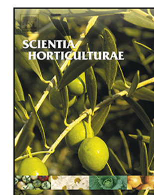
Scientia Horticulturae 195 (2015) 124–137

Contents lists available at ScienceDirect

Scientia Horticulturae

journal homepage: www.elsevier.com/locate/scihorti



Genetic analysis of citron (*Citrus medica* L.) using simple sequence repeats and single nucleotide polymorphisms

Chandrika Ramadugu[a,*], Manjunath L. Keremane[b], Xulan Hu[c], David Karp[a], Claire T. Federici[a], Tracy Kahn[a], Mikeal L. Roose[a], Richard F. Lee[b]

[a] Dept. of Botany and Plant Sciences, University of California, Riverside, CA, USA
[b] USDA-ARS, National Clonal Germplasm Repository for Citrus and Dates, Riverside, CA, USA
[c] Yunnan Provincial Dept. of Agriculture, Kunming, Yunnan, China

ARTICLE INFO

Article history:
Received 23 July 2015
Received in revised form 31 August 2015
Accepted 2 September 2015
Available online 20 September 2015

Keywords:
Citron
Microsatellite analysis
Single nucleotide polymorphism
Population structure

ABSTRACT

Citron (*Citrus medica* L.) is one of the three basic species of the genus *Citrus* L. that have contributed to the development of cultivated citrus. We analyzed the genetic diversity of 47 citrons (32 from Yunnan Province, China; and 15 of Mediterranean origin) to understand their diversity and relationships within the species. Genetic analysis was conducted using data from microsatellite markers, single nucleotide polymorphisms generated from sequences of a nuclear malate dehydrogenase gene and a chloroplast gene, *rps16*. Neighbor joining and maximum parsimony analyses were conducted. All three approaches found citron to be monophyletic. Population structure analysis clustered the 47 citrons into three distinct groups. The first group consisted of wild, non-fingered citrons generally having locules, juice sacs and seeds within the fruit. The second cluster consisted mostly of fingered citrons that lacked locules, juice sacs or seeds, and some non-fingered types with smaller locules and vestigial juice sacs, but with seeds. All accessions that clustered in groups I and II originated in China. The third cluster consisted of citrons cultivated in the Mediterranean region. Genetic distance between the clusters from population structure analysis indicated considerable diversity within the species. A citron-specific microsatellite marker was identified and characterized. We observed considerable heterozygosity in certain citrons, contrary to previous reports.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Citrus medica L. (citron) is the type species of the genus *Citrus*, and along with pummelo (*Citrus maxima* (Burm.) Merr.) and mandarin (*Citrus reticulata* Blanco) is considered to be one the three basic ancestral species of the genus *Citrus* (Barrett and Rhodes, 1976; Hodgson, 1967; Mabberley, 1997; Pang et al., 2007; Scora, 1975). Kumquat (*Fortunella* spp.) and papeda (*Citrus* spp.), along with the three basic species of citrus have contributed to the development of the vast majority of modern cultivated citrus (Nicolosi et al., 2000; Scora, 1975 and 1988). The greatest diversity of citron genotypes currently occurs in southwestern China, particularly in Yunnan Province, Northeastern India, and Southeast Asia (Hazarika, 2012; Hodgson, 1967; Gmitter and Hu, 1990). Concrete evidence regarding citron's area of origin is lacking, but it has generally been considered likely that the citron originated in Southeast Asia (Andrews, 1961; Gmitter and Hu, 1990; Kumar et al., 2010; Nair and Nayar, 1997; Scora, 1975). Citrons are valued for horticultural, medicinal, and religious uses and have been selected for desirable characteristics. In its place of origin, citron genotypes are usually referred to by local vernacular names. Citrons are monoembryonic and hence production of apomictic seeds is presumably absent (Hodgson, 1967). Most citrons are known to be self-compatible and hence hybrids with citron as the maternal parent are not common (Moore, 2001; Scora, 1975). Virtually all fingered citrons are totally seedless and would not have survived without human intervention. The pharmaceutical and food value of fingered citrons has resulted in constant selection for desirable characters and the generation of a large number of cultivars. The origin of these cultivars and their genetic relationships within the group remain unclear. Genetic characterization of citrons can help identify distinct genotypes for preservation and characterization.

* Corresponding author.
E-mail addresses: chandram@ucr.edu, ramaduguc@gmail.com (C. Ramadugu), Manjunath.keremane@ars.usda.gov (M.L. Keremane), xulan.hu@163.com (X. Hu), david.karp@ucr.edu (D. Karp), claire.federici@ucr.edu (C.T. Federici), tracy.kahn@ucr.edu (T. Kahn), Mikeal.roose@ucr.edu (M.L. Roose), rfleevirus@yahoo.com (R.F. Lee).

http://dx.doi.org/10.1016/j.scienta.2015.09.004
0304-4238/© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Citron rind consists of an outer pericarp with a glandular, aromatic, pigmented flavedo, and an internal firm white albedo. The endocarp has many locules filled with juice vesicles and 5–100 seeds around a central columella (Swingle and Reece, 1967). Some citron fruits are unusual compared to other citrus because of the complete absence of locules, juice vesicles, and seeds. Large citron fruits up to 50 cm in length, weighing up to 5.5 kg are common (Hodgson, 1967). Two of the authors of this article (Karp and Hu) found specimens of 'Ninger Giant' citron fruit weighing about 15 kg in Yunnan. In certain citrons, the fruit has an unusual shape and appears like a hand with fingers, unlike common citrus fruits (Hodgson, 1967). In some, the carpels develop separately and the whole fruit appears fingered, while other taxa have projections at the distal end only. Based on the type of endocarp, three groups of citrons are identified: acid or sour, acidless or sweet, and, pulpless and dry. In the pulpless group, the juice vesicles are either completely absent or very few and vestigial. In Yunnan, a dozen or more fingered citron varieties are cultivated for culinary or medicinal use. Although representative citron taxa have been traditionally included in most studies of citrus phylogeny, inadequate sampling of the group has resulted in insufficient understanding of the genetic diversity of the species (Barrett and Rhodes, 1976; Herrero et al., 1996a,b); very few studies have included more than 2 or 3 citrons (Barkley et al., 2006; Luro et al., 2012).

The objectives of this study were: (1) To study genetic diversity of a large number of citrons from Yunnan along with representative citrons from the Mediterranean region. We utilized DNA-based markers, microsatellites (simple sequence repeats or SSR) and single nucleotide polymorphisms (SNP) in nuclear and chloroplast (cp) genes, to conduct phylogenetic analyses. (2) To determine relationships among the citrons from Yunnan and the accessions originally cultivated in the Mediterranean region by studying population structure. (3) To evaluate the heterozygosity of the citron accessions and compare the findings with previous reports. The present study is significant as it is one of the few genetic studies analyzing a large number of citrons (47); many are collected from Yunnan, an important center of origin for the species, and several of the accessions are novel types that have not been studied before. The information generated regarding genetic diversity will be useful for preservation of genetically distinct accessions.

2. Materials and methods

2.1. Source of plant materials

A total of 47 citrons along with 7 non-citron controls (2 each of pummelos and papedas, 1 each of orange, mandarin, and kumquat) were included in the study. The taxa (CM1–CM58) and their presumed places of origin are listed in Table 1. Fruits of these citrons range in size from small to gigantic, have smooth or rough textures, are either fingered, round, oblong, ridged or long. Internal fruit characters range from juicy and seedy to pulpless and seedless (Fig. 1). Samples for CM1–26 were obtained as silica gel-dried leaf tissue from Yunnan; fresh tissue for CM27–30 was collected from the National Clonal Germplasm Repository for Citrus and Dates, Riverside, CA, USA. The open pollinated seeds for raising seedlings of CM27–30 were obtained from Yunnan. Accessions labeled CM31-50 and CM56–58 were acquired by the Citrus Variety Collection (CVC), University of California at Riverside (UCR), CA over a period of 100 years from many citrus-growing regions of the world (http://www.citrusvariety.ucr.edu/citrus; http://www.ars-grin.gov/). Accessions CM51, 53–55 were etrog-type citrons obtained from Lindcove Ranch, Exeter, CA, which cultivates citrons for sale for religious use. The Chinese collection activities were coordinated through Southwest Forestry University, Kunming, Yunnan. Budwood of the source plants was forwarded to the citrus collection, Citrus Research Institute of Chinese Academy of Agricultural Science, Beibei, Chongqing, P.R. China.

2.2. Extraction of DNA and PCR amplifications

Genomic DNA was extracted using the cetyl trimethyl ammonium bromide (CTAB) protocol (Doyle and Doyle, 1987). The extraction buffer consisted of 100 mM Tris buffer pH 8.0 containing 2% CTAB, 1.4 M NaCl, 20 mM EDTA, 2% polyvinyl pyrrolidone and 0.2% β-mercaptoethanol. 1–2 g of silica dried leaf tissue (CM1–26) or fresh leaf tissue (CM27-58) was ground to a fine powder and homogenized with four ml of extraction buffer. 40 μg of RNase A was added to the extractions, incubated at 65 °C for 30 min, kept on ice for 10 min, centrifuged at 13,000 rpm for 10 min, and the clear supernatant was extracted twice with an equal volume of chloroform:isoamylalcohol (24:1::v/v). DNA in the aqueous fraction was precipitated with an equal volume of cold isopropanol on ice. The pelleted DNA was dissolved in 1X TE (10 mM Tris, 1 mM EDTA, pH 8.0) and quantified using a Nanodrop 1000 UV–vis spectrophotometer (Thermo Fisher Scientific Inc., Wilmington, DE, USA).

PCR amplification of a 1600 bp fragment of the nuclear gene, malate dehydrogenase (MDH) was conducted using primers Cit 637 (5′GCTCCTGTGGAAGAGACCC) and Cit 653 (5′TTAACGATAGTTTCGGTAGAC) as described previously (Ramadugu et al., 2013). Amplification of chloroplast DNA was conducted using universal primers rpsF (5′GTGGTAGAAAGCAACGTGCGACTT) and rpsR2R (5′TCGGGATCGAACATCAATTGCAAC) designed to amplify a 941 bp fragment containing a group II intron from the *rps16* gene fragment (Oxelman et al., 1997). PCR amplification of the microsatellite marker was conducted using primers designed based on the CF-AT13 SSR primers (Table 2). Reaction conditions were similar to MDH PCR; annealing was at 54 °C for 45 s for *rps16* and at 60 °C for 45 s for CF-AT13.

2.3. Cloning and sequencing of PCR products

The products from MDH and CF-AT13 PCRs were electrophoresed on agarose gels, amplicons were purified using a QiaEX II kit (Qiagen Inc., Valencia, CA), cloned into pCR4 Topo vector (Invitrogen, Carlsbad, CA) and sequenced using vector-based or internal primers. The sequences were aligned using Clustal W (Thompson et al., 1994) and displayed using GeneDoc software (Nicholas et al., 1997).

