Identification of duplicates in mulberry germplasm using morphological descriptors and SSR markers

Original research article

ABSTRACT

Germplasm collections invariably contain duplicate accessions, both within and between genebanks. These redundancies are a burden for curators because they do not contribute to the diversity in the collection, but do require genebank budget for maintenance. Thus, both from a genetic and economic point of view, identification and elimination of redundancies should be an important genebank objective. In field gene banks, duplicate accessions are widespread and their identification is required to facilitate germplasm administration and lower the maintenance costs. The goal of this study was to identify the duplicates in mulberry germplasm that has been morphologically and agronomically evaluated. Until now the identification of duplicate accessions had to rely on comparison of morphological characters which are subject to environmental variation. However, it is possible to make routine use of molecular markers based on genomic DNA for the identification of duplicate accessions of vegetatively propagated species. In the present study, among the mulberry germplasm 312 accessions were characterized using 17 qualitative morphological descriptors and identifiied 84 suspected duplicates using multivariate cluster analysis. Further, the suspected duplicates were screened using 12 SSR markers and scored the polymorphic alleles using binary format. The data matrix was subjected to multivariate cluster analysis and 14 mulberry accessions were confirmed as true duplicates. The identification of redundant accessions enables curators to concentrate their efforts on the characterization, evaluation, and regeneration of distinct genetic material. The molecular markers aids in giving priorities for the evaluation and regeneration of unique genetic materials in field genebank.

Key words: Mulberry germplasm, Field gene bank, Morphological descriptors, Duplicate accessions, SSR markers, PAGE.

1. INTRODUCTION

The genus *Morus* L. is one of the most interesting taxonomic groups due to its genetic variability and commercial importance in the Sericulture industry. *Bombyx mori* L. which produces natural silk feeds only on its leaves.Mulberry grows in a wide range of environments around the world, from tropical to sub-arctic, resulting in a wide range of genetic resources (Sharma et al. 2000). The only way to keep the cultivars' uniqueness is through vegetative propagation. Long history of cultivation together with traditional breeding has resulted in a variety of types in mulberry at various places.

The Central Sericultural Germplasm Resources Centre (CSGRC) in India houses the world's biggest mulberry germplasm collection, with 1317 mulberry germplasm accessions maintained in the form of *ex situ* field gene bank. Larger phenotype-based selection and collection of a greater number of local accessions increased duplication in collections, making germplasm conservation more challenging (Das et al. 2020). Management and maintenance of such a huge germplasm collection is a difficult and costly task. To achieve long-term conservation, management practices must focus on safe multiplication and periodic rejuvenation of conserved germplasm to avoid genetic attrition during these operations. Unintentional sample duplication is one of the most common difficulties in gene banks.Various local names for the same cultivar could be the cause, as different people collecting the same plant material from different places (the same cultivar may be phenotypically very different) including sampling and maintenance errors (Sisko 2016). As a result, determining redundancy at the

molecular level within the germplasm prior to conservation in an *ex situ*centre has become critical. Naik and Dandin (2006) have identified duplicates in mulberry germplasm using both morphological characters and RAPD markers. They demonstrated the diversity pattern of mulberry collections ascertained by RAPD markers closely resembled that of morphological marker analysis. The study utilized nine morphological traits that are qualitative in nature and hence least influenced by environmental effects. The quantitative traits are vastly influenced by climate, agronomical inputs and soil conditions.

Stable markers are expected to improve the characterization of mulberry genetic resources. It is necessary for their effective management and efficient utilization. Moreover, available mulberry germplasm species exhibits much phenotypic diversity. Molecular markers are useful complements to morphological characters because they are plentiful, independent of tissue or environmental effects and allow accession identification in the early stages of development. Such techniques reveal polymorphisms at the DNA level and are very powerful tool for characterization and diversity estimation (Shabir et al., 2010). Many molecular marker techniques have been successfully used in identification and genetic diversity analysis in mulberry. The present study was undertaken in the above direction for exploring the possibility of utilization of SSR markers to identify the duplicates. The SSR markers were collected from CSRTI, Mysore developed under the project PIC01003CN. They have considered only 14 mulberry accessions for the screening using RAPD markers.In this study it was attempted to identify the possible duplicates among 312 mulberry germplasm. Hence, the present investigation was undertaken to identify the duplicates among 312 accessions using SSR markers.

2. MATERIAL AND METHODS

The mulberry accessions used in the present study were collected from ex situ field gene bank of CSGRC, Hosur. A total of 312 mulberry accessions belonging to the species *Morus indica* were selected for the study (Table 1). The fresh, young leaf samples were collected from mulberry field gene bank of CSGRC, Hosur. The leaves were surface sterilized with 70% ethanol and stored at -80°C until further use.

Table 1: List of *Morus indica* accessions

SN	Acc. No.	IC numbers
1	MI-0009	IC313677
2	MI-0010	IC313969
3	MI-0012	IC313971
4	MI-0018	IC313972
5	MI-0020	IC313683
6	MI-0021	IC313956
7	MI-0022	IC313913
8	MI-0035	IC313690
9	MI-0036	IC313980
10	MI-0040	IC313983
11	MI-0045	IC313988
12	MI-0047	IC313693
13	MI-0049	IC313693
14	MI-0051	IC313696
15	MI-0057	IC313702
16	MI-0058	IC313703
17	MI-0059	IC313704
18	MI-0062	IC313707
19	MI-0063	IC313990
20	MI-0067	IC313710
21	MI-0069	IC313712
22	MI-0070	IC313713
23	MI-0072	IC313715
24	MI-0073	IC313716
25	MI-0074	IC313717
26	MI-0078	IC313720
27	MI-0081	IC313722
28	MI-0084	IC313815
29	MI-0085	IC313816
30	MI-0086	IC313725
31	MI-0087	IC313726
32	MI-0088	IC313817
33	MI-0089	IC313818
34	MI-0095	IC313823
35	MI-0096	IC313824
36	MI-0101	IC313829
37	MI-0103	IC313728
38	MI-0110	IC313735
39	MI-0113	IC313738

40	MI-0114	IC313739
41	MI-0115	IC313740
42	MI-0116	IC313741
43	MI-0117	IC313831
44	MI-0121	IC313745
45	MI-0123	IC313991
46	MI-0125	IC313748
47	MI-0127	IC313750
48	MI-0130	IC313753
49	MI-0131	IC313754
50	MI-0134	IC313757
51	MI-0136	IC313759
52	MI-0138	IC313957
53	MI-0142	IC313764
54	MI-0144	IC313766
55	MI-0147	IC313769
56	MI-0148	IC313770
57	MI-0150	IC313772
58	MI-0151	IC313773
59	MI-0152	IC313774
60	MI-0154	IC313775
61	MI-0161	IC313780
62	MI-0163	IC313832
63	MI-0170	IC313784
64	MI-0171	IC313834
65	MI-0174	IC313785
66	MI-0175	IC313786
67	MI-0176	IC313787
68	MI-0177	IC313788
69	MI-0179	IC313837
70	MI-0180	IC313838
71	MI-0181	IC313839
72	MI-0183	IC313841
73	MI-0185	IC313843
74	MI-0187	IC313992
75	MI-0192	IC313994
76	MI-0193	IC313995
77	MI-0197	IC313851
78	MI-0198	IC313852
79	MI-0200	IC313854
80	MI-0201	IC313855
81	MI-0202	IC313856
82	MI-0203	IC313857
83	MI-0204	IC313858
84	MI-0205	IC313859
85	MI-0210	IC314109
86	MI-0213	IC313863
87	MI-0216	IC313866
88	MI-0217	IC313867
89	MI-0218	IC313868
90	MI-0221	IC313871
91	MI-0223	IC313873
92	MI-0225	IC313875
93	MI-0227	IC313877
94	MI-0232	IC313882
95	MI-0235	IC313885
96	MI-0236	IC313886
97	MI-0237	IC313887
98	MI-0238	IC313888
99	MI-0239	IC313889
100	MI-0240	IC313890
101	MI-0241	IC313891
102	MI-0243	IC313893
103	MI-0245	IC313895
104	MI-0248	IC313792
105	MI-0260	IC313804
106	MI-0264	IC313919
107	MI-0265	IC313920
108	MI-0271	IC313926
109	MI-0272	IC313927
110	MI-0273	IC313928
111	MI-0274	IC313929
112	MI-0278	IC313933
113	MI-0279	IC313934
114	MI-0280	IC313935
115	MI-0281	IC313662
116	MI-0282	IC313663
117	MI-0284	IC313665
118	MI-0287	IC313668
119	MI-0289	IC313670
120	MI-0297	IC314006
121	MI-0309	IC314021
122	MI-0311	IC314022
123	MI-0312	IC314023
124	MI-0316	IC314216
125	MI-0318	IC314152
126	MI-0319	IC314153
127	MI-0320	IC314154

