

**MOLECULAR DIVERSITY AND RELATIONSHIPS
OF SAFFRON AND WILD *CROCUS* SPECIES**

Thesis submitted for the degree of
Doctor of Philosophy
At the University of Leicester

By

Nouf Ahmad Fakieh Alsayied
Department of Biology,
University of Leicester, United Kingdom

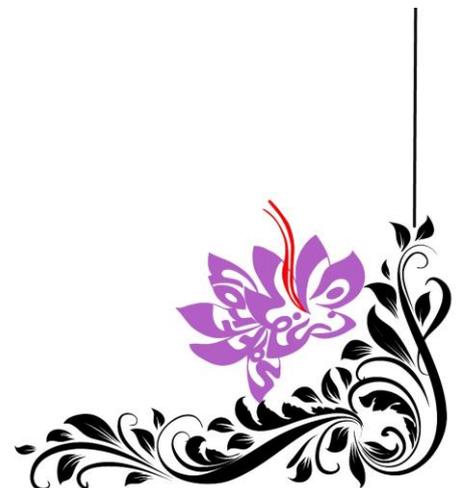
March, 2015

Abstract

MOLECULAR DIVERSITY AND RELATIONSHIPS OF SAFFRON AND WILD *CROCUS* SPECIES

Nouf A. Fakieh Alsayied

Crocus sativus L., saffron Crocus, is a sterile triploid ($2n=3x=24$) species of unknown relationship to other diploid and polyploid species in the genus *Crocus* (*Iridaceae*). The species have large genomes (typically 3000Mbp 1C), much repetitive DNA, and show high morphological diversity within and between species, with no clear phylogenetic patterns below the level of section *Crocus*, series *Crocus*. I aimed to examine molecular diversity in *C. sativus* and related species by analysis of sequences and IRAPs (Inter Retroelement Amplified Polymorphisms). Repetitive DNA sequences and genomic DNA from various species were used for *in situ* hybridization, with chromosome morphology, to infer relationships and ancestry of saffron. The IRAP analysis, involving 63 primer combinations and 4745 polymorphic bands, revealed no polymorphism within 17 saffron accessions obtained from across the world from Kashmir through Iran to Spain. In contrast, high levels of polymorphism were identified between accessions of six wild *Crocus* series *Crocus* species, with further variation between the species. Analysis of 123 sequences of the ATP-synthase gene and 107 TC₂₅ gene-SSR sequences from seven saffron accessions and eight wild species showed that the saffron accessions often carried three alleles, a result also found with clustering of published EST sequences. The analysis showed many alleles were shared by *Crocus* species and did enable a well-resolved phylogeny. Chromosome analysis grouped saffron chromosomes into 8 groups of 3, but one chromosome differed from the other two. It was concluded 1) Saffron crocus has minimal genotypic variation and the triploid hybrid species is most likely to have arisen only once; 2) Saffron is a allotriploid species, with the most likely ancestors being *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*.



Declaration

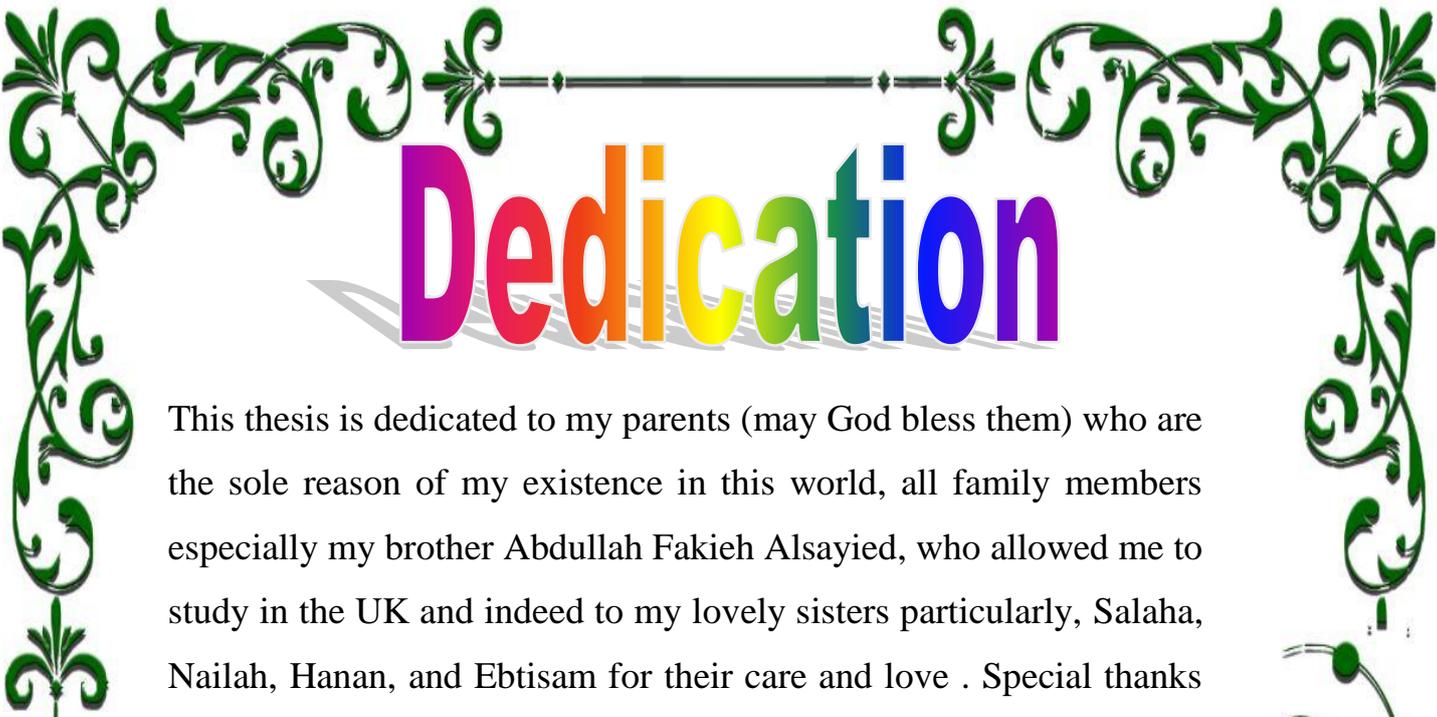
I hereby declare that no part of this thesis has been previously submitted to this or any other University as part of the requirements for a higher degree. The content of this thesis is result of my own experimentation and data analysis otherwise acknowledged in the text or by reference.

The work was conducted in the department of Biology, University of Leicester, during the period July 2009 to May 2014.

Signed

Nouf A. Fakieh Alsayied





Dedication

This thesis is dedicated to my parents (may God bless them) who are the sole reason of my existence in this world, all family members especially my brother Abdullah Fakieh Alsayied, who allowed me to study in the UK and indeed to my lovely sisters particularly, Salaha, Nailah, Hanan, and Ebtisam for their care and love . Special thanks for the closer sisters Fayzah and Fatmah who stood with me in all hard times and supported me both financially and morally and I will be faithful to them all my life.



Acknowledgements

Thanks to Almighty Allah for blessing me with the power and opportunity to be almost at the concluding point of my PhD study. My most sincere and infinite gratitude to **Prof. Pat Heslop-Harrison** (my PhD supervisor), for his excellent guidance, motivation and above all, his endless patience and never lasting support during this project. Without Pat's, I would have never solved many of my research problems and this study would have never ever ended. Today, I find it extremely difficult to express my feeling in words, but I want to thank him (Prof. Pat Heslop-Harrison), Dr. Trude Schwarzacher and Dr. Sinead Drea for the valuable suggestions and discussions to improve this work.

I thank Dr. John Bailey for his immense help in cytogenetics work. Special thanks to my best friends Dr. Niaz Ali for his help, support and guides to improve my lab skills and Emmanuele Ranieri from Italy who was part of my work on *Crocus* "saffron", and all members of lab 201 (those who graduated or still in the lab) who helped and supported me in during my study. All staff members of the Biology Department (those who are still in service, retired or passed away), friends from Saudi Arabia and UK.

I am thankful to all people who have extended their hands of help directly or indirectly throughout my life, and particularly those who helped me recently to complete my research and thesis.

Nevertheless, to King Abdullah and king Salman Bin Abdulaziz Al Saud, Kings of Saudi Arabia who set the trend and directed all Saudi Arabian universities to send their research scholars and faculty members for higher studies abroad.

I must thank Umm Al-Qura University, Makkah, KSA for the generous funding of my study, Department of Biology and particularly Head of the Department Dr. Hussein Abu-lreesh, all staff members in Saudi Cultural Bauru (UK) for the overall help and guidance during my study at the University OF Leicester, UK.

Finally, I want to thank the University of Leicester, UK for all the high quality research facilities and services, CrocusBank and Marcela Santaella-Tenorio (Biotechnology, IDR-UCLM, Albacete, Spain) for providing corms and DNA of several *Crocus* accessions and Mr. Saed Miri, the Selling Manager of Mashhad Union of Saffron (Edman Saffron Company. Pvt. Ltd. Iran).

(Note: The order of names mentioned does not mean more important or the level of impact to this work)

Signed
Nouf A. Fakieh Alsayied



Abbreviations

%	Percentage
°C	degree Celsius
AFLPs	Amplified fragment length polymorphisms
APG	Angiosperm phylogeny group
BCIP	5-bromo-4-chloro-3-indolyl-phosphate
BGV-CU	Bank of Plant Germplasm of Cuenca (Spain)
bp	Base pairs
BSA	Bovine serum albumin
cDNA	Complementary DNA
cm	Centimeter
COST	Co-operation in science and technology
CTAB	Cetyltrimethylammonium bromide
CV	Cultivar
DAPI	4',6-diamidino-2-phenylindole
dATP	Deoxyadenosinetriphosphate
dCTP	Deoxycytosinetriphosphate
dGTP	Deoxyguanosinetriphosphate
dH ₂ O	distilled water
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotidetriphosphates
DR	direct repeat
dTTP	Deoxythymidinetriphosphate
EDTA	Ethylenediamine tetra-acetic acid
ESTS	Expressed sequence tagged site
EtBr	Ethidium bromide
EtOH	Ethanol
EU	European union
FISH	Fluorescent in situ hybridization
FITC	fluorescein isothiocyanate
g	Gram
GISH	Genomic <i>in situ</i> hybridization
HCl	Hydrochloric acid
hr	Hour(s)
INT	2-(4-iodophenyl)-5-(4-nitrophenyl)-3-phenyltetrazolium chloride
IPTG	Isopropyl-β-Δ-thiogalactopyranoside ()
IRAP	Inter-retrotransposon amplified polymorphism
ITS	internal transcribed spacer
Kb	Kilo base
LINEs	long interspersed repetitive elements
LTRs	Long terminal repeats
M	Molar
M bp	Mega base pair
Mg	Milligram(s)
Min	minute(s)
ml	Millilitre(s)
mM	Millimolar
mm	Millimetre
NCBI	National Centre for Biotechnology Information
NJ	Neighbor Joining
NOR	Nucleolar organizer region
PCR	Polymerase chain reaction
PIC	polymorphism information content
PMC	pollen mother cells
PVP	Polyvinylpyrrolidone
RAPD	Random amplified polymorphic DNA
rDNA	Ribosomal DNA
REMAP	Retrotransposon-microsatellite amplified polymorphism
RFLP	Restriction fragment length polymorphism

RNA	Ribonucleic acid
RNAse	Ribonuclease
rpm	Rotations per minute
RT	Room temperature
s	Second(s)
SCARs	Sequence characterized amplified regions
SDS	Sodium dodecyl sulfate
SINE	short interspersed nuclear element
SNPs	Single nucleotide polymorphisms
sp, ssp	species, subspecies
SSC	Saline sodium citrate
SSRs	Simple sequence repeats or microsatellite
STRs	Short tandem repeats
STSMs	Short-Term Scientific Missions
TE	Tris-EDTA
TEs	Transposable elements
T _m , T _a	melting temperature, annealing temperature
U	Unit
v/v	Volume added to volume
VLPs	Virus-like particles
w/v	Weight added to volume
WGD	Whole genome duplications
μl	Microliter
μM	Micromolar

1	CHAPTER I: GENERAL INTRODUCTION	1
1.1	The family <i>Iridaceae</i> and genus <i>Crocus</i>	1
1.2	<i>Crocus</i> series <i>Crocus</i>	3
1.3	<i>Crocus sativus</i> (Saffron)	7
1.3.1	Morphology <i>C. sativus</i>	7
1.3.2	Origin and domestication of saffron	9
1.3.3	Genetic variation in saffron	10
1.3.4	Geographical distribution of <i>Crocus</i> species and cultivation of saffron	13
1.3.5	Quality and uses of saffron	19
1.3.5.1	Use of saffron in cooking	20
1.3.5.2	Medicinal uses of saffron.....	21
1.3.5.3	Miscellaneous uses of saffron.....	21
1.4	Taxonomy of the genus <i>Crocus</i>	22
1.5	Cytogenetic investigation of the genus <i>Crocus</i>	25
1.6	Ecology and evolution of plant mating systems	27
1.7	Genome analysis, diversity and evolution	29
1.8	Molecular DNA diversity: significance and exploitation of novel resources	30
1.8.1	RFLPs	31
1.8.2	RAPD markers.....	32
1.8.3	SCARs	32
1.8.4	AFLPs	32
1.8.5	Microsatellite DNA markers	33
1.8.6	ESTs and EST-SSR	33
1.8.7	IRAP and REMAP markers.....	33
1.8.8	DNA barcoding markers.....	35
1.8.9	RAD sequencing.....	35
1.8.10	SNPs	36
1.8.11	Whole genome sequencing.....	36
1.9	Polyploidy in flowering plants	36
1.10	Germplasm collection resources and its exploitation	38
1.11	Preserving genetic resources of <i>Crocus</i> and its allies.....	40
1.11.1	CROCUSBANK project.....	40
1.11.2	SAFFRONOMICS project.....	42
1.12	Aims of thesis	43
2	CHAPTER II: MATERIALS AND METHODS	44
2.1	Materials	44
2.1.1	Plant material.....	44
2.1.2	Germination of <i>Crocus</i> corms	44
2.1.3	Standard solutions and media used.....	47
2.2	Methods	49
2.2.1	Isolation of total genomic DNA	49
2.2.2	Agarose gel electrophoresis and purification of PCR products.....	51

2.2.3	DNA quantification	51
2.2.4	PCR markers and primer design	52
2.2.5	Polymerase Chain Reaction (PCR).....	52
2.2.6	Cloning of PCR products.....	53
2.2.6.1	Ligation reaction and transformation of competent E. coli cell	53
2.2.6.2	Screening for recombinant clones	54
2.2.6.3	Plasmid DNA purification, insert confirmation and storage of E. coli cells	54
2.2.7	Sequencing of PCR amplicons and sequence analysis.....	56
2.2.8	Probes used	56
2.2.9	Probe labelling	56
2.2.9.1	M13-PCR labelling.....	57
2.2.9.2	Random primers labelling.....	57
2.2.9.3	Testing of the incorporated labelled nucleotides in probes	57
2.2.10	Chromosome preparations	58
2.2.10.1	Collection and fixation of root tips.....	58
2.2.10.2	Metaphase chromosomes preparation.....	58
2.2.11	Fluorescent <i>in situ</i> hybridization	59
2.2.11.1	Pre -hybridization	59
2.2.11.2	Hybridization	60
2.2.11.3	Post-hybridization washes and detection.....	60
2.2.11.4	Mounting of slides, photography and image processing	61
2.2.12	Reprobing of slides	61
2.2.13	Isolation of repeated DNA sequences.....	62
2.2.13.1	Selection of plasmid clones for dot blot hybridization.....	62
2.2.13.2	Transfer of bacterial colonies onto charged nylon membrane.....	63
2.2.13.3	Membrane hybridization.....	63
2.2.13.4	Post-hybridization washes and detection.....	63
3	CHAPTER III: GENETIC VARIABILITY AND PHYLOGENY OF <i>CROCUS SATIVUS</i> L. (SAFFRON) BASED ON RETROELEMENT INSERTIONAL POLYMORPHISMS	65
3.1	Introduction.....	65
3.2	Material and methods	68
3.2.1	Plant material and genomic DNA extraction.....	68
3.2.2	IRAP markers	68
3.2.3	PCR amplification and gel electrophoresis	69
3.2.4	Genetic variability and phylogenetic analysis	69
3.3	Results	71
3.3.1	IRAP amplification and diversity within <i>Crocus</i> species.....	71
3.3.2	IRAP amplification and diversity within saffron accessions.....	72
3.3.3	Genome diversity and phylogenetic relationships among <i>Crocus</i> species.....	73
3.4	Discussion.....	85
4	CHAPTER IV: PHYLOGENETIC RELATIONSHIPS OF <i>CROCUS</i> SPECIES AND POLYPLOID NATURE OF SAFFRON INFERRED FROM EST-SSR, SNPS AND BARCODING GENES	92
4.1	Introduction.....	92

4.2	Material and methods	94
4.2.1	Plant materials and DNA extraction	94
4.2.2	PCR markers and primer design	94
4.2.3	PCR amplification and gel electrophoresis	95
4.2.4	Cloning and sequencing of PCR products	95
4.2.5	Sequence variability and phylogenetic analysis	95
4.3	Results	98
4.3.1	PCR markers analysis	98
4.3.2	Nucleotide sequence variation and phylogenetic tree of the ATPs region from saffron accession and <i>Crocus</i> species.....	107
4.3.3	Nucleotide sequence variation and phylogenetic tree of the TC ₂₅ region from saffron accession and <i>Crocus</i> species.....	113
4.3.4	Nucleotide sequence variation and phylogenetic tree of the barcoding genes 120	
4.3.4.1	<i>matK</i> 390F+1326R plastid gene sequence.....	120
4.3.4.2	<i>matK</i> XF+5R plastid gene sequence.....	120
4.3.4.3	<i>trnH</i> plastid gene sequence.....	120
4.3.4.4	<i>rbcL</i> plastid gene sequence.....	121
4.3.4.5	Phylogenetic analysis.....	121
4.4	Discussion.....	124
4.4.1	ATPs sequence variation	125
4.4.2	TC ₂₅ EST-SSR sequence variation	127
4.4.3	Variation in <i>matK</i> , <i>rbcL</i> and <i>trnH</i> chloroplast and mitochondrial (barcoding) genes 130	
4.4.4	Variation in <i>C. sativus</i>	132
5	CHAPTER V: RELATIONSHIPS OF <i>CROCUS</i> SPECIES BASED ON CYTOGENETIC INVESTIGATION AND ORGANIZATION OF REPETITIVE DNA SEQUENCES	135
5.1	Introduction.....	135
5.1.1	Cytogenetic structure of <i>Crocus sativus</i>	135
5.1.2	Repetitive DNA organization in <i>C. sativus</i>	137
5.2	Materials and Methods	138
5.2.1	DNA extraction, restriction enzyme digestion and gel electrophoresis	139
5.2.2	Isolation of repetitive DNA sequences from <i>C. sativus</i>	139
5.2.3	<i>In situ</i> hybridization.....	139
5.2.4	Probes used	139
5.3	Results	140
5.3.1	Karyotype of <i>C. sativus</i>	140
5.3.2	Meiotic chromosomes pairing in <i>C. sativus</i>	144
5.3.3	<i>In situ</i> hybridization and characterization of potential donor parents	149
5.3.3.1	<i>In situ</i> hybridization with genomic DNA from <i>C. cartwrightianus</i> and <i>C. pallasii</i> subsp. <i>pallasii</i>	149
5.3.3.2	<i>In situ</i> hybridization with genomic DNA from <i>C. thomasii</i> and <i>C. asumaniae</i>	150
5.3.3.3	<i>In situ</i> hybridization with genomic DNA from <i>C. hadriaticus</i> and <i>C. mathewii</i>	151
5.3.3.4	<i>In situ</i> hybridization with genomic DNA from <i>C. oreocreticus</i> and ‘ <i>C. sativus cartwrightianus</i> ’	152

5.4	Discussion.....	160
5.4.1	Ancestral species of <i>C. sativus</i>	160
5.4.2	Karyotype and allopolyploid nature of <i>C. sativus</i>	164
6	CHAPTER VI: GENERAL DISCUSSION.....	168
6.1	Genome studies and breeding in orphan crops “Cultigens”	168
6.2	Origin and genetic diversity in <i>Crocus sativus</i>	170
6.3	Molecular markers, genome diversity and evolution	172
7	REFERENCES	175
8	Appendix 4.....	198
9	Appendix 5.....	235



Courtesy of "Iran Saffron Book"

" We said: 'O Messenger of Allah! What is wrong with us that when we are with you our hearts are softened and we feel free of desire for this world, and we are of the people of the Hereafter. But when we depart from you and socialize with our families and our children, we do not recognize ourselves (i.e., we are changed persons)?' So the Messenger of Allah (s.a.w) said: 'If you were to be in that condition when you depart from me, the angels would have surely visited you in your houses. And if you did not sin, Allah would surely have brought anew creation that they may sin, so that then He may forgive them.'"He said: "I said: 'O Messenger of Allah! From what was the creation created?' He said: 'From water.' We said: 'Paradise, what is it constructed of?' He said,'Bricks of silver and bricks of gold. Its mortar is musk of a strong fragrance, and its pebbles are pearls and rubies, and its earth is saffron. Whoever enters it shall live and shall not suffer, and shall feel joy and shall not die, nor shall their clothes wear out, nor shall their youth come to an end.'

حَدَّثَنَا أَبُو كُرَيْبٍ، حَدَّثَنَا مُحَمَّدُ بْنُ فَضِيلٍ، عَنْ حَمْرَةَ الزَّيَّاتِ، عَنْ زِيَادِ الطَّائِبِيِّ، عَنْ أَبِي هُرَيْرَةَ، قَالَ قُلْنَا يَا رَسُولَ اللَّهِ مَا لَنَا إِذَا كُنَّا عِنْدَكَ رَفِقَتْ قُلُوبُنَا وَرَهَدْنَا فِي الدُّنْيَا وَكُنَّا مِنْ أَهْلِ الْآخِرَةِ فَإِذَا خَرَجْنَا مِنْ عِنْدِكَ فَاتَسْنَا أَهَالِيَنَا وَشَمَمْنَا أَوْلَادَنَا أَنْكَرْنَا أَنْفُسَنَا . فَقَالَ رَسُولُ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ " لَوْ أَنَّكُمْ تَكُونُونَ إِذَا خَرَجْتُمْ مِنْ عِنْدِي كُنْتُمْ عَلَى حَالِكُمْ ذَلِكَ لَزَارَتْكُمْ الْمَلَائِكَةُ فِي بُيُوتِكُمْ وَلَوْ لَمْ تُذْنِبُوا لَجَاءَ اللَّهُ بِخَلْقٍ جَدِيدٍ كَيْ يُذْنِبُوا فَيَغْفِرَ لَهُمْ " . قَالَ قُلْتُ يَا رَسُولَ اللَّهِ مِمَّ خُلِقَ الْخَلْقُ قَالَ " مِنْ الْمَاءِ " . قُلْنَا الْجَنَّةُ مَا بَنَاهَا قَالَ " لِبِنَّةٍ مِنْ فِصَّةٍ وَلِبِنَّةٍ مِنْ ذَهَبٍ وَمِلَاطُهَا الْمِسْكُ الْأَذْفَرُ وَحَصْبَاؤُهَا اللَّوْلُؤُ وَالْيَاقُوتُ وَتُرْبَتُهَا الرَّعْرَعَانُ مَنْ يَدْخُلُهَا يَنْعَمُ وَلَا يَبْئَسُ وَيُحْلَدُ وَلَا يَمُوتُ لَا تَبْلَى ثِيَابُهُمْ وَلَا يَفْنَى شَبَابُهُمْ.

<http://sunnah.com/urn/727120>

Grade : Da'if (Darussalam)

English reference : Vol. 4, Book 12, Hadith 2526

Arabic reference : Book 38, Hadith 2717

1 CHAPTER I: GENERAL INTRODUCTION

1.1 The family *Iridaceae* and genus *Crocus*

Iridaceae is one of the most species-rich and widely distributed families of herbaceous monocots (*Asparagales*). The family has a global distribution, and is represented by 65 to 75 genera and some 1800 species (Ali and Mathew, 2000; Goldblatt *et al.*, 2008). Many species of the family are highly valued as ornamental plants (Figure 1.1), and a few have medicinal uses. Although none of the *Iridaceae* is a significant food crop, saffron is cultivated agriculturally as a spice and dye (Mathew 1982; Petersen *et al.*, 2008; Harpke *et al.*, 2013), while many other species are prized horticultural specimens, many grown for the cut-flower industry. The family is characterized by rhizomes/corms, long unifacial leaves (both sides identical) oriented edgewise to the stem. Plants within the family have a wide diversity of flowers and can be recognized by their petaloid perianth, or corolla, with three tepals of the inner whorl and the three of the outer whorl usually alike in texture, shape and often in colour. *Iridaceae* is distinguished from other *Asparagalean* families by the three stamens in the androecium (Mathew, 1982, 1999; Goldblatt and Manning, 2008). The family's monophyletic origin is well defined. Earlier classifications have included the family within the *Liliales* or the Orchids, but modern treatments place it in the relatively new order *Asparagales*, now defined on the basis of DNA sequence analysis because of its high morphological diversity (Petersen *et al.*, 2008; Bremer *et al.*, 2009; Lovo *et al.*, 2012; Souza-Chies *et al.*, 2012).

