

Development and validation of a novel 12 to 23 Y-STR multiplex and its forensic casework application

Belén Navarro-López

University of the Basque Country UPV/EHU

Miriam Baeta

University of the Basque Country UPV/EHU

Eva Granizo-Rodríguez

University of the Basque Country UPV/EHU

Olatz Moreno-López

University of the Basque Country UPV/EHU

Tamara Kleinbielen

University of the Basque Country UPV/EHU

Joana Francesca Ferragut

Institut Universitari d'Investigació en Ciències de la Salut (IUNICS) i Laboratori de Genètica, Universitat de les Illes Balears

Antònia Picornell

Institut Universitari d'Investigació en Ciències de la Salut (IUNICS) i Laboratori de Genètica, Universitat de les Illes Balears

Marian M de Pancorbo (✉ mariancpancorbo@gmail.com)

University of the Basque Country UPV/EHU

Article

Keywords: Y chromosome, Y-STRs, forensics, validation, 23 Y-STRs

Posted Date: May 9th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1621099/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Y chromosomal short tandem repeats (Y-STRs) are used in forensic investigations as a useful complementary tool to autosomal markers. The ongoing development of new kits with an increasing number of markers makes it necessary to update populations typed in the Y-STR Haplotype Reference Database to reach at least 23 Y-STRs. A novel Y-STR multiplex panel was developed to offer a cost-efficient alternative to update Y-STR haplotypes from 12 to 23 loci. This panel includes the eleven markers, DYS448, DYS456, DYS458, DYS635, Y-GATA H4, DYS576, DYS481, DYS549, DYS533, DYS570 and DYS643, as well as DYS385a/b for traceability purpose. Developmental validation of this panel was conducted following the recommendations of the Scientific Working Group on DNA Analysis Methods (SWGDAM), showing highly sensitivity, tolerance to common inhibitors as well as high species specificity. It was efficient for degraded DNA samples and for detection of male mixtures. When applying it for extending the current data of the Ibiza population, both the discrimination capacity and the haplotype diversity increased from 0.5952 to 0.9048 and from 0.9808 to 0.9977, respectively. Together, the study demonstrates the suitability of this panel in forensic casework.

Introduction

Y chromosome polymorphisms have become a valuable tool in forensic, anthropological, evolutionary, and genealogical studies ^{1,2}. Due to its paternal inheritance, they elude recombination to a great extent, being able to discriminate between male lineages, preserving a simple record of their history and being useful tools for understanding the geographical distribution of current populations ³. Y-STRs, with mutation rates in the order of $1 \cdot 10^{-3}$ per locus per generation, have strong power of discrimination among paternal unrelated lineages and play a key role in the field of forensic genetics ^{2,4-7}.

Although Y-STRs do not allow individualization as autosomal markers do, they may be useful in certain situations. For instance, for male donor analysis in male-female DNA mixtures, such as in cases of sexual assault, paternity testing with male offspring ⁸⁻¹¹ and disaster victims or missing person investigations, involving males ¹². Furthermore, Y-STR information has proven to be helpful in rapidly determining the number of male contributors in mixtures. Being hemizygous, Y-STR profiles have mostly one allele in their loci, so the presence of multiple alleles at single-copy loci is a clear clue of the number of male contributors ^{11,13,14}. Y-STR analysis can also aid in reconstructing paternal relationship and even sometimes be used to infer the geographic region of paternal ancestry of a male DNA donor, useful in the case of missing persons ^{1,10,11}.

Over the past decades, thousands of Y-STRs have been identified and several commercially available Y-STR kits have been developed ⁸. Currently, most studies related to the analysis of Y-STRs aim to search for new markers to develop multiplex panels that offer greater discriminatory power, as it has been observed that a low number of Y-STRs may be insufficient for differentiating male lineages ¹⁵. This is the reason why more markers are needed than those included initially in the 9 Y-STRs of the Minimal

Haplotype, the 12 Y-STRs included in the PPY (PowerPlex®Y, Promega Corporation, Madison, WI, USA), or the 17 Y-STRs of the Yfiler® kit (Yfiler, Life Technologies, Foster City, CA)¹. Indeed, the development of new kits with a larger set of markers has been shown to provide higher genetic resolution by increasing both haplotype diversity and discriminatory power compared to other panels in numerous populations^{16–21}.

Despite the foregoing, where more markers analysed usually means greater discriminatory power, to date numerous Y-STR population data available in the Y-STR Haplotype Reference Database (www.yhrd.org) are still based on the 12 Y-STRs of the PPY. In fact, currently the YHRD (Release R67) still presents 1398 population samples including 9 Y-STRs (Minimal Haplotype), 1177 with 12 Y-STRs (PPY) and 1089 containing 17 Y-STRs (Yfiler). However, only 402 samples report the 23 Y-STRs of the PPY23 (PowerPlex® Y23 System, Promega Corporation, WI, USA) and 314 the 27 Y-STRs of the YfilerPlus²². The ideal scenario would be for the same samples to be regenotyped as new markers appear. This would allow the databases to be updated continuously, which would be of great advantage in the forensic field. However, re-analysing all populations with the new Y-STR kits that are becoming commercially available is a labour-intensive, time-consuming and expensive task.

