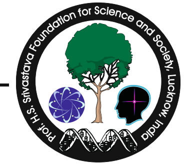
Physiol Mol Biol Plants (October–December 2013) 19(4):575–586
DOI 10.1007/s12298-013-0185-3



RESEARCH ARTICLE

Genetic diversity in *Capsicum* germplasm based on microsatellite and random amplified microsatellite polymorphism markers

Ved Prakash Rai · Rajesh Kumar · Sanjay Kumar · Ashutosh Rai · Sanjeet Kumar · Major Singh · Sheo Pratap Singh · Awadesh Bahadur Rai · Rajneesh Paliwal

Published online: 3 July 2013
© Prof. H.S. Srivastava Foundation for Science and Society 2013

Abstract A sound knowledge of the genetic diversity among germplasm is vital for strategic germplasm collection, maintenance, conservation and utilisation. Genomic simple sequence repeats (SSRs) and random amplified microsatellite polymorphism (RAMPO) markers were used to analyse diversity and relationships among 48 pepper (*Capsicum* spp.) genotypes originating from nine countries. These genotypes covered 4 species including 13 germplasm accessions, 30 improved lines of 4 domesticated species and 5 landraces derived from natural interspecific crosses. Out of 106 SSR markers, 25 polymorphic SSR markers (24 %) detected a total of 76 alleles (average, 3.04; range, 2–5). The average polymorphic information content (PIC) was 0.69 (range, 0.29–0.92). Seventeen RAMPO markers produced 87 polymorphic fragments with average PIC of 0.63 (range, 0.44–0.81). Dendrograms based on SSRs and RAMPOs generated two clusters. All 38 *Capsicum annuum* genotypes and an interspecific landrace clustered together, whereas nine non-*annuum* (three *Capsicum frutescens*, one *Capsicum chinense*, one *Capsicum baccatum* and four interspecific landraces) genotypes clustered separately. Genetic variation within non-*annuum* genotypes was greater than the *C. annuum* genotypes. Distinctness of interspecific derivative landraces grown in northeast India was validated; natural crossing between sympatric *Capsicum* species has been proposed as the mechanism of their origin.

Keywords *Capsicum* · SSR · RAMPO · Interspecific derivatives · Sympatric population

Electronic supplementary material The online version of this article (doi:10.1007/s12298-013-0185-3) contains supplementary material, which is available to authorized users.

V. P. Rai · S. P. Singh · R. Paliwal
Department of Genetics & Plant Breeding, Banaras Hindu University, Varanasi 221005, India

R. Kumar · A. Rai · M. Singh · A. B. Rai
Indian Institute of Vegetable Research, Varanasi 221305, India

S. Kumar
Department of Genetics & Plant Breeding, Jawaharlal Nehru Krishi Vishwavidyalaya, Jabalpur 482004, India

S. Kumar (✉)
AVRDC—The World Vegetable Center, P.O. Box 42, Shanhua Tainan 74199, Taiwan
e-mail: sanjeet.kumar@worldveg.org

Introduction

Pepper belongs to the genus *Capsicum*, which includes about 25 wild species and five (*Capsicum annuum* L., *Capsicum frutescens* L., *Capsicum chinense* Jacq., *Capsicum baccatum* L. and *Capsicum pubescens* Ruiz and Pavon) cultivated species (Kumar et al. 2006a). *C. annuum* species bearing pungent (chilli, chili or hot pepper) and non-pungent (sweet pepper) fruits are the most widely cultivated worldwide. Besides being one of the most commonly used spices, condiments and vegetables, peppers have several versatile and innovative food and non-food uses (Kumar et al. 2006a). Chilli is an important cash crop for small and marginal farmers in many developing countries across Asia (China, India, Pakistan, Bangladesh, Thailand and Indonesia) and Africa (Egypt, Ethiopia, Nigeria and Ghana). India is the largest producer of dry chilli fruits, accounting for more than 43 % of the world's total dry chilli production (FAOSTAT 2011).

Diversity analysis among the working collections of germplasm including elite lines complements and enhances the efficacy of breeding and germplasm management. *Capsicum* genetic diversity has been analysed using restriction fragment length polymorphism (Lefebvre et al. 1993),

amplified fragment length polymorphism (Aktas et al. 2009), random amplified polymorphic DNA (RAPD; Adetula 2006), microsatellites or simple sequence repeats (SSRs; Portis et al. 2007; Stágel et al. 2009; Pacheco-Olvera et al. 2012) and direct amplification of minisatellite DNA (DAMD-PCR; Ince et al. 2009). The SSRs are efficient and useful for diversity analysis, establishing linkage, mapping quantitative traits and marker assisted selection for useful traits (Varshney et al. 2005). SSRs are useful markers, as they are relatively abundant and evenly distributed over the genome, but lack of polymorphism of SSRs between genotypes may limit their application and also availability of SSR markers for pepper in public domain is comparatively lesser. Therefore, additionally to SSR markers, we examined the application of random amplified microsatellite polymorphism (RAMPO) markers for diversity analysis in the genus *Capsicum*. The other objectives of the study were to elucidate relationships among a set of *Capsicum* genotypes, and to assess the transferability of SSR markers between *Capsicum* species.

Materials and methods

Plant materials

Plant material for this study comprised of 48 genotypes selected on the basis of their origin (geographical and pedigree), cultivar type (improved or landrace), pungency, reaction to pepper leaf curl virus (Pep-LCV), male sterile cytoplasm, CMS (restorer or maintainer), as well as fruit shape and size (Table 1 and Fig. 1). The list comprised 38 *C. annuum*, three *C. frutescens*, one *C. chinense* and *C. baccatum* genotypes, and five natural interspecific hybrid derivative landraces from northeast India (Table 1).

DNA isolation and quantification

Leaf samples from five random plants of each genotype were collected and total genomic DNA was extracted following Doyle and Doyle (1987) with some modifications. Total genomic DNA was quantified using nanodrop (Thermo Scientific, Wilmington, DE, USA). For quality determination, DNA samples were electrophoresed on 0.8 % agarose (Bangalore Genei, India) gels stained with ethidium bromide and adjusted to 35 ng/μl final concentration.

