J. Amer. Soc. Hort. Sci. 142(1):47–56. 2017. doi: 10.21273/JASHS03834-16

Genetic Diversity in the Desert Watermelon *Citrullus colocynthis* and its Relationship with *Citrullus* Species as Determined by High-frequency Oligonucleotides-targeting Active Gene Markers

Amnon Levi[1], Alvin M. Simmons, Laura Massey, John Coffey, and W. Patrick Wechter
U.S. Department of Agriculture, Agricultural Research Service, U.S. Vegetable Laboratory, 2700 Savannah Highway, Charleston, SC 29414

Robert L. Jarret
U.S. Department of Agriculture, Agricultural Research Service, Plant Genetic Resources Conservation Unit, 1109 Experiment Street, Griffin, GA 30223

Yaakov Tadmor
Agricultural Research Organization (ARO) Ne'we Ya'ar Research Center, P.O. Box 1021, Ramat Yishai 30095, Israel

Padma Nimmakayala and Umesh K. Reddy
Department of Biology, Douglass Land-grant Institute, West Virginia State University, Institute, WV 25112

Additional index words. Cucurbitaceae, crop evolution, population structure, plant breeding, germplasm

Abstract. ***Citrullus colocynthis*** **(CC) is a viable source of genes for enhancing disease and pest resistance in common cultivated watermelon [*Citrullus lanatus* var. *lanatus* (CLL)] cultivars. However, there is little information about genetic diversity within CC or the relationship of CC accessions to *C. lanatus*. In this study, we examined genetic diversity and relationships among 29 CC accessions collected in northern Africa, the Middle East, and Asia, and their relationships to 3 accessions and 3 cultivars of CLL, 12 accessions of citron melon [*C. lanatus* ssp. *lanatus* var. *citroides* (CLC)], and 1 accession representing the desert perennial *Citrullus ecirrhosus* (CE). Twenty-three high-frequency oligonucleotides-targeting active gene (HFO–TAG) primers were used to produce a total of 431 polymorphic fragments that target coding regions of the genome. Cluster and multidimensional scaling plot analysis, separated the CC into five groups, in general agreement with their geographic origins. CC genotypes admixed with CLL and CLC also were identified. Major reproductive barriers resulted in significantly reduced fertility in CC × CLL hybridizations. However, several of the U.S. PIs of CC were successfully crossed with watermelon cultivars using traditional breeding procedures, and the seeds produced from these crosses were viable. This suggests that CC can be a viable source to introduce biotic and abiotic stress resistance genes into cultivated watermelon.**

Watermelon is an important vegetable crop throughout the world. It belongs to the xerophytic genus *Citrullus* and is cultivated in temperate and tropical regions of the world (Jarret et al., 1997; Paris, 2015; Wehner, 2008). The genus *Citrullus* includes several diploid ($2n = 22$) species (Shimotsuma, 1963), including *C. lanatus*, which gave rise to the red-fleshed sweet dessert watermelon, as well as the "egusi" type [also referred to as *Citrullus mucosospermus* (Fursa, 1972)], which is cultivated for its large oily seeds (Jarret and Levy, 2012). *Citrullus lanatus* also includes the "tsamma," citron or preserving melon, which is common in southern Africa and is known in the literature as CLC (Jeffrey, 2001; Whitaker and Bemis, 1976), *Citrullus vulgaris* var. *citroides* (Bailey, 1930), and *Citrullus lanatus* ssp. *lanatus* var. *citroides* (Fursa, 1972). Citron melons are also referred to in the literature as CLL (Jeffrey, 2001). Fursa (1972) recognized CLC as being distinct from *C. lanatus* ssp. *lanatus* var. *capensis* in which he placed *Citrullus amarus*. Recently, Chomicki and Renner (2014) indicated that *C. amarus* is the botanical name for citron melons and provided updated taxonomic names for *Citrullus* species. The Plant List (2013) does not yet acknowledge these botanical varieties, subspecies, or species-level (*C. amarus*) classifications for citron melon. However, in as much as they are distinct morphologically from the sweet dessert types, we have chosen to refer to them here as CLC. *Citrullus ecirrhosus* (CE) is a desert perennial, with a distribution limited to southern Africa and can be hybridized with *C. lanatus* (Navot et al., 1990; Sarafis, 1999; Shimotsuma, 1963).

Citrullus colocynthis (CC) is a perennial watermelon species, known as the "bitter apple," endemic to desert soils throughout northern Africa, the Middle East, and southwestern Asia (Burkill, 1985; Dane and Lang, 2004; Dane et al., 2006; Jarret et al., 1997; Paris, 2015; Zamir et al., 1984). Although major reproductive barriers resulting from wide differences in genome structure exist between CC and CLL, these two *Citrullus* species can be crossed with each other (Levi et al., 2006) and CC could be a useful source of genes for enhancing biotic and abiotic stress resistance in the watermelon cultivars (Levi et al., 2016). Several U.S. PIs of CC have been identified as having resistance

Received for publication 6 June 2016. Accepted for publication 28 Nov. 2016.
Mention of product or manufacturer names are for documentation purposes only and do not imply endorsement by the authors or USDA or the participating research institutions.
[1]Coresponding author. E-mail: amnon.levi@ars.usda.gov.

Table 1. Accession number (as shown in Fig. 1), accession name, the *Citrullus* species it belongs to, as designated by the U.S. Department of Agriculture (USDA), Agricultural Research Service (ARS), Germplasm Resources Information Network [GRIN (USDA, 2015)], the genotypic group(s) based on allele frequency found in each accession, as determined in the structure analysis (Fig. 2), and geographic location in which the *Citrullus* accession was collected as indicated at the USDA, ARS, GRIN (USDA, 2015).

Accession no.	Accession name	*Citrullus* species[z]	Genotypic group	Geographic location
1	ARO 21031	CC	CC-1	Negev Desert, Israel
2	ARO 23967	CC	CC-2	Negev Desert, Israel
3	ARO 18917	CC	CC-2	Negev Desert, Israel
4	ARO 18920	CC	CC-2	Negev Desert, Israel
5	ARO 20587	CC	CC-1	Negev Desert, Israel
6	ARO 22357	CC	CC-2	Negev Desert, Israel
7	ARO 22359	CC	CC-2	Negev Desert, Israel
8	ARO 23701	CC	CC-2	Negev Desert, Israel
9	ARO 22555	CC	CC-2	Jordan
10	PI 195927	CC	CLL and CC-3	Harage, Ethiopia
11	PI 220778	CC	CC-3	Farah, Afghanistan
12	PI 346082	CC	CC-3	Helmand, Afghanistan
13	PI 374216	CC	CC-3	Afghanistan
14	PI 386014	CC	CC-3 and CC-2	Iran
15	PI 386015	CC	CC-3, CC-2, and CLL	Iran
16	PI 386016	CC	CC-3, CC-2 and CC-5	Iran
17	PI 386018	CC	CC-2, CC-3 and CC-5	Iran
18	PI 386019	CC	CC-2 and CC-3	Iran
19	PI 386021	CC	CC-2, CC-1, CC-5 and CLL	Iran
20	PI 386024	CC	CC-4	Iran
21	PI 386026	CC	CC-3, CC-1, and CC-4	Iran
22	PI 388770	CC	CC-1	Morocco
23	PI 432337	CC	CC-1 and CLL	Cyprus
24	PI 525080	CC	CC-2, CC-3, CC-4, and CC-5	Qena, Egypt
25	PI 525082	CC	CC-5	Qena, Egypt
26	PI 537277	CC	CC-5, CC-1, CC-3 and CLL	Punjab, Pakistan
27	PI 537300	CC	CC-2 and CC-3	Ahal, Turkmenistan
28	PI 542616	CC	CC-1	Algeria
29	PI 549161	CC	CC-1	Chad
30	GRIF 16135	CLC	CLC, CLL and CC-1	France
31	GRIF 16945	CE	CLC, CLL. CC-1, CC-5	Cape Town, South Africa
32	GRIF 15896	CLC	CLC and CC-5	Russian Federation
33	'Black Diamond'	CLL	CLL, CL-5 and CL-1	U.S.
34	'Charleston Gray'	CLL	CLL	U.S.
35	'Sugar Baby'	CLL	CLL	U.S.
36	PI 482311	CLC	CLC	Zimbabwe
37	PI 482312	CLC	CLC	Zimbabwe
38	PI 596662	CLC	CLC	Transvaal, South Africa
39	PI 596670	CLC	CLC, CC-4 and CC-5	Cape Province, South Africa
40	PI 482252	CLC	CLC and CC-5	Zimbabwe
41	PI 248774	CLC	CLC and CC-4	Namibia
42	PI 482277	CLC	CLC	Zimbabwe
43	PI 169289	CLL	CLL and CC-4	Bursa, Turkey
44	PI 525083	CLC	CLL and CC-1	Egypt
45	PI 249010	CLL	CLL	Nigeria
46	PI 270562	CLC	CLL and CLC	Picketsburg, South Africa
47	PI 162667	CLL	CLL and CO-4	Buenos Aires, Argentina
48	PI 525081	CLC	CLL, CC-1, and CC-5	Qena, Egypt