Here, we present the development and validation of a novel 12 Y-STR multiplex, an extension of the one previously developed by Nuñez et al. (2017)¹⁵. The extended panel includes the eleven markers, DYS448, DYS456, DYS458, DYS635, Y-GATA H4, DYS576, DYS481, DYS549, DYS533, DYS570 and DYS643, which are contained in the PPY23 but absent in the PPY, as well as DYS385a/b for traceability purposes (Fig. 1). The number of markers has been extended to 23 because, as previously mentioned, among the kits with a high number of markers, PPY23 is the one with the largest proportion of samples analysed. This would allow us to perform comparative analyses of the updated populations with a greater number of individuals. The primer design was based on a miniSTR strategy in order to increase the performance success of this multiplex in degraded DNA. The developmental validation was performed following the guidelines of the SWGDAM²³. The new multiplex was applied in a population that was previously analyzed for the 12 Y-STRs of PPY with the aim to increase the current genetic data as well as to evaluate the capability of the multiplex.

Material And Methods

Selection of markers and primer design

The new multiplex panel includes the eleven markers DYS448, DYS456, DYS458, DYS635, Y-GATA H4, DYS576, DYS481, DYS549, DYS533, DYS570 and DYS643, which are contained in the PPY23 but absent in the PPY. Additionally, the loci DYS385a/b was included in the multiplex reaction as a traceability marker, as it is present in the aforementioned commercial kits.

Since the 12 Y-STR multiplex panel is an extension of the earlier one developed by Nuñez et al. (2017), for the eight markers that were already contained in that panel (DYS456, DYS576, DYS481, DYS549,

DYS533, DYS570, DYS643 and DYS385a/b) the same primers described were used¹⁵. For the remaining four Y-STR loci analyzed (DYS448, DYS458, DYS635 and Y-GATA H4), the primer design was performed using PerlPrimer v1.1.21²⁴. Potential primer cross-reactions were examined via this software, and the Y-chromosome specificity was evaluated using BLASTN alignment tool (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). Forward primers were labeled with a fluorescent dye at the 5' end (Supplementary Table S1).

Singleplex PCR reaction

Each primer pair was initially tested in a singleplex PCR reaction using the 2800M control DNA (Promega Corporation, Madison, WI). The singleplex reaction consisted of 12.5 µl of QIAGEN Multiplex PCR kit (Qiagen, Valencia, CA), 1 µl of primer (forward and reverse with an initial concentration of 10 µM), 1 ng of genomic DNA and 10.5 µl of Milli-Q water for a final reaction volume of 25 µl. PCR was performed in a GeneAmp 9800 (AB/LT/TFS) under the following cycle conditions: initial denaturation at 95°C for 15 min was followed by 30 cycles of 94°C for 30 s, 65°C for 90 s, and 72°C for 90 s, and a final extension at 72°C for 10 min. Evaluation of amplification performance was conducted by electrophoresis on 1.5% agarose gels and visualized with GelRed (3.00% µL/ml) (Biotium Inc., Hayward, USA) and UV light (UVitec Cambridge, USA or UK). PCR products were purified using 2 µl EXOSAP (0.5 µl EXO (Exonuclease I) and 2.5 µl SAP (Shrimp Alkaline Phosphatase) (Takara Bio Inc., Japan) to 5 µl of PCR product and sequencing was performed to confirm the specific amplification of each Y-STR. Sequencing was performed using the BigDye® TerminatorTM v1.1 Cycle Sequencing Kit (AB/LT/TFS: Applied Biosystems™, Life Technologies, ThermoFisher Scientific, Waltham, MA, USA). Sequencing products were purified with BigDye® XTerminatorTM Purification Kit (AB/LT/TFS) and capillary electrophoresis was done on a 3130 Genetic Analyzer (AB/LT/TFS). Sequencing results were analyzed with the Sequencing Analysis software v5.2 (AB/LT/TFS).

Multiplex PCR reaction and capillary electrophoresis

The 12 Y-STR loci under study were typed using multiplex reactions containing 5 µl of QIAGEN Multiplex PCR kit (Qiagen, Valencia, CA), 0.8 µl of primermix (primer concentrations in the mix are described in Supplementary Table S1), 1 ng of genomic DNA and Milli-Q water for a final reaction volumen of 10µl. PCR was performed in a GeneAmp 9800 (AB/LT/TFS) using the same foregoing cycle conditions. PCR products separation and detection were performed by capillary electrophoresis with an ABI3130 Genetic Analyzer (AB/LT/TFS) using the GeneScan 500 LIZ (AB/LT/TFS) as an internal size standard and fragment lengths were assessed with GeneMapper v4.0 (AB/LT/TFS). The analytical threshold was set at 50 RFUs for peak height minimum. Allelic nomenclature follows the recommendations of the International Society for Forensic Genetics (ISFG) (www.isfg.org).

