

Assessment of the genetic diversity of dessert watermelon (*Citrullus lanatus* var. *lanatus*) landrace collections of South Africa using SSR markers

Jacob Mashilo^{1*}, Hussein Shimelis^{1,2}, Alfred Odindo¹ and Beyene Amelework²

¹University of KwaZulu-Natal, Crop Science Discipline, Private Bag X01, Scottsville 3209, Pietermaritzburg, South Africa

²University of KwaZulu-Natal, African Centre for Crop Improvement (ACCI), Private Bag X01, Scottsville 3209, Pietermaritzburg, South Africa

*Corresponding author: jacobmashilo@yahoo.com

Abstract

Dessert watermelon landraces could provide useful germplasm sources for commercial watermelon breeding and/or conservation. This study assessed the genetic diversity present among dessert watermelon landrace collections of South Africa using simple sequence repeat (SSR) markers. Thirty one diverse dessert watermelon landraces were genotyped using 10 polymorphic SSR markers. A total of 94 alleles were amplified among the sampled population. The number of alleles ranged from 2 to 19 with a mean of 9.4 per locus. The number of effective alleles ranged from 1.07 to 9.94 with a mean of 4.61. Observed heterozygosity values ranged from 0.07 to 1.00 with a mean of 0.53. Expected heterozygosity, which measures gene diversity, was 0.66; which partitioned 75, 25 and 0% of the variation within individuals, among individuals and between populations, respectively. The mean gene fixation was 0.18 suggesting high levels of heterozygosity among the collections. The mean polymorphic information content (PIC) of the SSR loci was 0.65 suggesting their value in genetic diversity analysis of watermelon. The collections were allocated into three genetic groups using cluster analysis. The results revealed the presence of genetic diversity among South African dessert watermelon collections. Genetically unique landraces such as SWM-39, SWM-01, SWM-04, SWM-27, SWM-24, SWM-10 and SWM-36 from cluster I; SWM-35, SWM-21, SWM-02, SWM-34, SWM-07 and SWM-31 from cluster II; and SWM-22 and SWM-18 from cluster III were selected based on their high dissimilarity index. These are recommended for further phenotyping using horticultural attributes for effective breeding and strategic conservation.

Keywords: *Citrullus lanatus*, breeding, genetic diversity, simple sequence repeat markers, landrace.

Abbreviations: CTAB_mixed alkyltrimethyl-ammonium bromide protocol; H_o _observed heterozygosity; H_e _ expected heterozygosity; SSRs_simple sequence repeats; UPGMA_unweighted pair group method using arithmetic averages.

Introduction

Watermelon is a viny creeping crop belonging to the *Cucurbitaceae* family. Chomicki and Renner (2015) using molecular phylogenetic analysis proposed six species under the genus *Citrullus* namely: *C. lanatus*, *C. mucospermus*, *C. colocynthis*, *C. ecirrhosus*, *C. rehmi*, and *C. naudinianus*. These authors further ascribed species *C. amarus* Schrad. to the citron watermelon. The *Citrullus* species are mainly found in the temperate regions of Africa, central Asia and the Mediterranean Region (Levi et al., 2001a; Whitaker and Davis, 1962). The species *C. lanatus* encompasses two botanical varieties: *lanatus* (dessert watermelon) and *citroides* (citron watermelon). The citron watermelon is neither dessert nor bitter and elsewhere known as “cow-melon” or “tsamma” watermelon (Robinson and Decker-Walters, 1997). The southern Africa region is reported to be the center of diversity and probably center of origin of most of the species of *Citrullus* (Dane and Lang, 2004; Robinson and Decker-Walters, 1997; Rubatzky, 2001). Contrastingly, Chomicki and Renner (2015) reported West Africa to be the origin of the egusi-type dessert watermelon.

Dessert watermelon exhibits substantial variation in horticultural attributes such as fruit shapes, flesh colour and skin colour patterns (Maggs-Kölling et al., 2000; Levi et al., 2001; Levi et al., 2004; Ocal et al., 2014). The genetic bases

of the crop have been explored using diverse genetic markers. For example, Lee et al. (1996) and Solmaz et al. (2010) reported low genetic differences among dessert watermelon accessions using random amplification polymorphic DNA (RAPD) markers. Levi et al. (2001) reported high genetic similarities among watermelon cultivars using RAPD markers. Che et al. (2003) and Uluturk et al. (2011) reported low genetic differences among watermelon genotypes using RADP markers and sequence related amplified polymorphism (SRAP) markers, respectively. The reported markers have limitations for variety identification because they are dominant with low discriminatory power (Lee et al., 1996; Che et al., 2003; Kwon et al., 2010). Using simple sequence repeat (SSR) markers, Kwon et al. (2010) reported high genetic diversity among watermelon cultivars. Guerra-Sanz (2002) reported that SSR markers efficiently identified genetic differences of diverse watermelon germplasm. Munisse et al. (2013) also reported a wide variation among dessert watermelon cultivars using SSR markers. Jarret et al. (1997) reported genetic variation among plant introductions (PI) of *C. lanatus* var. *lanatus*, *C. lanatus* var. *citroides* and *C. colocynthis* using SSR markers. Sheng et al. (2012) reported genetic variability amongst Chinese, Japanese, Russian and USA watermelon cultivars using SSR markers.

