Forensic Science International: Genetics 21 (2016) 45–53

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Research paper

GHEP-ISFG collaborative simulated exercise for DVI/MPI: Lessons learned about large-scale profile database comparisons

Carlos M. Vullo[a,*], Magdalena Romero[a], Laura Catelli[a], Mustafa Šakić[b], Victor G. Saragoni[c], María Jose Jimenez Pleguezuelos[d], Carola Romanini[a], Maria João Anjos Porto[e], Jorge Puente Prieto[f], Alicia Bofarull Castro[g], Alexis Hernandez[h], María José Farfán[i], Victoria Prieto[j], David Alvarez[k], Gustavo Penacino[l], Santiago Zabalza[m], Alejandro Hernández Bolaños[n], Irati Miguel Manterola[o], Lourdes Prieto[p], Thomas Parsons[b]

[a] Equipo Argentino de Antropología Forense (EAAF), Córdoba, Argentina
[b] International Commission for Missing Persons (ICMP), USA
[c] Unidad de Genética Forense, SML, Santiago de Chile, Chile
[d] Unitat Central de Laboratori Biològic, Divisió de Policia Científica, Policia de la Generalitat—Mossos d'Esquadra, BCN, Spain
[e] Instituto Nacional de Medicina Legal e Ciências Forenses, I.P., Coimbra, Portugal
[f] LabGenetics, Madrid, Spain
[g] Instituto Nacional de Toxicología y Ciencias Forenses (INTCF), Barcelona, Spain
[h] Instituto Nacional de Toxicología y Ciencias Forenses, (INTCF), Canarias, Spain
[i] Instituto Nacional de Toxicología y Ciencias Forenses, (INTCF), Madrid, Spain
[j] Instituto Nacional de Toxicología y Ciencias Forenses, (INTCF), Sevilla, Spain
[k] Laboratorio de ADN, Unidad Central de Análisis Científicos, Comisaría General de Policía Científica, Madrid, Spain
[l] Unidad de Análisis de ADN (COFyBCF), CABA, Argentina
[m] Unidad de Laboratorio NASERTIC & Policia Foral, Navarra, Spain
[n] Análisis ADN Forense, Depto. Ciencias Forenses, Poder Judicial, Costa Rica
[o] Servicio de Genómica—SGIker–Universidad del País Vasco (UPV/EHU), Spain
[p] Instituto de Ciencias Forenses, Grupo de Medicina Xenómica, Universidade de Santiago de Compostela, Spain

ARTICLE INFO

Article history:
Received 21 September 2015
Received in revised form 13 November 2015
Accepted 17 November 2015
Available online 8 December 2015

Keywords:
Disaster victim identification
DVI
Missing person identification
MPI
Simulated identification exercise
Database comparisons

ABSTRACT

The GHEP-ISFG Working Group has recognized the importance of assisting DNA laboratories to gain expertise in handling DVI or missing persons identification (MPI) projects which involve the need for large-scale genetic profile comparisons. Eleven laboratories participated in a DNA matching exercise to identify victims from a hypothetical conflict with 193 missing persons. The post mortem database was comprised of 87 skeletal remain profiles from a secondary mass grave displaying a minimal number of 58 individuals with evidence of commingling. The reference database was represented by 286 family reference profiles with diverse pedigrees. The goal of the exercise was to correctly discover re-associations and family matches. The results of direct matching for commingled remains re-associations were correct and fully concordant among all laboratories. However, the kinship analysis for missing persons identifications showed variable results among the participants. There was a group of laboratories with correct, concordant results but nearly half of the others showed discrepant results exhibiting likelihood ratio differences of several degrees of magnitude in some cases. Three main errors were detected: (a) some laboratories did not use the complete reference family genetic data to report the match with the remains, (b) the identity and/or non-identity hypotheses were sometimes wrongly expressed in the likelihood ratio calculations, and (c) many laboratories did not properly evaluate the prior odds for the event. The results suggest that large-scale profile comparisons for DVI or MPI is a challenge for forensic genetics laboratories and the statistical treatment of DNA matching and the Bayesian framework should be better standardized among laboratories.

* Corresponding author at: EAAF Forensic Genetics Laboratory, Edificio EME-1, Independencia 644-4A, 5000 Cordoba, Argentina. Fax: +54 351 4240418.
E-mail addresses: cvullo@yahoo.com.ar, carlosvullo@gmail.com (C.M. Vullo).

http://dx.doi.org/10.1016/j.fsigen.2015.11.004
1872-4973/

1. Introduction

Human disaster victim identification (DVI) or large-scale missing persons identification (MPI) represents a challenge for forensic sciences. There are several recommendations which include different practices about how to proceed in the general framework for MPI [1,2], or DVI [3–5]. Natural [6,7] or unnatural disasters [8,9] can suddenly result in tens to thousands of victims, often overwhelming or greatly straining the available forensic services. Unlike DVI in which the episode occurs suddenly, in MPI hundreds to thousands of disappearances may accumulate over a period of time which may vary for different reasons: political violence or wars [8,9], enforced disappearances for political reasons [10,11], people forced into exile or migration [12,13]. Different factors may have an influence on the complexities of DVI or MPI. Whether the event is "open"[1] or "closed", the degree of disarticulation or fragmentation of the remains, and the availability of reference samples to be compared with the human remains are some of these variables [1]. On the other hand, the presence of family-related victims in the same episode may hinder individual identification as it decreases the power of 'discrimination to differentiate between relatives [14]. The degree of preservation of human remains may also influence the ability to produce full genetic profiles. Degraded, decomposed or burnt remains may yield partial profiles as a result of dropout, thus producing a lower power of discrimination of the genetic systems [15–18].

The type and characteristics of the ante-mortem reference samples (AM) to be used for the identification of victims may also have an impact on the power of discrimination [5,14,19]. There are two kinds of AM samples for human identification: (a) samples belonging to the victim or (b) samples from the victim's relatives. The possibility of having AM genetic profiles generated from the victim allows direct genetic comparisons with post-mortem sample profiles (PM), resulting in strong statistical significance and a high probability of identity. However, the possibility of collecting AM reference samples from the victim may not be possible in many MPI events, especially when a long period of time has passed after the disappearance. For these reasons, the use of victim's relatives as AM samples for kinship analysis is usually the preferred option in DVI or MPI, and is commonly used to validate direct reference samples when they are available. Kinship matching adds complexity and difficulty to the database comparisons because of different reasons: the quality of the family pedigrees (first degree vs second/third degree of the victim's relatives), the number of family members the pedigree is composed of, the possibility of mutational events as well as the higher complexity of the kinship's statistical calculations to analyze the match [19–21]. Additionally, differences between the actual genetic relationships and the reported genetic relationships of reference individuals may be found in nuclear families, making it difficult for forensic genetics laboratories to find a match, and also presenting ethical challenges for the agencies involved in victims' identification [3–5,22].

