Diversity and relationships of *Crocus sativus* and its relatives analysed by inter-retroelement amplified polymorphism (IRAP)

Nouf Fakieh Alsayied¹, José Antonio Fernández², Trude Schwarzacher¹ and J. S. Heslop-Harrison^{1,*}

¹Department of Biology, University of Leicester, Leicester LE1 7RH, UK and ²Laboratorio de Biotecnología, Instituto de Desarrollo Regional, Universidad de Castilla-La Mancha, Campus Universitario s/n, 02071 Albacete, Spain * For correspondence. E-mail: phh4@le.ac.uk

Received: 7 February 2015 Returned for revision: 8 April 2015 Accepted: 27 May 2015 Published electronically: 1 July 2015

• Background and Aims: Saffron (*Crocus sativus*) is a sterile triploid (2n = 3x = 24) cultivated species, of unknown origin from other diploid and polyploid species in the genus *Crocus* (Iridaceae). Species in the genus have high morphological diversity, with no clear phylogenetic patterns below the level of section *Crocus* series *Crocus*. Using DNA markers, this study aimed to examine the diversity and relationships within and between species of *Crocus* series *Crocus*.

• **Methods:** Eleven inter-retroelement amplified polymorphism (IRAP) primers were used in 63 different combinations with 35 single-plant accessions of *C. sativus* and related *Crocus* species in order to determine genetic variability and to conduct phylogenetic analysis.

• Key Results: A total of 4521 distinct polymorphic bands from 100 bp to approx. 4 kb were amplified; no fragment specific to all accessions of a single species was amplified. The polymorphic information content (PIC) values varied from approx. 0.37 to approx. 0.05 (mean 0.17 ± 0.1) and the major allele frequency had a mean of 0.87. High levels of polymorphism were identified between accessions of the six species of *Crocus* series *Crocus* related to *C. sativus*, with further variation between the species. In contrast, no polymorphisms were seen among 17 *C. sativus* accessions obtained in the region from Kashmir through Iran to Spain.

• Conclusions In contrast to the intraspecific variability seen in other *Crocus* species, *C. sativus* has minimal genetic variation, and it is concluded that the triploid hybrid species has most probably arisen only once. The data show that saffron is an allotriploid species, with the IRAP analysis indicating that the most likely ancestors are *C. cartwrightianus* and *C. pallasii* subsp. *pallasii* (or close relatives). The results may facilitate resynthesizing saffron with improved characteristics, and show the need for conservation and collection of wild *Crocus*.

Key words: *Crocus sativus*, saffron, inter-retroelement amplified polymorphism, IRAP, retrotransposons, markers, crops, polyploidy.

INTRODUCTION

Crocus is a genus in which 88-160 small corm-bearing perennial species are recognized; the genus is divided taxonomically into two subgenera, two sections and 15 series (Mathew, 1982; Petersen et al., 2008; Harpke et al., 2013, 2015). Species occur in the wild in Europe, the Middle East and North Africa, and are grown as ornamentals throughout the world. Saffron (*Crocus sativus*, 2n = 3x = 24) is cultivated as a spice and colorant, and the common name is applied both to the plant and to the spice. The spice, obtained from its dried stigmas, is the most expensive farmed agricultural product per gram. Saffron is grown in Kashmir, Iran, North Africa and Europe in environments characterized by cool winters and warm dry summers. Archaeological records indicate that saffron was cultivated and used as a spice and/or medicinal plant in the Mediterranean basin as early as the late Bronze Age. However, there is no consensus on where the first saffron plants were domesticated and grown (see Grilli Caiola and Canini, 2010; Molina et al., 2015).

Genetic diversity is crucial in all breeding programmes: crop improvement relies on new genes, new regulation of genes and new gene combinations. Desirable genes, which have been selected by either man or nature itself, are found within both domesticated and wild plant populations. Ancestral species are a major source of genetic diversity, and traits of interest may be introduced as chromosomal segments through direct crossing or through genetic manipulation techniques in crop improvement programmes (Vaughan *et al.*, 2007; Heslop-Harrison and Schwarzacher, 2012). With a basic chromosome number of x = 8, the sterile saffron is propagated exclusively by vegetative means (see Petersen *et al.*, 2008; Agayev *et al.*, 2009), although there are scattered reports of hybrids at least back to the work of Chappellier (1900).

The regions producing saffron each consider that they have a product with unique attributes. Variation in saffron product characters can be due to environmental effects, post-harvest processing and any genetic variation (Agayev *et al.*, 2006; Nehvi *et al.*, 2007; Ghaffari and Bagheri, 2009; Fluch *et al.*, 2010; Siracusa *et al.*, 2013; Babaei *et al.*, 2014). Saffron lines have been selected for better quality or higher yield from outperforming corms (Agayev *et al.*, 2009). As long ago as 1900, Chappellier reported 'for the saffron, there is only known a single and unique species; for ages it has not produced a single

© The Author 2015. Published by Oxford University Press on behalf of the Annals of Botany Company. All rights reserved. For Permissions, please email: journals.permissions@oup.com variety', writing that he was comparing bulbs obtained from Italy, Greece, Austria, Spain, Kashmir (as Cashmere) and China. Some authors have concluded that there is little or no genetic variation in saffron (Rubio-Moraga *et al.*, 2009; Fluch *et al.*, 2010), although other recent studies have indicated limited genetic diversity within the species (Álvarez-Ortí *et al.*, 2004; Sik *et al.*, 2008; Nemati *et al.*, 2012). A reduction in production area in Europe over the last 300 years and more recent global distribution of planting material could result in the loss of any variation present: an EU programme to collect saffron from multiple locations, CROCUSBANK, addresses the question systematically (Fernández *et al.*, 2011; http://www.crocusbank.org).

There is interest in understanding the relationships and diversity in the whole genus Crocus by examining its genomic structure and phylogenetic relationships. The genus, and in particular the sections Nudiscapus and Crocus, are well circumscribed by both morphological and DNA analysis (Petersen et al., 2008; Harpke et al., 2014, 2015). In section Crocus, morphological analysis has been used to separate species into different series. Despite relatively robust separation of species by morphology (e.g. of flower parts, corm tunics or floral and vegetative development, many not obviously single-gene, autapomorphic characters), no DNA markers have resolved the natural relationships [e.g. plastid and nuclear sequences (Seberg and Petersen, 2009; Harpke et al., 2014, 2015; Larsen et al., 2015), repetitive DNA (Frello and Heslop-Harrison, 2000) and anonymous polymorphic markers (Fluch et al., 2010)] and indeed DNA markers do not support all the series consistently and robustly. The interspecific hybrid garden-origin Crocus 'Golden Yellow' (3x) and 'Stellaris' (2x) (Ørgaard et al., 1995) and C. sativus (3x) are well known, easily propagated vegetatively and are successful in cultivation. However, among the large number of species for which there have been studies of chromosome number, morphology and fertility, apart from C. sativus there were few reports before Harpke et al. (2015) discussing hybrid species of evolutionarily recent origin, and there are few species recognized as tetraploids. There are diploid and tetraploid members of some single species; in C. vernus, Frello and Heslop-Harrison (2000) reported major polymorphisms in karyotypes of ten diploid accessions, but the chromosomes in a 2n = 16 accession differed from those in all diploids. There is also no consensus about the ancestors of saffron (Maw, 1886; Mathew, 1982; Frello et al., 2004; Petersen et al., 2008; Erol et al., 2013; Harpke et al., 2013; Izadpanah et al., 2014).

