X-chromosomal short tandem repeats (X-STRs) are an ideal complement to autosomal markers, particularly in cases of complex kinship analysis. This is especially true in cases where paternity is in question and the child in dispute is a girl. X-STRs can accurately determine sisterhood without needing the father's DNA and identify female DNA in mixed stains [1-3]. Many X-STRs have been evaluated for forensic use, and several X-STRs multiplex systems have been developed that can co-amplify three [4], four [5-7], five [8-9], six [10], seven [11], ten [12], 12 markers [13], 13 markers [14]. Multiplex with more markers has recently been developed to obtain a high degree of discrimination [15]. The approach of multiplex analysis of STR markers has proved worthwhile to minimize labor, genetic material and analysis time for casework in DNA testing laboratories [16]. One of the most critical parameters for the multiplex PCR reaction is the designing of primers. Primer pairs must have similar annealing temperatures to be amplified in a single reaction. Additionally, primers interactions must be examined to avoid the primer-dimer formation that can be reduced by minimizing the excessive regions of complementarily between primers. The stringent initial
primer selection reduces the costly optimization process of multiplex PCR [17]. Here, we describe developing a 14-plex PCR system with 13 X-chromosomal STR markers and sex-determining locus Amelogenin. Alleles for each locus were bi-directionally sequenced to determine the exact repeat size and repeat structure of the X-STRs. The estimated high combined power of discrimination for 13 X-STRs included in the multiplex strongly advocate for its potential to be used for parentage testing and forensic casework studies.

