Forensic Science, Department of Barcelona, Spain
[c] IPATIMUP – Institute of Molecular Pathology and Immunology of the University of Porto, Porto, Portugal
[d] Sección de Genetica Forense y Criminalística, Servicio de Laboratorio, Valencia, Spain
[e] Department of Biomedical Sciences and Public Health, Università Politecnica delle Marche, Italy
[f] Serviço de Genética e Biologia Forenses, Delegação do Centro, Instituto Nacional de Medicina Legal e Ciências Forenses, I.P. Coimbra, Portugal
[g] Universidade Estadual Paulista/UNESP, Laboratorio de Investigacao de Paternidade, Sao Paulo, SP, Brazil
[h] Genetic Forensic Service, North Branch of the National Institute of Legal Medicine and Forensic Sciences, IP, Porto, Portugal
[i] Servicio de Criminalística de la Guardia Civil, Laboratorio Central de Criminalística, Departamento de Biología, Madrid, Spain
[j] Forensic Science Unit, Forensic Genetics Section, Basque Country Police, Erandio, Bizkaia, Spain
[k] Laboratorio de Genética Molecular, Cruz Roja Ecuatoriana, Quito, Ecuador
[l] Laboratorio de Análisis Comparativo de ADN, de la Suprema Corte de Justicia de Buenos Aires, Argentina
[m] Laboratorio Identificación Genética, GENOMICA SAU, Grupo Zeltia, Madrid, Spain
[n] INTCF – National Institute of Toxicology and Forensic Science, Delegation of Canarias, Santa Cruz de Tenerife, Spain
[o] Laboratorio Genda SA., Ciudad Autónoma de Buenos Aires, Argentina
[p] Laboratorio Territorial de Biología – ADN, Brigada de Policía Científica, Jefatura Superior de Policía de Galicia, La Coruña, Spain
[q] Laboratorio de Identificación Genética – IdentiGEN, Universidad de Antioquia, Colombia
[r] Unitat Central de Laboratori Biològic, Àrea Central de Criminalística, Divisió de Policia Científica, Policia de la Generalitat – Mossos d'Esquadra, Barcelona, Spain
[s] INTCF – National Institute of Toxicology and Forensic Science, Department of Madrid, Spain
[t] Centro de Análisis Genéticos, Zaragoza, Spain
[u] Laboratorio de Genetica y Identificación humana, Facultad de Medicina, Universidad de Zaragoza, Spain
[v] Bioquimica, IACA Laboratorios, Bahía Blanca, Argentina
[w] Instituto de Genética Forense, Poder Judicial de Córdoba, Argentina
[x] Unidad de Genética Forense, Servicio Médico Legal, Santiago de Chile, Chile
[y] Laboratorio de Análisis de ADN para el CODIS, Dirección Nacional de Policía Técnica, Uruguay
[z] Navarra de Servicios y Tecnologias S.A.U. (NASERTIC), Pamplona, Spain
[A] NEODIAGNOSTICA, S.L., Lleida, Spain
[B] Laboratorio de ADN, Comisaría General de Policía Científica, Madrid, Spain
[C] LabGenetics, Laboratorio de Genética Clínica S.L., Madrid, Spain
[D] Sequencing and Genotyping Unit, Genomics Service – SGIker, University of the Basque Country, UPV/EHU, Bilbao, Spain
[E] Serviço de Genética e Biologia Forenses, Instituto Nacional de Medicina Legal e Ciências Forenses, Delegação Sul, Lisboa, Portugal
[F] Servicio de Huellas Digitales Genéticas, Fac. Farmacia y Bioquímica, University de Buenos Aires, Argentina
[G] INTCF – National Institute of Toxicology and Forensic Science, Department of Sevilla, Spain
[H] Unidad de Genética Forense, Departamento de Laboratorios, Servicio Médico Legal de Santiago de Chile, Santiago, Chile
[I] Genomic Engenharia Molecular, São Paulo, Brazil

* Corresponding author at: C/ La Mercè, 1, 08002 Barcelona, Spain. Tel.: +34 93 3174061; fax: +34 93 3182530.
E-mail address: manuel.crespillo@mju.es (M. Crespillo).

1872-4973/$ – see front matter
http://dx.doi.org/10.1016/j.fsigen.2014.01.009

ARTICLE INFO

Article history:
Received 16 November 2013
Received in revised form 30 December 2013
Accepted 27 January 2014

Keywords:
DNA analysis
Mixture interpretation
Autosomal STRs
Collaborative exercise
Spanish and Portuguese-Speaking Working Group

ABSTRACT

One of the main objectives of the Spanish and Portuguese-Speaking Group of the International Society for Forensic Genetics (GHEP-ISFG) is to promote and contribute to the development and dissemination of scientific knowledge in the area of forensic genetics. Due to this fact, GHEP-ISFG holds different working commissions that are set up to develop activities in scientific aspects of general interest. One of them, the Mixture Commission of GHEP-ISFG, has organized annually, since 2009, a collaborative exercise on analysis and interpretation of autosomal short tandem repeat (STR) mixture profiles. Until now, three exercises have been organized (GHEP-MIX01, GHEP-MIX02 and GHEP-MIX03), with 32, 24 and 17 participant laboratories respectively. The exercise aims to give a general vision by addressing, through the proposal of mock cases, aspects related to the edition of mixture profiles and the statistical treatment.

The main conclusions obtained from these exercises may be summarized as follows. Firstly, the data show an increased tendency of the laboratories toward validation of DNA mixture profiles analysis following international recommendations (ISO/IEC 17025:2005). Secondly, the majority of discrepancies are mainly encountered in stutters positions (53.4%, 96.0% and 74.9%, respectively for the three editions). On the other hand, the results submitted reveal the importance of performing duplicate analysis by using different kits in order to reduce errors as much as possible. Regarding the statistical aspect (GHEP-MIX02 and 03), all participants employed the likelihood ratio (LR) parameter to evaluate the statistical compatibility and the formulas employed were quite similar. When the hypotheses to evaluate the LR value were locked by the coordinators (GHEP-MIX02) the results revealed a minor number of discrepancies that were mainly due to clerical reasons. However, the GHEP-MIX03 exercise allowed the participants to freely come up with their own hypotheses to calculate the LR value. In this situation the laboratories reported several options to explain the mock cases proposed and therefore significant differences between the final LR values were obtained. Complete information concerning the background of the criminal case is a critical aspect in order to select the adequate hypotheses to calculate the LR value. Although this should be a task for the judicial court to decide, it is important for the expert to account for the different possibilities and scenarios, and also offer this expertise to the judge. In addition, continuing education in the analysis and interpretation of mixture DNA profiles may also be a priority for the vast majority of forensic laboratories.