The genus *Crocus* encompasses 88-100 small corm-bearing, perennial species, widely distributed in Central and Southern Europe, North Africa, and from Southwest Asia to western China, although the majority of taxa are restricted to Turkey and the Balkan Peninsula (Mathew 1982; Erol *et al.*, 2014 and Table 1.2). In Turkey, 35 *Crocus* species are reported as endemic (Mathew, 1984, 2000; Coskun *et al.*, 2012). Greece is the homeland of 40% of the world's wild *Crocus* diversity (Tsoktouridis *et al.*, 2009). Most species of the genus *Crocus* inhabit the Northern hemisphere, existing in the wild and are highly prized for ornamental purposes (Ørgaard *et al.*, 1995; Frello *et al.*, 2004; Petersen *et al.*, 2008; Fernandez *et al.*, 2011). A few varieties of *C. vernus*, *C. versicolor* and *C. aureus* are extensively used in gardening for their attractive flowers (Moraga *et al.*, 2010 and Figure 1.1). The importance of the genus is known to man for more than three

thousand years, primarily due to saffron's medicinal and nutritive importance (see section 1.2, 1.3 and Figure 1.3).

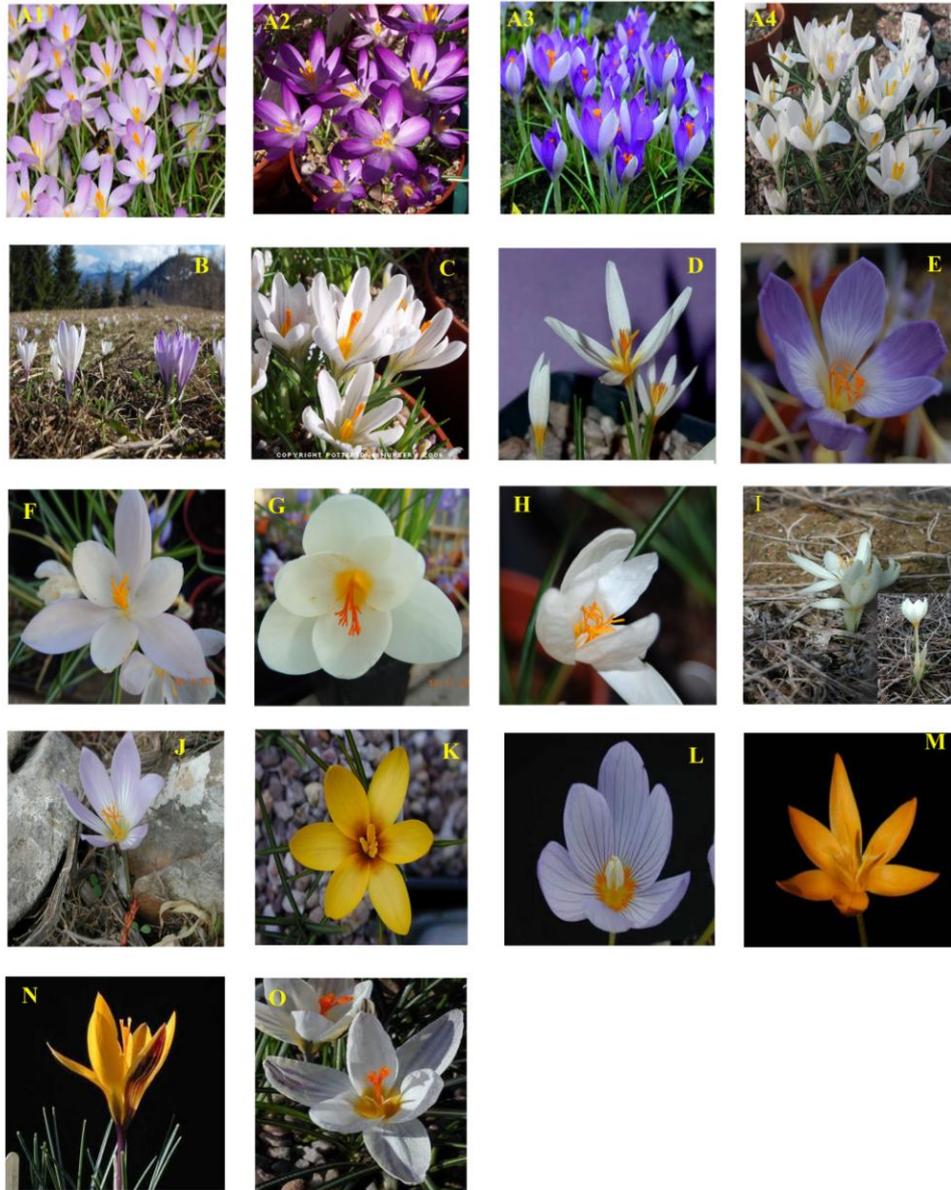


Figure 1.1: Members of the genus *Crocus* with beautiful flowers. A1). *C. tommasinianus* (Liliac beauty), A2) *C. tommasinianus* (barr purple), A3). *C. tommasinianus* (rubinetta), A4) *C. tommasinianus* cv. *albus*, B) *C. vernus*, C) *C. versicolor*, D) *C. veneris*, E) *C. speciosus*, F) *C. goulimyi* (leucanthus), G) *C. niveus*, H) *C. boryi*, I) *C. laevigatus*, J) *C. cancellatus*, K) *C. korolkowii*, L) *C. kotschyanus*, M) *C. flavus*, N) *C. angustifolius*, O) *C. biflorus biflorus*. (Source of A, B, C, D, I, J, K, L, M, N, and O are <http://www.alpinegardensociety.net/plants>).

Table 1.1: Taxonomic position of *Crocus*

Division	Spermatophyta
Sub-division	Angiospermae
Class	Monocotyledonae
Order	Asparagales
Family	Iridaceae
Subfamily	Crocoideae
Genus	<i>Crocus</i>
Species	<i>C. sativus</i>

1.2 *Crocus* series *Crocus*

Crocus series *Crocus* is one of the most well characterised and thoroughly studied series in the genus *Crocus* (Frello *et al.*, 2004; Petersen *et al.*, 2008; Harpke *et al.*, 2013). According to Mathew (1982, 1999) and there are several distinguishing features that differentiate members of the series. The corm tunics are finely fibrous and mostly reticulate; flowers in autumn, leaves are usually 5-30 but often numerous that appear with or shortly after flowers are formed; bracts flaccid, usually not closely sheathing the perianth tube; anthers are yellow in colour and style has branches 3, usually and often expanded at the apex; the seed coats are covered with a dense mat of papillae (reviewed in Caiola and Canini, 2010; Saxena, 2010). *Crocus* series *Crocus* includes 10 species all are diploids (Table 1.2, Figure 1.2), most with basic chromosome number of 8, while *C. sativus* is triploid (see result chapter V). Brief description of these species is given below while detail description of *C. sativus* is given in (section 1.3).

C. pallasii Gold. (1817), is one of the most variable species in the series, widespread from Balkans to Iran and from the Crimea to S. Jordan. At the moment at least four (Caiola and Canini, 2010) to five (Saxena, 2010) subspecies have been recognized (Table 1.2, Figure 1.2B). The corms are globose of about 10-25mm in diameter, with fibrous tunics that are finely reticulate and extended at the apex into a neck up to 2cm long. Flowers are fragrant, autumnal, 1-6, pale pinkish lilac to deep lilac-blue or purplish blue and usually slightly veined. Style is divided into three red branches, each branch 3-15 mm long (Saxena, 2010).

C. mathewii Kerndorff and Pasche (1994) has globose corms vary in size from 13-24mm in diameter and are flattened at the base. Tunics are made of fine fibers, which are parallel in the lower part, reticulate near the apex of corm, extended into a neck. Flowers

bracteolate, 1-3 in number, fragrant, autumnal with white or rarely pale lilac, often stained deep violet, throat violet and pubescent. Style divided into three orange to red branches, each branch 6–10 mm long, usually clearly exceeding, and rarely shorter than the anthers (see Caiola and Canini, 2010 and Figure 1.2G).

C. thomasii Tenore (1826), corms are 8-15mm in diameter, depressed globose and flattened at the base. Tunics fibrous, the fibers slender and finely reticulate, extended at the apex into a neck up to 1cm. Flowers bracteolate, fragrant, autumnal and generally but not always veined, violet towards the base, throat pale yellow and pubescent. Style is divided into three bright red branches each of 0.7–2 cm length (see Mathew, 1982 and Figure 1.2F).

C. cartwrightianus (Herb) Maw (1881), corms vary from 10-20mm in diameter. Further, the corms are depressed globose, with fibrous tunics, which are finely reticulated, extended into a neck. Flowers 1-5, bracteolate, autumnal, fragrant, pale to deep lilac-purple or white, strongly veined darker, sometimes stained darker at the base of the segments and on the tube, sometimes pure white with no veining (albinos are frequent in the species); throat white or lilac, pubescent. Style divided into three red branches, equaling or exceeding the anther (see Saxena, 2010 and Figure 1.2A).

C. oreocreticus B. L. Burtt (1949), corms are ovoid and are of approximately 10-15mm diameter, depressed globose with tunics made of finely reticulated fibers. Flowers 1-2, autumnal, mid-lilac to purple with darker veining (Figure 1.2E). Style divided into 3 red thickened branches, and about equaling the tips of the anthers, arising at a point at or just above the throat of the flower (Mathew, 1982; Caiola and Canini, 2010 and Figure 1.2E).

C. asumaniae B. Mathew and T. Bay Top (1976), corms are ovoid and have approximately 15-20mm diameter. Tunics of *C. asumaniae* are fibrous and the fibers finely reticulated, extended at the apex of the corm into a neck. Flowers 1-3, bracteolate, autumnal, white, occasionally white dark veins near the base of the segments, rarely very pale lilac; throat whitish or pale yellow (Mathew, 1982). Style divided into reddish-orange clavate branches, each considerably exceeding the anthers (Saxena, 2010 and Figure 1.2D).

C. hadriaticus Herbert (1845), corms are 10-15mm in diameters, which are depressed globose and flattened at the base, with tunics fibrous finely reticulated extended at the apex of the corm into a short neck. Flowers fragrant, autumnal and often stained externally brownish, yellowish or violet at the base of the segments, throat yellow or

rarely white, pubescent. Style divided into 3 slender branches, each branch slightly shorter than or exceeding the anthers. (Figure 1.2C)

C. moabiticus Bornmüller (1912), corms are approximately of 20-30mm diameter, subglobose, flattened at the base with fibrous tunics parallel at the base and weakly reticulate at the apex extending into a neck. Flowers 1-6, bracteole, fragrant, veined purple to varying degrees on all six segments on a white ground colour, sometimes so heavily as to appear purple, sometimes stained darker at the base of the segments and on the tube; throat white or purple, pubescent. Style divided into 3 deep red clavate branches, arising at a point well below the base of the anthers in the throat of the flower. (Figure 1.2H)

C. naqabensis Al-Eisawi and Kisawi (2001) has features similar to *C. pallasii* but its corms have reduced tunics that do not form a neck. Moreover, flowers have a globous throat. *C. naqabensis* also share homology to the endemic *C. moabiticus* and *C. cartwrightianus* from which it differs for the style branches which are not more than half as long as the perianth segments (Caiola and Canaini, 2010).

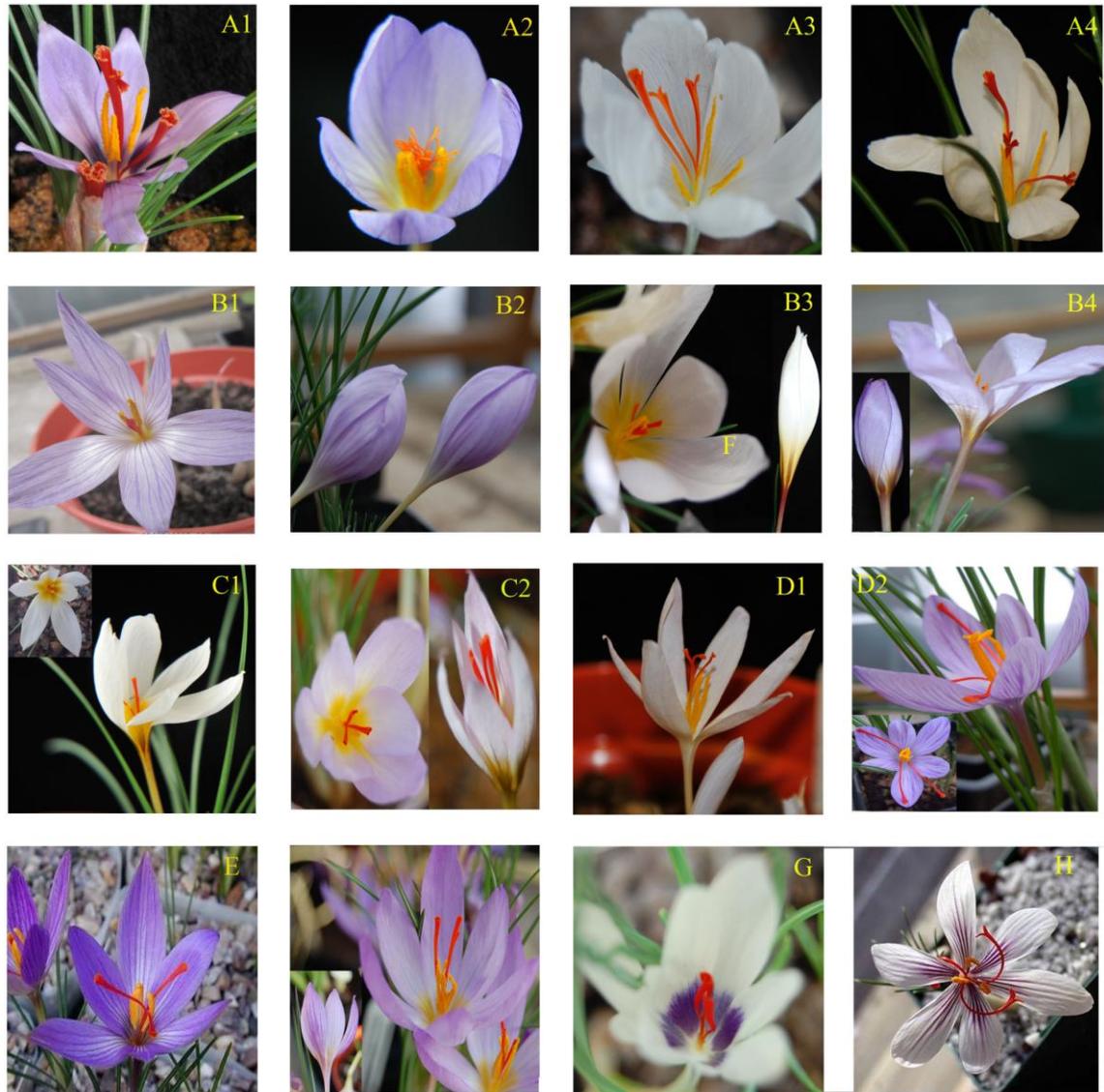


Figure 1.2: Flower morphology of the members of *Crocus* series *Crocus* (potential ancestors of saffron). The accession includes: A1) *C. cartwrightianus* (CEH613), A2) *C. cartwrightianus* (Dixexport), A3) *C. cartwrightianus* cv. *albus*, A4), *C. sativus* *cartwrightianus*, B1) *C. pallasii* subsp. *pallasii*, B2), *C. pallasii* subsp. *haussknechtii*, B3) *C. pallasii* subsp. *dispathaceus*, B4), *C. pallasii* subsp. *turcicus*, C1) *C. hadriaticus* cv. *Hadriaticus*, C2) *C. hadriaticus* cv. *lilacinus*, D1) *C. asumaniae* (white flower), D2) *C. asumaniae*, E) *C. oreoreticus*, F) *C. thomasii*, G) *C. mathweii*, H) *C. moabiticus*. Images of all accession except E and H are from this study. (E, H are modified from www.alpinegardensociety.net/plants).

1.3 *Crocus sativus* (Saffron)

All along the human history, *Crocus sativus* L. (*Iridaceae*) has been cultivated for obtaining saffron, the most expensive spice on earth (Kafi, 2006; Petersen *et al.*, 2008; Fernandez *et al.*, 2011). Its long history fluctuates between myths, legendary tales and traditions but man knew saffron from the Minoan period in Crete some 3000 years ago (Caiola and Canini, 2010; Fernandez *et al.*, 2011; Harpke *et al.* 2013). The term “Saffron” is probably originated from Arabic “zafran”, which means “yellow” (Arsalan *et al.*, 2007) and today the name saffron applies indistinctly to *C. sativus* as well as the spice obtained from its dried stigmas (Caiola *et al.*, 2004; Aytekin and Acikgoz, 2008). The scientific name of saffron dates back to Linnaeus who in 1762 named it *Crocus staivus* var. *officinalis* (see Caiola and Canini, 2010).

Saffron (stigmas) are harvested manually and the spice is mainly utilized for cooking (colour, flavour and unique aroma), dye or medicine (D’Agostino *et al.*, 2007; Fernandez *et al.*, 2007; Gresta *et al.*, 2009). It is the only plant whose product is sold by the gram; on average 1 kilogram (kg) of saffron is selling for over 2000 US dollars. However, the very high price of saffron is due to the great direct labour required for its cultivation, harvesting and handling (Aytekin and Acikgoz, 2008). Production of 1 kg of saffron requires 150,000 to 200,000 flowers (c. 500,000 stigmas), and the low productivity of 6 kg saffron per hectare limits saffron availability worldwide (Fernandez, 2007; Sharaf-Eldin *et al.*, 2008).

1.3.1 Morphology *C. sativus*

Saffron is a small autumn flowering perennial geophyte perennial, growing from an underground tuberous stem (corm) and may reach up to a height of 10-25cm. The plant is unknown in the wild and its corms are about 5cm in diameter with a globular to sub-oval shape, lightly flattened at the bottom that looks like an onion bulb in appearance (Molina *et al.*, 2004). Their sizes vary considerably between accessions (personal observations), and are well adapted to sustain harsh seasonal climates. Corms are covered with tunic (expanded leaf bases) of parallel fibres and extended at the apex of the corm into neck (Erol *et al.*, 2008). Under favorable growing conditions, the apical meristem of corm or one of the buds close to the apex and extends into an aerial flowering shoot usually bearing 10-15 foliage leaves. The leaves are needle like and may reach up to 40cm having green colour and a central white stripe, which is due to the lack of chlorophyll in some

cells (European Saffron White Book, 2006). Each shoot produces 1-4 flowers that are normally fragrant, purple in colour and usually pale lilac or mauve with darker colour veins (Figure 1.3). Flowers have an underground ovary, a style of 9-10cm long dividing into three intense red colour stigmas (25- 30 mm). Each flower has three yellow coloured anthers and six petals (perianth), three on the inner side and three on the outer side joined on the long pipe that comes out of the upper part of the ovary (Deo, 2003). Normally, each corm also produces numerous roots of white colour and variable lengths (usually 5-10cm).

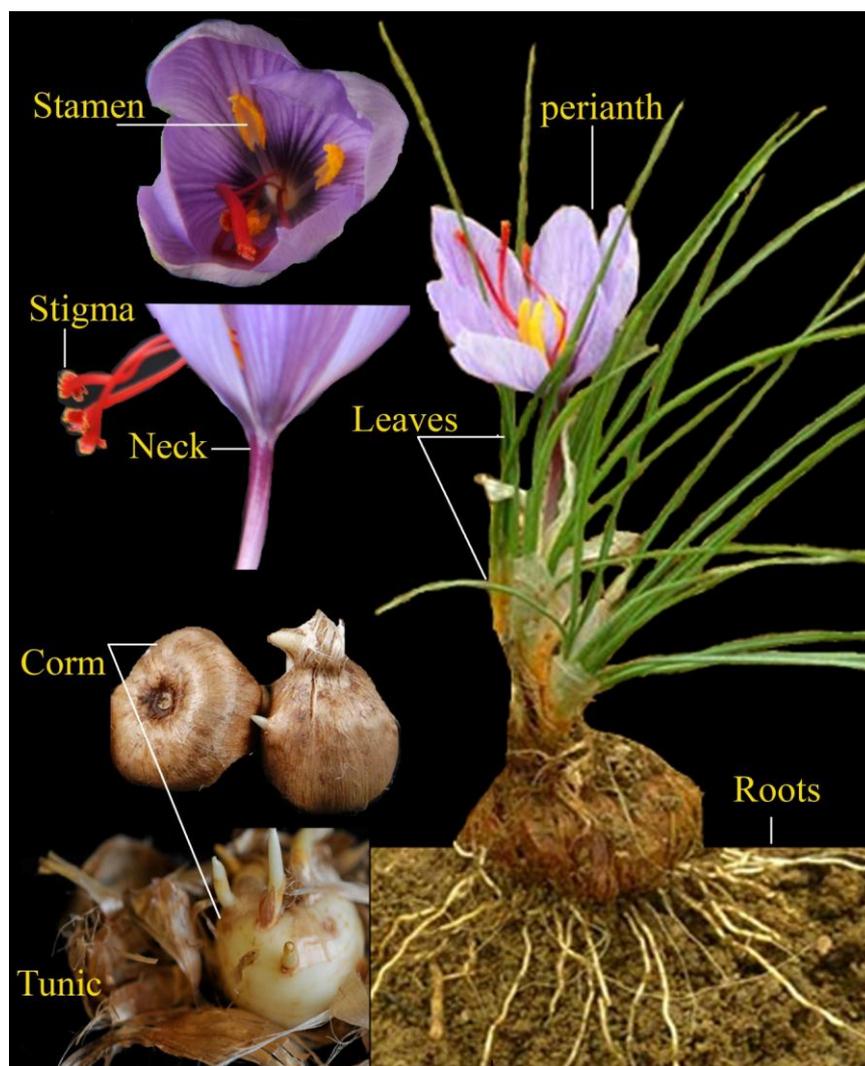


Figure 1.3: Plant body of *C. sativus* (saffron). Morphology of corm, tunic, roots, leaves, and floral parts are indicated.

1.3.2 Origin and domestication of saffron

To date, very little is known about the origin and domestication of saffron, and work on breeding of saffron for better quality or higher yield are non-existent, primarily due to the triploid nature ($2n = 3x = 24$). With basic chromosome number of $x = 8$, saffron never sets viable seeds and propagates exclusively by vegetative means (Kafi, 2006; Petersen *et al.*, 2008; Caiola and Canini, 2010). Despite of the extremely high market value and demands, saffron as a crop is facing the danger of extinction in many parts of the world (Carmona *et al.*, 2006; Fernandez, 2007; Fernandez *et al.*, 2011).

Saffron is native to the Mediterranean environments, characterised by cool winters and warm dry summers and archaeological records indicate that saffron was cultivated and used as a spice and/or medicinal plant in the Mediterranean basin since late Bronze Age (Negbi, 1999; Fernandez, 2007; Harpke *et al.*, 2013 and Figure 1.6). However the sites where the first saffron plants appeared differ according to the opinion of various authors (see Caiola and Canini, 2010). For most, saffron probably originated in Iran, Asia Minor or somewhere in Greece and later became widespread in India, China, the Mediterranean basin and Eastern Europe, and the domestication may have occurred during the Greek-Minoan civilization between 3,000 and 1,600 B.C (Negbi 1999; Caiola *et al.*, 2004; Fernandez, 2004; Ghorbani, 2007; D'Agostino 2007). The Romans introduced saffron into Great Britain, while the Arabs brought it to Spain (Fernandez, 2007).

Despite of the intensive studies for several decades, even today the evolutionary processes and the species involved in the origin of saffron are yet to be identified (Maw, 1886; Mathew, 1982; Fernandez *et al.*, 2011; Erol *et al.*, 2014). Today's domesticated saffron grown around the world could be one clone that was probably selected by man for its triploid vigour and long stigmas and has been maintained since then (Mathew 1982; Fernandez, 2004, 2007; Harpke *et al.*, 2013 and section 1.4). The wild source of domesticated *C. sativus* was probably *C. cartwrightianus*, and originated by fertilization of a diploid unreduced egg cell by a haploid sperm cell or a haploid egg by two haploid sperms cells (Caiola *et al.*, 2004; Caiola, 2005). Both *C. sativus* and *C. cartwrightianus* are morphologically very similar and even today, *C. cartwrightianus* is used as a wild source of saffron (Mathew, 1982 and result Chapters IV, V and Fig 1.2 A1, A2).

Classical studies based on morphology have revealed *C. cartwrightianus* to be the closest relative of *C. sativus* (Maw, 1886; Mathew, 1982). Molecular studies based on flow cytometry (Brandizzi and Caiola, 1998), RAPDs (Caiola *et al.*, 2004), ISSR (Sik *et*

al., 2008) and AFLPs (Zubor *et al.*, 2004) analysis supported *C. cartwrightianus* and *C. thomasi* as the most closely related species of *C. sativus*. Furthermore, based on IRAP markers, *C. almehensis* and *C. michelosnii* were shown to be the possible ancestral species of *C. sativus* (Alavi-Kia *et al.*, 2008). While the findings of Rubio-Moraga *et al.*, (2009) revealed *C. cartwrightianus* cv. *albus* to be more related to *C. sativus* than to *C. cartwrightianus* and it may be albino saffron. By and large, different species of the *Crocus* series *Crocus* have been suggested as possible ancestors for *C. sativus*. Studies using repetitive DNA (Frello *et al.*, 2004), chloroplast, ribosomal and nuclear single copy genes sequence focused on phylogeny and could not find a clear evidence for the site of domestication or ancestral species of saffron (Petersen *et al.*, 2008; Harpke *et al.*, 2013; Erol *et al.*, 2014).

It is very interesting that few studies (for example Caiola *et al.*, 2004) found *C. pallasii* and *C. asumaniae* as more distantly related species of saffron, however my findings suggest that *C. pallasii* could be one of the putative parents (see result Chapters III, IV, V). Plant domestication although, had set the road to human civilization, but despite of its immense role, our current understanding about domestication is very limited (Ross-Ibarra *et al.*, 2007). Although, sequencing large size genomes is still a problem (Doležel *et al.*, 2012), many sequencing projects are underway (<http://genomesonline.org>), so we hope very soon the whole genomic sequences of diploids and polyploid species of the series *Crocus* will be available; this may reveal some important details of the *Crocus* genomics and genes that had the most important role in the origin and domestication of saffron. These sequencing projects may enable us to answer some of the questions that we are currently unable to address.