128	MI-0321	IC313936
129	MI-0324	IC313939
130	MI-0325	IC313940
131	MI-0326	IC313941
132	MI-0327	IC314016
133	MI-0328	IC314232
134	MI-0329	IC314025
135	MI-0332	IC314028
136	MI-0333	IC314029
137	MI-0334	IC314030
138	MI-0335	IC314031
139	MI-0336	IC314032
140	MI-0338	IC314034
141	MI-0339	IC314035
142	MI-0342	IC254481
143	MI-0344	IC314114
144	MI-0345	IC314115
145	MI-0346	IC314116
146	MI-0349	IC314119
147	MI-0355	IC314124
148	MI-0356	IC314125
149	MI-0358	IC314127
150	MI-0359	IC314128
151	MI-0361	IC314130
152	MI-0367	IC314157
153	MI-0368	IC314158
154	MI-0370	IC314160
155	MI-0372	IC314162
156	MI-0373	IC314163
157	MI-0374	IC314164
158	MI-0377	IC314131
159	MI-0385	IC314039
160	MI-0386	IC314040
161	MI-0395	IC314140
162	MI-0398	IC313945
163	MI-0399	IC313946
164	MI-0401	IC314141
165	MI-0402	IC314142
166	MI-0403	IC314143
167	MI-0405	IC314172
168	MI-0406	IC314173
169	MI-0407	IC314174
170	MI-0413	IC314045
171	MI-0415	IC314046

172	MI-0422	IC314224
173	MI-0425	IC314179
174	MI-0432	IC314048
175	MI-0433	IC314049
176	MI-0434	IC314050
177	MI-0437	IC314185
178	MI-0441	IC314189
179	MI-0443	IC313673
180	MI-0444	IC314017
181	MI-0445	IC313674
182	MI-0446	IC314051
183	MI-0458	IC314239
184	MI-0459	IC314235
185	MI-0460	IC314240
186	MI-0461	IC314236
187	MI-0462	IC314237
188	MI-0463	IC314238
189	MI-0464	IC314057
190	MI-0466	IC314059
191	MI-0467	IC314060
192	MI-0468	IC314061
193	MI-0471	IC313999
194	MI-0472	IC314000
195	MI-0476	IC313947
196	MI-0477	IC313948
197	MI-0478	IC313949
198	MI-0479	IC313950
199	MI-0480	IC313951
200	MI-0482	IC313953
201	MI-0486	IC314062
202	MI-0487	IC314063
203	MI-0490	IC314066
204	MI-0492	IC314192
205	MI-0495	IC314069
206	MI-0496	IC314070
207	MI-0497	IC314226
208	MI-0500	IC314071
209	MI-0502	IC314073
210	MI-0503	IC314074
211	MI-0505	IC314076
212	MI-0506	IC313897
213	MI-0508	IC314078
214	MI-0509	IC314079
215	MI-0510	IC314227

216	MI-0515	IC314081
217	MI-0516	IC314082
218	MI-0519	IC314085
219	MI-0522	IC314002
220	MI-0526	IC313955
221	MI-0534	IC314229
222	MI-0536	IC314231
223	MI-0542	IC314208
224	MI-0545	IC314210
225	MI-0550	IC314090
226	MI-0552	IC314092
227	MI-0554	IC314093
228	MI-0555	IC313904
229	MI-0558	IC314095
230	MI-0559	IC314096
231	MI-0565	IC313909
232	MI-0567	IC313911
233	MI-0573	IC314262
234	MI-0575	IC314248
235	MI-0594	IC313960
236	MI-0627	IC405775
237	MI-0631	IC405779
238	MI-0634	IC405782
239	MI-0635	IC405783
240	MI-0639	IC405787
241	MI-0644	IC405792
242	MI-0649	IC405797
243	MI-0650	IC405798
244	MI-0652	IC405800
245	MI-0654	IC405802
246	MI-0657	IC405805
247	MI-0669	IC405817
248	MI-0670	IC405818
249	MI-0675	IC405823
250	MI-0678	IC405826
251	MI-0681	IC405829
252	MI-0683	IC405831
253	MI-0686	IC405833
254	MI-0695	IC405843
255	MI-0702	IC405850
256	MI-0706	IC405854
257	MI-0708	IC405856
258	MI-0710	IC405858
259	MI-0711	IC405859
260	MI-0712	IC405860
261	MI-0713	IC405861
262	MI-0714	IC405862
263	MI-0717	IC572938
264	MI-0724	IC572945
265	MI-0725	IC572946
266	MI-0726	IC572947
267	MI-0727	IC572948
268	MI-0728	IC572949
269	MI-0744	IC572965
270	MI-0752	IC572973
271	MI-0757	IC572978
272	MI-0759	IC572980
273	MI-0764	IC572985
274	MI-0770	IC572991
275	MI-0771	IC572992
276	MI-0775	IC572996
277	MI-0777	IC572998
278	MI-0778	IC572999
279	MI-0779	IC573000
280	MI-0783	IC573004
281	MI-0786	IC573007
282	MI-0789	IC573010
283	MI-0790	IC573011
284	MI-0792	IC573013
285	MI-0793	IC573014
286	MI-0797	IC573018
287	MI-0799	IC573020
288	MI-0806	IC573027
289	MI-0808	IC573029
290	MI-0820	IC573041
291	MI-0821	IC573042
292	MI-0823	IC573044
293	MI-0824	IC573045
294	MI-0827	IC573048
295	MI-0830	IC573051
296	MI-0831	IC573052
297	MI-0835	IC573056
298	MI-0836	IC573057
299	MI-0863	IC590595
300	MI-0864	IC590596
301	MI-0865	IC590597
302	MI-0868	IC590600
303	MI-0871	IC590603

304	MI-0873	IC590605
305	MI-0875	IC590607
306	MI-0876	IC590608
307	MI-0881	IC590613
308	MI-0882	IC590614

309	MI-0884	IC590616
310	MI-0886	IC590618
311	MI-0892	IC590622
312	MI-0893	IC590623

2.1 MORPHOLOGICAL CHARACTERIZATION

The selected accessions were characterized for 17 qualitative morphological parameters (Table 2). Four plants of each accession were characterized.

Table 2: Morphological descriptors used for duplicate identification

SN	Morphological descriptors
1	Branching nature
2	Curve or straightness of the branch
3	Color of young shoot
4	Color of mature shoot
5	Stipule nature
6	Stipule duration
7	Phyllotaxy
8	Lobation type
9	Lobation number
10	Leaf nature
11	Leaf color
12	Leaf surface
13	Leaf texture
14	Leaf apex
15	Leaf base
16	Leaf margin
17	Leaf shape

2.1.1 Isolation of genomic DNA

Fresh young leaves were harvested from four plants of each accession.Genomic DNA was isolated from the leaves of mulberry germplasm using a modified Cetyl-tri-methyl ammonium bromide (CTAB) method (Doyle and Doyle 1990). The DNA was dissolved in 50 µl of 10mM TE buffer (pH 8) and stored at -20^{0}C till further use. The quality of DNA was measured using 0.8% agarose gel electrophoresis containing ethidium bromide at a final concentration of 0.5 µg/ ml.

2.1.2 SSR primers

Four polymorphic SSR primers developed by Mathithumilan et al. (2013) were used for the study. All the primers were synthesized and purchased from Eurofins India Pvt. Ltd., Bengaluru. The list of SSR primers used for the study is given table 4.

Table4: List of SSR primers and their efficiency parameters

SN	Marker	Allelic range	No. of alleles	Frequency of alleles	PIC value
1	MISSR15	200-250	2	0.53	0.12
2	MISSR20	170-300	4	0.27	0.29
3	MISSR35	200-250	2	0.93	0.12
4	MULSSR29	150-220	3	0.78	0.31

5	**MULSSR39**	190-260	3	0.40	**0.36**
6	MULSSR96B	280-300	4	0.62	0.33
7	**MULSSR258**	150-240	3	0.29	**0.37**
8	**M2SSR68**	190-210	3	0.42	**0.46**
9	**M2SSR87**	220-330	3	0.48	**0.43**
10	M2SSR89A	150-230	2	0.71	0.29
11	M2SSR36	180-220	2	0.58	0.18
12	**M2SSR112A**	180-250	4	0.59	**0.39**

PIC > 0.5 (Highly informative); (Botstein *et al*., 1980)

2.1.3 Polymerase Chain Reaction (PCR)

PCR was carried out in Master Cycler Nexus Gradient (Eppendorff, USA) with a total volume of 10µl of reaction mixture (Takara Bio.) containing 5µl buffer, 1 µl each of 10 µM primer (forward and reverse), 2 µl of PCR grade water and 1 µl genomic DNA (~ 6ng). The DNA amplification was performed with an initial denaturation at 94^0C for 5 minutes followed by 35 cycles of denaturation at 94^0C for 30 seconds, annealing at 55^0C for 45 seconds, extension at 72^0C for 1 minute and a final extension at 72^0C for 8 minutes.

The PCR products were separated on 2% agarose gel in 1X Tris Acetic Acid and EDTA (TAE) containing 10 mg/ml ethidium bromide as stain. The gel images were recorded using Bio-Print Gel Documentation System (VilberLourmat, France). The primers which gave polymorphic bands were shortlisted for further Poly Acrylamide Gel Electrophoresis (PAGE) analysis.