1. Introduction

Currently, the analysis and interpretation of genetic profiles generated from autosomal short tandem repeat markers (STRs) is a well standardized practice in the field of forensics, which has been used by hundreds of laboratories around the world to aid the resolution of cases in the criminal and civil sphere [1–3], for several decades. At the same time, there have been important advances, namely technological and methodological improvements for the study and analysis of this type of markers (STRs). However, in certain complex DNA mixture profiles, where some of the components are found in limiting amounts or are affected by degradation, the interpretation and evaluation becomes a difficult task and not exempt of errors, and sometimes involving a subjective opinion due to such complexity.

Several scientific working groups have published recommendations and guidelines to address the analysis and assessment of this type of profile [4–11]. However, despite them, the interpretative difficulty and the lack of a single criterion constitute important challenges for the laboratories to cope with. Different computer programs have been developed over the past years (e.g. *DNAMIX* [12], *Grape* [13], *LRmix* [14,15]) that have favored, without any doubt, the evaluation of this type of profiles. Currently, most of the software are under a validation process by the forensic community.

GHEP-ISFG develops its scientific activity in various matters by organizing different collaborative exercises on specific issues (e.g. mtDNA, SNPs, mixture DNA analysis) which are coordinated by various working committees specifically created for that purpose. During the 2009 Annual Assembly of the GHEP-ISFG, members agreed upon the creation of a Commission (Mixture Commission of the GHEP-ISFG) with the aim of looking into the issue, which involves the analysis of mixture DNA profiles. Within the activities of this Commission, addressed to laboratories with GHEP-ISFG members, collaborative exercises based on mixture profiles are considered as an important tool toward standardization in the analysis and interpretation of this type of profile. The fourth edition of this type of exercise (results still in evaluation phase) was held during the year 2013. These exercises have been named with the acronym GHEP-MIX (GHEP-MIX01 in 2010, GHEP-MIX02 in 2011 and GHEP-MIX03 in 2012) for the first three editions.

The main objective of this exercise is to provide the opportunity for laboratories to contrast with each other their systematic analysis and interpretation of mixture profiles, as well as to check the statistical treatment used. Moreover, the exercise includes a didactic aspect, which tries to reveal some limiting factors in the interpretation of mixture profiles that can compromise the final result (proportion of contributors, thresholds values employed, approaches of hypothesis and mathematical treatment in the statistical assessment of the result). The exercises that were proposed had different designs, each one of them focusing on different basic aspects. This paper shows the results and conclusions that have been generated over the three editions of the exercise.

2. Materials and methods

2.1. Participants

This collaborative exercise was open to all laboratories with GHEP-ISFG members. Throughout the three editions carried out since 2009, a total of 43 laboratories from 12 different countries (Spain, Portugal, Costa Rica, Brazil, Argentina, Uruguay, Ecuador, Colombia, Chile, Italy, France and Venezuela), have been involved, many of which have participated in the 3 editions: 32 laboratories took part in the GHEP-MIX01; in the GHEP-MIX02, 24 laboratories; and in the latest edition of GHEP-MIX03, 17 laboratories. Although most of the participants developed their work in the criminal field,

there was a small group of them that exclusively performed paternity testing (4). The laboratories either belonged to public institutions (justice, health and university) (30) or to private companies (13).

2.2. Exercise scheme

The three exercises were organized and coordinated by the Mixture Commission of the GHEP-ISFG. All exercises had a common denominator, since in all cases raw data files (fsa format), containing the genetic typing of the samples under study, were made available to participating laboratories, as well as positive and negative controls, allelic ladders used in the electrophoresis and also matrix files for data analysis. Additionally, each exercise included a questionnaire where various issues were asked that dealt with the characteristics of the laboratory, as well as the technical criteria used in the interpretation of profiles and statistical estimation.

The structure of the three exercises varied slightly (Table 1). In the first exercise (GHEP-MIX01), organizers requested participants to report the results of two DNA mixtures under a dual perspective of profiles. Firstly, how the result would be informed in a judicial report and secondly, given the possibility of introducing the profile in a criminal database for investigation purposes, how this would be reported.

The second and third edition of the exercise (GHEP-MIX02 and GHEP-MIX03) included an additional block consisting of mock cases for evaluating issues related to the statistical treatment of the mixture profiles (calculation of LR values for a certain number of markers and the hypothesis approach).

2.3. Samples

A total of 9 samples were analyzed in the three editions of the exercise: 4 (GHEP-MIX01), 2 (GHEP-MIX02) and 3 (GHEP-MIX03) respectively. Mixtures were prepared artificially using DNA extracted from buccal epithelium samples provided by anonymous donors. DNA extracts had been previously quantified in duplicate (*Quantifiler® Duo*), in order to optimize as much as possible the correct ratio between the components of the mixture. The proportion and the number of contributors of the mixture samples were variable. Once the mixture sample was set up in the work proportion, the DNA extract was quantified again for estimating the optimal DNA input to be employed in the amplification reaction. These extracts were analyzed with different commercial kits depending on the edition of the exercise; therefore, all participants could freely analyze the kit or kits that are usually employed in their respective laboratories for the resolution of their routine forensic cases. It is also worth mentioning that the complexity of the sample mixture increased year after year, based on the relationship between contributors, the number of them and also depending on the specific profiles of the various components of the sample.

Table 1
General scheme of every collaborative exercise edition.

	GHEP-MIX01	GHEP-MIX02	GHEP-MIX03
General aspects. Participants characterization	X	X	X
Technical aspects. Edition tools	X	X	X
Statistical aspects	X	X	X
Results of mixture edition	X	X	X
Statistical treatment.		X	X
Mock cases			
Locked hypothesis		X	
Open hypothesis			X

In the GHEP-MIX01 exercise, the participants performed analysis for a total of 4 samples of two components: M1 (1:5, female–male), M2 (1:10, female–female), M3 (1:1, female–male) and M4 (5:1, female–male). The samples were typed with the following kits: *AmpFlSTR® Identifiler®* and *PowerPlex® 16 System*. In the second exercise (GHEP-MIX02), 2 mixture samples were analyzed, consisting of two and three contributors respectively, which were amplified exclusively with the *AmpFlSTR® Identifiler* kit®: M1 (1:5, male–female) and M2 (2:1:1, female–male–male). In the last edition (GHEP-MIX03), 3 samples were analyzed, which were tested for kits *AmpFlSTR® Identifiler® Plus* and *AmpFlSTR® NGM™*: M1 (1:5, female–male), M2 (1:10, female–female) and M3 (1:3:7, female–male–male).