1.3.3 Genetic variation in saffron

Genetic diversity is crucial in all breeding programmes as improvements in general could have been difficult without the existence of variation within these genes (Villalobos and Engelmann, 1995; Caiola *et al.*, 2004). Desirable genes, which have either been selected by man or nature itself, are dispersed within both domesticated and wild plant populations (Vaughan *et al.*, 2007). Thus the ancestral species remain the primary sources of genetic diversity (Akhunov *et al.*, 2010; Heslop-Harrison and Schwarzacher, 2012). These sources (wild and cultivated) offer the possibility of gene transfer and exploitation, and traits of interest may be reintroduced in the form of chromosomal segments through direct crossing

or through special manipulation techniques in crop improvement programmes (Vaughan *et al.*, 2007; Heslop-Harrison *et al.*, 2010).

Saffron is a male-sterile triploid plant that multiplies exclusively by vegetative means without any recombination except for mutation (Nehvi *et al.* 2007; Rubio-Moraga *et al.*, 2009; Fluch *et al.*, 2010). Genetic diversity in *C. sativus* is very limited or largely unknown (Fernandez, 2007; Harpke *et al.*, 2013). This lack of diversity may be partly attributed to its sterile nature, lacking pollination and homologous recombination but over the last 30-40 years land surface assigned to saffron cultivation has reduced significantly, thus exerting extra pressure and eroding its genetic base (Agayev *et al.*, 2009; Gresta *et al.*, 2008; Fernandez *et al.*, 2011).

Origin of Saffron is uncertain, but it grows in a wide geographical area since time immemorial and Saffron breeders believe the existence of limited genetic variability within the crop. For example, there exist morphological diversity in size of corm, flower, petal shape and colour intensity, number of style branches, stamens and stigmas in saffron collected from different areas (Alvarez-Orti *et al.*, 2004; Fernandez, 2004; Nehvi *et al.*, 2007; Agayev *et al.*, 2009). Still these variations occur at a low frequency and retaining the same numbers of chromosome (Nehvi *et al.* 2007; Agayev *et al.*, 2006; Ghaffari and Bagheri, 2009 and Figure 1.4). Further, there are several reports that describe diversity in the chemical constituents of the stigma too (Ordoudi *et al.*, 2004; Fernandez, 2007; Anastasaki *et al.*, 2009; Maggi *et al.*, 2011).

A number of studies carried out over recent years have suggested clonal origin of world's saffron. These studies hypothesised an ancient spontaneous hybridization event in nature resulted in a unique triploid clone of *C. sativus* or saffron (Mathew, 1977; Caiola *et al.*, 2004; Rubio-Moraga *et al.*, 2009; Fluch *et al.*, 2010). Due to its sterility it solely rely multiplication by corms and still continues to be propagated vegetatively (Dhar *et al.*, 1988, Piqueras *et al.*, 1999) and thus, saffron growing around the world today, may be just one clone (Jacobsen and Ørgaard, 2004 and section 1.3.1 above). Molecular studies that probed to understand the clonal origin of *C. sativus* and applied RAPDs, SSRs, ISSR, AFLPs, IRAPs, ESTs and chloroplast DNA markers, also assumed that there is just one saffron cultivar grown worldwide (Caiola *et al.*, 2004; Alavi-Kia *et al.*, 2008; Rubio-Moraga *et al.*, 2009; Fluch *et al.*, 2010; Izadpanah *et al.*, 2014). It seems very likely, that saffron has undergone artificial selection in the past, a practice that offers advantages in maintaining its genetic characteristics but causes reduction in genetic diversity (Agayev *et al.*, 2009; Heslop-Harrison, 2012). In addition, phenotypic differences stated above could

be attributed to differences in climate and cultivation practices (Caiola *et al.*, 2004; Rubio-Moraga *et al.*, 2009; Fluch *et al.*, 2010). While variation in saffron quality is mainly due to the methodology followed in processing the stigmas or may arise from adulterants (such as safflower) added to it, and is independent of the species origin (Rios *et al.*, 1996; Ordoudi *et al.*, 2004; Maggi *et al.*, 2011; Torelli *et al.*, 2014).

Several recent studies have shown the existence of genetic diversity on a limited scale (see Sik *et al.*, 2008; Nemati *et al.*, 2012). It is difficult to address, how or when saffron originated and naturalized but my PCR and cytogenetics results also suggest the existence of limited genetic variability within saffron of different geographical origin and will be discussed in the result chapters below.



Figure 1.4: Image showing minor morphological variation in *C. sativus* flowers collected from different geographical regions. Accession includes, A) *C. sativus* (Spain), B) *C. sativus* (Kashmir), C) *C. sativus* var. *Cashmirianus* (Dix Export), D) *C. sativus* (Pottertons), E) *C. sativus* (Dix Export).

1.3.4 Geographical distribution of *Crocus* species and cultivation of saffron

Members of the genus are adapted to a wide range of ecological habitats, however majority of them being endemic to the Mediterranean and Europe (see Table 1.2). Saffron is perhaps the most economically important and widely distributed species of the genus (Deo, 2003; Siracusa *et al.*, 2013 and Figure 1.6). With a broad ecological amplitude showing adaptability to a wide range of different soil types, temperatures, altitudes and day length over *i.e.* from South-Western Europe, throughout central Europe to Turkey and South-western parts of Asia, and as Far-East as Western China (Kafi *et al.*, 2006; Agayev *et al.*, 2006; Erol *et al.*, 2014).

Saffron possesses unique characteristics that are absent in other agricultural plants, and enable its cultivation under adverse climatic conditions (Deo, 2003; Molina *et al.* 2005; European Saffron White Book, 2006). In the autumn, when cultivated plants complete their growth phase and go through a dormant state, saffron corms begin to blossom as the mean air temperature falls below 15-17°C (Plessner *et al.*, 1989; Álvarez-Ortí *et al.*, 2003; Molina *et al.*, 2004). In the spring, when most plants usually begin a new growth cycle, the leaves of saffron turn yellow and dry up and enter dormancy. The plant survives the summer heat as an underground corm in its dormant stage but can withstand substantial frosts and a temperature as low as -10°C (Deo, 2003; Kafi, 2006; de Juan *et al.*, 2009). However, in control experimental conditions the dormancy may be released and optimal flower formation achieved 6 weeks earlier than in the open if corms are held at 25°C longer than 55 days followed by forcing at 17°C. Flowering could be further accelerated (up to 7 days) by curing the corms for 20 days at 30°C prior to 25°C storage (Molina *et al.*, 2004).

Saffron requires less water and is well adapted to arid or semiarid lands. Fertilizers and chemical inputs for saffron are very minor, and the overall cultivation of saffron has barely changed over the last 3,000 years (Fernandez *et al.*, 2007; Gresta *et al.* 2008). Each year, saffron passes through a distinct activity and a dormant phase (Figure 1.5) which can be sub-divided into six developmental stages *i.e.* C1 to C6. In C1 corms appear as latent buds attached to the surface of older corms, C2 is defined by floral stem sprouting and enlargement characterizes C3 corms while drying of the leaves defines C4. During this phase corms grow independently because the mother corm senesces. Corms at stage C5 maintain sprouting and growth of daughter corms and at stage C6 corms are senescent, and daughter corms advance to stage C4 and become independent (Álvarez-Ortí *et al.*,

2003). About approximately 40 days after cultivation, saffron flowers in the autumn, and depending upon the weather conditions the flowering period may last up to 15 days (Deo, 2003; European Saffron White Book, 2006; Fernandez, 2007).

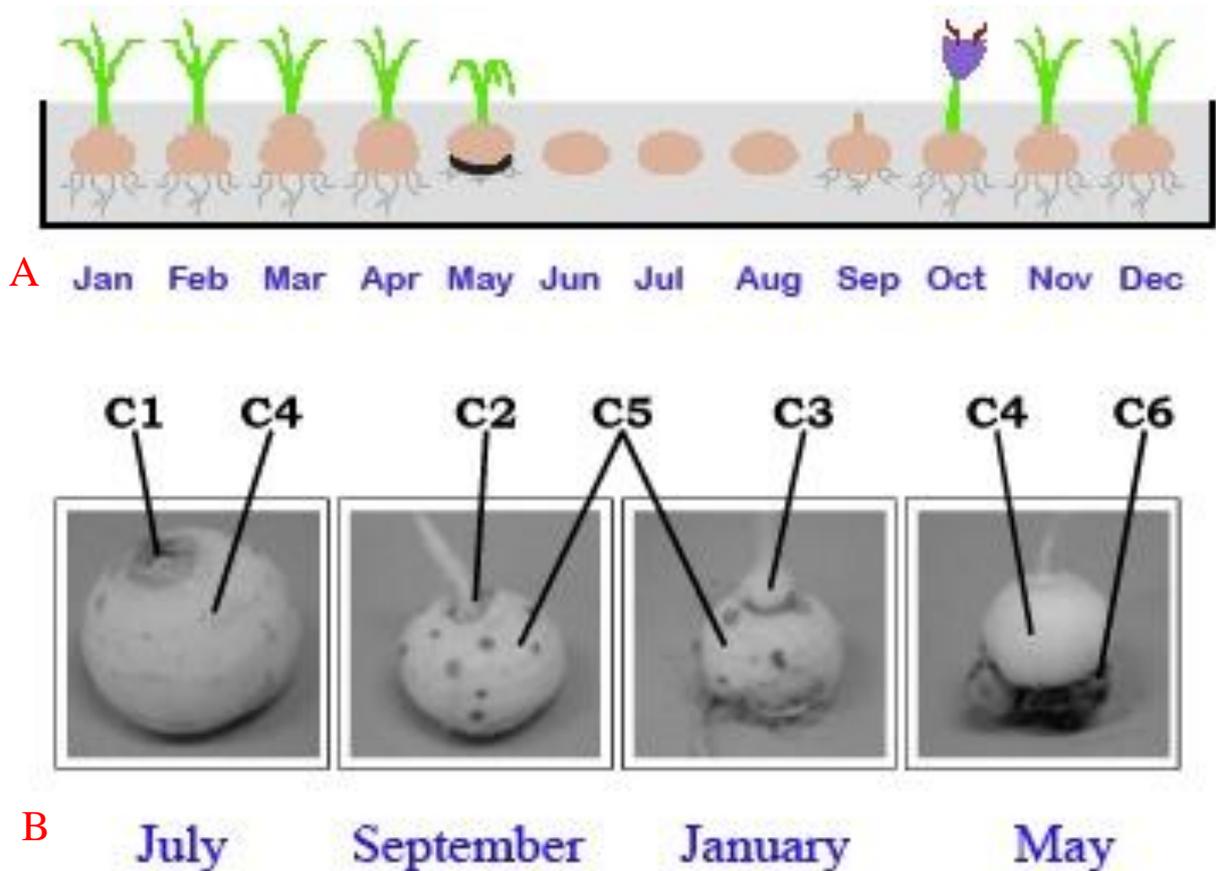


Figure 1.5: A) Diagrammatic representation of the production cycle of *C. sativus*, B) Developmental stage of saffron corm (modified from Álvarez-Ortí et al., 2003). C1: latent buds on the surface of mother corms, C2: base of the stem from sprouting to the end of flowering, C3: Daughter-enlarging corm, attached to the surface of mother corms, C4: daughter corm when become independent, dormancy, C5: corm supporting growth of daughter corms on its surface, C6: corms senescent.

Today Iran, Spain, Morocco, India, Greece, Italy, and China are important producers of saffron (Gresta *et al.*, 2008; Agayev *et al.*, 2009 and Figure 1.6). Lately, farmers in Afghanistan profitably started saffron cultivation and replaced the illegal production of Opium Poppy (Mollafilabi and Aslami, 2010). Small scale production of saffron is taking place in France, Switzerland, Turkey, Azerbaijan, Japan, Australia, Israel, Pakistan, Egypt and United Arab Emirates (Fernández, 2004; Schmidt *et al.* 2007; Agayev *et al.* 2009). However, for a country like Iran, saffron is one of the main sources of its income and is currently the largest producer of saffron (Gresta *et al.*, 2008; Agayev *et al.* 2009); in 2005 Iran produced 230 tons of saffron, which was 93.7% of the world's total saffron production, followed by Greece 5.7 tons, Morocco and then Kashmir producing 2.3 tons each (Carmona *et al.*, 2006; Fernandez *et al.*, 2007; Ghorbani, 2007).

Despite the desirable characteristics and high price, over the last few decades, area for saffron cultivation has reduced significantly in Europe (Gresta *et al.*, 2008, 2009; Kumar *et al.*, 2008; de Juan *et al.*, 2009). For example Spain, which was once the largest producer of the world's saffron, today in Spain the land surface assigned for saffron cultivation has dramatically reduced from 13,000 ha in 1914 to under 6000 ha in 1972, to 116 ha in 2006 (Fernandez, 2004; [http://: CrocusBank.org/](http://CrocusBank.org/)). Only 20 years ago Spain and Iran were producing almost the same amount of saffron, that is 35 to 40 tons (Fernandez, 2006; Kumar *et al.*, 2008), while in 2004, the whole of Europe produced only 4% of the international saffron (Fernandez, 2004; Gresta *et al.*, 2008). Arduous hand labour during cultivation and harvesting, lack of modern technology, urbanization and increasing labour costs may explain a few important reasons for saffron reduced cultivation (de Juan *et al.* 2003; Fernandez, 2007; Gresta *et al.*, 2008; Agayev *et al.* 2009).

SAFFRON ORIGIN AND PRODUCTION COUNTRIES

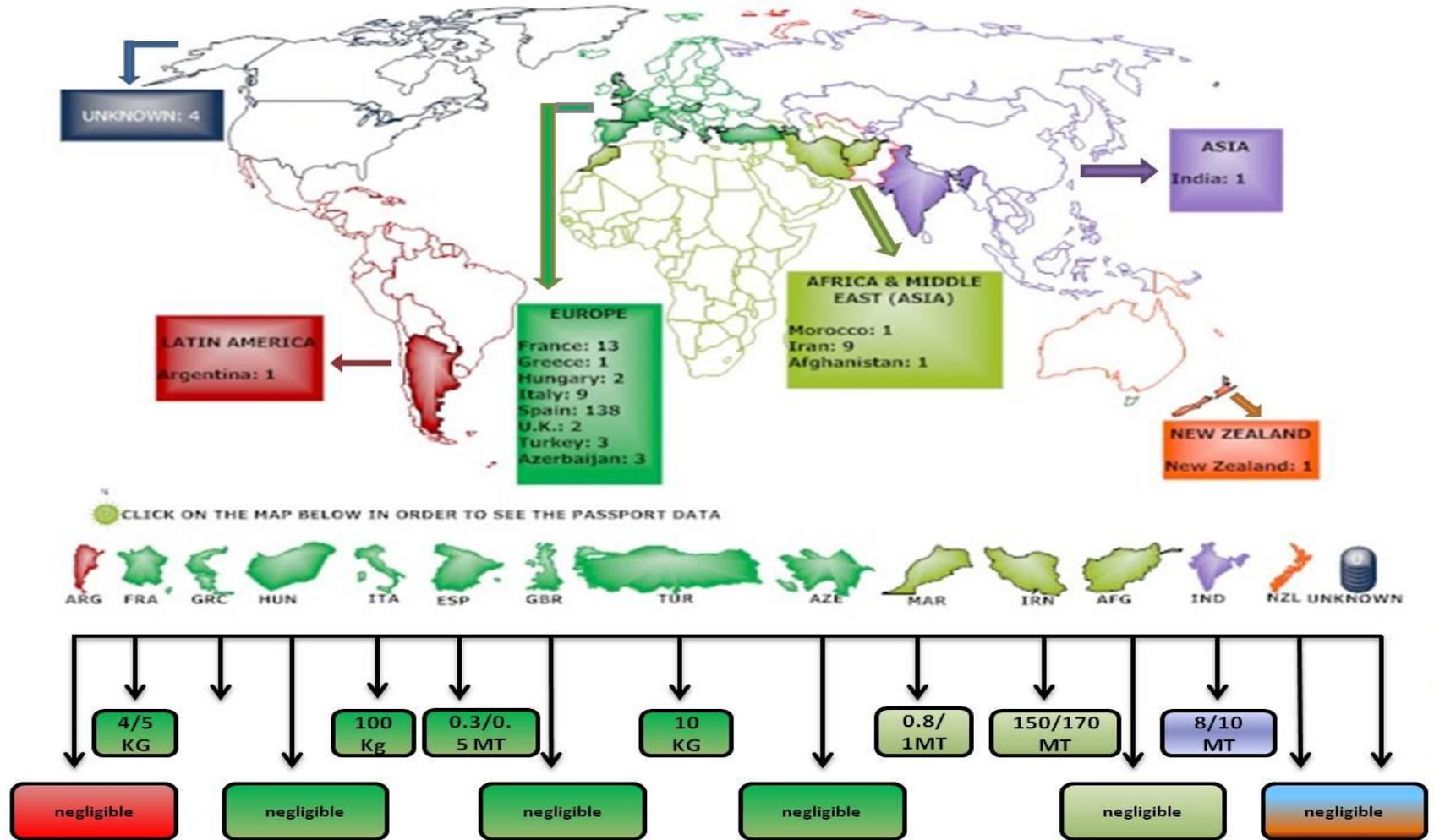


Figure 1.6: Origin and distribution of saffron in different countries of the world (modified from <http://www.crocusbank.org/>).

Table 1.2: Chromosome number, flowering habit, distribution and distinctive features of the *Crocus* species.

Serial	Species name	Chromosome number	Flowering time	Distribution	Native Climate	Distinctive features
1	<i>C. pallasii</i> subsp. <i>pallasii</i>	2n = 14	Autumn	S Serbia; SE Bulgaria; Crimea; Aegean Islands; Lebanon; Israel; Turkey	Mediterranean	Flowers in shades of lilac, often slightly veined darker. Corm tunic finely netted fibres
	<i>C. pallasii</i> subsp. <i>turcicus</i>	2n = 12	Autumn	SE Turkey; Lebanon; Syria	Cold winter, long, hot, dry summer	Flower shades of lilac, petals often narrow. Corm tunic finely netted fibres with long fibrous neck
	<i>C. pallasii</i> subsp. <i>hausknechtii</i>	2n = 16	Autumn	W Iran; NE Iraq; S Jordan	Cold winter, hot, dry summer	Flower shades of lilac, often veined darker. Corm tunic finely netted fibres with long fibrous neck
	<i>C. pallasii</i> subsp. <i>dispathaceus</i>	2n = 14	Autumn	S Turkey; N Syria	Dry Mediterranean	Flower deep reddish purple, petals very narrow. Corm tunic finely netted fibres
2	<i>C. mathewii</i>	2n = 16	Autumn	S Turkey (Restricted to few locations in Taurus mountains)	Mediterranean, long dry summer	Whitish flowers with dark purple zone in the throat. Corm tunic parallel fibres
3	<i>C. thomasii</i>	2n = 16	Autumn	S Italy; Serbia; Croatia	Mediterranean	Flower shades of lilac, pale yellow throat. Three red style branches up to half as long as the petals. Corm tunic thin fibres, finely netted
4	<i>C. cartwrightianus</i>	2n = 16	Autumn	Greece	Mediterranean	Three long bright red style branches. Flowers stay open at night. Finely fibrous corm tunic, reticulated
5	<i>C. cartwrightianus</i> cv. <i>albus</i>	2n = 16	Autumn	Greece, Turkey	Mediterranean	White flower with slightly veined, with intensive golden yellow anthers. Long style branching into three red coloured stigmas. Flowers are usually smaller than in <i>C. cartwrightianus</i> . Corm tunic finely netted fibres.
6	<i>C. sativus</i> <i>cartwrightianus</i> *	2n = 16	Autumn	Not known	Not known	Very similar to <i>C. cartwrightianus</i> cv. <i>albus</i>
7	<i>C. oreocreticus</i>	2n = 16	Autumn	C & E Crete	Dry Mediterranean	Lilac/purple flower with silvery or buff exterior to three outer petals. Style divides at or just above the throat of the flower into three red branches. Corm tunic finely netted fibres
8	<i>C. asumaniae</i>	2n = 26	Autumn	S Turkey	Cool winter, warm dry summer. In rain shadow of mountains to the south	Long style branches dividing at anther level. Fibrous corm tunic, fibres parallel towards base
9	<i>C. hadriaticus</i>	2n = 16	Autumn	W and S Greece	Mediterranean	White or pale lilac flower with yellow throat. Corm tunic fine netted fibres
10	<i>C. moabiticus</i>	2n = 14	Autumn	Jordan, Israel	Mediterranean	Whitish flower with lilac/purple throat. Stigma is divided into two branches. Corm tunic fine netted fibres
11	<i>C. naqabensis</i>	2n = 14	Autumn	Jordan, Israel		Corms have reduced tunics, no neck with globous throat. Style branches are not more than half as long as the perianth segments
12	<i>C. sativus</i>	2n = 24	Autumn	Not known as a wild plant	Not known (but grows in Mediterranean)	Large purple/lilac flower with three very long styles branching into three bright red stigmas. Corm tunic finely netted fibres

Table 1.2: continued

Serial	Species name	Chromosome number	Flowering time	Distribution	Native Climate	Distinctive features
13	<i>C. vernus</i>	2n=16	Spring	Austria; Czechoslovakia; Hungary; Italy; Balkan States; Poland; Romania; Ukraine	Cold winters with snow	Small variable, flower shades of purple, sometimes white or striped. Corm tunic fine fibres
14	<i>C. tommasinianus</i>	2n=16	Early spring	Croatia; Serbia; Bosnia; Montenegro; NW Bulgaria; S Hungary	Cold winter with snow, warm summer with regular rainfall	Flower pale lilac to deep purple. Corm tunic fine fibres, mostly parallel
15	<i>C. versicolor</i>	2n= 26	Spring	SE France; adjacent NW Italy	Mediterranean	Flower white, lilac, and purple, often with external stripes. Corm tunic membranous ageing to parallel fibres
16	<i>C. niveus</i>	2n= 28	Autumn	S Greece	Mediterranean	Large flower, usually white, occasionally pale lilac or bicoloured. Corm tunic finely netted fibres
17	<i>C. goulimy</i>	2n= 12	Autumn	S Greece (Mani peninsula, Peloponnese)	Mediterranean	Long tubed lilac flowers. Coriaceous corm tunic
18	<i>C. kotschyanus</i>	2n= 8,10	Autumn	C and S Turkey; NW Syria; Lebanon	Cold winter. Cool montane summer	Lilac flower with deep yellow blotches at the base of each petal. Corm tunic thin, membranous
19	<i>C. korolkowii</i>	2n= 20	Spring	Afghanistan; N Pakistan; Tajikistan; Uzbekistan	Cold winter with snow, dry summer	Glossy yellow flowers variously marked dark brown. Corm tunic membranous ageing to many parallel fibres
20	<i>C. flavus</i>	2n=8	Spring	Serbia; Greece; Bulgaria; Romania; NW Turkey	Mediterranean with continental influences	Bright yellow flowers, sometimes with brownish exterior markings. Corm tunic membranous with fibrous point at tip
21	<i>C. speciosus</i>	2n=6, 8, 10	Autumn	Crimea; Caucasus; N Iran; C & N Turkey	Cold winter, warm summer but with regular rainfall	Flower lilac blue, veined darker, many branched style. Corm tunic coriaceous splitting into rings at base
22	<i>C. laevigatus</i>	2n= 26,	Autumn to early spring	Greece (including Cyclades islands); Crete	Mediterranean	Flowers lilac or white usually with a dark vein or veins. Style much divided. Corm tunic coriaceous, older layers splitting into long pointed tipped teeth
23	<i>C. boryi</i>	2n= 24,30	Autumn	W and S Greece; Crete	Mediterranean	Large creamy white flower, prominent many branched orange style. Papery corm tunic
24	<i>C. veneris</i>	2n= 16	Autumn	Cyprus	Mediterranean	Flower small, starry, and white with some purple veining on outer petals. Corm tunic membranous with parallel fibres
25	<i>C. cancellatus</i>	2n=8,10, 12,16	Autumn	S Turkey; Lebanon; N Israel; Jordan	Mediterranean	Finely netted corm tunic, Much divided style which exceeds the anthers
26	<i>C. biflorus</i>	2n=8, 10,12,18,20,22	Spring	Italy; Sicily; Rhodes; NW Turkey	Mediterranean	Long leaves at flowering, striped flowers with a yellow throat. Sometimes with some lilac/violet staining on outer petals

* is a garden name, the accession was purchased under this unrecognized name and has similarities with *C. cartwrightianus* cv. *albus* (see results Chapter III, IV, V).

Sources: Brighton *et al.*, 1973; Mathew, 1999; Al-Eisawi, 2001; Agayev *et al.*, 2009; Saxena, 2010; Harpke *et al.*, 2013; Erol *et al.*, 2014.

1.3.5 Quality and uses of saffron

Saffron contains in excess of 150 volatile and aroma-yielding compounds, and the value of saffron is attributed to these characteristic phytochemical compounds (Kumar *et al.*, 2008; Srivastava *et al.*, 2010; Makri *et al.*, 2013). The chemistry of saffron has been thoroughly investigated and considerable insights have been gained about the constituents of saffron (Rios *et al.*, 1996; Winterhalter and Straubinger, 2000). The unique components of saffron are crocins, picrocrocin and safranal (Figure 1.7). Crocin is responsible for the colour of saffron, whereas picrocrocin is responsible for its bitter and exquisite taste while safranal is the main essential oil component responsible for odor and aroma (Rios *et al.*, 1996; de Juan *et al.*, 2009; Hosseinzadeh and Nassiri-Asl, 2012).

