Title: Analysis of uni and bi-parental markers in mixture samples: lessons from the 22nd GHEP-ISFG Intercomparison Exercise

Author: U. Toscanini L. Gusmão M.C. Álava Narváez J.C. Álvarez L. Baldassarri A. Barbaro G. Berardi E. Betancor Hernández M. Camargo J. Carreras-Carbonell J. Castro S.C. Costa P. Coufalova V. Domínguez E. Fagundes de Carvalho S. Ferreira S. Furfuro O. García A. Goios R. González A. González de la Vega A. Gorostiza A. Hernández S. Jiménez Moreno M.V. Lareu A. León Almagro M. Marino G. Martínez M.C. Miozzo N. Modesti V. Onofri S. Pagano B. Pardo Arias S. Pedrosa G.A. Penacino M.L. Pontes M.J. Porto J. Puente-Prieto R. Ramírez Pérez T. Ribeiro B. Rodríguez Cardozo Y.M. Rodríguez Lesmes A. Sala B. Santiago V.G. Saragoni A. Serrano E.R. Streitemberger M.A. Torres Morales S.A. Vannelli Rey M. Velázquez Miranda M.R. Whittle K. Fernández A. Salas

PII: S1872-4973(16)30130-2
DOI: http://dx.doi.org/doi:10.1016/j.fsigen.2016.07.010
Reference: FSIGEN 1557

To appear in: *Forensic Science International: Genetics*

Received date: 21-4-2016
Revised date: 14-7-2016
Accepted date: 17-7-2016

Please cite this article as: U.Toscanini, L.Gusmão, M.C.Álava Narváez, J.C.Álvarez, L.Baldassarri, A.Barbaro, G.Berardi, E.Betancor Hernández, M.Camargo, J.Carreras-Carbonell, J.Castro, S.C.Costa, P.Coufalova, V.Domínguez, E.Fagundes de Carvalho, S.Ferreira, S.Furfuro, O.García, A.Goios, R.González, A.González de la Vega, A.Gorostiza, A.Hernández, S.Jiménez Moreno, M.V.Lareu, A.León Almagro, M.Marino, G.Martínez, M.C.Miozzo, N.Modesti, V.Onofri, S.Pagano, B.Pardo Arias, S.Pedrosa, G.A.Penacino, M.L.Pontes, M.J.Porto, J.Puente-Prieto, R.Ramírez Pérez, T.Ribeiro, B.Rodríguez Cardozo, Y.M.Rodríguez Lesmes, A.Sala, B.Santiago, V.G.Saragoni, A.Serrano, E.R.Streitemberger, M.A.Torres Morales, S.A.Vannelli Rey, M.Velázquez Miranda, M.R.Whittle, K.Fernández, A.Salas, Analysis of uni and bi-parental markers in mixture samples: lessons from the 22nd

GHEP-ISFG Intercomparison Exercise, Forensic Science International: Genetics http://dx.doi.org/10.1016/j.fsigen.2016.07.010

Forensic Science International Genetics

Original paper

Analysis of uni and bi-parental markers in mixture samples: lessons from the 22nd GHEP-ISFG Intercomparison Exercise

U. Toscanini[1], L. Gusmão[2,3,4], M.C. Álava Narváez[5], J.C. Álvarez[6], L. Baldassarri[7], A. Barbaro[8], G. Berardi[1], E. Betancor Hernández[9], M. Camargo[10], J. Carreras-Carbonell[11], J. Castro[12], S.C. Costa[13], P. Coufalova[14], V. Domínguez[15], E. Fagundes de Carvalho[2], S. Ferreira[16], S. Furfuro[17], O. García[18], A. Goios[3,4], R. González[19], A. González de la Vega[20], A. Gorostiza[21], A. Hernández[22], S. Jiménez Moreno[23], M.V. Lareu[24], A. León Almagro[25], M. Marino[26], G. Martínez[27], M.C. Miozzo[28], N. Modesti[29], V. Onofri[30], S. Pagano[31], B. Pardo Arias[32], S. Pedrosa[33], G.A. Penacino[34], M.L. Pontes[35], M.J. Porto[36], J. Puente-Prieto[37], R. Ramírez Pérez[38], T. Ribeiro[39], B. Rodríguez Cardozo[40], Y.M. Rodríguez Lesmes[41], A. Sala[42], B. Santiago[43], V.G. Saragoni[44], A. Serrano[45], E.R. Streitemberger[46], M.A. Torres Morales[47], S.A. Vannelli Rey[48], M. Velázquez Miranda[49], M.R. Whittle[50], K. Fernández[43], A. Salas[24]

[1]*PRICAI-Fundación Favaloro, Buenos Aires, Argentina*

[2]*DNA Diagnostic Laboratory (LDD), State University of Rio de Janeiro (UERJ), Rio de Janeiro, Brazil*

[3]*IPATIMUP (Institute of Pathology and Molecular Immunology from de University of Porto), Porto, Portugal*

[4]*I3s (Instituto de Investigação e Inovação em Saúde, Universidade do Porto), Porto, Portugal*

[5]*Laboratorio de Genética Regional Bogotá del Instituto Nacional de Medicina Legal y Ciencias Forenses., Bogotá, Colombia*

[6]*Lab. de Identificación Genética. Depto. de Medicina Legal, Toxicología y Antropología Física. Facultad de Medicina. Universidad de Granada, Granada, Spain*

[7]*Institute of Public Sanity Section of Legal Medicine Catholic University of Sacred Heart, Rome, Rome, Italy*

[8]*Studio Indagini Mediche E Forensi (SIMEF), Reggio Calabria, Italy*

[9]*Laboratorio Genética Forense. Instituto de Medicina Legal de Las Palmas. ULPGC., Las Palmas, Spain*

[10]*Laboratorio de Genética Regional Suroccidente del Instituto Nacional de Medicina Legal y Ciencias Forenses., Cali, Colombia*

[11]*Policia de la Generalitat - Mossos d'Esquadra, Divisió de Policia Científica, Unitat Central del Laboratori Biològic, Sabadell, Barcelona, Spain*

[12]*Ministerio Público de Venezuela, Caracas, Venezuela*

[13]*Laboratório de Polícia Científica da Polícia Judiciária, Lisbon, Portugal*

[14]*Institute of Criminalistics Prague, Prague, Czech Republic*

[15]*Lab. Biológico de la Dirección Nacional de Policía Científica, Montevideo, Uruguay*

[16]*Instituto de Pesquisa de DNA Forense, IPDNA, Polícia Civil do Distrito Federal, PCDF, Brasília, Brazil*

[17]*Laboratorio de Análisis de ADN- Facultad de Ciencias Médicas- Universidad Nacional de Cuyo, Mendoza, Argentina*

[18]*Forensic Science Unit, Forensic Genetics Section, Basque Country Police-Ertzaintza, Erandio, Bizkaia, Spain*

[19]*Registro Nacional de ADN, Chile, Santiago de Chile, Chile*

[20]*CGC Genetics, Madrid, Spain*

[21]*GENOMICA, S.A.U., Madrid, Spain*

[22]*Instituto Nacional de Toxicología y Ciencias Forenses, Delegación en Canarias , Santa Cruz de Tenerife, Spain*

[23]*Laboratorio de Biología Forense. Dpto Patología y Cirugía. Universidad Miguel Hernández, Elche, Alicante, Spain*

[24]*Unidade de Xenética, Departamento de Anatomía Patolóxica e Ciencias Forenses, and Instituto de Ciencias Forenses, Grupo de Medicina Xenómica (GMX), Facultade de Medicina, Universidade de Santiago de Compostela, 15872, Galicia, Spain*

[25]*Comisaría General de Policía Científica - Laboratorio de ADN, Madrid, Spain*

[26]*Laboratorio de Genética Forense, Poder Judicial de Mendoza, Mendoza, Argentina*

[27]*Servicio de Genética Forense, Superior Tribunal de Justicia de Entre Ríos, Paraná, Argentina*

[28]*Laboratorio Regional de Genética Forense del NOA - Departamento Médico - Poder Judicial de Jujuy, Jujuy, Argentina*

[29]*Instituto de Genética Forense. Poder Judicial de Córdoba, Córdoba, Argentina*