2.4. Questionnaire design

Along with the electronic files (raw data), a questionnaire was also provided to the laboratories with the main goal of collecting information regarding general aspects related to the characteristics of the laboratory (e.g. type of routine cases analyzed in the laboratory, use of automated systems in the analysis of mixtures, validation of the method used to interpret profiles mixture), methodological issues that deal with the interpretation of the mixture profile (e.g. employment or not of threshold values and, if that is the case, values used for analytical, stochastic and stutter threshold, criteria for identification of a mixture profile, software used in the analysis of the profile), and also aspects related to the statistical treatment (use of the LR value or any other statistical parameter for statistical evaluation of the result, bibliographic references to carry out the statistical calculation; application or not of any software for the calculation of the LR). Regarding the results obtained for the different samples, the questionnaire also included tables for reporting their profiles.

With reference to the block concerning the statistical treatment and interpretation (GHEP-MIX02 and 03) the organizers proposed a number of practical cases, in which a hypothetical situation was described, and the genetic profiles from contributors to the mixture (victim and suspect/s) were also provided. Specifically, in the GHEP-MIX02 exercise, closed hypotheses were set out, and the participants only had to perform the statistical calculation of the partial LR for each STR marker, reporting the formulas usually used in their own laboratory.

However, in the GHEP-MIX03 exercise, the participating laboratories were free to outline their own hypotheses, which had to be justified. In both exercises, and to avoid excessive dispersion of the results generated by each laboratory, the organizers of the exercise provided participants with a table of allele frequencies to be used in the calculations.

2.5. Classification of discrepancies

To facilitate the analysis of the discrepancies reported by laboratories with respect to the expected result, they were classified into three groups, with variables in some of these groups (Slide 1, Appendix A):

Group A: those which take place at a stutter position ($n-4$, $n+4$ or $n-3$, $n+3$).
−A: No report of a real allele which is located at a stutter position, which represents a false allelic loss in the genetic profile.
+A: Report of a non-existent allele (actual stutter peak), which represents a false allelic gain in the genetic profile.
Group B: those which occur in a non-stutter position.
−B: No report of a real allele in a non-stutter position, which represents a false allelic loss in the genetic profile.

+B: Report of a non-existent allele in a non-stutter position, which represents a false allelic gain in the genetic profile.
Group C: discrepancies attributable to an incorrect transcription.

3. Results and discussion

3.1. Participants characterization

Extensive details about the questionnaire answers are given in Tables S1, S2 and S3, in the supplementary data of Appendix A.

Firstly, in accordance with the information provided through the questionnaire by the participating laboratories, the vast majority of them developed their activity both in the criminal and in the civil fields (paternity testing); GHEP-MIX01 −75%, GHEP-MIX02 −75% and GHEP-MIX03 −88.2%. On the other hand, in the first and second edition of the exercise the number of participating laboratories performing cases exclusively in civil matters was a minority, 9.4% and 12.5% respectively. In fact, these laboratories did not participate in the third edition.

Another question included in the questionnaire referred to the incorporation of genetic profiles in DNA databases. In the first two editions of the exercise, the great majority of the laboratories did not send their genetic profiles to any database (59.4% and 45.8%). However, in the last edition (GHEP-MIX03), these laboratories were just representing 23.5% (Table S1, Appendix A).

Participants were asked if in their routine casework results for mixture profiles were issued. Most of the laboratories answered this question affirmatively, 43.8%, 66.7%, and 70.6% for the three editions of the exercise, respectively. A minority of them claimed to carry out the interpretation of this type of profile exclusively in cases where reference samples were available (28.1%, 20.8% and 17.6%). On the other hand, laboratories were asked about what is done with such genetic mixture profiles, and most of the laboratories, in the three editions, replied that this kind of profile was only registered in the final report emitted to the court (46.9%, 41.7% and 47.1%). In the latest edition, 35.3% of participants apart from including the mixture DNA profile in the report, they would also send it to a national DNA database.

When participants were asked about how they had performed the allelic assignments of the mixture components throughout the whole three editions, only 19.2% had fulfilled the task by using a specific software, 43.8% had done it manually, and 34.2% had carried it out both manually as well as using a software. It is noteworthy that with regards to this question, an evolution was noticeable (Table S1, Appendix A). In the last edition of the exercise (GHEP-MIX03), only 11.8% of the participating laboratories would perform an allelic assignment manually, 35.3% corresponds to an exclusively automatic assignment, while 52.9% of the participants employed both strategies. This last information seems more consistent with the reality of the daily practice of a forensic laboratory.

Finally, the questionnaire of the exercise requested the participants if the criteria used to carry out the interpretation of mixture profiles had been validated according to international requirements [16,17]. The majority of the participating laboratories answered negatively (81.3, 87.5 and 29.4%, respectively) opposite to those which replied positively (15.6, 12.5 and 23.5%). However, the data show that in the third edition of the exercise there was a significant increase in the number of laboratories that were validating their methods of analysis for the study of mixture profiles. This fact gives an idea of the need and the awareness of the laboratories, to carry out a specific validation procedure for this type of analysis.

3.2. Profile characterization

3.2.1. Parameters used for the analysis of profiles

Table S2 (Appendix A) shows detailed information about this part of the questionnaire.

The questionnaire included a second block of questions related to technical and methodological aspects used to edit the mixture profiles of the exercise (i.e. threshold analysis used, software employed and criteria to characterize a profile as a mixture). Also, in this block of questions, participants were inquired about the main difficulties that they usually faced to carry out the interpretation of this type of mixture profiles.

Firstly, regarding the software used for editing the electropherograms (EPGs) of the samples, most of the participating laboratories employed one of the *GeneMapper*™ (Applied Biosystems) software versions. However, other programs were also used but in minor numbers (*GeneScan*®, *Genotyper*® or *Peak Scanner*™ software).