With its high value, adulteration of saffron with other plant products and dyes is a major problem in world trade and fraudulent mislabelling of the origin is also frequent (Torelli *et al.*, 2014); the Saffron trade is regulated by several quality standards, which take into account several parameters (ISO, 2003). The quality of saffron is dependent on its colouring power, odour and taste. Saffron is dry, glossy and greasy to the touch when freshly dried, turning dull and brittle with age. It is easily bleached if not stored in the dark, and also stores better under conditions of low temperature and low relative humidity (Winterhalter and Straubinger, 2000; Kumar *et al.*, 2008). The best quality saffron has high safranal content (Deo, 2003; de Juan *et al.*, 2009). During drying and processing saffron may lose some of the morphological, anatomical or chemical characteristics (Raina *et al.*, 1996; Ordoudi *et al.*, 2004). But owing to the incredible utilization and very high demand of saffron from different sectors (food, medicine, dye, and flavouring industries), saffron is sometimes mixed with other plants or plant parts intentionally (Fernandez, 2004). However, absorption spectrum analysis and molecular techniques using unique DNA sequences can be used to identify the high quality saffron if mixed up with adulterants (Ma *et al.* 2001; Agayev *et al.*, 2009).

It is very difficult to enumerate all the uses of saffron, as there is barely any sector of life where saffron has got no utilization. However, a few points summarizing the significant uses of saffron are given in brief below.

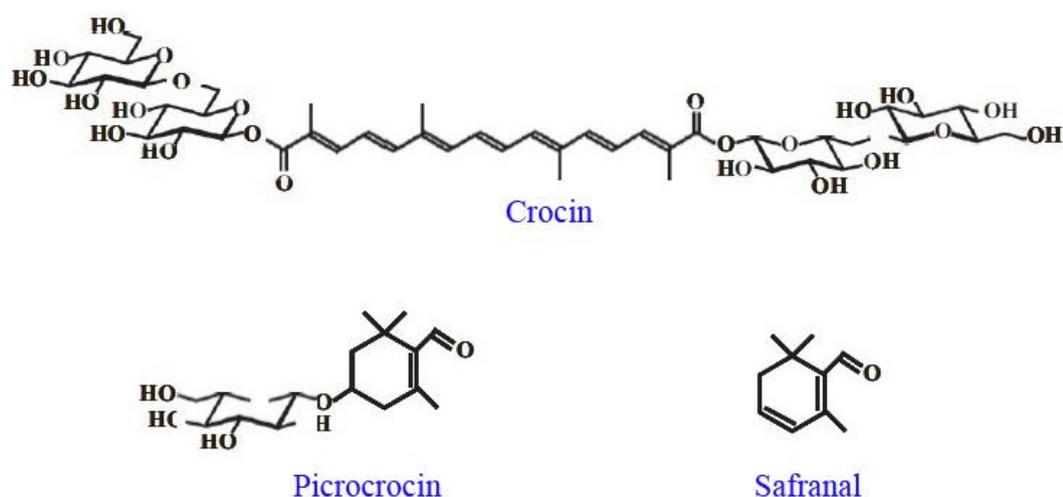


Figure 1.7: Chemical structure of crocin, picrocrocin and safranal (modified from Srivastava *et al.*, 2010)

1.3.5.1 Use of saffron in cooking

Saffron has a long history in European cuisine and there are no comparisons as far as the value of its spice is concerned (Agayev *et al.*, 2009). Its unique aroma and taste has no parallel and this property was recognized in ancient times, and has remained a valuable spice ever since (de Juan *et al.*, 2009; Hosseinzadeh and Nassiri-Asl, 2012). Saffron is used in traditional fish and seafood dishes (*Risotto alla Milanese* in Italy, *Bouillabaisse* in France or *Paella Valenciana* in Spain) as well as in *Gugelhupf* the traditional German saffron cake (reviewed in Winterhalter and Straubinger, 2000). Similarly in Iran, India, Pakistan, Middle East saffron and several other countries, it is used for seasoning, as a dye in confectionery, sweets, puddings, ice creams and in making *Biryane*, a very special type of rice cooked on important occasions (Kafi, 2006; Kumar *et al.*, 2008; Agayev *et al.*, 2009; Caiola and Canini, 2010).

1.3.5.2 Medicinal uses of saffron

As a curative plant, saffron is mentioned in the oldest available traditional medicines (Rios *et al.* 1996; Abdullaev *et al.*, 2004; Giaccio, 2004; de Juan *et al.*, 2009). Along with other uses, Hippocrates and Dioscorides mentioned saffron for treating ophthalmic disorders (see Makri *et al.*, 2013). Avicenna (980-1037 AD), one of the most influential philosophers and physicians in Islamic history, wrote a monograph “Al-Qanun Fe-Tib” or “Canon of Medicine” that describes more than 760 drugs, where saffron is described such as remedy for many disorders such as antidepressant, hypnotic, anti-inflammatory, hepatoprotective, bronchodilatory, aphrodisiac, inducer of labour, emmenagogue etc. (reviewed in Hosseinzadeh and Nassiri-Asl, 2012). Descriptions of saffron are available in the pharmacopoeias of many countries. Low doses of saffron work as a stimulant but high doses of saffron are reported to be toxic and abortifacient (Winterhalter and Straubinger, 2000). Saffron has been used in traditional medicines for insomnia, depression, bronchospasm, cardiovascular diseases, gastrointestinal disorders, menstrual pain, menopausal problems, as analgesic, anti poisonous, aphrodisiac, carminative, diaphoretic, diuretic, febrifuge, stimulant, sedative, retina-degeneration, immunomodulation and even against different kinds of tumours and cancers (Premkumar *et al.*, 2003; Abdullaev *et al.*, 2004; Giaccio, 2004; Fernandez, 2007; de Juan *et al.*, 2009; Poma *et al.*, 2012; Makri *et al.*, 2013; Siracusa *et al.*, 2013)

1.3.5.3 Miscellaneous uses of saffron

Saffron is mostly used as a spice, food colorant or in medicines. Nevertheless, to some extent it is used in textile, perfume making and flavoured tobacco industry (Fernandez, 2007; Ordoudi *et al.*, 2009; Poma *et al.*, 2012). Similarly on a small scale saffron is used in food processing units in sausage, butter, cheese, alcoholic, and non-alcoholic beverages (Kumar *et al.* 2008; Agayev *et al.*, 2009; Caiola and Canini, 2010).

1.4 Taxonomy of the genus *Crocus*

Being a genus of medicinally rich and attractive plants, the genus has always fascinated taxonomists (Figure 1.1). However, taxonomy of the genus has been extremely difficult, mainly due to the lack of absolute boundaries and unavoidable arbitrariness among different species along the wide range of habitats and heterogeneity of morphological traits (Caiola *et al.*, 2004; Rubio-Moraga *et al.*, 2009; Caiola and Canini, 2010). Based on multigene plastid DNA analysis, the genus *Crocus* was assigned to the Croceae tribe of the subfamily Crocoideae, and the three genera *Crocus*, *Romulea* Maratti, *Syringodea* D comprise the sub tribe Romuleinae (Goldblatt *et al.*, 2006).

Over the last 200 years a number of different classifications have been proposed for the genus. Most early attempts at *Crocus* taxonomy focused on morphological and anatomical characteristics or chromosome number (Maw, 1886; Mather, 1932; Karasawa, 1935; Mathew, 1977; Mathew, 1982; Rudall and Mathew, 1990; Goldblatt and Takei, 1997; Özhatay, 2002; Coşkun *et al.*, 2010). Haworth (1800) was among the first to classify the genus *Crocus* into two sections making presence/absence of hairs in the throat of the flower as the basis for his classification, later on Sabine (1829) proposed classification of the genus based on presence/absence of prophyll (basal spathe) and corm tunic. Similarly, Herbert's (1847) and Maw's (1886) classification of the genus relied on morphological and geographical features of the taxa, along with the presence/absence of basal spathe and flowering period (spring or autumn).

Based on Maw's (1886) classification, Mathew (1982) classified the genus *Crocus* into 80 species. Among these 80 species, 6 were identified in 1700, 54 in 1800 and 20 species were acknowledged in the 1900 (reviewed in Grilli Caiola and Canini, 2010). Further, Mathew (1982) divided the genus *Crocus* into two subgenera, the monotypic subgenus *Crociris* consisting only of *C. banaticus* and the subgenus *Crocus*. The subgenus *Crocus* was further divided into two sections and 15 series based on the existence of the prophyll, division of style, corm tunic and flowering time. One of the most interesting features of this classification was the introduction of subspecies concept for many species, an aspect that was missing in the earlier systems (see Mathew, 1982 and Figure 1.9).

Crocus is a highly complex genus with a wide range of variation in chromosome numbers and genome sizes (*i.e.* 11,000 Mbp 1C in *Crocus vernus*, $2n= 8$) (Frello *et al.* 2004; Candan *et al.* 2009). Among different species of the genus, chromosome number has been reported similar (Table 1.2) while intra-specific variation occurs in a number of

species and has been described in detail in the *C. vernus* aggregate (Brighton *et al.*, 1973; Caiola and Canini, 2010). Therefore, classification based on chromosome number alone might not be very helpful in understanding the relationships of the different species within the genus. Several attempts using molecular approaches have been made to underpin the phylogenetic relationship among *Crocus* species by analysing series and the whole genus (Frello and Heslop-Harrison, 2000a; Caiola *et al.*, 2004; Frello *et al.*, 2004; Petersen *et al.*, 2008; Harpke *et al.*, 2013). Petersen *et al.* (2008) classified the genus *Crocus* using nucleotide sequence data from three protein-coding (*ndhF*, *accD*, *rpoC1*) and two non-coding (*trnH-psbA*, *rpl36-rps8*) regions of plastid. The results obtained were contradictory to that of Mathew (1982) at subgenera and sections level. Interestingly the *C. banaticus* (subgenus *Crociris*) that is highly different morphologically was embedded within subgenus *Crocus* rather than keeping it as a sister group. The grouping of taxa into series is not entirely but considerably matched (Figure 1.9). Eight out of the fifteen series recognized by Mathew (1982), are confirmed as monophyletic in origin, and only one taxon altered the monophyletic origin of another two series. Species-rich series *Reticulati* and series *Biflori* that comprise almost one third of all species are clearly non-monophyletic. Today almost 100 species of *Crocus* have been recognized (Harpke *et al.*, 2013) and a variety of molecular analysis, including small and large scale genomic sequencing projects are underway (see also result chapters). Conclusions from these will be very helpful not only in the classification but also in understanding the affinities of different species within the genus (see also result Chapters III, IV and V).

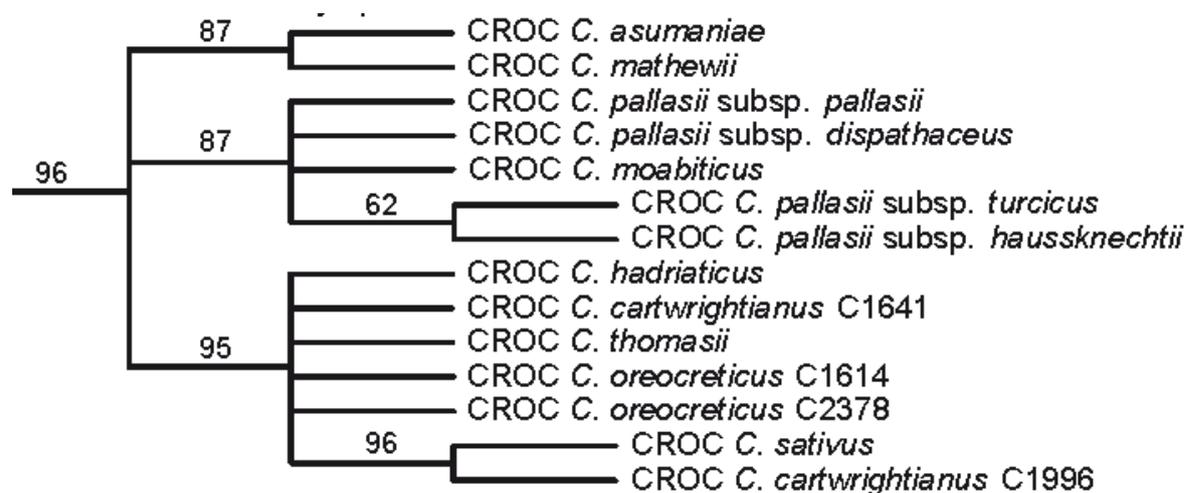


Figure 1.8: Phylogeny tree of *Crocus* series *Crocus* based on five plastid regions shows *C. sativus* cluster with *C. cartwrightianus* with 96 % nodal support (Source Petersen *et al.*, 2008).

- Subgenus *Crocus*
- Section *Crocus*
- Series *Verni* (VERN)
C. vernus, *C. tommasinianus*, *C. etruscus*,
C. kosaninii, *C. baytopiorum*
- Series *Scardici* (SCAR)
C. scardicus, *C. pelistericus*
- Series *Versicolores* (VERS)
C. versicolor, *C. malyi*, *C. imperati*, *C. minimus*,
C. corsicus, *C. cambessedesii*
- Series *Longiflori* (LONG)
C. longiflorus, *C. nudiflorus*, *C. serotinus*,
C. niveus, *C. goulimyi*, *C. ligusticus*** (syn.
C. medius hort, non Balbis) (Mariotti, 1988)
- Series *Kotschyani* (KOTS)
C. kotschyanus, *C. scharojanii*, *C. vallicola*,
C. autranii, *C. karduchorum*, *C. gilanicus*,
C. ochroleucus
- Series *Crocus* (CROC)
C. sativus, *C. pallasii*, *C. thomasi*, *C. cartwrightianus*,
C. moabiticus, *C. oreocreticus*, *C. asumaniae*,
C. hadriaticus, *C. mathewii** (Kerndorff &
Pasche, 1994), *C. naqabensis** (Al-Eisawi, 2001)
- Section *Nudiscapus*
- Series *Reticulati* (RETI)
C. reticulatus, *C. veluchensis*, *C. cvjicii*,
C. dalmaticus, *C. sieberi*, *C. robertianus*,
C. cancellatus, *C. hermoneus*, *C. abantensis*,
C. angustifolius, *C. ancyrensis*, *C. gargaricus*,
C. sieheanus, *C. rujanensis** (Randjelović & al.,
1990)
- Series *Biflori* (BIFL)
C. biflorus, *C. chrysanthus*, *C. almehensis*,
C. danfordiae, *C. pestalozzae*, *C. aeriis*,
C. cyprius, *C. hartmannianus*, *C. adanensis*, *C.*
leichlinii, *C. caspius*, *C. kerndorffiorum** (Pasche,
1993), *C. wattiorum** (Mathew, 1995; 2000), *C.*
*nerimaniae** (Yüzbasıoğlu & Varol, 2004)
- Series *Orientales* (ORIE)
C. korolkowii, *C. michelsonii*, *C. alatavicus*
- Series *Flavi* (FLAV)
C. flavus, *C. antalyensis*, *C. olivieri*, *C. candidus*,
C. vitellinus, *C. graveolens*, *C. hyemalis*,
*C. paschei** (Kerndorff, 1993)
- Series *Aleppici* (ALEP)
C. aleppicus, *C. veneris*, *C. boulosii*
- Series *Carpetani* (CARP)
C. carpetanus, *C. nevadensis*
- Series *Intertexti* (INTE)
C. fleischeri
- Series *Speciosi* (SPEC)
C. speciosus, *C. pulchellus*
- Series *Laevigati* (LAEV)
C. laevigatus, *C. boryi*, *C. tournefortii*
- Subgenus *Crociris* (CROCI)
C. banaticus
- Incertae cedis*
C. boissieri

Figure 1.9: Classification of *Crocus* species based on morphology following Mathew (1982) including seven species described (*) and one nomen novum (**). Yellow highlights the species used in the current study. (Source Petersen *et al.*, 2008).

1.5 Cytogenetic investigation of the genus *Crocus*

Over the last two decades, major advances have been made in the field of molecular biology, which essentially became possible due to better insight into the structural organization of the DNA and chromosomes (Heslop-Harrison, 1991, 2000b; Pires and Hertweck, 2008). Physical organization of DNA sequences and locating them on chromosomal regions is crucial for understanding genomic organization and evolution in plants (Kubis *et al.*, 2003; Contento *et al.*, 2005). The science concerned with genetic implications of chromosome structure and behaviour is referred to as cytogenetics. The discipline has deep roots in understanding of DNA sequences and the molecular structure of the chromosome and chromatin (Schwarzacher *et al.*, 1992; Schmidt and Heslop-Harrison, 1998; Gill and Friebe, 1999; Heslop-Harrison, 2010). Today, cytogenetics is an integral part of genome mapping projects as knowledge about karyotype and linkage groups comes from cytogenetics, and DNA sequence information is one part of the puzzle in such studies (Gill and Friebe, 1999; Heslop-Harrison and Schwarzacher, 2011).

A number of techniques such as C-banding, differential Giemsa staining, recognition of gene-rich and gene-poor regions as well early and late replicating regions have been developed over time and are still employed in cytogenetic research (see Schwarzacher, 2003b). However, the development of fluorescent DNA: DNA *in situ* hybridization techniques that exploit DNA sequence composition and molecular data to the structure and organization of chromosomes has redefined the field of molecular cytogenetics (Gall and Pardue, 1969; John *et al.*, 1969; Schwarzacher and Heslop-Harrison, 2000; Frello *et al.*, 2004). To date, fluorescent *in situ* hybridization (FISH) is the most powerful cytogenetic technique for chromosomal mapping and genomic analysis, which is available to scientists. It allows rapid identification of chromosomes and structural rearrangements in chromosomes such as deletions, duplications, translocations and inversions (Castilho *et al.*, 1995; Gill *et al.*, 2011; Patel *et al.*, 2011).

Karyologically the genus *Crocus* is very heterogeneous and the chromosome number varies from $2n = 6$ to $2n = 70$ (Mather, 1932; Karasawa, 1935; Brighton *et al.*, 1973; Goldblatt and Takei, 1997). At the moment parental species of saffron are unknown, although its cytogenetic structure and polyploid nature is well understood (see result Chapter VI). The first cytological studies of the genus *Crocus* were carried out 1930s, when Sugiura (1931) and Mather (1932) reported chromosome numbers for *C. sativus* as

$2n=24$ and $2n=15$ or $2n=14$, respectively. Karasawa (1935) analysed *C. sativus* and related species and reported $2n=24$ for *C. sativus*. Furthermore, Pathak (1940), Feinbrun (1958), Brighton *et al.*, (1973), Brighton (1977), Mathew (1977), Ghaffari (1986), Goldblatt and Takei, (1997), Ebrahimzadeh *et al.* (1998), Frello and Heslop-Harrison, (2000b), Frello *et al.*, (2004), Schneider *et al.*, (2013) have studied the chromosome number and karyotype evolution of *C. sativus* and its allied species. Their results confirmed *C. sativus* as triploid with $2n=3x=24$, $x=8$ (Figure 1.10), but whether *C. sativus* is auto or allotriploid is yet to be determined.

Since cytogenetic analysis provide direct insight into the plant genomes and karyotype evolution they play a critical role in reconstructing phylogenies (Schmidt and Heslop-Harrison, 1998; Schwarzacher, 2003a; Markova and Vyskot, 2009). Therefore, a variety of molecular cytogenetic approaches were applied to answer the fundamental question of saffron origin, evolution and diversity (see results Chapters IV, V, VI). I reason to believe that understanding the diversity and parental species of saffron will open an exciting era of saffron molecular breeding with more precise and predictable outcomes. Thus the current project will be helpful to renew optimism among saffron growers and to re-establish saffron as an important crop around the world particularly, where its cultivation is abandoned.

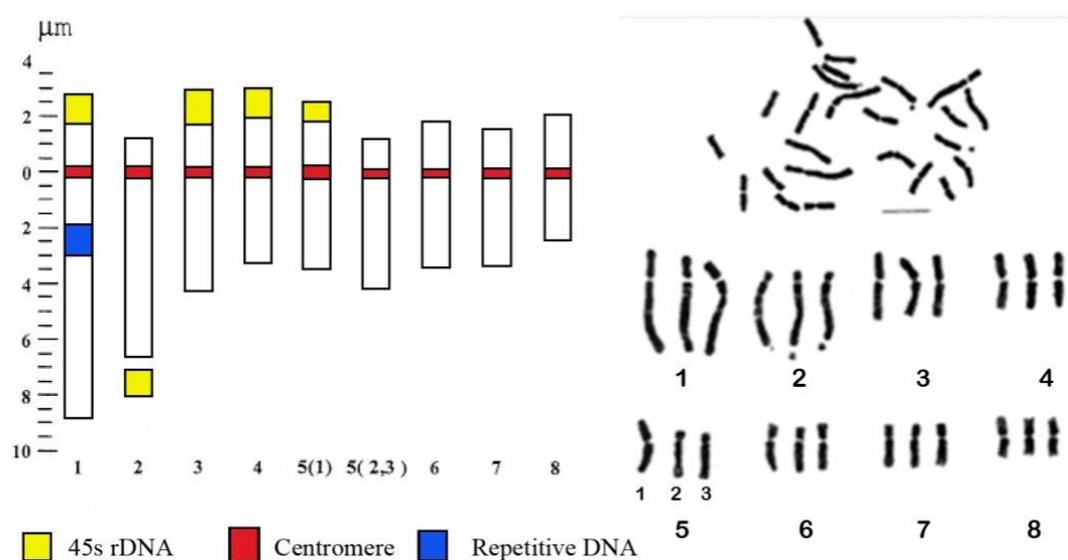


Figure 1.10: An ideogram of *C. sativus* Karyotype ($2n=3x=24$, $x=8$), showing relative size and position of heterochromatin. 1-8 indicating to morphologically similar chromosome (triplet 8), chromosomes 1,2,5(2,3): Subacrocentric, chromosomes 3,4,5(1),8: Metacentric, chromosomes 6,7: Submetacentric, (modified from Agayev, 2002) for further details (see section 5.3.1, and Table 5.1).

1.6 Ecology and evolution of plant mating systems

Angiosperms are intriguing models for studying the evolutionary biology of reproductive systems. In spite, most flowering plants are sessile in nature and have co-sexual (hermaphrodite) flowers still; the reproduction of vast majority is unique in a sense that they are heavily dependent on pollen vectors for their mating (Goodwillie *et al.*, 2004; Hiscock and Allen, 2008). This reliance on vectors for the transfer of pollen loads promotes the evolution of diverse floral adaptations associated with agents responsible for pollination. With biotic pollination, floral structures have evolved that facilitates both pollen delivery and pollen uptake with a single visit by the animal vector (Ollerton *et al.*, 2011; Devaux *et al.*, 2014). However, who is mating with whom is not random, but highly specific and genetically determined, and this is how the mating patterns have evolved (Hiscock and McInnis, 2003; Charlesworth, 2006).

The ratio of pollen to ovule is the most frequently used character to infer mating system (Scalone *et al.*, 2013). Angiosperms have developed two clearly opposite mechanisms for mating in their flowers. Selfing species, are defined by their ability to successfully self-fertilize, while outcrossing species have acquired morphological, physiological or genetic mechanisms that prevent selfing (Barret and Harder, 1996; Goodwillie *et al.*, 2005). Still, others have evolved mixed mating system, where a mixture of self and cross-fertilization determines the range between strict inbreeding and outbreeding (Carrio and Guemes, 2013; Devaux *et al.*, 2014). Previously, mixed mating system was regarded as a transient and non-adaptive state. But more recently, it has been shown that mixed mating system may become evolutionarily stable. Thus plants with mixed mating system have an edge, as it combines the advantages of both reproductive strategies by promoting outcrossing as well as assures reproductive success when outbreeding chances are limited (Goodwillie *et al.*, 2005; Carrio and Guemes, 2013).

The evolutionary transition from outbreeding to predominant selfing has occurred in many plant groups. Further, selfing has evolved numerous times in flowering plants, and has biologically critical and long lasting consequences (Goodwillie *et al.*, 2005; Charlesworth, 2006; Scalone *et al.*, 2013). This evolutionary shift toward selfing was associated with modification in floral biology, life history, and ecology (Devaux *et al.*, 2014). Several aspects that address the importance of selfing and outcrossing on fitness through inbreeding depression and heterosis are well documented. Similarly, the frequency of outcrossing is crucial to the overall population genetic structure, as it affects

the genetic diversity within a mating population. Nonetheless, the acquisition of selfing profoundly influences floral evolution, affecting floral design and resources allocation within sexes (Barret and Harder, 1996; Goodwillie *et al.*, 2005).

On the contrary, the aggregation of flowers in inflorescences, is likely to favour both intra-floral and geitonogamous self pollination. Therefore, it is not surprising, that many angiosperms have also acquired adaptations that prevent selfing (Charlesworth, 2006). For example, the physical separation of female and male gametophytes gave greater powers of maternal mate discrimination and has resulted in the evolution of a complex series of cellular and molecular interaction between the haploid pollen and the diploid pistil (Heslop-Harrison and Shivanna, 1977). These complex series of pollen-pistil interaction constitute a form of 'courtship' where; various recognition processes for the discrimination or rejection of 'incompatible' pollen take place. Such kind of self sterility may be regulated by various mechanisms, but in some plants the rejection of self-pollen or pollen tubes inhibition is due to genetically determined self-incompatibility. Among all pollen-pistil interactions studied so far at the molecular level, self-incompatibility is perhaps the best understood mechanism (Goodwillie *et al.*, 2004; Hiscock and Allen, 2008).