[30]*Universita' Politecnica Delle Marche, DSBSP, Section of Legal Medicine, Ancona, Italy*

[31]*Biko SA, Montevideo, Uruguay*

[32]*Instituto Nacional de Toxicología y Ciencias Forenses, Departamento de Sevilla, Sevilla, Spain*

[33]*NASERTIC, Villava, Spain*

[34]*Unidad de Analisis de ADN, Colegio Oficial de Farmaceuticos y Bioquímicos, Buenos Aires, Argentina*

[35]*Serviço de Genética e Biologia Forenses, Instituto Nacional de Medicina Legal e Ciências Forenses, I.P. – Delegação do Norte, Porto, Portugal*

[36]*Serviço de Genética e Biologia Forenses, Instituto Nacional de Medicina Legal e Ciências Forenses, I.P., Coimbra, Portugal*

[37]*LabGenetics. Laboratorio de Genética Clínica S.L., Madrid, Spain*

[38]*Dirección General de la Policía, Granada, Spain*

[39]*Serviço de Genética e Biologia Forenses, Instituto Nacional de Medicina Legal e Ciências Forenses, I.P.-Delegação Sul, Lisbon, Portugal*

[40]*Banco Nacional de Datos Genéticos, Buenos Aires, Argentina*

[41]*Laboratorio de Biología y Genética Regional Noroccidente del Instituto Nacional de Medicina Legal y Ciencias Forenses., Medellín, Colombia*

[42]*Servicio de Huellas Digitales Genéticas-Fac. Farmacia y Bioquímica-Universidad de Buenos Aires, Buenos Aires, Argentina*

[43]*Instituto Nacional de Toxicología y Ciencias Forenses, Departamento de Madrid. Servicio de Biología., Madrid, Spain*

[44] *Unidad de Genética Forense, Servicio Médico Legal, Santiago, Chile*

[45]*Instituto Nacional de Toxicología y Ciencias Forenses, Departamento de Barcelona, Barcelona, Spain*

[46]*IACA Laboratorios, Bahía Blanca, Argentina*

[47]*Laboratorio Genia Geo, Montevideo, Uruguay*

[48]*Laboratorio Regional Patagonia Norte de Genética Forense - Poder Judicial de Río Negro, Bariloche, Argentina*

[49]*Instituto de Medicina Legal de Valencia , Valencia, Spain*

[50]*Genomic Engenharia Molecular, Sao Paulo, Brazil*

Corresponding author: Ulises Toscanini, PhD: utoscanini@pricai.com.ar

FSIGEN- Highlights

- Analysis of mixture samples delivered as part of the 22nd intercomparison exercise of the Spanish and Portuguese Speaking Working Group of the ISFG (GHEP-ISFG).
- 79 participant laboratories from 17 countries.
- Autosomal, Y and X STRs, mtDNA results analyzed.
- Drop-out and drop-in alleles were the main cause of errors.
- The highest error rates among nuclear markers were observed for X-STRs (between 7.8 and 15.4%), and the lowest ones were for Y-STRs (between 0.0% and 1.2%), while autosomal STRs and amelogenin showed intermediate values (between 4.7% and 6.5%).
- Some problematic issues regarding microvariants occurring at autosomal STRs; in particular, in the ability to assign microvariant alleles of the minor profile.
- Disparity in the proportion of the major / minor components among the participants was observed, reflected by the peak heights / areas in the corresponding electropherograms.
- Nomenclature problems, contamination and sample mix-up issues detected in mtDNA analyses. Most laboratories failed to detect the haplotype from the semen in semen-saliva mixtures.

Abstract

Since 1992, the Spanish and Portuguese-Speaking Working Group of the ISFG (GHEP-ISFG) has been organizing annual Intercomparison Exercises (IEs) coordinated by the Quality Service at the National Institute of Toxicology and Forensic Sciences (INTCF) from Madrid, aiming to provide proficiency tests for forensic DNA laboratories. Each annual exercise comprises a Basic (recently accredited under ISO/IEC 17043: 2010) and an Advanced Level, both including a kinship and a forensic module. Here, we show the results for both autosomal and sex-chromosomal STRs, and for mitochondrial DNA (mtDNA) in two samples included in the forensic modules, namely a mixture 2:1 (v/v) saliva/blood (M4) and a mixture 4:1 (v/v) saliva/semen (M8) out of the five items provided in the 2014 GHEP-ISFG IE. Discrepancies, other than typos or nomenclature errors (over the total allele calls), represented 6.5% (M4) and 4.7% (M8) for autosomal STRs, 15.4% (M4) and 7.8% (M8) for X-STRs, and 1.2% (M4) and 0.0% (M8) for Y-STRs. Drop-out and drop-in alleles were the main cause of errors, with

laboratories using different criteria regarding inclusion of minor peaks and stutter bands. Commonly used commercial kits yielded different results for a micro-variant detected at locus D12S391. In addition, the analysis of electropherograms revealed that the proportions of the contributors detected in the mixtures varied among the participants. In regards to mtDNA analysis, besides important discrepancies in reporting heteroplasmies, there was no agreement for the results of sample M4. Thus, while some laboratories documented a single control region haplotype, a few reported unexpected profiles (suggesting contamination problems). For M8, most laboratories detected only the haplotype corresponding to the saliva. Although the GHEP-ISFG has already a large experience in IEs, the present multi-centric study revealed challenges that still exist related to DNA mixtures interpretation. Overall, the results emphasize the need for further research and training actions in order to improve the analysis of mixtures among the forensic practitioners.

Keywords: GHEP-ISFG; intercomparison exercise; forensic mixtures; autosomal STRs, Y-chromosome, mtDNA

1. Introduction

The Spanish and Portuguese Speaking Working Group of the International Society for Forensic Genetics, namely GHEP-ISFG (formerly GEP-ISFG; GEP-ISFH) (www.ghep-isfg.org) dates back its origin to the year 1989, when founder members first met in New Orleans (USA) during the 13th World Congress of the ISFG (ISFH at that time) [1]. One of the earliest initiatives of the representatives at that meeting was the development of a collaborative Intercomparison Exercise program (IE) to improve the standardization of theoretical and technical issues, and to allow the participants to evaluate the quality of their genetic tests [2]. The first IE took place in 1992; from then onwards it has occurred annually. Since its beginning, the IE continuously evolved as new technologies were incorporated into the field, and the number of GHEP-ISFG members increased. For instance, before 1995 the number of participants were about twenty, and the markers reported in those years included the “to-date discontinued” SLPs, DQA1 and Polymarker, as well as a few emergent STR markers [2, 3]. To date, there are 171 laboratories from 23 different countries represented by 354 members of the GHEP-ISFG (September, 2015). The majority of these laboratories participate actively in multi-centric scientific activities organized by the group.

The complexity of the IEs also increased over time. Thus, since 1997, mitochondrial DNA (mtDNA) profiling and the analysis of forensic samples were included [4-6]. Statistical and interpretation problems were also addressed in subsequent IEs. Therefore, in 1996 a theoretical module was implemented in the IE in order to assess the laboratories’ performance in the calculation of likelihood ratios (LR). Subsequently, a forensic theoretical module and a parentage paper challenge were incorporated in 2007 and 2009, respectively.

Since the beginning, the IE is coordinated by the Quality Department at the National Institute of Toxicology and Forensic Sciences (INTCF in Spanish) in Madrid. Since 2011 the IE comprises a basic and an advanced level. The “Basic Level” includes a kinship and a forensic module both comprising a practical and a theoretical study. The “Advanced Level” includes a kinship module comprising a kinship paper challenge, and a forensic module that includes both a practical study with samples of increased and diverse complexity, and a theoretical forensic paper challenge.

The objectives of the GHEP-ISFG and its IEs range from addressing technical or theoretical difficulties to developing intercomparison and collaborative exercises for validation of data and new methodologies, plus dealing with ethics and continuing education, among others. The results of several collaborative exercises, as well as different IEs have been published; e.g. [1, 3, 5-9] and have greatly contributed not only to the spread of knowledge but also to encourage many new forensic geneticists from different countries to become GHEP-ISFG members and to participate in the activities of the group.