Although DNA markers can often resolve questions about taxonomy and domestication (e.g. Parker *et al.*, 2014), Seberg and Petersen (2009) concluded that, for a plastid phylogeny alone, some 5800 bp of sequence would be needed to identify all *Crocus* species. Even this large amount of targeted sequencing would only identify the maternal parent in hybrids and would weight species delimitation to plastid genome evolution, and karyotype evolution, polyploidy, introgression or back-crossing would not be taken into consideration.

Apart from polyploidy which has played a significant role in plant speciation (see Levin, 2013), much of the DNA in the plant genome is associated with duplications or various classes of repetitive DNA including transposable elements (TEs) and satellite sequences (Kubis *et al.*, 2003; Heslop-Harrison and Schwarzacher, 2011; Estep *et al.*, 2013). TEs play an important role in the structure and evolutionary dynamics of the genomes, and retrotransposons are perhaps the most ancient components and make up the bulk of angiosperm genomes (Heslop-Harrison and Schmidt, 2012). Inter-retroelement amplified polymorphism (IRAP), using PCR primers facing outwards from terminal repeats of retroelements, allows measurement of polymorphisms arising from retrotransposon insertion. The ubiquitous nature, high copy number, diversification, amplification, movement and widespread chromosomal distribution of retrotransposons make these elements ideal for the development of such molecular markers (Teo *et al.*, 2005; Saeidi *et al.*, 2008; Kalendar *et al.*, 2011; Menzel *et al.*, 2014) to serve as biodiversity indicators, establish pedigrees of lines and allow inference of the evolutionary history and phylogeny of species.

Here, we aimed to measure the diversity in IRAP pattern in *Crocus* series *Crocus* and between individual accessions of saffron. We also aimed to find evidence for the single or multiple origins of *C. sativus*, to identify candidate ancestral species of saffron and to understand the relationships and genomic structures of species in the genus *Crocus*.

MATERIALS AND METHODS

Plant materials and genomic DNA extraction

The *Crocus* species, their sources and relevant CROCUSBANK accession numbers are listed in Table 1. Total genomic DNA was extracted from young leaves of single plants of the accessions using standard cetyltrimethylammonium bromide (CTAB) methods.

IRAP amplifications

Eleven IRAP primers previously designed to the conserved long terminal repeat (LTR) regions of retrotransposons were applied in the current study. Nucleotide sequences of the IRAP markers, GenBank accession number, position, orientation and original source are given in Table 2. IRAP primers were tested as single primers and in all 66possible combinations. PCR mixtures, amplification conditions and gel electrophoresis were modified from Teo et al. (2005). IRAP primers amplifying Crocus DNA are shown with experimentally optimized annealing temperatures in Table 2. DNA was amplified using a T professional Gradient Thermocycler (Biometra) in a $15 \,\mu L$ reaction mixture containing 50-100 ng of template DNA, 1 × Kapa Biosystems buffer A [750 mM Tris-HCl pH 8.8, 200 mM (NH₄)₂SO₄, 15 mM MgCl₂, 0.1 % Tween-20], 1.5 mM MgCl₂, 200 μ M dNTPs (Bioline), 0.6 μ M of each primer and 0.5 U of Kapa Taq DNA polymerase (Kapa Biosystems, USA). PCR conditions were: 95 °C for 2 min, followed by 30 cycles at 95 °C of 1 min, 40–62 °C for 1 min, ramp +0.5 °C to 72 °C, for 2 min and adding 3 s per cycle, with a final extension of 10 min at 72 °C followed by holding the block at 16 °C. Amplification of PCR products was confirmed on 2 % (w/v) agarose gels prepared by mixing normal (Bioline) and Hi-Res Super AGTC Agarose (Geneflow, UK) in ratios of 3:1 and run on at 7 V cm^{-1} for 45–60 min or a slow speed of 15 V for 15 h, visualized by staining with $0.5 \,\mu \text{g mL}^{-1}$ ethidium bromide. The reproducibility of amplified fragments was confirmed by repeating all reactions twice and using duplicate DNA extractions.