Concordance study

To compare allele designation concordance of the new assay, a set of 50 male DNA samples from the resident population living in the Basque Country previously analyzed with PPY23 in our laboratory, was

examined with the novel multiplex panel. These two assays have in common the 12 Y-STR markers included in the new multiplex.

Sensitivity and stability studies

In order to evaluate the minimum amount of DNA required to obtain a complete Y-STR profile, 2800M DNA control (Promega, Madison, WI) was used for amplification in triplicates in ascending quantities: 25 pg/µl, 50 pg/µl, 100 pg/µl, 200 pg/µl, 400 pg/µl, 1 ng/µl, 1.6 ng/µl, and 10 ng/µl.

Stability studies were conducted by including different concentrations of two common inhibitors, haematin (Sigma-Aldrich Corporation, St. Louis, MO, USA) and humic acid (Sigma-Aldrich Corporation, St. Louis, MO, USA) to the amplification reaction mix containing 1 ng of 2800M DNA. This study was performed in duplicate using ascending concentrations of the PCR inhibitors: 100 µM, 150 µM, 300 µM, 500 µM, 750 µM, 1000 µM, 1500 µM, 3000 µM, and 5000 µM of haematin and 25 ng/µl, 50 ng/µl, 100 ng/µl, 200 ng/µl, 250 ng/µl, 300 ng/µl, 500 ng/µl, 1000 ng/µl, 2000 ng/µl, and 3000 ng/µl of humic acid.

Artificially degraded DNA samples were prepared to evaluate stability of the Y-STR panel. 1 µg of a DNA in-house control sample was digested with 1 µL of DNase Reaction Buffer, 0.5 µL of DNase I and Milli-Q water for a final reaction volume of 10µl (DNase I, RNase-free, ThermoFisher Scientific, Waltham, MA, USA) for 15 min, 30 min, 1h, 2h, 4h and 16h at 37°C. To inactivate DNase I, 1 µL of 50 mM EDTA was added and then incubated at 65°C for 10 min. These samples were amplified in duplicate.

Species specificity

To analyze the species specificity, DNA samples from bull, goat, sheep, pig, cock, rabbit, dog, cat and mouse were tested. All samples were provided by the Bank of DNA of the BIOMICs Group at the University of the Basque Country (UPV/EHU). Of each animal species, 1 ng of DNA was used for each amplification reaction.

Repeatability and reproducibility

The parameter of repeatability was evaluated by analysing whether the peaks of different replicas were always located in the position where the allele of the corresponding STRs should appear.

Reproducibility was checked by performing the entire analysis by three different operators and by running amplification PCRs on three different thermal cyclers: GeneAmp PCR System 9700 Gold (AB/LT/TFS), GeneAmp 9800 PCR System (AB/LT/TFS) and C1000 thermal cycler (Bio-Rad, Hercules, CA, USA). Afterwards, the electropherograms resulting from these diverse analyses were compared to test the reproducibility of the results.

Mixture detection

The capacity of our panel to detect an admixed contribution was tested with a male:male and male:female mixtures. Male:male DNA mixtures were prepared using 2800M DNA control and DNA in-house control sample with the following ratios of 19:1, 9:1, 3:1, 1:1, 1:3, 1:9 and 1:19. To perform the

multiplex amplification reaction the total amount of mixed DNA input was maintained at 1 ng. Male:female DNA mixtures were established using a constant amount of 30 ng K562 DNA control (Promega) with a decreasing template of 2800M DNA control (1 ng, 400 pg, 200 pg, 100 pg, 50 pg and 25 pg). Each mixture analysis was repeated for two times.

Analysis of casework-type samples

To test the effectiveness of this panel with forensic-type samples, two samples from skeletal remains (bones) with degraded DNA and three casework samples from intercomparison exercises organized by the Spanish and Portuguese-Speaking Group (GHEP) of the International Society for Forensic Genetics (ISFG), were analyzed. Samples from skeletal remains (coded in this study as BADN1304 and BADN2049) belong to Spanish Civil War (1936–1939). These samples were selected as examples of forensic samples routinely analysed in our laboratory. The samples from GHEP were items M4 and M8 from exercise 2018 (coded in this study as M4-GHEP18 and M8-GHEP18) and item M4 from exercise 2020 (coded in this study as M4-GHEP20).