2.1.4 Poly Acrylamide Gel Electrophoresis (PAGE) analysis

A 6% polyacrylamide gel was prepared following the protocol described by Sambrook et al. (1989). Gel was run in a horizontal electrophoresis unit for 80 Volts, 3 hrs. The gel unit was covered with ice packs for better resolution of the amplicons.

2.1.5 Statistical analysis

The data matrix was subjected for cluster analysis. The polymorphic information content (PIC) values described by Botstein et al. (1980) were used to refer the relative value of each marker with respect to the amount of polymorphism exhibited. The PIC values for each primer (Table 4) were estimated using the formula given by Nei (1973).

$$\mathrm{PICi} = 1 - \sum_{f=1}^{n} (Pij)^2$$

3. RESULTS

3.1 Morphological characterization:

All the 312 accessions were characterized for 17 descriptors and are coded using 0-9 scale numerically for each character state as detailed below:

1. Branching nature: The branching nature of plants were recorded based on visual observations from 4 plants/ accession after 90 days of pruning and grouped into Erect, Semi-erect, and Spreading.

2. Curve or straightness of branch: Curve or straightness of the branches were grouped as Curved, Slightly Curved and Straight depending on plant habit.

3. Colour of young shoot: Young shoot colour were recorded on visual observations from 4 plants/ accession after 30 days of pruning and scored as Brown, Green and Purple.

4. Colour of mature shoot: Mature shoot colour on visual observations from 4 plants/ accession from the mature shoot after 90 days of pruning and grouped into Brown, Green, Greenish brown, Greenish grey and Purple brown.

5. Phyllotaxy: Phyllotaxy is the mode of arrangement of leaves on the stem. The phyllotaxy is called 1/2 rank (distichous) and followed by 1/3, 2/5. Different combinations of all (1/2, 1/3 and 2/5) are termed as mixed type.
6. Stipule Nature: Stipules are outgrowths from leaf base, which usually protects the young axillary buds. Depending on hanging arrangement of stipules nature, grouped into foliaceous, free-lateral and bud scale.
7. Stipule Duration: The attachment of stipules with leaf base was observed visually and depending on attachment duration, it was grouped into caducous and persistent.
8. Leaf lobation type: The leaf lobation was observed visually and grouped primarily into Unlobed and Lobed. Under lobed group it was again grouped into Deeply Lobed, Medium Lobed and Shallow Lobed.
9. Leaf lobation number: The lobation number on a leaf was counted and scored as No lobation (0), Lobation with (1-5) and Lobation with (6-10).
10. Leaf nature: The leaf nature was recorded based on unlobed, lobed and grouped into Homophyllous (All the leaves in the plant are of similar nature either entire or lobed), Heterophyllous (when some leaves are lobed and some are unlobed either in the same branch or in different branches of the same plant).
11. Leaf colour: The mature leaf colour was observed from each accession on 3 plants and scored visually as Green (G) and Dark Green (DG).
12. Leaf surface: The leaf surface was grouped based on feeling by touch on it and grouped into Rough, Slightly Rough and Smooth.
13. Leaf texture: The leaf texture was grouped based on feeling method and grouped into a) Chartaceous, opaque and like writing paper and b) Coriaceous, leathery, thick and stiff .
14. Leaf apex: It is the portion of leaf bounded by approximately the upper 25% of the leaf margin and based on observation, the leaf apex grouped into a) Acute (pointed and narrow), b) Acuminate (the apex is drawn out into a short tapering tail) and c) Caudate (when the apex shows long tapering tail).
15. Leaf margin: The margin of the lamina was grouped into a) Crenate (margin toothed and teeth are rounded without a pointed apex) b) Serrate (serrations are pointed with their axes inclined to the trend of margin), c) Dentate (if the margin is tooth-like edges), and d) Repand (when the margins form a smooth line or arc without noticeable projections).
16. Leaf base: It is the portion of the leaf bounded by approximately lower 25% of the margin. The observations were recorded into a) Cordate, leaf base embayed in a sinus whose sides are straight or convex, b) Truncate, leaf base terminating abruptly as if cut, margin perpendicular to the mid vein and nearly so, c) Lobate, leaf base small to large rounded projections whose inner margins towards the petiole and are in partly concave shape.
17. Leaf shape: The leaf shape is ovate in nature which represents the greatest width intersecting the leaf axis basal to the mid point of the later axis and based on (L/W) ratio of leaf it was grouped as Narrow Ovate (2:1), Ovate (1.5:1), Wide Ovate (1.2:1) and Cordate (<1:1).

A data matrix was prepared and subjected for the cluster analysis using PAST software (available as supplementary file S-1). The dendrogram (Fig.1) depicting 84 duplicate accessions are given. The suspected duplicate accessions were given in table 3. These suspected duplicates were subjected for further molecular characterization using 4 polymorphic SSR markers.

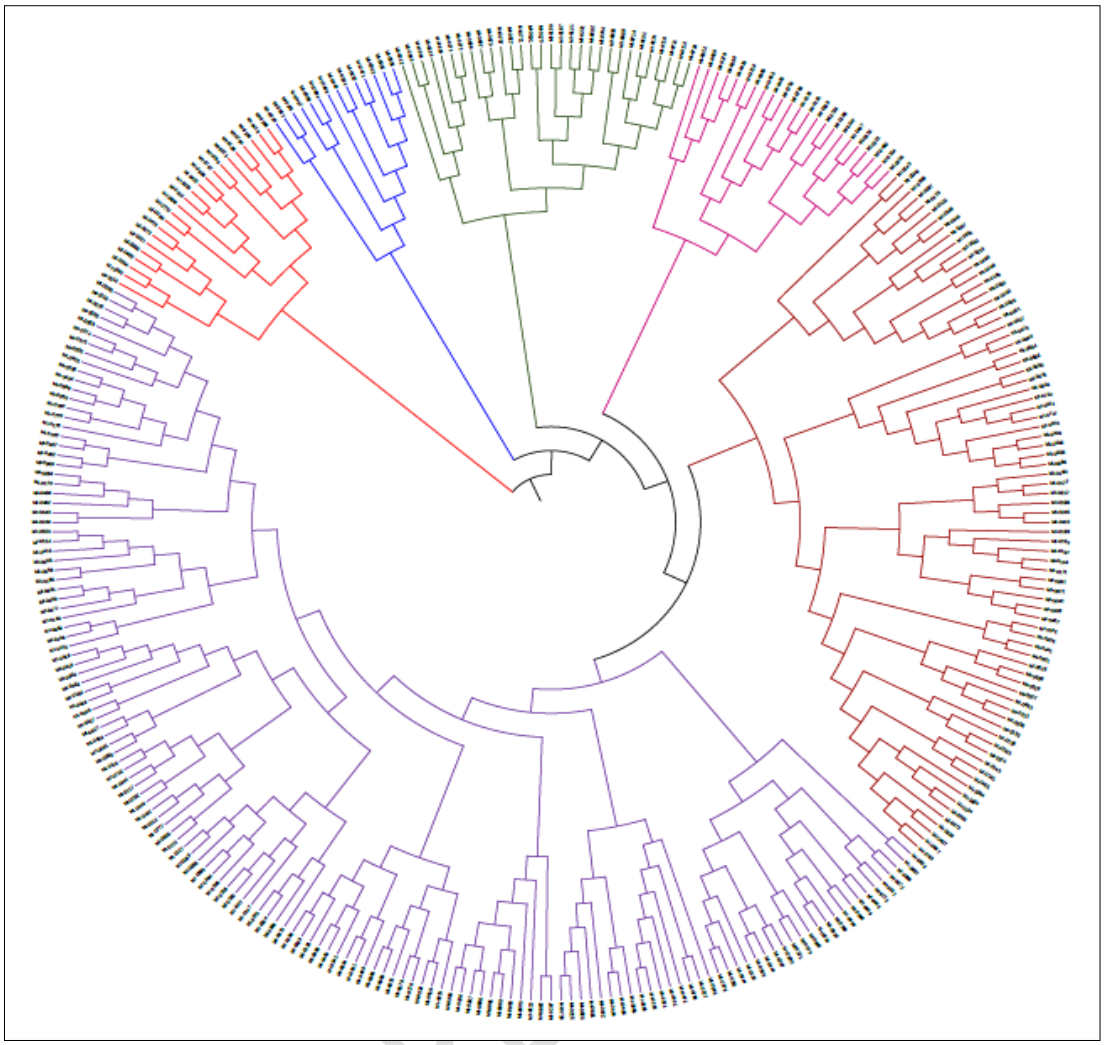

Fig. 1: Cluster analysis based on morphology
(312 accessions; 84 accs. with similar morphology)