No.	Section	Series	Species	Sub-taxon/ variety	CrocusBank accession	University of Leicester code	Source
1	Crocus	Crocus	C. sativus	_	BCU002746	CsatP09	Pottertons Nursery (UK)
2	Crocus	Crocus	C. sativus	-	BCU002744	CstVD09	JW Dix Export (The Netherlands)
3	Crocus	Crocus	C. sativus	-		CstPER09	J.Perez (Spain)
4	Crocus	Crocus	C. sativus	-		CstSUSD09	Suttons Nursery (UK)
5	Crocus	Crocus	C. sativus	cashmeriensis	BCU002584	CstCD09	JW Dix Export (The Netherlands)
6	Crocus	Crocus	C. sativus	Kashmir		Cstkf09	Srinagar, Kashmir
7	Crocus	Crocus	C. sativus cartwrightianus*	Albus	BCU002754	CstcP09	Pottertons Nursery (UK)
8	Crocus	Crocus	C. cartwrightianus	-	BCU002747	CcwBD09	JW Dix Export (The Netherlands)
9	Crocus	Crocus	C. cartwrightianus	Albus	BCU002766	CcwAD08	JW Dix Export (The Netherlands)
10	Crocus	Crocus	C. cartwrightianus	CEH.613	BCU002771	CcrCR09	Rareplant Nursery (UK)
11	Crocus	Crocus	C. pallasii	turcicus	BCU002748	CpltR09	Rareplant Nursery (UK)
12	Crocus	Crocus	C. pallasii	pallasii	BCU002767	CplVD09	JW Dix Export (The Netherlands)
13	Crocus	Crocus	C. pallasii	dicpataceus	BCU002759	CplDD09	JW Dix Export (The Netherlands)
14	Crocus	Crocus	C. mathewii	*		CmatD08	JW Dix Export (The Netherlands)
15	Crocus	Crocus	C. mathewii	HKEP.9291		CmtHR09	Rareplant Nursery (UK)
16	Crocus	Crocus	C. thomasii		BCU002751	CtmVD09	JW Dix Export (The Netherlands)
17	Crocus	Crocus	C. thomasii	MS 978		CtomI09	Matera, Italy
18	Crocus	Crocus	C. asumaniae	white	BCU002757	CasWD09	JW Dix Export (The Netherlands)
19	Crocus	Crocus	C. asumaniae	ʻalba'	BCU002760	CasAD09	JW Dix Export (The Netherlands)
20	Crocus	Crocus	C. asumaniae	S9104		CasAT09	Aseki Turkey
21	Crocus	Crocus	C. oreocreticus	VV.CR.114	BCU002774	CorVR09	Rareplant Nursery (UK)
22	Crocus	Crocus	C. oreocreticus		BCU002756	CorVD09	JW Dix Export (The Netherlands)
23	Crocus	Crocus	C. hadriaticus		BCU002764	ChdWD09	JW Dix Export (The Netherlands)
24	Crocus	Crocus	C. hadriaticus	'Indian summer'	BCU002770	ChaIR09	Rareplant Nursery (UK)
25	Crocus	Crocus	C. hadriaticus	Alepohori (AH8682)		ChdARD09	Rareplant Nursery (UK)
26	Crocus	Verni	C. vernus	1 ()	BCU001854	VER01	1 2 7
27	Crocus	Verni	C. tommasinianus	'lilac beauty'	BCU002765	CtmLD09	JW Dix Export (The Netherlands)
28	Crocus	Verni	C. tommasinianus	'barr purple'	BCU002768	CtmBD09	JW Dix Export (The Netherlands)
29	Crocus	Verni	C. tommasinianus	'rubinetta'	BCU002762	CtmTD09	JW Dix Export (The Netherlands)
30	Crocus	Verni	C. tommasinianus	'albus'	BCU002763	CtmAD09	JW Dix Export (The Netherlands)
31	Crocus	Versicolores	C. versicolor	'picturatus'	BCU002761	CvrPP09	Pottertons Nursery (UK)
32	Crocus	Longiflori (Verni [†])	C. niveus	1		CnivD08	JW Dix Export (The Netherlands)
33	Crocus	Longiflori (Verni [†])	C. goulimyi	'leucanthus'	BCU002755	CgulD08	JW Dix Export (The Netherlands)
34	Crocus	Kotschyani	C. kotschyanus	kotschyanus		CkotP09	Pottertons Nursery (UK)
35	Crocus	Kotschyani	C. kotschyanus	Zonatus		Ckot/z08	Garden Source
36	Nudiscapus	Reticulati	C. angustifolius			CangP09	Pottertons Nursery (UK)

TABLE 1. The taxonomic position of accessions and species from the genus Crocus used in the current study

*The accession was purchased under this unrecognized name. It has morphological similarities to *C. cartwrightianus* but is not this species. [†]Series revised to be Verni by Harpke *et al.* (2015); the revision would be consistent with the tree in Fig. 4.

Genetic variability and phylogenetic analysis

For each IRAP fragment, presence/absence was scored on gel images in Adobe Photoshop, and binary matrices were assembled as spreadsheets. Basic statistics including the total number of alleles, major allele frequency, genetic diversity and polymorphism information content (PIC) values were determined using PowerMarker version 3.25 (Liu and Muse, 2005). The data were also used to infer the relationships of *Crocus* species based on the UPGMA method (Saitou and Nei, 1987) with 1000 bootstrap replicates using PowerMarker. *Crocus angustifolius* (section *Nudiscapus*) was used as an outgroup. The consensus bootstrap tree was generated using Geneious v.7 (Biomatters Ltd, Auckland, New Zealand).

RESULTS

IRAP amplification and diversity within Crocus species

Out of 66 IRAP primers and primer combinations tested, 63 amplified multiple and distinguishable fragments from the genomic DNA of all *Crocus* species and accessions (Tables 1 and 2). In our analysis, the Sukkula primer, either alone or in

combination with other primers, produced the highest number of IRAP bands (Figs 1 and 2). The low number of fragments indicates that there are fewer elements of the respective retrotransposon in the *Crocus* genome (as might be expected given the primer design from heterologous species) or that the elements are distantly spaced. At least two PCRs were used for each primer combination. The analysis included 35 accessions, and 4521 distinct bands from 100 bp to approx. 4kb were amplified (Figs 1 and 2 show representative IRAP gel images). Every fragment was absent in one or more species; no fragment was specific to a single species and high levels of polymorphism were evident within all species (except *C. sativus*) and between species (Figs 1–3).

IRAP amplification and diversity within saffron accessions

DNA from multiple saffron accessions from different sources and representing diverse geographical collections (Table 1; Figs 2 and 3) was amplified with 11 IRAP markers. All bands were monomorphic, and no polymorphism could be confirmed. Other *Crocus* species and the garden-origin accession named

No.	Marker name	Retrotransposon name and orientation	Sequence (5'–3')	Accession	Position	Reference/source
1	LTR6150	BARE-1 ←	CTGGTTCGGCCCATGTCTATGTATCCACACATGGTA	Z17327	418-439	Kalendar et al. (1999)
2	LTR6149	BARE-1 \rightarrow	CTCGCTCGCCCACTACATCAACCGCGTTTATT	Z17327	1993-2012	Kalendar et al. (1999)
3	Nikita	Nikita →	CGCATTTGTTCAAGCCTAAACC	AY078073 AY078074 AY078075	1–22	Leigh <i>et al.</i> (2003)
4	IRAP Crocus Nikita	Nikita	CAGTTTTGATCAAGTCATAACC	AJ131448	15–36	Modified after Leigh <i>et al.</i> (2003) by Heslop-Harrison, Vikgren and Ørgaard (unpublished)
5	Sukkula	Sukkula \rightarrow	GATAGGGTCGCATCTTGGGCGTGAC	AY034376	10662-10685	Mannien et al. (2000)
6	IRAP Crocus Sukkula	Sukkula	AACAGAAGTAGTGGCAGCTTGAGAG	AY245374	1023	Modified after Leigh <i>et al.</i> (2003) by Heslop-Harrison, Vikgren and Ørgaard (unpublished)
7	ReverseTy1	W1, W3, W7, W8 ←	CCYTGNAYYAANGCNGT	AF416815 AF416816 AF416817 AF416818	1–17	Teo <i>et al.</i> (2005)
8	Reverse TY2	$\begin{array}{c} W1, W3,\\ W7, W8 \rightarrow \end{array}$	TRGTARAGRAGNTGRAT	AF416815 AF416816 AF416817 AF416818	252–269	Teo <i>et al.</i> (2005)
9	3' LTR	BARE-1 \rightarrow	TGTTTCCCATGCGACGTTCCCCAACA	Z17327	2112-2138	Teo et al. (2005)
10	IRAP Crocus 5' LTR		CCATAGCTTGTAGGGCGTCTCCCCA	AY245373		Modified after Leigh <i>et al.</i> (2003) by Heslop-Harrison, Vikgren and Ørgaard (unpublished)
11	5' LTR1	BARE-1 ←	TTGCCTCTAGGGCATATTTCCAACA	Z17327	1–26	Teo <i>et al.</i> (2005)

 TABLE 2. Characteristics of IRAP primers used for amplifications

 \rightarrow Primer direction with respect to the first open reading frame of each retrotransposon.