Concerning the criteria used to define profile as a mixture, there is great variability among the participating laboratories, and it is clear that there is no unique criterion (Table S2, Appendix A). In fact, the responses indicate that the majority of the laboratories did not employ a single criterion, but a combination of several to recognize a profile as a mixture. Most of the participants, namely 31.3%, 20.8% and 64.7% (GHEP-MIX01, 02 and 03, respectively) considered that two conditions should be met to characterize a profile as a mixture: on one hand, the presence of at least two genetic markers with at least 3 alleles each and; on the other hand, the existence of allele imbalance. This last criterion highlights the need for a validation of this type of profile in order to know what can help categorize a genetic profile as a mixture (e.g. heterozygous peak height ratios). However, it is well known that the use of thresholds for carrying out the evaluation of genetic profiles is not risk-free and it may sometimes lead to erroneous conclusions [9,18,19]. In spite of this, the use of thresholds helps laboratories make decisions when issuing a conclusion on a genetic profile. Accordingly, the laboratories were asked about the relative fluorescence units (RFU) value employed for the analytical threshold [9,20]. In this case, it seems that there was enough consensus among laboratories, since most of the participants in the three editions (65.5%, 75% and 88.2%) established that value at 50 RFUs.

It is very frequent that stutter positions in mixture profiles involve some interpretative difficulty because sometimes it is not easy to distinguish between a real allele or an artifact (stutter). In this regard, participants were questioned about when a possible peak detected in a stutter position was assigned as a true allele, regardless that these were in position $n-4$ ($n-3$ in the case of the D22S1045 marker). Once again in this question the variability of the responses from the participants (Table S2, Appendix A), showed the lack of a single criterion. Nevertheless, the main responses were: "Only those that exceed 15% of the main allele" – this represented 25% (GHEP-MIX01), 29.2% (GHEP-MIX02) and 11.8% (GHEP-MIX03), and "the assignment was variable depending on the STR marker", involving 40.6%, 37.5% and 47.1% respectively for the three editions. On the other hand, some participants employed different combinations of criteria simultaneously to be able to distinguish between real peaks and stutters (Table S2, Appendix A). Again, the answers highlight the important significant diversity of criteria used, that most certainly had an impact on the reported profile by the different laboratories on the three exercises. Finally, to close this section, when participants were asked about the main obstacles they encounter in interpreting mixed profiles, the answers denoted that the main problems were the lack of a single criterion within the same laboratory, as well as the lack of the necessary and suitable training.

Table 2
Discrepancies distribution on GHEP-MIX01 exercise. The keys for each group appear in point 2.5 in this paper.

	GHEP-MIX01 ($N=32$)									
	M1 (1:5)		M2 (1:10)		M3 (1:1)		M4 (1:5)		Total	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%
Group A										
−A	31	37.3	21	13.0	0	0.0	0	0.0	52	19.8
+A	10	12.0	69	42.6	5	71.4	4	40.0	88	33.6
Group B										
−B	1	1.2	70	43.2	0	0.0	0	0.0	71	27.1
+B	40	48.2	2	1.2	1	14.3	6	60.0	49	18.7
Group C										
C	1	1.2	0	0.0	1	14.3	0	0.0	2	0.8
Total	83	31.7	162	61.8	7	2.7	10	3.8	262	Discrepancies

3.2.2. Statistical aspects

The GHEP-MIX02 and GHEP-MIX03 exercises included a section consisting of carrying out a statistical assessment of the compatibility between undoubted genetic profiles (reference samples from the victim and/or suspect/s) and the mixture profiles corresponding to samples sent in the context of the mock cases. In this sense, following the recommendations of the ISFG [5], all participants employed the likelihood ratio statistic (LR) as the most appropriate approach for statistical evaluation. Taking into account this circumstance, one last question was outlined to the participating laboratories about how the calculation of the LR was performed (Table S3, Appendix A). A significant number of the participants (45.8% and 47.1%) exclusively employed commercial software (*GeneMapper® ID-X* Software, *Software package Grape*, *BDGen*), whereas the number of participants who used simultaneously manual and automatic methods corresponds to nearly a quarter of the laboratories (25% and 23.5%). Thus, laboratories reported that the formulas applied were taken mainly from Evett et al. [21], Weir et al. [22] and Gill et al. [5].

3.2.3. Discrepancies on edition

Tables 2–4 show detailed data regarding the types of discrepancies detected in the various editions of the exercise. The kind of mistake reported was varied and the percentage of their occurrence depended on different factors (number, proportion and genotype of contributors and usage or not of confirmation through various reviews of the same marker with different kits).

3.2.3.1. Collaborative exercise GHEP-MIX01. In the first edition of exercise, for the four samples analyzed, discrepancies of type A (position stutter) and B (position not stutter) were similarly distributed, 53.4% and 45.8% respectively. Considering the different type of profiles analyzed, most discrepancies of the exercise focused on M2 (61.8%), with a contributors ratio of 1:10. In this particular case, the main kind of discrepancies consisted in the wrongful inclusion of one allele in a stutter position (A) (42.6%), and a wrongful exclusion in a non-stutter position (B) (43.2%). Only for the markers Penta E and Penta D, was a consensus result reported for M2 by the participants. In the case of the STR Penta E, the result 8–9 was erroneously agreed since the expected result was 8-9-18 (Slide 2, Appendix A) and the allele 18 passed unnoticed for all participants (<50 RFU). In the case of the marker Penta D, the consensus value was the correct one (12–13). The result for this marker offered no problems for any of the participants since both contributors of the mixture had an identical genotype (12–13). Meanwhile, for the remaining markers in M2, at least two different results were collected. In the STR markers D21S11, CSF1PO, D3S1358, D16S539, D2S1338, D18S51 and D5S818, participants issued at least 4 different results (see Table 5).

Mixture samples M1 and M4 had a 1:5 contributor ratio, whereas sample M3 was 1:1. Samples M3 and M4 revealed a low discrepancy rate, 2.7% and 3.8% respectively. However, M1 (1:5) accumulated 31.7% of discrepancies from the whole exercise, mainly of the type −A (37.3%) and +B (48.2%). The participants reported unanimous results only for 4 of the analyzed 17 STR markers (D3S1358, D13S317, D2S1338, TPOX), and more discrepancies emerged in the STR D8S1179 where up to 4 different results were reported by the participants (Slide 3, Appendix A). Nearly all participants performed simultaneously the analyses of markers included in the two kits provided by the organizers (*AmpFlSTR® Identifiler®* and *PowerPlex® 16 System*). However, 4 of the participating laboratories analyzed exclusively *PowerPlex® 16 System* kit. These participants did not detect the presence of one of the alleles at two markers, the allele 6 (HUMTH01) and the allele 13 (D16S539) (Slide 4, Appendix A). Another interesting finding was presented at the marker Penta E. Only 2 laboratories of all

Table 3
Discrepancies distribution on GHEP-MIX02 exercise. The keys for each group appear in point 2.5 in this paper.