Table 3: List of suspected duplicates

SN	Acc. No.
1	MI-0144 & MI-0197
2	MI-0657 & MI-0670
3	MI-0710 & MI-0711
4	MI-0712 & MI-0713
5	MI-0714 & MI-0717
6	MI-0724 & MI-0422
7	MI-0010 & MI-0225
8	MI-0221 & MI-0232
9	MI-0744 & MI-0150
10	MI-0187 & MI-0193
11	MI-0374 & MI-0708
12	MI-0478 & MI-0789
13	MI-0067 & MI-0281
14	MI-0370 & MI-0502
15	MI-0503 &MI-0505
16	MI-0152 & MI-0161
17	MI-0461 & MI-0759
18	MI-0134 & MI-0335
19	MI-0824 & MI-0163
20	MI-0334 & MI-0490
21	MI-0495 & MI-0496
22	MI-0218 & MI-0820
23	MI-0213 & MI-0652
24	MI-0359 & MI-0425
25	MI-0329 & MI-0555
26	MI-0835 & MI-0836
27	MI-0567 & MI-0783
28	MI-0407 & MI-0202
29	MI-0238 & MI-0204

30	MI-0506 & MI-0138
31	MI-0764 & MI-0072
32	MI-0116 & MI-0117
33	MI-0344 & MI-0702
34	MI-0706 & MI-0355
35	MI-0779 & MI-0345
36	MI-0349 & MI-0284

37	MI-0486 & MI-0271
38	MI-0272 & MI-0318
39	MI-0319 & MI-0320
40	MI-0725 & MI-0726
41	MI-0325 & MI-0326
42	MI-0864 & MI-0865

3.2 Allele scoring and data analysis

The clear and reproducible alleles amplified by each SSR among 84 accessions (Fig.1 to 4) were scored according to their fragment size (bp) corresponding to 100 bp molecular weight marker (Renvik Bio). Only consistent and bright SSR bands were scored as 1 (for presence of band) and 0 (for absence of band) and each character was treated independently. The bands which were very faint were not considered for scoring. The data was transferred into the binary matrix (allele data available as supplementary file S-2).