Our current knowledge about saffron reproductive biology is limited and the mechanisms underlying its sterility are not fully understood. However, comparative studies on the pollen and pistil of *C. sativus*, *C. cartwrightianus*, *C. thomasii* and *C. hadriaticus* revealed highest percentage of anomalous and lower percentage of viable pollens in *C. sativus* (Grilli Caiola *et al.*, 2011). Only a small proportion of saffron pollens germinate on stigma, from where still a fewer number could penetrate the ovules, and a direct relationship exist between defective germination and abnormal pollens division (Chichiricco and Grilli Caiola, 1986). RNase and peroxidase activities are responsible for self incompatibility in dicots, both RNase and peroxidase analyses were carried out in *C. sativus* L. aggregate which revealed pollen growth was not inhibited on the stigma surface or style but in another region of the gynoecium (Zanier and Grilli Caiola, 2011). Besides saffron, microspore of the other diploid *Crocus* species may also germinate on saffron stigma. Further, interspecific hybridization is possible and viable seeds have been obtained from all diploid species after free and cross-pollination as well as from a cross between *C. sativus* with *C. cartwrightianus* (Grilli Caiola *et al.*, 2010; 2011). Therefore, it will be very interesting to see if sterility in saffron could be broken by doubling its chromosomes and to see if the pollens are still viable and self compatible. Such a possibility may open a new and exciting era for the development of genetically improved saffron varieties.

Nonetheless, how competition for pollination might be altered in future or new mechanisms evolves in scenario of a changing world, where alien species are spreading, climate is changing and pollinator declines.

1.7 Genome analysis, diversity and evolution

In the last two decades, study of the genomes of plants has become ubiquitous for understanding diversity and relationships within and between species. These results address the long-standing questions about relationships and evolution of species back to the origin of plants (Soltis *et al.*, 2009; Jiao *et al.*, 2011; Buggs, 2013). The nuclear DNA and its associated proteins in a nucleus are divided among the chromosomes, where a single unbroken linear DNA molecule runs from one end to the other (Heslop-Harrison and Schwarzacher, 2011). Chromosomes provide physical structure for genetic linkage groups that allows faithful transmission of hereditary characters (Schmidt and Heslop-Harrison, 1998; Schwarzacher, 2003a; Frello *et al.*, 2004). The packaging of DNA within chromosomes prevent DNA from becoming unmanageable, and ensures that it is readily available for various cellular processes of replication and repair (Heslop-Harrison, 2000b, Fuchs *et al.*, 2006; Heslop-Harrison and Schwarzacher, 2011).

The structural and functional description of eukaryotic chromosomes into heterochromatin and euchromatin, telomeres, and nucleolus organizing regions (NORs) has been known for a long time (Bedbrook *et al.*, 1980; Jiang *et al.*, 2003; Carvalho *et al.*, 2009). Euchromatin, stains lightly in cytological preparations and lies at the interstitial and distal regions of the chromosomes, while the heterochromatin has highly condensed chromatin that stains strongly in cytological preparations (Cuadrado and Jouve, 1995; Schwarzacher, 2003b; Ma *et al.*, 2007). The heterochromatic blocks generally lie at the telomeric and pericentromeric regions of chromosomes (Heslop-Harrison and Schwarzacher, 2011). Euchromatin is a gene dense region of the chromosome, with high meiotic recombination and transcriptional activity. It lies at the interstitial and distal regions of the chromosomes. By contrast, heterochromatin in general is rich in repetitive DNA and transposable elements (TEs). It lacks meiotic recombination and has relatively fewer numbers of genes (Burgess 2013; Senerchia *et al.*, 2013).

In higher plants, unlike genome size and chromosome number, which vary widely, gene content is uniform (Schmidt and Heslop-Harrison, 1998; Devos, 2010; Bennett and Leitch, 2011). For example, the nuclear DNA content shows as much as 2350 fold range

among the measured 6288 species and chromosome number varies between $n = 2$ to $n =$ approximately 600 chromosomes (Bennett and Leitch, 2011; Schwarzacher, 2011b). Most flowering plants have undergone a series of whole genome duplication (WGD) events that have played a significant role in their evolution and diversification (Heslop-Harrison, 2000a, 2000b; Soltis *et al.*, 2009; Jiao *et al.*, 2011; Buggs, 2013). The availability of nucleotide sequence data has added a lot to our understanding of plant genome duplications, their evolution and relationships (see section 1.7, 1.8). A variety of functional molecular marker systems have been developed over time, by exploiting the molecular DNA diversity within genomes (Korzun, 2002; Varshney and Dubey, 2009). These molecular markers are playing an increasingly important role in the management and utilization of plant genetic resources by allowing more objective and precise quantification of genetic diversity and deducing their phylogeny (Villalobos and Engelmann, 1995; Fu, 2003; Todorovska *et al.*, 2005). By 2009, the Angiosperm phylogeny group (APG) revised and updated the classification for families of flowering plants based on their affinities. Many formerly unplaced families were assigned into orders as well as new orders were adopted, that greatly reduced the number of unplaced taxa and resolved the everlasting shortcomings associated with morphological markers (Bremer *et al.*, 2009).

1.8 Molecular DNA diversity: significance and exploitation of novel resources

Global population levels are expected to reach from 6 to 9 billion by 2050. Handling this rapid increase in human population and shrinkage of agricultural land will remain a challenge for the mankind (<http://www.fao.org>). On the one hand, when efforts are underway to increase crops productivity, in several countries yield has been affected by global rise in temperature. Similarly, in crops that propagates vegetatively, lack of genetic variation is limiting our abilities to develop higher yielding varieties (Chakraborty and Newton, 2011; King *et al.*, 2013). Thus identification and exploitation of novel genetic variation for the development of superior varieties that are better adapted to the changing environment and new crop uses such as bio-fuels are critical (Takeda and Matsuoka, 2008; Heslop-Harrison and Schwarzacher, 2012; King *et al.*, 2013).

Useful genes were inherited from a common ancestor, and are distributed in both wild and domesticated taxa. Improvement in crops is possible by exploiting diversity

within these genes (see Takeda and Matsuoka 2008 and section 1.3.2 above). Maintaining this diversity is crucial to breeding programmes as the loss of it may result in susceptibility to different types of stresses and extinction (Holt and Pickard, 1999; Fernandez 2007; Gresta *et al.*, 2008). Conversely, decades of intensive selection in domesticated taxa for only few quality traits have narrowed down their genetic base and the process may be still ongoing (Li *et al.*, 2004; Andrabi and Maxwell, 2007; Heslop-Harrison, 2010; Fernandez *et al.*, 2011). Therefore, precise screening to identify and maintain the diverse germplasm resources is highly important.

The bulk of variation at the nucleotide level is often not visible at the phenotypic level and since phenotype is the interaction of genotype with environment and therefore, very often the same genotype (*Crocus sativus*, for example see Figure 1.4) display diversity when maintained in different ecological or growth regimes (Macchia *et al.*, 2013). Thus, selection purely based on phenotypic traits may be inaccurate (Heslop-Harrison, 2000b; Fu, 2003; Todorovska *et al.*, 2005). On the other hand molecular markers are numerous in every genome and can be selected to be polymorphic. Further, DNA-based markers are stable, reproducible and provide a diagnostic approach (Korzun, 2002; Saeidi *et al.*, 2008). Being DNA markers, they are neither affected by developmental stages nor environmental factors (Todorovska *et al.*, 2005; Heslop-Harrison, 2010). Application of the various types of molecular markers has definitely played a major role in our understanding of the heritable traits and has increased the efficiency of our crop species tremendously. Few examples of the molecular markers that are generally applied to perceive diversity and relationships or tagging genes of interest are described below:

1.8.1 RFLPs

Restriction fragment length polymorphism (RFLP) analysis was among the first large-scale methods to analyse DNA sequence diversity. This method relies on the variations in homologous DNA sequences, differing in restriction enzymes sites (Grodzicker *et al.*, 1974; Kubis *et al.*, 2003). Identification of the variable size DNA fragments is made by Southern blotting, a procedure whereby DNA fragments, separated by electrophoresis, are transferred to positively charged nylon membranes (Southern, 1975; Heitkam and Schmidt, 2009). The membrane is hybridized with radioactive or chemiluminescent homologous probes and exposed to an X-ray film, where the different fragments are visible on autoradiography (Staub *et al.*, 1996; Woolrab *et al.*, 2012).

1.8.2 RAPD markers

Random amplified polymorphic DNAs (RAPDs) are PCR-based short (commonly 10 bp) single primer technology of DNA finger printing (Williams *et al.*, 1990). RAPDs are produced by using genomic DNA with arbitrary primers to amplify DNA segments between closely spaced sequences (Yang *et al.*, 2013). Polymorphism results from changes in the primer-binding site in the target DNA sequence (Devos and Gale, 1992; Yang *et al.*, 2013). Amplified products can be separated by electrophoresis on agarose or polyacrylamide gels and visualized by staining with ethidium bromide (Staub *et al.*, 1996; Yu *et al.*, 2013).

1.8.3 SCARs

RAPDs are generally regarded as unreliable with many questions about reproducibility and application to DNA extracted at different times or reactions from different laboratories. However, the efficacy of RAPD markers can be enhanced by sequencing the specific RAPD generated products and designing primers within the sequenced amplicons, to obtain a more specific marker for amplification (Zhuang *et al.*, 2013). Such sequenced characterized amplified regions (SCARs) are similar to sequence-tagged sites (STS) in construction and application, and are very useful in genotyping (Talbert *et al.*, 1996; Geetha *et al.*, 2013).

1.8.4 AFLPs

Amplified fragment length polymorphisms (AFLPs) exploits selective amplification of the restriction enzyme digested DNA fragments (Vos *et al.*, 1995). Restricted genomic DNA with sticky ends are ligated with adaptors and amplified with specified primers that are complementary to the sequence on the adaptors (see Yang *et al.*, 2013). Multiple bands are generated in each amplification reaction that contains DNA markers of random origin. AFLPs are quantitative and genotypes can be differentiated by the intensity of the amplified bands. The ability of this technology to generate many markers and its high resolution are features that make AFLPs very attractive among the genetic markers (Vos *et al.*, 1995; Frascaroli *et al.*, 2013).

1.8.5 Microsatellite DNA markers

Microsatellites or simple sequences repeats (SSRs) or short tandem repeats (STRs) are a class of short (1-6 bp), repetitive DNA element dispersed in all organisms and are abundant in plants (Heslop-Harrison, 2000a, 2000b; Nemati *et al.*, 2012; Yang *et al.*, 2013). These repeats occurring on average every 6-7 kb and the di-, tri- or tetra-nucleotide repeats are arranged in tandem arrays consisting of 5-50 copies (Tautz, 1989; Cardle *et al.*, 2000; Ozkan *et al.* 2005). The repeat motifs are flanked by conserved nucleotide sequences from which forward and reverse primers can be designed to PCR-amplify the DNA section containing the SSR (Tautz, 1989; Cardle *et al.*, 2000; Frascaroli *et al.*, 2013).

1.8.6 ESTs and EST-SSR

Expressed sequence tags (ESTs) are fragments of cDNA sequences complementary to mRNA and represent parts of expressed genes (Adams *et al.*, 1991, Varshney *et al.*, 2005). However, because of their conserved nature of genes in related lineages, ESTs may reveal low levels of polymorphism (Qi *et al.*, 2003; Gao *et al.*, 2004; Xue *et al.*, 2008). *In silico* mining of the ESTs databases allows discovery of SSRs from the coding region of these genome. EST-SSR or genic microsatellites markers are thus potential candidates for gene tagging and comparative studies in related species. For *Crocus*, a large number of ESTs have been placed in the GenBank database by D'Agostino *et al.* (2007) and exploited in the current study (see Chapter IV).

1.8.7 IRAP and REMAP markers

Inter-retrotransposon amplified polymorphism is a PCR-based method for amplification of genomic DNA between closely located sequences of retrotransposons with the help of outward-facing primers designed from the conserved domains between the long terminal repeat (LTRs) of a retrotransposon (Teo *et al.*, 2005; Biswas *et al.*, 2010). The IRAP markers are generated by the proximity of two retrotransposons using outward facing primers annealing to their long terminal repeats. IRAP can be amplified with a single primer matching either the 5' or 3' end of the LTR but oriented away from the LTR itself, or with two primers (Figure 1.11). Thus the PCR products therefore, represent amplification of hundreds of the target sites within a genome.

The retrotransposon microsatellite amplified polymorphism (REMAP) method is similar to IRAP, but one of the two primers matches an SSR motif with one or more non-SSR anchor nucleotides present at the 3' end of the primer (Mandoulakani *et al.*, 2012). Microsatellites are abundant in plant genomes (see above) and are often associated with retrotransposons (Kalendar and Schulman, 2007). Of the various techniques IRAP and REMAP detect high levels of polymorphism and are more frequently used in diversity analysis (Teo *et al.*, 2005; Nair *et al.*, 2005; Kalendar and Schulman, 2007; Saeidi *et al.*, 2008; Mandoulakani *et al.*, 2012).

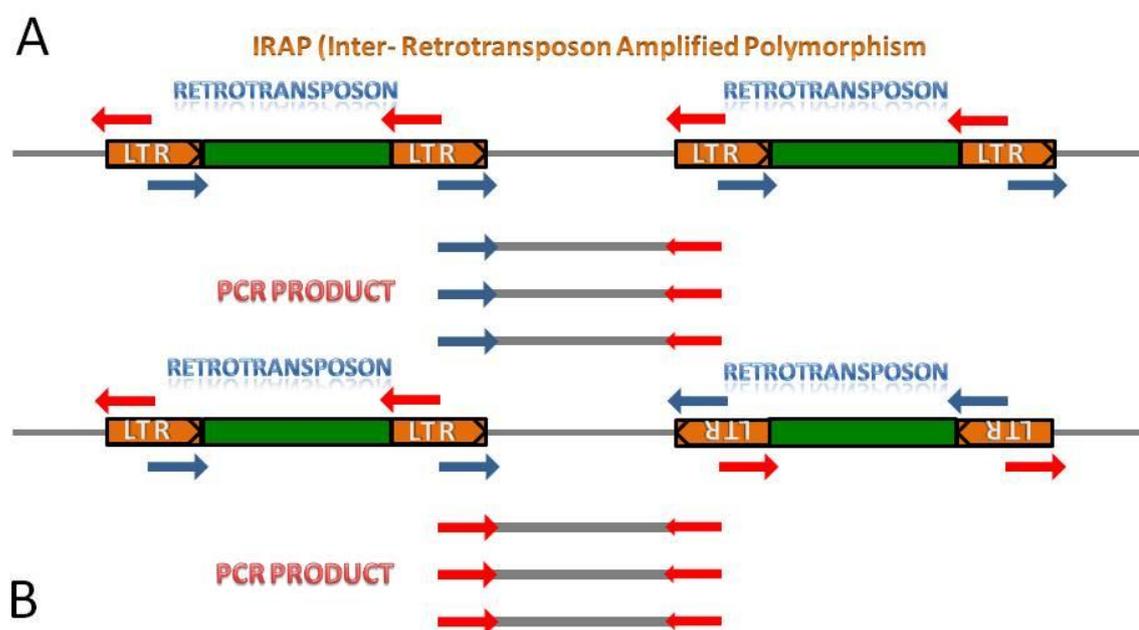


Figure 1.11: Schematic representation of IRAP markers amplification using two different primers (A) and the same primer as both forward and reverse (B). Arrows (grey and blue) in the right or leftward orientation indicating to the direction of amplification and priming sites of IRAPs marker within the LTR retrotransposons (modified from Barsirnia *et al.*, 2014).

1.8.8 DNA barcoding markers

DNA barcodes are short orthologous DNA sequences from a standardized region of the genome and are used with an objective, to use large-scale screening of genes or a standard short genomic region that is universally present in lineages (Hebert *et al.*, 2003; Kress *et al.*, 2005). Further, the target sequences are supposed to have enough sequence diversity and can discriminate among species (Kress and Erickson, 2007; Yang *et al.*, 2013). The nuclear ribosomal DNA internal transcribed spacer 2 (ITS2) region and MaturaseK gene (*MatK*) of the chloroplast are regarded as universal DNA barcodes (Hollingsworth *et al.*, 2009; Yao *et al.*, 2010). DNA barcoding genes including *rbcL*, ITS, *matK* and *trnH-psbA* are also applied in the current study to address the phylogeny of *C. sativus* and detailed results are given in (Chapter IV).

1.8.9 RAD sequencing

Restriction-site associated DNA (RAD) sequencing is a form of genotyping by sequencing strategy developed over the recent years for cost effective in-depth genomic analyses (Hipp *et al.*, 2014). RAD genotyping utilizes DNA fragments that lie adjacent to a particular restriction enzyme recognition site in a genome. Sequencing adapters are ligated to the restriction digested DNA prior to shearing and then sequencing of the tagged restriction sites is employed, rather than random sequencing of the whole genome. Further, RAD does not require the prior development of any genomic resources for the organism and allow parallel sequencing of millions of DNA fragments flanking individual restriction enzyme sites (Baird *et al.*, 2008). This approach has dramatically increased the coverage for a given sequenced site, and allows the discovery of thousands of single nucleotide polymorphisms (SNPs). So far, RAD genotyping has been effectively applied to resolve relationships among closely related as well as highly heterogeneous populations in both model and nonmodel organism, genetic mapping and quantitative trait loci (QTL) analyses, phylogeographic and population genomics (Hohenlohe *et al.*, 2010; Emerson *et al.*, 2010; Hipp *et al.*, 2014).

1.8.10 SNPs

Single nucleotide polymorphisms (SNPs) refer to genomic variation caused by a single nucleotide mutation at a specific locus (Lander *et al.*, 1996). SNPs may result from single base transitions, transversions, insertions or deletions and are distributed in both coding and non-coding regions of genomes (Vignal *et al.*, 2002; Yang *et al.*, 2013; Frascaroli *et al.*, 2013). A large number of SNP markers are now available for detailed analysis of genome structure, genome-wide association mapping and precision breeding. In both plants and animals. The 1000 Genome Project Consortium in the pilot phase revealed around 15 million SNPs in humans (see 1000 Human Genome Project Consortium, 2010 and Chapter IV).

1.8.11 Whole genome sequencing

Whole genome sequencing gives direct insights and provides a complete picture of the total genetic variation and gene content present in a population (Stein, 2007; Devos, 2010; Yang *et al.*, 2013). The current advancements in sequencing technologies at reduced costs have allowed the sequencing of 6870 complete genomes (August, 2013) including important species of plants, animals and microbes (<http://genomesonline.org>). Most progress in crop improvement over the few last decades and the better understanding of genomics today has been essentially possible due to the availability of these DNA sequences and it will continue to further enhance our understanding of the key biological phenomenon like domestication, hybridization and polyploidization (Varshney and Dubey, 2009; D'Hont, *et al.*, 2012).

1.9 Polyploidy in flowering plants

Most flowering plants have undergone one to several rounds of whole genome duplication (WGDs) events occurring near or at the time of their origins, suggesting that WGD has played a significant role in the origin of key novel traits that drove species evolution and diversification (Soltis *et al.*, 2009; Jiao *et al.*, 2011; Eric Schranz *et al.*, 2012; Buggs, 2013). For a long time the origin of novel traits in plant groups has fascinated scientists, including Darwin in his '*The Origin of Species*' (see Heslop-Harrison, 2012). Nearly 80% of the existing angiosperm species are considered polyploids, while in pteridophytes the

frequency of this polyploidy may be as high as 95% (Leitch and Bennet, 1997; Soltis and Soltis, 2000; Soltis *et al.*, 2010). Today, angiosperms represent the largest and most successful group of plants, with more than 300,000 living species. Their widespread occurrence and success may be attributed to the potential advantages they acquire through whole genome duplications (Bremer *et al.*, 2009; Harper *et al.*, 2012; King *et al.*, 2013). Besides WGDs, the subsequent evolutionary phenomena including later migration events, changing environmental conditions and differential extinction rates have also contributed to the ultimate success of angiosperms (Eric Schranz *et al.*, 2012).

Polyploidy is a recurrent, multiple-origin phenomenon, while the single-origin being the exception and not the rule (see Soltis and Soltis 1999, 2010). It is widely accepted that polyploidization either took place through genomic doubling or through non-reducing gametes, and during the process; the chromosomes doubled may involve a single genome (autopolyploidy) or a combination of two or more (allopolyploidy) genomes (Bardil *et al.*, 2011; Harper *et al.*, 2012; Buggs, 2013). However, based on meiotic behaviour analysis some allopolyploids such as *Leucaena confertiflora*, *L. pallida*, *L. leucocephala* and *L. involucrate* may have originated from somewhat similar species with some kind of genomic similarity, and are referred to as segmental allopolyploids (Boff and Schifino-Wittmann, 2003).

Most of the polyploidization events are several million years old (paleopolyploidy) while others may be relatively recent (neopolyploidy). However, both result in variation and in due course lead to speciation (Ortiz *et al.*, 2009; van der Peer and Pires, 2012). The first evidence of genomic duplication came from the analysis of the rice genome project (reviewed in Devos, 2010). Several of our important crop plants, such as bread wheat, oats, cotton, canola, coffee, and saffron are polyploids. In evolutionary terms, polyploid species have advantages over their diploid progenitors because it increases the net diversity (Heslop-Harrison, 2010; Jiao *et al.*, 2011; Arrigo and Barker, 2012; King *et al.*, 2013). However, inspite of the frequent incidence of polyploidy, recent analysis have found that newly formed polyploid species have higher extinction rates than their diploid relatives. These results suggest that despite leaving a substantial legacy in plant genomes, only rare polyploids survive over the long term (Arrigo and Barker, 2012). But how the possession of two or more than two complete sub-genomes from different parental species might contribute to their high yields is not fully understood, but for sure it adds to complexity of the genomes and transcriptomes of these crops (Harper *et al.*, 2012; Buggs, 2013). Nonetheless, a better understanding of the mechanisms of plant genome evolution

will be of critical value, as we could use that information to build a better and more sustainable crop and may conserve the huge biodiversity present within the plant kingdom (Heslop-Harrison, 2012; Van de Peer and Pires, 2012).

1.10 Germplasm collection resources and its exploitation

Genetic diversity in plants is present in natural resources, and during the recent years, failure to secure the local land races or substitution of local genotypes with improved varieties as well as changes in agricultural practices have caused a rapid erosion of this genetic diversity (Villalobos and Engelmann, 1995; Fu and Somers, 2009). Beside these, urbanization, pollution, habitat destruction, spread of invasive aliens and climatic changes are other factors that have endangered plant survival (Pitman and Jorgensen, 2002; Heslop-Harrison and Schwarzacher, 2012). In domesticated species of plants and animals lack of diversity and extinction may be related to their intense selection or inbreeding depression imposed by management techniques or market demands (Andrabi and Maxwell, 2007). A major step to prevent the loss and sustainable utilization of the valuable genetic resources is to ensure conservation of representative diversity of taxa (Fernandez *et al.*, 2011). A global response in the form of Convention on Biological Diversity (CBD 1993) is a milestone toward achieving this goal, as the convention's main goal is "conservation of biological diversity" (Maxted *et al.* 2007).

In general plant genetic resources may be secured in the form of *in situ* and *ex situ* conservation (Engelmann *et al.*, 2002; Paunescu, 2006; Fernandez *et al.*, 2011). *In situ* conservation is on site conservation, while *ex situ* conservation involves the maintenance of genetic resources away from their natural habitats in facilities called gene banks or botanical gardens (McGregor *et al.*, 2002; Vaughan *et al.*, 2007; Fernandez *et al.*, 2011). The genetic resources for *ex situ* collections may include local landraces, inbred lines and hybrids as well as open pollinated populations (see Paunescu, 2009). The principal aim of *ex situ* germplasm is to conserve the genetic diversity and stop the loss of potentially valuable material, which had barely been explored (Perrino *et al.*, 1991; Andrabi and Maxwell, 2007; Duc *et al.*, 2010). However, not only collection but, the accurate preservation of this germplasm and subsequent evaluation are highly important for future exploitation (Börner 2006; Heslop-Harrison, 2010).

Today the worldwide existing germplasm collection for food and agriculture comprises of around 7.5 million accessions (see Börner *et al.*, 2011). One of the most

challenging jobs in gene banks is proper maintenance of genetic variation that was initially reported in accessions (Vencovsky and Crossa, 1999). The major constraint is avoiding contamination by cross pollination especially for insect-pollinated outcrossing species (Ellstrand and Hoffman, 1990; Duc *et al.*, 2010). Molecular marker approaches are playing an increasingly important role in the management and utilization of plant genetic resources held in gene banks worldwide (Ayad *et al.*, 1997; Hodgkin and Rao, 2002). A major objective for the molecular germplasm characterization is to identify germplasm and duplicated accession (Villalobos and Engelmann, 1995; McGregor *et al.*, 2002). With the advent of new molecular tools and techniques, not only the precise genotypic characterisation but also accession that may have duplicates is becoming attainable (Karp *et al.*, 1997; McGregor *et al.*, 2002; Fu, 2003).

Although, extinction and speciation of living taxa is part of a natural process of evolution, the rate of extinction today, is much faster than speciation and human activities may be held responsible for this irreversible process (Holt and Pickard, 1999). In the latest *Red List* launched by the International Union for Conservation of Nature (IUCN-<http://www.iucnredlist.org/>) *Crocus cyprius* and *Crocus hartmannianus* are listed as vulnerable, while *Crocus etruscus* as nearly threatened (see CrocusBank project section 1.10.1). To some extent, the genetic diversity of many crops has been well preserved *ex situ*, but some vegetatively propagated crops such as saffron etc. have limited or unknown genetic diversity and require extra attention (Frello, and Heslop-Harrison, 2000; Fernandez *et al.*, 2007; Petersen *et al.*, 2008). Future food security is a major challenge for mankind, and efforts to maintain and secure our germplasm resources are vital. Fortunately, germplasm conservation is attracting more and more public concern, as it seems to be the only way to guarantee food supplies for future human generations (Villalobos and Engelmann, 1995; McGregor *et al.*, 2002; Börner *et al.*, 2011).

1.11 Preserving genetic resources of *Crocus* and its allies

The immense socioeconomic significance of saffron and potential threats to its survival implies that security of the worldwide *Crocus* germplasm is of paramount significance. Therefore, creation of a germplasm bank, to preserve the global genetic diversity of *Crocus* species may be considered as a great achievement (see section 1.9 above). The overall aim of such collection will be to maintain the required biodiversity which is a prerequisite for any future saffron breeding and improvement programmes to take place (Fernandez *et al.*, 2011). Further, every single species on earth has a unique role to play, and disappearance of it, affects functioning of the global ecosystem (Andrabi and Maxwell, 2007). Understanding the implications of *Crocus* germplasm security, the European Commission took initiatives to ensure preservation of the world's saffron germplasm diversity, and then to utilize the newly developed “omics” technologies for the better understanding and improvement of saffron at molecular level.