	GHEP-MIX02 ($N=24$)					
	M1 (1:5)		M2 (2:1:1)		Total	
	n	%	*n*	%	*n*	%
Group A						
−A	4	44.4	1	6.3	5	20.0
+A	4	44.4	15	93.8	19	76.0
Group B						
−B	0	0.0	0	0.0	0	0.0
+B	0	0.0	0	0.0	0	0.0
Group C						
C	1	11.1	0	0.0	1	4.0
Total	9	36.0	16	64.0	25	Discrepancies

Table 4
Discrepancies distribution on GHEP-MIX03 exercise. The keys for each group appear in point 2.5 in this paper.

	GHEP-MIX03 ($N=17$)							
	M1 (1:5)		M2 (1:10)		M3 (1:3:7)		Total	
	n	%	*n*	%	*n*	%	*n*	%
Group A								
−A	2	12.5	82	61.7	44	59.5	128	57.4
+A	7	43.8	7	5.3	25	33.8	39	17.5
Group B								
−B	4	25.0	44	33.1	2	2.7	50	22.4
+B	2	12.5	0	0.0	0	0.0	2	0.9
Group C								
C	1	6.3	0	0.0	3	4.1	4	1.8
Total	16	7.2	133	59.6	74	33.2	223	Discrepancies

Table 5
Distribution of different results obtained in each mixture for the samples analyzed in each GHEP-MIX exercise. Samples in proportion 1:10 concentrated the largest number of different results reported by laboratories. Also, results of GHEP-MIX03 exercise show that exclusive STR markers of a kit (e.g. D10S1248, D22S1045, D1S1656, D12S391, D2S441) and, therefore, non-replicated with another kit, also accumulate a great disparity of results. Key: (−), not analyzed.

	GHEP-MIX01				GHEP-MIX02		GHEP-MIX03		
	M1	M2	M3	M4	M1	M2	M1	M2	M3
	1:5	1:10	1:1	1:5	1:5	2:1:1	1:5	1:10	1:3:7
D8S1179									
D21S11									
D7S820									
CSF1PO									
D3S1358									
TH01									
D13S317									
D16S539									
D2S1338									
D19S433									
vWA									
TPOX									
D18S51									
D5S818									
FGA									
Penta E					-	-	-	-	-
Penta D					-	-	-	-	-
D10S1248	-	-	-	-	-	-			
D22S1045	-	-	-	-	-	-			
D1S1656	-	-	-	-	-	-			
D12S391	-	-	-	-	-	-			
D2S441	-	-	-	-	-	-			

Color key: 1 result | 2 different results | 3 different results | more than 3 results

participants reported the correct result as expected on the basis of the contributors of the mixture (12-13-14-17). Allele 13 was omitted for the most part of them (21 of 24) for being misidentified with a stutter (Slide 5, Appendix A).

Another issue that is worth mentioning and that was evident with this first M1, was the different behavior that different kits can present with the same mixture sample concerning the imbalance of the contributors (Slide 6, Appendix A). In certain situations, this issue can mislead us to an erroneous assignment regarding the proportion of each contributor in a mixture profile. In consequence, it might affect the final conclusions deduced during the evaluation of the results and so this may indicate the strong convenience of employing different commercial kits to ensure certain results and to confirm in this way the outcome of certain mixture profiles.

3.2.3.2. Collaborative exercise GHEP-MIX02. In this edition of the exercise, the number of discrepancies reported by the participants was much lower (25). The only kind of discrepancies was of type A, specifically +A (76%) and −A (20%). Sample M1 (1:5) showed fewer discrepancies regarding the expected result (36%), compare to 64% for M2 (2:1:1). Most of the participants reported the correct result, 75% (M1) and 87.5% (M2). On the other hand, two laboratories which developed their activity exclusively in paternity cases, accumulated most of the discrepancies: 72% and 89% respectively for both samples, M1 and M2.

3.2.3.3. Collaborative exercise GHEP-MIX03. In the same way as in collaborative exercise GHEP-MIX01, the majority of discrepancies was detected in the sample with a 1:10 ratio (M2), accumulating 59.6% of them. Sample M3 (1:3:7) mainly encountered 33.2% of discrepancies, while only 7.2% were in sample M1 (1:5).

Again, in the third edition, discrepancies which predominated were of type A (74.9%) most of them −A (57.4%) focused in the sample M2 (1:10), specifically on the markers D2S1388, FGA, TH01, D1S1656 and D12S391 (*AmpFlSTR® NGM*). For these markers, more than 70% of the participants did not detect any of the alleles expected in the mixture profile (Slide 7, Appendix A). For marker D1S1656, detection of the micro variable 16.3 allele was omitted, since in that profile an allele 17 also appeared. Therefore, the proximity of both alleles made the distinction between them difficult, making allele 16.3 unnoticed. In relation to allele 19 of D12S391, present in one of the contributors of the mixture, it was not detected by any participant. As it is known, the STRs D1S1656 and D12S391 markers are exclusive of *AmpFlSTR® NGM* kit, so participants could not have a second analysis with another kit so as to accept or discard the definitive result of the marker. However, the discrepancy rate (type −A) in the analysis of shared STR markers with the two studied kits was significantly lower when the analysis was carried out with the *AmpFlSTR Identifiler® Plus* kit than when they were analyzed with the *AmpFlSTR® NGM* kit (D2S1388, FGA and TH01) (Slide 8, Appendix A). In this way, the rate of discrepancies for the TH01 marker varied from 87.5% to

29.4%, while for FGA produced a decrease of 75–11.76% and for D2S1388 of 68.75–11.76%. These data, again, indicate the different behavior of different commercial kits to the same questioned sample and how the use of both may turn out to be very convenient.

3.3. Statistical treatment evaluation

Two editions of the exercise (GHEP-MIX02 and GHEP-MIX03) incorporated a block where it was intended to address statistical issues in the assessment of profile mixtures. The objectives in each issue were different, and so the approach was also distinct.

3.3.1. Locked hypothesis (GHEP-MIX02)

In the GHEP-MIX02 exercise, the material that was distributed to the participants consisted of a partial genetic profile (only 6 STR markers) in a *PDF* file, as well as a table containing the genetic profile of a hypothetical suspect and a population allele frequency table [23]. With that information, coordinators requested laboratories to calculate the LR value (likelihood ratio) in two different scenarios, for which the prosecution hypothesis (Hp) and the defense hypothesis (Hd) were provided and proposed by the coordinators of the exercise.