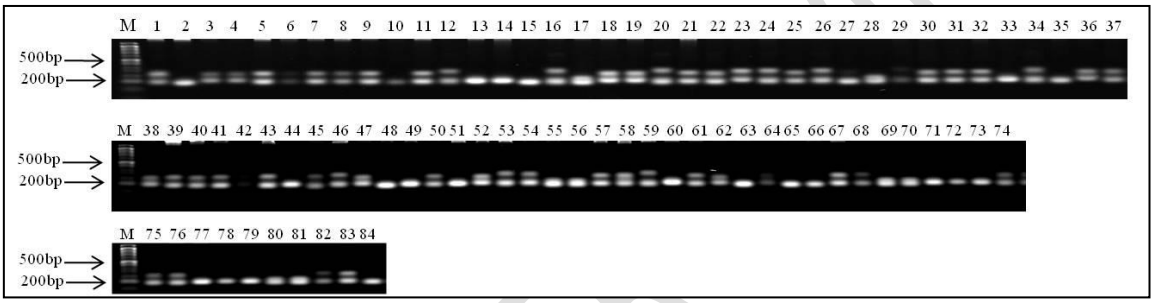

Fig. 2: PCR profile of 84 suspected duplicates for the marker MISSR15

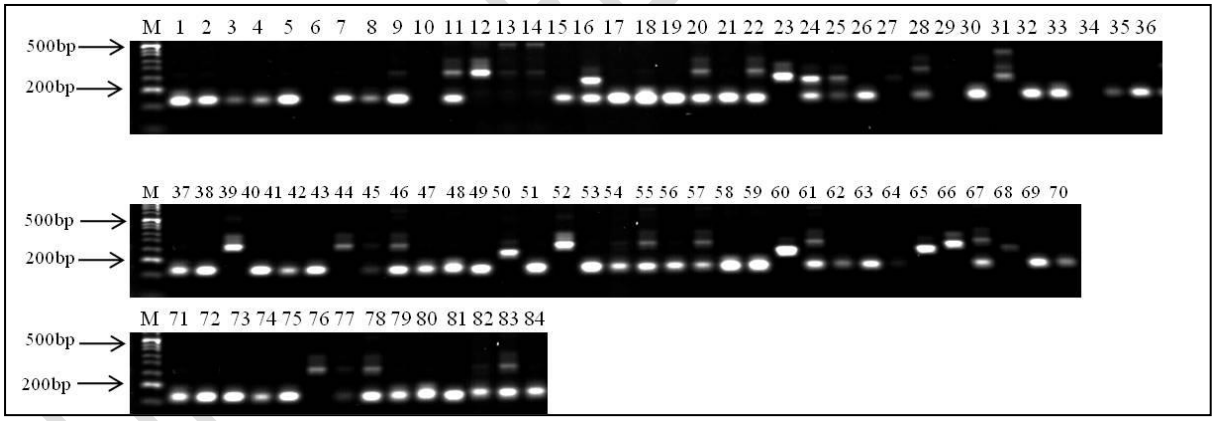

Fig. 3: PCR profile of 84 suspected duplicates for the marker MISSR20

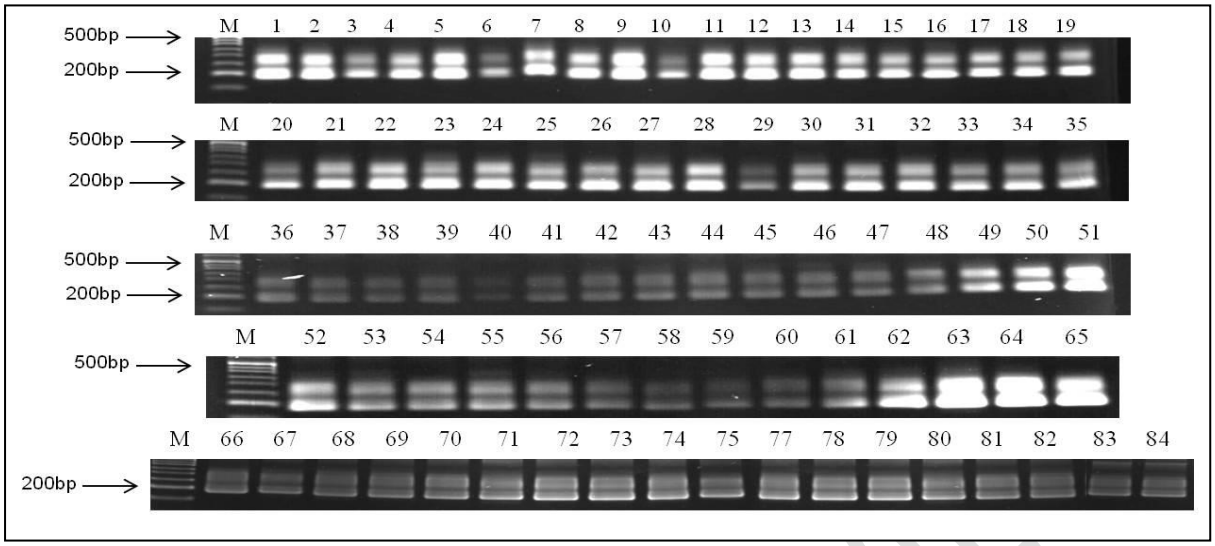

Fig. 4: PCR profile of 84 suspected duplicates for the marker MISSR35

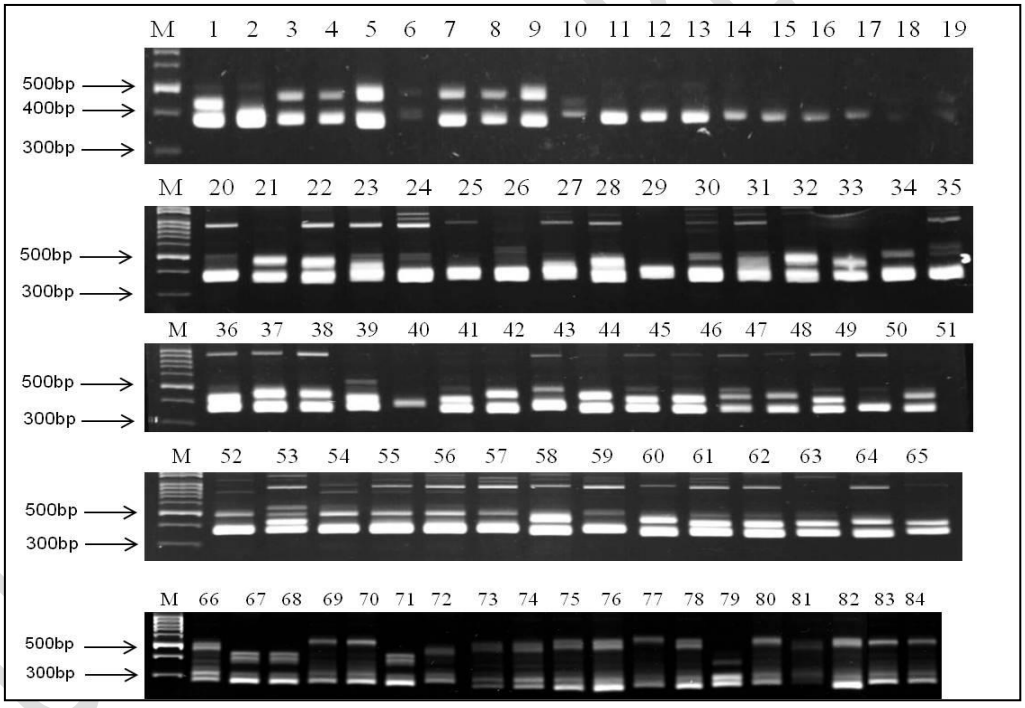

Fig. 5: PCR profile of 84 suspected duplicates for the marker MULSSR29

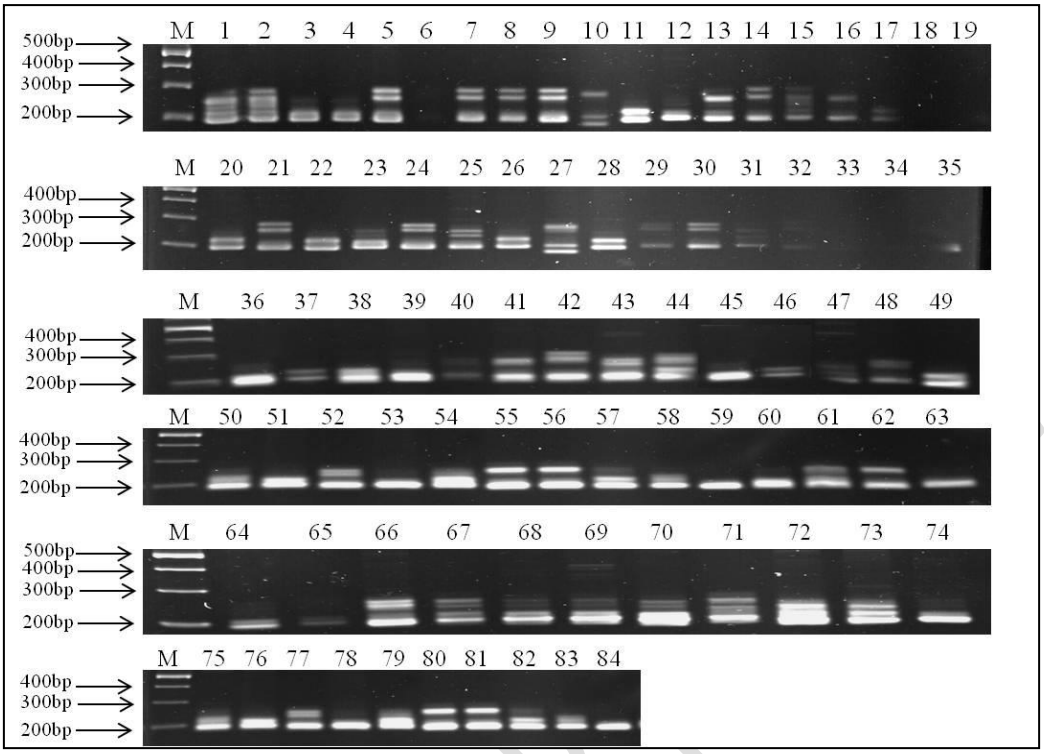

Fig. 6: PCR profile of 84 suspected duplicates for the marker MULSSR39

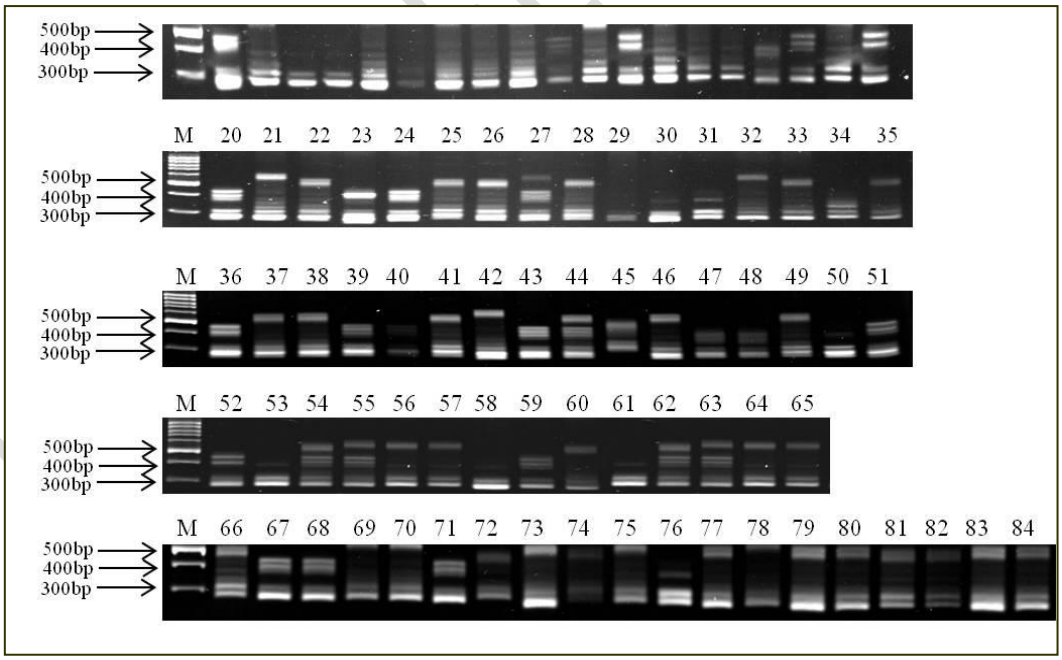

Fig. 7: PCR profile of 84 suspected duplicates for the marker MULSSR96B

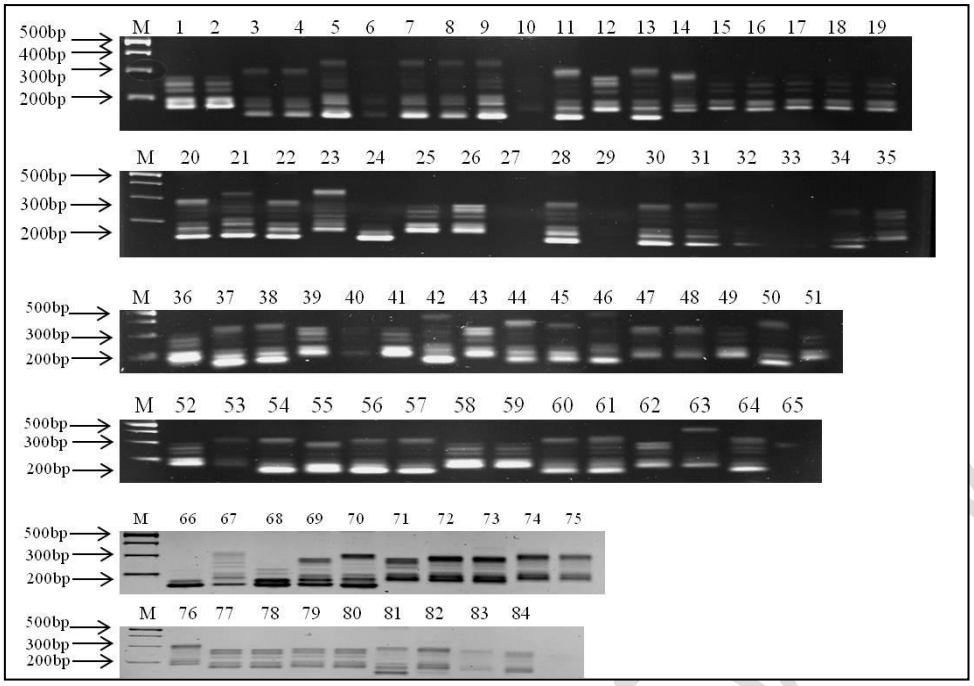

Fig. 8: PCR profile of 84 suspected duplicates for the marker MULSSR258

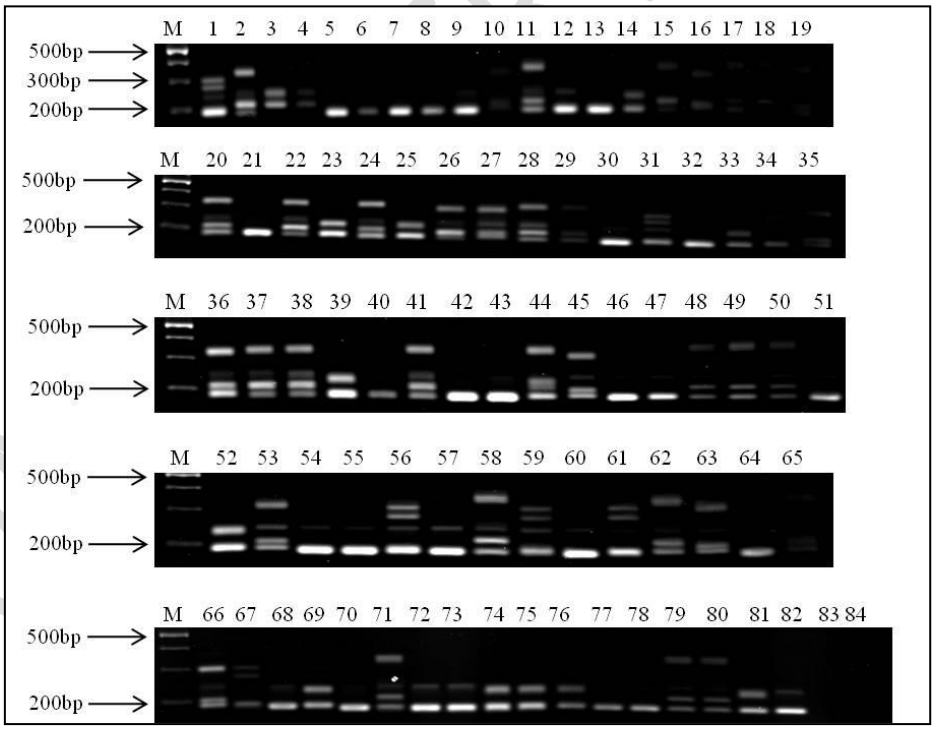

Fig. 9: PCR profile of 84 suspected duplicates for the marker M2SSR68

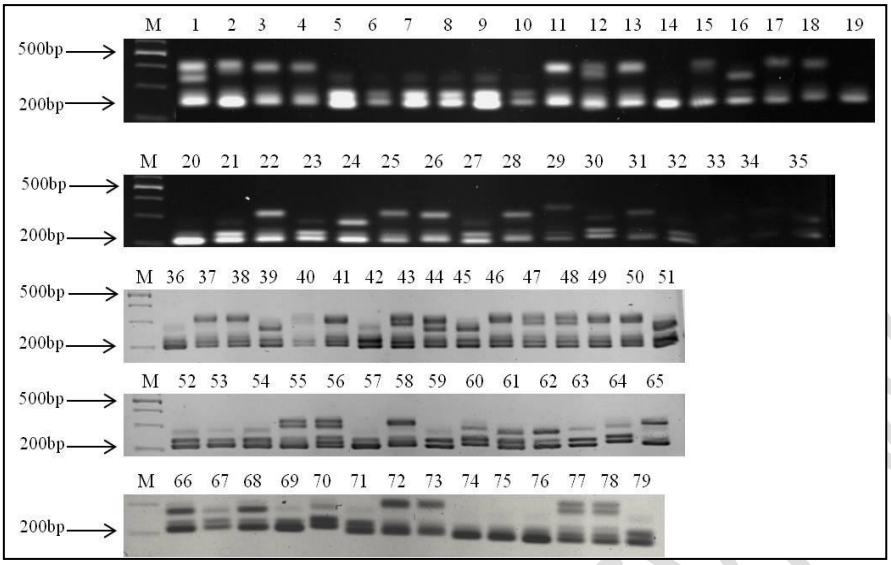

Fig. 10: PCR profile of 84 suspected duplicates for the marker M2SSR87

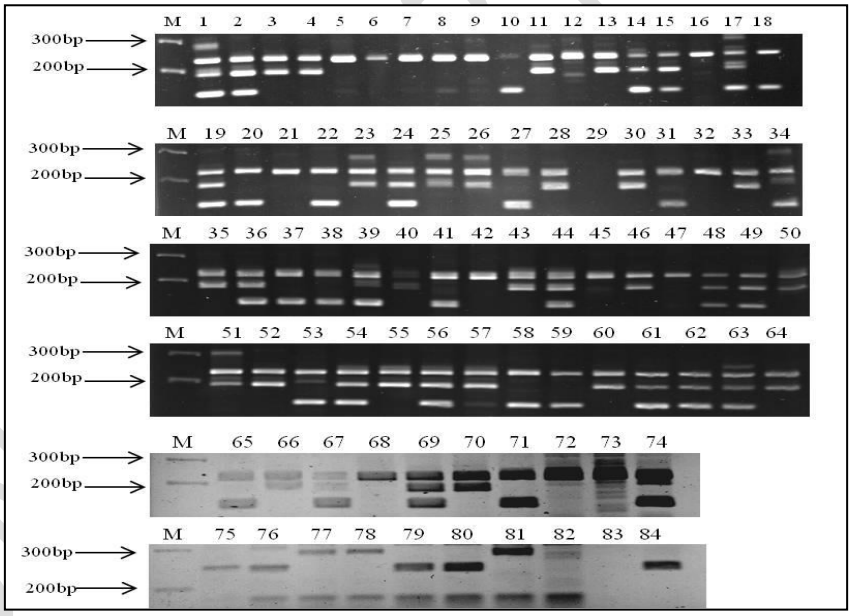

Fig. 11: PCR profile of 84 suspected duplicates for the marker M2SSR89A

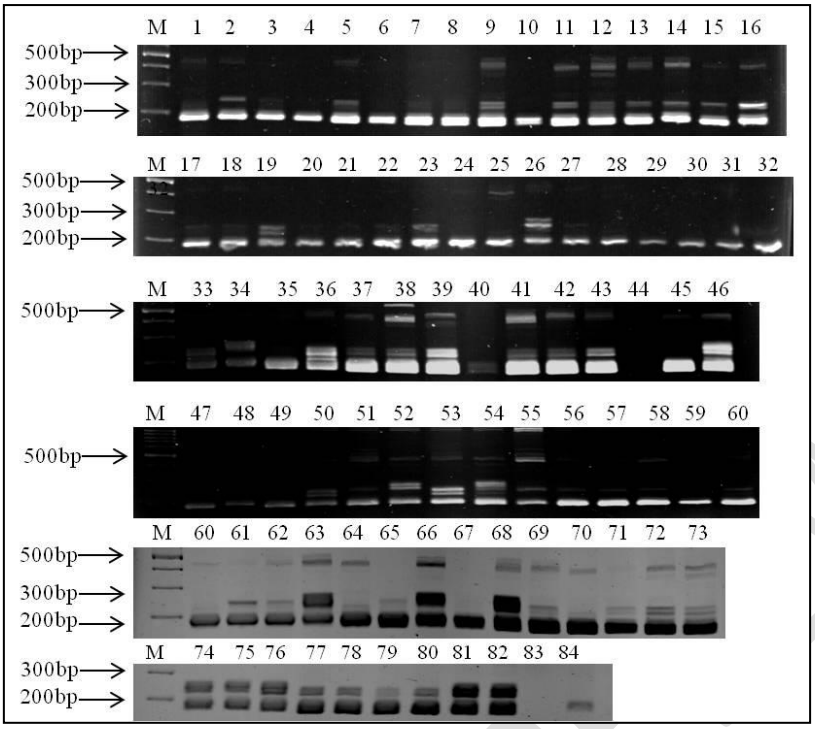

Fig. 12: PCR profile of 84 suspected duplicates for the marker M2SSR36

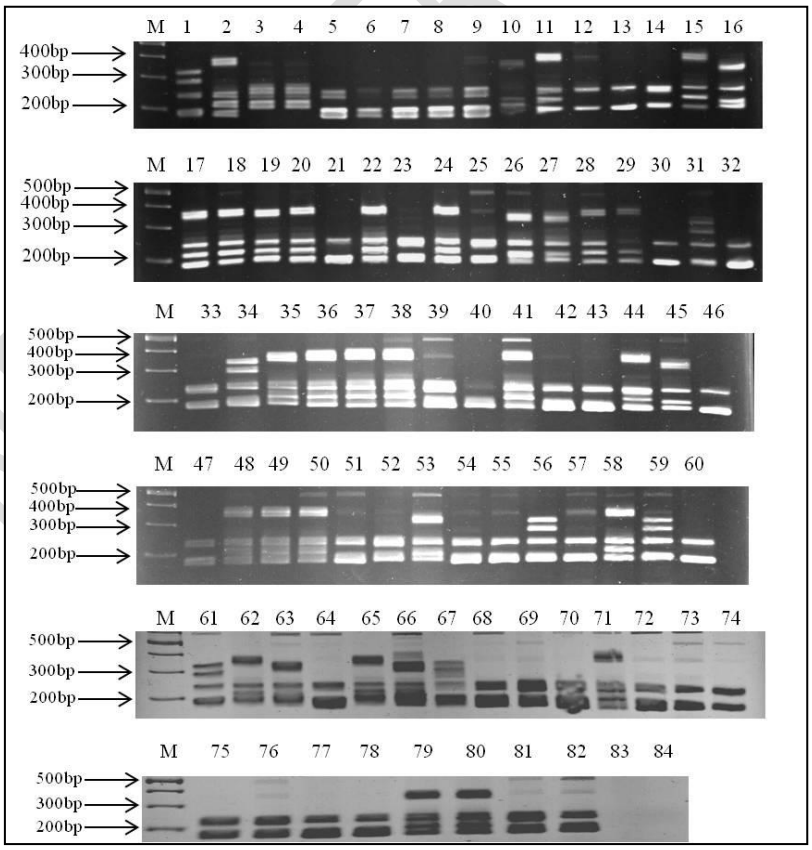