Once the genetic profiles of AM and PM samples have been gathered in proper databases, large-scale AM and PM genetic profile comparisons are possible, but effective detection of matches requires appropriate systematic approaches. Firstly, PM to PM profiles comparisons need to be done in cases of commingled remains in order to try to re-associate the human remains of each individual. Secondly, it is good practice to compare AM to AM profiles to validate the reported kinship relationships among the reference individuals, allowing detection of problems either in documentation or in unexpected biological relationships (e.g., non-paternity) that could complicate or obviate the ability to find matches. Finally, PM to AM profiles comparisons have to be performed in order to establish the identity of the remains.

With complete DNA profiles from human remains and the availability of multiple first or second degree relatives, kinship calculations can produce extremely high LRs to indicate identity with great surety. However, LR or posterior probability thresholds should be established that take into account the possibility of adventitious "matches" where elevated kinship indices are detected among unrelated individuals by chance [14]. The higher the number of profiles in the database, the greater the probabilities of finding adventitious matches, especially for kinship analysis [1], resulting the in the risk of false positives. To the same extent, when deficient family pedigrees are compared to PM genetic data, the probability of missing a true identification increases leading to false negatives [19]. An intermediate situation may occur when a match with a LR value below the pre-agreed identity threshold is obtained. In this case, the forensic genetics laboratories may carry out further analyses, such as typing more autosomal or genealogy markers (mtDNA, Y-STRs) in order to increase the power of discrimination.

The number of victims involved in DVI or MPI episodes has a profound influence on the certainty of identification. Each comparison between a victim profile and a pedigree creates an additional chance for an adventitious false positive, and therefore the "prior odds" adopted in a Bayesian approach should reflect the number of victims in the event. If prior odds are taken as the inverse of the number of missing, then LR thresholds can be established that correspond appropriately to defined levels of identification certainty (posterior probability) in each case [5–20].

Taking all this into consideration, the aim of this exercise was to introduce the participating laboratories of the GHEP-ISFG Working Group into the complexities of the large scale identification of victims through genetic data processing, i.e., statistical evaluation for large-scale genetic data comparisons through direct match or kinship analysis, the use of prior odds and the consideration of critical LR and posterior odds values.

As an initial attempt to assess the conduct of forensic laboratories in this area, and to provide participating laboratories with practical experience, the design of this exercise was kept comparatively simple, and did not include some of the complexities that can be encountered in practice. As such, the simulated data did not include instances of pedigree inconsistencies associated to victims, allelic drop-in or drop-out, or the presence of related victims. The GHEP-ISFG Commission on DVI/MPI has evaluated this first theoretical challenge and, based on this experience, may set out further challenges on more complex and realistic scenarios.

2. Materials and methods

2.1. Design of the preliminary exercise

Firstly, the Forensic Genetics Laboratory of the Argentinean Forensic Anthropology Team (EAAF), the organizing laboratory, sent a preliminary simulated matching exercise to the International Commission on Missing Persons (ICMP) for them to compare the concordance of results obtained from both laboratories. Results on both, direct match for "remains re-associations" and on kinship analysis for "identifications" were fully concordant between both laboratories. Based on this first preliminary result, a second exercise was designed and the laboratories were invited by GHEP-ISFG Working Group to participate. This second exercise is described below.

[1] "Closed" scenarios are those where the number and names of the missing are known, whereas in "open" scenarios there is not a defined list of possible victims to whom remains must correspond.

2.2. Proposed simulated scenario and materials distributed to the laboratories

In the exercise, the simulated scenario was the disappearance of 193 individuals after a violent confrontation in an open episode within the framework of a prolonged conflict. The missing individuals belonged to a metropolitan admixed population; there was no evidence that individuals involved in the episode were related, nor population sub-structure.

A common grave containing human skeletal remains was found years later and it was presumed that some victims from the above-mentioned conflict might be present in this common burial ground.

Once exhumed, and according to archeological and anthropological findings, the mass grave revealed that the minimum number of individuals (MNI) was 58 and commingled remains were evident, thus describing this burial as a secondary grave. Eighty seven (87) bone samples were selected for genetic analysis.

2.3. Bone sample profiles

An excel file with autosomal STR profiles (aSTR) for 15 markers included in the Identifiler® Kit for the 87 bone samples (BS) were distributed to the participating laboratories (partial and full profiles were included (Table S1)). The profiles used in this study were derived, with modifications to the genetic profiles conducted manually, from disparate cases of LIDMO. The participants were instructed that the aSTR profiles present in bone samples showing just one allele in any locus should be considered as a "true homozygous". Therefore, no dropout events must be considered, even in those samples rendering partial STR profiles. This criterion was proposed in order to prevent the laboratories from the need to use more complicated statistics, like probabilistic LRs for direct match [23,24] or dropout consideration for kinship analysis [25]. There were no profiles showing more than two allele patterns that could be interpreted as tri-allele patterns, contamination or drop-in events. As mentioned above, it was assumed that the victims were not related and nor belonged to any population sub-group.

Supplementary material related to this article found, in the online version, at http://dx.doi.org/10.1016/j.fsigen.2015.11.004.

2.4. Reference samples (victim's relatives)

Although there were 193 missing persons associated with the conflict, it was explained that only 286 relatives' reference samples were available and represented just 119 missing victims (average 2.4 reference samples per victim). The participants were provided with an Excel worksheet containing the genetic profiles from these 286 victim's relative references (RS) integrating diverse genealogies (Table S2); the relationship of each reference with the victim and the gender of the victim were also indicated. The profiles used were derived, with modifications to the genetic profiles conducted manually, from disparate cases of LIDMO. It was assumed that reference genealogies were consistent and that there were no differences between social and genetic pedigrees among nuclear families associated with the identifications of the human remains (i.e: unknown non-paternity events).

Supplementary material related to this article found, in the online version, at http://dx.doi.org/10.1016/j.fsigen.2015.11.004.

2.5. Simulated exercise: aim and challenges

2.5.1. Direct match comparisons for "remains re-associations"

As laboratories were informed that the grave was described as secondary with commingled remains, the possibility of "intra-skeletal re-associations" was present. For skeletal re-associations, a direct match among profiles rendering a LR value equal or greater than 1,0E + 06 (LR $\geq$ 1,000,000) was set in order to report any "re-association". Although the use of a prior odds value was not indicated for remains re-associations by direct match, the participating laboratories were asked whether they would use it or not, according to the following questionnaire: (a) would not use, (b) would use the minimum number of individuals (MNI = 58) set by anthropological methods, (c) would use the minimum number of profiles found among the PM samples and (d) would use the number of missing persons (MP) associated with the episode.