[z]*C. colocynthis* = CC; *C. lanatus* var. *lanatus* = CLL; *C. lanatus* var. *citroides* (CLC), *C. ecirrhosus* (CE).

to the sweetpotato whitefly [*Bemisia tabaci* (Coffey et al., 2015; Simmons and Levi, 2002)], twospotted spider mite [*Tetranychus urticae* (Cantu, 2014)], powdery mildew [*Podosphaera xanthii* race 2W (Tetteh et al., 2010)], or zucchini yellow mosaic virus (ZYMV) (Guner, 2004). In a recent study, we evaluated the CC PI collection and identified significant levels of resistance to papaya ringspot virus (PRSV) in several accessions collected in northern India and in northern Africa, indicating that CC might be a viable source of resistance to potyviruses (Levi et al., 2016). As a desert plant species, CC is endemic to arid environments,

Table 2. High-frequency oligonucleotides-targeting active genes' (HFO-TAG) primer sequences used in this study. The occurrence (frequency) of HFO-TAG primers in 4700 watermelon expressed sequence-tag unigenes (as shown by Levi et al., 2010), their guanine and cytosine (GC) content (1 = 100%), the number of polymorphic fragments (NPF) produced by each primer among the 45 *Citrullus* U.S. PIs and three watermelon cultivars evaluated in this study, and the size of each of the polymorphic fragments produced by the HFO-TAG primer.

Primer	Sequence	Occurrence (no.)	GC	NPF (no.)	Fragment sizes (bp)
HFO-13	TCCGCCGC	2,226	0.875	25	92, 110, 113, 114, 119, 122, 124, 125, 130, 132, 137, 138, 139, 161, 162, 165, 178, 188, 248, 266, 282, 283, 284, 332, 368
HFO-14	GCGGCGGA	2,226	0.875	18	85, 87, 96, 174, 193, 194, 208, 214, 229, 235, 246, 293, 307, 316, 324, 328, 333, 389
HFO-19	TCGCCGCC	1,991	0.875	24	90, 100, 109, 110, 112, 118, 120, 124, 143, 152, 159, 176, 178, 180, 194, 213, 215, 263, 265, 280, 282, 283, 331, 375
HFO-20	GGCGGCGA	1,991	0.875	9	80, 83, 87, 97, 110, 180, 208, 257, 272
HFO-23	ACGGCGGC	1,796	0.875	18	72, 86, 101, 127, 171, 176, 182, 190, 192, 201, 204, 205, 212, 218, 220, 223, 274, 323
HFO-31	CGCCGCCA	1,756	0.875	5	110, 122, 141, 183, 280
HFO-32	TGGCGGCG	1,756	0.875	13	88, 98, 109, 112, 123, 126, 145, 154, 175, 176, 181, 213, 276
HFO-37	CGGCGCCG	1,482	1	1	191
HFO-44	CGCCGGCG	1,418	1	8	81, 114, 125, 136, 146, 165, 194, 211
HFO-49	GCGGCGGT	1,494	0.875	40	82, 96, 106, 109, 125, 126, 127, 129, 130, 132, 133, 160, 165, 170, 172, 176, 204, 210, 211, 214, 217, 220, 223, 237, 243, 253, 258, 259, 260, 262, 274, 286, 288, 302, 324, 327, 329, 330, 333, 335
HFO-50	ACCGCCGC	1,494	0.875	15	113, 114, 125, 130, 132, 161, 165, 175, 177, 188, 242, 282, 283, 305, 309
HFO-55	CCGCCGCT	1,373	0.875	10	77, 91, 98, 122, 131, 142, 160, 241, 246, 276,
HFO-56	AGCGGCGG	1,373	0.875	9	110, 177, 213, 216, 219, 222, 225, 246, 275
HFO-60	TCGGCGGC	1,263	0.875	19	97, 99, 101, 110, 148, 157, 170, 176, 182, 212, 215, 218, 221, 224, 245, 264, 274, 278, 322
HFO-65	TCCGGCGG	1,213	0.875	13	102, 110, 120, 126, 142, 167, 177, 182, 230, 239, 246, 312, 427
HFO-66	CCGCCGGA	1,213	0.875	13	86, 97, 99, 106, 112, 115, 120, 128, 134, 171, 180, 193, 309
HFO-67	GCCGCTGC	1,098	0.875	27	82, 86, 88, 93, 95, 96, 103, 110, 111, 124, 129, 130, 136, 139, 147, 150, 157, 179, 191, 193, 233, 243, 253, 274, 299, 378, 406
HFO-68	GCAGCGGC	1,098	0.875	23	90, 101, 103, 112, 134, 137, 165, 167, 171, 176, 179, 182, 202, 206, 207, 210, 212, 240, 267, 274, 287, 296, 420
HFO-71	CCACCGCCG	1,237	0.889	32	87, 89, 94, 104, 111, 116, 117, 118, 123, 125, 130, 134, 151, 159, 167, 177, 181, 200, 204, 228, 242, 245, 262, 272, 283, 284, 285, 295, 307, 322, 330, 382
HFO-72	CGGCGGTGG	1,237	0.889	30	79, 88, 95, 101, 102, 104, 107, 122, 124, 132, 152, 158, 160, 169,171, 174, 190, 203, 211, 213, 216, 221, 222, 243, 257, 264, 276, 285, 301, 436
HFO-75	CCGCCGGC	1,011	1	9	82, 106, 109, 116, 140, 188, 319, 323, 324,
HFO-76	GCCGGCGG	1,011	1	38	79, 92, 94, 96, 100, 101, 106, 108, 109, 113, 119, 133, 138, 145, 149, 392 154, 159, 166, 174, 176, 181, 197, 199, 210, 212, 223, 228, 229, 230, 236, 246, 274, 284, 294, 361, 384, 390,
HFO-77	CCTCCGCCG	1,193	0.889	32	94, 95, 101, 103, 104, 114, 116, 118, 121, 124, 125, 131, 132, 148, 156, 158, 161, 163, 166, 188, 195, 215, 216, 227, 232, 250, 262, 283, 285, 334, 368, 382

and its root system rapidly spreads deep into the soil. For this reason, it also could be a useful source for enhancing drought tolerance in watermelon cultivars (Si et al., 2009; Wang et al., 2014).

A previous study using randomly amplified polymorphic DNA markers (Levi et al., 2001a, 2001b) indicated that high levels of genetic diversity exist among CC PIs. In a later study, we developed polymerase chain reaction (PCR) markers referred to as HFO–TAG using PCR primers designed to amplify oligonucleotides that exist in high frequency in expressed sequence-tag unigenes (Levi et al., 2010). The HFO-TAG markers are highly reproducible and polymorphic and are expected to depict genetic relationships based on coding regions (Levi et al., 2010, 2013; Paris, 2015). The objectives of this study were to use HFO-TAG markers to: 1) examine genetic relationships and diversity among CC PIs collected in northern Africa, Asia, and the Middle East and 2) examine genetic distances of the CC PIs from CLL, CLC, and CE.