In the present study we aim to present a detailed analysis of the results obtained for two forensic samples included in the 2014-IE, which were discussed during the XIX Annual Meeting on Forensic Genetics, held in Quito (Ecuador) on September 9-12, 2014. These two samples consisted of stains with mixtures of different biological fluids. Autosomal and sex-chromosomal STRs and mtDNA were analyzed. Here we critically discuss the different types of errors and evaluation criteria.

2. Materials and Methods

2.1. *Intercomparison exercise background and samples*

The present study focused on two out of five forensic items delivered as part of the forensic modules of the 2014-IE, namely M4 from the Basic Level, and M8 from the Advanced Level. The reference samples for comparison with these two items from the forensic modules were three items delivered as part of the *Basic Level Kinship Practical Study*, namely M1, M2 and M3.

The items consisted of: (*i*) M4: 50 µl of a mixture 2:1(v/v) of saliva from a male previously typed for autosomal and Y-chromosomal STRs, and blood from the female donor of sample M3, deposited on a napkin; (*ii*) M8: 50 µl of a mixture 4:1 (v/v) of saliva from an unknown female, and semen from the male donor of sample M1, deposited on a piece of T-shirt (**Figure 1**).

All items were carefully prepared to avoid contamination by (*i*) using sterilized material, (*ii*) using proper protection in the preparation of the samples, (*iii*) preparing all samples separately both in time and space, and (*iv*) vortexing the samples in order to assure homogeneity among the items. The preparation of the samples was undertaken by qualified personnel at the Quality Department of the INTCF (Madrid, Spain) as well as conditioning for shipment, and the delivery of all the items to the participant laboratories.

Participants had to report the genetic profiles obtained for the items using the methods regularly employed in their laboratories. They were asked to determine: (*i*) if any of the forensic items could correspond to mixtures of biological fluids; (*ii*) the number of possible contributors to each sample; (*iii*) the nature of the possible components (“saliva / semen / blood / undetermined”); and (*iv*) if any of the donors of the reference samples M1, M2, M3 (mentioned above) could have

contributed to their genetic profiles. Each laboratory uploaded the results and conclusions in an *ad hoc* online form, and sent a hard copy of the completed form signed by the authorized laboratory personnel to the IE coordinators. Furthermore, it was compulsory to attach copies, either in paper or electronic format (e.g. electropherograms) of the records upon which the reported results were based.

Additionally, details on DNA extraction methods (i.e. automated, manual, total / differential cell lysis, etc.), STR genotyping methodology (e.g. type of kits, detection methods, etc.), and other technical features were requested from the participants. Although this information was not taken into account for grading the performance of the laboratories, it may be useful for the discussion and interpretation of the results.

2.2. *Criteria for interpretation and grading*

The IE coordinator is in charge of the compilation and evaluation of the results. Occasionally, the coordinator may contact an expert in order to assist with the analyses and interpretation of particular aspects of the IE.

A certificate of participation and evaluation of results is issued, where the performance of the laboratory for the different analyses is stated. Thus, the results are assigned into different categories according to pre-established criteria that are stated in the “Participation Rules”, which can be consulted online at http://ghep-isfg.org/en/proficiency/participation. Briefly, the results of the practical exercises of the 2014 IE appeared in the certificates with the following codes:

- ***C***: It matches with the assigned value (formerly called reference value) obtained by consensus or with the known value.
- ***D***: Errors in typing, allelic loss or gain, etc.

- ***N***: Discrepancies due to the use of a nomenclature or a format other than those specified in the instructions.
- ***T***: Transcription errors in completing the form.
- ***NA***: not analyzed.
- ***SR***: no assigned value (formerly reference value).

Prior to grading the results the following definitions were established:

- ***Assigned value***: evaluation will be carried out with regards to assigned values. These values can be established by consensus from participants or from known values.
- ***Consensus values***: to agree on a result it requires a minimum participation of five laboratories and concordance of results of at least 70% of the participants. It is the experience of the group that a consensus value does not always reflect a correct result. Therefore, consensus values are generally discussed within the GHEP-ISFG framework.
- ***Known value***: value that is known.

For laboratories performing DNA extraction using differential lysis methods, the online form for the 2014 IE has separate columns to indicate the total alleles detected and the second (spermatic) fraction, for each marker analyzed in M4 and M8. However, evaluation of results and grading of performance was based only on the alleles reported for the total lysis.

Regarding mtDNA analysis of mixtures, although results obtained by the laboratories are reviewed and informed, they are not included in the certificate.

2.3. *Participants*

A total of 79 and 57 laboratories from seventeen countries enrolled for participation in the Basic and the Advanced Levels of the 2014-IE, respectively, considering the Kinship and the Forensic modules. **Figure 2** summarizes the number of participants that issued results for each type of marker for both M4 and M8.

2.4. Markers analyzed

a) *Autosomal STRs:* Results for at least seven CODIS STRs were necessary for evaluation of the participation, although laboratories were allowed to report results for any other marker used in their routine work. According to the pre-established criteria, 29 (M4) and 28 (M8) autosomal STRs reached the lower threshold for the number of reports required for consensus (**Figure 3**). These markers are the ones included in the commonly used commercial kits –e.g. PowerPlex® 16, PowerPlex® ESX17, PowerPlex® 21, PowerPlex® Fusion (Promega Corp., USA), AmpFLSTR Identifiler and AmpFLSTR Identifiler Plus, AmpFLSTR NGM (Applied Biosystems, USA), etc–. Less than five laboratories analyzed additional twelve and three markers for M4 and M8, respectively.

b) *Y-STRs:* Most laboratories genotyped the samples using the AmpFLSTR Y-filer kit (Applied Biosystems, USA). The second most used commercial kit was the PowerPlex Y23 (Promega Corp., USA), and only a few laboratories used the PowerPlex Y12 (Promega Corp., USA), or the Investigator Argus Y12 (Qiagen).

c) *X-STRs:* Eighteen laboratories analyzed the X-STR decaplex optimized by the GHEP-ISFG [10], and a minor group analyzed the commercial kit most

widely used in forensics, namely Investigator Argus X12 (Qiagen). These two multiplexes share four markers.

d) *mtDNA*: it was mandatory for the laboratories to report the first and the second hypervariable regions (HVS-I range 16024-16365 / HVS-II range 73-340). HVS-III (range 380-580) could also be voluntarily reported, as well as the entire control region. A total of 26 and 28 laboratories reported results for the M4 and the M8 sample, respectively. Most reported the sequence range 16024-16365 for the HVS-I and 73-340 for the HVS-II; only a few laboratories reported the sequence of the full control region or the HVS-I/II plus the third hypervariable region (HVS-III).

3. Results

3.1. Autosomal STRs and amelogenin

Records submitted by participants for autosomal markers and amelogenin summed-up to 1,521 and 1,116 genotypings for M4 and M8, respectively. Classification of these results according to the above mentioned categories is summarized in **Table 1**. A remarkably high percentage of the results (93.1% for M4 and 94.7% for M8) fell in the ***"C"*** (***correct***) category, either by agreement with the consensus or with a known value. For M4, one laboratory reported the number of alleles detected in each marker instead of the allele names. In a preliminary assessment, these results were assigned to the ***"D"*** (***discordant***) group. However, since the electropherograms submitted showed concordant results with the consensus, they were finally included in the ***"C"*** group, considering that this could be due to a misinterpretation of the required information.

Nomenclature errors (category ***"N"***) concerning non-adherence to the rules for reporting the genotypes were due to: (*i*) lack of separator between alleles in

heterozygote genotypes, (*ii*) use of an incorrect separator (e.g. use of a dash or a comma instead of the required slash symbol) in heterozygote genotypes, or (*iii*) double reporting the unique allele in homozygote genotypes. Cross-checking the results with the submitted electropherograms showed that transcription discrepancies (category ***"T"***) were either due to typos (e.g. "33.2" transcribed as "33.3") or to inversion in the order of markers (i.e. results of one locus transcribed into another locus and *vice versa*).