2.5.2. Kinship analysis for remains "identifications"

With the aim of identifying bone remains, genetic comparisons of the provided bone (BS) and reference profiles (RS) databases were required. Laboratories were asked to inform about the software used to compare the profiles (i.e: widely used software like CODIS [26], DNA-VIEW [27], Familias [28,29] or locally developed).

Regarding the Bayesian treatment of the results, laboratories were instructed that the event involved 193 missing persons. As a threshold to report a positive identification, a posterior odds (PO) value (prior odds $\times$ LR = posterior odds) of $\geq$1,0E + 04 was set (PO $\geq$ 10,000; probability of identity 99.99%).

2.5.3. Remains with matching PO values below threshold

Laboratories were asked to report any other bone remains which could not be excluded as related to a particular family group, but whose PO values were near the threshold but not enough to report the identification. This information was also collected and gathered in tables in order to report the matches below threshold.

The allelic frequencies of autosomal STR markers, according to the GHEP 2011 [30–34] proficiency testing exercise, were sent to all the participating laboratories. A one-step mutation rate of 1.0E-03 (μ = 0.001) was set for each genetic marker.

3. Results

Although 26 laboratories requested and received the material for the exercise, only 11 laboratories returned results. All the participating laboratories in this exercise informed that they had participated in at least the GHEP-ISFG proficiency tests, 4 laboratories informed that they also had participated in other proficiency tests and one laboratory reported that they had only participated in the ISFG proficiency test.

3.1. Software for profiles comparison and statistical analysis

Eight out of eleven participating laboratories (8/11 = 73%) used software for large-scale genetic comparisons: CODIS7 [26] or DNA-VIEW [27]; some laboratories used a combination of the above-mentioned software with different versions of Familias [28,29] or with other locally developed software for the database screening

Table 1
Results for direct matching profiles re-associations. There were 6 re-association groups involving 17 samples.

Group	Matching profiles	(%) of reporting laboratories
1	BS19, BS41, BS87	11/11 (100%)
2	BS1, BS16, BS18, BS81, BS86	11/11 (100%)
3	BS27, BS67 (F12)[a]	11/11 (100%)
4	BS10, BS71	11/11 (100%)
5	BS8, BS80	11/11 (100%)
6	BS3, BS17, BS57	11/11 (100%)

[a] Group 3 re-associated two bone samples BS27 and BS67; furthermore using the most informative profile, these re-associated samples matched with family group F12.

or statistical analysis of the results. Two laboratories manually compared the profiles and used versions of Familias software for statistics. Finally, one laboratory compared the profiles using a spreadsheet that was locally developed.

3.2. Reporting direct matches for remains (PM) re-associations

Albeit 87 postmortem (PM) profiles were sent, the exercise was designed to obtain 17 PM bone-to-bone profile matches in 6 re-association groups (Table 1).

All of the laboratories (11/11 = 100%) agreed on the profiles re-association by direct match under the established criterion (i.e.: report the match for re-association if LR ≥ 10E + 06).

Although laboratories were not instructed about the use of a "prior odds" value specifically with relation to re-associations, they were asked if they would use it to report a re-association according to the criteria described in methods. Eight out of eleven laboratories (73%) reported that they would use no prior odds value and one laboratory (9%) indicated that it would use 1/58 as this was the minimum number of individuals estimated by anthropological methods; only two laboratories (18%) specified that they would use 1/193, considering the number of victims associated to the conflict. In a secondary grave with extensively commingled remains relating to an episode with 193 missing persons, prior odds of 1/193 would be conservatively appropriate in a general DNA-centered evaluation. For this exercise, when laboratories used some other prior probability for re-association, this was classed as a "Type C" error.

3.3. Reporting matches between bones (BS) and family references (RS) by kinship

The exercise simulated 9 matches between bone remains (BS) and reference families (RS) under the above-mentioned criterion (Posterior odds ≥1.0E + 04 and probability of identity ≥99.99%). Five laboratories reported the 9 expected matches between bone samples (BS) and reference family samples (RS) with coincidental posterior odds values (labs 25, 15, 19, 7 and 5) (Table 2 top) and Supplementary Figs. 1–9. These results were consistent with the results expected by the organizing laboratory.

Supplementary material related to this article found, in the online version, at http://dx.doi.org/10.1016/j.fsigen.2015.11.004.

The remaining 6 laboratories showed different results (Table 2 bottom). Laboratories 12 and 17 did not report one match (error 11%); laboratories 8 and 20 did not report two matches (error 22%); laboratory 3 did not report 3 matches (error 33%) and laboratory 4 did not report 4 matches (error 44%). Altogether, 6/11 (55%) laboratories did not appropriately report all the matches under the specified criterion. Nevertheless, 5/6 laboratories that did not report all the matches by kinship analysis did indicate a "probable identification" but reporting a PO value below the threshold. Only laboratory 3 lost one identification (BS45 with F36) since they did not report any match at all.

Comments about the dispersion of PO values that are above the established threshold will be detailed in subsequent paragraphs. The only non-reported (NR) bone/family pair is shown in a dark grey box.

3.4. Types of errors detected in kinship analysis

After analyzing the results from the participants, two main types of errors were observed and we will refer to them as Type A and Type B errors.

Type A error: This error consisted in partially using the genetic information of the family references instead of using the whole pedigree for kinship calculation. Laboratories failing to detect the match between BS and RS due to this type of error reported several PO values (one for each family reference), or just used some

Table 2
Expected AM/PM matches with Post Odds ≥1.0E + 04. Top of table: laboratories with coincidental results for the 9 matchingbone/family groups. Bottom of table: laboratories with other results.