Fig. 13: PCR profile of 84 suspected duplicates for the markerM2SSR112A

Table 5: SSR primers and their efficiency parameters

SN	Marker	Allelic range	No. of alleles	Frequency of alleles	PIC value
1	MISSR15	200-250	2	0.53	0.12
2	MISSR20	170-300	4	0.27	0.29
3	MISSR35	200-250	3	0.93	0.12
4	MULSSR29	380-480	3	0.78	0.31
5	MULSSR39	190-260	3	0.40	0.36
6	MULSSR96B	280-400	3	0.62	0.33
7	MULSSR258	150-240	3	0.29	0.37
8	M2SSR68	190-210	3	0.42	0.46
9	M2SSR87	220-300	3	0.48	0.43
10	M2SSR89A	150-230	2	0.71	0.29
11	M2SSR36	180-220	2	0.58	0.18
12	M2SSR112A	180-250	4	0.59	0.39

PIC > 0.5 (Highly informative); (Botstein et al. 1980)

3.3 Polymorphism Information Content (PIC)

The PIC takes into account not only the number of alleles that are expressed but also the relative frequencies of those alleles (Smith et al. 1997). Cluster analysis was performed by subjecting the character state data using the software PAST (Fig.6). The Dendrogram was constructed using the unweighted pair group method with arithmetic average (UPGMA) procedure. A total of 7 clusters generated based on SSR data. The relationships were compared by visual examination of Dendrogram derived from clustering analysis. It is evident that 4 SSR markers are found to be polymorphic which could be used to screen the mulberry germplasm.