SSR analysis

One hundred two genomic SSR primer pairs designed for *Capsicum* (Minamiyama et al. 2006) were synthesised (Integrated DNA Technologies, CA, USA). The SSR markers were screened for amplification and polymorphism using six genotypes, namely, Bhut Jolokia, BS-35, GKC-29 (all interspecific hybrid derivative landraces from northeast India; Table 1), Kashi Sinduri, Kashi Anmol and Pusa Jwala (all improved *C. annuum*, Table 1). Finally, based on polymorphism and reproducibility, 25 SSRs (Table 2) were selected and used across all 48 genotypes. The polymerase chain reaction (PCR) mixture (15 μl) consisted of 35 ng genomic DNA, 10 pM each of forward and reverse primers, 100 μM dNTPs (Bangalore Genei, Bangalore, India), 1.5 mM of $MgCl_2$, 1× PCR buffer [10 mM Tris–HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$] and 0.75 U of *Taq* DNA polymerase (New England Biolabs, Ipswich, MA, USA). Amplifications were performed in a thermal cycler (Bio-Rad Laboratories, Hercules, CA, USA) with a profile of 94 °C for 5 min followed by 30 cycles of 30 s at 94 °C, 30 s at the annealing temperatures from 45 to 56 °C for individual SSR primers, and 30 s at 72 °C with a final extension for 5 min. The size of products were determined on 3.0 % Metaphor agarose (Lonza, USA) gels electrophoresis prepared in 1× TAE buffer [40.0 mM Tris-base, 16.65 M acetic acid, 0.5 M EDTA (pH 8.0)] at a constant voltage of 65 V for 3 h. Thereafter, gels were stained with ethidium bromide (0.5 μg/ml) and visualised and photographed in a gel documentation system (Alpha Imager 3400, Alpha Innotech Corporation, CA, USA).

RAMPO analysis

RAPD/PCR was performed and PCR products were used as template for inter-simple sequence repeats PCR (ISSR/PCR). For this purpose, four polymorphic RAPD primers, namely, OPG 19 (5′ GTCAGGGCAA 3′), OPH 11 (5′ CTTCCGCAGT 3′), OPJ 17 (5′ ACGCCAGTTC 3′) and OPL 19 (5′ GAGTGGTGAC 3′) were selected, as these were informative primers for hybrid purity testing and pepper leaf curl resistance (Rai 2010). Initially 50 ISSR primers were screened using two diverse genotypes (BS-35, interspecific derivatives and Kashi Anmol, *C. annuum*) and finally the five most polymorphic ISSRs, namely, BV 04 [5′ $(CT)_8$AC 3′], BV11 [5′ $(CT)_8$AT 3′], BV17 [5′ $(CA)_6$GT 3′], BV35 [5′ $(GT)_6$CC 3′] and BV 38 [5′ $(CAC)_3$GC 3′] were selected to develop 20 RAMPO combinations (4 RAPD primers×5 ISSR primers). RAPD/PCR reactions were performed in a 25 μl reaction mixture containing 35 ng DNA, 10 pM primer, 100 μM dNTPs, 2.5 μl 10× PCR buffer and 1 U *Taq* DNA polymerase. Afterwards, ISSR/PCR amplifications were performed in 25 μl reaction mixture containing 3 μl RAPD/PCR products, 10 pM ISSR primer, 100 μM dNTPs, 2.5 μl 10× PCR buffer and 1 U *Taq* DNA polymerase. The PCR profile for RAPD and ISSR was 94 °C for 5 min followed by 35 cycles of 94 °C for 1 min, 45 s at corresponding annealing temperatures, and 72 °C for 1 min, followed by a final extension at 72 °C for

Table 1 List and salient features of *Capsicum* genotypes used for the diversity analysis

Series no.	Genotype (sp.)	Origin	Salient known features (i.e. degree of pungency[a], reaction to pepper leaf curl[b], fertility restoration reaction to male sterile cytoplasm[c])
1	Bhut Jolokia (*f*×*c*)	India	Inter-specific derivative landrace from northeast, among hottest known, highly resistant to leaf curl, weak restorer
2	BS-35 (*f*×*c*)	India	Inter-specific derivative landrace from northeast, highly pungent, leaf curl symptomless, weak restorer
3	C00309 (*f*)	Taiwan	Highly pungent, weak restorer, highly resistant to leaf curl
4	Lankamura Collection (*f*×*b*)	India	Inter-specific derivative landrace from northeast, highly pungent, leaf curl resistant, weak restorer
5	C05635 (*b*)	Brazil	Non-pungent, susceptible to leaf curl
6	Punjab Lal (*a*)	India	Pungent, resistant to leaf curl, restorer
7	GKC-29 (*a*×*f*)	India	Inter-specific derivative landrace from northeast, highly pungent, leaf curl symptomless, reproductively isolated
8	IC-383072 (*f*)	India	Highly pungent, highly resistant to leaf curl, weak restorer
9	C00304 (*c*)	USA	Mild pungent, highly resistant to leaf curl
10	NMCA-40008 (*f*)	USA	Highly pungent Tabasco, resistant to leaf curl, weak restorer
11	NMCA-50003 (*a*)	USA	Sweet pepper, moderately susceptible to leaf curl, maintainer
12	AMK-11 (*a*)	India	Pungent, susceptible to leaf curl, restorer
13	ISC-9 (*a*)	Israel	Pungent, moderately susceptible to leaf curl, weak restorer
14	EC-519636 (*a*)	Unknown	Mild pungent, susceptible to leaf curl, restorer
15	Sel-4 (*a*)	India	Sweet pepper, susceptible to leaf curl, maintainer
16	Solan Local (*a*)	India	Improved sweet pepper developed in India, susceptible, maintainer
17	AKC-89/38 (*a*)	India	Improved variety, round and pungent fruits, moderately susceptible to leaf curl, restorer
18	DSL-2 (*a*)	India	Pungent, moderately susceptible to leaf curl, restorer
19	VR-339 (*a*)	India	Improved line, pungent, resistant to leaf curl, restorer
20	0337–7545 (*a*)	Taiwan	Pungent, moderately susceptible to leaf curl, restorer
21	Kashi Sinduri (*a*)	Indonesia	Suitable for oleoresin extraction, non-pungent, susceptible to leaf curl, weak restorer
22	Kashi Gaurav (*a*)	India	Improved line, mild pungent, susceptible, restorer
23	JapaniLongi (*a*)	India	Improved population (syn. Pusa Sadabahar), pungent, moderately susceptible to leaf curl (MS)
24	Taiwan-2 (*a*)	Taiwan	Highly pungent, moderately resistant to leaf curl, restorer
25	KTPL-19 (*a*)	Spain	Non-pungent, suitable for oleoresin extraction, moderately resistant to leaf curl, partial restorer
26	NuMex Twilight (*a*)	USA	Ornamental, mild pungent, susceptible to leaf curl, restorer
27	NuMex Centennial(*a*)	USA	Ornamental, mild pungent, susceptible to leaf curl, restorer
28	DC-16 (*a*)	India	Collection from north east, purple and pungent fruits, moderately susceptible to leaf curl, restorer
29	CM-334 (*a*)	Mexico	Landrace, pungent, resistant to *Phytophthora* and nematodes; moderately susceptible to leaf curl, restorer
30	Pant C-1 (*a*)	India	Improved variety, pungent, moderately susceptible to leaf curl, restorer
31	Pusa Jwala (*a*)	India	Improved variety, pungent, susceptible to leaf curl, restorer
32	PKM-51 (*a*)	India	Improved variety, pungent, moderately resistant to leaf curl, restorer
33	PDC-50 (*a*)	India	Pungent, moderately susceptible to leaf curl, restorer
34	LCA-235 (*a*)	India	Improved variety, pungent, moderately susceptible to leaf curl, restorer
35	California Wonder (*a*)	USA	Sweet pepper, highly susceptible to leaf curl, maintainer
36	Local Tripura (*a*×*f*)	India	Interspecific derivative landrace from north east, pungent, moderately susceptible to leaf curl, restorer
37	TC-6903 (*a*)	Mexico	Non-pungent, susceptible to leaf curl, maintainer
38	EC-119457 (*a*)	Unknown	Non-pungent, moderately resistant to leaf curl, restorer
39	MS-12 (*a*)	India	Nuclear genic male sterile line, pungent, moderately resistant to leaf curl
40	KDCS-810 (*a*)	India	Improved variety, pungent, moderately resistant to leaf curl, restorer
41	EC-458206 (*a*)	Unknown	Sweet pepper, highly susceptible to leaf curl, maintainer

