Parentage verification in field progeny testing program of Mehsana buffalo

S. J. Jakhesara,* D. N. Rank,† J. D. Kansara,† R. C. Parikh,* V. M. Patel,† P. H. Vataliya,† and J. V. Solanki§

*Dept. of Animal Biotechnology, Veterinary College, Anand Agricultural University, Anand, Gujarat, India-388001; †Dept. of Animal Genetics & Breeding, Veterinary College, Anand Agricultural University, Anand, Gujarat, India-388001; ‡Dudhsagar Research & Development Association, Mehsana, Gujarat, India; and §Ex. Dean & Principal, Veterinary College, Anand Agricultural University, Anand, Gujarat, India-388001

ABSTRACT: The present study was undertaken to construct a multiplex microsatellite panel for parentage testing in Mehsana buffalo (Bubalus bubalis). The study was based on a total of 212 Mehsana buffalos (100 dams, 100 daughters, and 12 sires). Genomic DNA was extracted from blood and semen samples. A panel of 10 microsatellite markers (CSSM61, ILSTS29, ILSTS17, ILSTS28, CSSM57, CSSM22, ILSTS61, CSSM8, ETH152, and ILSTS11) was amplified in a single multiplex reaction and analyzed by capillary electrophoresis on an automated DNA sequencer. The expected heterozygosity ranged from 0.642 to 0.833 (mean 0.762). The total exclusion probability using 10 microsatellite loci with 1 known parent was 0.993. Seven out of 10 microsatellite loci revealed relatively high polymorphic information content (>0.7). Eighty-one daughters out of 100 daughters qualified by compatibility according to Mendelism. The results suggest that multiplex microsatellite panel is a fast, robust, reliable, and economic tool to verify the parentage as well as to assign the putative sire to daughters under progeny testing with very high accuracy and hence can be used in routine parentage testing.

Key words: fragment analysis, Mehsana buffalo, microsatellites, multiplex polymerase chain reaction, parentage verification.

© 2012 American Society of Animal Science. All rights reserved.
doi:10.2527/jas2011-4353

INTRODUCTION

The successful and efficient use of a progeny testing program is a key factor for genetic improvement of dairy herds. Identification and exclusive use of proven sires is ideal to use them extensively for breeding programs; therefore, correct identification of sires in a breeding population is immensely important. However, wrong parentage information is a well-known problem in the estimation of breeding values of sires. Failure to record correct parentage can cause bias in the sire evaluation and introduce errors in estimates of heritabilities and breeding values. Therefore, verification of parentage may serve as a valuable tool for the success of progeny testing programs.

The dairy sector in India is largely dependent on buffalo milk, which contributes 60 to 65% of total milk production. Mehsana buffalo is one of the best milk breeds of buffalo in India (Gupta, 1997). Detailed information regarding breed characteristics has been exhaustively compiled by Oliver (1983). The Dudhsagar Research and Development Association (DURDA), Mehsana, has under taken a program for sire ranking based on progeny testing from 1985. The bulls under the 13th batch of progeny testing were included in the present study. Different techniques based on DNA markers are available for paternity testing including RFLP (Kashi et al., 1990), multilocus, minisatellite and oligosynthetic probes (Trommelen et al., 1993), and PCR-based minisatellites and microsatellites (Schnabel et al., 2000). The RFLP generally suffer from low heterozygosities and low polymorphic information content (PIC), making the DNA fingerprints difficult to interpret owing to the complex nature of the banding pattern, whereas microsatellite markers, being ubiquitous in nature, show high heterozygosity, follow Mendelian codominant inheritance, and thus have ease of scoring. These features make them markers of choice for

1The study was funded by grant from Dudhsagar Research and Development Association (DURDA), Mehsana, India. We thank D. K. Sharma, Director, Centre for Analysis and Learning in Livestock and Food (CALF), NDBD, Anand, Gujarat, India, for valuable suggestions and 2 anonymous reviewers for their comments on earlier version of this paper.

2Corresponding author: drsubhash82@gmail.com
Received June 9, 2011.
Accepted July 19, 2012.
paternity verification and individual identification. The present study was aimed at developing a panel of microsatellite markers and testing their potential for parentage verification of Mehsana buffaloes.

**MATERIALS AND METHODS**

Animal Care and Use Committee approval was not obtained for this study because no animals were used.

Experimental material for the present study comprised a total of 207 blood and 5 semen samples from 100 dams, 100 daughters, and 12 sires of Mehsana buffalo covered under the 13th batch of field progeny in the testing program operated by DURDA. Stored frozen semen samples were collected from 5 buffalo bulls used for AI under the program but not available at the time of blood sample collection. These blood and frozen semen samples were used for DNA extraction as per the method described by John et al. (1991) and Aravindakshan et al. (1998), respectively.