2.4. Amplification of microsatellite loci and gel electrophoresis

Simple sequence repeats (SSR) or microsatellites were amplified using primers shown in Table 2. PCR conditions, denaturing of the amplicons and electrophoresis on denaturing polyacrylamide gels were essentially as described (Barkley et al., 2006). Annealing temperatures for PCR varied from 45 to 58 °C, depending on the primers. The LI-COR gel images were visualized using Adobe Photoshop software (Fig. 2). Polymorphism information content (PIC) was calculated according to standard procedures (Nagy et al., 2012).

2.5. qPCR analysis

A quantitative real-time PCR (qPCR) analysis was conducted using representative citrus samples in an ABI ViiA 7 qPCR system (Applied Biosystems) as a preliminary test to evaluate the discriminatory power of the primers and to determine the optimal annealing temperature for amplification. ABI SYBR Green Master Mix was used to conduct PCR with template DNA and 100 nmoles of SSR primers. The cycling parameters were: 95 °C for 20 s (1

Table 1
Taxa used for the study.

No.	Code	Name of variety/cultivar group	Scientific name	Current location (presumed origin)
1	CM1	'Jinhua Qingpi' fingered citron	*Citrus medica* var. *sarcodactylis* (Hoola van Nooten) Swingle	Jinhua, Zhejiang province, China (China)
2	CM2	'Jinhua Dwarf' fingered citron	*Citrus medica* var. *sarcodactylis* (Hoola van Nooten) Swingle	Jinhua, Zhejiang province, China (China)
3	CM3	'Chuan' fingered citron	*Citrus medica* var. *sarcodactylis* (Hoola van Nooten) Swingle	Muchuan, Sichuan province, China (China)
4	CM4	'Guang' fingered citron	*Citrus medica* var. *sarcodactylis* (Hoola van Nooten) Swingle	Yangshuo, Guangxi, China (China)
5	CM5	'Goucheng' citron	*Citrus medica* var. *yunnanensis* S.Q. Ding	Weishan, Yunnan province, China (China)
6	CM6	'Octopus' fingered citron	*Citrus medica* var. *sarcodactylis* (Hoola van Nooten) Swingle	Qinghua, Weishan, Yunnan province, China (China)
7	CM7	'Maanshan' fingered citron	*Citrus medica* var. *sarcodactylis* (Hoola van Nooten) Swingle	Maanshan, Weishan, Yunnan province, China (China)
8	CM8	'Jinghong Water' citron	*Citrus medica* L.	Jinghong, Yunnan province, China (China)
9	CM9	'Yun' fingered citron	*Citrus medica* var. *sarcodactylis* (Hoola van Nooten) Swingle	Huaxi, Huaning, Yunnan province, China (China)
10	CM10	'Chuanjie' fingered citron	*Citrus medica* var. *sarcodactylis* (Hoola van Nooten) Swingle	Chuanjie, Lufeng, Yunnan province, China (China)
11	CM11	'Chuanjie Round' citron	*Citrus medica* L.	Chuanjie, Lufeng, Yunnan province, China (China)
12	CM12	'Honghe' papeda	*Citrus hongheensis* Y.M. Ye et al.	Yuanjiang, Yunnan province, China (China)
13	CM13	'India' lemon hybrid	*Citrus* spp.	Mengdian, Ruili, Yunnan province, China (China)
14	CM14	'Ruili Wild Fl' citron	*Citrus medica* L.	Forestry institute, Ruili, Yunnan province, China (China)
15	CM15	'Ruili Wild La' citron	*Citrus medica* L.	Laohuichan, Ruili, Yunnan province, China (China)
16	CM16	'Ruili Wild Hu' citron	*Citrus medica* L.	Hui huang, Ruili, Yunnan province, China (China)
17	CM17	'Mangshi Wild' citron	*Citrus medica* L.	Hexinchan, Mangshi, Yunnan province, China (China)
18	CM18	'Ruili Sour' pummelo	*Citrus maxima* (Burm.) Merr.	Huihuan, Ruili, Yunnan province, China (China)
19	CM19	'Yunmao Oval' citron	*Citrus medica* L.	Mangshi, Yunnan province, China (China)
20	CM20	'Suanmaliu' citrus hybrid	*Citrus* spp.	Shimaogang, Puer, Yunnan province, China (China)
21	CM21	'Ninger Giant' citron	*Citrus medica* L.	Toudaohe, Ninger, Yunnan province, China (China)
22	CM22	'Weishan Sweet' citron	*Citrus medica* L.	Lotus hill, Weishan, Yunnan province, China (China)
23	CM23	'Weishan Sour' citron	*Citrus medica* L.	Lotus hill, Weishan, Yunnan province, China (China)
24	CM24	'Tuanshan' fingered citron	*Citrus medica* var. *sarcodactylis* (Hoola van Nooten) Swingle	Tuanshan, Weishan, Yunnan province, China (China)
25	CM25	'Fist' fingered citron	*Citrus medica* var. *sarcodactylis* (Hoola van Nooten) Swingle	Yinchan, Qinghua, Weishan, Yunnan province, China (China)
26	CM26	'Bullet' citron	*Citrus medica* L.	Xiyao, Qinghua, Weishan, Yunnan, China (China)
27	CM27	'Weishan Sour' OPS citron	*Citrus medica* L.	Riverside (OPS[c] from Yunnan province, China)
28	CM28	'Ninger Giant' OPS citron	*Citrus medica* L.	Riverside (OPS from Yunnan province, China)
29	CM29	'Persistent Stigma' OPS citron	*Citrus medica* L.	Riverside (OPS from Yunnan province, China)
30	CM30	'Yunmao Oval' OPS citron	*Citrus medica* L.	Riverside (OPS from Yunnan province, China)
31	CM31	'Hart's Tardiff' Valencia orange	*Citrus sinensis* (L.) Osbeck	CRC570[a] (obtained from CA; origin India)
32	CM32	'Siamese Sweet' pummelo	*Citrus maxima* (Burm.) Merr.	CRC2240 (origin Thailand)
33	CM33	'Ponkan' mandarin	*Citrus reticulata* Blanco	CRC3849 (nucellar seedling; origin is India/China)
34	CM34	'Etrog' citron	*Citrus medica* L.	CRC3891; PI 508,265 (Israel)
35	CM35	'Buddha's Hand' fingered citron	*Citrus medica* var. *sarcodactylis* (Hoola van Nooten) Swingle	CRC3768 (China)
36	CM36	'Hanayu'	*Citrus hanaju* Siebold	CRC3469 (imported as OPS from Japan)
37	CM37	"Nagami" kumquat	*Fortunella margarita* Lour. (Swingle)	CRC3877; (China)
38	CM38	'Assads' citron	*Citrus medica* L.	CVC; (Morocco)
39	CM39	'Citron of Commerce'	*Citrus medica* L.	CRC3518 (Corsica)
40	CM40	'Corsican'citron	*Citrus medica* L.	IV No[b]. 3536 (Corsica)
41	CM41	'Diamante' citron	*Citrus medica* L.	IV No. 3539 (Italy)
42	CM43	'Yemen Temoni' citron	*Citrus medica* L.	IV No. 8477 (Israel)
43	CM44	'Italian' citron	*Citrus medica* L.	IV No. 9232 (Italy)
44	CM45	'Mexican' citron	*Citrus medica* L.	CRC3531 (origin unknown)
45	CM46	'Papuan' citron	*Citrus medica* L.	CRC3532; (origin unknown)
46	CM48	'Yunnanese' citron	*Citrus medica* L.	CRC3798 (China)
47	CM50	'Morning Song Temoni' citron	*Citrus medica* L.	Greenhouse, Riverside, CA (Yemen)
48	CM51	'Braverman' citron	*Citrus medica* L.	Jewish nursery in CA (Israel)
49	CM53	'Halperin' citron	*Citrus medica* L.	Jewish nursery in CA (Israel)
50	CM54	'Kivelevitz' citron	*Citrus medica* L.	Jewish nursery in CA (Israel)
51	CM55	'Temoni' citron	*Citrus medica* L.	Jewish nursery in CA (Yemen)
52	CM56	'Hiawassie' citron	*Citrus medica* L.	CRC3527 (origin unknown)
53	CM57	'Unnamed' citron CRC3819	*Citrus medica* L.	CRC3819; PI 539,440 (origin unknown)
54	CM58	'Unnamed' citron CRC3174	*Citrus medica* L.	CRC3174; PI 230,626 (origin Morocco)

Variety/cultivar names and the code used for the accessions in the manuscript are shown. Presumed origin of the cultivar, where known, is indicated in the last column.

[a] Accessions with CRC identification numbers are located in Citrus Variety collection, Riverside, CA, USA.

[b] Accessions with IV numbers are currently found in the greenhouses at University of California, Riverside, CA, USA.

[c] Open pollinated seedlings are designated as OPS.