Materials and Methods

Plant material and isolation of DNA. A total of 48 genotypes were chosen for analysis, including 29 CC PIs, 12 CLC PIs, 3 CLL PIs, 3 CLL cultivars (Charleston Gray, Sugar Baby, and Black Diamond), and one accession representing CE (Table 1). Five seedlings of each genotype were grown in the greenhouse at 26/20 °C (14/10 h day/night). The first true leaf was collected from each of five plants representing each of the 48 genotypes, and was stored at –80 °C for later DNA isolation. Because the PI plants are not derived from a true homozygous

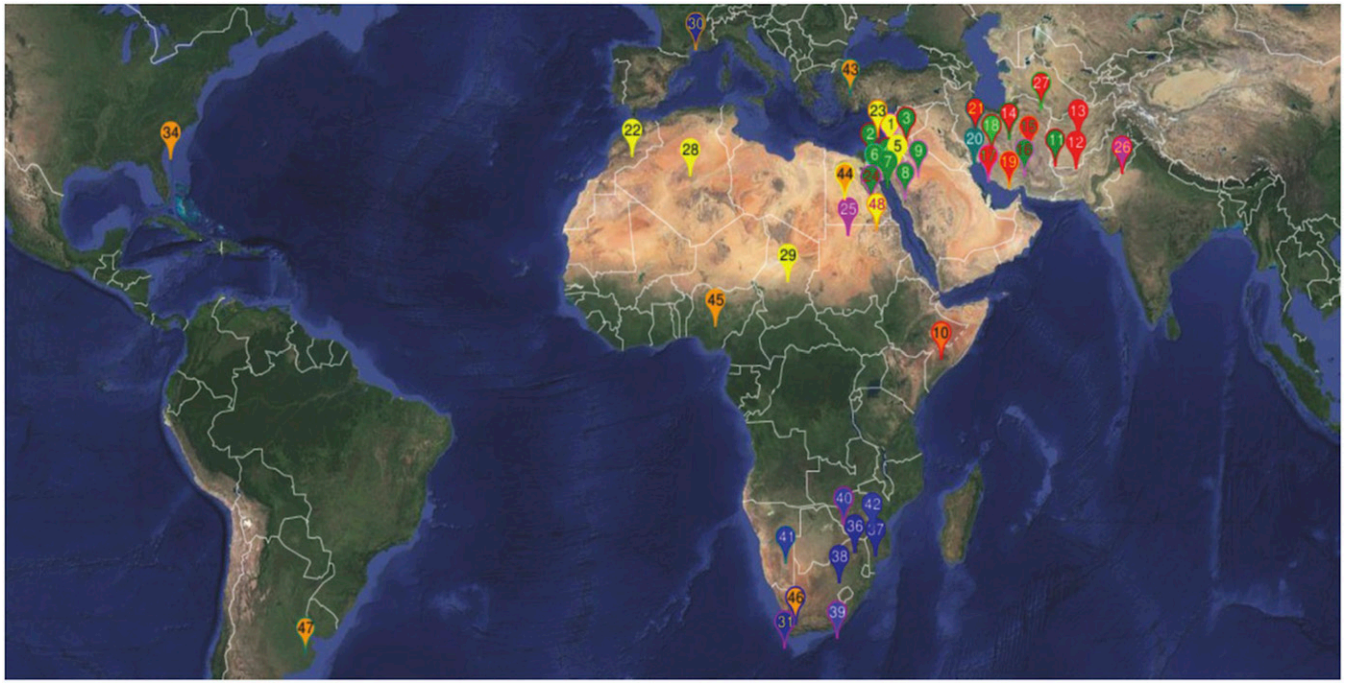

Fig. 1. Geographic locations in which *Citrullus* accessions included in this study were collected with map icon colors corresponding to their group color in the Structure analysis (Falush et al., 2007) in Fig. 2.

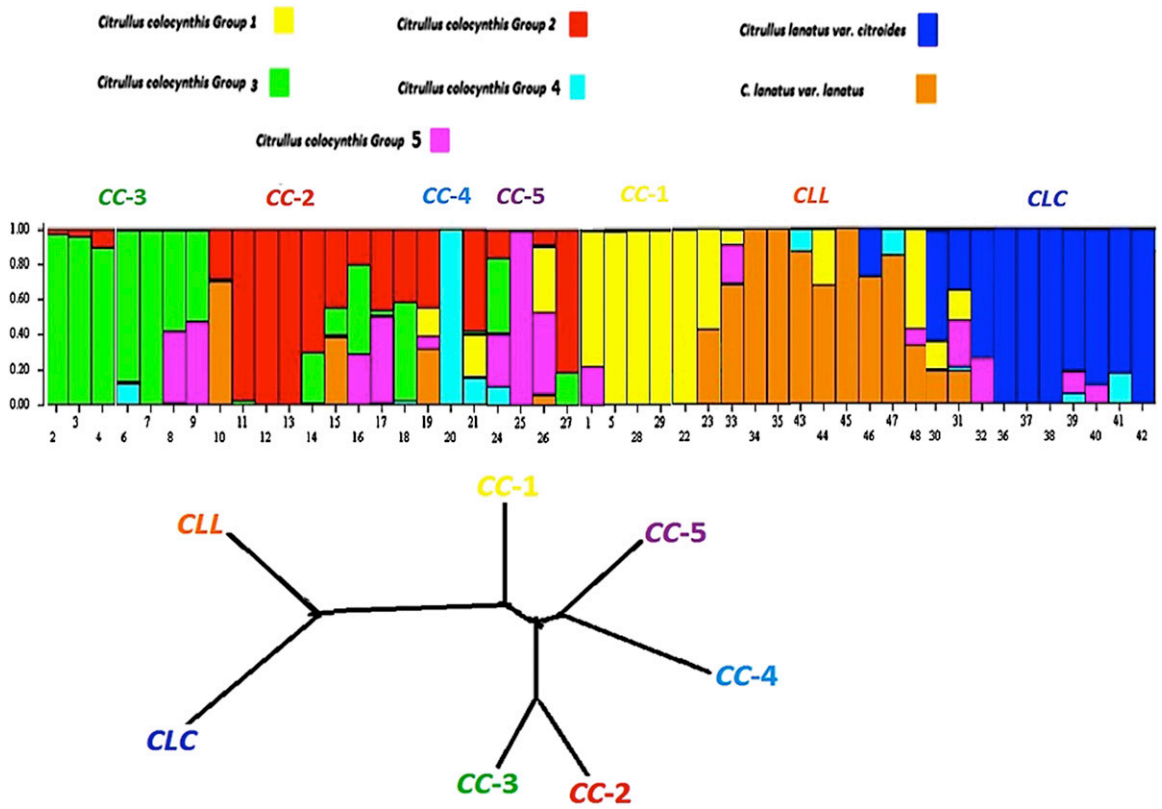

Fig. 2. Population structure analysis resolving seven groups (K = 7). Scale of *y* axis represents probability of log likelihood. Ancestry of the 48 *Citrullus* species genotypes, estimated based on the Structure analysis (Falush et al., 2007). The ancestry from the inferred *Citrullus lanatus* var. *citroides* (CLC) and *Citrullus lanatus* var. *lanatus* (CLL) gene pools are shown in dark blue and mustard color, respectively. The ancestry from the inferred *Citrullus colocynthis* (CC) gene pools are shown in yellow (Group 1), red (Group 2), green (Group 3), light blue (Group 4), and purple (Group 5).

Table 3. U.S. PI accessions of *Citrullus colocynthis* (CC) (and the group they belong to in Fig. 1) and watermelon cultivar used for producing interspecific F_1 hybrid and consequent F_2, and BC_1 seeds in a greenhouse at the U.S. Department of Agriculture, Agricultural Research Service U.S. Vegetable Laboratory, Charleston, SC. The approximate number of viable F_1, F_2, and BC_1 seeds per fruit produced in plants derived from the crosses. Approximate number of seeds is based on two to six fruit collected from two to four plants of each cross in the greenhouse.