Simple sequence repeat markers are currently the marker of choice for genetic diversity analysis studies because of their high degree of polymorphism and random distribution across the genome (Varshney et al., 2005). Watermelon breeders regularly acquire new germplasm from diverse geographical regions for genetic improvement of the crop (Zhang et al., 2012; Yang et al., 2016). However, there is limited information on the magnitude of genetic variation of watermelon genetic resources (Hashizume et al., 2003; Weng et al., 2016). Various studies reported that South African watermelon germplasm have been useful in disease resistance breeding programs in the USA (Gusmini et al., 2005; Tetteh et al., 2010; Thies et al., 2010; Ma and Wehner, 2015; Levi et al., 2016). However, information is scanty regarding genetic diversity of dessert watermelon germplasm from South Africa as there have been no prior studies on genetic clustering using molecular data.

Dessert watermelon is economically important crop in South Africa. During the 2011/12 season, a total of 58 955 tons of dessert watermelon was produced in the country (Department of Agriculture Forestry and Fisheries, 2013). These level of production was comparatively lower than previous years, 2009/10 (69, 001 tons) and 2010/11 (104 852 tons, respectively). Despite the low levels of production, South African watermelon exports increased significantly by 251% between 2010 and 2012 with an estimated value of R19.9 million during 2012 (National Agricultural Marketing Council, 2013). The reasons for low yields perhaps point to the use of genetically inferior genotypes, hence the need to develop genetically superior genotypes. Small-holder farmers in South Africa grow unimproved dessert watermelon landraces which exhibit great morphological diversity such as variation in fruit shape, exocarp colouring patterns and seed morphology (Hashizume et al., 2003). Dessert watermelon landraces grown and maintained by small-holder farmers in the country could be useful germplasm to improve yield and yield related traits, quality, biotic and abiotic stresses. Therefore, it is important to develop systematic genetic groupings using diverse genetic pool of watermelon adapted to local growing conditions. This will assist in developing varieties in a reduced timeline in South Africa or other breeding programs. The efficiency of genetic classification of genotypes based on genetic distance estimates depends on the availability of polymorphic SSR markers. Therefore, the objective of this study was to determine the genetic diversity present among dessert watermelon landraces widely grown in South Africa using SSR markers and to select genetically diverse and complimentary genotypes for breeding and/or conservation.

Results

Polymorphism and allelic diversity of SSR markers

Estimates of genetic parameters are presented in Table 3. The SSR markers generated a total of 94 putative alleles (different fragment sizes) among the dessert watermelon landraces. Number of alleles ranged from 2 for the marker MCPI-03 to 19 for marker BVWS00433 with a mean of 9.4 per locus. Number of effective alleles ranged from 1.07 to 9.94 with a mean of 4.61. Observed heterozygosity (H_o) values ranged from 0.07 (MCPI-03) to 1.00 (MCPI-37) with a mean of 0.53, suggesting almost half of the loci were heterozygous and the remaining half of the loci reached acceptable level of homozygosity. Expected heterozygosity (H_e), as a measure of total gene diversity; ranged from 0.07 (MCPI-03) to 0.91 (BVWS00433) with a mean of 0.66, suggesting that 66% of

the genetic individuals are expected to be heterozygous at a given locus under random mating conditions. Fixation index (F) exhibited contrasting values ranging from -0.47 to 0.62 with a mean of 0.18. F represents the average deviation of the population's genotypic proportions from Hardy-Weinberg equilibrium for a locus and the values range from 0 to 1. A negative F value represents an excess of heterozygotes. For example, for locus MCPI-37, 69% of the dessert watermelon landraces are expected to be heterozygous at the specific locus under random mating conditions; however, 100% of the genotypes at this locus were heterozygotes. This may be due to the high outcrossing nature of dessert watermelon or mutations at the specific loci. The polymorphic information content for the 10 SSR markers ranged from 0.06 to 0.90, with a mean of 0.65. Over 80% of the SSR-loci had PIC values of > 0.65, indicating an adequate discriminatory power of individual SSR loci used in the study. This suggested the potential of these loci for genetic diversity analysis in watermelon. The mean highest allele frequency was 0.47. Genetic differentiation (F_{ST}) ranged from 0.01 to 0.04 with a mean of 0.02. A low mean value of F_{ST} implies that the dessert watermelon landraces are not genetically different. An indirect estimate of gene flow (N_m) ranged from 3.68 to 63.42 with a mean of 18.98, suggesting high exchange of genes between the studied materials.

Genetic variability within and among populations

Genetic parameter estimates among dessert watermelon populations based on district of collection are presented in Table 4. The mean observed and effective number of detected alleles was higher for Waterberg District compared to Capricorn District, suggesting the existence of genetically diverse and unique landraces from Waterberg District. Shannon's information index was much higher in Waterberg District (1.56) compared to Capricorn District (1.31). The mean observed and expected heterozygosity was 0.51 and 0.63 for Capricorn District and 0.55 and 0.69 for Waterberg District, respectively. No differences were observed between the two Districts with regards to fixation index. The lowest F-value detected in Waterberg and Capricorn Districts suggests the environment may favor out crossing in this districts. Highest number private alleles per population were detected in Waterberg District (34) compared Capricorn District (15). Polymorphic information content was 0.60 and 0.67 for Capricorn and Waterberg Districts, respectively (Table 4).

Analysis of molecular variance

Analysis of molecular variance (AMOVA) among dessert watermelon populations are shown in Table 5. Non-significant ($P > 0.05$) differences of molecular variation were detected among populations. However, highly significant differences ($P < 0.001$) of molecular variation were observed among and within individuals. The largest genetic variability (75%) was attributed to variation within individuals, whereas 25% of the total variation was explained by variation among individuals (Table 5). These suggest that dessert watermelon populations had slightly higher and significant contribution to the total molecular variances observed hence the much higher variation detected within individuals.