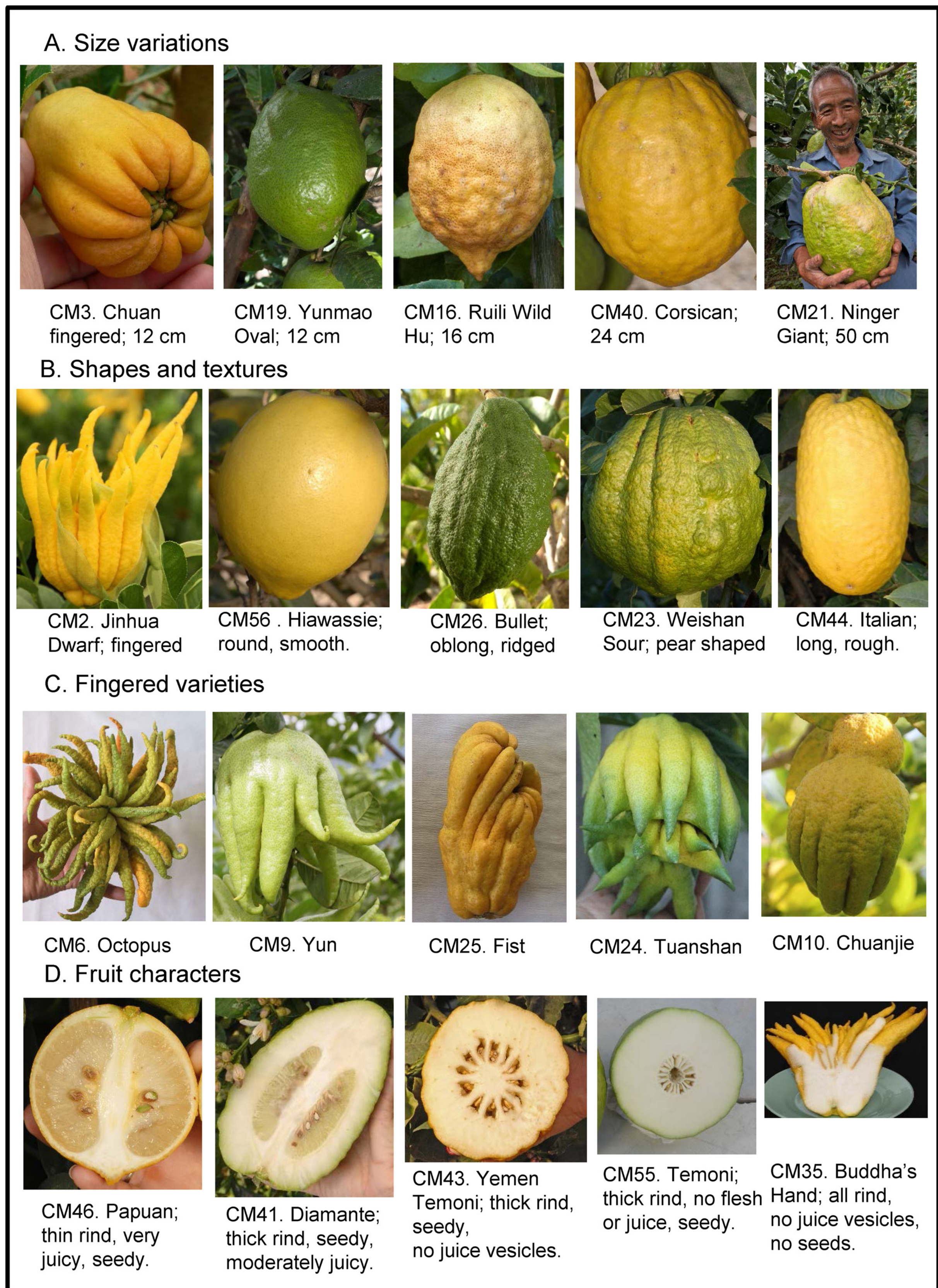


Fig. 1. Morphological diversity of representative citrons.

Table 2
Primers used for microsatellite analysis. Forward and reverse primer sequences are shown.

No.	SSR primer name	Sequence of SSR primer; forward and reverse.	M13 F/R tail	Scaffold number and genomic coordinates[a]	No. of alleles	Size range of amplicons	PIC[d]	Repeat unit	Reference/primer designed by
1	CF-AG06	F: TGTTTTGCTTTGTGCATGGT; R: ACCATGCAAGGAGTTTCCAC	F	3:49,318,530–49,318,711	6	198–212	0.43	AG	Roose lab
2	CF-CTC04	F: CGGCTGGTTACTTGGTTCAT; R: GATTCTGGTGCCTTGGTGAT	R	2:29,433,113[b]	6	163–186	0.57	CTC	Roose lab
3	JI-CTT01	F: TCGCTCTACTCCAATGGCTT; R: GGCTTTATGTCCGATTCTGC	F	7:615,708–615,910	5	224–241	0.41	CTT	Roose lab
4	JI-AC03	F: AAGGGTTACCACCATCACCA; R: AGCCACAACCACCAAACAAT	F	3:6,656,705[b]	12	210–242	0.65	AC	Roose lab
5	CF-CAG06	F: AGCAACCACAGCAACAACAG; R: TCTGAAGTGGGAGAGAGGGA	F	7:7,400,372[b]	6	206–219	0.56	CAG	Roose lab
6	CF-ACA01	F: ACAATGGATTCATTCCTCGC; R: TCGATTTCGAGCACTCCTCT	F	4:14,526,184–14,526,365	3	186–198	0.40	ACA	Roose lab
7	CX6F19	F: ATTCTCATGTCATGCGTACCTCG; R: TGAATCGTGAGAGACGAGTTGAAG	R	4:24,521,165 and3:3,622,202[c]	6	150–184	0.51		Chen et al., (2006).
8	CF-GA07	F: CACAGTCACATAGCACATGCC; R: CAACGTTCCAGTCTTGACGA	R	3:7,854,245-7,854,554	9	289–336	0.67	GA	Roose lab
9	CF-GT02	F: AATAAAACCGTTGGGCTGTG; R: GCATAAGGCAAGTGAAGGGA	R	8:31,559–31,759	6	196–224	0.75	GT	Roose lab
10	CF-AAG15	F: TTCTTCTGCGGAACAAGTGA; R: AGCCAATGGTAGCTCAAAGC	F	5:32,919,716 –32,919,893	6	191–208	0.57	AAG	Roose lab
11	CF-TC07	F: CTTTTTGCAAACCCTTCTGG; R: TCGTCGTCAAAGATCACAGG	R	7:3,432,915 –3,433,106	4	199–211	0.46	TC	Roose lab
12	CF-AT13	F: TGCACAGAAAGCATGGACTC; R: AATGGTTACACGAAGGGACG	F	7:2,357,131–2,357,320	13	201–379	0.68	AT	Roose lab
13	CF-TA03	F: TGGTGGTCGATTTAAGGAGG; R: TTGCGCATCATCAGATCAAT	R	4:584,371[b]	13	200–256	0.65	TA	Roose lab
14	CF-CCT01	F: ATCAAGGTCGCAGCTGAAGT; R: AGATTGAACTGATGGCCCTG	F	5:39,929,084–39,929,266	5	194–205	0.44	CCT	Roose lab
15	CF-TCA03	F: ACAACGGCAACAAGTCCTTC; R: CGAACACAACGCAAAAGCTA	R	8:831,750 –831,953	8	207–229	0.63	TCA	Roose lab
16	JI-AC1	F: TTTATTCACCGCTCAAGGACT; R: TTAGGGGTGGAAAACATGGA	F	2:24,501,500–24,501,711	4	227–240	0.34	AC	Roose lab
17	JI-TCT1	F: CAATCAACTTTCCCACCACC; R: ACCAAGGAGCAGGCTGACTA	R	9:16,971,669[b]	8	236–261	0.41	TCT	Roose lab
18	JI-GAA02	F: CGTGTCGCTCAAGAAAATGA; R: TTCGCTGAAGAGCATGTCAC	F	2:29,599,027–29,599,173	8	165–198	0.56	GAA	Roose lab
19	JK-cAGG9	F: AATGCTGAAGATAATCCGCG; R: TGCCTTGCTCTCCACTCC	R	5:12,849,747[b]	7	126–145	0.28	CAGG	Kijas et al. (1997).
20	NB-CT19	F: ACCACTCACTACTTAATTACCCTTTT; R: AGGGTTGCCACGATTTGTAG	F	2:10,566,017–10,566,165	8	147–186	0.33	CT	Roose lab
21	JK-CAC23	F: ATCACAATTACTAGCAGCGCC; R: TTGCCATTGTAGCATGTTGG	F	3:209,668–209,915	5	261–271	0.23	CAC	Kijas et al. 1997.
22	NB-CTT01	F: TCAGACATTGAGTTGCTCG; R: TAACCACTTAGGCTTCGGCA	R	2:23,985,179–23,985,325	5	153–182	0.53	CTT	Barkley et al. (2006).
23	CMS07	F: CAGGATGCTTGTTGGTGATG; R: ACAGTGGATACAAACATGCTGC	F	1:3,443,674 –3,443,824	8	158–183	0.62	CT	Ahmad et al. (2003).

M13 forward primer (CACGACGTTGTAAAACGAC) and M13 reverse primer sequences (GGATAACAATTTCACACAGG) were used as a tail to the forward SSR primer sequence and are indicated by F and R. Number of alleles per primer, polymorphism information content and microsatellite repeat regions are shown.

[a] Genomic coordinates based on Citrus clementina haploid mandarin genome sequence (www.phytozome.org).

[b] Genomic coordinates found for only one SSR primer.

[c] Genomic coordinates for the two primer sequences were found on two different scaffolds in the Citrus clementina genomic sequence.

[d] Polymorphism information content (PIC).

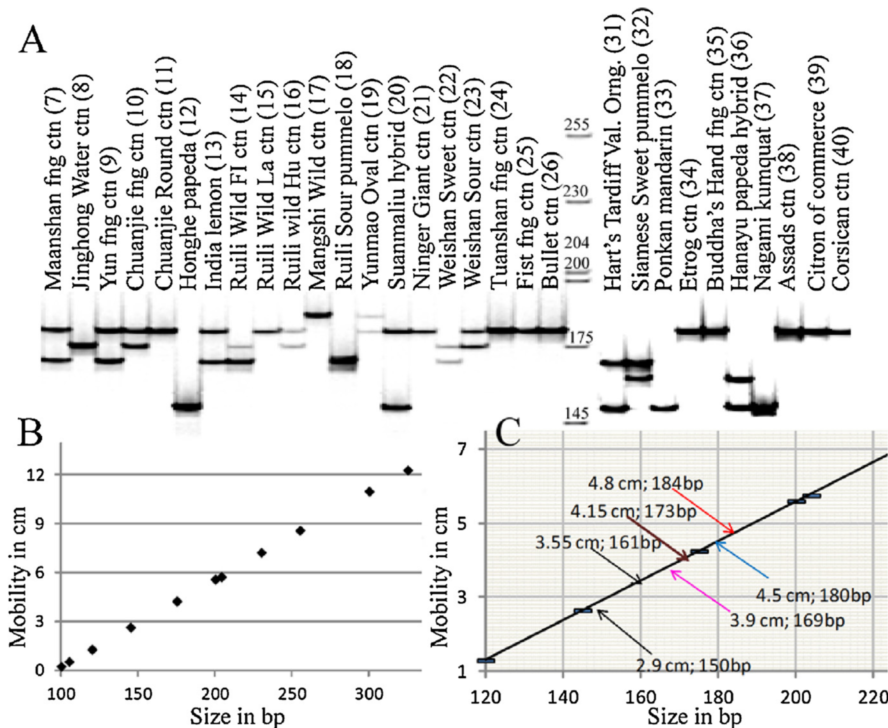


Fig. 2. LI-COR gel image of samples amplified using CX-6F19 primer pair. Panel A: a representative portion of a 7% denaturing polyacrylamide gel. Molecular weight markers (50–350 bp, labeled with IRDye 700 obtained from LI-COR) loaded in the middle. Panel B: mobilities of molecular weight marker fragments were calculated, a scatter plot was created and the approximate sizes of amplicons were calculated using Microsoft Excel. Panel C: size estimation of bands based on scatter plot data indicated by arrows. Horizontal bars on the trend line indicate molecular weight markers (120, 145, 175, 200 and 204 bp). "Fingered" and "citron" abbreviated as fng and ctn. CM numbers used for accessions in the study are indicated in parenthesis.

cycle), 40 cycles of 95 °C for 1 s, annealing at variable temperatures (45–60 °C) and dissociation (65–95 °C) for 1 min.