CC accession crossed with cultivar	Male parent	Female parent	F_1 seeds (no./fruit)	F_2 seeds (no./fruit)	BC_1 to male seeds (no./fruit)[z]	BC_1 to female seeds (no./fruit)[z]
PI 388770 (CC Group 1)	'Sugar Baby'	PI 388770	30–55	Few	60–100	40–60
	PI 388770	'Sugar Baby'	Few[y]	Few	—[x]	—
PI 525080 (CC Group 1)	'Sugar Baby'	PI 525080	40–75	30–70	50–90	20–40
	PI 525080	'Sugar Baby'	Few	Few	—	—
ARO 25555 (CC Group 2)	ARO 25555	'Charleston Gray'	Few	50–70	40–70	40–60
	'Charleston Gray'	ARO 25555	Few	Few	—	35–60
ARO 22357 (CC Group 2)	'Sugar Baby'	ARO 22357	60–100	Few	Few	Few
	ARO 22357	'Sugar Baby'	Few	50–70	40–70	40–60
PI 537300 (CC Groups 2 and 3)	'Sugar Baby'	PI 537300	60–100	—	—	35–60
	PI 537300	'Sugar Baby'	Few	—	Few	Few
PI 195927 (CC Groups 3 and CLL)	PI 195927	'Charleston Gray'	60–150	60–80	40–70	40–70
	'Charleston Gray'	PI 195927	70–150	Few	—	—
PI 386024 (CC Group 4)	PI 386024	'Charleston Gray'	25–40	Few	30–70	25–40
	'Charleston Gray'	PI 386024	25–40	Few	20–40	25–40
PI 386026 (CC Groups 3, 1, and 4)	PI 386026	'Charleston Gray'	20–50	Few	30–60	25–40
	'Charleston Gray'	PI 386026	25–40	Few	25–40	25–40
PI 537277 (CC Groups 5, 1, 3, and CLL)	'Sugar Baby'	PI 537277	Few	—	Few	—
	PI 537277	'Sugar Baby'	60–100	Few	Few	40–70
PI 386021 (CC Groups 3, 1, 5, and CLL)	'Sugar Baby'	PI 386021	20–35	Few	20–45	20–35
	PI 386021	'Sugar Baby'	Few	—	20–35	20–40

[z]BC_1 using the watermelon cultivar as a male or female recurrent backcross parent.
[y]4–10 seeds.
[x]No seeds.

plant, leaf samples from each of the five PI plants were bulked for DNA extraction to detect minor alleles that may not exist in all plants representing the PI. DNA extractions were conducted using the method described by Levi et al. (2013).

PCR amplification and analysis using HFO-TAG primers. The PCR amplification conditions for the 23 HFO-TAG primers (Table 2) were as described by Levi et al. (2010, 2013). The HFO-TAG markers were scored and analyzed using a DNA analysis system (CEQ 8800; Beckman Coulter, Fullerton, CA). For visualization of DNA fragments, the forward primers were labeled with dye labels (WellRED, D2, D3, or D4; Proligo, Boulder, CO) as previously described (Levi et al., 2009, 2010).

Marker data collection and analysis. The HFO-TAG fragments (Table 2) were scored based on their presence or absence using the vendor-supplied software accompanying the CEQ-8800 system. A similarity matrix for the HFO-TAG data was generated using the Nei–Li similarity index (Nei and Li, 1979). A dendrogram (Fig. 1) was created based on the unweighted pair-group method with arithmetic average using the SAHN module in NTSYS-PC, version 2.02j (Rohlf, 1998). Bootstrap support for clusters was conducted in FreeTree using 5000 permuted datasets (Pavlicek et al., 1999). A cophenetic matrix was generated from the dendrogram by using the COPH module in NTSYS-PC. A population structure analysis procedure and admixture model for clarifying genotypic ambiguity was also used (Fig. 2) by means of the computer program Structure, version 2.200 (Falush et al., 2007; Pritchard et al., 2000), and optimum K value was determined based on L(K) (mean ± SD) graph, estimated using structure harvester.

Plant material for cross-pollination attempts of CC with watermelon cultivars. U.S. PIs of CC and heirloom watermelon cultivars (Charleston Gray and Sugar Baby) were chosen for the interspecific hybridization experiment. Reciprocal cross attempts were carried out in a greenhouse using plants of heirloom cultivars and CC PIs. The greenhouse was maintained at a temperature of 26/20 °C (14/10 h day/night), with a light intensity of 140–175 $\mu mol \cdot m^{-2} \cdot s^{-1}$ with light-emitting diodes (PRO 325; LumiGrow, Novato, CA). All plants were started from seed in standard 50-cell black plastic trays and transplanted into 0.004-m^3 nursery pots using a commercial potting mix consisting of sphagnum peatmoss, pine bark, sand, and lime. Two to four plants of each PI or cultivar or F_1 and BC_1 plants (Table 3) were grown in 12-L pots arranged randomly on benches in the greenhouse and were fertilized biweekly with 150 $mg \cdot L^{-1}$ 20N–8.8P–16.6K water-soluble fertilizer (Peters Professional; Scotts-Sierra, Marysville, OH).

All pollination attempts were carried out between 0700 and 1000 HR, immediately following the collection of three male flowers from the pollen donor parent plant that were used to pollinate female flowers on the recipient plant. The ripe fruit of each plant of the heirloom cultivars and CC PI plants were harvested at 40- to 45-d postpollination.

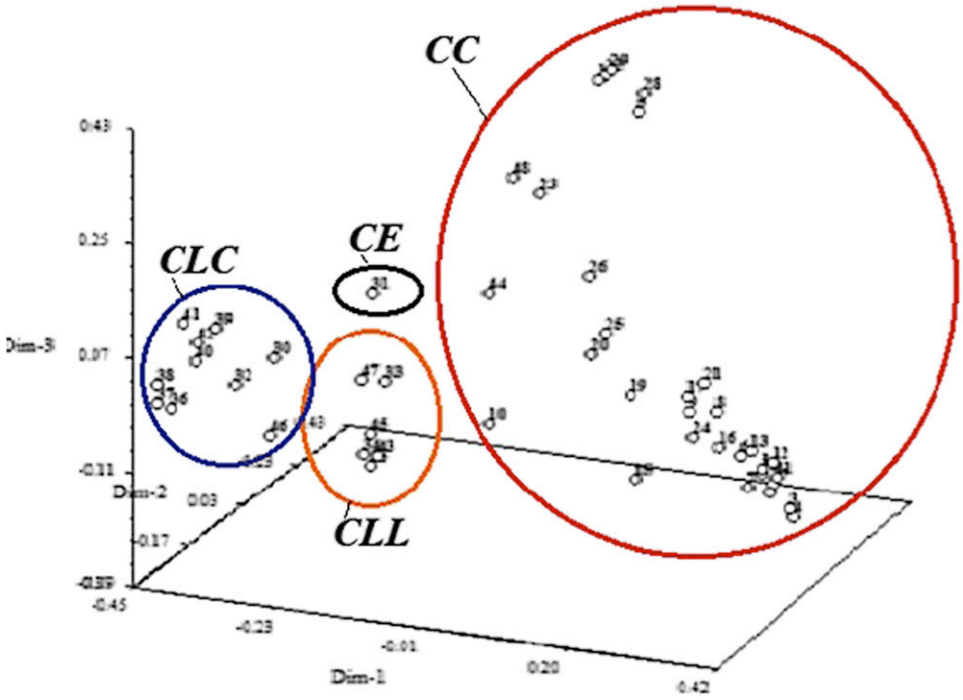

Fig. 3. Two-dimensional plot of *Citrullus colocynthis* (CC), *Citrullus lanatus* var. *lanatus* (CLL), *Citrullus lanatus* var. *citroides* (CLC) accessions using multidimensional scaling based on the 431 polymorphic high-frequency oligonucleotides-targeting active genes markers. The CC accessions are clustered closer to CLL and further from CLC.