**METHODS**

Sequences of all primer-pairs, GATA172D05, DXS8378, DXS8801, DXS6793, DXS6810, DXS7132, GATA31E08, DXS9902, HPRTB, DXS6789, DXS7423, DXS8377 and DXS981 were obtained from the Genome Database (http://www.gdb.org). Primer sequences for GATA172D05, DXS8378, DXS8801, DXS6793, DXS6810, DXS7132, GATA31E08 and DXS9902 were used without modification. In contrast, primer sequences for HPRTB, DXS6789, DXS7423, DXS8377 and DXS981 were modified to avoid primer dimer formation and have similar melting temperatures for all markers included in the multiplex. Taq polymerase adds an extra Adenosine nucleotide to the 3’ end of each PCR product, and this PCR product is termed “plus A”. In contrast, the PCR product without the additional adenosine is called “minus A” [18]. It has been shown that if the 5’-end of un-labelled primer is a guanosine nucleotide, it favors the plus A form of PCR product and reduces the split peaks that arise for two PCR products differing in length by one base pair [19]. GATA172D05, DXS6810, and DXS9902 sequences were modified by adding guanine to the 5’-end of either forward or reverse primer sequences. The fragment size of all markers labelled with the same fluorescent dye in a multiplex reaction must be sufficiently apart to eliminate overlaps that may lead to ambiguities in assigning the alleles, especially with the discovery of new allele(s). The fragment size for each newly designed primer pair was determined in silico PCR as previously described [20]. A multiplex layout schematic of 13 X-STRs and Amelogenin was prepared (Figure 1). All primer pairs were examined for the possible interaction with Autodimer software to avoid primer dimer formation in multiplex PCR as described [21]. Ten cc blood samples were collected from healthy, unrelated individuals of Pakistani descent and genomic DNA was extracted from white blood cells using the phenol-chloroform method described previously [22]. The genomic DNAs were quantified spectrophotometrically, and the multiplex PCR was performed in a 20 µl volume containing 10 ng of genomic DNA, 75 mM Tris HCl, 20 mM (NH4)2SO4, 2 mM MgCl2, 6.25 µM Spermine Tetra-hydrochloride, 200 µM each of dNTP, and 1 unit of Taq DNA polymerase. Primer concentrations, magnesium concentrations and annealing temperature were optimized for proper amplification. The optimized primer concentrations are shown in Table 1. The amplification program consisted of pre-amplication at 95 °C for 5 min, followed by 28 cycles of denaturation at 94 °C for 1 min, annealing at 57 °C for 1 min and 30 sec, and extension at 65 °C for 2 min. To terminate the “minus A” amplification sequence, a final extension for 60 min at 60 °C was added. DNA samples were genotyped in a 16-capillary ABI PRISM 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA). The loading mixture was prepared by mixing 1 µl of the multiplex PCR product with 15 µl de-ionized Formamide (Applied Biosystems) and 0.2 µl GeneScan 500 LIZ size standard (Applied Biosystems). 9947 A DNA sample from Promega (Promega, Madison, WI) was used as a control [23]. GeneScan Analysis software version 3.7 (Applied Biosystems) analyzed the resolved PCR products. Different alleles of all markers were sequenced on purified fragments using male genomic DNA to determine the allele size and the repeat sequence. Purified products were sequenced using the Big Dye Sequencing Kit from Applied Biosystems according to the manufacturer’s instructions. The amplification process consisted of a pre-denaturation step at 96°C for 1 min, followed by 35 cycles of denaturation at 96°C for 20 sec, annealing at 50°C for 15 sec, extension at 60°C for 4 min and a final extension at 60 °C for 5 min. Sequencing products were separated on an ABI 3100 Genetic Analyzer (Applied Biosystems) and were assembled with Seq Scape software (Applied Biosystems). Sensitivity studies of 14-plex were performed using serial dilutions (20 ng, 10 ng, 5 ng, 2.5 ng, 1.25 ng, 0.625 ng, 0.312 ng) on female DNA samples to evaluate the minimum quantity required to obtain the complete DNA profile. The higher amount of template DNA (40-150 ng) was also tested for co-amplification of markers in the multiplex to observe if there were any non-specific products. The Gene Scan...
Analysis threshold was set at 100 RFU (relative fluorescence units). The per cent stutter was calculated by dividing the peak height at the stutter position by the height of the true allele. We examined 432 DNA samples from five ethnic groups living in Pakistan and calculated the combined power of discrimination of all 13 markers [24]. We constructed an allelic ladder consisting of all variants identified in the Pakistani population. PCR amplification of all markers was performed in a single multiplex reaction, and the resulting peaks of alleles were balanced by mixing the markers in appropriate ratios. The amplification protocol included pre-denaturation at 95°C for 5 min, followed by 35 cycles of denaturing at 94°C for 1 min, annealing at 57°C for 1 min, and extension at 72°C for 1 min. The final extension was performed at 72°C for 30 min. The PCR product was diluted to a ratio of 1/100. PCR conditions were included using 1 ul of the diluted allelic mixture as a template to re-amplify all alleles. The process consisted of pre-denaturation at 96°C for 2 min, followed by 10 cycles of denaturing at 94°C for 1 min, annealing at 57°C for 1 min, and extension at 70°C for 1 min and 30 sec followed by 28 cycles of denaturing at 90°C for 1 min, annealing at 57°C for 1 min, and extension at 70°C for 1 min and 30 sec. The final extension was done at 60°C for 30 min.

RESULTS
A multiplex PCR system was developed for paternity testing and complex kinship analyses with 13 X-STR markers and sex-determining locus amelogenin. The 14-plex PCR system included DXS7132, DXS6789, DXS9902, DXS8377, GATA172D05, DXS8378, DXS6810, DXS7423, GATA31E08, DXS981, DXS6793, HPRTB and DXS6801, and all of them were optimized to be amplified in a single multiplex reaction. These 13 X-STR markers were selected based on their genetic localization, high polymorphism and span of the X-chromosome representing all the 4-linkage groups categorized previously by Szibor R. and his co-workers [3].