Cluster analysis

Genetic relationships among the dessert watermelon landraces was examined using neighbour-joining algorithm using the unweighted pair group method (UPGMA). The

Table 1. List of 31 dessert watermelon landraces used in the study with collection districts in Limpopo Province of South Africa.

Sr. No.	Entry	Location	District	Sr. No.	Entry	Location	District
1	SWM-01	Ga-Kgoroshi	Capricorn	17	SWM-26	Ga-Kgoroshi	Capricorn
2	SWM-02	Kgohloane	Capricorn	18	SWM-27	Ga-Kgoroshi	Capricorn
3	SWM-04	Kgohloane	Capricorn	19	SWM-29	Ga-Kgoroshi	Capricorn
4	SWM-05	Kgohloane	Capricorn	20	SWM-28	Ga-Kgoroshi	Capricorn
5	SWM-06	Kgohloane	Capricorn	21	SWM-30	Ga-Kgoroshi	Capricorn
6	SWM-07	Kgohloane	Capricorn	22	SWM-31	Ga-Kgoroshi	Capricorn
7	SWM-10	Maeteletsa	Waterberg	23	SWM-32	Thabaleshoba	Waterberg
8	SWM-13	Maeteletsa	Waterberg	24	SWM-33	Thabaleshoba	Waterberg
9	SWM-14	Maeteletsa	Waterberg	25	SWM-34	Thabaleshoba	Waterberg
10	SWM-15	Maeteletsa	Waterberg	26	SWM-35	Thabaleshoba	Waterberg
11	SWM-18	Maeteletsa	Waterberg	27	SWM-36	Thabaleshoba	Waterberg
12	SWM-20	Ga-Molepo	Capricorn	28	SWM-38	Thabaleshoba	Waterberg
13	SWM-21	Ga-Molepo	Capricorn	29	SWM-39	Thabaleshoba	Waterberg
14	SWM-22	Ga-Molepo	Capricorn	30	SWM-40	Thabaleshoba	Waterberg
15	SWM-24	Kgohloane	Capricorn	31	SWM-41	Thabaleshoba	Waterberg
16	SWM-25	Ga-Kgoroshi	Capricorn				

Sr. No: Serial number

Table 2. Description of the simple sequence repeats (SSR) primers used for dessert watermelon genetic diversity analysis.

SSR Primer	Forward primer (5'→3')	Reverse primer (3'→5')	PIC
MCPI-03	GCATAAACCCACCTGTGAGTGG	ATGGCTTTGCGTTTCATTTC	0.80 ⁽¹⁾
MCPI-12	GGAGTAGTGGTGGAGACATGG	TCCTTCTCTTTTCGCAAACCTTC	0.80
MCPI-37	AATCTTCCCATGCCAAAAC	GACTTCCAAACCCTCCCTTC	0.87
MCPI-21	AAAGTTTTTCATGCCAACGTATC	TCAGCCAATATGGTCAAATAGC	0.88
MCPI-13	TTCCTGTTTCATGATTCTCCAC	TCAGAATGGAGCCATAAATTG	0.88
MCPI-14	TCAAATCCAACCAAAATATTGC	GAGAAGGAAACATCACCACG	0.97
BVWS01734	AAAATTACATCTTAAATGCGCC	GGAACATTGACTTCAATCAGCA	0.74 ⁽²⁾
BVWS00433	TCTTTTAAGTTTTGAGGGAGAGC	TTCCAAGCTAGCCTTTTCA	0.82
BVWS00209	TGCTTCAAATCTATTACAATTTGC	TTCTTGGTTTCGGGTTTCTTTACA	0.79
BVWS00228	GGAAGAGTGAGGTGATAAATCAATATGT	AATTGGCCCAAATATCCATATGAC	0.74

PIC: Polymorphic information content; ⁽¹⁾ Joobeur et al. (2006); ⁽²⁾ Zhang et al. (2016)

Table 3. Genetic parameters generated by 10 SSR markers among dessert watermelon landrace collections of South Africa.

Loci	Fragment size range	Genetic parameters								
		N _a ⁽¹⁾	N _e	H _o	H _e	F	PIC	HAF	F _{ST}	N _m
MCPI-03	225-255	2	1.07	0.07	0.07	-0.03	0.06	0.97	0.04	6.50
MCPI-12	157-255	5	3.00	0.77	0.68	-0.15	0.67	0.50	0.00	63.42
MCPI-37	120-250	8	3.15	1.00	0.69	-0.47	0.68	0.45	0.01	24.37
MCPI-21	171-250	7	2.51	0.23	0.61	0.62	0.60	0.58	0.01	34.48
MCPI-13	195-250	10	4.46	0.38	0.79	0.51	0.78	0.40	0.06	3.68
MCPI-14	198-310	4	1.35	0.16	0.27	0.38	0.26	0.85	0.02	13.91
BVWS01734	120-278	13	6.04	0.93	0.85	-0.12	0.83	0.25	0.02	13.28
BVWS00433	217-306	19	9.94	0.83	0.91	0.07	0.90	0.22	0.02	13.53
BVWS00209	135-233	13	7.11	0.37	0.87	0.57	0.86	0.23	0.03	7.73
BVWS00228	135-233	13	7.50	0.53	0.88	0.38	0.87	0.22	0.03	8.87
Mean	-	9.4	4.61	0.53	0.66	0.18	0.65	0.47	0.02	18.98
SE	-	1.6	0.93	0.11	0.09	0.12	0.09	0.08	0.01	5.73

⁽¹⁾N_a: Number of alleles per locus; N_e: number of effective alleles per locus; H_o: observed heterozygosity; H_e: expected heterozygosity; F: Fixation index, PIC: polymorphic information content; HAF: the highest allele frequency per locus; F_{ST}: Genetic differentiation; N_m: Gene flow; SE: Standard error.