2.6. Citron population structure analysis

Data generated for 47 accessions (citrons and citron hybrids) from 23 discriminating nuclear microsatellite markers was used to conduct population structure analysis using STRUCTURE 2.3 (Pritchard et al., 2000). The parameters considered for this analysis were: (a) Admixture model (assumes that each individual may have a part of the genome from each of the K populations) with the option that assumes that "allele frequencies are correlated" among populations (Pritchard et al., 2000; Falush et al., 2003). (b) Based on test runs, the allele frequency prior, lambda, was set to 0.62. (c) Number of populations tested (K) varied from 2 to 8. (d) For each K value, multiple runs, each consisting of 2.5 million iterations, were conducted. The first 0.5 million iterations were discarded as "burnin." (e) Stable and consistent results were obtained when K value was three; at $K=3$, the Dirichlet parameter for calculating the degree of admixture, alpha, varied between 0.069 and 0.078. (f) The recommended probability value, Pr (X/K) is a negative value very close to zero (Pritchard et al., 2000). (g) Although the approximate locations of the microsatellite markers are known for most regions (based on the haploid Clementine mandarin genome, listed in Table 2; available from http://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Cclementina; for the current analysis, we did not consider inter-marker distances. (h) STRUCTURE program can utilize prior population information to cluster individuals; however, we did not pre-define groups to avoid possible bias.

2.7. Phylogenetic analysis

Data obtained from three different approaches was used for phylogenetic analysis: (a) SSR data from 23 markers was utilized for constructing neighbor joining trees using Mega version 6.0; (b) *rps16* sequences (901 bp) were used to construct maximum parsimony (MP) trees; (c) sequences from the nuclear gene MDH (1600 bp) were used for MP analysis.

At least eight clones were sequenced for each sample for all three sequenced gene fragments (MDH, *rps16* and CF-AT13). To confirm SNPs and rule out polymerase error, DNA from two citron samples was used to conduct 10 separate PCR reactions for the MDH gene, products were cloned and 96 colonies were sequenced. The sequences were analyzed using MEGA 6.0. (Tamura et al., 2013). Although most accessions had two distinct haplotype sequences, if the SNPs were not parsimony informative, only one sequence was used for construction of the MP tree. If the two haplotypes of one sample clustered in two separate clades during preliminary analysis, we included both sequences for phylogenetic analysis. Subtree-pruning-regrafting algorithm (Nei and Kumar, 2000) was utilized with 1500 iterations for bootstrap analysis (Felsenstein, 1985). All sites and gaps were included in the MDH-SNP analysis.

For the chloroplast gene *rps16* analysis, an MP tree was constructed using MEGA 6.0 with 54 sequences generated in this study along with 14 other *rps16* sequences. Bootstrap analysis included 500 replicates and subtree-pruning-regrafting algorithm (Tamura et al., 2013; Nei and Kumar, 2000; Felsenstein, 1985) similar to MDH analysis. All positions involving gaps and missing data were eliminated from the rps16 dataset.

2.8. Analysis of citron-specific marker sequences

Amplicons generated using CF-AT13 primers were purified, and cloned into pCR4-TOPO cloning vector. Sequences were analyzed using Clustal W (Thompson et al., 1994), GeneDoc (Nicholas et al., 1997) and Sequencher 5.0 (Gene Codes Corporation, Ann Arbor, MI). For certain accessions, two amplicons corresponding to the

two haplotypes were cloned and sequenced. For others, only one fragment was cloned. The sequences were aligned and displayed using GeneDoc software.

3. Results

3.1. Microsatellite marker analysis

Amplification of SSRs was conducted using a total of 40 primer pairs distributed throughout the citrus genome. Data from 23 primer pairs that were able to discriminate the citron accessions was utilized for the study (Supplementary Table 1 shows amplicon sizes). In total 161 putative alleles were detected in the dataset, with an average of 7 alleles per locus. The number of unique alleles recorded was 49 for all taxa and 6 for citrons only (3 in putative hybrid citron accessions). Five of the six unique citron alleles were observed in accessions from China. Microsatellite loci CF-AT13 and CF-TA03 had a maximum of 13 alleles each; locus CF-ACA01 had 3 alleles, the lowest number recorded for this dataset (Table 2). The amplicon sizes ranged from 126 to 379 bp. The maximum number of alleles (43) was observed in 'Hart's Tardiff' orange. Among citrons, the total number of alleles in the microsatellite analysis ranged from 22 ('Ruili Wild Hu') to 35 ('Citron of Commerce'). We repeated each analysis with at least three independent PCRs and gels for each marker. The discriminating power of each microsatellite marker was determined by calculating the PIC value (Nagy et al., 2012). For the primers used in this study, the PIC value ranged from 0.23 (JK-CAC23) to 0.75 (CF-GT02) (Table 2). Thirteen of the 23 loci were considered to be informative, since they had a PIC value greater than 0.5 (Ram et al., 2007).

Initial testing of the SSR primers in qPCR analysis was useful in selecting markers for the analysis of citrons and determining optimum temperatures for amplification. The dissociation peaks at the end of the run were valuable in determining if all the samples had similar dissociation (indicating the possible absence of variation) or if the primers would be discriminatory in SSR analysis.

3.2. Population structure

Analysis of microsatellite marker data for citrons using the program STRUCTURE (Pritchard et al., 2000) indicated the presence of three distinct clusters (Fig. 3). Cluster 1 consisted of many Chinese wild citrons and certain putative natural hybrids. All the fingered citrons and certain non-fingered Chinese citrons grouped in cluster 2. Most other citrons presumed to be from the Mediterranean basin grouped together in cluster 3. Average distances between individuals in each cluster were: 0.5349 (cluster 1), 0.2192 (cluster 2) and 0.3178 (cluster 3). Mean values of Fst (fixation index) for the three clusters were: 0.0012 (cluster 1), 0.4990 (cluster 2) and 0.4864 (cluster 3).

3.3. Neighbor joining tree based on microsatellite marker data

A consensus neighbor joining (NJ) tree was generated from the SSR data (Fig. 4). Three population clusters observed in the NJ tree were similar to the STRUCTURE analysis (Fig. 3), as expected. All the accessions with citron ancestry, including the hybrids 'Suanmaliu' and 'India' lemon hybrid, formed a distinct clade with a bootstrap value of 84%. However, most citron sub-clades did not have good bootstrap support. Citrons from China formed two clades and the etrog-type of citrons formed a separate clade. Fingered and non-fingered citrons did not necessarily form separate groups. Some fingered citrons formed tight clusters. Cultivars 'Kivelevitz' and 'Etrog' formed a well-supported cluster (Fig. 4). Among the non-citrons in the dataset, the two pummelos formed a very distinct cluster. Papedas, mandarin and sweet orange formed a well-supported clade. 'Nagami' kumquat and the unnamed citron CRC3819 formed a separate clade. Based on the distance matrix created, the distance between 'Yun' fingered citron (representative for cluster 2) and 'Ruili Wild Hu' citron (cluster 1) was 0.5; between 'Morning Song Temoni' citron (cluster 3) and 'Ruili Wild Hu', the distance was 0.44; between 'Yun' fingered citron and 'Morning Song Temoni' citron, the distance was 0.39.

3.4. Heterozygosity

The heterozygosity observed with the microsatellite data in all the taxa included in the study ranged from 0 to 86.96% (Table 3). In general, citrons had a lower level of heterozygosity than pummelos, papedas, mandarins or kumquats. In citrons, the heterozygosity ranged from 0 to 52%. However, contrary to previous reports (Barkley et al., 2006), certain citrons showed a high level of heterozygosity. 'Citron of Commerce' had 52.17% heterozygous SSR markers; 'Italian' citron had 43.38% heterozygosity. One known citron hybrid (CRC3819) had a high percentage of heterozygosity at 82.61%. SNPs of the nuclear gene MDH were also analyzed for heterozygosity. In the 1600 bp fragment of the MDH gene, 13 citron accessions had no heterozygosity and two accessions had six heterozygous bases (Table 3).

3.5. Single nucleotide polymorphism analysis of a nuclear gene, malate dehydrogenase (MDH)

Microsatellite markers can differentiate among alleles that have altered numbers of repeats. Point mutations will not bedetected by such markers. A sequence generated from the single-copy gene MDH was used to study SNP patterns in the data matrix. The sequences used for MDH analysis were obtained from eight colonies for each clone, so we believe that the SNPs recorded are accurate. Since MDH is a nuclear gene, most accessions have two sequences corresponding to the two haplotypes (Genbank Accession numbers KT175609–KT175700).

The length of the MDH sequences in the 54 taxa studied was about 1615 bp. Eight non-citron accessions had one or both haplotypes with a deletion of 10 bp in the intron region of the MDH gene fragment. Two non-citron accessions included in the dataset, the lemon and the kumquat, did not have this deletion. The number of SNPs in the citron accessions ranged from 0 to 9 ('Ruili Wild FI' citron had the maximum number of SNPs). The MP tree shown in Fig. 5 has a citron clade with strong bootstrap support of 98. The 'India' lemon hybrid clustered with the other citrons and the hybrid accession 'Suanmaliu', clustered with the non-citrons. Three main clusters of citrons were observed (Fig. 5).