This exercise sought to minimize as much as possible the variables which could cause dispersion in the final data. The main objective of this exercise was simply to evaluate the LR value and show the formulas employed and deductions carried out for this purpose.

The locked hypotheses proposed were as follows:

First scenario:
Hp: The suspect and another unknown person provide the genetic material of the sample.
Hd: Two unrelated unknown individuals provide the genetic material of the sample.
Second scenario:
Hp: The suspect and the victim provide the genetic material of the sample.
Hd: The victim and an unrelated unknown individual provide the genetic material of the sample.

As expected, in both theoretical cases a high degree of consensus was achieved (Table S5, Appendix A), 78.3% (18 over 23 participants) for the first scenario and 73.9% (17 over 23 participants) for the second one. Laboratories that agreed on the result of the partial LR reported had used the formulas developed by Weir et al. [22] (Table S5, Appendix A). There are basically two reasons that explain the non consensus results. Firstly, a block of discrepancies was attributable to erroneous transcriptions of the provided allele frequencies thus affecting some of the partial LR values. This type of error was detected in one participant (scenario 1) and in two participants (scenario 2). The second reason, there were also errors attributable to improper use of the formulae employed for the calculation of the partial LR, specifically in 3 laboratories (scenario 1) and 3 laboratories (scenario 2). It is worth pointing out that the errors were mainly encountered in 5 laboratories, being the same participants those who made the same mistakes for both scenarios.

3.3.2. Open hypothesis (GHEP-MIX03)

In the third edition of the exercise, similar material of the previous edition was provided to the participating laboratories: partial genetic profile (only 6 STR markers), a table containing the genetic profile of two possible suspects and an allele frequency table from the Spanish population [23]. This exercise focused its interest on getting to know how the laboratories carry out the approach of freely making the hypotheses and know how the LR value calculation is carried out.

The first of the two cases proposed a scenario in which a condom containing traces of semen had been found in the house of a victim of a sexual assault, generating a mixture profile as a result of the analysis. Participants were required to weigh the contribution of two suspects by calculating the LR value. The second mock case proposed a different scenario, in this case, a condom was found at the scene of crime. From its external side a mixture profile was obtained. In addition, the suspect and the victim's profile are also available. The laboratories were asked if the suspect was involved in the attack, requiring the calculation of the value of LR.

3.3.2.1. Mock case 1. With respect to the first hypothetical case, all participants except one carried out a statistical assessment of the compatibility of the suspects regarding the analyzed mixture. According to the type of hypothesis approaches issued by the participating laboratories, they were classified into three groups (Table 6). A first group consisting of 5 laboratories which considered as the only option the joint participation of the two suspects by proposing to this pair of hypotheses Hp = S1+S2 and Hd = 2U (unknown) for the calculation of the LR value. In contrast, the larger group of participants (11) reported 3 different LR values, taking into account different pairs of hypotheses. These participants argued that depending on other data concerning the background of the case, the use of one or another approach could be more indicated. However, in the absence of these data, the participants would decide which of the more reasonable options would be addressed in the final report. Thus, like the first group of participants, they evaluated the possibility of joint participation of the two suspects (Hp = S1+S2, Hd = 2U). However, they also considered weighing separately the contribution of each suspect, proposing to suspect 1 this pair of hypotheses Hp = S1+U and Hd = 2U and this one for suspect 2, Hp = S2+U and Hd = 2U. Only one participant considered the unique option of evaluating the LR for each of the two suspects separately. For this participant the joint assessment of both suspects in the Hp against two unknown persons (Hd) was not conservative. In addition, this approach increases artificially and erroneously the LR value that would correspond to the participation of every suspect if it was made separately. In fact, in the proposed case the laboratories obtained a consensus LR value of 1.12×10^{11} (Hp = S1+S2 and Hd = 2U) against 1.48×10^{5} (Hp = S1+U and Hd = 2U) values for suspect 1 or 3.14×10^{3} (Hp = S2+U and Hd = 2U) for suspect 2 (Table S7, Appendix A).

56.3% of the participants (Table S7, Appendix A) whom in some of their options considered the joint assessment of both suspects, agreed upon a final LR value (1.12×10^{11}). However, some laboratories showed final LR values discrepant in various degrees of magnitude with respect to the consensus (minimum value 6.17×10^{10} and maximum value of 2.64×10^{17}). These discrepant values were promptly focused on two participants. The cause seems attributable to errors in the use of the formulas employed or the wrong transcription of allele frequencies.

More agreement (83.3%) was obtained on the value of the final LR (1.48×10^{5}) when the participation of suspect 1 was assessed separately to the contribution of the mixture profile problem. Again, deviations from the consensus LR value were focused in a few (2) laboratories. These deviations were due to changes in the value of some partial LR for a certain marker, which resulted in changes in the final LR value. Errors, as in the previous case, were mainly encountered in the same participating laboratories. Furthermore, in assessing suspect 2 separately, the value of the LR consensus was 3.14×10^{3} and it was reported by 50% of the participants. In this situation the LR values ranged from 7.74×10^{2} to 3.8×10^{8}. More details can be seen in Table S7 (Appendix A).

Table 6
Descriptive statistics for the distribution of established hypotheses by participants on GHEP-MIX03 exercise. Keys: Hp, prosecution hypothesis; Hd, defense hypothesis; S, suspect; V, victim; U, unknown.

Case 1 (N = 17)			Case 2 (N = 17)		
Hypothesis couples (LR = H_p/H_d)	n (labs)	%	Hypothesis couples (LR = H_p/H_d)	n (labs)	%
2S/2U	5	29.4	(S+V)/(V+U)	8	47.1
(2S/2U) + [(S1+U)/2U] + [(S2+U)/2U]	11	64.7	[(S+V)/(V+U)] + [(S+V)/2U]	2	11.8
[(S1+U)/2U] + [(S2+U)/2U]	1	5.9	[(S+V)/(V+U)] + [(S+V)/2U] + [(S+V)/(S+U)]	1	5.9
			(S+V)/2U	3	17.6
			[(S+V)/(V+U)] + [(S+V)/2U] + [(V+U)/2U]	1	5.9
			[(S+V)/(V+U)] + [(S+V)/2U]+ [(S+U)/2U]	1	5.9
			[(S+U)/2U] + [(V+U)/2U]	1	5.9