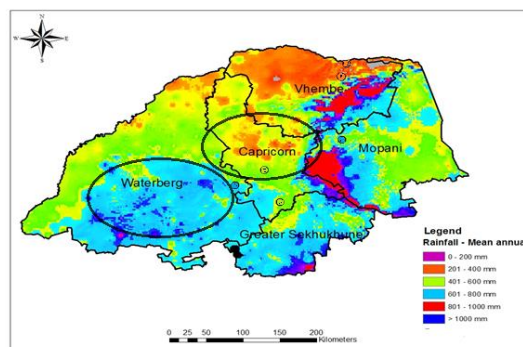


Fig 1. Map of the Limpopo Province of South Africa showing the two districts (circled in black) where the dessert watermelon landraces were collected.

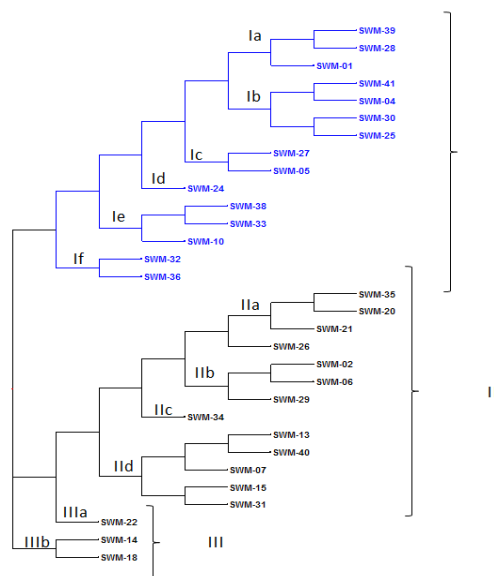


Fig 2. Dendrogram using neighbor-joining algorithm using the unweighted pair group method (UPGMA) revealing genetic relationships among dessert watermelon landraces based on SSR markers. Ia, Ib, Ic, Id, Ie and If; IIa, IIb, IIc and IId; and IIIa and IIIb denote subgroups within the clusters. See codes of genotypes in Table 1.

Table 4. Genetic parameter estimates of dessert watermelon populations based on districts of collection.

Populations/Districts	Genetic parameters								
	$N^{(1)}$	N_a	N_e	I	H_o	H_e	F	PA	PIC
Capricorn	16.4	6.00	3.70	1.31	0.51	0.63	0.18	15.00	0.60
Waterberg	13.7	7.90	4.85	1.56	0.55	0.69	0.17	34.00	0.67
Mean	15.1	6.95	4.27	1.44	0.53	0.66	0.17	-	0.64
SE	0.3	0.86	0.58	0.16	0.07	0.06	0.08	-	0.03

⁽¹⁾ N : Number of observations; N_a : number of alleles per locus; N_e : number of effective alleles per locus; I: Shannon's information index; H_o : observed heterozygosity; H_e : expected heterozygosity; F: Fixation index; PA: Private allele per population; PIC: Polymorphic information content.

Table 5. Analysis of molecular variance among and within the dessert watermelon landraces.

Source of variation	df ⁽¹⁾	SS	MS	Est. Var.	Perc. Var.	F-statistics
Among populations	1	4.520	4.520	0.009	0%	0.309
Among individuals	29	123.061	4.243	0.848	25%	0.001
Within individuals	31	79.000	2.548	2.548	75%	0.001
Total	61	206.581	-	3.405	100%	

⁽¹⁾df: degrees of freedom; SS: sum of squares; MS: mean squares; Est. var: estimated variance; Perc. Var: percentage variance.

analyses indicated the presence of three distinct clusters namely: cluster I, cluster II and cluster III (Fig. 2). Cluster I was further sub-divided into six sub-clusters consisting of 15 landraces. Cluster II consisted of four sub-clusters with 13 landraces. Cluster III contained two sub-clusters consisting of three landrace collections. Overall, the cluster analysis allowed selection of unique and genetically complementary landraces based on their high dissimilarity index for breeding and conservation. Landraces selected from cluster I included SWM-39, SWM-01, SWM-04, SWM-27, SWM-24, SWM-10 and SWM-36, selections from cluster II were: SWM-35, SWM-21, SWM-02, SWM-34, SWM-07 and SWM-31; and SWM-22 and SWM-18 selected from cluster III. These selections are recommended for subsequent phenotypic characterization to select complementary genotypes with promising horticultural attributes for effective breeding.

Discussion

The present study assessed genetic diversity of South African dessert watermelon landraces using SSR markers. The SSR markers generated a total of 94 putative alleles (Table 3).