3.6. A citron specific marker

All the citrons had CF-AT13 PCR products estimated to be about 361–379 nucleotide base pairs based on mobility on LI-COR gels. All the non-citrons included in the analysis yielded PCR products estimated to be about 201–221 bp and lacked the higher MW band (~370 bp). Presumed hybrid accessions had one band around 365 bp and a second band around 211–213 bp indicating an admixture of citron and non-citron alleles (Supplementary Fig. 1). To ensure that the larger band is not a dimer on LI-COR gels, the products of representative sizes were recovered, cloned and sequenced. There was a homozygous insertion in all the citrons and a heterozygous insertion in citron hybrids (Table 3). Supplementary Fig. 1 shows an alignment of the sequences of the CF-AT13 marker region from the representative samples. Since the microsatellite marker is in a nuclear gene, we obtained two sequences for each accession indicating haplotypes 1 and 2. All the citron accessions had an insertion of about 146–175 bp that was missing from the

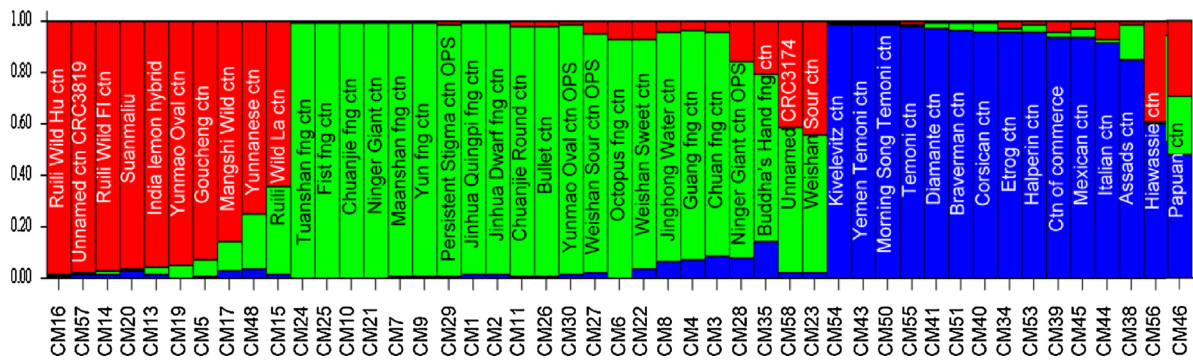


Fig. 3. Population structure analysis of citrons using data from 23 microsatellite markers. Three major populations were inferred. Y axis represents probable admixture in each accession. Cluster 1 (Red) primarily represents non-fingered wild citrons from China, including many likely hybrids. Cluster 2 (green) consists of fingered and some non-fingered citrons from China. Cluster 3 (blue) represents non-fingered citrons from many locations like Israel, Italy, Morocco, United States, etc. Overall proportion of membership in the three inferred clusters were: 0.240, 0.449, and 0.311. Citron, fingered and open pollinated seedling are abbreviated as ctn, fng and OPS.

Fig. 4. Neighbor joining tree showing relationships of citrons. SSR marker data wasconverted into a distance matrix (POSA—percentage of shared alleles) using POPULATIONS software. The 500 distance matrices generated were fed into MEGA 6.0 and 500 NJ trees were created. A consensus tree obtained from these 500 datasets is shown. Bootstrap values are indicated on the branches. Two clades had support values of 84 and 87—as indicated. Other clades had lower support values. Cluster 1 (red), 2 (green), or 3 (blue, hollow oval design) indicated in parentheses represents the clustering observed in STRUCTURE analysis (shown in Fig. 3, analyzed using the same SSR data). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3
Heterozygosity observed in the taxa. Microsatellite marker data were used to calculate percentage of heterozygosity. Number of heterozygous bases observed in the MDH amplicon is indicated. The presence or absence of a citron-specific insertion in CF-AT13 microsatellite marker region was used to designate citron/non-citron status of accessions. CM numbers indicated in parenthesis.

Name of the cultivar	Microsatellite analysis		SNP analysis (MDH)	CF-AT13 analysis
	Heterozygous loci (total)	Heterozygosity observed (%)	Number of heterozygous bases	CF-AT13 alleles
“Jinhua Qingpi” fingered citron	1 (23)	4.35	0	Citron/Citron
“Jinhua Dwarf” fingered citron	1 (23)	4.35	1	Citron/Citron
“Chuan” fingered citron	3 (22)	13.64	3	Citron/Citron
“Guang” fingered citron	4 (23)	17.39	0	Citron/Citron
“Goucheng” citron	3 (23)	13.04	5	Citron/Citron
“Octopus” fingered citron	4 (23)	17.39	4	Citron/Citron
“Maanshan” fingered citron	3 (23)	13.04	0	Citron/Citron
“Jinghong Water” citron	0 (23)	0.00	5	Citron/Citron
“Yun” fingered citron	3 (23)	13.04	6	Citron/Citron
“Chuanjie” fingered citron	2 (23)	8.70	0	Citron/Citron
“Chuanjie Round” citron	2 (23)	8.70	0	Citron/Citron
“Honghe” papeda	6 (23)	26.09	2	Non-citron/Non-citron
“India” lemon hybrid	11 (23)	47.83	3	Non-citron/Citron
“Ruili Wild Fl” citron	4 (20)	20.00	6	ND
“Ruili Wild La” citron	5 (22)	22.73	0	Citron/Citron
“Ruili Wild Hu” citron	4 (18)	22.22	1	ND
“Mangshi Wild” citron	1 (23)	4.35	3	Citron/Citron
“Ruili Sour” pummelo	11 (22)	50.00	4	Non-citron/Non-citron
“Yunmao Oval” citron	4 (22)	18.18	5	Citron/Citron
“Suanmaliu” citrus hybrid	18 (23)	78.26	20	Citron/Non-citron
“Ninger Giant” citron	2 (22)	9.09	0	Citron/Citron
“Weishan Sweet” citron	6 (22)	27.27	1	Citron/Citron
“Weisham Sour” citron	6 (21)	28.57	4	ND
“Tuanshan” fingered citron	1 (23)	4.35	0	Citron/Citron
“Fist” fingered citron	1 (23)	4.35	3	Citron/Citron
“Bullet” citron	2 (23)	8.70	0	Citron/Citron
“Weishan Sour” OPS citron	5 (22)	22.73	2	Citron/Citron
“Ninger Giant” OPS citron	7 (23)	30.43	0	Citron/Citron
“Persistent Stigma” OPS citron	3 (23)	13.04	3	Citron/Citron
“Yunmao Oval” OPS citron	1 (23)	4.35	5	Citron/Citron
“Hart’s Tardiff” Valencia orange	20 (23)	86.96	3	Non-citron/Non-citron
“Siamese Sweet” pummelo	9 (23)	39.13	3	Non-citron/Non-citron
“Ponkan” mandarin	13 (23)	56.52	5	Non-citron/Non-citron
“Etrog” citron	5 (23)	21.74	0	Citron/Citron
“Buddha’s Hand” fingered citron	5 (23)	21.74	2	Citron/Citron
“Hanayu” papeda	14 (23)	60.87	7	Non-citron/Non-citron
“Nagami” kumquat	8 (23)	34.78	0	Non-citron/Non-citron
“Assads” citron	0 (23)	0	3	Citron/Citron
“Citron of Commerce”	12 (23)	52.17	0	Citron/Citron
“Corsican” citron	1 (22)	4.55	3	Citron/Citron
“Diamante” citron	6 (23)	26.09	2	Citron/Citron
“Yemen Temoni” citron	2 (23)	8.7	1	Citron/Citron
“Italian” citron	10 (23)	43.48	3	Citron/Citron
“Mexican” citron	0 (23)	0	3	Citron/Citron
“Papuan” citron	0 (23)	0	2	Citron/Citron
“Yunnanese” citron	1 (23)	4.35	4	Citron/Citron
“Morning Song Temoni” citron	2 (23)	8.7	2	Citron/Citron
“Braverman” citron	1 (23)	4.35	1	Citron/Citron
“Halperin” citron	0 (22)	0	0	Citron/Citron
“Kivelevitz” citron	6 (23)	26.09	2	Citron/Citron
“Temoni” citron	1 (23)	4.35	5	Citron/Citron
“Hiawassie” citron	0 (23)	0	5	Citron/Citron
“Unnamed” citron CRC3819	19 (23)	82.61	3	Citron/Non-citron
“Unnamed” citron CRC3174	0 (23)	0	5	Citron/Citron

pummelo, papeda, mandarin and kumquat accessions. Supplementary Fig. 1 shows that there are two regions with AT repeats in the CF-AT13 microsatellite region. The CF-AT13 microsatellite marker sequences generated in this study were deposited in Genbank (Accession numbers KT149800–KT149815).

Sequence comparison of chloroplast gene, rps16

We have sequenced a 901 bp fragment of the chloroplast gene, ribosomal protein S16 (*rps16*), and analyzed sequence variation. There were 43 loci with polymorphic bases of which 20 were present only in the non-citron accessions. We recorded 23 variant sites specific to the citrons in the dataset. We found one parsimony informative SNP present in 26 of the 56 accessions tested. Three other SNPs were present in two or three citron accessions. The *rps16* sequences generated in this study were deposited in Genbank (Accession numbers KJ364661–KJ364716).

Analysis of the cp gene *rps16* identified 43 nucleotide positions with variable bases. There were 83 alleles in citrons; 42 in Chinese citrons (15 in pure citrons and 27 in citron hybrids) and 41 in citrons of Mediterranean origin. Seventeen SNPs were unique to specific citrons. To construct an MP tree with citron and a large number of non-citron accessions, we used 68 *rps16* sequences generated from all the accessions listed in Table 1 and included 14 additional sequences retrieved from Genbank (Supplementary Fig. 2). ‘Assads’ citron (CM38) had seven unique polymorphisms in the