3.3.2.2. Mock case 2. The second case (details in Table 6 and Table S8, Appendix A) generated a variety of proposals with respect to the assumptions made. Participants reported a total of 5 different combinations of hypotheses to explain the case proposed. As in the first mock case, several participants considered various possibilities simultaneously. A total of 13 laboratories out of 17, considered the Hp = S+V and Hd = U+V hypotheses as the best that explained the situation that arose. For this pair of hypotheses, 84.6% of the laboratories (11) agreed on an LR value of 3.52×10^7, however, minor variations were observed with respect to that value (3.30×10^7 and 6.96×10^8) for the remaining two labs. On the other hand, 8 out of 17 participants considered that it was possible to explain the scenario proposed in this case through this pair of hypotheses Hp = S+V and Hd = 2U. In this case, all but one laboratory obtained a consensus result of 1.12×10^{11}. Some examples of justifications for these hypotheses were based on statements of the type "*Being a sign not coming from the body's victim we might consider as disputed both the victim and the suspect*" or "*The prosecution will present the hypothesis that the observed remains belong to the victim and the suspect while the defense will question the suspect intervention attempting to show that the profile mix detected has been caused by two unknown*". Although with minor representation, other pairs of hypotheses were raised by the participants to explain this case (e.g. Hp = S+V and Hd = S+U, Hp = S+V and Hd = 2U, Hp = S+U and Hd = 2U, Hp = V+U and Hd = 2U).

4. Conclusion

To our knowledge, this is the first collaborative exercise that offers you to perform a complete evaluation of editing, interpretation and statistical assessment of DNA mixture profiles. The training stage in the forensic field is essential in order to thrive and transmit with certainty the result of the analysis to the courtroom. The analysis and interpretation of profiles is one of the fields of forensics that currently raises greater interest in the forensic community [4–11,24,25]. In our opinion, the main objective of training, initially delineated by the GHEP-ISFG Mixture Commission, has been fulfilled. The different degrees in complexity of samples analyzed in the various exercises has supplied a collection of data and experiences that offer forensic laboratories a useful tool in the analysis of mixture DNA profiles.

The data collected in this study point to the need that the forensic laboratories, as part of ISO/IEC 17025:2005 [17], must validate the methods used for analysis by the laboratory and, in particular, for the analysis of mixture profiles. It is well known that behavior of mixture profiles at different phases of analysis (amplification, electrophoresis, results generation, interpretation and editing) is completely different from that of single profiles. The data points out to a growing trend toward the validation of this type of analysis, although often this type of validation is carried out in laboratories with "controlled" samples which sometimes do not accurately reflect the reality of samples received in forensic cases. It becomes especially noticeable that the highest rate of discrepancies is spotted in stutter positions, where alleles are assigned or omitted erroneously (error type A). A thorough knowledge of the kits used as well as a strong internal validation by the laboratory could significantly reduce these kinds of discrepancies.

Similarly, these data allow us to conclude that concerning the disproportion between the contributors of a mixture, the greater it is, the higher the error rates are. The samples at 1:10 proportions present more than 60% of the total number of discrepancies for each respective exercise. However, for a contributor ratio at 1:5 total error rates were variable, ranging from 36% (GHEP-MIX02) to 3.8% (GHEP-MIX01). These data highlight the importance of the genotypic combination of the contributors to the mixture's profile.

Due to the different behavior of different commercial kits with the same DNA extract, the results obtained in some of the exercises emphasize the need to analyze the sample using more than one kit, in order to be able to duplicate the result for confirmatory purposes.

With respect to the statistical treatment section, some interesting conclusions can also be drawn. All laboratories used the LR value as the statistical parameter and the formulas used to carry out such statistical treatment were similar [5,21,22]. The organizers' proposal of the employment of fixed hypotheses to assess the value of the LR in the hypothetical cases (GHEP-MIX02) generated a broad consensus in the reported data. The discrepancies observed in the calculation of the LR were attributable to transcription errors of the data (clerical mistakes) (e.g. values of allele frequencies) and were mainly encountered in a few laboratories (2) anyway. However, when the approach to formulate the hypotheses was left open to enable the calculation of the LR value (according to each laboratory's criteria), the results obtained were much more scattered and quite varied. In this sense, detailed knowledge of the background of the case will certainly allow to adjust more accurately the assumptions to work on. In the absence of this information, it seems advisable to offer to the judge different possibilities of assumptions and hypotheses, in cases where these have not already been pinpointed by the prosecutor or defense.

At this point, the results shown in this survey indicate the need and the importance of continuing training in terms of analysis and interpretation of mixture profiles. This has been continuously pointed out by several international institutions and organizations (NIST, ENFSI, ISFG, GHEP-ISFG) which have been permanently contributing to this respect in the recent years.

Conflict of interest statement

None declared.

Acknowledgments

The coordinators of the exercises would like to express their sincere gratitude to every participating laboratory for their interest. Besides, they would like to express their sincere gratitude to INTCF (National Institute of Toxicology and Forensic Science, Department of Barcelona, Spain) and GHEP-ISFG for their institutional support. This gratitude is also directed especially to Marta Federico, Estel Enreig and Georgina Marcos for their help during the processing of the data. The authors wish to thank the reviewers for their commentaries, in order to improve the contents of the original manuscript.

IPATIMUP is an Associate Laboratory of the Portuguese Ministry of Science, Technology and Higher Education and is partially supported by FCT (Foundation for Science and Technology).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.fsigen.2014.01.009.

References

[1] Z.M. Budimlija, M.K. Prinz, A. Zelson-Mundorff, J. Wiersema, E. Bartelink, G. MacKinnon, B.L. Nazzaruolo, S.M. Estacio, M.J. Hennessey, R.C. Shaler, World Trade Center human identification project: experiences with individual body identification cases, Croat. Med. J. 44 (3) (2003) 259–263.

[2] M.D. Coble, O.M. Loreille, M.J. Wadhams, S.M. Edson, K. Maynard, C.E. Meyer, H. Niederstätter, C. Berger, B. Berger, A.B. Falsetti, P. Gill, W. Parson, L.N. Finelli, Mystery solved: the identification of the two missing Romanov children using DNA analysis, PLoS ONE 4 (3) (2009) e4838, http://dx.doi.org/10.1371/journal.pone.0004838.

[3] Y.J. Deng, Y.Z. Li, X.G. Yu, L. Li, D.Y. Wu, J. Zhou, T.Y. Man, G. Yang, J.W. Yan, D.Q. Cai, J. Wang, H.M. Yang, S.B. Li, J. Yu, Preliminary DNA identification for the tsunami victims in Thailand, Genomics Proteomics Bioinformatics 3 (3) (2005) 143–157.