Case	BS9/F90	BS12/F96	BS27/F12	BS45/F36	BS56/F44	BS62/F63	BS76/F110	BS77/F85	BS82/F72	Error rate
Reference's—relationship with victim[a]	RS210-Ch RS211-Ch	RS222-Ch RS223-Sb	RS26-Mo RS27-Sb	RS77-Sb	RS96-Ch RS97-Ch	RS145-Mo RS146-Sb	RS256-Ch RS257-Ch RS258-Ch RS259-Sb RS260-Sb	RS199-Sb RS200-Mo	RS168-Wf RS169-Ch RS170-Ch	
Average[b]	3,52E + 06	1,64E + 05	3,58E + 05	1,16E + 06	3,56E + 06	1,30E + 08	7,80E + 18	3,24E + 05	9,82E + 10	
25	2.60E + 06	1.49E + 05	2.90E + 05	9.48E + 05	2.27E + 06	6.48E + 07	3.93E + 18	1.32E + 05	8.60E + 10	0/9
15	6.37E + 06	3.11E + 05	4.84E + 05	1.63E + 06	8.19E + 06	3.62E + 08	1.69E + 19	1.04E + 06	1.35E + 11	0/9
19	3.03E + 06	1.14E + 05	7.64E + 05	1.26E + 06	2.54E + 06	8.13E + 07	1.39E + 19	1.75E + 05	1.14E + 11	0/9
7	3.03E + 06	7.29E + 04	1.35E + 05	9.77E + 05	2.54E + 06	8.13E + 07	2.46E + 18	1.43E + 05	7.17E + 10	0/9
5	2.59E + 06	1.74E + 05	1.15E + 05	9.90E + 05	2.25E + 06	6.22E + 07	1.79E + 18	1.29E + 05	8.45E + 10	0/9
17	1.05E + 05	5.60E + 01	1.49E + 07	1.51E + 06	1.00E + 05	3.38E + 07	1.12E + 18	1.26E + 04	1.45E + 11	1/9
12	3.00E + 06	1.29E + 05	5.95E + 00	1.12E + 05	2.50E + 06	8.02E + 07	6.94E + 29	1.73E + 05	3.59E + 13	1/9
20	5.40E + 09	1.40E + 04[c] (RS222)	6.53E + 01	5.11E + 04	8.10E + 04	8.34E + 02[c]	2.91E + 04[c] 6.44E + 04 1.20E + 07	1.07E + 04[c] (RS200)	6.94E + 04[c] (RS170)	2/9
8	5.28E + 09	1.60E + 07	1.40E + 05	1.40E + 03	4.04E + 07	8.34E + 06	1.24E + 29	3.11E + 01	3.79E + 13	2/9
3	2.06E + 05	1.36E + 03	4.61E + 03	Not reported	2.62E + 04	2.50E + 05	2.21E + 15	8.96E + 04	1.08E + 09	3/9
4	4.19E + 03[c] 6.23E + 03	1.42E + 04[c] (RS222)	6.52E + 01[c] 7.84E + 01	1.26E + 06	7.15E + 02[c] 8.93E + 02	5.02E + 03[c] 8.35E + 02	3.10E + 08[c] 6.48E + 04 2.92E + 04	1.08E + 04[c] (RS200)	9.23E + 04[c] 8.26E + 06	4/9

BS/F = bone sample/family group, i.e: BS9/F90 means bone sample BS9 matching with family group F90.

[a] Reference's relationship with victims; Mo = mother; Sb = sibling; Ch = child; Wf = wife.

[b] Average posterior odds value obtained from the 5 correct coincidental laboratories. Light grey color boxes = reported as probable match (under threshold 1.0E + 04). Dark grey color box = non reported match.

[c] Laboratory reported several posterior odds (and LR) comparing different family references or using some relatives instead of the whole pedigree for statistical calculations. i.e: (RS222) = laboratory used only reference RS222.

selected references for the kinship analysis. For this reason, since the whole genetic information of the pedigree was not used, these laboratories frequently did not reach the pre-agreed PO value to report the match. Table 2 shows for example, that lab #4 reported two PO values for case BS62/F63, using RS145 (victim's mother) and RS146 (victim's sibling) separately (PO values 5.02E + 03 and 8.35E + 02 respectively). As both PO values did not reach 1.0E + 04 necessary to report a match, the laboratory missed the identification although they indicated it as 'probable'. Other laboratories used only certain references for the kinship analysis, and disregarded others. It can be seen in Table 2 that lab #20 reported a PO value of 8.34E + 02 for case BS62/F63 although they did not indicate which reference was used to report this PO value. However, this PO value is identical to the one reported by lab #4 for RS145 (victim's mother). The laboratories that reported PO values by selecting only some family references generally prioritized the use of parent/child relationships over siblings/siblings. Fig. 1 graphically shows the results expressed in Table 2 for case BS62/F63.

Type B error: The other reasons why laboratories reported dispersed PO values for kinship analysis were due to either the wrong definition of the "Identity" hypothesis (H_1) or the wrong expression of the "Non-identity" hypothesis (H_2) or both. In general, this type of error may artificially increase or decrease the LR and consequently the PO value and the probability of identity as well.

Fig. 2 presents an example of this error for case BS09/F90. References integrating the family group F90 comprise RS210 and RS211 (victim's children). The most suitable "Identity" hypothesis would be H_1: BS9 belongs to the father of references RS210 and RS211, *who are full siblings*; this introduces the need to add a wife of the missing person to the pedigree for the kinship analysis. The most suitable "Non-identity" hypothesis would be H_2: BS9 belongs to someone else not related to references RS210 and RS211, *who are full siblings*. For case BS09/F90, laboratories #8 and #20 set the "Identity" hypothesis (H_1) as: BS09 belongs to the father of RS210 and RS211 but without considering the references as being full siblings (the victim's wife was not taken into account). For the "Non-identity" hypothesis, both laboratories set H_2 as BS09, RS210 and RS211 as being not related. Considering the hypothesis of 3 individuals that are not related in H_2 produces LR and PO values artificially higher than the correct ones.

Other examples of Type A and B errors can be seen in Table 2 for case BS76/F110 (see also Supplementary Fig. 6). Graphics showing positive matches among Bone Samples (BS) and Family Groups indicating the PO values reported by individual laboratories are represented in Supplementary Figs. 1–9.

3.5. Reporting matches with posterior odds (PO) below threshold 10E ± 04 (probable match)

The simulated data set also included 3 expected matches between bone samples (BS) and family reference samples (RS) but with PO values below 10,000 and higher than 50. These PO values correspond to LR values from 9650 to 1930,000 (with prior odds of 1/193). This feature of the data set was included to represent frequent situations when AM and PM profiles database are compared through kinship analysis, i.e.: obtaining matches with significantly elevated PO values that are below the pre-established threshold for reporting a true match.

Table 3 shows the results obtained for the 3 expected matches below threshold: BS23/F31, BS49/F118 and BS32/F113 (Supplementary Fig. 10–12). Again, type A and B errors already mentioned, influenced the matching results. The laboratories which committed Type A error (did not use all the family genetic information), lost the "match below threshold" in some cases; e.g: case BS23/F31 indicated as not reported (NR) in Table 3 (Supplementary Fig. 10).

Supplementary material related to this article found, in the online version, at http://dx.doi.org/10.1016/j.fsigen.2015.11.004.