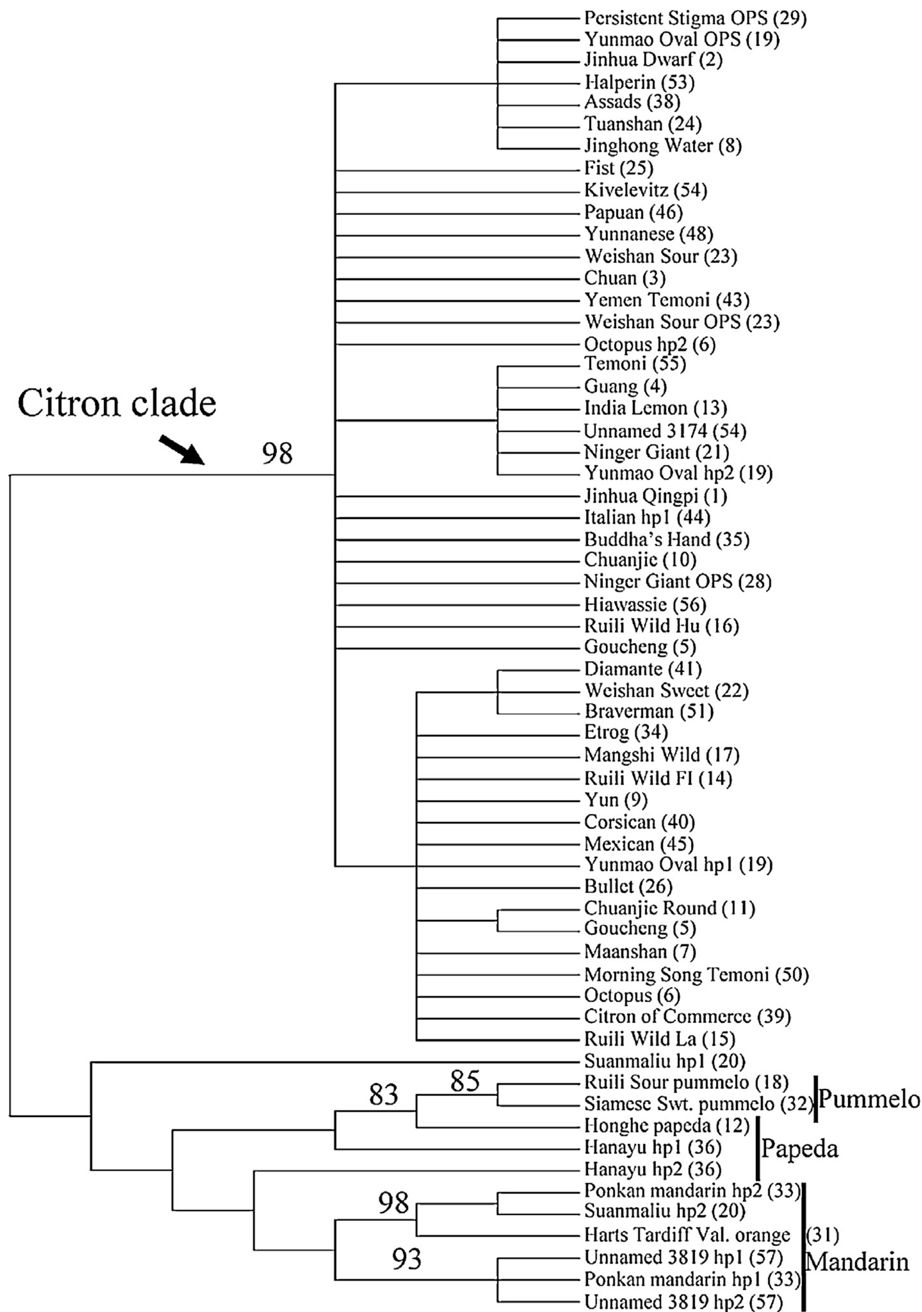


Fig. 5. Maximum parsimony tree constructed from 60 MDH sequences. Two haplotype sequences of 1615 bp length were obtained for most accessions. Only haplotypes of an accession that cluster in separate clades are shown. For other taxa, only one haplotype is represented. Tree length = 124. Bootstrap values obtained from 1550 replicates shown on the branches. Major clades (citron, pummelo, papedas and mandarin) are indicated. The aligned matrix had a total of 487 alleles in 171 nucleotide positions; there were 120 unique alleles, of which 104 were from citrons (67 alleles in citrons of direct Chinese origin and 37 alleles in citrons from other parts of the world; out of 120 unique alleles, 109 were from pure citrons and 11 alleles were from putative citron hybrids). 34 SNPs were parsimony informative.

rps16 sequence that were not shared with other citrons or any of the other citrus groups. Mean diversity for the whole population (consisting of 68 sequences) based on *rps16* sequence was 0.008. When only citrons were considered (44 accessions), the diversity was 0.002. The citrons in the dataset formed a well-supported clade (bootstrap of 99, indicated by an arrow in Supplementary Fig. 2). Pairwise distances between *rps16* sequences of 68 accessions were calculated using MEGA 6.0 software. Variance analysis was conducted using the maximum composite likelihood model (Tamura et al., 2004) using 500 bootstrap replications. Maximum divergence was between 'Assads' and *Poncirus* (2.8%). Divergence between non-citron accessions and the citron accessions in the dataset (excluding 'Assads') was between 0.006 to 2.1%. Divergence between citron accessions and 'Assads' was about 1.2%. Divergence between various citron accessions (excluding 'Assads') ranged from 0 to 0.6%. Most citrons had zero divergence from each other in the *rps16* sequences.

4. Discussion

Understanding the genetic diversity of *C. medica*, one of the three basic species in the genus *Citrus*, is essential for germplasm preservation, for protection of rare cultivars and for enhancing future breeding efforts. Citron is believed to have contributed as a male parent to the development of many cultivars (Barrett and Rhodes, 1976; Nicolosi et al., 2000). Most molecular studies of the genus *Cit-*

rus include one to three representatives of *C. medica*, usually 'Etrog', 'Buddha's Hand' and 'Diamante' (Abkenar et al., 2008; Bayer et al., 2009; Cheng et al., 2005; Deng et al., 2007; Federici et al., 1998; Herrero et al., 1996b; Jannati et al., 2009). Insufficient sampling of *C. medica* results in a lack of understanding of the population variation and the genetic diversity of the group, leading to an erroneous conclusion that *C. medica* is a very homogeneous group (Herrero et al., 1996a,b). We found only two previous studies that included an adequate number of citrons in their analysis (Barkley et al., 2006; Luro et al., 2012)

We used three molecular marker techniques to study different aspects of citron diversity. The SSRs provide genome-wide molecular data from mostly unlinked markers. SNP study of the nuclear gene MDH can be highly informative because of the large number of data points. Information is generated for each nucleotide position for both haplotypes, so the implications concerning both maternal and paternal ancestry are embedded in the data file. In comparison, SNP analysis of the chloroplast gene, *rps16* is of limited value since only maternal ancestry is evident from the sequence information. A combination of all three methods is useful to understand the ancestry of, and relationships between, different accessions in the study.

The taxonomy, phylogeny and diversity of the genus *Citrus* have been studied previously by several methods. Analysis of essential oils, phytochemistry profiles and isozyme studies have indicated that citrons have very little heterozygosity (Esen and Scora, 1977; Ogawa et al., 2000; Scora and Malik, 1970; Torres et al., 1978, 1982). Citrus phylogeny has been studied extensively using cp sequences to understand kinship and infer parentage of hybrids (Abkenar et al., 2004; Araújo et al., 2003; Bayer et al., 2009; Chase et al., 1999; Jena et al., 2009; Morton et al., 2003; Penjor et al., 2010, 2013; Samuel et al., 2001). Chloroplast DNA evolves slowly and is used widely to study taxonomic and phylogenetic relationships between plants belonging to various groups (Gielly and Taberlet, 1994; Olmstead and Palmer, 1994). The region of chloroplast DNA selected for the present study is an *rps16* group II intron sequence, a region used for comparative study of diverse groups of plants at higher taxonomic levels (Golenberg et al., 1993; Oxelman et al., 1997). In a comprehensive study including 59 genera and 65 species of Rutaceae (about one third of all Rutaceae), significant discrimination was observed among different groups of Rutaceae using chloroplast sequences (Groppo et al., 2008). For our study aimed at resolving differences among members of basically one species (*C. medica*), the cp sequences provided limited resolution. However, we reported 17 unique SNPs found in individual citron accessions (hence not considered informative) and these SNPs may add greater diversity to the citron *rps16* sequence database. The phylogenetic tree constructed from *rps16* sequences had a distinct citron clade with a bootstrap of 99, useful in distinguishing citrons from non-citrons.

Because of the predominant maternal inheritance of cp DNA and the prevalence of hybridization in citrus, phylogenetic relationships may be better understood using nuclear molecular markers such as internal transcribed spacer sequences (ITS), SNP and SSR. SNPs are more prevalent than microsatellites and have been used for studying relationships of several citrus groups (Garcia-Lor et al., 2012; Novelli et al., 2004; Ramadugu et al., 2013). Single nucleotide polymorphisms generally have very low mutation rates, making them very useful to study very deep genealogies (thousands of generations) (Walsh, 2001). In a previous study involving 32 taxa of the Aurantioideae and analyses of 6 nuclear genes, we determined that the single copy nuclear gene MDH is a good candidate for the phylogenetic study of *Citrus* and its close relatives (Ramadugu et al., 2013). The MDH sequences of various groups of citrus show distinct differences. In the current MDH sequence analysis aimed at differentiating between members of *C. medica*, about 104 SNPs were unique to certain accessions, perhaps indicating wild-type germplasm in the taxa utilized for the study. Although we observed a considerable amount of variability, our phylogenetic analysis does not reflect this variation, as the SNPs were not parsimony-informative. If more accessions were included in the analysis, many of these singletons might be considered as parsimony-informative, and the phylogenetic tree would better represent relationships among the accessions. For the MDH region, we have 0–9 SNPs in the citron accessions and up to 32 SNPs in the non-citron accessions. The citrons from Yunnan had 67 unique alleles and the citrons from the Mediterranean region had 37 accession-specific alleles indicating a significant genetic variability in the accessions studied.

The clustering pattern of citrons in the neighbor joining tree based on microsatellite marker data and the MDH maximum parsimony (MP) tree based on nuclear gene SNPs were different. This is an expected result since the neighbor joining tree is based on 23 genome-wide markers, while the MDH SNP data is based on 34 parsimony-informative linked data points from a single gene fragment. Our phylogenetic analysis using both chloroplast *rps16* sequences and nuclear MDH sequences placed citrons in a separate group from other citrus, including kumquat (Fig. 4, Supplementary Fig. 2). Since we included wild citron-like types from a remote location, these analyses added evidence to support the assumption that the accessions studied are citrons and not some other wild forms of citrus. The *C. medica* group appears to be monophyletic in this study.

Microsatellites have a high mutation rate that results in altered length of the repeats; this variability makes SSRs ideal candidates for certain genealogy studies involving multiple generations (Walsh, 2001). To analyze the data from 23 microsatellite loci we used the program STRUCTURE, which utilizes a model-based clustering, Bayesian approach to infer population structure (Evanno et al., 2005; Falush et al., 2003; Pritchard et al., 2000). In the present study we have identified three distinct genetic populations and certain admixed individuals (Fig. 3). Based on the genomic information available (Table 2), the SSR loci are distributed over 8 of the 9 linkage groups and only a few appear to be closely linked (Ollitrault et al., 2012; http://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Cclementina genomic coordinates indicated in Table 2). Comparison with Clementine mandarin genome indicates that the markers used are present as single copy markers in the genome. However, the locations of SSR markers on citron genomes may not be identical to their locations on mandarins. In addition, the physical distances may not represent the genetic distances between markers. Based on available information, we have made the assumption that linkage disequilibrium is not a significant concern for this dataset.

The microsatellite markers that we used for the study had PIC values of 0.23–0.75. The PIC value depicts the discriminatory power of a microsatellite marker and considers both the number of alleles present and the relative frequencies in a particular dataset. Thirteen of the markers had a PIC value higher than 0.5, indicating good discriminating ability (DeWoody et al., 1995). Since we are primarily analyzing accessions belonging to one species (*C. medica*), the SSR data generated is considered informative and adequate to discriminate among most accessions included in the study. Most of the fingered citrons are seedless and typically propagated by vegetative means. It is likely that genetically some of them are very similar to each other. In our analysis with 23 pairs of microsatellites, we could not distinguish between CM 1 and 2, between CM 7 and 9, or between CM 24 and 25. About 4% of the microsatellite data constituted unique alleles for specific accessions and hence were considered uninformative in the phylogenetic analysis. Five of the six unique SSR alleles were observed in the Chinese citrons and may represent putative wild-type germplasm (Barkley et al., 2006).