Population study and forensic parameters

A sample of 84 unrelated males from Ibiza (Balearic Islands, Spain), previously genotyped with the 12 Y-STR PPY²⁵ was combined with the new Y-STRs in order to assess the forensic parameters of the resulting 23 Y-STR haplotypes.

All donors gave their informed consent prior to inclusion in the study, following the ethical principles and guidelines of the Declaration of Helsinki for the protection of human subjects. Blood samples were collected under the approval of the the Danish local ethical committee (KF-01-037/03) and the Direcció General de R + D + I (Government of Balearic Islands, Spain) (procedure AAEE 12099/2003). DNA was extracted by using QIAamp spin columns (Qiagen, Hilden, Germany) following the manufacturer's recommendations.

Allele frequencies, single-marker genetic diversity (GD), haplotype diversity (HD), as well as different, unique and population specific haplotypes, were assessed using the software Arlequin v3.5.2.2²⁶. For these calculations, the DYS385a/b alleles were treated as haplotypes. The global discrimination capacity (DC) was calculated by dividing the number of different haplotypes by the total number of individuals in the population sample.

Results And Discussion

Primer set and PCR optimization

The 12 Y-STR multiplex panel allows the analysis of the markers DYS448, DYS456, DYS458, DYS635, Y-GATA H4, DYS576, DYS481, DYS549, DYS533, DYS570, DYS643 and DYS385a/b. Primer design was carried out following a miniSTR approach by locating primers as close to the Y-STR repetition units as possible. The final twelve primer pairs generate PCR fragments 105 to 316 bp long (Fig. 1 and

Supplementary Table S1). This design has the advantage of reducing the amplicon size for markers DYS533, DYS549, DYS481, DYS643 and Y-GATA H4 with regard to the PPY23 system²⁷ (Supplementary Figure S1).

Concordance study

A total of 50 samples from the resident population living in the Basque Country were compared with haplotypes obtained in previous analyses performed in our laboratory employing the PPY23 kit. The results showed that the profiles of the same 12 Y-STR markers were completely identical.

Sensitivity and stability studies

Sensitivity was evaluated through the analysis of positive control 2800M at different concentrations. The results indicated that the minimum quantity of DNA necessary to obtain complete genetic profiles, with peak heights above 50 RFUs, in triplicate reactions was 100 pg/μl. Lower amounts experienced allele dropouts of the larger amplicons (Supplementary Table S2 and Supplementary Figures S2 and S4). However, given the results for 50 pg/μl and 25 pg/μl, by increasing certain conditions within the genotyping step, such as the amount of DNA or the injection time, complete profiles might be obtained. Taken together, these results suggest that 12 Y-STR multiplex panel under study could be applied to forensic remains with high limitation in their DNA content.

Stability studies consisted of PCR reactions that included inhibitors such as haematin or humic acid. Complete genetic profiles were obtained with concentrations of ≤ 100 μM of haematin and ≤ 50 ng/μl of humic acid for all replicas (Supplementary Table S2 and Supplementary Figures S3 and S4). Additionally, to determine the efficiency of amplification with degraded samples, artificially degraded DNA was prepared by digestion with DNase I at different times. Complete genetic profiles were obtained with incubation times < 1 h for all replicas. After 1 h incubation, due to the shortage of target DNA, the results start to randomise. For instance, the 1 h incubation time resulted in 83.33% and 50.00% partial profiles, for each replica respectively. Another example occurs after 4 h incubation, where 50.00% and 16.67% of partial profiles were obtained for each replicate respectively (Supplementary Figure S5). All stability test results are indicative of the robustness of the novel panel at moderate DNA degradation levels.

Species specificity

Evidence collected at the crime scene may be exposed to non-human biological sources, such as those of animal origin, and cross-contamination with DNA from other species may occur. In this regard, it is important to assess the species specificity of the assay, ensuring that primers designed to amplify the Y-STRs in this 12 Y-STR multiplex panel are human-specific and do not amplify animal DNA. Here, we have analysed some common livestock and domestic animal species, assuming that these might be the most frequently found at a crime scene.

The results showed the absence of DNA amplification results for all non-human samples tested (Supplementary Figure S6). Therefore, our panel produces no cross-reaction in the tested animal samples, which shows its specificity in terms of species that are common in the human habitat.

Repeatability and reproducibility

Analysis of different replicas proved the repeatability of the panel. No profile experienced a change of allele in any of the replicates, and the position in which the peaks corresponding to each marker appeared did not suffer displacements in the different electropherograms. Since samples are analysed in different capillaries and different runs, the background noise may differ between replicas, resulting in electropherograms with slightly different RFUs.