Type B error, in which H_1 and/or H_2 hypotheses are erroneously considered, is represented in Table 3 as bold numbers in grey colored boxes. Two laboratories reported the "match below threshold" of case BS49/F118 as a true match with PO values artificially higher than 10E + 04 (Supplementary Fig. 11). Family F118 is represented by references RS281, RS282 and RS283 as full

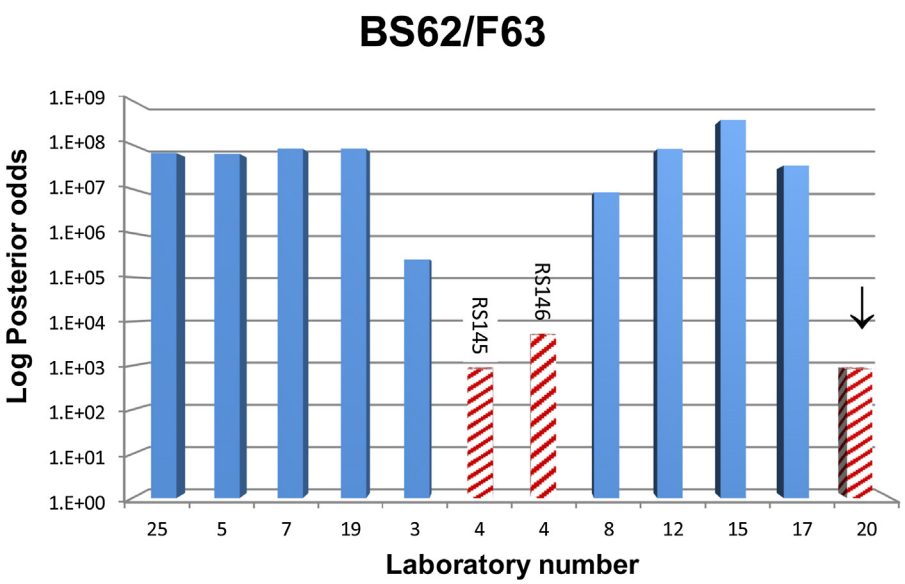

Fig. 1. Graphic representation of posterior odds values (PO) for the match between BS62 and family F63 comprising references RS145 (victim's mother) and RS146 (victim's brother). Laboratories reporting consensus values informed PO of approximately 1.0E + 08. Laboratories #4 and #20 reported PO below threshold 1.0E + 04. Lab #4 reported two PO values, one to each reference. Lab #20 did not indicate the reference used to the kinship analysis (right arrow) although they reported the same PO for RS145 (victim's mother) as lab #4. Hatched bars = individual PO below threshold 1.0E + 04 reported by both laboratories.

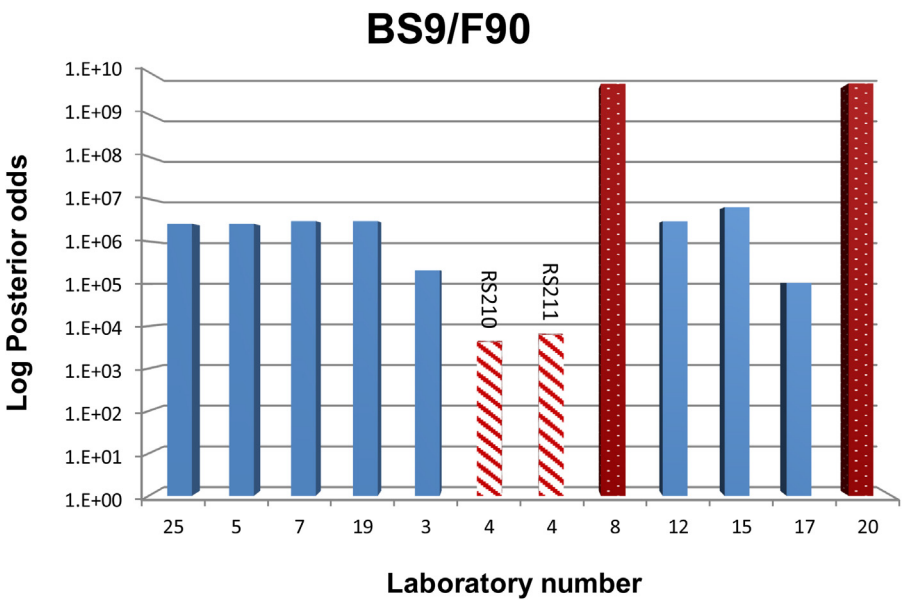

Fig. 2. Graphic showing different posterior odds for the match between BS9 and family group F90 which comprises references RS210 (victim's daughter) and RS211 (victim's daughter). Hatched bars = individual PO reported for each victim's daughters produced PO values below threshold of 1.0E + 04 (Type A error). Dark bars = artificially higher reported PO values by two laboratories (Type B error).

siblings of the victim. So, the most appropriate H_1 would be that BS49 belongs to a full brother of references RS281, RS282, RS283 (who are full siblings). This raises the necessity of adding both parents of RS281, RS282 and RS283 when analyzing the pedigree to integrate the victim to it. If no genetic inconsistences are detected, the most appropriate H_2 expression is that BS49 belongs to someone else who is not related to RS281, RS282 and RS283 *who are all full siblings*. Laboratory #17, reported a PO value five degrees of magnitude higher than the consensus. This laboratory set the H_1 as BS49, RS281, RS282 and RS283 being all half siblings from the same parent and the H_2 as all BS49, RS281, RS282 and RS283 are not related. For this reason, the PO reported by this lab was 3.64E + 08, fivefold higher than the correct consensus of ~10E + 03.

Another example of "match below threshold" is the case BS32/F113. When references RS268 and RS269 (children of victim) were compared with the bone samples database there was a match with BS32 as father of both references, but all laboratories reported a PO below 10E + 04 as expected. This PO under the threshold was due to an inconsistency between RS268 and BS32 that could be considered as a one-step mutation at locus CSF1PO (mutation rate indicated for all markers = 0.001). BS32 for CSF1PO is 11/13 while references RS268 (son) 12/12 and RS269 (son) are 10/13 (Table S1). Considering sample RS268 (son) as being 12/13, the LR for this BS/RS match would have been 1.8E + 08 and the PO = 9.4E + 05, exceeding the threshold for reporting a genuine match. In Table 3, laboratory #4 reports two results (Type A error) one of them with a PO below 1 (2.39E-02 for sample RS268) due to the mutation previously mentioned in CSF1PO, (see also Supplementary Fig. S12).

Regarding the use of software with facilities for large-scale profile databases comparisons, the 5 laboratories that presented correct results for profiles re-association as well as for kinship analysis, used available software for massive comparisons: CODIS7 or DNA-VIEW. However, 3/6 laboratories that made errors, also used these available software. Furthermore, 2/4 laboratories that made 85% of overall errors used CODIS7. Fig. 3 shows the proportion of success per laboratory indicating the type of software used. While some laboratories failed to report all the matches that should have been made, none reported false matches that would have resulted in mis-identification.