Results and Discussion

The main objective of this study was to evaluate genetic diversity among CC PIs collected in northern Africa, the Middle East, and Asia and assess their relationships with CLC and CLL PIs using HFO-TAG primers. The 431 HFO-TAG markers produced by 23 primers, ranged in molecular weight from 70 to 420 bp (Table 2). A large number of the HFO-TAG markers differed by only one or a few nucleotides (Table 2) and could represent forms of the same sequence (Levi et al., 2010, 2011). As in previous studies (Levi et al., 2010, 2013), the HFO-TAG markers proved useful for population structure analysis and for differentiating among closely related genotypes. As depicted in Figs. 1–5, wide genetic and phenotypic diversity exists among CC PIs, while several of the CC PIs share a considerable number of alleles with CLL or CLC.

The population structure analysis differentiated the CC PIs into five distinct groups (Fig. 2) and identified CC genotypes admixed with CLL and/or CLC forms (Figs. 3 and 4). The first major CC group includes PIs collected in northern Africa or in the adjacent Negev Desert, Israel. Several PIs in this first group have a set of alleles unique to CC group 1, whereas other PIs in this group share alleles with CLL PIs (Figs. 1 and 2). The second and third CC groups include PIs collected mainly in the Middle East (the Negev Desert, Israel, Jordan, and Iran). The third group includes CC PIs collected in Iran and Afghanistan. The fourth and the fifth groups are represented by CC PIs 386024 and 525082, collected in Iran and Egypt, respectively (Table 1; Figs. 2 and 3). Each of these latter two PIs may represent isolated CC populations with unique alleles. However, in a previous study evaluating genetic diversity among CLC PIs collected in southern Africa, PI 386024 and PI 525082 were shown to have alleles of both CC and CLL (Levi et al., 2013). The difference in the allele profile in the present study vs. that in Levi et al. (2013) is likely the result of using different HFO-TAG primers, which amplify different genomic regions. Still, the analysis in this study classified both PI 386024 and PI 525082 as CC (Figs. 2–4), as described by Levi et al. (2013).

While several CC PIs have a higher number of unique alleles, most of the CC PIs are admixed and share alleles of the different CC groups or of CLL and CLC (Figs. 2 and 3). Dane et al. (2006) examined genetic relationships among CC accessions using a chloroplast DNA and a single-copy nuclear gene sequence, and reported the existence of several CC groups, suggesting the possible migration of CC from the African continent into the Middle East and Asia. Here, the HFO-TAG fragments mostly represent coding region alleles, and facilitated the identification of distinct CC groups, in congruence with the chloroplast and the single-copy nuclear gene sequence results of Dane et al. (2006).

The genetic analysis presented here suggests that several of the CC genotypes might be more closely related to CLL than to CLC or CE (Figs. 2–4). Several of the CC PIs collected in northern Africa or in the Middle East share alleles with the CLL genotypes (Fig. 2). PI 386015 and PI 386021 (collected in Iran), PI 195927 (collected in Ethiopia), and PI 432337 (collected in Cyprus), share a large number of alleles with CLL PIs (Fig. 2). It should be noted that the latter two PIs are classified as CC on the Germplasm Resources Information Network [U.S. Department of Agriculture (USDA), 2015], but are clustered here in the CLL group (Figs. 2–4). Two PIs (PI 525081 and PI 525083) (collected in Qena, Egypt) are classified as CLC on the Germplasm Resources Information Network (USDA, 2015). However, in this study they appeared to be admixture of CC and CLL alleles (Table 1; Figs. 2 and 3). In a previous study evaluating genetic relationships among CLC PIs (Levi et al., 2013), PI 525081 and PI 525083 also showed an admixture of CLL and CC alleles and were clustered with the CLL group. Here, PI 525083 was clustered with the CLL group, as shown in an earlier study (Levi et al., 2013). However, PI 525080 was clustered with the CC group collected in northern Africa (Figs. 2–5). Because most of the HFO-TAG primers used in this study are different from these used in our previous study (Levi et al., 2013), they likely reveal different gene sequences, and because of the high number of admixed CC and CLL alleles, the classification of PI 525081 is intermediate between these two *Citrullus* species.

Several CC accessions, including ARO 23701, ARO 22555, PI 286016, PI 386018, PI 386024, PI 386026, PI 525080, PI 525082, and PI 537277 share alleles with CLC and/or with CE (GRIF 16945). It is worth noting that this CE accession, collected in South Africa, is clustered together with CLC, but

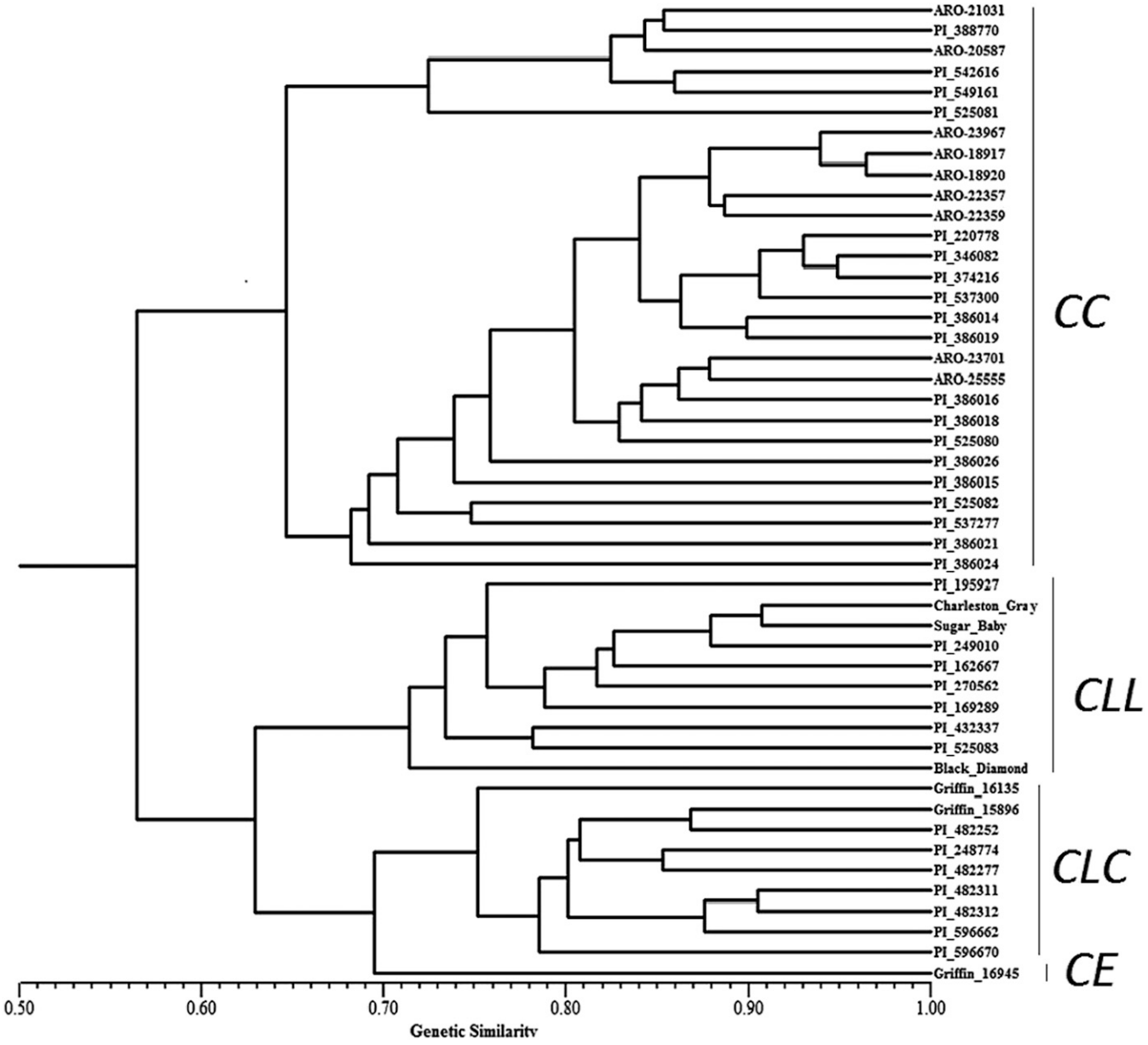

Fig. 4. An unweighted pair group method with arithmetic mean dendrogram, based on 431 high-frequency oligonucleotides-targeting active genes markers, revealing the major groups of accessions including *Citrullus colocynthis* (CC), *Citrullus lanatus* var. *lanatus* (CLL), *Citrullus lanatus* var. *citroides* (CLC), and *Citrullus ecirrhosus* (CE).