The 11 fingered citrons in this study clustered together in population structure analysis. Certain non-fingered citrons like 'Ninger Giant', 'Ninger Giant' OPS, 'Persistent Stigma' OPS, 'Chuanjie Round', 'Bullet', 'Yunmao Oval' OPS, 'Weishan Sour' OPS, 'Weishan Sweet' and 'Jinghong Water' also clustered with the fingered cultivars. 'Weishan Sweet' was an exceptional accession in this group since it has normal locules, juice vesicles and seeds; many other non-fingered accessions in this group have rudimentary locules and juice vesicles. All non-fingered varieties in this study have seeds. According to STRUCTURE analysis, the average distance between individuals in this cluster was 0.219 and the mean Fst was about 0.5, indicating variability within the group. Considerable genomic diversity was observed in fingered citrons that do not produce any viable seed and are asexually propagated. It is probable that many of the fingered varieties either evolved independently of each other, were selected for this unusual fruit character, or, resulted from a cross involving a fingered citron as a pollen parent.

Many non-fingered citrons that have normal locules, juice vesicles and seeds grouped in cluster 1 (Fig. 3). The Fst in this cluster was 0.0012, indicating a high level of similarity between the genotypes in this cluster. The Mediterranean citrons formed a cluster with a mean Fst of 0.486.

Two hybrids found wild in Yunnan, 'India' lemon and 'Suanmaliu', probably have citron/non-citron parentage and had unique alleles when amplified with the JI-AC03 primer set (Supplementary Table 1). Since these alleles were not observed in any of the citrons analyzed, it is probable that the alleles originated from the non-citron parent. In the *rps16* phylogenetic tree, these accessions clustered with the non-citrons, indicating a non-citron as the maternal parent. Three microsatellite marker alleles that were unique and specific to citrons were observed in 'Octopus' fingered citron and 'Ruili Wild FI' citron with the CF-AT13 primer, and 'Yunmao Oval' citron accession with the CF-TA03 primer. SSR loci are known to have a high rate of mutation per locus per generation, ranging from 2.5×10^{-5} to 1×10^{-2}, and it is possible that these unique alleles are derived from recent mutation events (Weber and Wong, 1993). The presence of unique microsatellite alleles and SNPs may indicate novel genotypes.

Microsatellite marker CF-AT13 will be useful for quick identification of citron genotypes and citron/non-citron hybrids by a simple PCR followed by agarose gel electrophoresis to estimate amplicon sizes. This marker amplified a much larger fragment from all citrons in the dataset compared to non-citrons. A single copy of the CF-AT13 locus was detected on scaffold 7 of the genome sequence of the haploid Clementine mandarin genome (Table 2) (http://Phytozome.jgi.doe.gov). The fragment amplified by the CF-AT13 primer set codes for the 3′ terminal region of the gene encoding anthocyanidin 3-*O*-glucosyl transferase, a region present in both citron and non-citron accessions. The intergenic sequence downstream of the stop codon has an insertion of 146-175 nucleotides observed only in the citron group. Indel polymorphisms are considered to be more informative about citrus phylogeny than SNPs (Garcia-Lor et al., 2012).

Citrons are generally assumed to be pure since they are self-compatible (Hodgson, 1967; Scora, 1975). In many studies involving isozymes (Herrero et al., 1996a), RFLP analysis (Federici et al., 1998) and microsatellite markers (Barkley et al., 2006), citrons exhibit lower levels of heterozygosity than other citrus groups. We made a similar observation in the current microsatellite dataset for most accessions. However, two citron cultivars, 'Citron of Commerce' and 'Italian', had 52% and 43% heterozygosity, respectively (Table 3). The higher heterozygosity of these citron accessions does not appear to originate from interspecific hybridization, as their level of admixture is low (Fig. 3). The heterozygosity recorded for these accessions is comparable to the level of heterozygosity observed in 'Ponkan' mandarin (56%), 'Ruili Sour' pummelo (50%), 'India lemon' hybrid (48%) and 'Siamese Sweet' pummelo (39%). About 12 citrons had a heterozygosity level of 20–30%. These included wild citrons found in China and one fingered citron ('Buddha's Hand'). Eight citrons (mostly fingered types from China) had heterozygosity levels from 10 to 20%. Six citrons had heterozygosity levels of 8–9%. Ten citrons, including some Mediterranean types such as 'Corsican', 'Temoni' and 'Braverman', had 4% heterozygosity. Since the focus of this work is on *C. medica*, we only used primers that discriminated among citrons. Hence our data is skewed to reflect citron diversity. Overall, the current study, along with Barkley et al. (2006); indicates that some citrons have a higher level of heterozygosity than previously estimated.

In a study of 24 citrons, Luro et al. (2012) identified 13 true citrons including 'Etrog', 'Diamante', 'Buddha's Hand', 'Corsican' and certain other accessions not included in the present study. Data from mitochondrial and chloroplast microsatellite markers in addition to chemical composition of the leaf essential oils was used for analysis. In general, the true citrons had a low level of heterozygosity. A heterozygosity frequency exceeding 0.35 exhibited by 'Etrog' and 'Diamante' were assumed to be because of multiple introductions of these cultivars and cross hybridization between the different varieties of these citron cultivars (Luro et al., 2012). We observed 20–52% heterozygosity in our study in 11 citrons that were not hybrids of citron (Table 3).

Heterozygosity values for 'Ninger Giant' citron (9%) and an open pollinated seedling of 'Ninger Giant' citron (30%) were different in microsatellite analysis. A comparison of the two 'Ninger Giant' samples revealed presence of five variable bases in the MDH sequence. The five homozygous SNPs indicate that it is unlikely that 'Ninger Giant' OPS is directly derived from the 'Ninger Giant' genotype that constituted the CM21 sample. Similarly, 'Yunmao Oval' and 'Yunmao Oval' OPS had heterozygosity values of 18% and 4% in microsatellite analysis. Heterozygous SNPs were recorded in the two accessions in ten nucleotide positions. 'Weishan' Sour and 'Weishan Sour' OPS citron had heterozygosity values of 28% and 22% respectively in microsatellite analysis. A comparison of these two accessions showed presence of seven heterozygous SNPs. We conclude that these OPS samples may have originated from pollination with another type of citrus.

Citron was the first type of citrus to be cultivated and to be brought to the Mediterranean region, very likely in the first millennium BCE, and was later dispersed to other parts of Europe and then to the Americas (Hodgson, 1967). Historical trade routes suggest that India may have been a source of citrons grown in the Mediterranean region. According to Andrews, based on archaeological evidence and the testimony of agricultural writers, the Mediterranean region did obtain citrons from India (Andrews, 1961). The rest of Europe and the Americas obtained their citrons from the Mediterranean regions. Although certain accessions currently maintained in the CVC were obtained from various sources in the United States, the assumption is that all of these were originally procured from various Mediterranean countries. Molecular marker analysis of citrons in this study points out that the citrons from the Mediterranean region cluster separately from the Chinese citrons. If these Mediterranean citrons originated in India, they are likely to be distinct from the Chinese citrons and there is substantial scope for improving the diversity of citrus germplasm by using the under-utilized, diverse Chinese citron germplasm resources. Currently cultivated citrus varieties have been subjected to selection over a long period of time and most are clonally propagated and hence have a narrow genetic diversity, making them vulnerable to diseases like citrus Huanglongbing (Bové, 2006). Sustainable crop production will require exploiting the genetic resources available for creating new cultivars capable of disease resistance, growth under suboptimal environmental conditions, and horticultural acceptability (Talon and Gmitter, 2008). The mag-

nitude of genetic heterogeneity in wild citrus genotypes growing in China and India has not yet been catalogued. The study of the genetic diversity of citron, an important parent species of cultivated limes and lemons, will be useful to exploit the genetic potential of novel germplasm, which can be utilized to develop new cultivars and to implement appropriate conservation strategies.

5. Conclusions

We have conducted genetic characterization of 47 citrons from China and the Mediterranean region. Most of the Chinese citrons included in the study have not been analyzed before. We report significant genetic diversity between the three major groups of citrons in this study. Citrons of Mediterranean ancestry are genetically distinct from citrons of Chinese origin. We provide evidence that there exist multiple cultivars of fingered citrons that are morphologically, and genetically, distinct; that other fingered citron accessions, which appear to be different morphologically, are the same or very closely related. Understanding the diversity of this basic citrus species is essential for proper conservation efforts.

Acknowledgements

Financial support for the study was provided through a specific cooperative agreement between University of California Riverside and United States Department of Agriculture. The sponsors had no involvement in the study design, analysis or interpretation of data. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. The authors declare that they do not have any conflict of interest.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.scienta.2015.09.004.

References

Abkenar, A.A., Isshiki, S., Matsumoto, R., Tashiro, Y., 2008. Comparative analysis of organelle DNAs in acid citrus grown in Japan using PCR-RFLP method. Genet. Resour. Crop Evol. 55, 487–492.

Abkenar, A.A., Isshiki, S., Tashiro, Y., 2004. Phylogenetic relationships in the true citrus fruit trees revealed by PCR-RFLP analysis of cpDNA. Sci. Hortic. 102, 233–242.

Ahmad, R., Struss, D., Southwick, S.M., 2003. Development and characterization of microsatellite markers in Citrus. J. Am. Soc. Hortic. Sci. 128, 584–590.

Andrews, A.C., 1961. Acclimatization of citrus fruits in the Mediterranean region. Agri. Hist. 35, 35–46.

Araújo, E.F., de Queiroz, L.P., Machado, M.A., 2003. What is Citrus?: Taxonomic implications from a study of cp-DNA evolution in the tribe Citreae (Rutaceae subfamily Aurantioideae). Org. Divers. Evol. 3, 55–62.

Barkley, N.A., Roose, M.L., Krueger, R.R., Federici, C.T., 2006. Assessing genetic diversity and population structure in a citrus germplasm collection utilizing simple sequence repeat markers (SSRs). Theor. Appl. Genet. 112, 1519–1531.

Barrett, H.C., Rhodes, A.M., 1976. A numerical taxonomic study of affinity relationships in cultivated Citrus and its close relatives. Syst. Bot. 1, 105–136.

Bayer, R.J., Mabberley, D.J., Morton, C., Miller, C.H., Sharma, I.K., Pfeil, B.E., Rich, S., Hitchcock, R., Sykes, S., 2009. A molecular phylogeny of the orange subfamily (Rutaceae: Aurantioideae) using nine cpDNA sequences. Am. J. Bot. 96, 668–685.

Bové, J., 2006. Huanglongbing: a destructive, newly-emerging, century-old disease of citrus. J. Plant. Pathol. 88, 7–37.

Chase, M.W., Morton, C.M., Kallunki, J.A., 1999. Phylogenetic relationships of Rutaceae: a cladistic analysis of the subfamilies using evidence from rbcL and atpB sequence variation. Am. J. Bot. 86, 1191–1199.