4. Discussion

Missing people identification through large-scale comparisons of genetic profiles in AM and PM databases involves some issues that laboratories should carefully consider [1]. This exercise has revealed interesting results which can guide forensic genetics laboratories on how to cautiously consider the complexity involved in large-scale victims identification. The most relevant lessons left by this exercise are: (i) considering odds within the Bayesian framework, (ii) the correct analysis of the family pedigree in both "Identity" and "Non-identity" hypotheses to calculate the likelihood ratio, and (iii) using as much genetic information provided by the family pedigree as possible to identify the victim by kinship analysis [35].

Table 3
Results reported by laboratories on the 3 expected bone samples(BS)/family reference (RS) match with post odds below 1.0E + 04 (5.0E + 01 < posterior odds < 1.0E + 04).

Bone sample (BS)		BS23/F31	BS49/F118	BS32/F113
Reference sample		RS66 (Sb)	RS281 (Sb)	RS268 (Ch)
		RS67 (Sb)	RS282 (Sb)	RS269 (Ch)
			RS283 (Sb)	
Lab[a]	25	4.67E + 02	1.13E + 03	3.31E + 00
	15	3.51E + 03	9.64E + 04	1.52E + 02
	19	5.30E + 02	1.30E + 03	2.08E + 01
	7	5.24E + 02	1.30E + 03	4.17E + 01
	5	4.69E + 02	1.13E + 03	7.15E + 00
	17	3.76E + 03	3.64E + 08	7.64E + 02
	12	5.28E + 02	1.35E + 03	1.71E + 01
	20	NR	6.11E + 01	4.48E + 02
	8	NR	1.37E + 03	5.54E + 03
	3	2.43E + 02	3.22E + 00	7.41E + 00
	4	5.31E + 02	1.30E + 03	2.74E + 02
				2.39E − 02[b]

BS/F = bone sample/family group, i.e:BS23/F71 means bone sample BS23 matching with family group F71.

[a] Reference relationship with victims; Sb = sibling; Ch = child. Light grey color boxes with bold numbers = match wrongly reported (Post Odds higher 1.0E + 04). NR = match not reported by the laboratory.

[b] Laboratory reporting two PO, one to each family reference for kinship analysis. The negative PO value of 2.39E-02 is due to a possible single step mutation in one of the two references when analyzing the integrated pedigree.

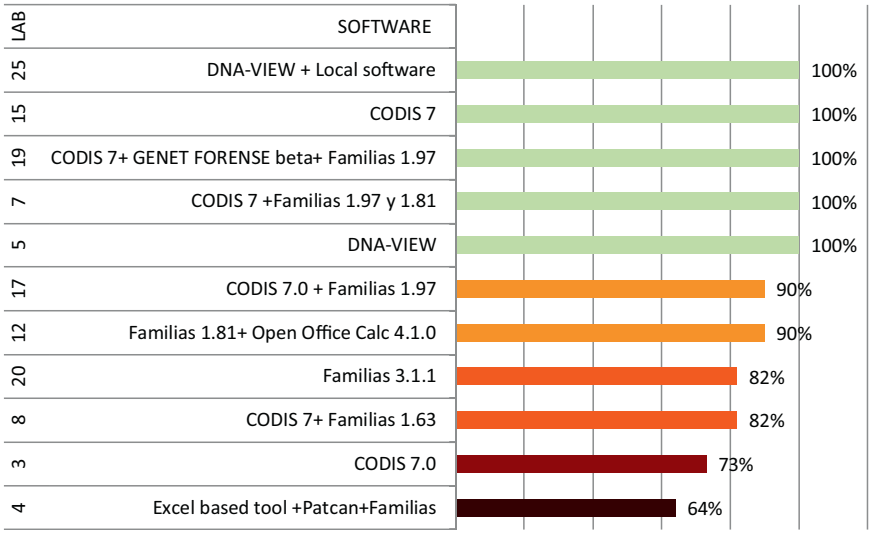

Fig. 3. Software used by different laboratories in relation to correct reported results. Top of figure: laboratories with correct consent results (light bars). Labs with decreasing correct results are shown in darker bars.

One of the main variables influencing the complexity in DVI and/or MPI identification is the number of victims involved. A large number of victims represents a big challenge to the identification of missing persons since considering the total amount of victims involved in the disaster within the Bayesian framework is mandatory [14,19]. This exercise focused on a fairly small number of profiles: 87 evidence profiles and 286 family references profiles corresponding to 119 victims within an episode involving a total amount of 193 missing people. Likewise, the exercise included a low number of PM profiles re-associations by direct match as well as few PM and AM profiles matches by kinship analysis. This small amount of profiles simplified the report of the results by participants as well as the analysis of the whole reported results and general conclusions.

Another source of complexity in DVI/MPI scenarios is the degree of disarticulation or fragmentation of the victims [8,9] or the presence of commingled remains [37]. Regarding this, the exercise considered a secondary grave with commingled remains and the possibility of PM profiles re-associations. All laboratories properly re-associated the 17 matching profiles in 6 re-association groups.

The exercise also asked the participating laboratories for evaluating the use of prior odds since the exercise simulated a mass grave with commingled remains related to 193 missing people. Although the evaluation of metadata other than DNA results is not the aim of genetic laboratories, the exercise included a questionnaire to evaluate how the participants considered the Bayesian framework for the remains re-association by direct match in the common grave analyzed. Surprisingly, most of laboratories (8/11 = 73%) considered unnecessary to use prior odds for remains re-association, even when dealing with a secondary grave with commingled remains associated to an episode of 193 missing people (Type C error), where the correct prior odds is 1/(number of victims). LRs obtained from direct matches may render so high values that turn negligible the impact of prior odds for this number of victims on the posterior odds. However, the prior odds should still be considered within the Bayesian context of the episode under investigation.

Regarding the use of prior odds for "identification" matches by comparisons between PM profiles (BS) and AM reference profiles (RS), the laboratories were instructed to combine a prior odds of 1/193 with the LR to report a posterior odds $\geq$10E + 04 as a true identification. This criterion was used to avoid dispersion of results due to the use of prior odds by some labs and not by others.

The state of preservation of biological evidences may influence the success for producing full genetic profiles and challenges the laboratory's expertise in the use of special protocols on these difficult samples. Decomposed [15], burnt [16] or severely degraded bone remains [17] may decrease the recovery of DNA [35,36] thus, the interpretation of profiling results on such degraded or low template DNA [23,38,39] might be an extra-variable when several agencies cooperate in DVI or MPI projects. The exercise considered the above mentioned difficulties including partial PM profiles. However, laboratories were instructed to consider markers showing one allele as a "true homozygous" in order to simplify the statistics for the participants and to avoid dispersion of results.