 \leftarrow Y = C + T, N = A + G + C + T, R = A + G nucleotides.

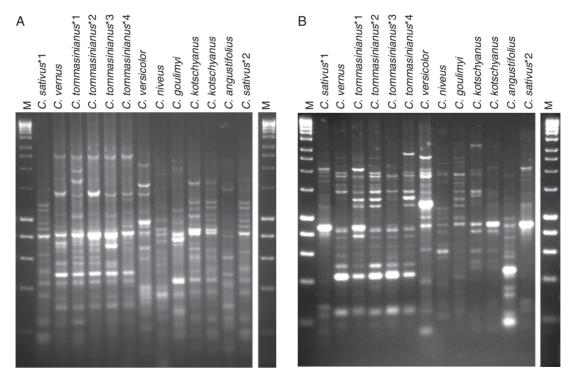


Fig. 1. IRAP amplification from eight *Crocus* species amplified with (A) Nikita and IRAP *Crocus* Sukkula primers and (B) IRAP *Crocus* 5' LTR primer. Species names are given; *C. sativus**1 (CsatPER09), *C. sativus**2 (CstVD09), *C. tommasinianus**1 (CtmLD09), *C. tommasinianus**2 (CtmBD09), *C. tommasinianus**3 (CtmTD09), *C. tommasinianus**4 (CtmAD09), *C. kotschyanus**1 (subsp. *kotschyanus*, CkotP09) and *C. kotschyanus**2 (var. Zonatus, Ckot/z08). M, DNA size marker HyperLadder I (200 bp steps to 1 kb, then 1.5, 2, 3, 4 and 5 kb).

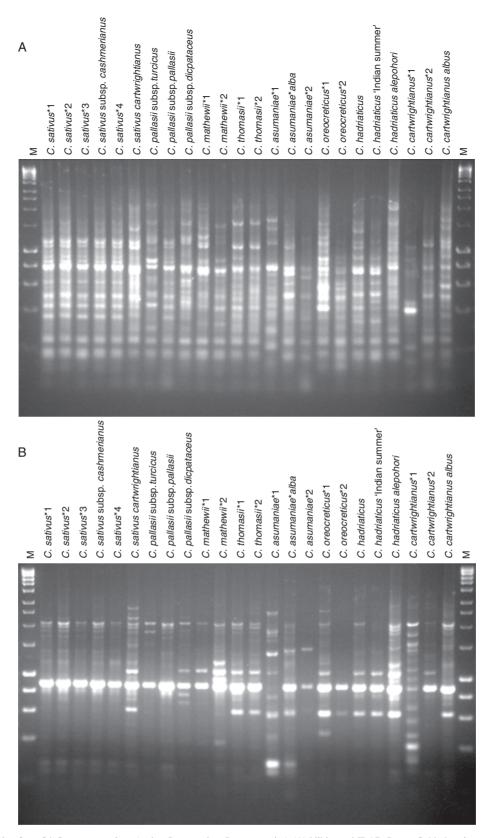


FIG. 2. IRAP amplification from 24 Crocus accessions (series Crocus, nine Crocus species) (A) Nikita and IRAP Crocus Sukkula primers and (B) IRAP Crocus 5 'LTR primer. Crocus sativus*1 (CsatPER09), C. sativus*2 (Cstkf09), C. sativus*3 (CstVD09), C. sativus*4 (CsatP09), C. mathweii*1 (CmatD08), C. mathweii*2 (HKEP.9291, CmtHR09), C. thomasii*1 (CtmVD09), C. thomasii*2 (MS978, CtomI09), C. asumaniae*1 (CasWD09), C. asumaniae*2 (CasAT09), C. oreocreticus*1 (CorVR09), C. oreocreticus*2 (CorVD09), C. cartwrightianus*1 (CcwBD09) and C. cartwrightianus*2 (CcrCR09). M, DNA size marker HyperLadder I as in Fig. 1.

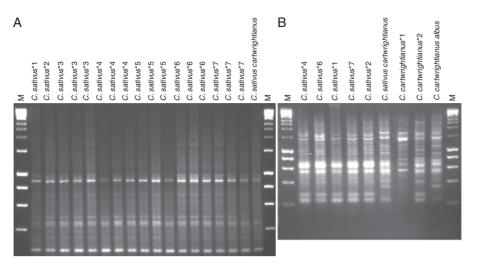


Fig. 3. Examples of the IRAP pattern from multiple saffron accessions and duplicate DNA extractions using the reverse TY2 primer (A) and Nikita (B) with three *C. cartwrightianus* and a hybrid accession. Saffron accessions represent global collections (Table 1), but no variation was evident (except for small differences ascribed to amplification). *1, JW Dix Export, The Netherlands (2007); *2, Pottertons Nursery, UK; *3, JW Dix Export, The Netherlands (2010); *4, CrocusBank, Spain; *5, Suttons Nursery, UK; *6, Kashmir, India; *7, var. *cashmeriensis* from JW Dix Export, The Netherlands (2009). M, HyperLadder I (200 bp steps, then 1·5 and 2 kb).

C. sativus cartwrightianus were used in parallel amplifications and showed evident polymorphism (Figs 2 and 3).

C. hadriaticus) of bands. Species from the other series in the genus had many fewer shared bands (8.4-15.7 %).

Genome diversity and relationships among the Crocus species

To evaluate the relationships, the binary data obtained from 185 913 IRAP fragment scores were used for the reconstruction of distances between species by the UPGMA method with C. angustifolius from section Nudiscapus as an outgroup. The IRAP tree, including 35 accessions of 16 taxa (including one of garden-origin and three subspecies), was divided into three well-supported clusters A1, A2 and B (Fig. 4) with three outlying accessions (cluster A). Cluster A1 included five species (C. cartwrightianus, C. oreocreticus, C. hadriaticus, C. asumaniae and C. thomasii) and cluster A2 comprised 11 accessions belonging to five taxa and three species: C. mathewii, two subspecies of C. pallasii, C. sativus and a hybrid. Cluster B comprises nine accessions from six Crocus species: C. vernus and C. tommasinianus from Crocus series Verni; C. kotchvanus, C. niveus, C. goulimyi and C. versicolor from other series of Crocus (Harpke et al., 2015, merged Longiflori into Verni).