Regarding discordant results (category ***"D"***), a deep scrutiny was carried out in order to reveal the possible reason for the genotyping errors. Thus, by individually analyzing the records that neither matched the consensus nor the known values, a number of causes of error could be identified (**Table 2**). Due to the different nature of the mixtures M4 and M8, causes of errors were separated into different classes. Thus, drop-out of the minor component (single allele or complete genotype) was the main source of error in both samples, followed by the inability to discriminate stutter peaks from true alleles.

It is noteworthy that discrepant results for M4 from one participant laboratory were due to sample mix-up (with an exogenous sample to the IE), since the profile reported –and the supporting electrophoretic records– were completely different from the reference profile. Considering that this laboratory reported results for 20 markers, this sample mix-up accounted for nearly 20% of the errors in category ***"D"*** (20 out of 98; **Table 1**). If this error is counted as a single inconsistency (that is, as a single contamination event) the discrepancies observed in category ***"D"*** would fall-down to 79 (instead of 98), decreasing from 6.5% to 5.2%. Similarly, one laboratory only reported the profile of the main male contributor in M8, then concentrating 14 discrepancies in the same laboratory.

Concerning the number of laboratories that reported incorrect results, 46% (*n* = 33) showed discordant results for M4, while 35% (*n* = 18) reported discordant results for M8. **Figure 4** shows that the majority of the laboratories reported only one discrepancy, both in M4 and M8; the two outliers with 20 and 14 discrepancies correspond to the sample mix-up for M4 plus the participant that detected only the male contributor in M8 mentioned above.

Analyzing the discrepancies by autosomal STR marker, it was observed for M4 that a considerable proportion (19%) was concentrated in locus D12S391 (**Figure 5**). As discussed below, this was mainly due to difficulties in assigning microvariant alleles in the profile of the minor contributor.

3.2. Y-STRs

Only 13 discrepancies were observed in M4, of which 12 were due to the single sample mix-up event mentioned above in regards to the autosomal STRs. The remaining discrepancy for M4 was due to a transcription error in completing the form, and notably no discrepancies were observed for M8 (**Tables 1, 2** and **3)**.

3.3. X-STRs

For sample M4, complete loss of the minor component was the main cause of discrepancy for X-STRs (16/35), followed by the lack of capacity to discriminate stutter peaks from true alleles (13/35) (**Table 2**). The latter was concentrated in two markers, namely GATA31E08 and DXS6809 (12/35).

The situation was similar for M8, although the proportion of discrepancies for all the genotypes observed for M8 was roughly half of that for M4 (7.8% *vs.* 15.4%) (**Table 1**).

3.4. Mitochondrial DNA

It is worth mentioning that mtDNA in mixture samples is not included in the evaluation of the IE. This is owed to the difficulties in interpreting mixtures from haplotype sequences from both technical [11] and theoretical [12] points of view.

As expected, the mtDNA results were problematic. For M4, the consensus haplotype of the blood donor (female reference sample M3 of this exercise) was 73G-189G-194T-195C-199C-204C-207A-263G-315.1C for the HVS-II and 16223T-16292T for the HVS-I (then ascribed to haplogroup W3b/W3a1b), and the haplotype of the saliva donor (unknown man) was 72T-150T-263G-309.1C-309.2C-315.1C for the HVS-II and 16183del-16298C for the HVS-I (haplogroup HV0). The expected haplotype pattern for the mixture M4 should be 72Y-73R-150Y-189R-194Y-195Y-199Y-204Y-207R-263G-315.1C for the HVS-II and 16183a-16223Y-16292Y-16298Y; 16183a denoting an heteroplasmic-like pattern at this site where an haplotype with an adenine at this position mixes with another haplotype with a deletion at the same site; see updated nomenclature rules in [13]. Only one participant reported this mixed profile, and therefore, a consensus could not be reached. Some laboratories reported only one haplotype, or haplotypes different than expected. In addition, several nomenclature problems were also common; e.g. heteroplasmies indicated as C/T instead of Y [13].

For M8, the haplotype from the saliva donor (unknown woman) was 73G-152C-263G-309.1C-315.1C for the HVS-II and 16126C-16294T-16304C for the HVS-I, and the haplotype from the semen donor (male reference sample M1 of this IE) was 263G-315.1C for the HVS-II and 16129A for the HVS-I. Most laboratories (22 out of 28) detected only the saliva profile. Few laboratories reported incomplete haplotypes of only one region and one laboratory detected only the semen profile.

4. Discussion

In this study we analyzed the autosomal, sex-chromosomal and mtDNA results submitted by the participants of the 2014 GHEP-ISFG IE for two mixture samples of the Forensic Module of the Basic and Advanced Level.

Overall, the highest error rates among nuclear markers were observed for X-STRs (15.4% for M4 and 7.8% for M8), and the lowest ones were for Y-STRs (1.2% for M4 and 0.0% for M8), while autosomal STRs and amelogenin showed intermediate values (6.5% for M4 and 4.7% for M8).

Most problems associated with X-chromosome markers occurred at GATA31E08 and DXS6809. It is noteworthy that the performance of these two markers in non-mixed samples is usually correct, but it turns to be challenging in mixtures. The high numbers of uninterrupted repeats of the alleles present at these two loci were responsible for the prominent stutter peaks. Indeed, all errors detected in these two markers were due to an incorrect attribution of the stutter peak being a real allele or *vice-versa*. Overall, the results suggest that, although these markers are useful in kinship or individual identification involving unique samples, special attention should be paid in mixture interpretation because of stutter ratio variation among alleles with different sequences and sizes. Furthermore, these kind of discrepancies could be reduced, at least in part, if laboratories perform their own validation studies to determine the stutter thresholds for each kit or group or markers.

With respect to the Y-chromosome markers, it is important to consider that there was only one male contributor in M4 and in M8. Therefore, it was expected that neither of these samples should present the same level of complexity as for

autosomal and X-STRs. This was reflected by a remarkably good performance of the laboratories for the male lineage markers.

For autosomal STRs, similar proportions for each of the pre-defined grading categories (***"C"***, ***"D"***, ***"T"***, ***"N"*** and ***"SR"***) were observed for M4 and M8. It was anticipated that sample M8 would be more complex than M4. However, comparing the proportion of records in category ***"D"*** for M4 and M8, a slightly lower value was observed for M8, even counting the sample mix-up event for M4 as a single record (instead of 20; see above). Also, errors in M8 were concentrated in a lower fraction of participants than for M4 (35% *vs.* 46%, respectively). In addition, 67% of the laboratories (n = 8) that reported incorrect results for M8 (n = 12) also reported errors in M4, which represents 36% of the participants that reported errors in M4. It could be tentatively argued that this outcome might be related to a greater proportion of laboratories contributing to the Forensic Module of the Advanced Level with better expertise than the proportion of laboratories involved in the Basic Level.

The main causes of inconsistencies at autosomal STRs were due to drop-out problems –either allelic or genotypic– related to the minor component of the mixture. Several studies have reported difficulties in assigning complete genotypes in complex samples and / or highly unbalanced mixtures (e.g. [14-16]). It is important to note that items M4 and M8 were prepared in a way to allow easy detection of both components of the mixtures; this is reflected on the considerable number of laboratories that reported full consensus profiles, and the overall high quality of the submitted electropherograms. Therefore, incorrect results obtained by a few laboratories can most probably be attributed to a sub-optimum procedure of DNA extraction in mixture samples; however, although DNA extraction is one of

the critical issues that could influence the quality of the profiles, we cannot be certain about this because the information needed to evaluate potential problems related to DNA extraction methods was not available for review (e.g. quantitation data). Also plausible would be the existence of some kind of stochastic effects related with the time of arrival of the samples to the different laboratories, PCR, genotyping kit employed, etc.