[4] T.M. Clayton, J.P. Whitaker, R. Sparkes, P. Gill, Analysis and interpretation of mixed forensic stains using DNA STR profiling, Forensic Sci. Int. 91 (1) (1998) 55–70.

[5] P. Gill, C.H. Brenner, J.S. Buckleton, A. Carracedo, M. Krawczak, W.R. Mayr, N. Morling, M. Prinz, P.M. Schneider, B.S. Weir, DNA commission of the International Society of Forensic Genetics: Recommendations on the interpretation of mixtures, Forensic Sci. Int. 160 (2–3) (2006) 90–101. , http://dx.doi.org/10.1016/j.forsciint.2006.04.009.

[6] P. Gill, R.M. Brown, M. Fairley, L. Lee, M. Smyth, N. Simpson, B. Irwin, J. Dunlop, M. Greenhalgh, K. Way, E.J. Westacott, S.J. Ferguson, L.V. Ford, T. Clayton, J. Guiness, National recommendations of the Technical UK DNA working group on mixture interpretation for the NDNAD and for court going purposes, Forensic Sci. Int. Genet. 2 (1) (2008) 76–82. , http://dx.doi.org/10.1016/j.fsigen.2007.08.008.

[7] P.M. Schneider, R. Fimmers, W. Keil, G. Molsberger, D. Patzelt, W. Pflug, T. Rothämel, H. Schmitter, H. Schneider, B. Brinkmann, The German Stain Commission: recommendations for the interpretation of mixed stains, Int. J. Legal Med. 123 (1) (2009) 1–5. , http://dx.doi.org/10.1007/s00414-008-0244-254.

[8] P. Stringer, J.W. Scheffer, P. Scott, J. Lee, R. Goetz, V. Ientile, C. Eckhoff, G. Turbett, D. Carroll, S. Harbison, Interpretation of DNA mixtures—Australian and New Zealand consensus on principles, Forensic Sci. Int. Genet. 3 (2) (2009) 144–145. , http://dx.doi.org/10.1016/j.fsigen.2008.09.003.

[9] B. Budowle, A.J. Onorato, T.F. Callaghan, A. Della Manna, A.M. Gross, R.A. Guerrieri, J.C. Luttman, D.L. McClure, Mixture interpretation: defining the relevant features for guidelines for the assessment of mixed DNA profiles in forensic casework, J. Forensic Sci. 54 (4) (2009) 810–821. , http://dx.doi.org/10.1111/j.1556-4029.2009.01046.

[10] A.J. Meulenbroek, T. Sijen, C.C.G. Benschop, A.D. Kloosterman, A practical model to explain results of comparative DNA testing in court, Forensic Sci. Int. Genet. Suppl. Ser. 3 (1) (2011) e325–e326. , http://dx.doi.org/10.1016/j.fsigss.2011.09.025.

[11] C.C.G. Benschop, H. Haned, T.J.P. de Blaeij, A.J. Meulenbroek, T. Sijen, Assessment of mock cases involving complex low template DNA mixtures: a descriptive study, Forensic Sci. Int. Genet. 6 (6) (2012) 697–707. , http://dx.doi.org/10.1016/j.fsigen.2012.04.007.

[12] DNAmix: http://www.biostat.washington.edu/∼bsweir/DNAMIX3/webpage/.

[13] GRAPE: http://dna-soft.com/.

[14] H. Haned, Forensim: an open-source initiative for the evaluation of statistical methods in forensic genetics, Forensic Sci. Int. Genet. 5 (4) (2011) 265–268. , http://dx.doi.org/10.1016/j.fsigen.2010.03.017.

[15] P. Gill, H. Haned, A new methodological framework to interpret complex DNA profiles using likelihood ratios, Forensic Sci. Int. Genet. 7 (2) (2013) 251–263. , http://dx.doi.org/10.1016/j.fsigen.2012.11.002.

[16] Council framework Decision 200 9/905/JHA of 30 November 2009 on Accreditation of forensic service providers carrying out laboratory activities, Off. J. Eur. Union L 322 (2009) 14–16, Available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L: 2009:322:0014:0016:EN:PDF.

[17] ISO/IEC 17025, General Requirements for the Competence of Testing and Calibration Laboratories, 2005.

[18] SWGDAM Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories, 2010 http://www.fbi.gov/about-us/lab/biometric-analysis/codis/swgdam.pdf.

[19] J.A. Bright, J. Turkington, J. Buckleton, Examination of the variability in mixed DNA profile parameters for the identifiler multiplex, Forensic Sci. Int. Genet. 4 (2) (2010) 111–114. , http://dx.doi.org/10.1016/j.fsigen.2009.07.002.

[20] J.M. Butler, Validation Workshop. HID University/Future Trends in Forensic DNA Technology, 2006, http://www.cstl.nist.gov/strbase/pub_pres/ValidationWorkshop_May2006.pdf.

[21] I.W. Evett, C. Buffery, G. Willot, D. Stoney, A guide to interpreting single locus profiles of DNA mixtures in forensic cases, J. Forensic Sci. Soc. 31 (1) (1991) 41–47.

[22] B.S. Weir, C.M. Triggs, L. Starling, K.A.J. Stowell, J. Buckleton, Interpreting DNA mixtures, J. Forensic Sci. 42 (2) (1997) 213–222.

[23] O. García, J. Alonso, J.A. Cano, R. García, G.M. Luque, P. Martín, I. Martínez de Yuso, S. Maulini, D. Parra, I. Yurrebaso, Population genetic data and concordance study for the kits Identifiler, NGM, PowerPlex ESX 17 System and Investigator ESSplex in Spain, Forensic Sci. Int. Genet. 6 (2) (2012) e78–e79. , http://dx.doi.org/10.1016/j.fsigen.2011.05.010.

[24] P. Gill, L. Gusmão, H. Haned, W.R. Mayr, N. Morling, W. Parson, L. Prieto, M. Prinz, H. Schneider, P.M. Schneider, B.S. Weir, DNA commission of the International Society of Forensic Genetics: recommendations on the evaluation of STR typing results that may include drop-out and/or drop-in using probabilistic methods, Forensic Sci. Int. Genet. 6 (6) (2012) 679–688. , http://dx.doi.org/10.1016/j.fsigen.2012.06.002.

[25] J.M. Butler, M.D. Coble, R.W. Cotton, B.J. Heidebrecht, C.J. Word, DNA Analyst Training on Mixture Interpretation, National Institute of Standards and Technology (NIST), 2013, http://www.nist.gov/oles/forensics/dna-analyst-training-on-mixture-interpretation.cfm.