1.11.1 CROCUSBANK project

In 2007, the European Commission AGRI GEN RES 018 action approved a framework for World Saffron and Crocus Collection (WSCC), the “CROCUSBANK” project (<http://www.crocusbank.org>). The objective of this project was to create, characterise and exploit a germplasm collection (bank) of *Crocus* species, including saffron at a global scale. This was perhaps the first ever mega project of its nature in Europe assigned to *Crocus* conservation and diversity. The CROCUSBANK project included 14 best research groups at the World scale with expertise in saffron Agronomy, Breeding, Botany, Genetics, Chemistry, Food Technology and Commercialisation. The total budget of CROCUSBANK project was €2,634,137 and the project was co-financed with a contribution of €929,507 from the European Union. The project started on 01 June 2007 and the final technical report was submitted by December 2011 (<http://www.crocusbank.org>).

At present the germplasm collection of CROCUSBANK project is housed at the Bank of Plant Germplasm of Cuenca (BGV-CU, Spain), and currently consists of 761 effectively preserved accessions, 225 correspond to saffron material and 536 to other *Crocus*, including 62 different species (see Figure 1.12). The preserved biodiversity of saffron (*Crocus sativus* L.) covers a wide range of the genetic variability of the crop from

15 countries, 169 of these accessions come from European cultivation countries, 18 from commercial areas in non EU countries, 26 from regions of minimal or relict production and/or from abandoned fields and 7 from commercial nurseries. The non-saffron *Crocus* collection currently comprises 352 accessions: 179 were collected from the wild in 12 countries of natural distribution, 24 from donations of public and private institutions, 91 from commercial nurseries and 58 acquired from BGV-CU collection management (Fernandez *et al.* 2011; <http://www.crocusbank.org>). The CROCUSBANK project for the first time has resulted in a relational database with all sorts of data generated during the conservation and characterisation of *Crocus*. There was no list of descriptors for saffron and other *Crocus*, therefore lists of descriptors were elaborated and published to describe the variability observed in the most important crops and wild related species at World scale (Fernandez *et al.*, 2011; <http://www.crocusbank.org>). Being part of the CROCUSBANK project, my thesis also addressed characterization of genetic diversity and ancestry of saffron and my conclusions are given in the results chapters below.



Figure 1.12: A view of various *C. sativus* accessions maintained at Bank of Plant Germplasm of Cuenca (BGV-CU, Spain).

1.11.2 SAFFRONOMICS project

Once the CROCUSBANK project provided the baseline for *Crocus* research, it was highly important to maintain, carryover and exploit the *Crocus* germplasm available at BGV-CU, Spain (see <http://www.crocusbank.org>). Further, the remarkable progress in crop improvement would have never been possible without the development of new genomic technologies like next generation sequencing, high-throughput marker genotyping, omics and an understanding of the variation at the DNA, RNA and protein level (Varshney and Dubey, 2009; Heslop-Harrison and Schwarzacher 2012). Therefore, the European Cooperation in Science and Technology Food and Agriculture approved “SAFFRONOMICS COST Action FA1101” with an aim to build up a network of collaborative research in order to increase our knowledge of the structural organization of saffron genome DNA fingerprinting, chemical fingerprinting, proteomics, transcriptomics, and metabolomics. This integrated knowledge will be the basis for the development of saffron genetic improvement, and the maturity of reliable techniques for traceability applications, determination of authenticity, and for fighting against fraud of origin in saffron (<http://www.saffronomics.org/>).

To achieve the objective of the COST Action, international coordination, cooperative research, and a multidisciplinary approach has been adopted. This project join together geneticists, molecular biologists, biochemists, biotechnologists, analytical chemists, food technologists, plant breeders, but also manufacturers and experts in Saffron business. Furthermore, SAFFRONOMICS COST Action FA1101 allows the exchange of experts, scientists, and graduate students for training, especially through Short-Term Scientific Missions (STSMs), that for surely will facilitate the Action to achieve its objectives and to re-initiate saffron cultivation as a highly important and beneficial crop in the European countries.

1.12 Aims of thesis

As a curative plant, saffron is mentioned in the oldest available traditional medicines. On the other hand saffron is the most expensive spice on earth, and has been consumed since antiquity. Since the origin of agriculture, crop improvement has been a continuous process driven by the human needs for improved quality, yield, resistance and adaptation to new and changing climates. However, limited knowledge of saffron genetic diversity and shrinkage of land surface assigned to saffron cultivation are the main bottlenecks in saffron improvement and these made me to design a project to confront the potential challenges of future saffron. I was also interested to understand the possible genomic and evolutionary implications of the saffron polyploid nature and the genomes that comprise today's saffron. Chief aims of the project include:

The aims of this thesis are thus: -

- ❖ Assessment of potential diversity in saffron accessions and understanding the relationships of saffron with wild *Crocus* species to pinpoint ancestral species of saffron.
- ❖ Characterization of potential diversity and relationships of wild *Crocus* species and saffron using IRAPs markers.
- ❖ To exploit online *Crocus* EST databases and design novel PCR markers to understand the diversity and relationship among wild *Crocus* species and saffron.
- ❖ Exploit cloning and sequencing technology to identify new SNPs markers within *Crocus* ESTs and universal barcoding genes.
- ❖ Based on the results of IRAPs, ESTs, SNPs data, apply GISH with genomic DNA as a probe from potential parents, to identify the ancestor of saffron.
- ❖ Characterization of novel repeated DNA sequences from wild *Crocus* species and saffron.
- ❖ Use repetitive DNA sequences to understand the genomic organization, diversity, and chromosomal localization of repeated DNA sequences and to understand the phylogeny of saffron.

2 CHAPTER II: MATERIALS AND METHODS

2.1 Materials

2.1.1 Plant material

Forty-four different accessions from 24 species of the genus *Crocus* were used in the current study. Corms of all accessions were purchased commercially or were obtained from the CrocusBank germplasm collection (www.crocusbank.org/). For a few *Crocus* species and accessions DNA was kindly provided by Marcela Santaella-Tenorio (Biotechnology, IDR-UCLM, Albacete, Spain). All these *Crocus* species and accessions along their University of Leicester identification code, CrocusBank accession number and source they were obtained from, are listed in Table 2.1. The taxonomy (and authorities for naming) follows Petersen *et al.* (2008). When plants were received and grown, corm and flower morphology were generally checked to be in accordance with published photographs and descriptions (Philips and Rix, 1989).

2.1.2 Germination of *Crocus* corms

Two to five healthy corms per plant (both *Crocus* species and accessions) were grown at a depth of about 10-15 cm in labelled pots, on a mixture of compost (Scotts Professionals, UK) and sand (Sinclair Horticulture Ltd. UK) mixed in a ratio of 2:1. The corms were grown in a green house at the University Of Leicester, UK on 20-25°C temperature. Since most of the *Crocus* species and hybrids grow on well drained soils grit (Horticulture Grit, UK) was added to these pots to improve drainage. With few exceptions all corms sprouted very well in the laboratory conditions.

Table 2.1: List of *Crocus* species and accessions along CrocusBank accession number, University of Leicester identification code and source of origin used in the current study.

No.	Sub genus	Section	Series	Species	Sub taxa / variety	CrocusBank accession	University of Leicester Number	Source
1	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. sativus</i>	-	BCU002746	CsatP09	Pottertons Nursery (United Kingdom)
2	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. sativus</i>	-	BCU002744	CstVD09	JW Dix Export (The Netherlands)
3	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. sativus.</i>			CstPER09	J.Perez (Spain)
4	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. sativus.</i>			CstSUSD09	Suttons Nursery (United Kingdom)
5	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. sativus</i>	<i>cashmeriensis</i>	BCU002584	CstCD09	JW Dix Export (The Netherlands)
6	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. sativus</i>	Kashmir		Cstkf09	Srinagar, Kashmir
7	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. sativus cartwrightianus*</i>	Albus	BCU002754	CstcP09	Pottertons Nursery (United Kingdom)
8	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. cartwrightianus</i>	-	BCU002747	CcwBD09	JW Dix Export (The Netherlands)
9	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. cartwrightianus.</i>	albus	BCU002766	CcwAD08	JW Dix Export (The Netherlands)
10	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. cartwrightianus</i>	CEH.613	BCU002771	CcrCR09	Rare plant Nursery(United Kingdom)
11	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. pallasii</i>	<i>turcicus</i>	BCU002748	CpltR09	Rare plant Nursery(United Kingdom)
12	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. pallasii</i>	<i>pallasii</i>	BCU002767	CplVD09	JW Dix Export (The Netherlands)
13	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. pallasii</i>	<i>dispathaceus</i>	BCU002759	CplDD09	JW Dix Export (The Netherlands)
14	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. mathewii</i>			CmatD08	JW Dix Export (The Netherlands)
15	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. mathewii</i>	HKEP.9291		CmtHR09	Rare plant Nursery(United Kingdom)
16	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. thomasii</i>		BCU002751	CtmVD09	JW Dix Export (The Netherlands)
17	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. thomasii</i>	MS 978		CtomI09	Matera Italy
18	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. asumaniae</i>	white	BCU002757	CasWD09	JW Dix Export (The Netherlands)
19	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. asumaniae</i>	'alba'	BCU002760	CasAD09	JW Dix Export (The Netherlands)
20	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. asumaniae</i>	S9104		CasAT09	Aseki Turkey
21	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. oreoreticus</i>	VV.CR.114	BCU002774	CorVR09	Rare plant Nursery(United Kingdom)
22	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. oreoreticus</i>		BCU002756	CorVD09	JW Dix Export (The Netherlands)
23	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. hadriaticus</i>		BCU002764	ChdWD09	JW Dix Export (The Netherlands)
24	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. hadriaticus</i>	'Indian summer'	BCU002770	ChaIR09	Rare plant Nursery(United Kingdom)
25	<i>Crocus</i>	<i>Crocus</i>	Crocus	<i>C. hadriaticus</i>	Alepohori (AH8682)		ChdARD09	Rare plant Nursery(United Kingdom)

Table 2.1: continued.

No.	Sub genus	Section	Series	Species	Sub taxa	Crocus bank Acc.Number	University of Leicester Number	Source
26	<i>Crocus</i>	<i>Crocus</i>	Verni	<i>C. vernus</i>		BCU001842	VER01/10	
27	<i>Crocus</i>	<i>Crocus</i>	Verni	<i>C. tommasinianus</i>	'lilac beauty'	BCU002765	CtmLD09	JW Dix Export (The Netherlands)
28	<i>Crocus</i>	<i>Crocus</i>	Verni	<i>C. tommasinianus</i>	'barr purple'	BCU002768	CtmBD09	JW Dix Export (The Netherlands)
29	<i>Crocus</i>	<i>Crocus</i>	Verni	<i>C. tommasinianus</i>	'rubinetta'	BCU002762	CtmTD09	JW Dix Export (The Netherlands)
30	<i>Crocus</i>	<i>Crocus</i>	Verni	<i>C. tommasinianus</i>	'albus'	BCU002763	CtmAD09	JW Dix Export (The Netherlands)
31	<i>Crocus</i>	<i>Crocus</i>	Versicolores	<i>C. versicolor</i>	'picturatus'	BCU002761	CvrPP09	Pottertons Nursery (United Kingdom)
32	<i>Crocus</i>	<i>Crocus</i>	Longiflori	<i>C. niveus</i>			CnivD08	JW Dix Export (The Netherlands)
33	<i>Crocus</i>	<i>Crocus</i>	Longiflori	<i>C. goulimyi</i>	'leucanthus'	BCU002755	CgulD08	JW Dix Export (The Netherlands)
34	<i>Crocus</i>	<i>Crocus</i>	Kotschyani	<i>C. kotschyanus</i>	kotschyanus		CkotP09	Pottertons Nursery (United Kingdom)
35	<i>Crocus</i>	<i>Crocus</i>	Kotschyani	<i>C. kotschyanus</i>	Zonatus		Ckot/z08	Garden Source
36	<i>Crocus</i>	<i>Nudiscapus</i>	Reticulati	<i>C. angustifolius</i>			CangP09	Pottertons Nursery (United Kingdom)
37	<i>Crocus</i>	<i>Nudiscapus</i>	Orientalis	<i>C. korolkowii</i>	gold-colored		CkrGD10	JW Dix Export (The Netherlands)
38	<i>Crocus</i>	<i>Nudiscapus</i>	Flavi	<i>C. flavus</i>			CflaP09	Pottertons Nursery (United Kingdom)
39	<i>Crocus</i>	<i>Nudiscapus</i>	Speciosi	<i>C. speciosus</i>	speciosus	BCU002753	CspP09	Pottertons Nursery (United Kingdom)
40	<i>Crocus</i>	<i>Nudiscapus</i>	Laevigati	<i>C. laevigatus</i>			Clae08	JW Dix Export (The Netherlands)
41	<i>Crocus</i>	<i>Nudiscapus</i>	Laevigati	<i>C. boryi</i>			Cbor08	JW Dix Export (The Netherlands)
42	<i>Crocus</i>	<i>Nudiscapus</i>	Aleppici	<i>C. veneris</i>	Creamy		CvenD10	JW Dix Export (The Netherlands)
43	<i>Crocus</i>	<i>Nudiscapus</i>	Reticulati	<i>C. cancellatus</i>	cancellatus		CcanD10	JW Dix Export (The Netherlands)
44	<i>Crocus</i>	<i>Nudiscapus</i>	Biflori	<i>C. biflorus</i>			CbfaD10	JW Dix Export (The Netherlands)

* The accession was purchased under this unrecognized name. It has similarities to *C. cartwrightianus* but is probably not this species.

2.1.3 Standard solutions and media used

Table 2.2: Unless indicated otherwise all solutions were made up with purified water (ELGA LabWater, High Wycombe, UK), autoclaved and stored at room temperature (RT).

Solution	Preparation/final concentration
CTAB buffer (pH 7.5 - 8.0)	2% (w/v) cetyltrimethylammonium bromide, 100mM Tris-HCl, 1.4M NaCl, 20mM EDTA.
DNA Wash buffer	76 % ethanol, 10mM ammonium acetate. No autoclaving.
10x TE buffer ^{*1} (pH 8.0)	100mM Tris (tris-hydroxymethylamino-methane)-HCl, 10mM EDTA (ethylene-diamine-tetra-acetic acid).
6x Gel loading buffer	0.25% Bromophenol blue, 0.25% Xylene cyanol FF, 60% Glycerol. No autoclaving and stored at 4°C.
50x TAE ^{*1} (pH 8.0)	242g of Tris-base, 57.1ml of glacial acetic acid, 100ml of 0.5M EDTA. Final volume 1000ml with sterile distilled water.
Ethidium Bromide (10 mg/ml)	1g Ethidium bromide, 100ml of sterile distilled water. No autoclaving and stored at 4°C.
Ampicillin	10mg/ml (dissolved in distilled water). No autoclaving and stored at -20°C.
20x SSC (saline sodium citrate, pH 7.0) ^{*1}	0.3M NaCl, 0.03M sodium citrate.
0.5 M EDTA (pH 8.0)	186.1g disodium ethylenediamine tetraacetate.2H ₂ O into 800ml of distilled water. Adjust pH to 8.0 with NaOH. Final volume 1 litre.
Detection buffer (FISH)	4x SSC, 0.2% (v/v) Tween 20.
20% SDS	2g Sodium dodecyl sulfate (SDS) with 8ml water Not autoclaved
10x Enzyme buffer (pH 4.6) ^{*1}	40mM citric acid, 60mM tri-sodium citrate. No autoclaving, stored at 4°C.
1x Enzyme solution	3% (w/v) pectinase (Sigma), 1.8% (w/v) cellulase (Calbiochem), 0.2% (w/v) cellulase (Onozuka RS) in 1x enzyme buffer. No autoclaving and stored at -20°C.
2 mM 8-hydroxyquinoline	2 mM 8-hydroxyquinoline: 0.15 g 8-hydroxyquinoline, 500 ml ddH ₂ O. Store in the dark at 4°C.
Alpha bromonaphthalene	1 litre of standard α -bromonaphthalene solution in sterile distilled water
4% Paraformaldehyde (pH 7.0)	4g paraformaldehyde (Agar Scientific) dissolved in distilled water. Final volume 100ml, no autoclaving and used fresh.
McIlvaine's buffer (pH 7.0)	0.1M citric acid, 0.2M di-sodium hydrogen phosphate.
Blocking DNA ^{*2}	Autoclaved at 114°C for 5min

Table 2.2: continued

100µg/ml DAPI ^{*3}	5g of DAPI (4',6-diamidino-2-phenylindole) dissolved in Sigma water. Final volume 50ml. No autoclaving and stored at -20°C.
50x Denhardt's solution	1% Ficoll type 400 (Sigma), 1% polyvinylpyrrolidone (Sigma) and 1% bovine serum albumin (Amersham Biosciences). Filter sterilized and stored at -20°C.
Southern denaturing solution	0.25M NaOH, 1M HCl.
Southern depurinating solution	0.25M HCl.
Southern neutralizing solution (pH 7.5)	0.5M Tris-HCl, 3M NaCl.
Southern Transfer buffer	0.4M NaOH.
Buffer 1 (probe detection, pH 7.5)	100mM Tris-HCl, 15mM NaCl
Buffer 2 (probe detection)	0.5% (w/v) Blocking Reagent (Roche Diagnostics) in buffer 1
Buffer 3 (probe detection, pH 9.5)	100mM Tris-HCl, 100mM NaCl, 50 mM MgCl ₂
Salmon sperm DNA ^{*4}	1mg/ml of sheared salmon sperm DNA.
Wash buffer 1 (Southern hybridization, pH 7.5)	0.1M Maleic acid, 0.15 M NaCl, 0.3% (v/v) Tween 20
Buffer 1 (Southern hybridization, pH 7.5)	0.1M Maleic acid, 0.15 M NaCl
Buffer 2 (Southern hybridization)	1% (w/v) Blocking Reagent (Roche Diagnostics) in buffer 1
Buffer 3 (Southern hybridization, pH 9.5)	0.1M Tris-HCl, 0.1 M NaCl
SOB medium (super optimal broth, pH 7.0)	20g of Tryptone, 5g Yeast extract, 0.5g NaCl, 10ml 250mM KCl. Final volume 1000ml with sterile distilled water.
LB medium (Luria-Bertani, pH 7.0)	10g Tryptone, 5g Yeast extract, 10g NaCl. Final volume 1000ml with sterile distilled water and autoclaved
LB medium Agar	10g Tryptone, 5g Yeast extract, 10g NaCl. Final volume 1000ml with sterile distilled water and 1.5% Agar.
200 mM IPTG	476mg/ml isopropyl-B-D-thiogalacto-pyranoside (dissolved in 10ml distilled water). Filter sterilized and stored at -20°C.
40mg/ml Xgal	1g of 5-Bromo-4-chloro-3-indolyl β-D-galactopyranoside with 25ml of Dimethylformamide. Filter sterilized, stored at -20°C.
50% Dextran sulfate	50 gm Dextran sulfate with 100 ml distilled water, Filter sterilized and stored at -20°C.

*1 Diluted with distilled water to appropriate concentration

*2 Genomic DNA from *Crocus sativus* was sheared into pieces and applied 4-20x of the probe concentration to block the repetitive DNA sequences.

*3 DAPI was diluted in water for stock of 100µg/ml and then diluted with McIlvaine's buffer to final concentration of 4µg/ml.

*4 Salmon sperm DNA was denatured in boiling water for 10 min and placed on ice for 10 min before adding it to the hybridization mixture.

2.2 Methods

2.2.1 Isolation of total genomic DNA

Genomic DNA was extracted from freshly collected green young leaves of *Crocus sativus* and related species (Table 2.1) using cetyltrimethylammonium bromide (CTAB) method modified from Doyle and Doyle (1990). The figures in this chapter give a simplified graphic overview of the key points of the procedures, with Figure 2.1 showing the steps in DNA extraction. One to two grams of young leaves were collected from a single individual and ground to a fine powder using a pestle and mortar in liquid nitrogen to prevent enzymatic degradation. A small amount of fine sand was added to help the grinding process. Half spatula of PVP (Poly Vinyl Pyrrolidone, Sigma, c. 0.2g) was added before the powdered leaf was taken into a 50ml Falcon tube with 5ml of pre-heated CTAB buffer (Table 2.2) containing 50 μ l of β -mercaptoethanol. Tubes were incubated at 60°C for 1hr in a shaking water bath. An equal volume of absolute chloroform : isoamyl alcohol (24:1) was added to each tube and mixed by repeated inverting for 3 min, followed by centrifugation at 5000 rpm at RT for 10 min. The aqueous supernatant was carefully transferred to a new Falcon tube using 1ml blue tip cut at the end. The chloroform : isoamyl alcohol washing and centrifugation steps were repeated and the DNA was precipitated with 0.6 volume of pre-chilled isopropanol added to the supernatant, mixed gently by inverting and then kept on ice for 10 min.

Precipitated DNA was spooled out with a sterile glass rod or spun down at 2000 rpm for 2 min, dried and washed with 5ml of wash buffer (Table 2.2) for 20 min, and then air dried before resuspending DNA in 1ml of 1x TE buffer (Table 2.2). The extracted DNA was then incubated at 37°C for 1 hr with 1 μ l of 10mg/ml RNase A (Bioline) to get rid of RNA. A diagrammatic representation of the process is given as Figure 2.1 below. Adequate measures were taken at all the times to avoid contamination of the genomic DNA samples from any DNA or dust present in the surrounding. Stocks of genomic DNA were stored in a -20°C freezer.

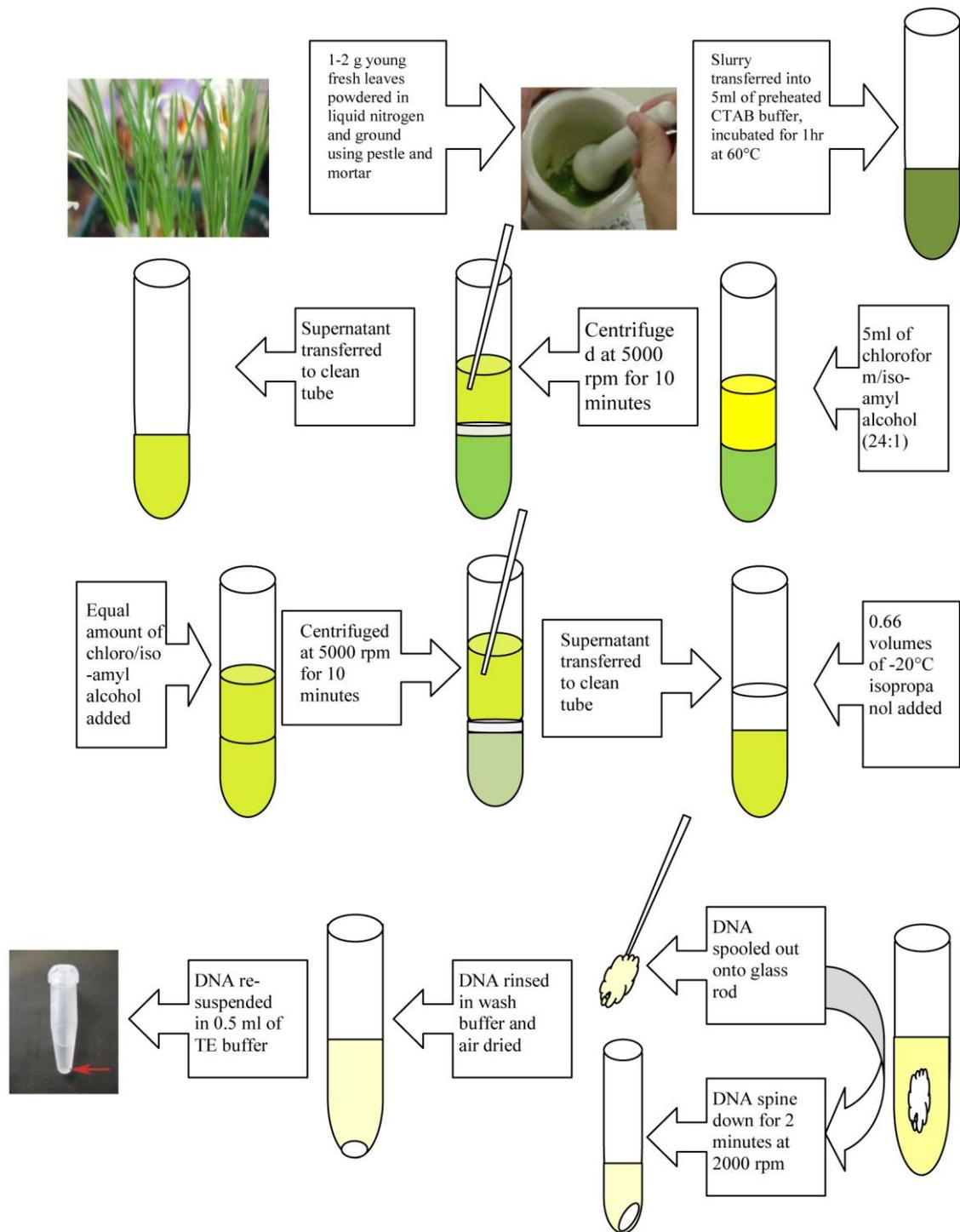


Figure 2.1: Flow diagram showing the standard extraction of genomic DNA procedure using CTAB method, modification of Doyle and Doyle (1990).