There were some problematic issues regarding microvariants occurring at autosomal STRs; in particular, in the ability to assign microvariant alleles of the minor profile, when the microvariants occurred one or two bases away from the complete allele. This was particularly evident at D12S391 and FGA loci in M4. Specifically, the consensus profile at D12S391 in M4 was “17 / 17.3 / 18 / 20”. As an example, the electropherograms in **Figure 6** indicate that the electrophoresis resolution was insufficient to allow discrimination of allele “17.3” from “18”, thus accounting for almost 20% of the inconsistencies. Interestingly, one participant discussed the allele “18” discrepancy at D12S391. This laboratory reported that the analyses of this marker for M4 were carried out using both the PowerPlex ESX17 (Promega Corp., USA) and the AmpFLSTR NGM kit (Applied Biosystems USA), and reported that the D12S391 marker was individually analyzed in a single-plex reaction. Moreover, they created an artificial mixture in proportion 3:1 with “M3” (“17.3 / 20”) and the positive control 007 of the NGM kit (18/19). This laboratory demonstrated that the “18” allele could be assigned by the software only when using the Powerplex ESX17 kit. Besides, some laboratories did not detect allele “19.2” at FGA, for a consensus profile “19.2 / 20 / 23 / 25” where the alleles “19.2” and “23” correspond to the minor contributor. However, in most of

the cases, this allele was present under the analytical threshold of the laboratories and consequently was not reported.

Regarding M8, some particular features emerged from a detailed inspection of the electropherograms. Although it was not subject to evaluation, a marked disparity in the proportion of the major / minor components among the participants was observed, reflected by the peak heights / areas in the corresponding electropherograms. An illustrative example is shown in **Figure 7**, where the same section of the electropherograms for the AmpFLSTR Identifiler Kit (Applied Biosystems, USA) submitted by two laboratories is shown. The X / Y chromosome peak area ratio in amelogenin might be indicative that differences could be due to the nature of the biological fluids in sample M8 –saliva and semen 4:1 (v/v)– coupled with the DNA extraction procedure used by different laboratories. It is worth mentioning that special care was taken during preparation of the samples in order to maintain homogeneity as much as possible among and within the items.

While it is interesting to see the significant variation in electropherograms between laboratories when analyzing the same material, one can speculate that similar differences could be observed within the same laboratory when repeating analyses, and especially if the samples are re-extracted.

It is also important to note that no peak heights are indicated in the electropherograms of “Laboratory 1” in **Figure 7**, although this is mandatory for participants. This is not an isolated case; there are some other laboratories that did not adhere to this requirement. Finally, the analyses of electropherograms from laboratories reporting genotyping problems revealed that, in general, most of the inconsistencies were due to low quality results associated, in some cases, to manual allele calling.

The 22nd GHEP-ISFG IE on DNA mixtures allowed elaborating several conclusions. Among nuclear markers, the best performance was by far for Y-STR typing, while the X-STR markers were the most problematic. There was a high proportion of consensus results for autosomal markers. The main errors were due to single and complete drop-out of minor contributors' alleles, followed by difficulties in differentiating stutter peaks from true alleles, and a low capillary electrophoresis resolution for particular alleles.

The results of mtDNA analysis were very problematic, in good agreement with previous exercises out by the GHEP-ISFG on mtDNA mixtures [11]. Nomenclature problems are common in these exercises. Some contamination and sample mix-up issues were also detected in previous editions of the IE and as described in the literature [6, 17, 18]. In semen-saliva mixtures, most of the laboratories were unable to detect the haplotype from the semen.

As in previous editions of the GHEP-ISFG IE, an important effort is being carried out in order to detect the main difficulties of laboratories in the analyses of complex samples. This effort could focus, among others, on the organization of training courses, and the elaboration of specific documentation by experts; overall a set of actions aiming to prevent these issues in forensic casework.

5. Acknowledgements

The authors declare no conflicts of interest.

6. References

[1] A. Alonso, and C. Albarrán, The Spanish and Portuguese ISFG Working Group: Ten Years Coordinating DNA Typing In Spain, Portugal and Latin America, Profiles in DNA (2000)

[2] J. Gómez, and Á. Carracedo, The 1998-1999 collaborative exercises and proficiency testing program on DNA typing of the Spanish and Portuguese Working Group of the International Society for Forensic Genetics (GEP-ISFG), Forensic Sci. Int. 114 (2000) 21-30.

[3] J. Gómez, M.S. Rodriguez-Calvo, C. Albarran, A. Amorim, J. Andradas, C. Cabrero, R. Calvet, D. Corach, M. Crespillo, C. Doutremepuich, et al., A review of the collaborative exercises on DNA typing of the Spanish and Portuguese ISFH Working Group. International Society for Forensic Haemogenetics, Int J Legal Med 110 (1997) 273-277.

[4] A. Alonso, A. Salas, C. Albarrán, E. Arroyo, A. Castro, M. Crespillo, A.M. di Lonardo, M.V. Lareu, C.L. Cubria, M.L. Soto, et al., Results of the 1999-2000 collaborative exercise and proficiency testing program on mitochondrial DNA of the GEP-ISFG: an inter-laboratory study of the observed variability in the heteroplasmy level of hair from the same donor, Forensic Sci. Int. 125 (2002) 1-7.

[5] L. Prieto, M. Montesino, A. Salas, A. Alonso, C. Albarrán, S. Álvarez, M. Crespillo, A.M. Di Lonardo, C. Doutremepuich, I. Fernández-Fernández, et al., The 2000-2001 GEP-ISFG Collaborative Exercise on mtDNA: assessing the cause of unsuccessful mtDNA PCR amplification of hair shaft samples, Forensic Sci. Int. 134 (2003) 46-53.

[6] A. Salas, L. Prieto, M. Montesino, C. Albarrán, E. Arroyo, M.R. Paredes-Herrera, A.M. Di Lonardo, C. Doutremepuich, I. Fernández-Fernández, A.G. de la Vega, et al., Mitochondrial DNA error prophylaxis: Assessing the causes of errors in the GEP'02-03 proficiency testing trial, Forensic Sci. Int. 148 (2005) 191-198.

[7] M. Montesino, A. Salas, M. Crespillo, C. Albarrán, A. Alonso, V. Álvarez-Iglesias, J.A. Cano, M. Carvalho, D. Corach, C. Cruz, et al., Analysis of body fluid mixtures by mtDNA sequencing: An inter-laboratory study of the GEP-ISFG working group, Forensic Sci. Int. 168 (2007) 42-56.

[8] M. Crespillo, M.R. Paredes, L. Prieto, M. Montesino, A. Salas, C. Albarrán, V. Álvarez-Iglesias, A. Amorin, G. Berniell-Lee, A. Brehm, et al., Results of the 2003-2004 GEP-ISFG collaborative study on mitochondrial DNA: focus on the mtDNA profile of a mixed semen-saliva stain, Forensic Sci. Int. 160 (2006) 157-167.

[9] M. Crespillo, P.A. Barrio, J.A. Luque, C. Alves, M. Aler, F. Alessandrini, L. Andrade, R.M. Barretto, A. Bofarull, S. Costa, et al., GHEP-ISFG collaborative exercise on mixture profiles of autosomal STRs (GHEP-MIX01, GHEP-MIX02 and GHEP-MIX03): results and evaluation, Forensic Sci Int Genet 10 (2014) 64-72.

[10] L. Gusmão, P. Sánchez-Diz, C. Alves, I. Gomes, M.T. Zarrabeitia, M. Abovich, I. Atmetlla, C. Bobillo, L. Bravo, J. Builes, et al., A GEP-ISFG collaborative study on the optimization of an X-STR decaplex: data on 15 Iberian and Latin American populations, Int J Legal Med 123 (2009) 227-234.

[11] L. Prieto, A. Alonso, C. Alves, M. Crespillo, M. Montesino, A. Picornell, A. Brehm, J.L. Ramírez, M.R. Whittle, M.J. Anjos, et al., 2006 GEP-ISFG Q1 collaborative exercise on mtDNA. Reflections about interpretation, artefacts, and DNA mixtures, Forensic Sci. Int. Genet. 2 (2008) 126-133.

[12] T. Egeland, and A. Salas, A statistical framework for the interpretation of mtDNA mixtures: forensic and medical applications, PLoS One 10 (2011) e26723.

[13] W. Parson, L. Gusmão, D.R. Hares, J.A. Irwin, W.R. Mayr, N. Morling, E. Pokorak, M. Prinz, A. Salas, P.M. Schneider, et al., DNA Commission of the International Society for Forensic Genetics: Revised and extended guidelines for mitochondrial DNA typing, Forensic Sci. Int. Genet. 13 (2014) 134-142.