Table 1 (continued)

Series no.	Genotype (sp.)	Origin	Salient known features (i.e. degree of pungency[a], reaction to pepper leaf curl[b], fertility restoration reaction to male sterile cytoplasm[c])
42	NCCH-143 (*a*)	India	Sweet pepper hybrid, moderately susceptible to leaf curl
43	No. 8 (*a*)	India	Sweet pepper, susceptible to leaf curl, maintainer
44	CCA-4261 (*a*)	Taiwan	Pungent, susceptible to leaf curl, cytoplasmic male sterile (CMS)
45	PBC-473 (*a*)	Taiwan	Pungent, susceptible to leaf curl, restorer of CMS
46	AC Assam(*a*)	India	Collection from northeast, pungent, moderately susceptible to leaf curl, restorer
47	Kashi Anmol (*a*)	Sri Lanka	Improved variety reselected in India, pungent, moderately susceptible to leaf curl, restorer
48	9955-15 (*a*)	Taiwan	Non-pungent, moderately resistant to leaf curl, restorer

a annuum, *b baccatum*, *c chinense*, *f frutescens*

[a] Pandey (2006) and unpublished data

[b] Kumar et al. (2011) and Rai (2010)

[c] Kumar et al. (2007, 2009) and unpublished data

10 min. The PCR products were separated using 1.5 % agarose gel in 1× TAE buffer (pH 8.3). The instruments and other conditions were same as above. Amplifications and separation of amplified products were carried out twice to test reproducibility.

Data analysis

For SSR and RAMPO analyses, for each primer pair, only reproducible bands with high intensity were scored either as present (1) or absent (0). Polymorphic information content (PIC) of each marker was calculated according to Nei's statistic (Nei 1973) as following:

$$PIC = 1 - \sum (p_I)^2$$

Where p_I is the frequency of the *I*th allele at the locus. For each marker, three PIC were calculated, for all 48 genotypes (PIC), for 38 *C. annuum* genotypes (PIC_a) and for 10 non-*annuum* genotypes (PIC_n). Cluster analysis was performed using NTSYSpc version 2.1 (Rohlf 1998). Pairwise combinations of genotypes were employed to calculate

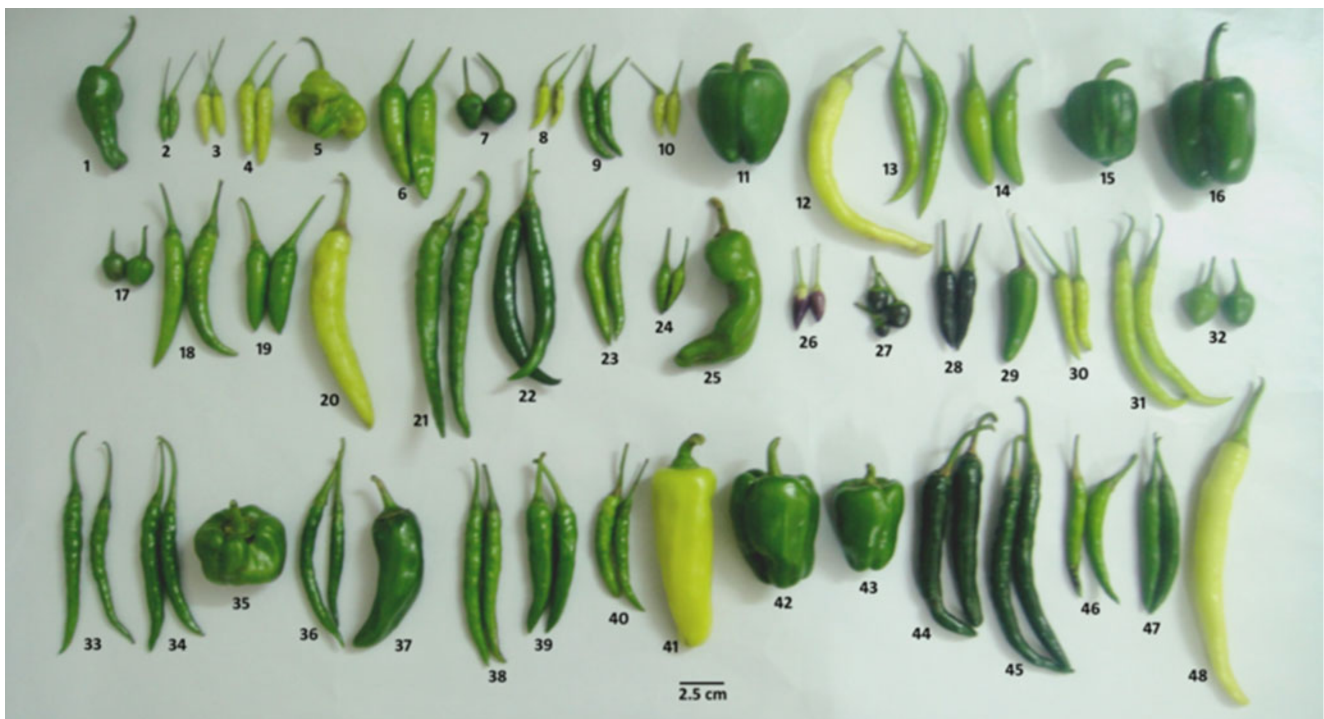


Fig. 1 Fruits size, shape and colour (immature) of 48 *Capsicum* genotypes, numbers (*1* to *48*) corresponds to serial numbers and names in Table 1

Table 2 List of polymorphic SSR primers, their linkage group, allele size range, alleles per locus and polymorphism information content among all (PIC), *C. annuum* (PCA_a) and non-*annuum* (PIC_n) genotypes