List of duplicates confirmed based on both morphological descriptors and SSR markers:

1. MI-0713 (Gujarat) & MI-0712 (Gujarat)
2. MI-0670 & MI-0657 (N. India HP)
3. MI-0490 (Kerala) & MI-0334 (Tamil Nadu)
4. MI-0325 & MI-0326 (ERRC, KL)
5. MI-0725 & MI-0726 (Arunachal Pradesh)
6. MI-0271 & MI-0272 (Kerala)
7. MI-0318 & MI-0319 (Rajasthan)

In the present study, the qualitative morphological descriptors were characterized for the selected 312 *Morus indica* accessions. These descriptive data was converted to binary matrix which was subjected to clustering analysis using PAST (4.03 version) software. The similar accessions were considered as suspected duplicates and were shortlisted for SSR marker analysis. 12 SSR markers were employed to screen the suspected duplicates and confirmed 14 accessions as true duplicates which similar both morphological and genetically.

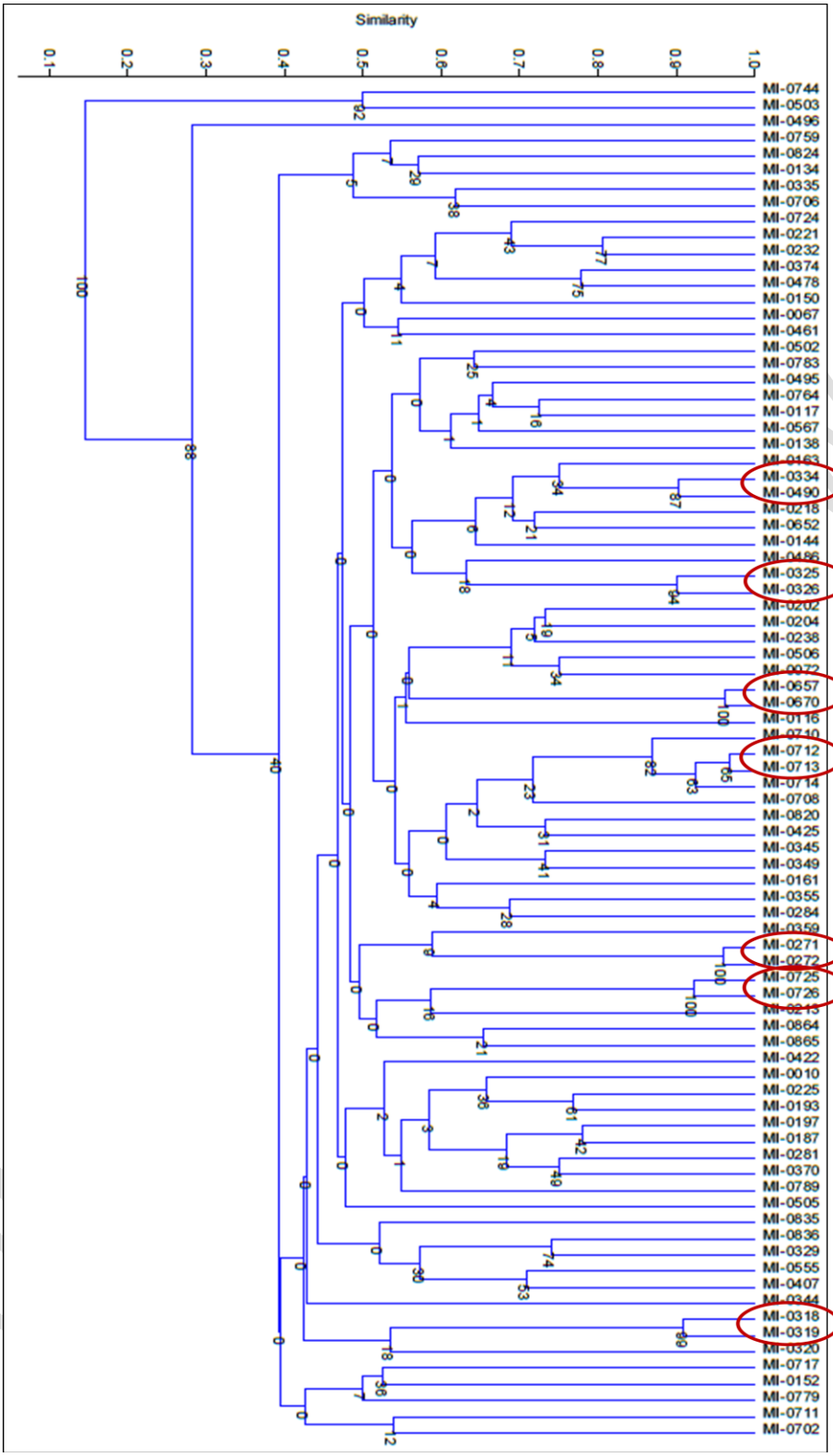

Fig. 14: Dendrogram based on SSR markers
(Circled accessions are confirmed duplicates)

4. DISCUSSION

The identification of duplicates in germplasm banks of vegetatively propagated plants seems to be common (Irish *et al*., 2010; van Treuren*et al*., 2010). This may be due to the usual exchange of propagules among farmers of different regions, especially when the species has economical importance. In the new locality, the genotype may receive a new name which may lead to confusion on samples and maintenance of accessions. Sometimes the accessions collected from different geographical regions might have been originally migrated from their original location to other areas.

Over a period of time, the accessions collected from the migrated place may turn out to be duplicated after characterization and evaluation which might be the case in the present study. However duplicate identification is important to cluster accessions and to avoid crossings between them. The study showed the occurrence of duplicates in mulberry germplasm composed mainly of landraces collected from Gujarat, Rajasthan, Berhampur, Maharashtra, South India, Assam and Meghalaya. The re-sampling of the same genotype or variety over many years in different places may reflect the importance of those genotypes for farmers.

Passport data and morphological descriptors are very important for any crop germplasm in deciding the probable duplicates. Virk *et al*. (1995) identified duplicates in Rice germplasm at IRRI using 22 morphological traits and 7 RAPD markers. They proposed both passport and RAPD markers as two procedures for the identification of duplicates in Rice germplasm. They have suggested that the passport data including both qualitative and quantitative should be considered preliminarily to identify probable duplicates. The probable duplicates may be further subjected to marker analysis. In the mulberry field gene bank, both the qualitative morphological descriptors and SSR markers were used to identify the probable duplicates and their confirmation.

Hintum and Knupffer (1995) identified the probable duplicates in barley germplasm based on passport data. Though passport data is not reliable to decide the duplicates, they emphasized that passport data can be used to identify the probability of duplication. Only to a lesser extent, the probable duplication can be predicted using passport data. Similarly, in the present study, morphological data was considered to identify the suspected duplicates.

Le Clerc *et al*. (2005) evaluated the carrot germplasm through molecular markers and identified duplicate accessions. They have identified 21 presumed duplicates and only accessions which were not distinguished on a morphological basis were subjected to molecular analysis with the help of AFLP markers.

Naik and Dandin (2006) have identified duplicates in mulberry germplasm using RAPD markers. Based on passport and morphological data, they have identified 14 suspected duplicates and the same were subjected to RAPD analysis.

Molecular characterization of Apricot germplasm was carried out by Martin et al. (2011) where two set of SSRs were used for screening. They have evaluated the variability of cultivars preserved in *ex situ* collections. They have emphasized that SSR markers are most useful in studying the genetic diversity of other species in different areas. In the present study SSR markers were used to characterize the mulberry accessions for the identification of duplicates.