Table 1: Primer sequences, dye labels, fragment lengths, primer concentrations and physical characteristics of the 13 X-STR markers and sex-determining marker amelogenin

<table>
<thead>
<tr>
<th>Marker</th>
<th>Primer labelled</th>
<th>Dye Labeling</th>
<th>Primer Sequences</th>
<th>Fragment Length</th>
<th>Multiplex PCR (Primer conc. µM)</th>
<th>Melting temperature (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GATA172D05</td>
<td>F</td>
<td>6-FAM</td>
<td>F-tagtggtgatggtgcacag</td>
<td>108-132</td>
<td>0.105</td>
<td>58.1</td>
</tr>
<tr>
<td>HPRTEB</td>
<td>R</td>
<td>6-FAM</td>
<td>F-gtgcttttcacctgcttcc</td>
<td>158-186</td>
<td>0.1</td>
<td>58.6</td>
</tr>
<tr>
<td>DXS8378</td>
<td>F</td>
<td>6-FAM</td>
<td>F-cacagaggttggcatgtt</td>
<td>191-219</td>
<td>0.11</td>
<td>57.8</td>
</tr>
<tr>
<td>DXS6801</td>
<td>F</td>
<td>VIC</td>
<td>F-agtagcctctctcaagcttcc</td>
<td>118-146</td>
<td>0.067</td>
<td>57.2</td>
</tr>
<tr>
<td>DXS6793</td>
<td>F</td>
<td>VIC</td>
<td>F-acacacgttggtagacacctgt</td>
<td>178-199</td>
<td>0.125</td>
<td>58.7</td>
</tr>
<tr>
<td>DXS6810</td>
<td>F</td>
<td>VIC</td>
<td>F-acagaaaaacctttggacc</td>
<td>209-225</td>
<td>0.115</td>
<td>57.8</td>
</tr>
<tr>
<td>DXS7132</td>
<td>F</td>
<td>VIC</td>
<td>F-agcgcctttcataaaacctgcct</td>
<td>273-301</td>
<td>0.2</td>
<td>56.3</td>
</tr>
<tr>
<td>DXS6789</td>
<td>F</td>
<td>NED</td>
<td>F-ccctcgtcatgtaagttg</td>
<td>121-161</td>
<td>0.06</td>
<td>58.09</td>
</tr>
<tr>
<td>DXS7423</td>
<td>F</td>
<td>NED</td>
<td>F-caacctgcccctttaacctcc</td>
<td>185-201</td>
<td>0.5</td>
<td>57.1</td>
</tr>
<tr>
<td>GATA31E08</td>
<td>F</td>
<td>NED</td>
<td>F-aggaggagagcctcagagta</td>
<td>224-252</td>
<td>0.07</td>
<td>57.9</td>
</tr>
<tr>
<td>DXS9902</td>
<td>F</td>
<td>PET</td>
<td>F-ttagtgctctgggtggaagag</td>
<td>160-176</td>
<td>0.07</td>
<td>57.9</td>
</tr>
<tr>
<td>DXS8377</td>
<td>F</td>
<td>PET</td>
<td>F-acacacactctacgcttcc</td>
<td>225-276</td>
<td>0.15</td>
<td>57.3</td>
</tr>
<tr>
<td>DXS981</td>
<td>F</td>
<td>PET</td>
<td>F-caggtatgtggatcagagta</td>
<td>299-331</td>
<td>0.16</td>
<td>58.8</td>
</tr>
<tr>
<td>Amelogenin</td>
<td>F</td>
<td>PET</td>
<td>F-cctctgggctctgtaaagata</td>
<td>ChrX=119 ChrY=125</td>
<td>0.042</td>
<td>58.3</td>
</tr>
</tbody>
</table>

DOI: https://doi.org/10.54393/pjhs.v4i07.931

Tariq MA and Tahir MA

13 X-STRs Multiplex PCR System

PJHS VOL. 4 Issue. 7 July 2023

Copyright © 2023. PJHS, Published by Crosslinks International Publishers

This work is licensed under a Creative Commons Attribution 4.0 International License.
The markers used in this study were spaced sufficiently apart on the X chromosome to minimize linkage disequilibrium (LD). The analysis revealed no indication of LD between the 13 markers [24]. As a result, all markers investigated in this study have been identified as independent markers suitable for use in forensic practice. An example of a 14-plex assay showing the female DNA profile is in Figure 2.