Series no.	SSR name	Repeat motif	Linkage group	Allele size range (bp)	Alleles per locus	PIC	PIC_a	PIC_n
1	CAMS-63-2	(tc)6…(ac)4…(tc)4(ac)6…(tc)3ct(tc)3	2	200–255	2	0.73	0.77	0.00
2	CAMS-72	(ac)13	5	150–185	3	0.69	0.46	0.86
3	CAMS-89	(tc)19	11	210–255	3	0.77	0.62	0.85
4	CAMS-90	(ca)3a(ac)15	8	225–245	3	0.51	0.08	0.82
5	CAMS-101	(tg)4ta(tg)10	6	210–240	2	0.66	0.00	0.00
6	CAMS-117	(tg)21(ta)3	11	210–245	4	0.90	0.82	0.90
7	CAMS-142	(ta)3…(ac)7…(ac)12a(ta)8	7	230–265	3	0.41	0.08	0.51
8	CAMS-163	(at)7(gt)14	5	240–270	4	0.77	0.70	0.87
9	CAMS-311	(ga)3g(ga)4g(ga)7g(ga)3	6	225–245	2	0.52	0.48	0.66
10	CAMS-313	(ag)10…(ac)3…(ta)3	11	215–240	2	0.44	0.38	0.52
11	CAMS-327	(tc)7	13	235–255	2	0.61	0.53	0.70
12	CAMS-348	(ag)11	3	200–250	5	0.69	0.05	0.76
13	CAMS-351	(tg)3…(ag)26	4	190–240	3	0.84	0.73	0.59
14	CAMS-378	(tc)6…(tc)4…(ct)3	1	155–175	3	0.65	0.38	0.83
15	CAMS-424	(ag)16	6	180–235	3	0.81	0.74	0.82
16	CAMS-644	(tg)3…(ag)26	4	205–230	2	0.76	0.73	0.68
17	CAMS-647	(tat)6tg(tta)3…(tat)21	3	220–245	3	0.86	0.82	0.91
18	CAMS-811	(aag)3…(gaa)3…(gaa)7	9	250–270	2	0.75	0.65	0.68
19	CAMS-826	(gaa)6ga(gga)	8	230–285	5	0.82	0.68	0.88
20	CAMS-844	(gaa)6	1	220–230	2	0.29	0.22	0.42
21	CAMS-846	(tct)10	3	190–230	5	0.92	0.83	0.93
22	CAMS-855	(agt)14a(gaa)9	8	240–265	2	0.54	0.00	0.66
23	CAMS-864	(aga)32	7	220–245	4	0.83	0.80	0.80
24	CAMS-876	(cct)3tc(tct)4…(tct)7	9	240–255	2	0.67	0.05	0.59
25	CAMS-885	(gaa)28	2	240–265	5	0.82	0.62	0.81
	Average				3.04	0.69	0.49	0.68

Jaccard's similarity coefficient (GS; Jaccard 1908). The similarity matrix was used to generate an unweighted pair group method with arithmetic averages (UPGMA)-based dendrogram (Sneath and Sokal 1973) using the *sequential agglomerative hierarical nested* cluster analysis (SAHN) module of NTSYSpc. The average similarity coefficient was used as cut-off value for defining number of clusters. To investigate the correlation between the SSR and RAMPO datasets, the Mantel matrix correspondence coefficient was estimated, based on their distance matrices (Mantel 1967) using the MXCOMP module. The binary data was subjected to principal component analysis (PCA) using the EIGEN and PROJ modules of NTSYSpc.

Population structure analysis

A Bayesian-based cluster analysis was performed using SSR and RAMPO data with STRUCTURE version 2.3.1 (Pritchard et al. 2000). This method uses Markov chain Monte Carlo methods (MCMC) to estimate allele frequencies and to identify the optimum number of population (K) subgroups. Analyses were performed using admixture model assumptions with correlated alleles, K was presumed to be 2–10, selected after five independent runs. Each run consisted of a burn-in period of 50,000 steps followed by 100,000 MCMC replicates (Pritchard and Wen 2003). STRUCTURE HARVESTER software (Earl and vonHoldt 2012) was used to collate the results obtained from STRUCTURE following Evanno et al. (2005) and maximum value of ΔK associated with each K value were analysed to identify the number of clusters that best described the data.

Results

SSR and RAMPO analyses

From 102 SSR primers, 25 were finally selected based on their ability to reproduce polymorphic fragments (Table 2).

A total of 76 alleles were detected; the number of alleles at each locus varied from two to five with an average of 3.04 alleles. The size of the amplified products ranged from ~150 to ~270 bp. Among the 25 SSR loci used, PIC ranged from 0.29 (for CAMS-844) to 0.92 (for CAMS-846) with an average of 0.69 (Table 2). Among *C. annuum* genotypes, the average number of alleles per locus was 2.32 with an average PIC of 0.49, whereas, among non-*annuum* genotypes, the average number of alleles per locus was 2.8 with an average PIC of 0.68.

A total of 20 RAMPO primer combinations were used for the first time in *Capsicum*. Of these, 17 generated 106 scorable and reproducible fragments (Table 3). The fragment size ranged from ~300 to ~2,500 bp and among 48 genotypes, 87 fragments out of 106 were polymorphic. The number of fragments per primer combination ranged from three (OPL 19–BV 35) to nine (OPJ 17–BV 17 and OPL 19–BV 17) with an average of five fragments per primer combination. The PIC ranged from 0.44 (OPL 19–BV 04) to 0.81 (OPH 11–BV 11) with an average of 0.63 (Table 3). The RAMPO primer combinations generated an average of 5.12 alleles per locus with an average PIC of 0.48 in *annuum* genotypes, while an average of 5.88 alleles per locus with PIC of 0.66 were obtained with non-*annuum* genotypes.

Cluster and principal component analyses

Dendrogram obtained by UPGMA-based cluster analysis using SSR markers revealed average GS of 0.58, which ranged between 0.26 and 0.89 (Fig. 2a). All the 48 genotypes were grouped into two major clusters. The first cluster consisted exclusively of non-*annuum* genotypes such as interspecific derivative landraces (Bhut Jolokia, BS-35, Lankamura Collection and Local Tripura), *C. frutescens* genotypes (C00309, NMCA-40008 and IC-383072), *C. chinense* (C00304) and *C. baccatum* (C05635) genotypes. However, C05635, the only *C. baccatum*, was distinctly separated from the remaining eight genotypes within same cluster. The second cluster comprised of remaining 38 genotypes, exclusively belonging to *C. annuum* (Table 1) with the exception of the natural interspecific derivative landrace, GKC-29 (Fig. 2a). GKC-29 was separated from the rest of the *C. annuum* genotypes in the second cluster, indicating its genomic distance to *C. annuum*. Likewise, 0337–7545 (AVPP0304), an improved C. annuum line developed for global use (Keatinge et al. 2012) by Asian Vegetable Research and Development Center, was also found to be distinctly related to other *C. annuum* genotypes (Fig. 2a). There was no grouping observed among *C. annuum* genotypes according to geography, pungency and other traits,

Table 3 List of RAMPO marker combinations (4 RAPDs×5 ISSRs), their amplification details, percentage polymorphism and polymorphic information content among all (PIC), *C. annuum* (PCAa) and non-*annuum* (PICn) genotypes