Two kinds of AM references may be useful for missing person identifications: biological specimens from the victim and victim's relatives. Genetic profiles from the victim produce strong statistical significance and high probability of identity and simpler statistic calculations. The GHEP-ISFG exercise evaluated the direct match statistics through PM profiles re-associations; all laboratories agreed on correct results.

However, in many DVI/MPI cases, only victims' relatives are available as reference samples so that kinship analysis is needed. This involves more complex statistics that depends on the degree of relatedness of the references with the victim and these comparisons usually render a lower power of discrimination. Different studies have attempted to determine which reference relatives are preferred for kinship analysis to identify the victim [14,19–21]. The exercise included diverse family genealogies for kinship analysis to report the PM/AM match as a challenge to the participating laboratories. A substantial type of error found (Type A) for PM/AM kinship analysis was due to the use of genetic information from only selected family references disregarding others, instead of using the full pedigree. A variable of this error was the report of several individual PO values, one for each reference; both errors led some laboratories not to report the match. However, all laboratories except one, reported these true matches as probable matches, and recommended further genetic studies, or adding more references to solve the case. This point is relevant since a laboratory considering a near-threshold match, may perform further actions such as analyzing more autosomal genetic systems, and/or genealogy systems as mtDNA and/or Y-STR markers if applicable.

Another difficulty present in kinship analysis to report AM/PM matching was the erroneous description of the family pedigree or the incorrect integration of the victim to the pedigree proposed in the "Identity" hypothesis H_1, in the "Non-identity" hypothesis H2 or in both (Type B). These miscalculations caused some dispersion in the results with lower PO values and the corresponding loss of matches or the artificially higher PO values leading to "false matches". Both difficulties described may cause the victim's identification to be missed, or adventitious matches with the consequent profound impact on the families in a real DVI/MPI scenario.

In this exercise, the participants were also asked to report AM/PM near-threshold matches. Type A and B errors also influenced in the results by missing the weak matches in some cases or considering them as a true match due to an artificially high LR. Special care should be taken on this issue as weak true matches as described above may overlap with random matches when performing kinship analysis raising the need to perform further studies to distinguish among them [5,21].

Appropriate software is an essential tool for large-scale genetic profile database comparisons in DVI or MPI. All the participating laboratories used some kind of software to carry out the comparisons and match calculations. Eight out of eleven participating laboratories used software such as CODIS and DNA-VIEW, but even so, many laboratories reported erroneous results. Consequently, the exercise showed that the use of powerful software did not always ensure correct results so special care should be taken for proper validation, training, and proficiency testing prior to the use of the software for kinship analysis.

Other factors that increase the complexity for victim identifications in DVI/MPI were not included in this exercise, e.g.: related victims, where multiple relatedness scenarios must be considered to distinguish amongst the individuals [14]. Likewise, differences between social and genetic pedigrees in nuclear families, considering dropout for LR calculation, population sub-structure or analysis of genealogy markers such as mtDNA or Y-STR and their statistical evaluation, were not included.

In short, within the usual complexity in DVI or MPI, the main variables faced by laboratories in this exercise were: some partial PM profiles, diverse family genealogies as AM information, the presence of commingled remains, one genetic inconsistency in a pair AM/PM family group which could be interpreted as a mutation and the use of software for profiles comparisons. Although other simulations on DVI and MPI using CODIS 6 have been published previously with interesting results [40], this GHEP-ISFG exercise aimed at a collaborative effort carried out by different participants to gain experience on the proceedings in DVI or MPI scenarios. Topics such as the use of prior odds for human remains re-associations or identifications, the selection of the most suitable hypotheses to properly weigh the biological evidence, and the strategies that can be followed when near-threshold AM/PM matches are obtained were examined in this study. In conclusion, the major contribution of this exercise was to introduce the participating laboratories into the challenge of identifying missing people in DVI or MPI through the genetic profiles database comparisons. The detection of errors that may be committed when dealing with data management, the biostatistical interpretation and the Bayesian treatment of the results is a learning experience intended to prompt forensic genetics laboratories to focus on the potential pitfalls and required expertise necessary to the challenging task of identifying victims in mass disasters or missing person identification programs.

References

[1] Scientific Working Group on DNA Analysis Methods (SWGDAM). Guidelines for Missing Persons Casework. http://www.swgdam.org/ (consulted in June, 2015).

[2] Missing People, DNA Analysis and Identification of Human Remains, A guide to best practice in armed conflicts and other situations of armed violence. International Committee of the Red Cross (ICRC), http://www.icrc.org/eng/resources/documents/publication/p4010.htm (consulted in June, 2015).

[3] Interpol DVI guide. http://www.interpol.int/es/INTERPOL-expertise/Forensics/DVI-Pages/DVI-guide (consulted in June, 2015).

[4] A. Sozer, M. Baird, M. Beckwith, B. Harmon, D. Lee, G. Riley, S. Schmitt, Guidelines for Mass Fatality DNA Identification Operations, AABB, 2015 http://www.aabb.org/programs/disasterresponse/Documents/aabbdnamassfatalityguidelines.pdf (consulted in June, 2015).

[5] M. Prinz, A. Carracedo, W.R. Mayr, N. Morling, T.J. Parsons, A. Sajantila, R. Scheithauerg, H. Schmitter, P.M. Schneider, DNA commission of the International Society for Forensic Genetics (ISFG): recommendations regarding the role of forensic genetics for disaster victim identification (DVI), Forensic Sci. Int.: Genet. 1 (2007) 3–12.

[6] R. Lessig, K. Thiele, J. Edelmann, Tsunami 2004—experiences, challenges and strategies, Int. Congress Ser. 1288 (2006) 747–749.

[7] S. Donkervoort, S.M. Dolan, M. Beckwith, T.P. Northrup, A. Sozer, Enhancing accurate data collection in mass fatality kinship identifications: lessons learned from Hurricane Katrina, Forensic Sci. Int.: Genet. 2 (2008) 354–362.

[8] L.G. Biesecker, J. Bailey-Wilson, J. Ballantyne, H. Baum, F. Bieber, C. Brenner, B. Budowle, J.M. Butler, G. Carmody, P.M. Conneally, B. Duceman, A. Eisenberg, L. Forman, K.K. Kidd, B. Leclair, S. Niezgoda, T.J. Parsons, E. Pugh, R. Shaler, S.T. Sherry, A. Sozer, A. Walsh, DNA identifications after the 9/11 World Trade Center attack, Science 310 (2005) 1122–1123.

[9] E. Huffine, J. Crews, B. Kennedy, K. Bomberger, A. Zinbo, Mass identification of persons missing from the break-up of the former Yugoslavia: structure, function, and role of the International Commission on Missing Persons, Croat. Med. J. 42 (June (3)) (2001) 271–275.