We considered pooling the data by species, but there was substantial intraspecific variation between accessions of some species, particularly *C. cartwrightianus* and *C. hadriaticus*, and the analysis grouped these as well-supported branches with the remaining species. The analysis would give such clusters where some bands of similar sizes in different species were not identical by descent, or else some genetic introgression had occurred so species were of hybrid origin.

Because the UPGMA tree does not account for shared bands due to hybridity, we examined the bands shared between *C. sativus* and all other *Crocus* accessions. Of the 477 bands in the profile of *C. sativus*, 270 (56.6 %) were shared with *C. pallasii* subsp. *pallasii* and 41.1 % with *C. pallasii* subsp. *turcicus*. The remaining series *Crocus* accessions shared between 36.9 % (an accession of *C. thomasii*) and 26.2 % (an accession of

DISCUSSION

The IRAP primers used here were designed from LTR sequences of retroelements found in species other than Crocus (Table 2); the great majority were able to amplify multiple loci from *Crocus* species, indicating the transferable nature of the retrotransposon-based markers. Here, as elsewhere, they proved to give good markers for genome-wide assessment of diversity and relationships (Kalendar et al., 1999, 2011; Teo et al., 2005; Saeidi et al., 2008). IRAPs do not analyse nuclear copies of plastid or mitochondrial genes, polymorphisms in possible multicopy genes and genome duplications or potential variation in priming sites; unlike amplification with shorter primers, they are also relatively insensitive to the DNA extraction protocol and amplification conditions. It is frequently noted that polyploids do not give the sum of bands of their parents, presumably because of primer competition, nesting or mismatches [e.g. see Teo et al. (2005) with AA, BB and hybrid bananas, or Saeidi et al. (2008) with diploid and hexaploid wheats (their table 2 shows no correlation of band number with ploidy); band patterns were more additive in cell fusion hybrids of tobacco (Patel et al., 2011)].

The relatedness of *C. sativus* (saffron) to other members of the *Crocus* series *Crocus* was evident (Figs 2 and 4, cluster A). IRAPs did not generate a tree position suggesting origination from one diploid species, and autopolyploidy was not supported (Figs 1–3). Among the *C. sativus* accessions analysed here, there were no bands unique to *C. sativus* (Figs 1 and 2). The lack of variation within saffron (Fig. 3) agrees with reports from Chappellier (1900) and earlier from morphology. Recent reports with a range of DNA markers (Álvarez-Ortí *et al.*, 2004; Nehvi *et al.*, 2007; Agayev *et al.*, 2009; Fluch *et al.*, 2010; Fernandez *et al.*, 2011; Nemati *et al.*, 2012; Larsen *et al.*, 2015) show low variation, although differences in genes,

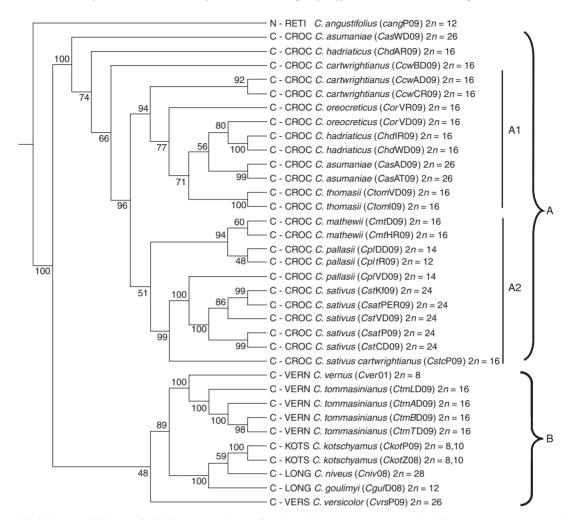


FIG. 4. Consensus UPGMA tree of IRAP data for 35 *Crocus* accessions (16 species, with chromosome number 2*n*). The bootstrap consensus tree is inferred from 1000 replicates computed by PowerMarker software (percentage support shown at nodes) with well-supported branches for clusters A1, A2 and B. The phylogenetic analysis included 11 primers in 63 combinations (4521 bands). C section *Crocus*; N section *Nudiscapus* set as outgroup; CROC series *Crocus*; VERN series *Verni*; LONG series *Longiflori* (revised to series *Verni* by Harpke *et al.*, 2015); VERS series *Versicolores*; KOTS series *Kotschyani*; SPEC series *Speciosi*; RETI series *Reticulati*; FLAV series *Flavi*; LAEV series *Laevigatae*.

regulatory sequences or repetitive DNA may be present outside the IRAP sequences surveyed. These may give differences that can be selected and propagated. Phenotypic differences in saffron can also be attributed to differences in climate and cultivation practices, and 'quality' of saffron is therefore more likely to be defined by growing conditions (e.g. soil, water, temperature and altitude), collection and processing techniques, and is less dependent on the origin of the corm (Álvarez-Ortí *et al.*, 2004; Agayev *et al.*, 2009; Maggi *et al.*, 2011; Torelli *et al.*, 2014).

Subgenus *Crocus* comprises two sections, *Crocus* and *Nudiscapus*, both probably monophyletic (Harpke *et al.*, 2013, 2015). Section *Crocus* series *Crocus* is a strongly supported monophyletic group (Petersen *et al.*, 2008). The IRAP analyses (UPGMA based on shared bands, Fig. 4) supports separation of some species of series *Crocus* (cluster A) from other species of series *Crocus* (cluster B). The tree topology (Fig. 4) is in accordance with Mathew (1982) and Petersen *et al.* (2008), and the position of all species is satisfied at series level. This indicates the close relationship of the two; IRAP, reflecting retroelement

diversification and inheritance, supports morphological analyses.

Within series *Crocus*, phylogenetic relationships are not well resolved, compatible with the uncertainty of taxonomy and phylogeny in the literature with many methods [Mathew (1982) with chromosome number and morphology; Seberg and Petersen (2009) with nuclear and plastid DNA markers; and Frello *et al.* (2004) with repetitive DNA]. Further, sub-branching at the accession level indicates genomic diversity within accessions, and high bootstrap support values show the confidence and usefulness of IRAPs in discriminating species and accessions (see Figs 3 and 4).