Primer combination	Number of fragments		Percentage polymorphism	PIC	PIC_a	PIC_n
	Total	Polymorphic				
OPG 19–BV 04	4	4	100	0.69	0.47	0.84
OPG 19–BV 11	8	6	75	0.72	0.70	0.78
OPG 19–BV 17	7	7	100	0.60	0.43	0.79
OPG 19–BV 35	6	5	83.33	0.59	0.41	0.70
OPG 19–BV 38	5	4	80	0.61	0.49	0.46
OPH 11–BV 04	4	4	100	0.69	0.47	0.80
OPH 11–BV 11	7	6	85.71	0.81	0.77	0.90
OPH 11–BV 17	Smear	–	–			
OPH 11–BV 35	6	4	66.67	0.59	0.42	0.70
OPH 11–BV 38	Smear	-	-			
OPJ 17–BV 04	5	5	100	0.72	0.50	0.64
OPJ 17–BV 11	7	5	71.42	0.69	0.65	0.80
OPJ 17–BV 17	9	7	77.78	0.64	0.54	0.40
OPJ 17–BV 35	5	5	100	0.61	0.49	0.46
OPJ 17–BV 38	7	5	71.42	0.64	0.51	0.74
OPL 19–BV 04	6	4	66.67	0.44	0.28	0.52
OPL 19–BV 11	Smear	–	–			
OPL 19–BV 17	9	6	66.67	0.55	0.32	0.50
OPL 19–BV 35	3	3	100	0.48	0.38	0.69
OPL 19–BV 38	8	7	87.5	0.60	0.35	0.53
Total	106	87	71.60	0.63	0.48	0.66

Fig. 2 UPGMA-based dendrogram showing genetic relationship among 48 *Capsicum* genotypes based on SSR (**a**) and RAMPO (**b**) markers

although pungent and non-pungent genotypes of *C. annuum* are grouped in small subclusters; e.g. SSR markers were able to differentiate closely related genotypes CCA-4261 (CMS) from PBC-473 (restorer), although within the same subcluster.

Cluster analyses were successfully performed using permutations of 25 SSR markers. Twelve polymorphic SSRs (CAMS-142, CAMS-644, CAMS-647, CAMS-327, CAMS-351, CAMS-89, CAMS-117, CAMS-424, CAMS-

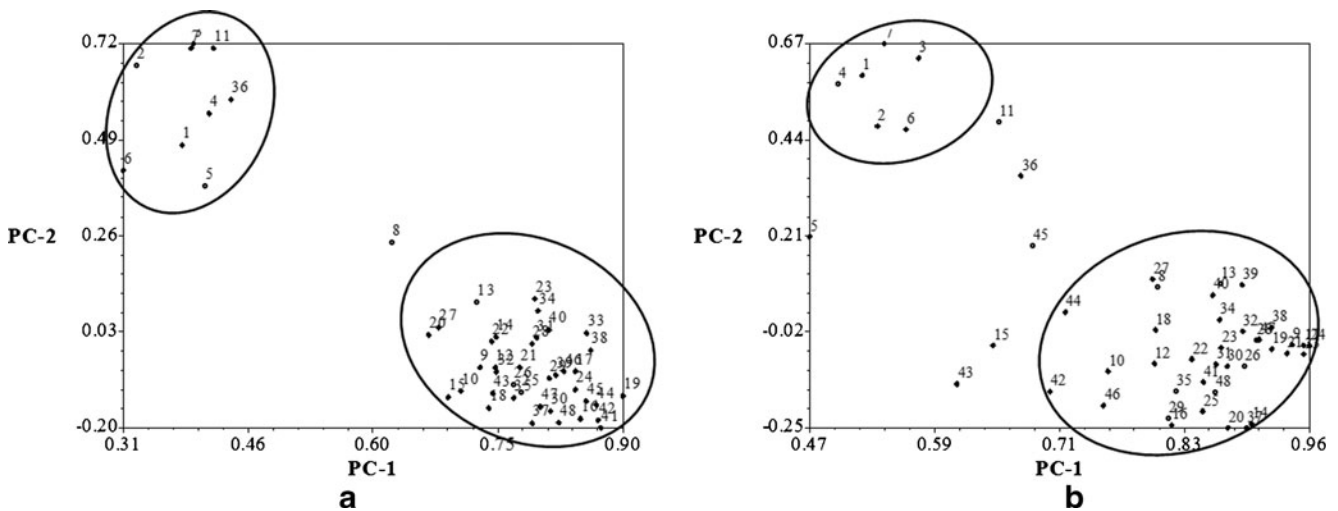


Fig. 3 PCA plot of the first two principal components showing the spatial distribution of 48 *Capsicum* genotypes based on SSR (**a**) RAMPO (**b**) markers

876, CAMS-811, CAMS-63-2 and CAMS-844) were found to separate 48 genotypes in a similar pattern as with 25 primers (figure not shown).

The UPGMA-based cluster analysis using RAMPO markers grouped 48 genotypes into three clusters with genetic similarity ranging from 0.37 to 0.98 (Fig. 2b). C05635 (*C. baccatum*) was separated from all other genotypes in a separate cluster. Two other clusters consisted of nine non-*annuum* (first cluster) and 38 *C. annuum* (second cluster) genotypes, respectively. In the first cluster, BS-35 ($f \times c$) was most divergent, while C00309 and IC-383072 were most closely related (GS>0.73). In the second cluster, the genotype no. 8 was most divergent. Unlike the SSR-based dendrogram, GKC-29 grouped with *C. annuum* genotypes with hot peppers in a subcluster, but stood separately. The Mantel matrix correspondence also revealed significant correlation (r=0.73) between the two matrices of SSR and RAMPO. Thus, besides few exceptions, there was strong positive correlation for the results of the analyses of genetic relationships among *Capsicum* genotypes through both marker systems.

PCA using SSR and RAMPO markers also clustered the genotypes into two clusters (Fig. 3). The eigenvalues obtained from the SSRs revealed that the first three principal components cumulatively accounted for 64.97 % of the total variation, in which 53.47 % was accounted for by component 1 and 7.7 % by component 2 (Fig. 3a). The *C. annuum* genotypes formed a distinct cluster from genotypes belonging to the other species and interspecific derivative landraces. PCA using the RAMPO data (Fig. 3b) showed that the eigenvalues of the first three components cumulatively accounted for 75.94 % of the total variation with 64.2 and 6.32 % accounted for by components 1 and 2, respectively. Like SSRs, PCA of RAMPOs also divided the 48 genotypes into two distinct clusters.

Population structure

SSR and RAMPO markers data, separately and combined, were used to estimate the exact number of subpopulations on the basis of the LnP(D) value. Initially, two groups were formed at K=2, which corresponds to *annuum* and non-*annuum* namely OI and OII, respectively (Fig. 4). Further subgroups were formed at K=3, K=4 and K=5 to determine whether any intermediate groups were formed. To collate the results from STRUCTURE, the exact number of population was assessed using STRUCTURE HARVESTER. The maximum ΔK value for both marker systems and for combined data was observed for K=2 (Fig. 5). The LnP(K) value for SSR (−2,516.85), RAMPO (−3,039.16) and combined data (−5,609.85) were also highest at K=2. The genotypes from the northeast region of India clustered into first group, while *C. annuum* genotypes clustered into second group (Fig. 4). Of the 48 genotypes, 38 genotypes were clearly clustered in one of the two groups, while 10 genotypes (IC-383072, NMCA-40008, NMCA-50003, Sel-4, Numex Centennial, Local Tripura, NCCH-143, No. 8, CCA-4261 and PBC-535) showed an admixed population of origin, with less than 80 % of inferred ancestry at K=2. The exact membership proportion of K=2 is provided in Electronic supplementary material (ESM) Table S1.