[14] P. Gill, J. Whitaker, C. Flaxman, N. Brown, and 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.

[15] P. Gill, L. Gusmão, H. Haned, W.R. Mayr, N. Morling, W. Parson, L. Prieto, M. Prinz, H. Schneider, P.M. Schneider, et al., 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.

[16] H. Haned, T. Egeland, D. Pontier, L. Pene, and P. Gill, Estimating drop-out probabilities in forensic DNA samples: a simulation approach to evaluate different models, Forensic Sci Int Genet 5 (2011) 525-531.

[17] Q.-P. Kong, A. Salas, C. Sun, N. Fuku, M. Tanaka, L. Zhong, C.-Y. Wang, Y.-G. Yao, and H.-J. Bandelt, Distilling artificial recombinants from large sets of complete mtDNA genomes, PLoS One 3 (2008) e3016.

[18] H.-J. Bandelt, M. van Oven, and A. Salas, Haplogrouping mitochondrial DNA sequences in Legal Medicine/Forensic Genetics, Int J Legal Med 126 (2012) 901-916.

Figure Legend

Figure 1. The forensic samples M4 and M8. M4 was prepared as follows: 50 µl of a mixture 2:1(v/v) of saliva from a male previously typed for autosomal STRs, and blood from the female donor of sample “M3”, deposited on a napkin. M8 was prepared as follows: 50 µl of a mixture 4:1 (v/v) of saliva from an unknown female, and semen from the male donor of sample “M1”, deposited on a piece of T-shirt

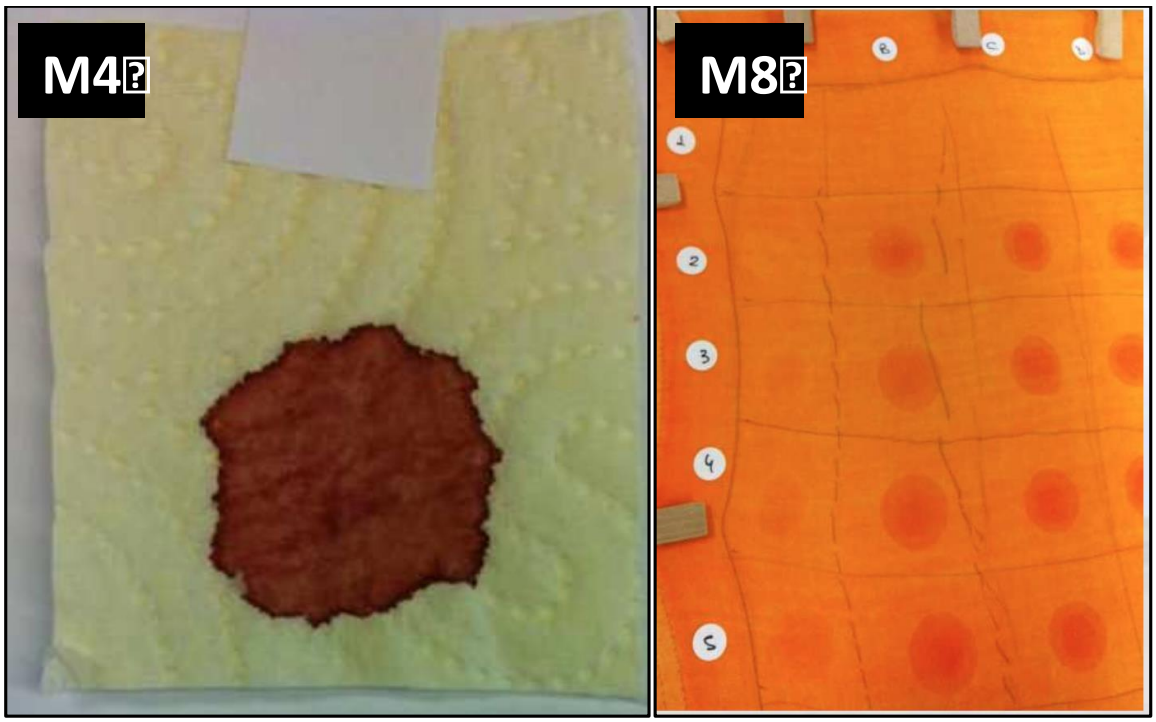

Figure 2. Laboratories that reported results for M4 and M8.

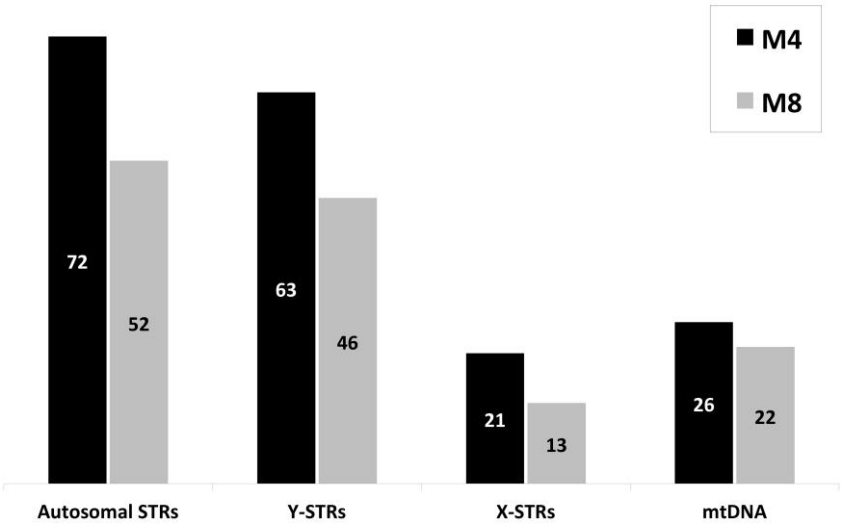

Figure 3. Autosomal STRs and number of laboratories that analyzed each marker for M4 and M8.

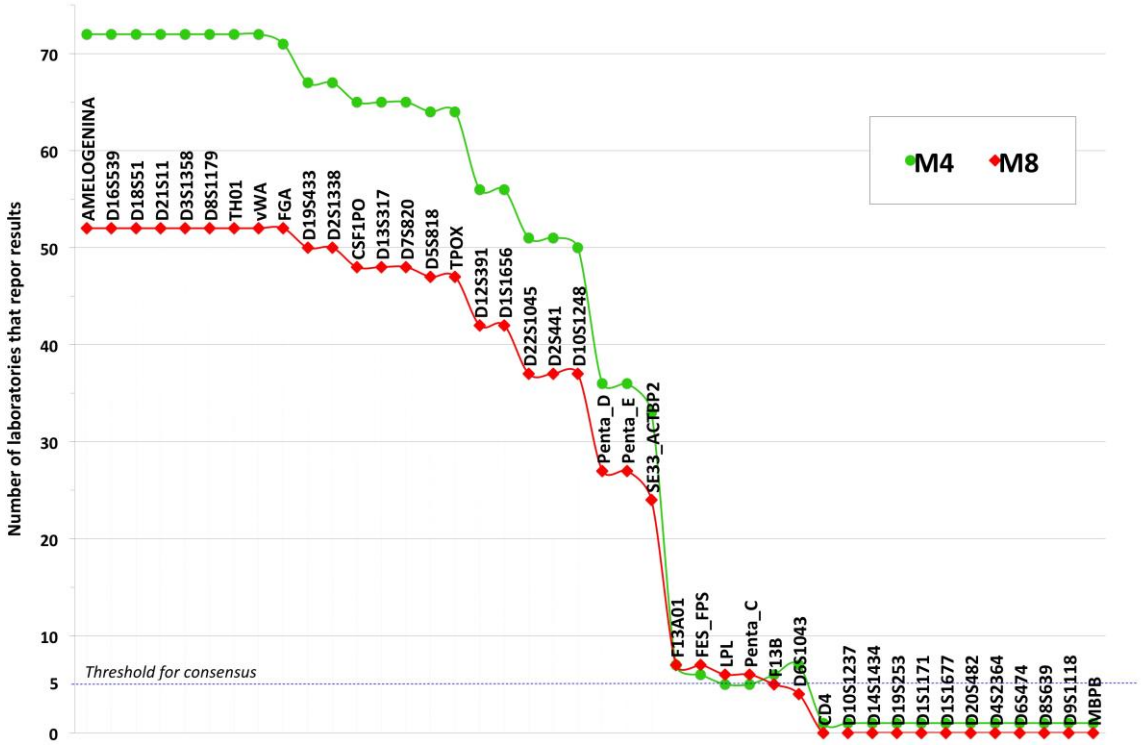

Figure 4. Number of laboratories with different number of discrepancies (category ***"D"***) for autosomal STRs.