A panel of 10 microsatellite markers was developed for present study (Table 1). Primers with more than 4 reported alleles with good heterozygosity and high PIC were preferred. Microsatellite primers were synthesized from MWG Biotech (Ebersberg, Germany). The 5’-end of the forward primer was labeled with one of the 4 fluorescent dyes: Carboxyfluorescein (FAM), Carboxyhexachlorofluorescein (HEX), Carboxytetramethylrhodamine (TAMRA), or Carboxyrhodamine (ROX).

**Polymerase Chain Reaction Optimization**

The present study was aimed to develop a panel containing maximum number of microsatellite markers that can be amplified in a single PCR reaction. A total of 22 microsatellite markers were tested for desired PCR amplification. The primers giving nonspecific amplification and multiple bands were excluded. The remaining primers were amplified at different annealing temperatures, for example, 52, 54, 56, and 58°C, with the strategy to find the temperature at which maximum numbers of primers gave desired and specific amplification. Fifteen primers giving amplification at 58°C were selected and tested for compatibility in co-amplification by multiplex PCR and analysis by capillary electrophoresis on an ABI PRISM 310 automated DNA sequencer (Applied Biosystems, Foster City, CA).

It is necessary that routine parentage analysis be carried out by single injection of PCR amplicons in a capillary electrophoresis. Currently, multiple fluorescent dyes can be used with a commercial automatic DNA sequencer. This enables simultaneous analysis of the overlapping microsatellites within a similar size range. The analysis of microsatellites with the automated DNA sequencer generally shows a good resolution for PCR products ranging from 100 to 300 bp, resulting in a setup

<p>| <strong>Table 1. Microsatellite markers used to verify parentage in Mehsana buffalo</strong> |</p>
<table>
<thead>
<tr>
<th>LOCUS name</th>
<th>Chromosomal location</th>
<th>5’ labeling</th>
<th>Primer sequences (5’ to 3’)</th>
<th>Repeats</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSSM61</td>
<td>(10: 55,424,742 - 55,424,921)</td>
<td>FAM</td>
<td>F-AGGCCATATAAGGAGCCAGGCTTAC &lt;br&gt; R-TTCAGAAGAGGGCACGAATACAC</td>
<td>(GT)20</td>
<td>Barker et al., 1997</td>
</tr>
<tr>
<td>ILSTS17</td>
<td>(X: 45,444,764 - 45,444,889)</td>
<td>HEX</td>
<td>F-GTCCCCTAAAAATCGAATGCC &lt;br&gt; R-GCTATCTCTTAAACCCTGTTCC</td>
<td>(GT)17</td>
<td>Kemp et al., 1995</td>
</tr>
<tr>
<td>CSSM57</td>
<td>(7: 52,951,320 - 52,951,430)</td>
<td>TAMRA</td>
<td>F-GTCGCCGTTGATAAACAATTTAAGT &lt;br&gt; R-TGTTGGTTGTTTAAACCTCTCTATCCT</td>
<td>(GT)16</td>
<td>Moore and Vankan, 1994</td>
</tr>
<tr>
<td>ILSTS28</td>
<td>(11: 96,769,846 - 96,769,976)</td>
<td>HEX</td>
<td>F-TCCAGATTTTGTACACGACC &lt;br&gt; R-GTCATGTATACCTTCTGAGC</td>
<td>–</td>
<td>Kemp et al., 1995</td>
</tr>
<tr>
<td>ILSTS61</td>
<td>(15: 58,787,060 - 58,787,211)</td>
<td>ROX</td>
<td>F-AAATTTAGGGGCTATACGG &lt;br&gt; R-TGGCCCTACCCCTACCTGTTCC</td>
<td>(CA)14</td>
<td>Kemp et al., 1995</td>
</tr>
<tr>
<td>CSSM22</td>
<td>(5: 66,423,476 - 66,423,699)</td>
<td>TAMRA</td>
<td>F-TCTCTCTAATGGAGGTTGTTTTT &lt;br&gt; R-ATATCCCCACTAGGATAGAAATTC</td>
<td>(CA)19</td>
<td>Moore and Vankan, 1994</td>
</tr>
<tr>
<td>ILSTS29</td>
<td>(3: 51,415,436 - 51,415,595)</td>
<td>FAM</td>
<td>F-CTGTTGTTGATGGACACAGCC &lt;br&gt; R-TGAGATTTGACCGATAGTTG</td>
<td>(CA)18</td>
<td>Moore and Vankan, 1995</td>
</tr>
<tr>
<td>ETH152</td>
<td>(5: 114,885,285 - 114,885,451)</td>
<td>FAM</td>
<td>F-TATCTGCTAGCGGCTGTCG &lt;br&gt; R-GAGACCTCAAGGTTGCGATCAG</td>
<td>(CA)17</td>
<td>Kemp et al., 1995</td>
</tr>
<tr>
<td>ILSTS11</td>
<td>(14: 11,778,542 - 11,778,810)</td>
<td>FAM</td>
<td>F-GCTTGCTACATGGAGAACGTG &lt;br&gt; R-CATAAATGCGAGGCGTACC</td>
<td>(TC)2(CA)11</td>
<td>van Hooft et al., 1999</td>
</tr>
</tbody>
</table>