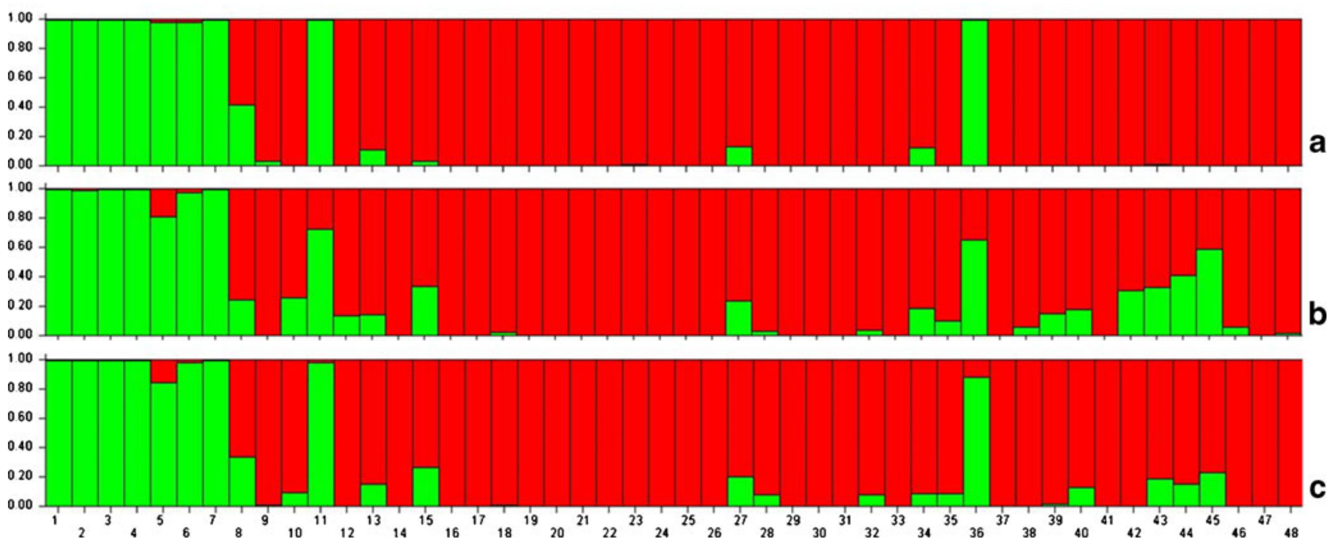


Fig. 4 The population structure bar plots generated by STRUCTURE software at K=2. 48 genotypes of *Capsicum* (*1–48* corresponds to Table 1) grouped into two genetic groups or subpopulations using SSR (**a**), RAMPO (**b**) and combined (**c**) data. Each *solid bar* represents single genotype and two different colours (*red* and *green*) represent genotypes belonging to two different subpopulations. The proportions of the colour bars represent the admixtures in the varieties. The *Y*-axis shows the estimated ancestry of each genotype from a particular subpopulation (colour figure online)

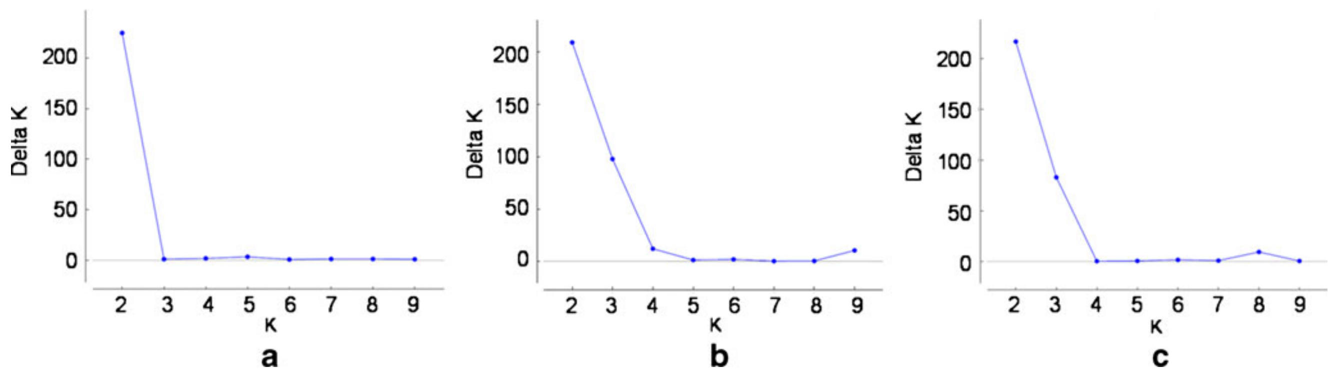


Fig. 5 Comparison of ΔK values for *Capsicum* genotypes using SSR (**a**), RAMPO (**b**) and combined data (**c**); the maximum value of ΔK at $K=2$ was considered to be the appropriate number of population

Discussion

Cluster analyses: SSR vs. RAMPO

SSR markers in *Capsicum* are not in abundance in public domain compared with crops like tomato, rice, corn, etc. Therefore, for the first time, we used RAMPO markers along with SSRs to analyse diversity in *Capsicum*. Both SSR and RAMPO markers produced similar results and differentiated all the 48 genotypes broadly into *annuum* and non-*annuum* groups (Fig. 2). This was expected as both SSR and RAMPO detect differences at simple sequence repeat motifs, although SSR depicts the variation in the number of repeats while RAMPO depicts the variations in presence or absence of SSR sequences within an RAPD-amplified region. In the case of SSRs, except GKC-29, all eight non-*annuum* genotypes clustered separately from all 38 *C. annuum* genotypes. The clustering pattern obtained by RAMPO markers was similar to that of SSR markers; however, C05635 (*C. baccatum*) in this case, stood alone as a separate cluster (first cluster). This result is in agreement with the results of DAMD-PCR marker (Ince et al. 2009) where *C. baccatum* accessions did not cluster with the accessions of *C. annuum* and *C. pubesense* species complexes. In the second cluster, all 38 *C. annuum* genotypes grouped together along with an interspecific derivative (GKC-29) and the remaining non-*C. annuum* genotypes were grouped separately in a third cluster (Fig. 2b). The PCA using SSR and RAMPO data also clustered the genotypes into two groups that support the grouping shown by both dendrograms. The Mantel matrix correspondence test indicated the RAMPO data set was highly consistent with the SSR data set. The significant positive correlations indicated that RAMPO can provide a useful measure of genetic diversity among *Capsicum* genotypes as SSR markers. Although SSR markers showed superiority over RAMPO markers, RAMPO markers may be useful in those crop species where SSR markers are not available in abundance.