Here, a few accessions of *C. hadriaticus 'alepohori'* clustered with *C. oreocreticus*, not *C. hadriaticus* 'Indian summer' (sub-cluster A1, Fig. 4), and one of the three *C. asumaniae* accessions 'white' also did not cluster with *C. asumaniae 'alba'* and 'S9104' (see sub-clusters A1 and outlying species). Both accessions of *C. asumaniae* 'white' and '*alba'* were obtained from cultivation in The Netherlands, whereas 'S9104' originated from Aseki, Turkey, and *C. hadriaticus 'alepohori'* and

'Indian summer' had the same nursery origin (The Netherlands). Thus the variation in accessions may partly be attributed to their different origin or hybridization stress imposed under cultivation. It is known that stress or unusual environmental stimuli may induce heritable changes, and in plants is associated with the accumulation and rise in the activities of TEs (Kubis *et al.*, 2003; Cullis, 2005; Ågren and Wright, 2011; Heslop-Harrison and Schwarzacher, 2011). The CROCUSBANK collection is a comprehensive collection of the genus (Fernández *et al.*, 2010; see www.crocusbank.org). One of the three *C. cartwrightianus* accessions (CcwBD09 from The Netherlands) was unrelated to the true *C. cartwrightianus* accessions, and '*C. sativus cartwrightianus*' had an invalid name and was different.

In the analysis, high variation between accessions within each of the species (other than *C. sativus*) is evident, and it is clear that much more extensive collections will be required to circumscribe the taxonomic units, as reported by Larsen *et al.* (2015). Many wild collections, although locally abundant in their native range, are difficult to maintain in cultivation. The evidence also suggests that interspecific hybridization occurs occasionally, with consequences allowing gene flow, homogenization and hybrid speciation, leading to uncertain delimitation of species.

Most current approaches to phylogenetic tree construction based on multiple co-dominant DNA markers are not able to identify species of hybrid origin because of their reliance on monophyly: with a hybrid of two species with few shared markers, the hybrid will be resolved on a separate branch to the parents. IRAP markers are not exclusively dominant or co-dominant (in common with most other DNA marker systems when amphipolyploid or hybrid species are compared with diploids), so a hybrid will not normally share all the bands of both progenitors. In some cases, it is apparent that there is rapid loss of DNA sequences in new hybrids (Ozkan *et al.*, 2001; Ma and Gustafson *et al.*, 2008).

No unequivocal parents (ancestors) of C. sativus emerge from the IRAP analysis of shared bands. Are there other Crocus species that remain to be discovered either in the wild or misnamed in herbarium collections? Or are the two species we have not included from series Crocus (C. moabiaticus and C. nagabensis) the ancestors? Crocus moabiticus is a newly identified species with a narrow range in Jordan. It is unlikely that there are significant new species as the regions of occurrence in Europe and the Middle East are well collected for conspicuous plants, although Harpke et al. (2015) suggested substantial changes to the taxonomy recognizing many new species previously included within other species, not least the widespread C. vernus. Our results support the comments of Larsen et al. (2015) that it is important to include many individual plants from different populations in future molecular studies, and that there is likely to be substantial gene flow between populations. Given the likelihood of taxonomic revision of the series and recognition of more taxa (Harpke et al., 2015), it will also be important to ensure documentation of the collections and ascertainment of chromosome numbers.

Different species of *Crocus* series *Crocus* have been suggested as potential ancestors of *C. sativus*. *Crocus cartwrightianus* shows morphological similarity to *C. sativus*, and studies that used morphology and karyotype analysis of the species

allied to C. sativus demonstrated that C. cartwrightianus is one of the progenitors of C. sativus (Maw, 1886; Mathew, 1982; Grilli Caiola et al., 2004; Larsen et al., 2015). Further, the diploid C. oreocreticus is similar to C. cartwrightianus and has also been considered as a possible ancestor of C. sativus (Burtt, 1948). Repetitive DNA sequences have also been employed in phylogenetic analysis of the genus, but their contribution to the understanding of Crocus phylogenetics was limited (Frello and Heslop-Harrison, 2000; Frello et al., 2004). Their results did not support all parts of Mathew's classification, leading them to discuss the possibility of far-reaching hybridization and rapid speciation in the genus. In the case of allotriploid saffron, C. cartwrightianus, C. hadriaticus, C. oreocreticus (Jacobsen and Ørgaard, 2004; Agayev et al., 2010) or C. thomasii and C. pallasii or C. cartwrightianus and C. pallasii (Tammaro, 1990) have been proposed as candidate ancestral species, with each contributing the basic set of x = 8 chromosomes.

Although the overall IRAP profile of C. sativus was different from that of all species analysed here, among the analysed species, it was most similar to that of C. pallasii subsp. pallasii (Figs 2 and 3) and hence this subspecies (albeit with 2n = 2x = 14 chromosomes) is suggested to be a candidate ancestor for C. sativus. Sanei et al. (2007) also reported that C. pallasii (without specifying the subspecies, which were substantially different in our analysis) is one of the close relatives of C. sativus based on karyotype data; Erol et al. (2013) also found the maximum similarity between an accession of C. pallasii subsp. pallasii and C. sativus. Amplified fragment length polymorphism (AFLP) fingerprinting revealed C. cartwrightianus and C. thomasii to be the closest relatives of C. sativus (Zubor et al., 2004), like the random amplified poymorphic DNA (RAPD) data of Grilli Caiola et al. (2004). Flow cytometric analysis of diploid species of Crocus series Crocus including C. cartwrightianus and C. thomasii found C. cartwrightianus to be the most likely ancestor of C. sativus (Brandizzi and Grilli Caiola, 1998). Here, C. sativus is close to these diploid species on separate branches (Fig. 4). Based on IRAP markers, C. almehensis and C. michelosnii were shown to be possible ancestors of C. sativus (Alavi-Kia et al., 2008). Petersen et al. (2008) analysed five plastid regions; their analysis included 86 recognized species of the genus and their study also found C. cartwrightianus to be closely related to C. sativus. The results here show considerable variation between accessions of C. cartwrightianus in contrast to the lack of variation between geographically diverse C. sativus accessions. Notably, our accession from the UK nursery Rare Plants (CcartRP07) shared most bands with the saffron accessions in most of the IRAP primer combinations, and hence it is suggested that it is most similar to one of the donors of the C. sativus genome. Plastid, ribosomal and nuclear single-copy gene sequences, focused on those used for phylogenetic analysis, suggested C. cartwrightianus and C. pallasii as ancestral species of C. sativus (Harpke et al., 2013), and their results are in general agreement with the findings here. Recent intersimple sequence repeat (ISSR) analyses established that C. cartwrightianus 'albus' is more closely related to C. sativus than to C. cartwrightianus (Rubio-Moraga et al., 2009). Clearly more accessions of C. cartwrightianus must be genotyped to find that closest to C. sativus.