**Figure 2:** The DNA profile of a female generated using a 14-plex X-STR multiplex system developed during the course of this study

The formation of primer-dimer is a critical parameter, especially in multiplex amplification. All primer pairs were analysed by Auto-dimer software to eliminate primer dimers, and any primer pairs that exhibited an alignment score of 8 or greater were redesigned to obtain optimal primer pairs. Similarly, melting temperatures (Tm) is also another critical parameter for optimal amplification, and 5 X-STR markers HPRTB, DXS6789, DXS7423, DXS8377 and DXS981 were redesigned for Tm values in a small range (56.3-59.6°C). The primer (Tm) values and primer sequences are shown in Table 1. Finally, the primer annealing specificity to the target region is an equally important parameter. All primer pairs were examined with BLAST algorithms. Only those primer pairs were selected that showed specific binding to target regions. Smaller products provide more advantages than their larger counterparts for optimizing the analysis of degraded DNA [25]. Therefore, the total length of the amplified fragments in the multiplex was set at a maximum of 325 base pairs. We examined 432 DNA samples from five ethnic groups of the Pakistan population, and their profiles in multiplex reactions were similar to profiles obtained in PCR with single X-STR markers [24]. All markers’ amplification was balanced by empirically varying the number of primers. DXS6801, DXS6789 and DXS9902 of multiplex PCR system showed preferential amplification due to their smaller product size, as reported previously [26]. Therefore, the concentration of primers of these three markers was reduced to approximately half to obtain balanced peaks in the DNA profile (Table 1). The markers labelled with NED dye (DXS7423, GATA31E08) also resulted in higher amplification and thus, the primers were used in low concentration to obtain a balanced DNA profile, as shown in Table 1. We examined a range of template DNA concentrations in our 14-plex system and obtained optimized results with 1-2 ng template DNA when used in 28 cycles of PCR. Amplification with 1-2 ng template DNA produced >100 RFU threshold, as mentioned in previous studies for the Y-STR ‘megaplex’ system [27], whereas low input of template DNA amounts (< 1 ng) resulted in more variable results and partial profiles. Moreover, markers with considerable fragment lengths (DXS7132, DXS981) started dropping at a low amount of template DNA (< 1 ng), consistent with previous reports [28, 29]. On the other hand, nonspecific products were observed with a higher amount of template DNA (40–150 ng), particularly in the markers labelled with VIC and NED dyes, consistent with previous reports [27].
contains the most frequent alleles for a particular short tandem repeat (STR) found in a human population [30]. It controls the size (migration) variations due to temperature fluctuations during electrophoresis [31]. We constructed an allelic ladder that includes all the common alleles of the 13 X-STRs identified in the Pakistani population, except for some rare alleles, as recommended in the literature [31]. In male subjects, the PCR product was sequenced to determine the exact size of the allele and to repeat the sequence. Ladder alleles were placed between rare alleles of adjacent markers with no overlapping. A minimum difference of 8 base pairs between two adjacent markers was also maintained (Table 1). However, there was an exception of difference between rare alleles of marker HPRTB and DXS8378 (5-bp) in the 6-FAM dye channel, but it was according to the recommendation of IFSH [32]. Therefore, overlapping between the multiplex markers was only observed in some populations even after discovering three new alleles in marker DXS6793 and two in marker DXS981 (Table 2) [24]. The Allelic ladder consisted of 102 alleles for 13 X-STRs alleles, as shown in Figure 3 and listed in Table 2.