This was higher than 41 alleles detected by de Gama et al. (2013) among dessert watermelon accessions. Sheng et al. (2012) detected a total of 94 putative alleles, similar to the current study. However, this was lower than 133, 101 and 795 alleles detected by Zhang et al. (2016), Wang et al. (2015) and Reddy et al. (2015) reported among watermelon accessions, respectively. Nimmakayala et al. (2009) reported a total of 169 alleles amplified in watermelon accessions. In the current study the number of alleles per locus ranged from 2 to 19, with a mean of 9.4. This was higher than the range from 2 to 7 alleles (mean = 3.7) reported by Zhang et al. (2012). Jarret et al. (1997) reported an average of 4.7 alleles per locus in 32 watermelon accessions which was lower than those observed in the current findings. Similarly, Guerra-Sanz (2002) reported 1 to 8 alleles (average = 3.6) in eight watermelon genotypes which was also lower than those observed in the study. The number of effective alleles (N_e) ranged from 1.07 to 9.94 with a mean of 4.61. Reddy et al. (2015) reported mean N_e value of 1.54 ranging from 1.00 to 3.16 lower than current study. Observed heterozygosity (H_o) values ranged from 0.07 to 1.00 with a mean of 0.53. Expected heterozygosity (H_e), as a measure of total gene

diversity, ranged from 0.07 to 0.91 with a mean of 0.66. The H_o and H_e values observed in the current study are higher than the mean value of 0.31 and 0.38 reported by Minsart et al. (2011) among watermelon accessions. Reddy et al. (2015) reported a mean gene diversity (H_e) of 0.34 among 90 dessert watermelon genotypes lower than the current study. These differences are probably due to higher number of detected alleles per locus (Nyaligwa et al., 2015) due to the outcrossing nature of dessert watermelon. The high mean value for expected heterozygosity in this study suggests a high level of heterozygosity among the tested dessert watermelon landrace collections. The mean fixation index recorded in the present study was 0.18 suggesting that watermelon populations were relatively heterozygous, which is explained by the outcrossing nature of the crop. The polymorphic information content (PIC) ranged from 0.06 to 0.90, with a mean of 0.65. Kwon et al. (2010) reported PIC values ranging from 0.22 to 0.80 with a mean of 0.50 lower than the current study. Zhao et al. (2014) reported PIC values ranging from 0.04 to 0.67 with a mean of 0.39 lower than the current study. Minsart et al. (2011) and de Gama et al. (2013) also reported mean PIC values of 0.37 and 0.39, respectively also lower than the current study. The high levels of polymorphism observed in the current study could be due to large genome size, outcrossing nature and heterozygosity of watermelon. Also, differences in PIC values may be due to the polymorphism of the SSR markers used and genetic differences among the South African dessert watermelon landraces. The PIC values of loci provide an estimate of the discriminatory power of loci, considering the number of alleles and their relative frequencies (Smith et al., 2000). In this study, over 80% of the SSR-loci in this study had PIC value of > 0.65, indicating an adequate discriminatory power of individual SSR loci used in the study. Analysis of molecular variance (AMOVA) among wild watermelon populations revealed 0, 25 and 75% of the variation was attributable to among populations, among and within individuals, respectively (Table 5). Results of the present study agree with Mujaju et al. (2011) and Minsart et al. (2011) who reported similar trends among and between watermelon accessions. de Gama et al. (2013) also reported 68 and 32% within and among watermelon accessions respectively. The magnitudes of between and within population differentiation were quantified using the F-statistics according to Wright (1951). Genetic differentiation and gene flow are important parameters for evaluating the genetic structure of populations. A low genetic differentiation (0.02) was detected among the tested landraces (Table 3). According to Wright (1978), F_{ST} values ranging from 0 to 0.005 indicates low, 0.05–0.15 moderate, 0.15–0.25 high, and above 0.25 very high genetic differentiations. The low genetic differentiation could be attributed to the high level of gene flow among the landraces. In this study, the mean flow among the landraces was 18.98, suggesting a high genetic introgression (Table 3). Similar results were reported by Reddy et al. (2015) with a low genetic differentiation (F_{ST}) value of 0.04 and high gene flow value (N_m) of 69.63 among dessert watermelon genotypes. According to Morjan and Rieseberg (2004), gene flow < 1 is considered to be low whereas $N_m = 1$ is considered to be moderate. Moderate or relatively low levels of gene flow can significantly reduce loss of genetic diversity (Djè et al., 1999; Aguilar et al., 2008). The high level of gene flow observed may be attributed to an exchange of genetic materials (Bhawna et al., 2014) between farmers leading to low levels of genetic differentiation (Bhawna et al., 2015). Also, high gene flow could be attributed to a high degree of movement of

germplasm probably through frequent seed exchange among farmers. This practice results in low genetic variability among individuals within populations (Bhawna et al., 2015). Cluster analysis revealed three main genetic groups among the tested dessert watermelon landraces revealing great genetic diversity for breeding and strategic conservation (Fig. 2). Dessert watermelon has a narrow genetic base (Levi et al., 2001; Hwang et al., 2011) probably due to many years of domestication and selection for desirable horticultural traits (Levi et al., 2001; Dane and Liu, 2007; Minsart et al., 2011). Therefore, continued search, identification and selection of unique genotypes is important for sustainable production of dessert watermelon (Sheng et al., 2012; Zhang et al., 2012; Yang et al., 2016).

Materials and methods

Plant materials

Thirty one dessert watermelon landraces collected from Capricorn and Waterberg districts of the Limpopo Province of South Africa were used in the study. Table 1 summarises information related to the collection sites of landraces. The map showing the collection sites of the dessert watermelon samples is also shown in Fig. 1.