Moreover, in order to check the reproducibility of the technique when using different thermocyclers, the procedure was performed in a GeneAmp PCR System 9700 (AB/LT/TFS), in a GeneAmp PCR System 9800 (AB/LT/TFS) and in a C1000 Thermal Cycler (Bio-Rad, Hercules, CA, USA). No differences were observed in the results, demonstrating the reproducibility of the technique when different thermocyclers were used.

Additionally, no further differences were detected when the technique was performed by three different operators. In all cases, identical electropherograms were obtained, resulting in complete profiles of 12 Y-STRs.

Mixture detection

Mixtures of multiple male samples or female and male are frequently encountered in forensic crime scenes. Here, the ability of the panel to reveal the presence of mixed DNA samples was tested by two types of mixtures containing male:male and male:female.

For male:male DNA mixtures, as the ratios increased, the proportion of minor alleles that could be identified decreased. The minor component at 3:1 and 1:3 ratios was completely identified. At 9:1 the minor component resulted also on complete profiles. However, the minor component at 1:9 resulted in a 95.83% partial profile on average. At ratios of 19:1 and 1:19, the minor component resulted in 29.17% and 54.17% partial profiles on average, respectively (Fig. 2a, Supplementary Table S3 and Supplementary Figure S7a).

For male:female DNA mixtures, 12 Y-STR multiplex panel was tested in the presence of large amounts of female DNA. The results showed that complete profiles were obtained when the male DNA amount was varied from 1 ng to 400 pg. When the male template DNA amount was reduced to 200 pg, 100 pg, 50 pg and 25 pg the average loci detection rates declined due to allele dropout (Fig. 2b, Supplementary Table S3 and Supplementary Figure S7b). A higher percentage of alleles recovered in 100 pg compared to 200 pg, as well as in 25 pg compared to 50 pg, is due to the fact that primer annealing during PCR is a random process. In cases where there is a small amount of DNA, this randomness can lead to such results. Furthermore, obtaining complete profiles with a higher amount of DNA than in the sensitivity test is explained by the fact that sensitivity might be affected in mixtures. When the target sequence represents only a small fraction of the total DNA, as in this case where 30 ng of female DNA control is present, the

annealing frequency between the primers and the target sites in the first cycles can be reduced, leading to a decrease in sensitivity²⁸.

Identifying the number of contributors in the mixtures and defining the minor and major contributor genotyping, as well as obtaining the Y-STR profile of male contributor in sexual assault cases are some of the challenges of forensic investigation. The 12Y-STR multiplex panel has proven to overcome these difficulties successfully.

AB

Figure 2. Percentages of allele recovery for the minor contributor with ratios varying from 19:1 to 1:19 in male:male mixture study (A). Percentages of allele recovery in male:female mixtures decreasing the male DNA from 1 ng to 25 pg and keeping constant the female DNA to 30 ng (B).

Analysis of casework-type samples

The effectiveness of this panel for analyzing forensic-type samples was tested over two skeletal remains with degraded DNA, as well as over three items from intercomparison exercises organized by the GHEP-ISFG. The analysis of skeletal samples from Spanish Civil War (1936–1939), BADN1304 and BADN2049 resulted on complete profiles. These results were compared with the results obtained with the PPY23, and in both cases the same Y-STR profiles were obtained. The analysis of samples M4-GHEP18 and M8-GHEP18 resulted in complete profiles. These results were checked with the consensus results obtained in the intercomparison exercises of GHEP-ISFG, confirming the Y-STR profile of these samples and reaffirming the reliability of this panel. M4-GHEP20 is a mixture of blood and semen from two males being possible to observe both complete profiles in its analysis. Overall, these results show that the electropherograms obtained with this panel can be informative enough for forensic-type samples, even if they contained degraded DNA.

Population study and forensic parameters

The 12 Y-STR multiplex panel was applied to expand the current data available for the population of Ibiza, previously analyzed with the PPY. Forensic parameters were compared for the 12 loci previously analyzed (PPY) and the resulting 17 Y-STRs (contained in Y-Filer) and 23 Y-STRs (contained in PPY23), both using PPY data extended with the novel 12 Y-STR multiplex panel.

Allelic frequencies and haplotypes

Allele frequencies and gene diversity (GD) of the 23 Y-STRs (12 Y-STRs previously analyzed and the new 11 Y-STRs here genotyped) for Ibiza are summarized in the Supplementary Table S4. Supplementary Table S5 provides the 23-loci haplotypes of the 84 individuals from this population.

Forensic and population genetic parameters

At the single-locus level, DYS456, DYS458 and DYS576 loci were the most discriminative of the eleven incorporated Y-STRs included in the multiplex, generating the highest GD values (0.8181 for DYS576, 0.7533 for DYS458 and 0.7372 for DYS456) (Supplementary Table S4). The lowest GD value in this case was 0.4079 for DYS533. These results are maintained if DYS385a/b loci is included in the computation.