For most crops, domestication is seen as a bottleneck reducing genetic variation; further artificial selection has advantages in maintaining its genetic characteristics, but causes reduction in genetic diversity. Given the high levels of polymorphism between the species and even individual accessions, minimal if any variation was evident in *C. sativus*, despite accessions from a broad geographical range being included. Thus it is most likely that a single hybrid gave rise to the saffron corms now grown. Wide genetic diversity is of importance for the development of improved varieties, and it will certainly be valuable to resynthesize *C. sativus* as a triploid from crossing *C. cartwrightianus* and *C. pallasii* subsp. *pallasii* for improvement of saffron cultivars, exploitation of genetic diversity and conservation of the *Crocus* germplasm.

ACKNOWLEDGEMENTS

We are grateful to the European Commission through the 018-AGRI GEN RES Action 'Genetic Resources of Saffron and Allies (*Crocus* spp.): CROCUSBANK'; and the COST FA1101 Action 'Omics Technologies for Crop Improvement, Traceability, Determination of Authenticity, Adulteration and Origin in Saffron' for support of this project. We thank all the partners for many useful discussions. N.A. thanks the University of Umm al-Qura, Saudi Arabia, for support.

LITERATURE CITED

- Agayev YM, Shakib AM, Soheilivand S, Fathi M. 2006. Breeding of saffron (*Crocus sativus*): possibilities and problems. *Acta Horticulturae (ISHS)* **739**: 203–207.
- Agayev YM, Fernández JA, Zarifi E. 2009. Clonal selection of saffron (Crocus sativus L.): the first optimistic experimental results. Euphytica 169: 81–99.
- Agayev YM, Zarifi E, Fernández JA. 2010. A study of karyotypes in the Crocus sativus L. aggregate and origin of cultivated saffron. Acta Horticulturae 850: 47–54.
- Ågren JA, Wright SI. 2011. Co-evolution between transposable elements and their hosts: a major factor in genome size evolution? *Chromosome Research* 19: 777–786.
- Alavi-Kia SS, Mohammadi SA, Aharizad S, Moghaddam M. 2008. Analysis of genetic diversity and phylogenetic relationships in *Crocus* genus of Iran using inter-retrotransposon amplified polymorphism. *Biotechnology and Biotechnological Equipment* 22: 795–800.
- Álvarez-Ortí M, Gómez LG, Rubio A, et al. 2004. Development and gene expression in saffron corms. Acta Horticulturae 650: 141–153.
- Babaei S, Talebi M, Bahar M, Zeinali H. 2014. Analysis of genetic diversity among saffron (*Crocus sativus*) accessions from different regions of Iran as revealed by SRAP markers. *Scientia Horticulturae* 171: 27–31.
- Brandizzi F, Grilli Caiola M. 1998. Flow cytometric analysis of nuclear DNA in *Crocus sativus* and allies (Iridaceae). *Plant Systematics and Evolution* 211: 149–154.
- Burtt RL. 1948. Crocus oreocreticus. Phyton 1: 224-225.
- Chappellier P. 1900. Creation of an improved variety of Crocus sativus. Journal of the Royal Horticultural Society XXIV Hybrid Conference Report 275– 277.
- Cullis CA. 2005. Mechanisms and control of rapid genomic changes in flax. Annals of Botany 95: 201–206.
- Erol O, Kaya HB, Şik L, Tuna M, Can L, Tanyolaç MB. 2013. The genus Crocus, series Crocus (Iridaceae) in Turkey and 2 East Aegean islands: a genetic approach. Turkish Journal of Biology 38: 48–62.
- Estep MC, DeBarry JD, Bennetzen JL. 2013. The dynamics of LTR retrotransposon accumulation across 25 million years of panicoid grass evolution. *Heredity* 110: 194–204.
- Fernández JA, Santana O, Guardiola J-L, et al. 2011. The world saffron and Crocus collection: strategies for establishment, management, characterisation and utilisation. Genetic Resources and Crop Evolution 58: 125–137.
- Fluch S, Hohl K, Stierschneider M, Kopecky D. 2010. Crocus sativus L. Molecular evidence on its clonal origin. Acta Horticulturae 850: 41–46.