Chen, C., Zhou, P., Choi, Y.A., Huang, S., Gmitter, F.G., 2006. Mining and characterizing microsatellites from citrus ESTs. Theor. Appl. Genet. 112, 1248–1257.

Cheng, Y., De Vicente, M.C., Meng, H., Guo, W., Tao, N., Deng, X., 2005. A set of primers for analyzing chloroplast DNA diversity in Citrus and related genera. Tree Physiol. 25, 661–672.

Deng, Z., Malfa, S.L., Xie, Y., Xiong, X., Gentile, A., 2007. Identification and evaluation of chloroplast uni- and trinucleotide sequence repeats in citrus. Sci. Hortic. 111, 186–192.

DeWoody, J.A., Honeycutt, R.L., Skow, L.C., 1995. Microsatellite markers in white-tailed deer. J. Hered. 86, 317–319.

Doyle, J.J., Doyle, J.L., 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19, 11–15.

Esen, A., Scora, R.W., 1977. Amylase polymorphism in Citrus and some related genera. Am. J. Bot. 64, 305–309.

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.

Falush, D., Stephens, M., Pritchard, J.K., 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164, 1567–1587.

Federici, C.T., Fang, D.Q., Scora, R.W., Roose, M.L., 1998. Phylogenetic relationships within the genus *Citrus* (Rutaceae) and related genera as revealed by RFLP and RAPD analysis. Theor. Appl. Genet 96, 812–822.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Garcia-Lor, A., Curk, F., Snoussi-Trifa, H., Morillon, R., Ancillo, G., Luro, F., Navarro, L., Ollitrault, P., 2012. A nuclear phylogenetic analysis: SNPs, indels and SSRs deliver new insights into the relationships in the 'true citrus fruit trees' group (Citrinae, Rutaceae) and the origin of cultivated species. Ann. Bot. 111 (1), 1–19.

Gielly, L., Taberlet, P., 1994. The use of chloroplast DNA to resolve plant phylogenies: noncoding versus *rbcL* sequences. Mol. Biol. Evol. 11, 769–777.

Gmitter, F.G., Hu, X., 1990. The possible role of Yunnan, China, in the origin of contemporary *Citrus* species (Rutaceae). Econ. Bot. 44, 267–277.

Golenberg, E.M., Clegg, M.T., Durbin, M.L., Doebley, J., Ma, D.P., 1993. Evolution of a noncoding region of the chloroplast genome. Mol. Phylogenet. Evol. 2, 52–64.

Groppo, M., Pirani, J.R., Salatino, M.L.F., Blanco, S.R., Kallunki, J.A., 2008. Phylogeny of Rutaceae based on two noncoding regions from cpDNA. Am. J. Bot. 95, 985–1005.

Hazarika, T.K., 2012. Citrus genetic diversity of north-east India, their distribution, ecogeography and ecobiology. Genet. Resour. Crop Evol. 59, 1267–1280.

Herrero, R., Asins, M.J., Carbonell, E.A., Navarro, L., 1996a. Genetic diversity in the orange subfamily Aurantioideae. I. Intraspecies and intragenus genetic variability. Theor. Appl. Genet. 92, 599–609.

Herrero, R., Asins, M.J., Pina, J.A., Carbonell, E.A., Navarro, L., 1996b. Genetic diversity in the orange subfamily Aurantioideae. II. Genetic relationships among genera and species. Theor. Appl. Genet. 93, 1327–1334.

Hodgson, R.W., 1967. Horticultural varieties of citrus. In: Reuther, W., Webber, H.J., Batchelor, L.D. (Eds.), The Citrus Industry, vol. I. University of California Press, Berkeley, pp. 431–591, http://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Cclementina http://www.ars-grin.gov http://www.citrusvariety.ucr.edu/citrus.

Jannati, M., Fotouhi, R., Abad, A.P., Salehi, Z., 2009. Genetic diversity analysis of Iranian citrus varieties using microsatellite (SSR) based markers. J. Hortic. For. 1, 120–125.

Jena, S.N., Kumar, S., Nair, N.K., 2009. Molecular phylogeny in Indian *Citrus* L. (Rutaceae) inferred through PCR-RFLP and trnL-trnF sequence data of chloroplast DNA. Sci. Hortic. 119, 403–416.

Kijas, J.M.H., Thomas, M.R., Fowler, J.C.S., Roose, M.L., 1997. Integration of trinucleotide microsatellites into a linkage map of *Citrus*. Theor. Appl. Genet. 94, 701–706.

Kumar, S., Jena, S.N., Nair, N.K., 2010. ISSR polymorphism in Indian wild orange (*Citrus indica* Tanaka Rutaceae) and related wild species in North-east India. Sci. Hortic. 123, 350–359.

Luro, F., Venturini, N., Costantino, G., Paolini, J., Ollitrault, P., Costa, J., 2012. Genetic and chemical diversity of citron (*Citrus medica* L.) based on nuclear and cytoplasmic markers and leaf essential oil composition. Phytochemistry 77, 186–196.

Mabberley, D.J., 1997. A classification for edible *Citrus* (Rutaceae). Telopea 7, 167–172.

Moore, G.A., 2001. Oranges and lemons: Clues to the taxonomy of *Citrus* from molecular markers. Trends Genet. 17, 536–540.

Morton, C., Grant, M., Blackmore, S., 2003. Phylogenetic relationships of the Aurantioideae inferred from chloroplast DNA sequence data. Am. J. Bot. 90, 1463–1469.

Nagy, S., Poczai, P., Cernák, I., Gorji, A.M., Hegedűs, G., Taller, J., 2012. PICcalc: an online program to calculate polymorphic information content for molecular genetic studies. Biochem. Genet. 50, 670–672.

Nair, K.N., Nayar, M.P., 1997. Rutaceae. In: Hajra, P.K., Nair, V.J., Daniel, P. (Eds.), Flora of India, vol. IV. Botanical Survey of India, Calcutta, India, pp. 229–407.

Nei, M., Kumar, S., 2000. Molecular Evolution and Phylogenetics. Oxford University Press, New York, NY, pp. 1–333.

Nicholas, K.B., Nicholas Jr., H.B., Deerfield II, D.W., 1997. GeneDoc: analysis and visualization of genetic variation. EMBNEW NEWS 4, 14.

Nicolosi, E., Deng, Z.N., Gentile, A., La Malfa, S., Continella, G., Tribulato, E., 2000. Citrus phylogeny and genetic origin of important species as investigated by molecular markers. Theor. Appl. Genet. 100, 1155–1166.

Novelli, V.M., Takita, M.A., Machado, M.A., 2004. Identification and analysis of single nucleotide polymorphisms (SNPs) in citrus. Euphytica 138, 227–237.

Ogawa, K., Kawasaki, A., Omura, M., Yoshida, T., Ikoma, Y., Yano, M., 2000. 3',5'-Di-C-β-glucopyranosylphloretin, a flavonoid characteristic of the genus *Fortunella*. Phytochemistry 57, 737–742.
Ollitrault, P., Terol, J., Chen, C., Federici, C.T., Lotfy, S., et al., 2012. A reference genetic map of *C. clementina* hort. Ex Tan.; citrus evolution interferences from comparative mapping. BMC Genomics 13, 593.
Olmstead, R.G., Palmer, J.D., 1994. Chloroplast DNA systematics: a review of methods and data analysis. Am. J. Bot. 81, 1205–1224.
Oxelman, B., Lidén, M., Berglund, D., 1997. Chloroplast rps16 intron phylogeny of the tribe *Sileneae* (*Caryophyllaceae*). Plant Syst. Evol. 206, 393–410.
Pang, X.-M., Hu, C.-G., Deng, X.-X., 2007. Phylogenetic relationships within *Citrus* and its related genera as inferred from AFLP markers. Genet. Resour. Crop Evol. 54, 429–436.
Penjor, T., Anai, T., Nagano, Y., Matsumoto, R., Yamamoto, M., 2010. Phylogenetic relationships of *Citrus* and its relatives based on rbcL gene sequences. Tree Genet. Genomes 6, 931–939.
Penjor, T., Yamamoto, M., Uehara, M., Ide, M., Matsumoto, N., Matsumoto, R., Nagano, Y., 2013. Phylogenetic relationships of *Citrus* and its relatives based on *matK* gene sequences. PLoS One 8, e62574.
Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.
Ram, S.G., Thiruvengadam, V., Vinod, K.K., 2007. Genetic diversity among cultivars, landraces and wild relatives of rice as revealed by microsatellite markers. J. Appl. Genet. 48, 337–345.
Ramadugu, C., Pfeil, B.E., Keremane, M.L., Lee, R.F., Maureira-Butler, I.J., Roose, M.L., 2013. A six nuclear gene phylogeny of *Citrus* (Rutaceae) taking into account hybridization and lineage sorting. PLoS One 8, e68410.
Samuel, R., Ehrendorfer, F., Chase, M.W., Greger, H., 2001. Phylogenetic analyses of Aurantioideae (Rutaceae) based on non-coding plastid DNA sequences and phytochemical features. Plant Biol. 3, 77–87.
Scora, R.W., 1975. On the history and origin of *Citrus*. Bull. Torrey Bot. Club 102, 369–375.
Scora, R.W., 1988. Biochemistry, taxonomy and evolution of modern cultivated *Citrus*. Margraf, Weikershein, Germany In: Goren, R., Mendel, K. (Eds.), Proc. 6th Intl. Citrus Congr, vol 1, pp. 277–289.
Scora, R.W., Malik, M.N., 1970. Chemical characterization of *Citrus* as a tool in phylogeny. Taxon 19, 215–228.
Swingle, W.T., Reece, P.C., 1967. The botany of Citrus and its wild relatives of the orange subfamily. In: Reuther, W., Webber, H.J., Batchelor, L.D. (Eds.), The Citrus Industry, vol I. University of California Press, Berkeley, pp. 190–423.
Talon, M., Gmitter Jr., F.G., 2008. Citrus genomics. Int. J. Plant Genomics., Article ID 528,361, 17 pages.
Tamura, K., Nei, M., Kumar, S., 2004. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc. Natl. Acad. Sci. U. S. A 101, 11030–11035.
Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.
Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.
Torres, A.M., Soost, R.K., Diedenhofen, U., 1978. Leaf isozymes as genetic markers in *Citrus*. Am. J. Bot. 65, 869–881.
Torres, A.M., Soost, R.K., Mau-Lastovicka, T., 1982. Citrus isozymes genetics and distinguishing nucellar from zygotic seedlings. J. Hered. 73, 335–339.
Walsh, B., 2001. Estimating the time to the most recent common ancestor for the Y chromosome or mitochondrial DNA for a pair of individuals. Genetics 158, 897–912.
Weber, J.L., Wong, C., 1993. Mutation of human short tandem repeats. Hum. Mol. Genet. 2, 1123–1128.