1Chromosomal location in cattle is given in parenthesis as chromosome: base pair on the basis of *Bos taurus* (cattle) reference genome build 5.2. Location in buffalo was determined by conversion of cattle chromosomal location on the basis of Amaral et al. (2008).

25’ labeling = Fluorescent dye labels are most commonly added to the 5’-position of an oligonucleotide

3F = Forward primer, R = Reverse primer
Parentage testing in Indian buffalo

The primers giving multiple amplicons or producing amplicons with overlapping size range with other primers of the same dye label were excluded. Finally, 10 primers were selected after testing them thrice on 10 samples for repeatability and reproducibility of results. Reasons for excluding 12 microsatellites are summarized in Table 2.

Table 2. Reasons for discarding 12 microsatellite markers found unsuitable for multiplex designing

<table>
<thead>
<tr>
<th>Serial number</th>
<th>Microsatellite</th>
<th>Reason for exclusion from primer panel</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ILSTS52</td>
<td>Not amplified at 58°C</td>
</tr>
<tr>
<td>2</td>
<td>ILSTS30</td>
<td>Not amplified at 58°C</td>
</tr>
<tr>
<td>3</td>
<td>ILSTS58</td>
<td>Not amplified at 58°C</td>
</tr>
<tr>
<td>4</td>
<td>ILSTS56</td>
<td>Multiple peaks at 58°C</td>
</tr>
<tr>
<td>5</td>
<td>ETH3</td>
<td>Multiple peaks at 58°C</td>
</tr>
<tr>
<td>6</td>
<td>ILSTS33</td>
<td>Range overlap with ILSTS29</td>
</tr>
<tr>
<td>7</td>
<td>CSSM32</td>
<td>Failed to amplify in multiplex at 58°C</td>
</tr>
<tr>
<td>8</td>
<td>CSSM45</td>
<td>Failed to amplify in multiplex at 58°C</td>
</tr>
<tr>
<td>9</td>
<td>INRA63</td>
<td>Failed to amplify in multiplex at 58°C</td>
</tr>
<tr>
<td>10</td>
<td>BM1818</td>
<td>Failed to amplify in multiplex at 58°C</td>
</tr>
<tr>
<td>11</td>
<td>ILSTS22</td>
<td>Failed to amplify in multiplex at 58°C</td>
</tr>
<tr>
<td>12</td>
<td>CSSM43</td>
<td>Null allele</td>
</tr>
</tbody>
</table>

Primers excluded at time of multiplex primer panel testing and optimizing

Statistical Analysis

The data for all the samples were analyzed for genetic variability, allele counts, frequencies, and observed and expected heterozygosities. Dam–daughter–sire trio was verified with respect to Mendelian inheritance. Every daughter should show Mendelian inheritance of microsatellite alleles (i.e., every allele of a daughter should be traced back to either dam or sire). Parentage was verified manually by matching alleles at 10 microsatellite loci in the daughter with that of dam and sire. After manual parentage verification, parentage was verified using Cervus software version 3.0 (Field Genetics Ltd, London, UK). Exclusion probability was calculated using the formula originally given by Chakravarti and Li et al. (1998). Combined EP takes account of individual EP and was calculated by following formula:

$$P = 1 - \prod_{i=1}^{n} (1 - P_i)$$

in which \( n \) = number of markers, \( P_i \) = individual exclusion probability for single marker, \( \Pi \) = symbol of product, and \( P \) = combined exclusion probability.

RESULTS

After amplification and capillary electrophoresis, all selected markers in the multiplex panel gave nonambiguous peaks for parentage assignment with confidence. A representative figure showing multiplex microsatellite data is shown as Figure 1.