- Frello S, Heslop-Harrison JS. 2000. Chromosomal variation in *Crocus vernus* Hill (Iridaceae) investigated by *in situ* hybridization of rDNA and a tandemly repeated sequence. *Annals of Botany* 86: 317–322.
- Frello S, Ørgaard M, Jacobsen N, Heslop-Harrison JS. 2004. The genomic organization and evolutionary distribution of a tandemly repeated DNA sequence family in the genus *Crocus* (Iridaceae). *Hereditas* 141: 81–88.
- **Ghaffari SM, Bagheri A. 2009.** Stigma variability in saffron (*Crocus sativus* L.). *African Journal of Biotechnology* **8**: 601–604.
- Grilli Caiola M, Canini A. 2010. Looking for saffron's (Crocus sativus L.) parents. In: Husaini AM, ed. Saffron. Functional Plant Science and Biotechnology 4: 1–14.
- Grilli Caiola M, Caputto P, Zanier R. 2004. RAPD analysis in Crocus sativus L. accessions and related Crocus species. Biologia Plantarum 48: 375–380.
- Harpke D, Meng S, Rutten T, Kerndorff H, Blattner FR. 2013. Phylogeny of Crocus (Iridaceae) based on one chloroplast and two nuclear loci: ancient hybridization and chromosome number evolution. *Molecular Phylogenetics* and Evolution 66: 617–627.
- Harpke D, Peruzzi L, Kerndorff H, et al. 2014. Phylogeny, geographic distribution and new taxonomic circumscription of the Crocus reticulatus species group (Iridaceae). Turkish Journal of Botany 38: 1182–1198.
- Harpke D, Carta A, Tomović G, et al. 2015. Phylogeny, karyotype evolution and taxonomy of Crocus ser. Verni (Iridaceae). Plant Systematics and Evolution 301: 309–325.
- Heslop-Harrison JS, Schmidt T. 2012. Plant nuclear genome composition. eLS.
- Heslop-Harrison JS, Schwarzacher T. 2011. Organisation of the plant genome in chromosomes. *The Plant Journal* 66: 18–33.
- Heslop-Harrison JS, Schwarzacher T. 2012. Genetics and genomics of crop domestication. In: Altman A, Hasegawa PM, eds. *Plant biotechnology and agriculture: prospects for the 21st century*. Dordrecht: Elsevier, 3–18.
- Izadpanah F, Kalantari S, Hassani ME, Naghavi MR, Shokrpour M. 2014. Variation in saffron (*Crocus sativus* L.) accessions and *Crocus* wild species by RAPD analysis. *Plant Systematics and Evolution* 300: 1941–1944.
- Jacobsen N, Ørgaard M. 2004. Crocus cartwrightianus on the Attica Peninsula. Acta Horticulturae 650: 65–69.
- Kalendar R, Grob T, Regina M, Suoniemi A, Schulman A. 1999. IRAP and REMAP: two new retrotransposon-based DNA fingerprinting techniques. *Theoretical and Applied Genetics* 98: 704–711.
- Kalendar R, Flavell AJ, Ellis THN, Sjakste T, Moisy C, Schulman AH. 2011. Analysis of plant diversity with retrotransposon based molecular markers. *Heredity* 106: 520–530.
- Kubis SE, Castilho AM, Vershinin AV, Heslop-Harrison JS. 2003. Retroelements, transposons and methylation status in the genome of oil palm (*Elaeis guineensis*) and the relationship to somaclonal variation. *Plant Molecular Biology* 52: 69–79.
- Larsen B, Orabi J, Pedersen C, Ørgaard M. 2015. Large intraspecific genetic variation within the Saffron–Crocus group (*Crocus L.*, Series *Crocus*; Iridaceae). *Plant Systematics and Evolution* 301: 425–437.
- Levin DA. 2013. The timetable for allopolyploidy in flowering plants. Annals of Botany 112: 1201–1208.
- Liu K, Muse SV. 2005. PowerMarker: integrated analysis environment for genetic marker data. *Bioinformatics* 21: 2128–2129.
- Ma XF, Gustafson JP. 2008. Allopolyploidization-accommodated genomic sequence changes in triticale. *Annals of Botany* 101: 825–832.
- Maggi L, Carmona M, Kelly SD, Marigheto N, Alonso GL. 2011. Geographical origin differentiation of saffron spice (*Crocus sativus* L. stigmas) – Preliminary investigation using chemical and multi-element (H, C, N) stable isotope analysis. *Food Chemistry* 128: 543–548.
- Mathew B. 1982. *The Crocus. A revision of the genus Crocus (Iridaceae).* Portland, OR: Timber Press.
- Maw G. 1886. A monograph of the genus Crocus. London: Dulau & Co.
- Menzel G, Heitkam T, Seibt KM, et al. 2014. The diversification and activity of hAT transposons in Musa genomes. Chromosome Research 22: 559–571.
- Molina RV, Guardiola JL, García-Luis D, et al. 2015. Descriptors for crocus (Crocus spp.). Rome: Bioversity International.
- Nehvi FA, Wani SA, Dar SA, Makhdoomi MI, Allie BA, Mir ZA. 2007. Biological interventions for enhancing saffron productivity in Kashmir. *Acta Horticulturae* **739**: 25–32.
- Nemati Z, Zeinalabedini M, Mardi M, Pirseyediand SM, Marashi SH, Nekoui SMK. 2012. Isolation and characterization of a first set of polymorphic microsatellite markers in saffron, *Crocus sativus* (Iridaceae). *American Journal of Botany* 99: e340–e343.
- Ørgaard M, Jacobsen N, Heslop-Harrison JS. 1995. The hybrid origin of two cultivars of *Crocus* (Iridaceae) analysed by molecular cytogenetics

including genomic Southern and *in situ* hybridization. Annals of Botany 76: 253–262.

- Ozkan H, Levy AA, Feldman M. 2001. Allopolyploidy-induced rapid genome evolution in the wheat (*Aegilops–Triticum*) group. *The Plant Cell* 13: 1735–1747.
- Parker KC, Trapnell DW, Hamrick JL, Hodgson WC. 2014. Genetic and morphological contrasts between wild and anthropogenic populations of *Agave parryi* var. *huachucensis* in south-eastern Arizona. *Annals of Botany* 113: 939–952.
- Patel D, Power JB, Anthony P, Badakshi F, Heslop-Harrison JS, Davey MR. 2011. Somatic hybrid plants of *Nicotiana* × sanderae (+) *N. debneyi* with fungal resistance to *Peronospora tabacina*. Annals of Botany 108: 809–819.
- Petersen G, Seberg O, Thorsøe S, Jørgensen T, Mathew B. 2008. A phylogeny of the genus *Crocus* (Iridaceae) based on sequence data from five plastid regions. *Taxon* 57: 487–499.
- Rubio-Moraga A, Castillo-López R, Gómez-Gómez L, Ahrazem O. 2009. Saffron is a monomorphic species as revealed by RAPD, ISSR and microsatellite analyses. *BMC Research Notes* 2: 189–193.
- Saeidi H, Rahiminejad MR, Heslop-Harrison JS. 2008. Retroelement insertional polymorphisms, diversity and phylogeography within diploid, D-genome Aegilops tauschii (Triticeae, Poaceae) sub-taxa in Iran. Annals of Botany 101: 855–861.
- Saitou N, Nei M. 1987. The Neighbor–Joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4: 406–425.
- Sanei M, Rahimyan H, Agayev YM, Soheilivand S. 2007. New cytotype of Crocus pallasii subsp. haussknechtii from west of Iran. Acta Horticulturae 739: 107–111.

- Seberg O, Petersen G. 2009. How many loci does it take to DNA barcode a crocus? *PLoS One* 4: e4598.
- Sik L, Candan F, Soya S, Karamenderes C, Kesercioglu T, Tanyolac B. 2008. Genetic variation among *Crocus sativus* L. species from western Turkey as revealed by RAPD and ISSR markers. *Journal of Applied Biolgical Science* 2: 73–78.
- Siracusa L, Gresta F, Avola G, et al. 2013. Agronomic, chemical and genetic variability of saffron (*Crocus sativus* L.) of different origin by LC-UV-vis-DAD and AFLP analyses. *Genetic Resources and Crop Evolution* 60: 711–721.
- Tammaro F. 1990. Crocus sativus L. cv di Navelli (L'Aquila saffron): environment cultivation, morphometric characteristic, active principles, uses. In: Tammaro F, Marra L, eds. Lo zafferano: Proceedings of the International Conference on Safforn (Crocus sativus L.). L'Aquila, Italy, 27–29 October 1989, 47–96.
- Teo CH, Tan SH, Ho CL, et al. 2005. Genome constitution and classification using retrotransposon-based markers in the orphan crop banana. Journal of Plant Biology 48: 96–105.
- Torelli A, Marieschi M, Bruni R. 2014. Authentication of saffron (Crocus sativus, L.) in different processed, retail products by means of SCAR markers. Food Control 36: 126–131.
- Vaughan DA, Balázs E, Heslop-Harrison JS. 2007. From crop domestication to superdomestication. Annals of Botany 100: 893–901.
- Zubor ÁA, Surányi G, Gyóri Z, Borbély G, Prokisch J. 2004. Molecular biological approach of the systematics of *Crocus sativus* L. and its allies. *Acta Horticulturae* 650: 85–93.