DNA extraction, purification and quantification

Seed of diverse dessert watermelon landraces were planted in 2 litre capacity polyethylene pots at the Controlled Research Facility (CEF), University of KwaZulu-Natal, Pietermaritzburg (29°37'51.75" S; 30°23'59.10" E), South Africa. Young fresh leaves were harvested from 40 plants per landrace six weeks after planting. The leaf samples were sent to INCOTEC PROTEIOS laboratory (Incotech, SA Pty Ltd, Mkhondeni, Pietermaritzburg, South Africa) for SSR analysis. The DNA was extracted following the CTAB (mixed alkyltrimethyl-ammonium bromide protocol) as described by CIMMYT (2005). The concentration of the extracted DNA was determined using 0.7% Tris-Borate-EDTA (TBE) agarose gel. A working concentration of 10 ng μl^{-1} was standardized for all extracted DNA (Erasmus, 2008). The samples were bulked and used in SSR amplification.

PCR and SSR analyses

All samples were used in bulked amplification, using DNA extracted from the leaf material. SSR sequences were amplified through polymerase chain reaction (PCR) using SSR primers specific for watermelon. Ten SSR markers were used for the analysis (Table 2) and were selected based on their high polymorphic information content and being developed specifically for watermelon (Joobeur et al., 2006; Zhang et al., 2012). High polymorphic information content values suggest markers may have high discriminatory power to distinguish differences between genotypes. PCR was performed using 12 μl of reaction mixture containing 1 x PCR buffer, 2.5 mM Mg^{++} , 0.2 μl each of dNTPs (Bioline), 1 unit of Taq polymerase (Bioline) and 5–10 ng of genomic DNA. Primers were labeled with a 104 fluorescent dye. Two primers were provided for the amplification of each SSR locus: one tailed forward primer (0.25 μmol) and one normal reverse primer (0.25 μmol). The initial denaturation step was performed at 94°C for 2 minutes, followed by 33 cycles at 94°C for 30 seconds. Annealing of primer at primer specific 3°C for 30 seconds and 72°C for 45 seconds with a final extension for 20 m minutes (Erasmus, 2008). PCR products

were fluorescently labelled and separated by capillary electrophoresis on an ABI 3130 automatic sequencer (Applied Biosystems, Johannesburg, South Africa).

Data analysis

Genetic diversity analysis

Two approaches were adopted to investigate the genetic structure and diversity among the dessert watermelon landraces. In the first approach, polymorphisms were treated as binary data (presence or absence). However, to determine the genetic structure within and among landraces, a second approach based on the co-dominant nature of the marker was adopted using GenAlex version 6.5 (Peakall and Smouse, 2012). Genetic diversity parameters, such as number of alleles per locus (N_a), number of effective alleles per locus (N_e), observed (H_o) and expected (H_e) heterozygosity and Shannon's Information Index (I) were calculated using GenAlex version 6.5 (Peakall and Smouse, 2012) according to the protocol described by Nei and Li (1979). Further, an indirect estimate of the level of gene flow (N_m) between the genotypes was calculated using the formula: $Nm = 0.25 (1 - F_{ST}/F_{ST})$ using GenAlex. Polymorphic information content (PIC) was calculated using the formula: $PIC = \sum P_{ij}^2$, where P_{ij} is the frequency of j^{th} allele of the i^{th} locus (Nagy et al., 2012). The F-statistics such as genetic differentiation (F_{ST}), fixation index or inbreeding coefficient (F_{IS}) and overall fixation index (F_{IT}) were calculated according to Wright's original derivation (Wright, 1951).

Genetic distance and cluster analysis

The binary data was used to obtain dissimilarity matrix using the Jaccard index. The matrix was then used to perform cluster analysis based on neighbor-joining algorithm using the unweighted pair group method (UPGMA) using Darwin 5.0 (Perrier and Jacquemoud-Collet, 2006). A dendrogram was then generated on the dissimilarity matrix. Bootstrap

Aguilar R, Quesada M, Ashworth L, Herrerias-Diego Y, Lobo J (2008) Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. *Mol Ecol*. 17:5177-5188.

Bhawna AMZ, Arya L, Ram C, Sureja AK, Verma M (2015) Development of novel gene-based microsatellite markers for robust genotyping purposes in *Lagenaria siceraria*. *Sci Hortic*. 191:15-24.

Bhawna AMZ, Arya L, Saha D, Sureja AK, Pandey C, Verma M (2014) Population structure and genetic diversity in bottle gourd [*Lagenaria siceraria* (Mol.) Standl.] germplasm from India assessed by ISSR markers. *Plant Syst Evol*. 300:767-773.

Che KP, Liang CY, Wang YG, Jin DM, Wang B (2003) Genetic assessment of watermelon germplasm using the AFLP technique. *HortScience*. 38:81-84.

Chomicki G, Renner SS (2015) Watermelon origin solved with molecular phylogenetics including Linnaean material: another example of museomics. *New Phytol*. 205 526-532.

CIMMYT (2005) Laboratory Protocols: CIMMYT applied molecular genetics laboratory 3rd edn. CIMMYT, Mexico.

Dane F, Liu J (2007) Diversity and origin of cultivated and citron type watermelon (*Citrullus lanatus*). *Genet Resour Crop Evol*. 54:1255-1265.

de Gama RNC, Santos CAF, Dias RCS (2013) Genetic variability of watermelon accessions based on microsatellite markers. *Genet Mol Res*. 12:747-754.

analysis was performed for node construction using 10,000 bootstrap values to estimate the reliability of the clustering pattern.