[10] Guatemala: The Commission for Historical Clarification (CEH). https://hrdag.org/wp-content/uploads/2013/01/CEHreport-english.pdf (consulted in June, 2015).

[11] Latin American Initiative for the Identification of the Disappeared (LIID), Argentinean Forensic Anthropology Team. http://eaaf.typepad.com/iniciativa_en/ (consulted in June, 2015).

[12] L.E. Baker, E.J. Backer, Reuniting families: an online database to aid in the identification of undocumented immigrant remains, J. Forensic Sci. 53 (2008) 50–53.

[13] Amnesty International. http://www.amnestyusa.org/sites/default/files/amr410142010eng.pdf (consulted in June, 2015).

[14] C.H. Brenner, Some mathematical problems in the DNA identification of victims in the 2004 tsunami and similar mass fatalities, Forensic Sci. Int. 157 (2006) 172–180.

[15] A.Z. Mundorff, E.J. Bartelink, E. Mar-Cash, DNA preservation in skeletal elements from the World Trade Center disaster: recommendations for mass fatality management, J. Forensic Sci. 54 (4) (2009) 739–745.

[16] B. Leclair, C.J. Fregqau, K.L. Bowen, R.M. Fourney, Enhanced kinship analysis and STR-based DNA typing for human identification in mass fatality incidents: the Swissair flight 111 disaster, J. Forensic Sci. 49 (2004) 1–14.

[17] A. Alonso, S. Andelinovic, P. Martin, D. Sutlovic, I. Erceg, E. Huffine, et al., DNA typing from skeletal remains: evaluation of multiplex and megaplex STR systems on DNA isolated from bone and teeth samples, Croat. Med. J. 42 (3) (2001) 260–266.

[18] S.M. Edson, J.P. Ross, M.D. Coble, T.J. Parsons, S.M. Barritt, Naming the dead: confronting the realities of rapid identification of degraded skeletal remains, Forensic Sci. Rev. 16 (1) (2004) 63–90.

[19] C. Brenner, B. Weir, Issues and strategies in the DNA identification of World Trade Center victims, Theor. Popul. Biol. 63 (2003) 173–178.

[20] J. Ge, R. Chakraborty, A. Eisenberg, B. Budowle, Comparisons of familial DNA database searching strategies, J. Forensic Sci. 56 (2011) 1448–1456.

[21] W. Goodwin, C. Peel, Theoretical value of the recommended expanded European standard set of STR loci for the identification of human remains, Med. Sci. Law 52 (2012) 162–168.

[22] L. Parker, A. London, J. Aronson, Incidental findings in the use of DNA to identify human remains: an ethical assessment, Forensic Sci. Int.: Genet. 7 (2013) 221–229.

[23] P. Gill, L. Gusmão, H. Haned, W.R. Mayr, N. Morling, W. Parson, L. Prieto, M. Prinz, M. Schneider, P. Schneider, B. 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 (2012) 679–688.

[24] D. Balding, J. Buckleton, Interpreting low template DNA profiles, Forensic Sci. Int.: Genet. 4 (2009) 1–10.

[25] J. Buckleton, C. Triggs, Dealing with allelic dropout when reporting the evidential, Forensic Sci. Int. 160 (2006) 134–139.

[26] Federal Bureau of Investigation. http://www.fbi.gov/about-us/lab/biometric-analysis/codis/codis_brochure (consulted in June, 2015).

[27] C. Brenner, http://dna-view.com/ (consulted in June, 2015).

[28] T. Egeland, P. Mostad, B. Mevåg, M. Stenersen, Beyond traditional paternity and identification cases. Selecting the most probable pedigree, Forensic Sci. Int. 110 (2000) 47–59.
[29] D. Kling, A.O. Tillmar, T. Egeland, Familias 3—extensions and new functionality, Forensic Sci. Int. Genet. 13 (2014) 121–127.
[30] P. Martin, A. Alonso, B. Budowle, C. Albarran, O. Garcia, M. Sancho, Spanish population data on 7 tetrameric short tandem repeat loci, Int. J. Legal Med. 108 (3) (1995) 145–149.
[31] P. Martin, O. Garcia, C. Albarran, P. Garcia, M. Sancho, A. Alonso, Spanish population data on the four STR loci D8S1179, D16S539, D18S51 and D21S11, Int. J. Legal Med. 112 (5) (1999) 340–341.
[32] L. Fernández, O. García, P. Martín, C. Albarrán, P. García, M. Sancho, A. Alonso, A Spanish population study of the STR loci D2S1338, D19S433, Penta D, and Penta E, J. Forensic Sci. 48 (5) (2003) 1182.
[33] P. Martín, O. García, C. Albarrán, M. Sancho, A. Alonso, A Spanish population study of the STR loci HumLPL, D5S818, D7820 and D13S317, Int. J. Legal Med. 112 (1) (1999) 70–71.
[34] P. Martin, O. García, A. Alonso, C. Albarrán, M. Sancho, in: B. Olaisen, B. Brinkmann, P.J. Lincoln (Eds.), Alleles Frequencies of D3S1358 and Humfibra (FGA) in a Central Spain Population, Elsevier Science B.V., 1997, pp. 309–311.
[35] I.W. Evett, B.S. Weir, Interpreting DNA evidence, Statistical Genetics for Forensic Scientists Sunderland, Sinauer, MA, 1998 ISBN 0-87893-155-44.
[36] O. Loreille, T. Diegoli, J. Irwin, M. Coble, T. Parsons, High efficiency DNA extraction from bone by total demineralization, Forensic Sci. Int.: Genet. 1 (2007) 191–195.
[37] B.J. Adams, J.E. Byrd (Eds.), Commingled Human Remains Methods in Recovery, Analysis, and Identification, Elsevier, 2014 ISBN: 978-0-12-405889-7.
[38] P. Gill, J. Whitaker, C. Flaxman, N. Brown, J. Buckleton, An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA, Forensic Sci. Int. 112 (2000) 17–40.
[39] B. Caddy, A. Lineacre, G. Taylor, A review of the science of low template DNA analysis. Available from the National Archive. http://tna.europarchive.org/20100419081706/http:/www.police.homeoffice.gov.uk/publications/operational-policing/Review_of_Low_Template_DNA_1.pdf, 2008 (consulted in June, 2015).
[40] L. Bradford, J. Heal, J. Anderson, N. Faragher, K. Duval, S. Lalonde, Disaster victim investigation recommendations from two simulated mass disaster scenarios utilized for user acceptance testing CODIS 6.0, Forensic Sci. Int.: Genet. 5 (2011) 291–296.