Heterozygosities and the Number of Alleles

The heterozygosities and EP value for different markers are presented in Table 3. The number of alleles varied from 5 for marker ILSTS11 to 16 for marker ILSTS61. Allele frequencies observed in the present study at each microsatellite locus is also included in Supplementary Table 1 (see online version of the article at http://journalofanimalscience.org). The observed heterozygosity ranged between 0.58 for marker ILSTS11 and 0.82 for marker ILSTS61. Calculated PIC values ranged from 0.61 (least) for marker CSSM57 to 0.81 (most).
(greatest) for marker CSSM61. Exclusion probability value was greatest for marker CSSM61 (0.49) and least for marker CSSM57 (0.25). Combined EP for the selected 10 markers was 0.993 indicating parentage assignments with 99% of confidence.

In addition, EP were calculated using the most polymorphic marker, then subsequently including the 2 most polymorphic markers, and then including 3 and 4 and so on up to 10 polymorphic markers (Table 4). It is evident from Table 4 that the 5 most polymorphic markers were sufficient to assign parentage with 95% of exclusionary power. The inclusion of additional markers resulted in very small gains.

**Parentage Verification**

**Parentage Analysis with Cervus 3.0.** Parentage verification was done using software Cervus version 3.0 (Field Genetics, Ltd.) by analyzing each candidate father (12 in number) as the potential father for the daughter in question. Simulation of parentage analysis was done for likelihoods using the corrected likelihood equations of Kalinowski et al. (2007). The corrected likelihood equations of Kalinowski et al. (2007) are superior because they have greater power to assign parentage and more accurately take account of typing errors. The logarithm of odds (LOD) [logarithm (base 10) of odds] score was calculated for each candidate parent and parentage was assigned to the sire having the maximum LOD score for 95 daughters with strict confidence, leaving 5 parentages unassigned. The software detected 19 daughters with wrong parentage and assigned the most probable correct sire for each of the daughter.

**Manual Parentage Verification.** After allele scoring and genotyping of each individual, the results obtained with Cervus (Field Genetics, Ltd.) were confirmed manually by matching samples of each daughter with her respective dam and sire. A mismatch at more than 1 locus was interpreted as wrong parentage. A confirmed match at all loci was considered as true parentage. Parentage was verified for each daughter likewise by comparing daughter–dam–sire trio and marked as correct or incorrect. We found 19 daughters having wrong parentage out of 100 daughters using the above procedure. Each result was also checked manually using an Excel document (Microsoft, Redmond, WA). To verify if the wrong parentages observed were due to genotyping error, samples from each daughter–dam–sire trio showing wrong parentages were reanalyzed by performing monoplex PCR and results were found to be in accordance with multiplex PCR.

**DISCUSSION**

Our objective of this study was to develop and test a suitable multiplex panel consisting of a battery of microsatellites for parentage verification in Mehsana buffalo. Microsatellites suitable for this purpose are those showing multiple alleles as well as high heterozygosity. It is important to note that no microsatellite markers were available for parentage verification in Indian buffalo. Due to nonavailability of confirmed markers, there is difficulty in the selection and setup of a microsatellite panel useful for the purpose. We selected 22 microsatellites with more than 5 reported alleles, high heterozygosity, and high PIC from the list of 25 markers recommended by the National Bureau of Animal Genetic Resources for genetic characterization of buffalo and 25 cattle specific microsatellites suggested by the Interna-
Table 4. Probability of exclusion including successively most polymorphic markers

<table>
<thead>
<tr>
<th>No. of markers</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marker name</td>
<td>CSSM61</td>
<td>ILSTS61</td>
<td>ETH152</td>
<td>CSSM8</td>
<td>ILSTS28</td>
<td>CSSM22</td>
<td>ILSTS17</td>
<td>ILSTS11</td>
<td>CSSM57</td>
<td>ILSTS29</td>
</tr>
<tr>
<td>Exclusion probability</td>
<td>0.496</td>
<td>0.743</td>
<td>0.862</td>
<td>0.921</td>
<td>0.954</td>
<td>0.971</td>
<td>0.982</td>
<td>0.986</td>
<td>0.99</td>
<td>0.993</td>
</tr>
</tbody>
</table>

Accurate estimation of genetic parameters, for a population and an individual, are dependent on correct recording of genetic relationships among individuals in the population. Misidentification reduces genetic gain with sire models (Geldermann et al., 1986) and may have an even greater effect with animal models that account for all assumed genetic relationships (Wiggans et al., 1988). Often wrong sire information may compromise the results of the progeny testing program and therefore, correct identification of sires is immensely important.