2.2.2 Agarose gel electrophoresis and purification of PCR products

Both genomic and PCR-amplified DNA products (section 2.2.5) were separated by agarose gel electrophoresis. Agarose gels [1-3% (w/v)] were prepared by boiling agarose in a microwave oven (Molecular Grade, Bionline or Hi-Res Super AGTC Agarose, Geneflow, UK) in 1x TAE buffer (Table 2.2) and poured into sealed gel trays after adding ethidium bromide (final concentration of 0.5µg/ml). Gel combs were placed to make wells for loading DNA samples and then left at RT to solidify. Genomic DNA samples were mixed with appropriate amount of 6x gel loading buffer (Table 2.2). While for PCR products 6x gel loading buffer was diluted 1:3 with 70% glycerol and loaded along DNA length marker, 5µl/lane Hyperladder I (Bionline) or Q-Step 2 (YorkBio) and run on 7V/cm for 45-60 min or at a slow speed of 15V for 15hrs and visualized with GeneFlash (Syngene, UK) gel documentation system. For IRAPs markers 2% (w/v) agarose gels were prepared by mixing normal (Molecular Grade, Bionline) and Hi-Res Super AGTC Agarose, (Geneflow, UK) in ratios of 3:1 respectively, while SNPs, EST-SSRs, barcoding markers electrophoresis were carried out on 2-3% (w/v) agarose gels (Molecular Grade, Bionline).

After analyzing agarose gels, selected PCR bands were excised and purified with the QIAquick Gel extraction kit (Qiagen) following the manufacturer's protocol (www.qiagen.com). Once the residual contaminants were removed, 1µl of the recovered DNA was reloaded on 1% (w/v) agarose gel to confirm the size and concentration, before using the eluted DNA in probe labelling, cloning or for direct sequencing.

2.2.3 DNA quantification

The concentration and quality of DNA (genomic and eluted) was assessed through gel electrophoresis and using a NanoDrop 8000 Spectrophotometer (Thermo Scientific). For electrophoresis 1µl DNA was loaded on 1% (w/v) agarose gel (section 2.2.2), while for the NanoDrop, 1µl of the genomic DNA was measured directly using default wavelength spectrum that ranged from 220-750nm. High molecular weight DNA samples with no visible shearing on gels and NanoDrop spectrophotometer O.D.260/O.D.280 ratio of 1.8 or above was used for subsequent PCR amplifications and restriction digestion experiments (section 2.2.13).

2.2.4 PCR markers and primer design

PCR markers including, IRAPs (Inter Retroelements Amplified Polymorphism), SNPs (Single Nucleotide Polymorphism), EST-SSRs (Expressed Sequence Tags Simple Sequence Repeat) and barcoding genes were applied in the current study. *Crocus* ESTs were downloaded from National Centre for Biotechnology Information (NCBI) and microsatellites (SSR) sequences were identified using Tandem Repeat Finder package (<http://tandem.bu.edu/trf/trf.html>) Unless obtained from published sources, primer pairs were designed using the online program Primer 3 (http://www.frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) with the annealing temperature set from 50-60°C, and optimal length of 20 bases preferably with 50% GC content for the amplification of products between 100-400bp size, and ordered from Sigma (www.sigmaaldrich.com/). Complete details of markers are given in the respective results chapters.

2.2.5 Polymerase Chain Reaction (PCR)

DNA was amplified using a Tprofessional Gradient Thermocycler (Biometra) in a 15µl reaction mixture containing 50-100ng of template DNA, 1x Kapa Biosystems buffer A [750mM Tris-HCl pH 8.8, 200mM (NH₄)₂SO₄, 15mM MgCl₂, 0.1% Tween 20], 1.5mM MgCl₂, 200µM of dNTPs (Bioline), 0.6µM of each primer and 0.5U of Kapa Taq DNA polymerase (Kapa Biosystems, USA). IRAP primer amplification was carried out as described in Teo *et al.* (2005) and PCR conditions were: 95°C for 2 min, followed by 30 cycles at 95°C of 1 min, 40-60°C for 1 min (depending upon optimized annealing temperature of different primer sets), 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 was followed by holding the block at 16°C.

PCR conditions for EST-SSR, SNPs and barcoding primers were: 94°C for 4 min, followed by 30 cycles at 94°C for 1 min, 45-60°C (depending upon the annealing temperature of different primer sets) for 45 s, 72°C for 2 min, and final extension of 72°C for 7 min was followed by holding the block at 16°C. Similarly, plasmid and colony DNA was amplified in a final volume of 50µl containing 1× PCR buffer [16mM (NH₄)₂SO₄, 67mM Tris-HCl, 0.1% Tween 20 (Bioline)], 1.5mM MgCl₂, 200µM of dNTPs (Bioline), 0.4µM of each M13 primers (forward: 5'-GTA AAA CGA CGG CCA GT-3', reverse: 5'-GGA AAC AGC TAT GAC CAT-3'), 0.5U of Taq DNA

Polymerase (Bioline) and 0.5µl of recombinant plasmid DNA. PCR cycling conditions were: 94°C for 5 min, 35 cycles of 94°C 30 s, 50°C for 30 s and 72°C for 45 s, followed by 72°C for 5 min and holding the block at 16°C. Amplification and polymorphism of the PCR products were analyzed by electrophoresis on 2 % (w/v) agarose gels as described above (section 2.2.2).

2.2.6 Cloning of PCR products

Purified PCR fragments (section 2.2.2) were cloned in pGEM[®]-T Easy vectors, using pGEM[®]-T Easy Vector System I kit (Promega) following the manufacturer's protocol with little modification. The cloning site of pGEM[®]-T Easy vector has a single overhanging 3' deoxythymidine (T) nucleotide that can be ligated to a single base deoxyadenosine (A) to the 3' end of the PCR products generated by Taq polymerase. Since both *Hae*III and *Dra*I produces blunt end fragments, thus deoxyadenosine was added through a single step PCR in a total of 20µl solution containing 1x Kapa Biosystems buffer A, 2mM MgCl₂, 0.2µM of dATP (Bioline), 10U of Kapa Taq DNA polymerase (Kapa Biosystems, USA) and 6µl of eluted DNA. The reaction was incubated at 70°C for 30 min followed by holding the block at 16°C.

2.2.6.1 Ligation reaction and transformation of competent *E. coli* cell

Ligation reactions of 10µl were set up in a small 300µl tube, that comprised of 5µl of 2x Rapid Ligation Buffer [60mM Tris-HCL pH 7.8, 20mM MgCl₂, 20mM DTT, 2mM ATP, 10% PEG (Promega)], 0.9µl of the pGEM-Teasy vector, 1.2µl of T4 DNA Ligase and 5.4µl of purified PCR product were mixed and incubated at RT for 1 hr, or at 4°C overnight. The insert: vector ratio was calculated below following guidelines provided by Promega (<http://www.promega.com/>). For transformation, 5µl of the ligation reaction was added to 50µl of the competent *E. coli* cells (α -Select Bronze Efficiency, Bioline) and was kept on ice for 20 min before a heat shock of 42°C for 45 s, which was again followed by 2 min on ice. Pre-warmed 750µl of Super Optimal Broth media SOB (Table 2.2) was added to each reaction tube on ice and then incubated at 37°C for 1.5 hr in an orbital shaker at 230 rpm to allow the growth of transformed competent cells. After the incubation, 50µl, 100µl & 200µl of culture was plated on three LB agar plates, containing 100µg/ml Ampicillin, 40µg/ml 5-bromo-4-chloro-3-indolyl- β -D-

galactosidase (X-gal) and 500 μ M isopropyl- β - Δ -thiogalactopyranoside (IPTG). Plates were incubated on 37°C for 14-16 hrs (Figure 2.2).

$$\frac{\text{ng of vector} \times \text{kb size of insert}}{\text{kb size of vector}} \times \text{insert: vector molar ratio} = \text{ng of insert}$$

2.2.6.2 Screening for recombinant clones

Recombinant clone selection was based on screening for blue and white colonies. The pGEM[®]-T Easy vector contains lacZ gene encoding for β -galactosidase that breaks down the chromogenic X-gal substrate and results in blue colonies. Successful transformation results in the disruption of the plasmid β -galactosidase gene (lacZ) and colonies appear white due to their inability to metabolize X-gal. Single white colonies were picked with a sterile toothpick and inoculated in 5ml LB medium (Table 2.2) with 40 μ g/ml of Ampicillin and incubated overnight at 230 rpm in an orbital shaker at 37°C. For sequencing five white colonies per plate were selected and commercially sequenced (see below). To recover transformed *E. coli* cells, 750 μ l of medium were spin down in a 1.5ml Eppendorf tube at 13000 rpm for 1 min, the supernatant was carefully decanted and this process was repeated 3-4 times until a pellet of appreciable size was obtained.

2.2.6.3 Plasmid DNA purification, insert confirmation and storage of *E. coli* cells

Recombinant plasmid DNA was recovered from the pellet of *E. coli* cells with a Minprep Kit (QIAGEN) following manufacturer's protocol. Size of insert was confirmed either with PCR (section 2.2.5) using universal M13 primers or by digesting the plasmid DNA with *EcoRI*, to release the cloned fragment. Both M13 and *EcoRI* sites are located near the multiple cloning site in pGEM[®]-T Easy vector (Figure 2.2). For restriction ~300ng of plasmid DNA (pGEM[®]-T Easy vector) was digested with *EcoRI* (New England BioLabs) in a final volume of 20 μ l, according to manufacturer guidelines in the presence of appropriate NEB buffer at 37°C for at least 2 hrs. Once the clone sizes were confirmed, 500 μ l of the overnight culture was mixed with 500 μ l of sterilized 50% glycerol in a 1.5ml Eppendorf tube and stored in -80°C freezer.

2.10.

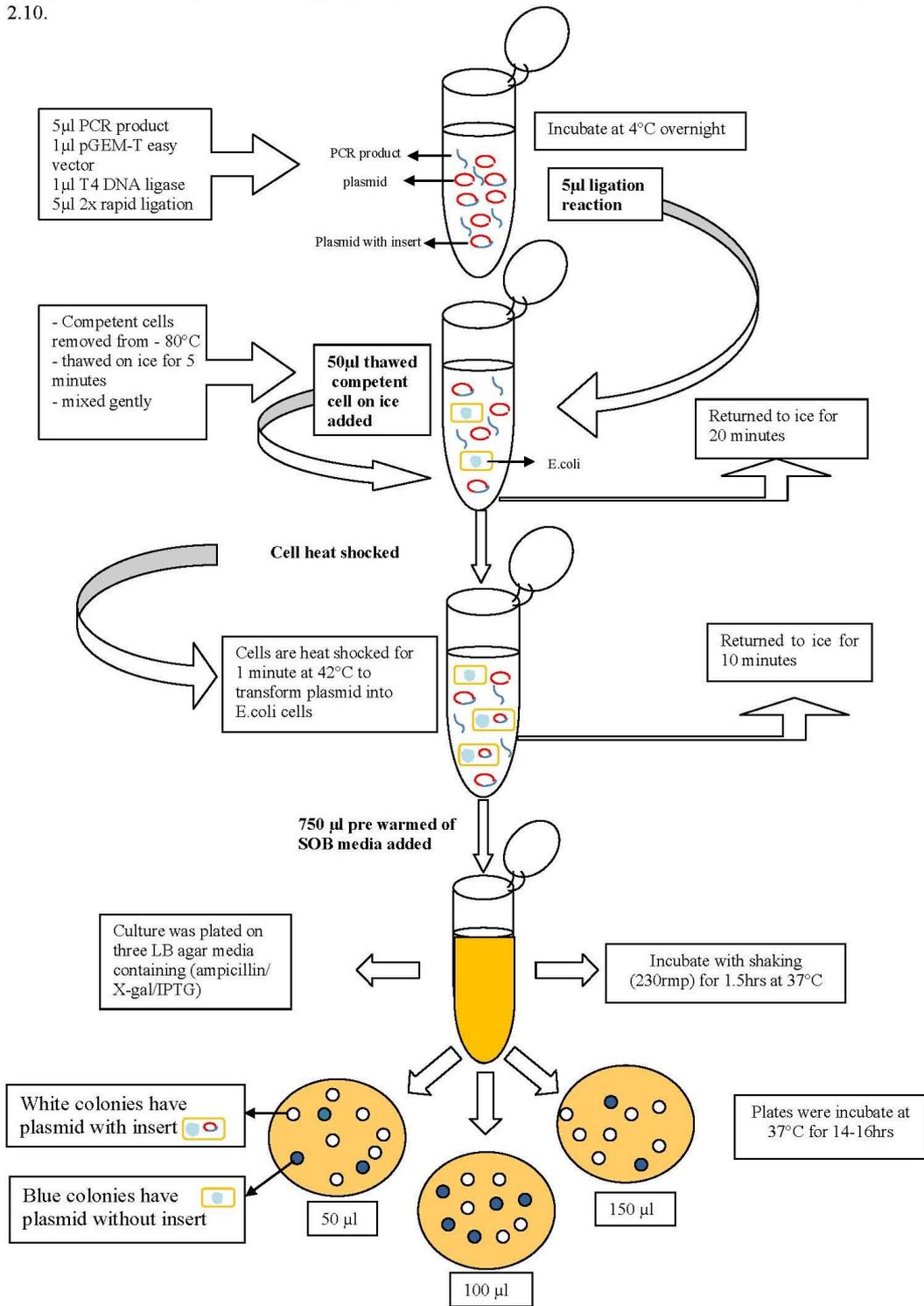


Figure 2.2: Flow diagram showing the standard cloning and colony selection procedure (Maniatis et al., 1982) modified in the Molecular Cytogenetics laboratory.

2.2.7 Sequencing of PCR amplicons and sequence analysis

Purified DNA fragments were sequenced commercially at Source Biosciences (Nottingham, UK) either by sending the PCR products directly using custom primers or with universal M13 forward or reverse primers, using recombinant plasmid DNA and following the company's guidelines for dilution and concentration etc.

DNA sequences in the form of chromatograms were obtained from Source Biosciences, and opened using bioinformatics software Geneious R6 (Kearse *et al.*, 2012). The DNA sequences were copied and saved in FASTA format, the pGEM[®]-T Easy vector sequences flanking the inserts were identified and deleted from the FASTA file. Multiple sequence alignment of the sequences was performed using ClustalW multiple alignment tools embedded in Geneious R6, and improved by eye when necessary. Phylogenetic reconstruction and estimation of nucleotide variability were carried out using GENEIOUS R6 or MEGA5 program (Tamura *et al.*, 2011). The evolutionary history was inferred with Neighbour joining method (NJ) method based on Tamura 3-parameter model (Tamura, 1992). Nodal support was assessed via bootstrapping, and the bootstrap consensus tree was inferred from 1000 replicates (Felsenstein, 2005) using default settings of the software applied (for details see result Chapter V).

2.2.8 Probes used

Probe used Included;

pTa71 contains a 9kb *EcoRI* fragment of the repeat unit of 25S-5.8S-18S rDNA isolated from *Triticum aestivum* (Gerlach and Bedbrook, 1979) and linearised with *EcoRI* before labelling (see section 2.2.6.3 above).

Total genomic DNA was sheared to 3-5kb pieces by autoclaving before labelling. Details of the genomic probe used are given in result (Chapter V, Table 5.2)

2.2.9 Probe labelling

For *in situ* hybridization both cloned and genomic DNA was labelled with biotin-16-dUTP and digoxigenin-11-dUTP (Roche Diagnostics) in separate reactions as described by Schwarzacher and Heslop-Harrison (2000).

2.2.9.1 M13-PCR labelling

Cloned repetitive DNA, such as pTa794, was labelled through PCR amplification using universal M13 primers, by adding 1µl of biotin-16-dUTP or digoxigenin-11-dUTP (1mM, Roche Diagnostics) or 1µl of water as a control to the standard PCR mixture and amplified as described above in (section 2.2.5).

2.2.9.2 Random primers labelling

Total genomic DNA, large clones such as pTa71, was labelled with BioPrime® Array CGH Labelling System (Cat. No. 18095-011, www.invitrogen.com). Labelling reactions were performed in a final volume of 50µl, following manufacturer's instruction with little modifications. Genomic DNA was sheared to 3-5kb pieces by autoclaving at 110°C for 4 min before labelling. The fragment sizes were estimated by running the autoclaved DNA on 1% agarose gel (section 2.1.5). Labelling was achieved with 200ng of the purified clone DNA (section 2.2.2.) or 1µg of sheared genomic DNA mixed with 20µl of 2.5x Random Primer Solution, denatured in boiling water for 5 min and then placed on ice for 5 min. To this mixture, 5µl of 10x dNTP Mix, 1µl of labelled nucleotids and 1µl 40U Exo- Klenow Fragment was added and incubated at 37°C for 2 hrs or kept at RT overnight. The polymerization reactions were stopped by adding 5µl of Stop Buffer (0.5M EDTA pH 8.0) to each tube. Labelled probes were purified to remove any unincorporated nucleotides, enzyme and salts using NucleoSpin® Extract II Kit (MACHERY-NAGEL), following manufacturer's instructions (<http://www.mn-net.com/tabid/1452/default.aspx>) and stored at -20°C freezer.

2.2.9.3 Testing of the incorporated labelled nucleotides in probes

Efficiency of the incorporated labelled nucleotide was estimated by a colorimetric dot blot reaction (Schwarzacher and Heslop-Harrison, 2000). Positively charged nylon membrane (Hybond-N⁺, Amersham Biosciences) of appropriate size was marked with pencil at the edge and soaked in buffer 1 (Table 2.2) at RT for 5 min, and dried between filter papers. Labelled probes (1µl), along with a positive control, were micro-pipetted on to the membrane, air-dried again for 5 min and then re-soaked in buffer 1, for 2 min. The membrane was incubated at RT for 30 min in buffer 2 (Table 2.2). Excess of buffer 2, was drained and the membrane was then incubated under a plastic cover slip at 37°C

for 30 min, with 0.75U/ml of conjugated antibody solution (anti-biotin-alkaline phosphatase and anti-digoxigenin-alkaline phosphatase, Roche Diagnostics) diluted 1:500 in buffer 1. During incubation the membrane was slowly agitated from time to time and then washed with buffer 1 for 15 min. The membrane was equilibrated in buffer 3 (Table 2.2) for 3 min and then detected with INT/BCIP (Roche Diagnostics). The stock solution of INT/BCIP [33mg/ml 2-(4-iodophenyl)-5-(4-nitrophenyl)-3-phenyltetrazolium chloride and 33mg/ml 5-bromo-4-chloro-3-indolyl-phosphate, toluidine-salt in DMSO] was diluted 1:500 in buffer 3 and the membrane detected at RT for 10-15 min in the dark. Labelled probe(s) appeared as a dark brown dot on the membrane due to the colorimetric reaction of the detection reagents. Labelling efficiency was judged by colour intensity in comparison to the control. The efficiency of unclean PCR-labelled probes was checked by agarose gel electrophoresis, 1µl of PCR product was loaded on a 1% agarose gel (section 2.2.2). Successful incorporation of label nucleotide was revealed by retardation of the same size band in biotin and digoxigenin incorporated probes compared with the unlabelled control reaction.

2.2.10 Chromosome preparations

2.2.10.1 Collection and fixation of root tips

Actively growing saffron roots of around 1-2cm length were collected from corms growing at the University of Leicester between 9:00AM-5:00PM. However, roots collected between 11-11:30 AM were the best with highest mitotic division, and put into the metaphase arresting agent, 2mM 8-hydroxyquinoline and α -Bromonaphthalene (Table 2.2) for 2-3 hrs at room temperature before transferring the roots to 24 hrs at 4°C. The root tips were partially blotted dry on filter paper before transferring into freshly prepared fixative consisting of 100% (v/v) ethanol and glacial acetic acid in 3:1 ratio and transferring the roots to 4°C for short term or -20°C for long term storage.

2.2.10.2 Metaphase chromosomes preparation

Fixed root tips were washed twice for 10 min in 1x enzyme buffer (Table 2.2) to get rid of the fixative and then digested at 37°C for 45 min, with 3% (w/v) pectinase (Sigma, 450U/ml), 1.8 % (w/v) cellulase (Calbiochem, 4000U/g) and 0.2 % (w/v) cellulase (Onozuka RS, 5000U/g). The digestion enzyme was not discarded and stored at -20°C for re-use. After digestion, root tips were washed in 1x enzyme buffer for 15 min.

Chromosomal preparations were made on clean glass slides (SuperFrost®, Menzel-Glaser, Thermo Scientific) under a stereo microscope as per Schwarzacher and Heslop-Harrison (2000). A single root tip was put in a drop of 60% glacial acetic acid, the root cap and other differentiated tissues were removed by using fine needles and forceps. The meristematic tissue was dissected, separated and then squashed under a No. 1, 18mm x 18mm cover slip by applying thumb pressure.

For meiotic chromosomes, buds of 7mm containing 2-3 florets were dissected from 11:00AM till 3:00PM. Anthers of 2.5-3.5 mm were checked with 45% acetic acid for appropriate stages and then fixed directly in absolute ethanol : glacial acetic acid (3:1) at RT for at least 4hrs. These anthers were processed on the same day or kept at -20 °C in a freezer. Chromosomal preparations were made from single anther by dissecting an anther and squeezing the content out with the help of a needle in 60% acetic acid and squashed as described above. The cover slips (from both mitotic and meiotic slides) were removed with a razor blade after freezing the slides on dry ice for 5-10 min. Slides were air dried at RT, scanned and then used for *in situ* hybridization or stored at -20°C in slide boxes together with silica gel for future use.

2.2.11 Fluorescent *in situ* hybridization

Fluorescent *in situ* hybridization (FISH) followed the protocol of Schwarzacher and Heslop-Harrison (2000) and is summarized in the following sub-headings.

2.2.11.1 Pre -hybridization

Slides were re-fixed in fresh absolute ethanol : glacial acetic acid (3:1) for 10-15 min and dehydrated with absolute ethanol twice for 5 min. Slides were allowed to air-dry and then incubated under a plastic cover slip with 200µl of RNase A (100µg/ml, Bioline) diluted in 2x SSC (Table 2.2) At 37°C for 1 hr. The slides were washed in 2x SSC at RT for 5 min and re-fixed in freshly prepared 4% (w/v) paraformaldehyde (Table 2.2) at RT for 10 min and then washed twice in 2x SSC for 5 min. Slides were then dehydrated in a series of 70%, 85% and absolute ethanol for 2 min. Slides were air dried and re-scanned before probing for the possible loss of cells that may occur during the above steps.

2.2.11.2 Hybridization

A total of 40µl probe mixture was applied per slide, containing 50% (v/v) formamide, 20% (w/v) dextran sulphate, 2x SSC, 25-100ng probe, 0.025µg/ul of salmon sperm DNA and 0.125% SDS as well as 0.125mM EDTA . In certain genomic *in situ* hybridization (GISH) experiments, autoclaved genomic DNA from *Crocus sativus* was added to the mixture as blocking DNA (Table 2.2). The hybridization mixture was denatured at 80°C for 10 min followed by immediate cooling on ice for 10 min. Probe and chromosomal DNA was denatured together on a Hybaid Omniblock at 72°C for 7 min under a plastic cover slip and slowly cooled to the hybridization temperature of 37°C for 16-20 hrs.

The formamide concentration, Na⁺ ions amount in SSC and temperature of the probe mixture determine stringency of hybridization. Unlabelled blocking DNA and salmon sperm DNA out-compete nonspecific hybridization or binding of probe, while dextran sulphate increases the volume without diluting the probe. SDS helps the penetration of probe and EDTA stops nucleases (Schwarzacher and Heslop-Harrison 2000). The above concentrations of formamide and salt at 37°C for hybridization and washing allowed sequences of 75-80% homology to form duplexes.

2.2.11.3 Post-hybridization washes and detection

Post-hybridization washes were carried out to remove the hybridization mixture and any unbound probe from the slides. Both high and low stringency washes were carried, depending upon the probes. A slightly higher stringency than the hybridization stringency was used to remove non-specific or weakly bound probes and to minimize background signal. Cover slips were floated off by incubating the slides in 2x SSC at 35-40°C. For genomic DNA probes two high stringency washes were carried out with 20% (v/v) formamide and 0.1x SSC at 42°C of 5 min each, an equivalent to 85% stringency (Schwarzacher and Heslop-Harrison, 2000). While, for repetitive DNA probes (microsatellites) low stringency washes with 0.1x SSC at 42°C of 5 min each, an equivalent to 73% stringency were carried out. Slides were then washed twice in 2x SSC at 42°C for 5 min, followed by cooling down to RT. Slides were incubated in detection buffer (Table 2.2) for 5 min and then blocked at 37°C for 30 min with 200µl of 5% (w/v) BSA (bovine serum albumin, Sigma) made in detection buffer. Hybridization sites were detected with 40-50µl of 2µg/ml streptavidin conjugated to Alexa594 (Molecular Probes) and 4µg/ml antidigoxigenin conjugated to FITC (flourescein isothiocyanate,

Roche Diagnostics) made up in 5% BSA solution. Slides were incubated at 37°C for 1hr in humid chamber, followed by two washes in detection buffer at 42°C for 10 min each.

2.2.11.4 Mounting of slides, photography and image processing

Chromosomes were counterstained with 100µl of 4µg/ml DAPI (Table 2.4) diluted in McIlvaine's buffer (Table 2.2) for 30 min in dark. The slides were then rinsed in detection buffer before final mounting in 80µl of antifade solution (Citifluor, Agar Scientific) under a No. 0, 24mm x 40mm coverslip. The slides were stored at 4°C overnight, to allow binding of the antifade solution to the fluorophores that stabilizes the fluorescence when viewed under the microscope. The processed slides were analyzed on Nikon ECLIPSE N80i fluorescent microscope (Nikon, Japan) with single band pass filters equipped with a DS-QiMc monochromatic camera (Nikon) and an X-Cite Series 120Q xenon lamp (Lumen Dynamics Group, Mississauga, ON, Canada). NIS-Elements BR3.1 software (Nikon) and Adobe Photoshop CS3 software were used for chromosomal analysis. Only those functions that treat all pixels of the image equally were used for colour balance, contrast/gain adjustment, and other processing without noise or other filters.