Das *et al*. (2020) have identified duplicates in the Ginger germplasm collection. They employed 9 ISSR and 22 SSR markers for the analysis of variation among 60 accessions. Based on the mean genetic distance, out of 60 accessions, 31 were found to be distinct and the rest 29 accessions were considered as potential duplicates. In the current study, both morphological descriptors and SSR markers were consideredfor the identification of duplicates. Both phenotypic and SSR markers were employed in identifying the duplicate accession in lettuce germplasm (Sochor *et al*., 2019). They have screened 39 accessions of *Lactuca sativa* using 17 morphological descriptors and 23 SSR markers. Among them, only 10 markers gave polymorphism and 19 were identified as duplicates. Anil *et al*. (2020) analyzed fifty five accessions of sweet potato using morphological and 11 ISSR markers and identified 3 pairs of duplicates based on 100% similarity. In the current study, 7 pairs of

duplicates were identified with help of 17 qualitative morphological descriptors and 12 polymorphic SSR markers.

Palme *et al*. (2020) have evaluated duplicate holdings of *Brassica oleracea* using 10 morphological traits and 11 SNP markers. They have suggested the usefulness of SNP markers for distinguishing the cabbage duplicates and also emphasized the necessity of duplicate identification within the germplasm especially cross pollinated species such as cabbage. In accordance with this, mulberry is also an open pollinated species where the chances of duplication are high. In the present investigation, duplicates were identified using both morphological traits and SSR markers. The study also confirmed the availability of polymorphic SSR markers for future screening of accessions.

Rosell *et al*. (2022) have identified 197 duplicates out of 360 sweet potato accessions. They have considered 33 morphological descriptors and AFLP markers for the confirmation of duplicates. Similarly, both morphological descriptors and SSR markers were employed to identify 84 suspected duplicates among mulberry germplasm. Out of these, 14 duplicates were confirmed based on 12 polymorphic SSR markers.

Conclusion

Germplasm collections may contain duplicate accessions due to various reasons such as the presence of identical material registered under different identifiers. Accessions might have been collected from common or identical collection areas. Sometimes expeditions may have been carried out without prior knowledge about the distribution of genetic variation across the areas. The diversity in the germplasm collections is a prerequisite for the effective utilization of genetic resources. In the present study, both morphological descriptors and SSR markers were considered for the identification of duplicates. The results of the study highlighted that out of 84 suspected duplicates, 14 accessions were confirmed as duplicates. Among 86 SSR primers screened, 12 are found to be polymorphic. These polymorphic markers can be used to screen mulberry germplasm. Moreover, most of the accessions collected from Gujarat, Rajasthan, Assam, Arunachal Pradesh and South Indian states like Tamil Nadu, Karnataka and Kerala are found to be duplicates. It is evident that both morphological and molecular characterization with highly informative markers is important to determine duplications if any and gives valuable trait information for future researchers and breeders.

Disclaimer

The authors declare that NO generative AI technologies were used during writing or editing the manuscript.

Competing interests

Authors declare that no competing interests exist.

5. REFERENCES

Banerjee, R., Roychoudhury, S. S., Das, B., Saha, A.K., Saratchandra, B.& Bajpai, A.K. (2011). Phenotypic divergence in mulberry (Morus spp.) germplasm collections. Journal of Crop Improvement, 25 (5),459-471.

Botstein, D., White, R.L., Skolnick, M. & Davis, R.W.(1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. American Journal of Human Genetics, 32(3), 314-331.

Das, A., Sahoo, R.K., Barik, D.P. & Subudhi, E.(2020). Identification of duplicates in ginger germplasm collection from Odisha using morphological and molecular characterization. Proceedings of the National Academy of Sciences, India Section B: Biological Sciences, 3, 10.

Dobrovolskaya, O., Saleh, U., Malysheva, O.,Rodder, M.S.&Borner, A.(2005). Rationalization of germplasm collections: a case study of wheat. Theoretical and Applied Genetics,111, 1322-1329.
Doyle, J.J., & Doyle, J.L.(1990). Isolation of plant DNA from fresh tissue. Focus, 12, 13-15.
Elisa, F.M., Tome, F.N., Jose, E.S., Diehgo, T.D., &Girena, F.R.(2013). Identification of duplicates of cassava accessions sampled o the North Region of Brazil using microsatellite markers. Acta Amazonica, 43(4), 461-468.
Hintum, T.J.L.,&Knupffer, H.(1995). Duplication within and between germplasm collection. I. Identifying duplication on the basis of passport data. Genetic Resources and Crop Evolution, 42,127-133.
Irish, B.M., Goenaga, R., Zhang, D., Schnell, R., Brown, J.S., &Motamayor, J.C.(2010). Microsatellite fingerprinting of the USDA-ARS tropical agriculture research station cacao (Theobroma cacao L.) germplasm collection. Crop Science, 50, 656-667.
Jayeoba, F.M., Ogunbanjo, O.R., &Awotoye, O.O.(2014). Characterization of mulberry (Morus alba L.) varieties in South Western Nigeria. Continental Journal of Agricultural Science, 8(1), 13-19.
Khurana, P., & Checker, V.(2011). The advent of genomics in mulberry and perspectives for productivity enhancement. Plant Cell Reports, 30, 825-838.
Martin, C., Herrero, M., & Hormaza, J.I.(2011). Molecular characterization of apricot germplasm from an old stone collection. PLoS One, 6 (8), e23979.
Mathithumilan, B., Kadam, N.N., Biradar, J., Reddy, S.H., Ankaiah, M., Narayan, M.J., Markarla, U., Khurana, P., & Sreeman, S.M.(2013). Development and characterization of microsatellite markers for Morus spp. and assessment of their transferability to other closely related species. BMC Plant Biology,13, 194.
Naik, V.G., & Dandin, S.B. (2006).Identification of duplicate collections in the mulberry (Morus spp.) germplasm using RAPD analysis. Indian Journal Of Genetics And Plant Breeding, 66(4), 287-292.
Nei, M.,(1973). Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences (PNAS), 70, 3321-3323.
Plalme, A.E., Hagenland, J., Solberg, S.O., Aloisi, K. and Artemyeva, A.(2020). SNP markers and evaluation of duplicate holding of Brassica oleracea in two European gene banks. Plant Basel, 9, 925.
Rossel, G., Robles, R. &Ay, D.(2009). Identification of duplicate accessions within a sweet potato germplasm collection using morphological characterization and AFLP markers. In: 15th triennial ISTRC symposium: 2-6 November, Peru. International Society for Tropical Root Crops, Peru pp. 25-30.
Sambrook, J., Fritsch, E. R.,& Maniatis, T.(1989). Molecular Cloning: A Laboratory Manual (2nded.). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
Sharma, A., Sharma, R., &Machii, H.(2000). Assessment of genetic diversity in a Morus germplasm collection using fluorescence-based AFLP markers. Theoretical and Applied Genetics, 101, 1049-1055.
Sisko, M. (2016). Identification of hypothetical duplicate accessions of plums (Prunusdomestica L.) within the slovene plant gene bank collection using molecular markers. Agricultura 13(1-2), 57-64.
Smith, J.S.C, Chin, E.C.L., Shu, H., Smith, O.S., Wall, S.J., Senior, M.L., Mitchell, S.E., Kresovich, S., &Ziegle, J.(1997). An evaluation of the utility of SSR loci as molecular markers in maize (Zea mays L.): comparisons with data from RFLPS and pedigree. Theoretical and Applied Genetics, 95, 165-173
Sochor, M., Jemelkova, M., &Dolezalova, I.(2019). Phenotyping and SSR markers as a tool for identification of duplicates in lettuce germplasm. Czech Journal of Genetics and Plant Breeding, 55, 110-119.
Suresh, K., Ghosh, M.K., Banerjee, R., Chakravarty, D.,&Trivedy, K.(2018). Multivariate analysis of indigenous and exotic mulberry (Morus spp.) germplasm for identifying diverse genotypes under humid subtropical region. Indian Journal of Pure & Applied Biosciences, 6(1), 618-627.

Thangavelu, K., Tikader, A., Ramesh, S.R., Rao, A.A., Naik, V.S., Sekar, S., &Deole, A.K.(2000). Catalogue on mulberry (Morus spp.) germplasm CSGRC, Hosur, India 2, pp 1-225.
Thriveni, M.C., Thanavendan, G., Raju Mondal, Jhansilaksmi, K., Saraswathi, P., Ravikumar, G., & Sekar, S.(2021). Catalogue on Mulberry (Morus spp.) germplasm, vol. 6, CSGRC, Hosur.
Tikader, A., & Kamble, C.K.(2008). Mulberry wild species in India and their use in crop improvement: a review. Australian Journal of Crop Science, 2(2), 64-72.
Tikader, A.,& Rao, A.A.(2002). Phenotypic variation in mulberry (Morus spp.) germplasm. Sericologia, 42, 221-233.
Van Treuren, R., Kemp, H., Ernsting, G., Jongejans, B., Houtman, H., & Visser, L.,(2010). Microsatellite genotyping of apple (Malus x domestica Borkh) genetic resources I the Netherlands: application in collection management and variety identification. Genetic Resources and Crop Evolution, 57, 853-865.
Virk, P.S., Newbury, H.J., Jackson, M.T., & Ford-Lloyd, B.V.(1995). The identification of duplicate accessions within a rice germplasm collection using RAPD analysis. Theoretical and Applied Genetics, 90, 1049-1055.