still distinct from all other genotypes of this form (Figs. 2–4). This CE accession comprises alleles present in CLC, but also alleles present in CLL, and in CC [Groups 1, 2, and 3 (Fig. 2)]. These results indicate the possibility of alleles in an ancient common ancestor of CC, CLL, CLC, and CE which facilitated parallel or convergent evolution of analogous features (Lorts et al., 2008; Williams et al., 2013), vital for adaptation of CE and CC in the deserts of southern and northern Africa, respectively. Dane and Liu (2007) examined relationships among *Citrullus* species using chloroplast DNA sequences and suggested that CLC and CLL "have split from a common ancient ancestor followed by haplotype fixation." Our study using HFO-TAG markers (Levi et al., 2013) indicates that the CLC PIs collected in southern Africa are differentiated into two distinct groups based on different sets of alleles.

There is wide phenotypic variation in leaf shape of the CC PIs. The narrow serrated leaves and sharp lobes of CC resemble these of CLL, adapted to dry conditions. Most of the CLC PIs collected in southern Africa are adapted to milder conditions and consequently have wide leaves with wide-rounded lobes (Fig. 5). The leaves of CE are small and thick adapted to the desert conditions in southern Africa.

The CLC and CLL share the same reproductive features and are readily crossed with each other to produce fertile progeny using traditional breeding procedures (Levi et al., 2011, 2013). In contrast, crosses of CLC or CLL with CC show some directionality and frequently result in significantly reduced fertility of progeny (Levi et al., 2006) (Table 3). Still, wide differences in gene sequences (Guo et al., 2013) exist between CLC and CLL. These differences are the result of evolutionary

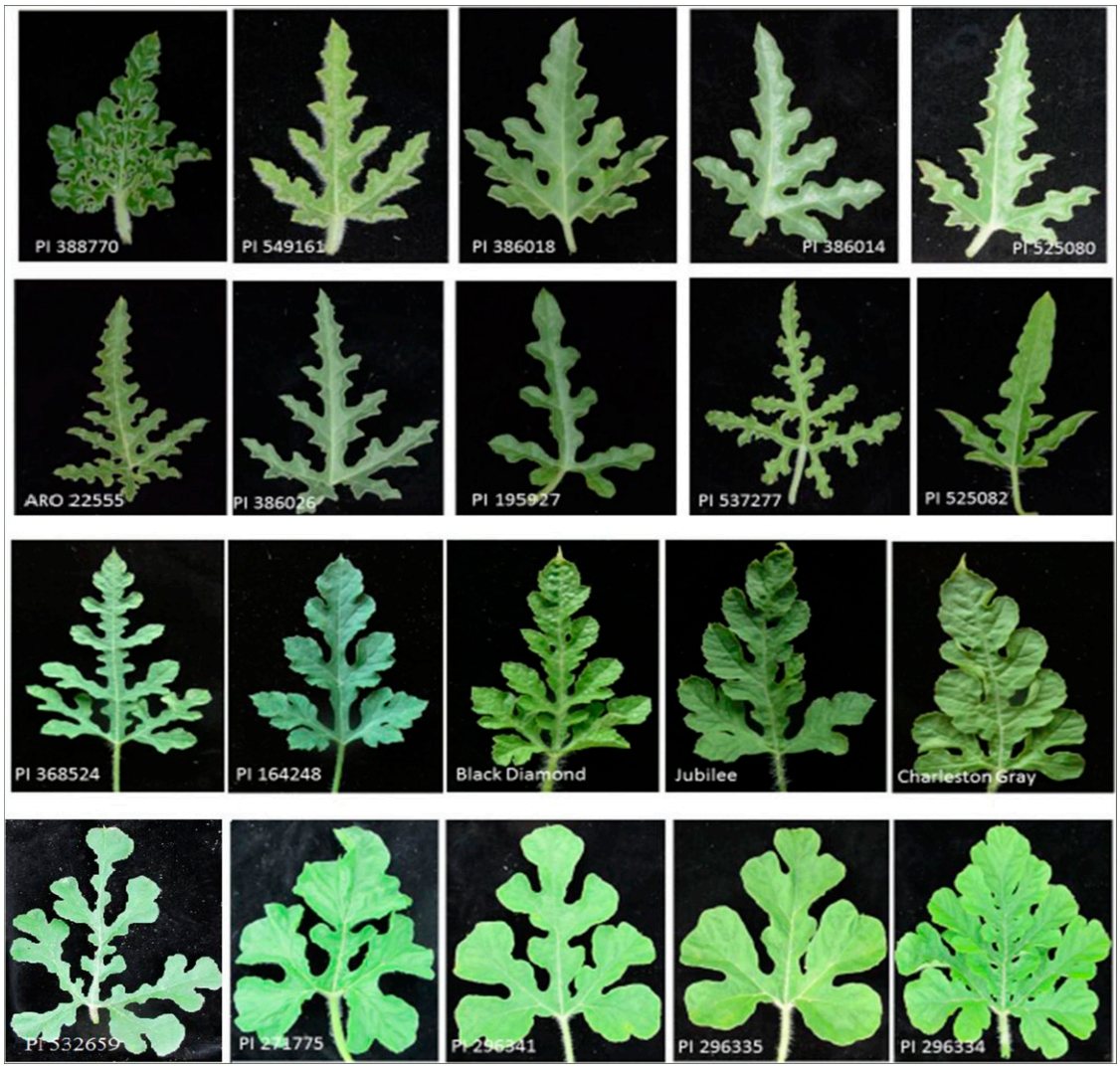

Fig. 5. Leaf samples of plants representing *Citrullus colocynthis* [CC (upper two rows)], *Citrullus lanatus* var. *lanatus* [CLL (third row from top)], and *Citrullus lanatus* var. *citroides* [CLC (bottom row)]. Leaves of CC have a narrow and pointed shape and serrated lobes with a dry waxy surface. Leaves of CLL are also narrow and pointed but with larger serrated lobes. On the other hand, CLC leaves have wide and round lobes.

events, but also the result of selective sweeps of chromosomal regions associated with the cultivation of CLL (Nimmakayala et al., 2014; Reddy et al., 2014, 2015). Our previous study using HFO-TAG markers indicated a close relationship of CC to watermelon cultivars [CLL (Levi et al., 2013)]. The CC PIs that show resistance to whiteflies (Coffey et al., 2015; Simmons and Levi, 2002), spider mites, or PRSV (Levi et al., 2016) are being crossed with heirloom watermelon cultivars (Table 3) with the objective of transferring that resistance into a CLL genomic background.

To overcome reproductive barriers between CC and watermelon cultivars, we conducted cross-pollination tries under controlled environmental conditions in the greenhouse. Plants of watermelon cultivars produced one fruit while the CC PI plants produced two to three fruit in the greenhouse. Plants of watermelon cultivars (Charleston Gray or Sugar Baby) or CC PIs [PI 388770, PI 525080, PI 537300, PI 195927, PI 386021, PI 386024, PI 386026, PI 537277, ARO 22357 (Table 1; Fig. 2)] were self-pollinated in the greenhouse and produced 50–170 seeds. At the same time, cross-pollination attempts of watermelon cultivars (CLL) with CC accessions produced variable numbers of F_1 or BC_1 seeds and a few or no F_2 seed (Table 3), indicating that reproductive barriers exist between these two *Citrullus* species. The interspecific pollination attempts produced viable F_1 and BC_1 plants in cross-pollinations of watermelon cultivars and CC accessions (Table 3). In a previous study, we were able to produce reciprocal crosses between watermelon cultivars and CC PIs, and released watermelon breeding lines (BC_6F_6 and BC_8F_8) containing the mitochondrial and chloroplast genomes of CC and most of the nuclear genome of their recurrent parent cultivar (Levi et al., 2006, 2010).