The genetic variation detected between 48 genotypes was found to be greater with RAMPO markers than with SSR markers. In both marker systems, clustering of genotypes according to geographical origin, fruit pungency, size and shape could not be observed. However, genetic relationships between analysed genotypes were in agreement with their known taxonomic classification in both marker systems. Hence, diversity analysis results obtained from 25 SSRs were equivalent to the 17 RAMPO markers. However, we have demonstrated that out of 25 SSRs, only 12 SSRs can give similar taxonomic and diversity results, especially with respect to clustering patterns. Although SSRs are more robust and easier to use, successful use of RAMPO markers has validated their usefulness in phylogenetic studies in *Capsicum*.

Genetic diversity and relationship among the tested genotypes

Both SSR- and RAMPO-based UPGMA dendrograms clearly separated the *C. annuum* genotypes from the non-*annuum* genotypes (*C. frutescens*, *C. chinense*, *C. baccatum* and interspecific landraces) into two clusters (Fig. 2). However, *C. baccatum* accession C05675 distinctly separated from all the other accessions belonging to the *C. annuum* species complex (i.e. *C. frutescens*, *C. chinense* or their interspecific derivatives/landraces; Fig. 2). Clustering of these genotypes in subgroups (with GS more than 60) was correlated with species differentiation between four cultivated species and their landraces derived through natural interspecific hybridization. With both marker systems, genetic variation within genotypes belonging to interspecific derivative landraces and three less commonly cultivated species (*C. frutescens*, *C. chinense* and *C. baccatum*) was significantly higher than genetic variation within most commonly cultivated *C. annuum* genotypes (Fig. 2), although sample size of non-*annuum* genotype was smaller than the *C. annuum*. Genetic diversity between commonly grown

improved *C. annuum* genotypes has been found to be less than the diversity between semi-wild and landrace genotypes (Oyama et al. 2006). This is expected, as during and after domestication nearly all domesticated crop species have gone through a decline in genetic diversity (Gepts 2004). The frequent use of selected elite breeding lines in commercial breeding worldwide has further narrowed genetic diversity in many crop plant species. This genetic bottleneck is also evident within *C. annuum* species (Ince et al. 2009), where diversity (molecular marker based) between more commonly bred and commercialised non-pungent (sweet pepper) genotypes has been found to be narrower than the diversity between pungent (chilli) genotypes (Ortiz et al. 2010). The magnitude of the observed genetic bottleneck, however, depends on the type of marker (molecular or phenotypic) used to measure genetic diversity.

The utility of both marker systems to differentiate individual accessions within *C. annuum* also became apparent, as markers could separate even isogenic lines and many genetically closely related accessions within the same subcluster. For example, CCA-4261 (CMS) and PBC-473 (restorer) (isogenics and isoplasmic lines) were differentiated with both SSR and RAMPO markers (Fig. 2). Likewise, Kashi Anmol, a commercial Sri Lankan variety (KA-2) reselected and released in India, shares Pant C-1 in its parentage (Table 1) and was separated apart from Pant C-1, although within the same cluster.

Cross-species transferability and unique alleles

The genomic SSRs used in the study were primarily designed and tested for *C. annuum* (Minamiyama et al. 2006). Out of 102 SSRs, only 25 (24 %) were polymorphic between six *Capsicum* genotypes, indicating low SSR polymorphism rate in the germplasm. Flanking regions of SSRs are highly conserved and usually display a high level of transferability across the species (Varshney et al. 2005). Therefore, the markers could successfully be applied also on other species of the *Capsicum* genus. Out of 25 SSRs, CAMS-89, CAMS-117, CAMS-163, CAMS-351, CAMS-424, CAMS-644, CAMS-647, CAMS-811, CAMS-826, CAMS-846 and CAMS-885 showed comparatively high PIC in both groups, and thus could be used for mapping and MAS in inter- and intra-specific crosses. Interestingly, 24 % SSRs were polymorphic, which is highly comparable and consistent with previous reports in *Capsicum* (24–26 %) regardless of the use of genotypes (intra- vs. inter-specific) and SSR types (EST/gene vs. genomic; Minamiyama et al. 2006; Portis et al. 2007). In this study, we also identified useful informative markers that could facilitate our ongoing breeding efforts. For instance, CAMS-89 (~200 bp) and CAMS-424 (~210 bp) were specific to the Pep-LCV resistant genotype (Bhut Jolokia) and were used to map Pep-LCV locus in F_2 population derived from Bhut Jolokia× PBC-535 (Rai 2010).

Population structure

The present study exhibits a population structure with two clusters which corresponds to *annuum* and non-*annuum* populations, and the overall membership proportion explains the admixture present in *C. annuum* and interspecific derivatives. The admixture is the representation of diverse parents, which themselves have diverse ancestry in breeding history and domestication, the main reason for variation present in the population. The maximum ΔK value was observed for $K=2$ and further groups appeared at $K=3$, $K=4$ and $K=5$ which indicated some extent of subgrouping in each group as was also evidenced by UPGMA cluster analysis. The two clusters based on genetic structure and UPGMA analysis with molecular markers for all the genotypes correspond mostly with one another with respect to their pedigree relationships, but the genetic structure explains the relationships better because of the higher degree of simulation (the exact membership proportion of $K=2$ is provided in ESM Table S1). Earl and vonHoldt (2012) described that grouping based on larger value of ΔK could described number of subpopulations best fitted with the data rather than the higher LnP(K) value. But in this study, LnP(K) value were also found larger at $K=2$ than at $K=3$–10 for SSR, RAMPO and combined data, which further supports the grouping at $K=2$. Recently, population structure of *Capsicum* germplasm has been reported with genetic subpopulations ranging from three to six (Albrecht et al. 2012).

Existence of unique diversity

The northeast region of India has long been known as a hot spot of hot pepper biodiversity (Purkayastha et al. 2012). This region has received renewed attention in the past decade due to the discovery of India's hottest pepper (Naga Jolokia) and later a variant of this landrace (Bhut Jolokia) was reported to be the world's hottest pepper, and was taxonomically assigned to a putative interspecific hybrid between *C. chinense* and *C. frutescens* (Bosland and Baral 2007). Many highly pungent genotypes are found in the region (Sanatombi et al. 2010). Because all five landraces with different fruit morphology (Fig. 1) and parental species (Table 1) were found to be genetically distinct from all the *C. annuum* accessions from within and outside the region (Table 1; Fig. 2), these landraces are believed to have originated in the region from sympatric domesticated species in the past; we preclude their chances of introduction from elsewhere. The occurrence and existence of interspecific hybrid derivative landraces within *Capsicum* in

northeast India should be viewed as analogous to interspecific hybrids domesticated in *Phaseolus*, *Gossypium* and *Vigna*, where more than one species has been domesticated in a given genus (Gepts 2004). Our study has shown that besides being highly pungent, these landraces (e.g. BS-35, GKC-29 and Bhut Jolokia; Table 1) are also resistant or symptomless to Pep-LCV when artificially challenged with viruliferous whiteflies (Kumar et al. 2006b; Rai 2010). Reproductive isolation mechanisms, viz., pre-hybridization with GKC-29 and post-hybridization hybrid breakdown with BS-35, which are typical features of speciation has been demonstrated with these landraces when we tried to cross them with *C. annuum* accessions (Rai 2010). Nevertheless, the taxonomic status and reproductive isolation mechanisms operating between these interspecific derivatives and *C. annuum* species need to be studied in greater detail. Local communities have played a very important role in the evolution of these landraces by maintaining them under in situ conditions. In the post-Colombian era of *Capsiucm* evolutionary studies, for the second time, very recently, a naturally occurring allotetraploid pepper has been discovered in the region (Jha et al. 2012). It would be worthwhile to conduct a regional, *Capsicum*-specific germplasm exploration that recognises the role of local communities of East Himalayan region. Such exploration would be very relevant also to ensure conservation of precious pepper genetic resources, before they get extinct due to the recent high investments and efforts on the agricultural intensification in the region.