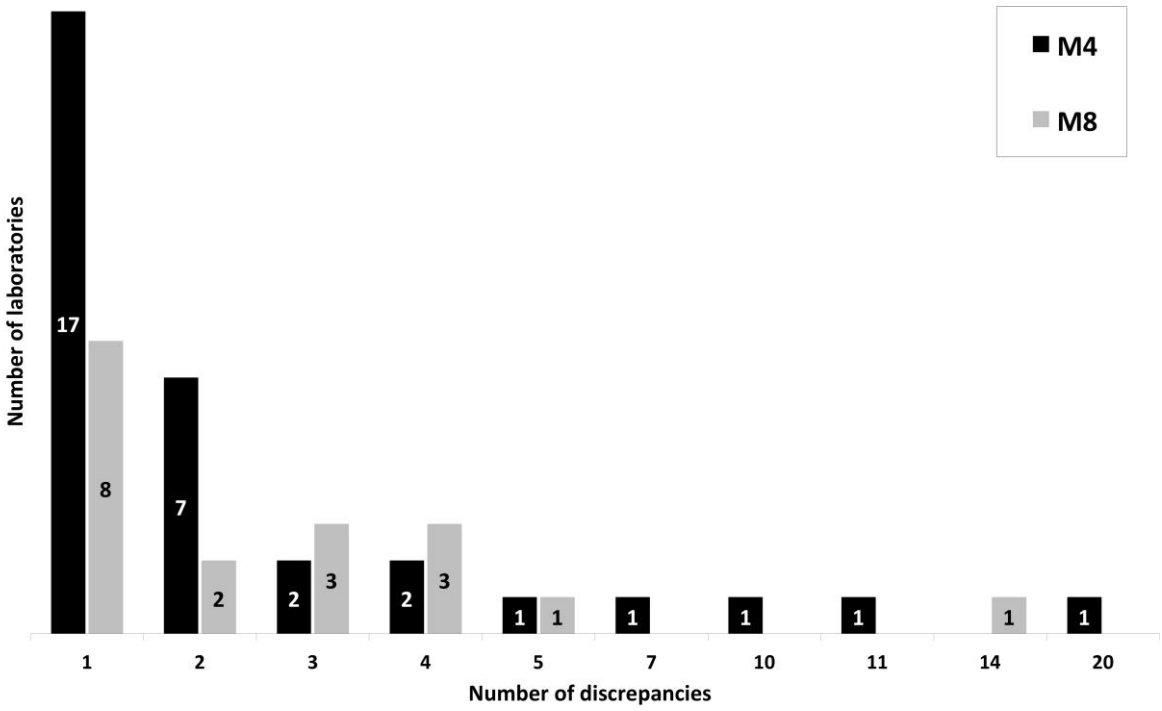

Figure 5. Records in category ***"D"*** grouped by autosomal STR markers.

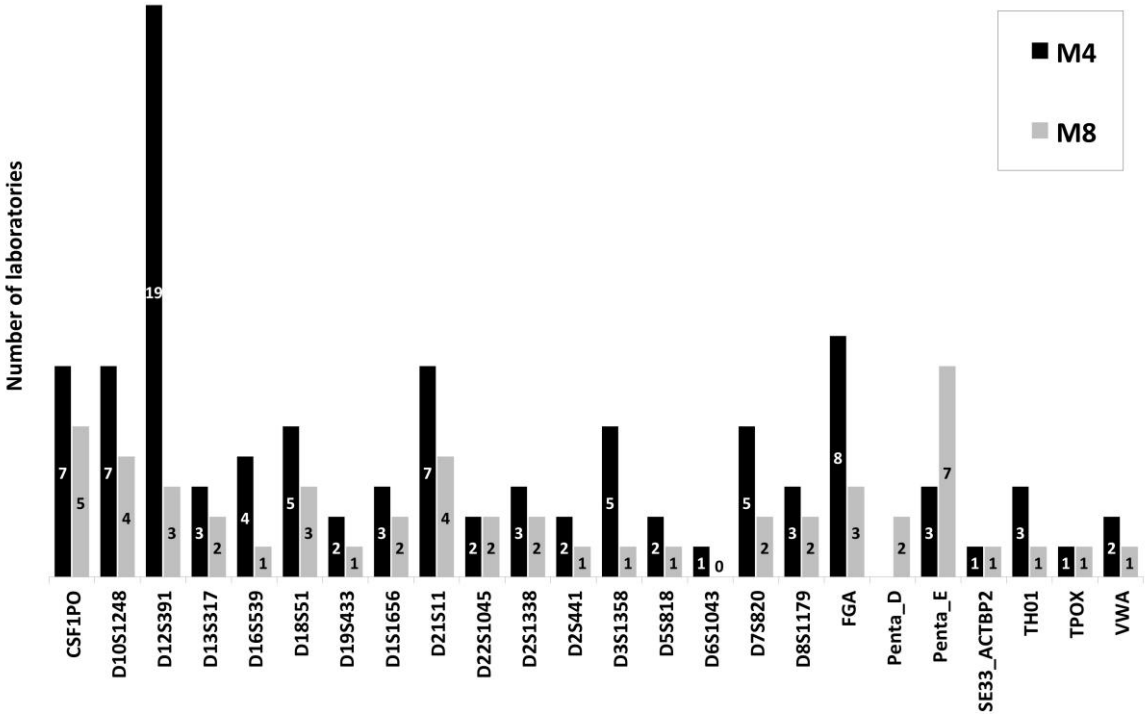

Figure 6. Examples of microvariants in minor contributor not reported at locus D12S391 in M4.