At the haplotype level, the forensic and population genetics parameters of the Ibiza population based on 12-loci contained in PPY (12 loci previously analyzed with PPY), 17-loci contained in Y-Filer (12 loci previously analyzed with PPY and new 5 loci here analyzed with the 12 Y-STR multiplex panel) and 23-loci contained in PPY23 (12 loci previously analyzed with PPY and new 11 loci here analyzed with the 12 Y-STR multiplex panel) haplotypes are compiled in Table 1. As expected the values for number of haplotypes, unique haplotypes, haplotype diversity and discrimination capacity augmented as the number of loci increased. When 23-loci haplotypes were considered, Ibiza exhibit high levels of haplotype diversity, with value of 0.9977, and the discrimination value was 0.9048.

Table 1
Diversity parameters obtained for the population analyzed considering 12, 17 and 23 loci. N = sample size, HD = haplotype diversity, DC = discriminatory capacity.

Ibiza	12 loci (previously analysed with PPY)	17 loci (12 loci previously analysed with PPY and new 5 loci here analysed with the 12 Y-STR multiplex panel)	23 loci (12 loci previously analysed with PPY and new 11 loci here analysed with the 12 Y-STR multiplex panel)
N	84	84	84
Different haplotypes	50	64	76
Unique haplotypes	32	49	68
HD	0.9808 ± 0.0056	0.9923 ± 0.0034	0.9977 ± 0.0022
DC	0.5952	0.7619	0.9048

Conclusion

In the present study, a novel 12 Y-STR multiplex panel has been developed and validated to expand to 23-loci Y-STR haplotypes population samples already analyzed with the 12-Y-STRs (PPY). Indeed, the application of this new panel has allowed to update the current Y-STR database of the Ibiza population previously based on 12 loci genotyped by using the PPY, increasing the resulting haplotype diversity and discriminatory capacity, and consequently providing evidence of the suitability of these Y-STR markers for forensic purposes. Furthermore, the validation studies have also shown that the new panel is valuable for forensic applications as it overcomes the typical difficulties found at a crime scene. In conclusion, the 12 Y-STR panel here described represents an efficient and affordable alternative to update to 23 Y-STRs

numerous populations with only 12 Y-STRs contained in the databases, currently available for population genetics and forensic use.

Declarations

Acknowledgments

Authors acknowledge funding support from Basque Government under predoctoral fellowship PRE_2021_2_0157 to BNL, PRE_2021_2_0252 to EGR and PRE_2019_2_0005 to TK. The authors are grateful to PhD Maite Álvarez for her human and technical assistance provided on the DNA Bank Service (SGIker) of the University of the Basque Country (UPV/EHU, European funding: ERDF and ESF).

Author contributions

MMP conceived the study. AP and JFF contributed to sample collection. BNL, MB, EG, OML and TK carried out the genotyping experiment. BNL, MB and MMP performed data analyses. BNL, MB and MMP wrote the main manuscript text and prepared figures with input from all authors. All authors reviewed the manuscript.

Availability of data and material

All data generated during this study are included in this published article and its supplementary information files. It is also available in the YHRD repository (accession number: YA003484), <https://yhrd.org>

Funding: Funds were provided by the Basque Government (Grupo Consolidado IT-1271-19).

Conflict of interest: The authors declare no competing interests.

Ethic declarations: All donors gave their informed consent prior to inclusion in the study, following the ethical principles and guidelines of the Declaration of Helsinki for the protection of human subjects. Blood samples were collected under the approval of the the Danish local ethical committee (KF-01-037/03) and the Direcció General de R + D + I (Government of Balearic Islands, Spain) (procedure AAEE 12099/2003).

References

1. Vermeulen, M. *et al.* Improving global and regional resolution of male lineage differentiation by simple single-copy Y-chromosomal short tandem repeat polymorphisms. *Forensic Sci. Int. Genet.* **3**, 205–213 (2009).
2. Kayser, M. Forensic use of Y-chromosome DNA: a general overview. *Hum. Genet.* **136**, 621–635 (2017).