In the present study, we were able to produce F_1, but not F_2 seeds in most CLL × CC crosses in the greenhouse. F_1 plants derived from crossing 'Sugar Baby' with CC PI 537277 were recalcitrant to produce F_2 or BC_1 seeds in the greenhouse. However, when the same F_1 plants were placed in the field, they produced a large number of fruit (25–70 fruit per plant) with 7–18 viable seeds per fruit (Fig. 6). F_1 plants derived from crossing 'Charleston Gray' and CC PI 525080 (collected in northern Africa) produced sufficient number of F_2 and BC_1 seeds. These results indicate that reproductive barriers exist

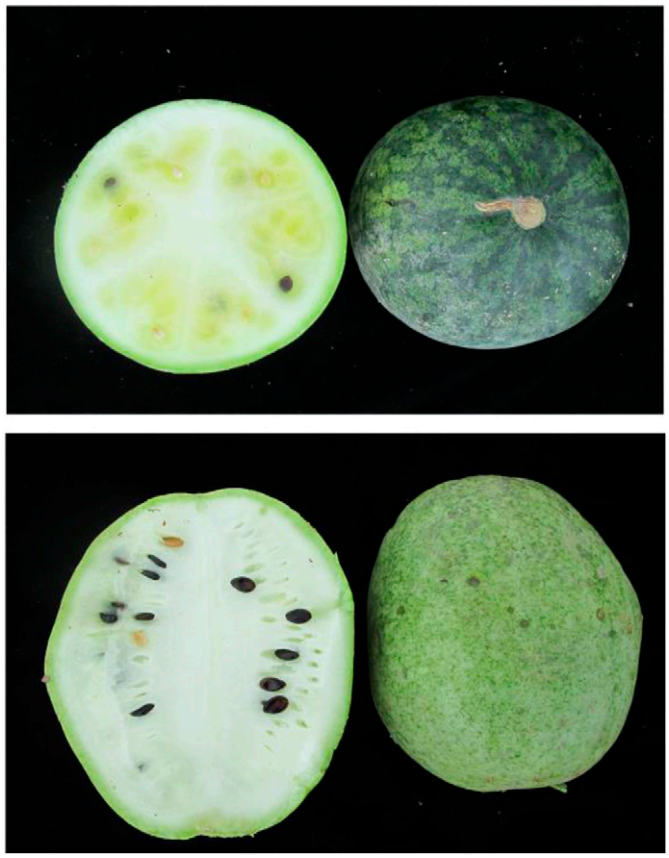

Fig. 6. Mature fruit of an F_1 hybrid (*Citrullus colocynthis* ♂ PI 537277 × 'Sugar Baby' ♀) with only a few seeds [7–18 seeds/fruit (upper)], and a fruit of an F_1 (*C. colocynthis* ♂ PI 525082 × 'Charleston Gray' ♀) plant containing a large number of seeds [60–120 seeds/fruit (bottom)]. Plants were open pollinated in the field in Charleston, SC during Summer 2015.

between watermelon cultivars and CC PIs collected in the deserts of southern Asia (PI 537277) compared with the CC PIs collected in northern Africa (PI 525080). The reproductive barriers are likely the results of wide differences in genome structure between watermelon cultivars and CC PIs (Guo et al., 2013; Reddy et al., 2013). The PIs collected in northern Africa might be more easily used in breeding programs aimed to enhance watermelon with disease or pest resistance [e.g., both PI 525080 and PI 537277 exhibit resistance to PRSV (Levi et al., 2016)]. Studies combining genetic analyses and advanced genomic sequencing technologies (Lambel et al., 2014) are needed to identify gene loci in CC that could be useful for enhancing watermelon cultivars. Additional studies are needed to estimate pollen viability and compatibility of plants representing the different CC groups, to better assess their use in efforts to enhance genetic diversity within and among watermelon cultivars.

Overall, the five CC groups are distinct from CLC and CLL (Figs. 3 and 4). Each of the five CC groups contains unique alleles, but also shares alleles with CLL, CLC and/or CE, implying evolution from a common ancestor. The question of whether CC evolved from CLL or from a common ancestor of CE, CLL, and CLC, remains to be determined. Rapid advances in next generation sequencing technologies and the possibility of sequencing and assembling the genomes of a large number of genotypes should provide higher resolution for assessing genetic relationships among *Citrullus* species and subspecies.

Literature Cited

Bailey, L.H. 1930. Three discussions in *Cucurbitaceae*. Gentes Herbarum 2:175–186.

Burkill, H.M. 1985. The useful plants of west tropical Africa. 2nd ed. Royal Botanic Gardens, Kew, UK.

Cantu, H. 2014. An evaluation of watermelon (*Citrullus* spp.) germplasm for additional sources of resistance to the two-spotted spider mite (*Tetranychus urticae* Koch). MS Thesis, Univ. of Nebraska, Lincoln, NE.

Chomicki, G. and S. Renner. 2014. Watermelon origin solved with molecular phylogenetics including Linnaean material: Another example of museomics. New Phytol. 205:526–532.

Coffey, J.L., A.M. Simmons, B.M. Shepard, Y. Tadmor, and A. Levi. 2015. Potential sources of whitefly (Hemiptera: Aleyrodidae) resistance in desert watermelon (*Citrullus colocynthis*) germplasm. HortScience 50:13–17.

Dane, F. and P. Lang. 2004. Sequence variation at cpDNA regions of watermelon and related wild species: Implications for the evolution of *Citrullus* haplotypes. Amer. J. Bot. 91:1922–1929.

Dane, F., J. Liu, and C. Zhang. 2006. Phylogeography of the bitter apple, *Citrullus colocynthis*. Genet. Resources Crop Evol. 54:327–336.

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

Falush, D., M. Stephens, and J.K. Pritchard. 2007. Inference of population structure using multilocus genotype data: Dominant markers and null alleles. Mol. Ecol. Notes 7:574–578.

Fursa, T.B. 1972. On the taxonomy of the genus *Citrullus* Schad. Bot. Z. 57:31–34.

Guner, N. 2004. *Papaya ringspot virus* watermelon strain and *Zucchini yellow mosaic virus* resistance in watermelon. North Carolina State Univ., Raleigh, PhD Thesis.

Guo, S., J. Zhang, H. Sun, J. Salse, W.J. Lucas, H. Zhang, Y. Zheng, L. Mao, Y. Ren, Z. Wang, J. Min, X. Guo, F. Murat, B.K. Ham, Z. Zhang, S. Gao, M. Huang, Y. Xu, S. Zhong, A. Bombarely, L.A. Mueller, H. Zhao, H. He, H. Zhang, Z. Zhang, S. Huang, T. Tan, E. Pang, K. Lin, Q. Hu, H. Kuang, P. Ni, B. Wang, J. Liu, Q. Kou, W. Hou, X. Zou, J. Jiang, G. Gong, K. Klee, H. Schoof, Y. Huang, X. Hu, S. Dong, D. Liang, J. Wang, K. Wu, Y. Xia, X. Zhao, Z. Zheng, M. Xing, X. Liang, B. Huang, T. Lv, J. Wang, Y. Yin, H. Yi, R. Li, M. Wu, A. Levi, X. Zhang, J.J. Giovannoni, J. Wang, Y. Li, Z. Fei, and Y. Xu. 2013. The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. Nat. Genet. 45:51–58.

Jarret, R.L., L.C. Merrick, T. Holms, J. Evans, and M.K. Aradhya. 1997. Simple sequence repeats in watermelon (*Citrullus lanatus* (Thunb.) Matsum. & Nakai). Genome 40:433–441.

Jarret, R. and I. Levy. 2012. Oil and fatty acid contents in seed of *Citrullus lanatus* Schrad. J. Agr. Food Chem. 60:5199–5204.