Acknowledgments Financial support provided by Indian Council of Agricultural Research (ICAR), New Delhi, India is gratefully acknowledged. The authors also wish to thank AVRDC–The World Vegetable Canter, Taiwan for several germplasm lines used in this study.

References

Adetula OA (2006) Genetic diversity of *Capsicum* using random amplified polymorphic DNAs. Afr J Biotech 5:120–122

Aktas H, Abak K, Sensoy S (2009) Genetic diversity in some Turkish pepper (*Capsicum annuum* L.) genotypes revealed by AFLP analyses. Afr J Biotech 8:4378–4386

Albrecht E, Zhang D, Saftner RA, Stommel JR (2012) Genetic diversity and population structure of *Capsicum baccatum* genetic resources. Genet Resour Crop Evol 59:517–538

Bosland PW, Baral JB (2007) 'Bhut Jolokia'—the world's hottest known chile pepper is a putative naturally occurring inter specific hybrid. HortSci 42:222–224

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

FAOSTAT (2011) http://faostat.fao.org/default.aspx

Gepts P (2004) Crop domestication as a long-term selection experiment. Plant Breed Rev 24:1–44

Ince AG, Karaca M, Onus AN (2009) Development and utilization of diagnostic DAMD-PCR markers for *Capsicum* accessions. Genet Resour Crop Evol 56:211–221

Jaccard P (1908) Nouvellesrecherchessur la distribution florale. Bull Soc Vaud Sci Natur 44:223–270

Jha TB, Dafadar A, Ghorai A (2012) New genetic resource in *Capsicum* L. from Eastern Himalayas. Plant Genet Resour 10:141–144

Keatinge JDH, Ledesma DR, Keatinge FJD, Hughes JDA (2012) Projecting annual air temperature changes to 2025 and beyond: implications for vegetable production worldwide. The J Agricultural Science. doi:10.1017/S0021859612000913

Kumar S, Kumar R, Singh J (2006a) Cayenne/American pepper (*Capsicum species*). In: Peter KV (ed) Handbook of herbs and spices. Woodhead, Cambridge, pp 299–312

Kumar S, Kumar S, Singh M, Singh AK, Rai M (2006b) Identification of host plant resistance to pepper leaf curl virus in chilli (*Capsicum* species). Sci Hortic 110:359–361

Kumar S, Singh V, Singh M, Rai S, Kumar S, Rai SK, Rai M (2007) Genetics and validation of fertility restoration associated RAPD markers in pepper (*Capsicum annuum* L.). Sci Hortic 111:197–202

Kumar R, Kumar S, Dwivedi N, Kumar S, Rai A, Singh M, Yadav DS, Rai M (2009) Validation of SCAR markers, diversity analysis of male sterile cytoplasms (S-) and isolation of an alloplasmic S-cytoplasm in *Capsicum*. Sci Hortic 120:167–172

Kumar S, Kumar R, Kumar S, Singh AK, Singh M, Rai AB, Rai M (2011) Incidence of leaf curl disease on *Capsicum* germplasm under field conditions. Indian J Agric Sci 80:187–189

Lefebvre V, Palloix A, Max R (1993) Nuclear RFLP between pepper cultivars (*Capsicum annuum* L.). Euphytica 71:189–199

Mantel NA (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209–220

Minamiyama Y, Tsuro M, Hirai M (2006) An SSR-based linkage map of *Capsicum annuum*. Mol Breed 18:157–169

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci USA 70:3321–3323

Ortiz RF, De la Flor D, Alvarado G, Crossa J (2010) Classifying vegetable genetic resources-A case study with domesticated *Capsicum* spp. Sci Hortic 126:186–191

Oyama K, Verdugo SH, Sanchez C, Rodrıguez AG, Pena PS, Tiznado JAG, Casas A (2006) Genetic structure of wild and domesticated populations of *Capsicum annuum* (Solanaceae) from northwestern Mexico analyzed by RAPDs. Genet Resour Crop Evol 53:553–562

Pacheco-Olvera A, Hernández-Verdugo S, Rocha-Ramírez V, González-Rodríguez A, Oyama K (2012) Genetic diversity and structure of pepper (*Capsicum annuum* L.) from Northwestern Mexico analyzed by microsatellite markers. Crop Sci 52:231–241

Pandey J (2006) Characterization of *Capsicum* germplasm, inbreds and CMS based crosses with special reference to oleoresin and capsaicin contents. Ph.D. Thesis, Udai Pratap Autonomous College, Varanasi, India

Portis E, Nagy I, Sasvari Z, Stágel A, Barchi L, Lanteri S (2007) The design of *Capsicum* spp. SSR assays via analysis of in silico DNA sequence, and their potential utility for genetic mapping. Plant Sci 172:640–648

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Pritchard JK, Wen W (2003) Documentation for STRUCTURE software: Version 2. Available from http://pritch.bsd.uchicago.edu

Purkayastha J, Alam SI, Gogoi HK, Singh L, Veer V (2012) Molecular characterization of 'Bhut Jolokia' the hottest chilli. J Bioscience 37:757–768

Rai VP (2010) Genetic and molecular analyses of pepper leaf curl resistance in chilli (*Capsicum annuum* L.). Ph.D. Thesis, Banaras Hindu University, Varanasi, India

Rohlf FJ (1998) On applications of geometric morphometrics to studies of ontogeny and phylogeny. System Bio 47:147–158

Sanatombi K, Sen-Mandi S, Sharma GJ (2010) DNA profiling of *Capsicum* landraces of Manipur. Sci Hortic 124:405–408

Sneath PHA, Sokal RR (1973) Numerical taxonomy. Freeman, San Francisco, CA, USA

Stágel A, Gyurjan I, Sasvari Z, Lanteri S, Ganal M, Nagy I (2009) Patterns of molecular evolution of microsatellite loci in pepper (*Capsicum* spp.) revealed by allele sequencing. Plant System Evol 281:251–254

Varshney RK, Graner A, Sorrells ME (2005) Genic microsatellite markers: their characteristics, development and application to plant breeding and genetics. Trends Biotech 23:48–55