3. Navarro-López, B. *et al.* Phylogeographic review of Y chromosome haplogroups in Europe. *International Journal of Legal Medicine* vol. 135 (2021).
4. Batini, C. *et al.* Large-scale recent expansion of European patrilineages shown by population resequencing. *Nat. Commun.* **6**, 7152 (2015).
5. Myres, N. M. *et al.* A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe. *Eur. J. Hum. Genet.* **19**, 95–101 (2011).
6. Jobling, M. A. & Tyler-Smith, C. Human Y-chromosome variation in the genome-sequencing era. *Nat. Rev. Genet.* **18**, 485–497 (2017).
7. Shang, L. *et al.* A novel multiplex of 12 multicity Y-STRs for forensic application. *J. Forensic Sci.* **66**, 1901–1907 (2021).
8. Zhou, Y. *et al.* The validation study of a novel assay with 30 slow and moderate mutation Y-STR markers for criminal investigation and database applications. *Int. J. Legal Med.* **134**, 491–499 (2019).
9. Gusmão, L. *et al.* DNA Commission of the International Society of Forensic Genetics (ISFG): An update of the recommendations on the use of Y-STRs in forensic analysis. *International Journal of Legal Medicine* vol. 120 191–200 (2006).
10. Roewer, L. Y chromosome STR typing in crime casework. *Forensic Science, Medicine, and Pathology* vol. 5 77–84 (2009).
11. Hanson, E. K. & Ballantyne, J. An ultra-high discrimination Y chromosome short tandem repeat multiplex DNA typing system. *PLoS One* **2**, 688 (2007).
12. Baeta, M. *et al.* Digging up the recent Spanish memory: Genetic identification of human remains from mass graves of the Spanish Civil War and posterior dictatorship. *Forensic Sci. Int. Genet.* **19**, 272–279 (2015).
13. Hanson, E. K., Berdos, P. N. & Ballantyne, J. Testing and evaluation of 43 'Noncore' Y chromosome markers for forensic casework applications. *J. Forensic Sci.* **51**, 1298–1314 (2006).
14. Syndercombe Court, D. The Y chromosome and its use in forensic DNA analysis. *Emerg. Top. life Sci.* **5**, 427–441 (2021).
15. Núñez, C. *et al.* 17 to 23: A novel complementary mini Y-STR panel to extend the Y-STR databases from 17 to 23 markers for forensic purposes. *Electrophoresis* **38**, 1016–1021 (2017).
16. Gao, T. *et al.* Phylogenetic analysis and forensic characteristics of 12 populations using 23 Y-STR loci. *Forensic Sci. Int. Genet.* **19**, 130–133 (2015).
17. Purps, J. *et al.* A global analysis of Y-chromosomal haplotype diversity for 23 STR loci. *Forensic Sci. Int. Genet.* **12**, 12–23 (2014).
18. Coble, M. D., Hill, C. R. & Butler, J. M. Haplotype data for 23 Y-chromosome markers in four U.S. population groups. *Forensic Sci. Int. Genet.* **7**, e66–e68 (2013).
19. Nuñez, C. *et al.* Highly discriminatory capacity of the PowerPlex® Y23 System for the study of isolated populations. *Forensic Sci. Int. Genet.* **17**, 104–107 (2015).

20. Tomas, C. *et al.* Results for five sets of forensic genetic markers studied in a Greek population sample. *Forensic Sci. Int. Genet.* **16**, 132–137 (2015).
21. Turrina, S., Caratti, S., Ferrian, M. & De Leo, D. Are rapidly mutating Y-short tandem repeats useful to resolve a lineage? Expanding mutability data on distant male relationships. *Transfusion* **56**, 533–538 (2016).
22. YHRD: Y-Chromosome STR Haplotype Reference Database. <https://yhrd.org/>.
23. *Scientific Working Group on DNA Analysis Methods Validation Guidelines for DNA Analysis Methods SWGDAM Validation Guidelines for DNA Analysis Methods*.
24. Marshall, O. J. PerlPrimer: Cross-platform, graphical primer design for standard, bisulphite and real-time PCR. *Bioinformatics* **20**, 2471–2472 (2004).
25. Rodríguez, V. *et al.* Genetic sub-structure in western Mediterranean populations revealed by 12 Y-chromosome STR loci. *Int. J. Legal Med.* **123**, 137–141 (2009).
26. Excoffier, L. & Lischer, H. E. L. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10**, 564–567 (2010).
27. *PowerPlex ® Y23 System for Use on the Applied Biosystems ® Genetic Analyzers Instructions for Use of Products DC2305 and DC2320*.
28. Prinz, M., Boll, K., Baum, H. & Shaler, B. Multiplexing of Y chromosome specific STRs and performance for mixed samples. *Forensic Sci. Int.* **85**, 209–218 (1997).