Analysis of molecular variance

The partitioning of total genetic variation using Analysis of Molecular Variance (AMOVA) was performed to estimate population genetic structure and differentiation among and within dessert watermelon landrace collections. AMOVA uses the estimated F- statistics such as F_{ST} , F_{IS} , and F_{IT} to compare the genetic structure among and within populations. For easy management and utilization, the total molecular variance was dissected into within and among population variations. The AMOVA procedure was performed using GenAlex 6.5 according to Nei (1978).

Conclusion

The present study revealed the presence of genetic diversity among South African dessert watermelon landrace collections. Genetically unique landraces such as SWM-39, SWM-01, SWM-04, SWM-27, SWM-24, SWM-10 and SWM-36 from cluster I; SWM-35, SWM-21, SWM-02, SWM-34, SWM-07 and SWM-31 from cluster II; and SWM-22 and SWM-18 from cluster III were selected based on their dissimilarity values. These are recommended for further phenotyping using horticultural attributes for breeding and systematic conservation.

Acknowledgements

The University of KwaZulu-Natal, Limpopo Department of Agriculture and Rural Development and the National Research Foundation of South Africa are acknowledged for financial support of this study.

References

- Department of Agriculture Forestry and Fisheries (2013) Abstract of Agricultural Statistics, Pretoria, South Africa, 43p.
- Djè Y, Forcioli D, Ater M, Lefèbvre C, Vekemans X (1999) Assessing population genetic structure of sorghum landraces from North-western Morocco using allozyme and microsatellite markers. *Theor Appl Genet*. 99:157-163.
- Erasmus TE (2008) Genetic diversity of proprietary inbred lines of sunflower, determined by mapped SSR markers and total protein analysis. PhD thesis, School of Agricultural Sciences and Agribusiness, University of KwaZulu-Natal, Pietermaritzburg.
- Guerra-Sanz JM (2002) *Citrullus* simple sequence repeats markers from sequence databases. *Mol Ecol Notes*. 2:223-225.
- Gusmini G, Song G, Wehner TC (2005) New sources of resistance to Gummy stem blight in watermelon. *Crop Sci*. 45:582-588.
- Hashizume T, Shimamoto I, Hirai M (2003) Construction of a linkage map and QTL analysis of horticultural traits for watermelon [*Citrullus lanatus* (Thumb.) Matsum & Nakai] using RAPD, RFLP and ISSR markers. *Theor Appl Genet*. 106:779-785.
- Hwang JH, Ahn SG, Oh JY, Choi YW, Kang JS, Park YH (2011) Functional characterization of watermelon (*Citrullus lanatus* L.) EST-SSR by gel electrophoresis and high resolution melting analysis. *Sci Hortic*. 130:715-724.
- Jarret RL, Merrick LC, Holms T, Evans J, Aradhya MK (1997) Simple sequence repeats in watermelon (*Citrullus lanatus* (Thumb) Matsum. & Nakai). *Genome*. 40:433-441.
- Jobeur T, Gusmini G, Zhang X, Levi A, Xu Y, Wehner TC, Oliver M, Dean RA (2006) Construction of a watermelon BAC