**Table 2:** Alleles of all 13 X-STR markers combined to develop an Allelic ladder

<table>
<thead>
<tr>
<th>Locus</th>
<th>Alleles included in allelic ladder</th>
</tr>
</thead>
<tbody>
<tr>
<td>GATA172D05</td>
<td>6, 8, 9, 10, 11, 12</td>
</tr>
<tr>
<td>HPRTB</td>
<td>9, 11, 12, 13, 14, 15, 16</td>
</tr>
<tr>
<td>DXS8378</td>
<td>8, 9, 10, 11, 12, 13, 14</td>
</tr>
<tr>
<td>DXS6801</td>
<td>8, 9, 10, 11, 12, 13, 14, 15</td>
</tr>
<tr>
<td>DXS6793</td>
<td>9, 10, 11, 12, 13, 14, 15, 16</td>
</tr>
<tr>
<td>DXS6910</td>
<td>16, 17, 18, 19, 20</td>
</tr>
<tr>
<td>DXS7132</td>
<td>11, 12, 13, 14, 15, 16, 17, 18</td>
</tr>
<tr>
<td>DXS6789</td>
<td>14, 15, 16, 17, 18, 19, 20, 21, 22, 23</td>
</tr>
<tr>
<td>DXS7423</td>
<td>13, 14, 15, 16, 17</td>
</tr>
<tr>
<td>GATA31E08</td>
<td>6, 7, 8, 9, 10, 11, 12, 13</td>
</tr>
<tr>
<td>DXS9902</td>
<td>9, 10, 11, 12, 13</td>
</tr>
<tr>
<td>DXS8377</td>
<td>39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56</td>
</tr>
<tr>
<td>DXS981</td>
<td>10, 12, 13, 14, 15, 16, 17</td>
</tr>
</tbody>
</table>

Stutter for tetra-nucleotide repeat markers has been reported at approximately <10% of the central peak, and the amount of stutter observed for STR marker was inversely correlated with the length of the core repeat unit [33]. Therefore, the stutter is always higher with dinucleotide repeats but to a lesser extent with tri- and tetra-nucleotide repeat markers [34]. These results were further confirmed in our analyses as the stutter peaks for DXS7132 and DXS981 were negligible (<10%) [13, 35]. In contrast, the stutter products with tri-nucleotide repeat marker DXS8377 were intensive (up to 27%), consistent with previous reports [4, 36]. Another tri repeat marker, DXS6793, also produced up to 13% of stutter product, which was higher than stutter product (up to 10%) for tetra repeat markers (DXS7132 and DXS981). However, typing the samples was easy due to stutter peaks at all the markers, including DXS8377. This parameter should be considered during the mixture analysis because it has been proven essential to distinguish between stutter and true alleles [37].

**D I S C U S S I O N**

Developing the 13 X-STR multiplex PCR system for paternity testing was a remarkable achievement in forensic genetics. Markers were selected based on their genetic localization, high polymorphism, and representation of all 4-linkage groups on the X-chromosome. Successful amplification of these markers in a single multiplex reaction further validated their suitability for paternity
testing and complex kinship analyses. Primer design was essential to maximize amplification and minimize primer dimer formation. Autodimer software was used to analyses all primer and redesigned pairs with an alignment score of 8 or greater. The melting temperature (Tm) was carefully considered to achieve optimal amplification, and some primer pairs were redesigned to have Tm values in the smaller range. The specificity of primer annealing was also evaluated using the BLAST algorithm to ensure specific binding. These considerations increase the efficiency and reliability of multiplex PCR systems [25].

Quantitative optimization of primer pairs and template DNA concentration further improved the performance of the multiplex PCR system. Balancing the amplification of markers was achieved by adjusting primer concentrations for some markers, considering their product size and amplification characteristics. The optimal template DNA concentration was determined to be in the range of 1-2 ng. These improvements ensured consistent and reliable results for paternity testing. Designing and constructing an allelic ladder specific to the Pakistani population was another critical aspect of the multiplex PCR system. The allelic ladder is a size standard that allows accurate allele sizing and translocation variation during electrophoresis. The ladder was carefully constructed to include all common alleles of the 13 X-STR markers identified in the population and some rare alleles [31]. The allelic ladder consisted of 102 alleles, providing a comprehensive reference for allele identification in paternity testing.