Figures

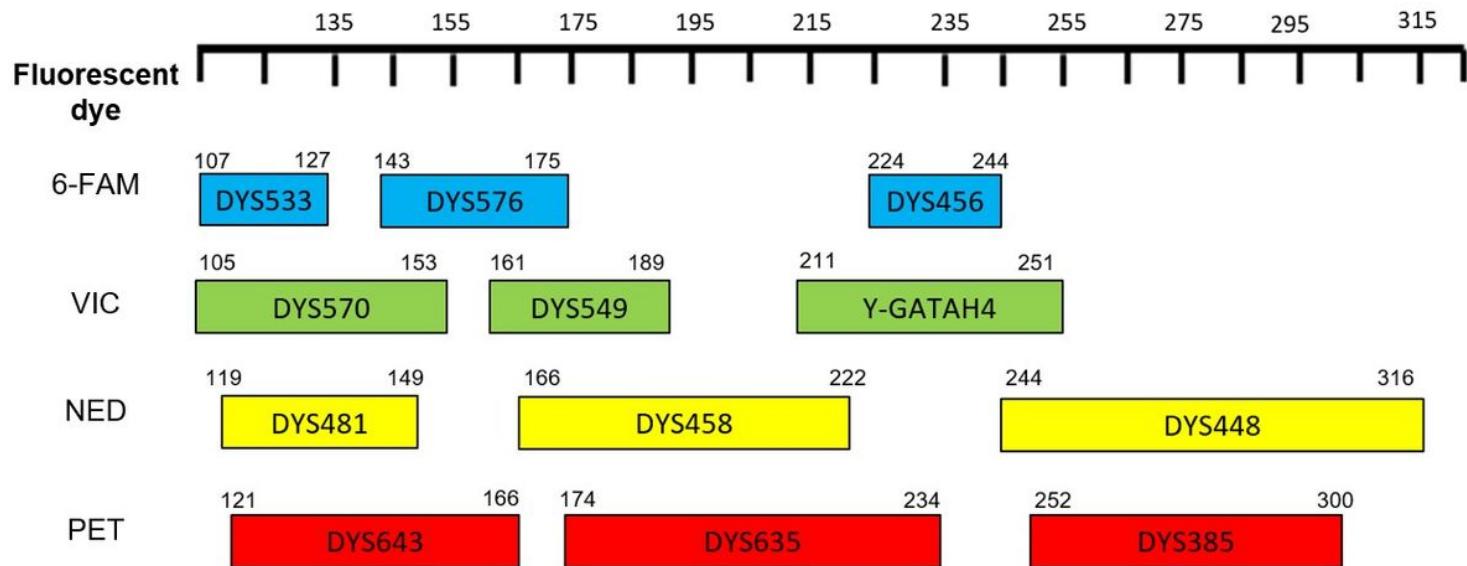


Figure 1

Design of the 12 Y-STR multiplex panel developed in the present study. The boxes represent the expected fragment sizes for each locus and the color of the fluorescent dye with which forward primers were

labeled. The scale at the top represents the size (bp) and the fluorescent dye for the corresponding loci is shown on the left.

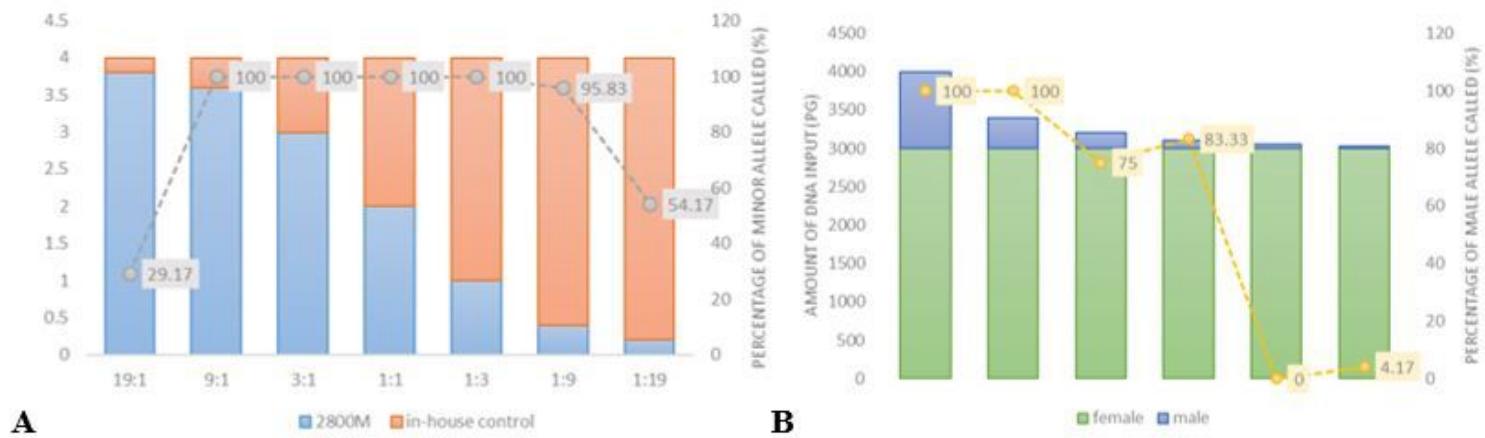


Figure 2

Percentages of allele recovery for the minor contributor with ratios varying from 19:1 to 1:19 in male:male mixture study (A). Percentages of allele recovery in male:female mixtures decreasing the male DNA from 1 ng to 25 pg and keeping constant the female DNA to 30 ng (B).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFigures13YSciRep.docx
- SupplementaryTables13YSciRep.xlsx