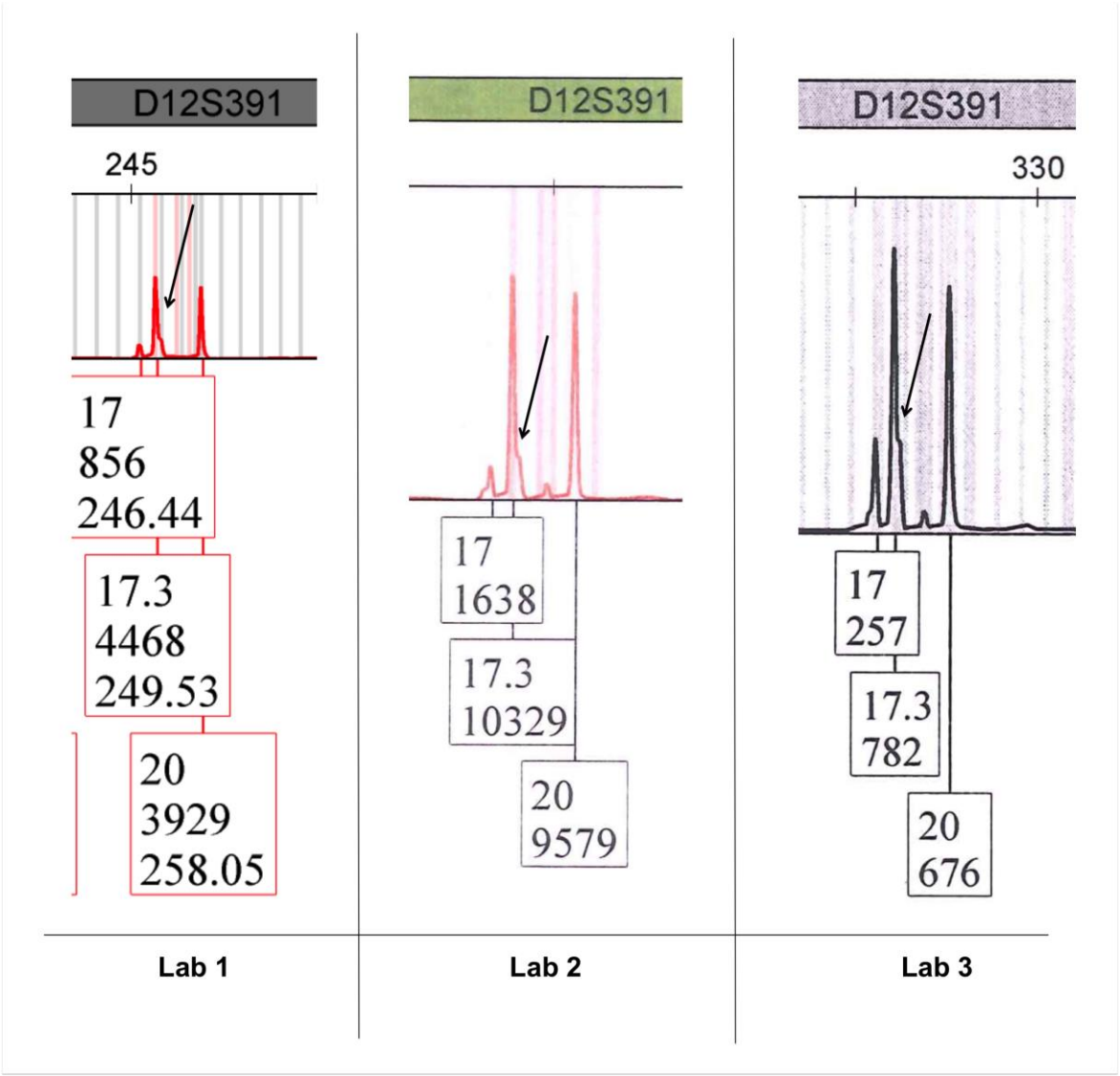

Figure 7. Examples of different proportions of contributors detected among two participants for M8 (results of the same section of the electropherogram submitted by the two laboratories are shown, AmpfLSTR Identifiler kit, Applied Biosystems, USA).

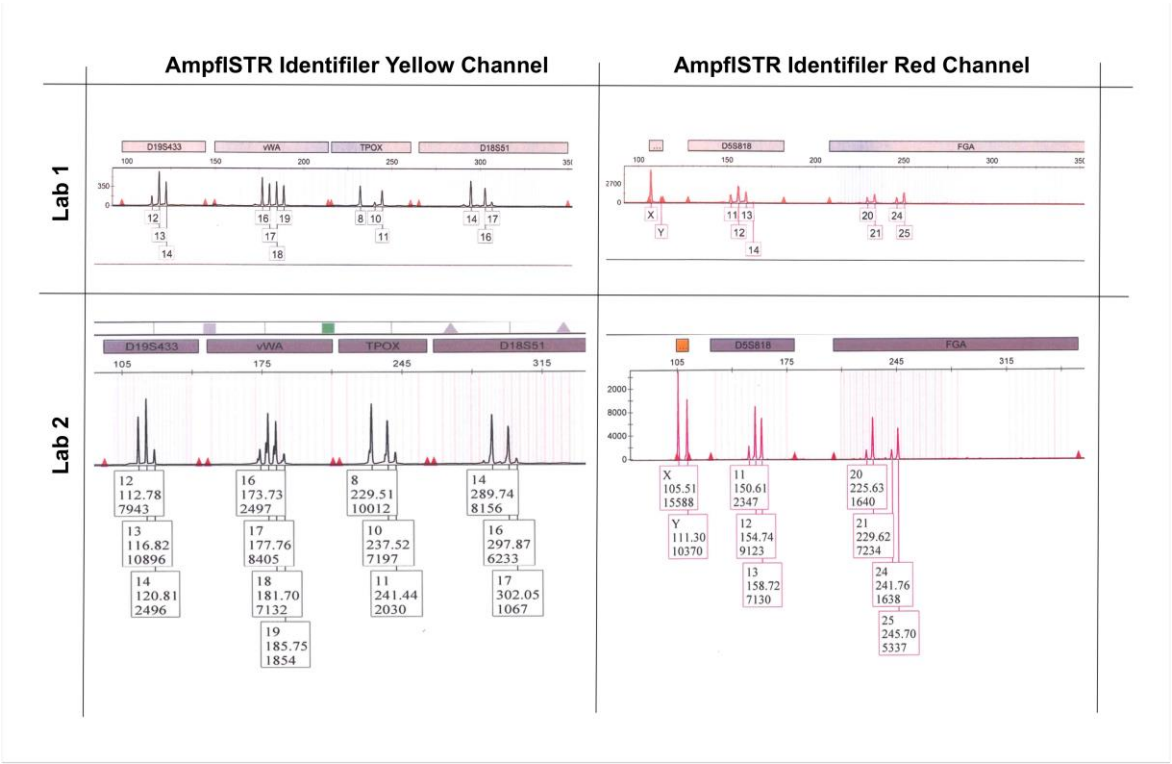

Table 1. Results for M4 and M8 classified into pre-established categories.

	Autosomal STRs		***Y - STRs***		***X - STRs***	
	M4	***M8***	***M4***	***M8***	***M4***	***M8***
Category	***N (%)***	***N (%)***	***N (%)***	***N (%)***	***N (%)***	***N (%)***
C	*1405 (93.1)*	*1051 (94.7)*	*1032 (98.6)*	*790 (99.9)*	*192 (84.6)*	*106 (92.2)*
D	*98 (6.5)*	*52 (4.7)*	*13 (1.2)*	*0 (0.0)*	*35 (15.4)*	*9 (7.8)*
N+T	*6 (0.4)*	*7 (0.6)*	*2 (0.2)*	*1 (0.1)*	*0 (0.0)*	*0 (0.0)*
Subtotal	*1509*	*1110*	*1047*	*791*	*227*	*115*
SR	*12 (0.8)*	*6 (0.5)*	*6 (0.6)*	*3 (0.4)*	*9 (3.8)*	*34 (22.8)*
TOTAL	*1521*	*1116*	*1053*	*794*	*236*	*149*

"C": matches with the assigned value; "***D"***: Errors in genotyping, allelic loss or gain, etc., "***N+T"***: Discrepancies due to the use of a nomenclature or a format other than those specified in the instructions, or to transcription errors in completing the form. "***SR"***: no assigned value.

Table 2. Classification of discrepancies (***"D"*** *Category*) for M4 for autosomal and sex chromosomal STRs.

Identified cause of discrepancy for M4	***Autosomal STRs***	***Y - STRs***	***X - STRs***
Single allele drop-out of the minor contributor	*37*	*-*	*-*
Peak in stutter position assigned as allele	*21*	*-*	*13*
Sample mix-up (one participant)	*20*	*12*	*-*
Only major contributor genotype detected / assigned	*12*	*-*	*16*
Allele drop-out & allele drop-in at the same marker	*2*	*-*	*-*
Allele drop-in	*2*	*-*	*-*
Allele of minor contributor in stutter position not assigned	*2*	*-*	*5*
Other	*2*	*1*	*1*
TOTAL	*98*	*13*	*35*
Number of consensus / known genotypes reported for each type or marker	*1509*	*1047*	*227*

Table 3. Classification of discrepancies (***"D"*** *Category*) for M8 for autosomal and sex chromosomal STRs.

Identified cause of discrepancy for M8	***Autosomal STRs***	***Y - STRs***	***X - STRs***
Only Male (semen) contributor detected / assigned	*15*	*-*	*3*
Single allele drop-out of the female (saliva) contributor profile	*10*	*-*	*3*
Peak in stutter position assigned as allele	*9*	*-*	*1*
Allele drop-in	*9*	*-*	*2*
Allele drop-out of one allele of the male (semen) contributor	*4*	*-*	*-*
Only female (saliva) contributor detected	*2*	*-*	*-*
Microvariant next to complete allele of main contributor not assigned	*2*	*-*	*-*
Allele of minor contributor in stutter position not assigned	*1*	*-*	*-*
TOTAL	*52*	*0*	*9*
Number of consensus / known genotypes reported for each type or marker	*1110*	*791*	*115*