The results of this study show a relatively high pedigree error rate of 19%. Under Indian field conditions such a high error rate can be attributed to many reasons, such as 1) errors in data recording, 2) poor management practices, 3) error in identification of semen straws by AI technicians, 4) buffaloes fertilized by previous insemination, 5) errors in labeling semen straws, 6) the use of natural-service buffalo bulls leading to pregnancies of previously inseminated buffaloes, and 7) mistaken interchange of buffalo calves at birth. Similar reasons have also been reported in earlier studies (Christensen et al., 1982). Some of these errors can be prevented by alert technicians, awareness of the owner about the program, and good AI recording and verification systems. For example, if insemination records and milk recording (including date of calving) are linked and a single buffalo identification is assigned throughout her lifetime, causes 4 and 6 could be avoided. Cause 5 could be minimized by strict quality control measures at the AI stations.

Christensen et al. (1982) reported misidentification rates between 5 and 15% in Danish dairy cattle, Geldermann et al. (1986) found a misidentification rate of 13% in German dairy cattle, and Ron et al. (1996) found a sire error rate of 5% in Israeli Holstein cows, based upon genotyping 12 microsatellite markers on 173 cows and their putative 4 sires. However, exercising exclusion based on at least 2 mismatched markers, the error rate was reduced to 2%. In the present study exclusion was practiced with at least a single mismatched marker out of 10. When we relaxed exclusion based on at least 2 and later with 3 markers, a 1 and 7% decrease in error rate was noticed, respectively.

Estimation of accurate genotyping error detection is important for accuracy of paternity testing. Various approaches have been discussed in the literature for detection of genotyping error, namely 1) inconsistencies in Mendelian inheritance in parent–offspring pairs, 2) comparison of high-quality reference samples with error-prone genotypes, and 3) counting mismatches between duplicated genotypes. The last approach has been more recommended because it is feasible, but it also has some limitations (Pompanon et al., 2005; Dewoody et al., 2006). It is appropriate to perform the first approach whenever possible as it will give a nearly correct representation of genotyping error. Considering the first approach in our dataset of 100 dam–daughter pairs, no instance of Mendelian inconsistency was observed. Therefore, in the present study genotyping error is likely to be 0. We have also used Cervus 3.0 program (Field Genetics Ltd.), which is reported to accommodate genotyping error during paternity assignment (Kalinowski et al., 2007). In addition, a PCR was performed with each single marker for each wrong parentage assigned and cross-verified on the capillary sequencer. Hence, it is very likely that all the paternity errors observed in the study are due to pedigree errors and not due to genotyping error.

It would be relevant to note that a set of minimum guidelines and recommendations for use of animal DNA in forensic and identity testing has been developed to assure quality service from service providers (Budowle et al., 2005; Linacre et al., 2011). Basic guidelines for standard laboratory practice would be the same but important guidelines pertaining to paternity testing briefly are 1) primers used in the study should have references showing map positions, 2) preferentially, tetra nucleotide repeat markers should be used except widely used dinucleotide repeat markers in animals, 3) the primers used in the study should be reproducible in results, 4) the number of repeats should be the basis for result reporting, 5) reported alleles should be sequenced and an allelic ladder should be prepared by sequencing, and 6) population and forensic genetic parameters including allele frequencies should be estimated. Accordingly, guidelines 2, 3, and 6 were fulfilled in our study. Furthermore, for cross-laboratory comparison of results, PCR products of homozygous animals should be sequenced, which is the only way to establish true allele lengths. Under Indian conditions, meeting the remaining guidelines may be somewhat difficult because the Indian buffalo is genetically less studied and validated markers for individual characterization are scarce. In the absence of specific markers for buffalo, the markers from cattle are used frequently for characterization studies. Nonetheless these guidelines are very important for high quality genetic practices and it is recommended to use these in Indian conditions also.

Declaring nonpaternity when a single marker shows...
exclusion would increase the pedigree error detection rate but would also increase the false positive rate. These problems could easily be resolved if more markers were used in the parentage test. If 20 to 30 markers were used, there should be negligible ambiguity in the outcome, even in the presence of genotyping errors. However, it should be noted that this would increase the time and cost involved in the study.

In conclusion, a robust, fast, and reliable multiplex primer panel containing 10 primers has been developed successfully for parentage verification in Mehsana buffalo. The combined EP of all 10 markers was 0.993, ensuring parentage assignments with a 99% confidence level. These microsatellites are highly polymorphic and have proved very useful for parentage verification in the Mehsana buffalo population. With primer multiplexing and fragment analysis on a genetic analyzer, throughput is increased to a larger extent, saving time as well as cost.

LITERATURE CITED


Oliver, A. 1983. A brief survey of some of the important breeds of cattle in India. Pages 17–45 in Miscellaneous Bulletin, Council of Agricultural Research, New Delhi, India.