Jeffrey, C. 2001. Cucurbitaceae, p. 1510–1557. In: P. Hanelt (ed.). Mansfeld's encyclopedia of agricultural and horticultural crops. 3. Springer, Berlin, Germany.

Lambel, S., B. Lanini, E. Vivoda, J. Fauve, W.P. Wechter, K. Harris-Shultz, L. Massey, and A. Levi. 2014. A major QTL associated with *Fusarium oxysporum* race 1 resistance identified in genetic populations derived from closely related watermelon lines using selective genotyping and genotyping-by-sequencing for SNP discovery. Theor. Appl. Genet. 127:2105–2115.

Levi, A., C.E. Thomas, A.P. Keinath, and T.C. Wehner. 2001a. Genetic diversity among watermelon (*Citrullus lanatus* and *Citrullus colocynthis*) accessions. Genet. Resources Crop Evol. 48:559–566.

Levi, A., C.E. Thomas, T.C. Wehner, and X. Zhang. 2001b. Low genetic diversity indicates the need to broaden the genetic base of cultivated watermelon. HortScience 36:1096–1101.
Levi, A., C.E. Thomas, J.A. Thies, A.M. Simmons, K. Ling, and H.F. Harrison, Jr. 2006. Novel watermelon breeding lines containing chloroplast and mitochondrial genomes derived from the desert species *Citrullus colocynthis*. HortScience 41:463–464.
Levi, A., P. Wechter, and A. Davis. 2009. EST-PCR markers representing watermelon fruit genes are polymorphic among watermelon heirloom cultivars sharing a narrow genetic base. Plant Genet. Resour. 7:16–32.
Levi, A., W.P. Wechter, K.R. Harris-Shultz, A.R. Davis, and Z. Fie. 2010. High-frequency oligonucleotides in watermelon expressed sequenced tag-unigenes are useful in producing polymorphic polymerase chain reaction markers among watermelon genotypes. J. Amer. Soc. Hort. Sci. 135:369–378.
Levi, A., W.P. Wechter, L.M. Massey, L. Carter, and D. Hopkins. 2011. Genetic linkage map of *Citrullus lanatus* var. *citroides* chromosomal segments introgressed into the watermelon cultivar Crimson Sweet (*Citrullus lanatus* var. *lanatus*) genome. Amer. J. Plant Sci. 2:93–110.
Levi, A., J.A. Thies, W.P. Wechter, H.F. Harrison, A.M. Simmons, U.K. Reddy, P. Nimmakayala, and Z. Fei. 2013. High frequency oligonucleotides: Targeting active gene (HFO-TAG) markers revealed wide genetic diversity among *Citrullus* spp. accessions useful for enhancing disease or pest resistance in watermelon cultivars. Genet. Resources Crop Evol. 60:427–440.
Levi, A., J. Coffey, L. Massey, N. Guner, E. Oren, Y. Tadmor, and K. Ling. 2016. Resistance to *Papaya ringspot virus*-watermelon strain (PRSV-W) in the desert watermelon *Citrullus colocynthis*. HortScience 51:4–7.
Lorts, C., T. Briggeman, and T. Sang. 2008. Evolution of fruit types and seed dispersal: A phylogenetic and ecological snapshot. J. Syst. Evol. 46:396–404.
Navot, N., M. Sarfatti, and D. Zamir. 1990. Linkage relationships of genes affecting bitterness and flesh colour in watermelon. J. Hered. 81:162–165.
Nei, M. and W.H. Li. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA 1979:5269–5273.
Nimmakayala, P., A. Levi, L. Abburi, V.L. Abburi, Y.R. Tomason, T. Saminathan, V.G. Vajja, S. Malkaram, R. Reddy, T.C. Wehner, S.E. Mitchell, and U.K. Reddy. 2014. Single nucleotide polymorphisms generated by genotyping by sequencing to characterize genome-wide diversity, linkage disequilibrium, and selective sweeps in cultivated watermelon. BMC Genomics 15:767.
Paris, HS. 2015. Origin and emergence of the sweet dessert watermelon, *Citrullus lanatus*. Ann. Bot. (Lond.) 116:133–148.
Pavlicek, A., S. Hrda, and J. Flegr. 1999. Free-Tree: Freeware program for construction of phylogenetic trees on the basis of distance data and bootstrap/jackknife analysis of the tree robustness. Application in the RAPD analysis of genus *Frenkelia*. Folia Biol. (Praha) 45:97–99.
Pritchard, J.K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.
Reddy, U.K., N. Aryal, N. Islam-Faridi, Y. Tomason, A. Levi, and P. Nimmakayala. 2013. Cytomolecular characterization of rRNA gene sequences among *Citrullus* species and subspecies using fluorescent in situ hybridization (FISH) technology. Genetica 60:427–440.
Reddy, U.K., P. Nimmakayala, A. Levi, V.L. Abburi, T. Saminathan, Y.R. Tomason, G. Vajja, R. Reddy, L. Abburi, T.C. Wehner, Y. Ronin, and A. Karol. 2014. High-resolution genetic map for understanding the effect of genome-wide recombination rate on nucleotide diversity in watermelon. Genes Genomes Genet. 4:2219–2230.
Reddy, U.K., L. Abburi, V.L. Abburi, T. Saminathan, R. Cantrell, V.G. Vajja, R. Reddy, Y.R. Tomason, A. Levi, T.C. Wehner, and P. Nimmakayala. 2015. A genome-wide scan of selective sweeps and association mapping of fruit traits using microsatellite markers in watermelon. J. Hered. 106:166–176.
Rohlf, F.J. 1998. NTSYS-PC numerical taxonomy and multivariate analysis system, ver. 2.2. Exeter Publ., Setauket, NY.
Sarafis, V. 1999. Cucurbit resource s in Namibia, p. 400–402. In: J. Janick (ed.). Perspectives on new crops and new uses. ASHS Press, Alexandria, VA.
Shimotsuma, M. 1963. Cytogenetic and evolutionary studies in the genus *Citrullus*. Seiken Ziho 15:24–34.
Si, Y., C. Zhang, S. Meng, and F. Dane. 2009. Gene expression changes in response to drought stress in *Citrullus colocynthis*. Plant Cell Rep. 28:997–1009.
Simmons, A.M. and A. Levi. 2002. Sources of whitefly (Homoptera: Aleyrodidae) resistance in *Citrullus* for the improvement of cultivated watermelon. J. Amer. Soc. Hort. Sci. 3:581–584.
Tetteh, A.Y., T.C. Wehner, and A.R. Davis. 2010. Identifying resistance to powdery mildew race 2W in the USDA-ARS watermelon germplasm collection. Crop Sci. 50:933–939.
The Plant List. 2013. The Plant List. Version 1.1. 28 Nov. 2016. <http://www.theplantlist.org/>.
U.S. Department of Agriculture. 2015. Germplasm Resources Information Network. 28 Nov. 2016. <http://www.ars-grin.gov>.
Wang, Z., H. Hu, L.R. Goertzen, J.S. McElroy, and F. Dane. 2014. Analysis of the *Citrullus colocynthis* transcriptome during water deficit stress. PLoS One 9:e104657, doi: 10.1371/journal.pone.0104657.
Wehner, T.C. 2008. Watermelon, p. 381–418. In: J. Prohens and F. Nuez (eds.). Handbook of plant breeding; Vegetables I: Asteraceae, Brassicaceae, Chenopodiaceae, and Cucurbitaceae. Springer, New York, NY.
Whitaker, T.W. and W.P. Bemis. 1976. Cucurbits. *Cucumis, Citrullus, Cucurbita, Lagenaria*, p. 64–69. In: N.W. Simmonds (ed.). Evolution of crop plants, Longman, London, UK.
Williams, B.P., I.G. Johnston, S. Covshoff, and J.M. Hibberd. 2013. Phenotypic landscape inference reveals multiple evolutionary paths to C4 photosynthesis. eLife 2:E00961.
Zamir, D., N. Navot, and J. Rudich. 1984. Enzyme polymorphisms in *Citrullus lanatus* and *Citrullus colocynthis* in Israel and Sinai. Plant Syst. Evol. 146:163–170.