Stutter formation, a well-known phenomenon in STR analysis, was also investigated in the context of multiplex PCR systems. The extent of stutter formation was influenced by the length of the repeat unit of the markers, with tetranucleotide repeat markers exhibiting lower levels of stuttering than dinucleotide repeat markers. This excellent 13 X-STR multiplex PCR system offered a reliable and efficient tool for complex paternity testing and kinship analyses. Careful selection of markers, primer design, quantitative optimization, and consideration of perturbation formation contributed to the robustness and accuracy of the system. The International Society of Forensic Haemogenetics (ISFH) Commission guidelines were followed to assign the alleles of X-STR markers [32]. However, there are a few discrepancies in the reported nomenclature. In the case of GATA172D05, the repeat motif was first described as GATA[38] and later as the addition of an extra repeat of TAGA by Edelmann and colleagues [39, 40]. The flanking region of the TAGA repeat motif has a TATA sequence that added one repeat in nomenclature [41]. This research study used the nomenclature Edelmann et al., 2002 concerning GATA172D05. The repeat motif (TCTA) of the coding strand was used by Hearne colleagues [42] for HPRTB. In contrast, Edwards and colleagues used the complementary strand resulting in an extra repeat motif AGAT [38], although it was according to the recommendation of ISFG [32]. Therefore, the AGAT repeat motif from the non-coding strand was used in allele calling in this study, which is very similar to Edwards et al., [38]. A simple repeat motif (AGAT) has been described previously for marker GATA31E08 [43], followed by the marker’s allele designation during this study. However, recently a change in nomenclature at this particular locus has been proposed with the inclusion of “AGGG” in the repeat motif [(AGGG),(AGAT),] [41]. Therefore, this nomenclature adds two extra repeats to the previously described nomenclature [43]. The alleles were designated according to the suggestion of Edelmann et al., for marker DXS7132 and DXS8377 [40], Hering et al., for marker DXS6789 [44], Edelmann et al., for marker DXS9902 [39] and Edelman et al., DXS6801 [45]. The allele designation of DXS8378, DXS6810, DXS7423 and DXS981 was according to Shin et al., [43], whereas the allele designation of DXS6793 was according to Jia et al., [46]. All 13 X-STR primer pairs showed high discriminatory powers ranging from 0.758-0.978 for females and 0.613-0.905 for males in our population[24]. The observed alleles, repeat size and motif structure of 13 X-STR markers are shown in Table 2. We examined DNA profiles generated with genomic DNAs of 432 individuals of Pakistani descent, and the combined power of discrimination of 13 X-STRs was determined to be 2.96 x 10^13 and 2.58 x10^10 in females and males, respectively, which strongly advocates for their utility in complex kinship testing especially in deficiency paternity testing, grand-mother grand-daughter testing and analysis of mixed stains to identify the female DNA. Multiplex PCR-based STR analysis was the most effective technique for forensic DNA analysis. A recent study evaluated the efficacy of indirect multiplex kits for direct amplification of 103 saliva samples without pretreatment using thirteen non-direct multiplex kits, resulting in an 80% reduction in turnaround time. The developed protocol was cost-effective, time-efficient, and did not compromise the quality of the DNA profile. All tested samples produced complete DNA profiles matching the required quality parameters. That was the first report of direct DNA amplification with non-direct multiplex STR kits without pretreatment[47].

**Conclusions**

In conclusion, we have developed and optimized a 13 X-STR system with a combined power of discrimination 1-2.96 x10^13 and 1-2.58 x10^10 in females and males, respectively, due to its significant combined power of discrimination and sensitivity. This system can be efficiently used in parentage testing and forensic casework. To our
knowledge, this was the first study that describes a combination of 13 X-STR markers in a single amplification system.

**Authors Contribution**

Conceptualization: MAT1

Methodology: MAT2

Formal analysis: MAT2

Writing-review and editing: MAT1, MAT2

All authors have read and agreed to the published version of the manuscript.

**Conflicts of Interest**

The authors declare no conflict of interest.

**Source of Funding**

The authors received financial support for the research from CEMB, Indigenous PhD fellowship, HEC, Islamabad.

**Reference**


[38] Edwards A, Civitello A, Hammond HA, Caskey CT. DNA