- library and identification of SSRs anchored to melon or *Arabidopsis* genomes. *Theor Appl Genet.* 112:1553–1562.
- Kwon YS, Oh YH, Yi SI, Kim HY, An JM, Yang SY, Ok SH, Shin JS (2010) Informative SSR markers for commercial variety discrimination in watermelon (*Citrullus lanatus*). *Genes Genome.* 32:115–122.
- Lee SJ, Shin JS, Park KW, Hong YP (1996) Detection of genetic diversity using RAPD-PCR and sugar analysis in watermelon [*Citrullus lanatus* (Thunb.) Mansf.] germplasm. *Theor Appl Genet.* 92:719–725.
- Levi A, Harris-Schulz KR, Ling KS (2016) USVL-370, a Zucchini yellow mosaic virus-resistant watermelon breeding line. *HortScience.* 51:107–109.
- Levi A, Thomas CE, Wehner TC, Zhang X (2001) Low genetic diversity indicates the need to broaden the genetic base of cultivated watermelon. *HortScience.* 36:1096–1101.
- Levi A, Thomas CE, Newman M, Reddy O, Zhang X, Xu Y (2004) ISSR and AFLP markers differ among American watermelon with limited genetic diversity. *J Am Soc Hortic Sci.* 129:553–558.
- Ma S, Wehner TC (2015) Flowering stage resistance to bacterial blotch in the watermelon germplasm collection. *Crop Sci.* 55:727–736.
- Maggs-Kölling GL, Madsen S, Christiansen JL (2000) A phenetic analysis of morphological variation in *Citrullus lanatus* in Namibia. *Genet Resour Crop Evol.* 47:385–393.
- Minsart LA, Zoro-Bi IA, Dje` Y, Baudoin JP, Jacquemart AL, Bertin P (2011) Set up of simple sequence repeat markers and first investigation of the genetic diversity of West- African watermelon (*Citrullus lanatus* ssp. *vulgaris* oleaginous type). *Genet Resour Crop Evol.* 58:805–814.
- Morjan CL, Rieseberg LH (2004) How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. *Mol Ecol.* 13:1341–1356.
- Mujaju C, Zborowska A, Werlemark G, Garkava-Gustavsson L, Andersen SB, Nybom H (2011) Genetic diversity among and within watermelon (*Citrullus lanatus*) landraces in Southern Africa. *J Hortic Sci Biotechnol.* 86:353–358.
- Munnisse P, Jensen BD, Andersen SB (2013) Genetic differentiation of watermelon landraces in Mozambique using microsatellite markers. *Afr J Biotechnol.* 12:5513–5521.
- Nagy S, Pocza P, Cernák I, Gorji AM, Hegedús G, Taller J (2012) PIC calc: an online program to calculate polymorphic information content for molecular genetic studies. *Biochem Genet.* 50:670–672.
- National Agricultural Marketing Council (2013) Markets and Economic Research Centre: SA fruit flow, Pretoria, 1–15p.
- Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genet Mol Rep.* 89:583–590.
- Nei M, Li W (1979) Mathematical method for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci USA.* 76:5256–5273.
- Nimmakayala P, Tomason YR, Jeong J, Ponniah SK, Karunathilake A, Amnon L, Perumal R, Reddy UK (2009) Genetic reticulation and interrelationships among citrullus species as revealed by joint analysis of shared AFLPs and species-specific SSR alleles. *Plant Genet Resour.* 8:16–25.
- Nyaligwa L, Shimelis H, Beyene A, Habteab G (2015) Genetic diversity analysis of elite maize inbred lines of diverse sources using SSR markers. *Maydica.* 60:1–9
- Ocal N, Akbulut M, Osman GO, Yetisir H, Solmaz I, Sari N (2014) Genetic diversity, population structure and linkage disequilibrium among watermelons based on peroxidase gene markers. *Sci Hortic.* 176 151–161.
- Peakall R, Smouse PE (2012) GenAlex 6.5: genetic analysis in Excel. Population software for teaching and research-an update. *Bioinformatics.* 28:2537–2539.
- Perrier X, Jacquemoud-Collet JP (2006) DARwin software. Dissimilarity Analysis and Representation for windows. Available from <http://www.darwin.cirad.fr/darwinhtml>.
- Reddy UK, Abburi L, Abburi VL, Saminathan T, Cantrell R, Vajja VG, Reddy R, Tomason YR, Levi A, Wehner TC, Nimmakayala P (2015) A genome-wide scan of selective sweeps and association mapping of fruit traits using microsatellite markers in watermelon. *J Heredity.* 106:166–176.
- Sheng Y, Luan F, Zhang F, Davis RA (2012) Genetic diversity within Chinese watermelon ecotypes compared with germplasm from other countries. *J Am Soc Hortic Sci.* 137:144–151.
- Smith JSC, Kresovich S, Hopkins MS, Mitchell SE, Dean RE, Woodman WL, Lee M, Poster K (2000) Genetic diversity among elite sorghum inbred lines assessed with simple sequence repeats. *Crop Sci.* 40:226–232.
- Solmaz I, Sari N, Aka-Kacar Y, Yalcin-Mendi NY (2010) The genetic characterization of Turkish watermelon (*Citrullus lanatus*) accessions using RAPD markers. *Genet Resour Crop Evol.* 57:763–771.
- Tetteh AY, Wehner TC, Davis AR (2010) Identifying resistance to powdery mildew race 2W in the USDA-ARS watermelon germplasm collection. *Crop Sci.* 50:933–939.
- Thies JA, Ariss JJ, Hassell RL, Olson S, Kousik CS, Levi A (2010) Grafting for management of Southern root-knot nematode, *Meloidogyne incognita*, in watermelon. *Plant Dis.* 94:1195–1199.
- Uluturk ZI, Frary A, Doganlar S (2011) Determination of genetic diversity in watermelon [*Citrullus lanatus* (Thunb.) Matsum & Nakai] germplasm. *Aust J Crop Sci.* 5:1832–1836.
- Varshney RK, Graner A, MES (2005) Genomics-assisted breeding for crop improvement. *Trends Plant Sci.* 10:621–630.
- Wang P., Li Q., Hu J., Su Y. (2015) Comparative analysis of genetic diversity among Chinese watermelon germplasms using SSR and SRAP markers, and implications for future genetic improvement. *Turk J Agric For.* 39:322–331.
- Weng Y, Davis A, Xu Y (2016) Genetic diversity, population structure, and formation of a core collection of 1197 *Citrullus* accessions. *HortScience.* 51:23–29.
- Wright S (1951) The genetical structure of populations. *Ann Eugen.* 15:323–354.
- Wright S (1978) Evolution and the genetics of populations: variability within and among natural populations. University of Chicago Press, Chicago.
- Yang X, Ren R, Ray R, Xu J, Li P, Zhang M, Liu G, Yao X, Kilian A (2016) Genetic diversity and population structure of core watermelon (*Citrullus lanatus*) genotypes using DArTseq-based SNPs. *Plant Genet Resour.* 20:226–233.
- Zhang H, Fan J, Guo S, Ren Y, Gong G, Zhang J (2016) Genetic diversity, population structure, and formation of a core collection of 1197 citrullus accessions. *HortScience.* 51:23–29.
- Zhang H, Wang H, Guo S, Ren Y, Gong G, Weng Y, Xu Y (2012) Identification and validation of a core set of microsatellite markers for genetic diversity analysis in watermelon, *Citrullus lanatus* Thunb. Matsum. & Nakai. *Euphytica.* 186:329–342.
- Zhao SJ, Zhu HJ, Lu XQ, He H, Liu WG (2014) Studies on DNA fingerprinting and genetic diversity of seedless watermelon (*Citrullus lanatus*) varieties using core simple sequence repeat (SSR) markers. *J Agric Biotechnol.* 22:188–194.