2.2.12 Reprobing of slides

FISH/GISH slides were re-probed as per Schwarzacher and Heslop-Harrison (2000) to see probes with different labels and label combinations sequentially on the same cell with little modification. Traces of immersion oil were carefully wiped from cover slips of selected slides. Slides were kept at 37°C for 10 min to reduce the viscosity of the antifade mount and coverslips were removed by lifting them with a razor blade. Slides were washed in 4x SSC containing 0.2% (v/v) Tween 20 at RT once for 5 min and then twice for 30-60 min, followed by two washes in 2x SSC at RT for 5 min. Preparations were denatured with 70% formamide 2x SSC at 70°C for 2 min and then dehydrated in an ice-cooled ethanol series of 70%, 85% and absolute for 2 min and air dried. Hybridization, washes and detection then followed the standard protocol from (section 2.2.11.2 onward).

2.2.13 Isolation of repeated DNA sequences

For characterization of novel repetitive DNA sequences in genus *Crocus*, about 15-30 µg good quality genomic DNA (section 2.2.2) was digested with *Hae*III, *Hind*III, *Bam*HI, *Sau*3A I, *Dra*I and *Eco*RI restriction enzymes (New England BioLabs) in the presence of appropriate buffers following manufacturer's instructions in a final volume of 20µl (see also Table 2.3). Although, all restriction enzymes could digest genomic DNA and given clear band(s) but *Hae*III and *Dra*I derived fragments were further analysed (see below). Digested DNA was loaded on 1-2% agarose gels and electrophoresis was carried out at a slow speed of 30V in 1x TAE buffer for 2-4hrs, and then visualized by staining with 0.5µg/ml of ethidium bromide as described above. Clear bands of high molecular weight were eluted and cloned into pGEM-Teasy vectors as described (see section 2.2.6 and Appendix 5,).

Table 2.3: Restriction endonucleases, source, restriction recognition sequence site, DNA cut end type and enzyme type.

Enzyme	Source	Recognition Sequence	Restriction recognition site in double-strand DNA	Cut DNA end	Endonuclease enzyme
<i>Eco</i> RI	<i>Escherichia coli</i>	5'GAATTC 3'CTTAAG	5'---G AATTC---3' 3'---CTTAA G---5'	Sticky	type II restriction enzyme
<i>Hae</i> III	<i>Haemophilus influenzae</i> biogroup <i>aegyptius</i>	5'GGCC 3'CCGG	5'---GG CC---3' 3'---CC GG---5'	Blunt	type II restriction enzyme
<i>Hind</i> III	<i>Haemophilus influenzae</i>	5'AAGCTT 3'TTCGAA	5'---A AGCTT---3' 3'---TTCGA A---5'	Sticky	type II restriction enzyme
<i>Bam</i> HI	<i>Bacillus amyloliquefaciens</i>	5'GGATCC 3'CCTAGG	5'---G GATCC---3' 3'---CCTAG G---5'	Sticky	type II restriction enzyme
<i>Sau</i> 3AI	<i>Staphylococcus aureus</i>	5'GATC 3'CTAG	5'--- GATC---3' 3'---CTAG ---5'	Sticky	type II restriction enzyme
<i>Dra</i> I	<i>Deinococcus radiophilus</i>	5' TTTAAA 3' AAATTT	5' ---TTT AAA--- 3' 3' ---AAA TTT--- 5'	Blunt	type II restriction enzyme

2.2.13.1 Selection of plasmid clones for dot blot hybridization

About 50-100 white colonies were chosen and replicates of LB agar plates were obtained by labelling Petri plates from below and inoculating them with the same colony at a respective position in both plates using a sterile tooth pick. Control blue colonies were also inoculated at several respective positions in both plates to check efficiency of the experiment. The bacterial colonies were allowed to grow at 37°C in an oven for 12-

16 hrs. The purpose of replicated plates was to use one plate for colony transfer and the second plate for selection of potential colonies for plasmid DNA isolation (Appendix 5).

2.2.13.2 Transfer of bacterial colonies onto charged nylon membrane

Positively charged nylon membrane (Hybond N⁺, Amersham Biosciences) of appropriate size (90mm Petri dish) was marked with a pencil at three asymmetric locations to identify the orientation of the membrane in the Petri dish. The membrane was then placed carefully upside down on the surface of LB-agar plate and ensured contact of the membrane with the bacterial colonies until the membrane became completely wet. Four pieces of 3MM Whatman filter papers of Petri dish size were cut and placed inside clean Petri dishes containing 5ml of 10% SDS, denaturing solution, neutralization solution and 2x SSC (Table 2.2) to soak the 3MM Whatman filter papers. Excess of solution was drained and the membrane was placed (colony side facing up) for 3 min in the Petri dish with 10% SDS, followed by 5 min into the denaturing solution, neutralization solution and 2x SSC respectively. The membrane was dried for 30 min between two sheets of 3MM Whatman filter papers at RT before it was heated at 80°C oven for 3 hrs and stored at 4°C overnight (Appendix 5).

2.2.13.3 Membrane hybridization

The membrane was brought to a working desk from 4°C and kept at RT for 10min before being re-hydrated with 2x SSC for 5 min and washed with 0.1x SSC, 0.1% (w/v) SDS for 1 min, before it was pre-hybridized at 55°C for 5 hrs in a Thermohybrid Hybridization oven (Ashford, UK) using 5ml of pre-hybridization solution containing 1x Denhardt's solution (Table 2.4), 5x SSC, 0.2% (w/v) SDS, 10mM EDTA and 100µg/ml of sheared Salmon sperm DNA (Table 2.4). 1ml of pre-hybridization solution was taken and mixed with 3-4µl (corresponding to ~150ng) of digoxigenin-labelled probe and freshly denatured salmon sperm DNA and replaced along with 1ml of 50% (w/v) dextran sulphate and then hybridized at 55°C for 16 hrs with constant rotation .

2.2.13.4 Post-hybridization washes and detection

High stringency washes were carried out by washing the membrane twice in 2x SSC x 0.1% (w/v) SDS at 56°C for 5 min (64% stringency) and then twice in 0.5x SSC x 0.1%

(w/v) SDS for 15 min each at 56°C (equivalent to 82% stringency). Detection was carried out at RT. Membranes were washed briefly for 5 min in 10ml of washing buffer 1 (Table 2.2), followed by 10ml of buffer 2 (Table 2.2) for 30 min and then incubated for 30 min with 10ml of antibody conjugate solution [anti-digoxigenin conjugated to alkaline phosphatase (Roche Diagnostics)] with a final dilution of 150U/ml (1:5000) in buffer 2. After antibody incubation, the membranes were washed twice for 15 min with 10ml of buffer 1 (Table 2.2), and then equilibrated for 5 min with buffer 3 (Table 2.2).

The membranes were finally incubated in the dark for 5 min in with 500µl of CDP-star solution (Roche Diagnostics) diluted 1:100 in buffer 3. The excess of CDP-Star solution was drained and then the membrane was wrapped in a cling film and transferred to an autoradiographic cassette in complete darkness. The chemiluminescence was recorded by keeping X-ray film (Fuji Medical X-Ray film) of appropriate size below the membrane. Different exposure times from 1-15 min, were given to detect all possible signals. X-ray films were developed using an automatic photographic developing machine and scanned with EPSON Expression Pro 1600, and images were processed with Adobe Photoshop CS3 (Appendix 5).

3 CHAPTER III: GENETIC VARIABILITY AND PHYLOGENY OF *CROCUS SATIVUS* L. (SAFFRON) BASED ON RETROELEMENT INSERTIONAL POLYMORPHISMS

3.1 Introduction

Complete genome sequencing of several important plant species such as rice, maize, sorghum, brassica and banana etc. has dramatically improved our understanding of the organization of angiosperm genomes. These studies have revealed relatedness of the gene order and content, and also the enormous instability at the level of repetitive DNA within plant genomes (Bennett and Leitch, 2011; Heslop-Harrison and Schwarzacher, 2011). The number of genes within angiosperms is typically 25,000 to 40,000 but the amounts of DNA present in plant genomes show over 2300 folds variation (Heslop-Harrison and Schmidt, 2012). The majority of the difference is associated with genome duplication or various classes of repetitive DNA, and that could possibly be the main evolutionary force responsible for biological diversification (Kubis *et al.*, 2003; Bennett and Leitch, 2011; Estep *et al.*, 2013). Therefore, understanding the role and nature of these repeated DNA elements are pivotal, particularly for investigating organizational and phylogenetic relationships as well as evolutionary dynamics of the genomes (Heslop-Harrison, 2000b; Kalendar *et al.*, 2011).

Repetitive DNA is broadly classified by the way it is organized structurally and in plants, a significant proportion of repetitive DNA comprises sequences of various motifs and lengths, which are tandemly organised and form long arrays extending from few to tens of kilobases (Schwarzacher, 2003a; Heslop-Harrison and Schwarzacher, 2011). These sequences concentrate at one or more distinct genomic locations and are referred to as satellite DNAs (Vershinin *et al.*, 1996; Schmidt, 1999; Contento *et al.*, 2005). Similarly, transposable elements (TEs) are also part of the repetitive DNA, but they have a more dispersed distribution and found throughout the genome (Kapitonov and Jurka, 2008; Sergeeva *et al.*, 2010). TEs are dynamic in nature and are capable of shuffling their locations within the genomes; they are divided into two main types based on their transposition intermediate (Kazazian, 2004; Salina *et al.*, 2011). Class I, also known as retrotransposons, replicate via reverse transcription of an RNA intermediate before integrating into the genome; Class II, or DNA transposons, transpose directly

from DNA to DNA, these elements excise from one region and reintegrate into another genomic location, following “cut and paste” mechanism (Finnegan, 1989; Wessler, 2006; Sergeeva *et al.*, 2010). Variability of the non-genic sequences that make the bulk of angiosperm nuclear DNA is primarily due to these TEs (Gaut and Ross Ibarra, 2008; Bennet and Leitch, 2011). Moreover, TEs have the ability to affect genome composition and functioning through genetic (directly changing the nucleotide order) as well as through epigenetic repatterning (Slotkin and Martienssen 2007; Fedoroff, 2012) and thus to what extent TEs may be regarded as “junk or selfish DNA” when compared with coding regions for their role in genome evolution needs further addressing (Charles *et al.*, 2008; Senerchia *et al.*, 2013).

Within plant genomes retrotransposons are perhaps the most ancient components and are present in virtually all eukaryotes (Brandes *et al.*, 1997; Kalendar *et al.*, 2011). They are either flanked by long-terminal repeats (LTR) at both their ends (LTR retrotransposon) or terminate at a poly-A tail (non-LTR retrotransposon) at their 3' end (Kumar and Bennetzen, 1999, Kazazian, 2004; Burgess, 2013). In addition to LTRs, these elements also contain regulatory sequences for *gag* and *pol* genes, whose protein products are essential for the formation of virus-like particles (VLPs) where replication of the elements take place (Syomin and Ilyin 2005; Begum *et al.*, 2013). *Gag* gene codes for a VLP structural protein while the *pol* gene codes for several enzymatic activities that include protease, reverse transcriptase, RNase H, and integrase. Products of these genes are required for making cDNA copy and then reintegration of this element into a new genomic site (Kazazian, 2004; Kalendar *et al.*, 2011). The common non-LTR retrotransposons include SINEs (short interspersed repetitive elements) and LINEs (long interspersed repetitive elements), where SINEs are non-autonomous and rely on LINEs machinery for their transposition (Kumar and Bennetzen, 1999; Dewannieux *et al.*, 2003; Ohshima, 2013).

The LTR-retrotransposons make up the bulk of angiosperms genomes. They contribute over 75% of the nuclear DNA even in a genome of moderate size such as maize (Schnable *et al.*, 2009) and show a direct correlation with genome size (Devos, 2010; Estep *et al.*, 2013). Although most LTR-retrotransposon families exist in low copy number, amplification of certain families may contribute individually >100Mb of DNA to a genome and it responsible for ‘genomic obesity’ in plants (see Estep *et al.*, 2013). Integration of LTR-retrotransposons typically produces a new insertion of 5-10kb cDNA (reviewed in Carvalho *et al.*, 2010) in which the LTRs, conserved within a

retrotransposon family, lie next to the anonymous host sequences (Saeidi *et al.*, 2008; Salina *et al.*, 2011). With time, the accumulation, fixation and incomplete excision of retrotransposon insertion cause genomic diversification (Wessler, 2006; Heitkam *et al.*, 2009). Furthermore, the ubiquitous nature, high copy number and widespread chromosomal distribution of retrotransposons make these elements ideal for the development of molecular markers (Flavell *et al.*, 1992; Teo *et al.*, 2005). Since the insertional polymorphisms due to retrotransposons allow one to infer the evolutionary history and phylogeny of species, they may be applied to establish pedigrees of lines as well as serving as biodiversity indicators (Saeidi *et al.*, 2008; D'Onofrio *et al.*, 2010).

In recent years, several molecular markers based on retrotransposons have been developed. However, among them inter-retrotransposon amplified polymorphism (IRAP) detects high levels of polymorphism and does not require DNA digestion, ligations or probe hybridization and are frequently used in genomic analysis (D'Onofrio *et al.*, 2010; Kalendar *et al.*, 2000, 2011). The IRAPs marker system exploits two basic principles within the LTR-retrotransposon *i.e.* the large insertions due to their transpositional activity and presence of the conserved domains from which PCR primers can be designed. The primers are generally designed within the LTRs near to the insertion site (see section 1.8.7 and Figure 1.11). The LTR sub-domains are conserved within a retrotransposon family, but differ between retrotransposon families (Teo *et al.*, 2005; Kalendar and Schulman 2006). Although regions internal to the LTR also contain conserved regions, that may be PCR amplified, but generally the LTRs are chosen to minimize the size of the target to be amplified and to analyze insertion site polymorphism rather than internal sequences of the element (Ellis *et al.*, 1998; Kalendar *et al.*, 2011).

In spite of the enormous socioeconomic impact of saffron on various cultures, there are potential threats to its survival and it is mostly due to its narrow genetic base, which may be explained by its sterile nature along shrinkage of the land surface assigned to saffron cultivation (see Chapter I for details). Moreover, polyploids with known ancestral species, offer a possibility of re-introducing genetic variation from the diploid progenitor species. However, in the case of saffron, the parental species are poorly defined and known genetic variability within saffron accessions itself is very limited, and this further undermines our abilities for genetic improvement (Caiola and Canini 2010; Fernandez *et al.*, 2011 and Chapter I).

The IRAPs method has been successfully applied to genome mapping, diversity and phylogeny analysis in fungi, wheat, cereals, banana, grapevines as well as saffron (Flavell *et al.*, 1998; Kalendar *et al.*, 1999, 2000; Vicient *et al.*, 2001; Teo *et al.*, 2005; Nair *et al.*, 2005; Kalendar and Schulman, 2006; Alavi-Kia *et al.*, 2008; Saeidi *et al.*, 2008; Carvalho *et al.*, 2010; D'Onofrio *et al.*, 2010; Mandoulakani *et al.*, 2012; Santana *et al.*, 2013). Therefore, in this chapter, IRAPs makers have been applied to various species and accessions of the genus *Crocus* including saffron, to (1) detect diversity within saffron accessions and between *Crocus* species (2) identify closely related species to saffron. The results shown here will be particularly helpful in identifying diversity and tracing potential ancestry of saffron. Further, the current work in the long run will be helpful in re-synthesizing saffron from its potential diploid ancestors, and thus novel useful diversity may be introduced into saffron germplasm.

3.2 Material and methods

3.2.1 Plant material and genomic DNA extraction

Thirty nine accession of saffron and other sexual *Crocus* listed in (Table 2.1), except *C.korolkowii*, *C. veneris*, *C. cancellatus* and *C. biflorus* were not used in the current study. Corms of all *Crocus* species were grown in glasshouse at the University Of Leicester, UK. Total genomic DNA was extracted from the young leaves and floral buds all plants, using standard CTAB technique. Details about plant growing conditions and DNA extraction procedure are given above in material and methods (Chapter 2).

3.2.2 IRAP markers

A total of 11 IRAP primers (also referred to as IRAP markers) previously designed to the conserved LTR regions of retrotransposons were applied in the current study. Nucleotide sequences of the IRAP markers, GenBank accession number, position and orientation along the original sources are given below (Table 3.1). IRAP primers were tested alone as well as in all possible 66 combinations.

3.2.3 PCR amplification and gel electrophoresis

PCR reaction mixtures, amplification conditions and gel electrophoresis were as described in materials and methods (Chapter II section 2.2.5). All primers except three are designed from non-*Crocus* species and therefore, melting temperatures given here varied from the published sources. IRAP markers alone and in combinations successfully amplified multiple loci from saffron and related species. Optimum melting temperatures for successful amplification are given in (Table 3.2). The reproducibility of amplified fragments was confirmed by repeating all reactions twice.

3.2.4 Genetic variability and phylogenetic analysis

A presence/absence analysis of clear and distinguishable IRAP fragments was performed for all markers. Presence of a DNA band was considered (1) and absence as (0). Gels were scored manually from gels images opened in Adobe Photoshop CS3 and binary matrices were assembled as Excel sheets. Basic statistics including the total number of alleles, major allele frequency, genetic diversity and polymorphism information content (PIC) values were determined by categorizing the data set into two levels as accessions, and species using PowerMarker version 3.25 (Liu and Muse, 2005). PIC values provide an estimate of the discriminating power of a locus by taking into account the number of alleles generated by each reaction unit and their frequency distribution in the population. PIC values range between 0 to 1, where 0 indicating to minimum and 1 to maximum diversity and the values above 0.5 are considered as useful and informative. The allele molecular weight data was also used to determine the phylogenetic relationship of the *Crocus* species based on the Neighbor Joining (NJ) method (Saitou and Nei, 1987) with 1000 bootstrap replicates PowerMarker and the tree viewed using TreeView (Page 1996). The consensus 1000 bootstrap tree was generated using PHYLIP (Phylogeny Inference Package) version 3.69 programs (Felsenstein, 2005).

Table 3.1: List of IRAP primers, orientation, sequence, GenBank accession number, source and crops they are applied.

No.	Marker name	Retrotransposon name and orientation	Sequence (5'-3')	Accession	Position	Reference/source	Applied/tested in
1	LTR6150	<i>BARE-1</i> ←	CTGGTTCGCCCCATGTCTATGTATCCACACATGGTA	Z17327	418-439	Kalendar <i>et al.</i> , 1999	Barley, Banana, goat grass, wheat, saffron
2	LTR6149	<i>BARE-1</i> →	CTCGCTCGCCCACTACATCAACCGCGTTTATT	Z17327	1993-2012	Kalendar <i>et al.</i> , 1999 (in Barley)	Barley, Banana, goat grass, wheat
3	Nikita	Nikita →	CGCATTTGTTCAAGCCTAAACC	AY078073 AY078074 AY078075	1-22	Leigh <i>et al.</i> , 2003	Barley, Banana, goat grass, wheat, saffron
4	IRAP Crocus Nikita	Nikita	CAGTTTTGATCAAGTCATAACC	AJ131448	15-36 of pCvKB4	Modified after Leigh <i>et al.</i> (2003) by Heslop-Harrison, Vikgren and Ørgaard (unpublished)	<i>Crocus</i>
5	Sukkula	Sukkula →	GATAGGGTCGCATCTTGGGCGTGAC	AY034376	10662-10685	Mannien <i>et al.</i> , 2000	Banana, goat grass, wheat, saffron
6	IRAP Crocus Sukkula	Sukkula	AACAGAAGTAGTGGCAGCTTGAGAG	AY245374	1023	Modified after Leigh <i>et al.</i> (2003) by Heslop-Harrison, Vikgren and Ørgaard (unpublished)	<i>Crocus</i>
7	ReverseTy1	W1, W3, W7, W8 ←	CCYTGNAYYAANGCNGT	AF416815 AF416816 AF416817 AF416818	1-17	Teo <i>et al.</i> , 2005	Banana, goat grass
8	Reverse TY2	W1, W3, W7, W8 →	TRGTARAGRAGNTGRAT	AF416815 AF416816 AF416817 AF416818	252-269	Teo <i>et al.</i> , 2005	Banana
9	3' LTR	<i>BARE-1</i> →	TGTTTCCCATGCGACGTTCCCAACA	Z17327	2112-2138	Teo <i>et al.</i> , 2005	Banana, wheat, goat grass
10	IRAP Crocus 5'LTR		CCATAGCTTGTAGGGCGTCTCCCA	AY245373	5100	Modified after Leigh <i>et al.</i> (2003) by Heslop-Harrison, Vikgren and Ørgaard (unpublished)	<i>Crocus</i>
11	5' LTR1	<i>BARE-1</i> ←	TTGCCTCTAGGCATATTCCAACA	Z17327	1-26	Teo <i>et al.</i> , 2005	Banana, wheat, goat grass

→
←
Arrows indicate primer direction with respect to the first open reading frame of each retrotransposon;
Y= C + T, N = A + G + C + T, R = A + G nucleotides

3.3 Results

3.3.1 IRAP amplification and diversity within *Crocus* species

Out of 66 IRAP primers combinations tested, 63 allowed successful amplification of multiple and distinguishable fragments from the genomic DNA of all *Crocus* species and accessions (Table 3.2). The overall analysis included 40 assays and resulted in amplification of a total of 4745 IRAP fragments (bands); all of them were polymorphic and were absent in one or more accessions (Table 3.3) and the average Major Allele Frequency was 0.87. On average, the LTR primer alone produced 75 distinct bands, while the two primers combination resulted in 76 bands. Among the primers, Sukkula and Nikita produced consistently the largest number of clearly separated DNA fragments. The Sukkula and Nikita primer combination yielded the maximum number of 113 bands, while 5'LTR1 and RTY1 primer combination produced the minimum 40 bands. The Sukkula and Nikita primer combination produced the maximum number of polymorphic bands; while 5'LTR1 and RTY1 combination was least polymorphic and produced only 40 bands from the *Crocus* species (see Table 3.3). The polymorphic information content (PIC) values that reflect allele diversity and frequency among the *Crocus* species used, varied from ~0.4 to ~0.05 with an average PIC value of 0.17 ± 0.1 (see section 3.2.4 and Table 3.3). Similarly, the genetic diversity for each IRAP marker associated with the number of alleles amplified. The higher the PIC value for a locus, the higher the number of alleles detected (Table 3.3). These results are consistent with the previous reports obtained for banana, rice and wheat (Teo *et al.*, 2005; Lapitan *et al.*, 2007; Saeidi *et al.*, 2008).

Although, only eight representative gel images are given below (Figures 3.1-3.3), four for the 12 *Crocus* species (17 accessions) representing 9 series of both section *Crocus* and *Nudiscapus* (see Figure 1.9), and four for the 24 *Crocus* accessions (9 species, including *C. sativus*) confined to the *Crocus* series *Crocus* only (Figures 3.1-3.3). All gel images have been treated equally for analysis, and a summary of all gel images is given (Table 3.3) and used for understanding the phylogenetic relationships of the genus *Crocus* (Figures 3.5, 3.6).

Reproducible IRAP banding patterns were achieved for *Crocus* species and accessions. IRAP bands ranging in size from 100bp to ~4kb were obtained (Figure 3.1 and Table 3.3). Several of the IRAP amplified fragments were shared among saffron and

other sexually reproducing diploid *Crocus* species and indicated conservation of the organization of retroelement insertion among members of the genus (arrow in Figures 3.1-3.3). Further, the relatedness of *C. sativus* with other members of the *Crocus* series *Crocus* was evident, and a significant number of the amplified bands were shared among the series (for example see arrow in Figure 3.1C, 3.2B). High levels of polymorphism between species (Figures 3.1-3.3) as well as within accessions of the sexually reproducing *Crocus* accessions (see *C. tommasinianus* in Figure 3.1B) were evident. In a few cases IRAP markers alone as well as in combination with other IRAP primers produced relatively fewer but strong PCR bands, amplified from almost all accessions and species (arrow in Figure 3.2B). Furthermore, the banding pattern of section *Crocus* was different from that of the section *Nudiscapus*. Further, Nikita and IRAP *Crocus* Sukkula primers combination resulted in maximum bands from the section *Crocus* (Figure 3.1B) while LTR6150 from the section *Nudiscapus* (Figure 3.1C). IRAPs could also produce unique bands limited to species other than *C. sativus* (arrow head in Figure 3.2B). By and large, the PCR profile of *C. sativus* was different from all species applied. However, informal observation revealed greatest similarity to *C. pallasii* subsp. *pallasii* (compare banding in Figure 3.2 and 3.3). Similarly, “*C. sativus cartwrightianus*”, purchased under this unrecognized name, had a banding pattern more similar to *C. cartwrightianus albus* than any other species used here (compare symbol # in Figure 3.2, 3.3). Furthermore, it was remarkable to see most of these LTR-markers are designed within non-*Crocus* species (Table 3.1) still all markers alone as well as in combination could successfully produce multiple loci from *Crocus* species, and indicating to the transferable nature of the retrotransposon-based marker within angiosperms (Table 3.3).

3.3.2 IRAP amplification and diversity within saffron accessions

To estimate the exact level of diversity within *C. sativus* (saffron), 17 saffron accessions originating from different geographical areas were subjected to IRAP analysis (Table 3.1). As *C. sativus* was applied in both analyses that included 12 *Crocus* species as well as 9 species (see above and Figures 3.1-3.3), therefore, IRAP primers in combination were not used here. All the 11 IRAP markers generated multiple bands of distinguishable sizes from the genomic DNA of all saffron accessions; four (04) representative gel images are given below for the 17 saffron accessions (Figure 3.4). The range of IRAP bands varied in size from 100bp to ~4kb. Further, all bands obtained with

11 IRAP markers were monomorphic and no polymorphism within the 17 accessions of saffron could be confirmed (Figure 3.4). By comparing the banding patterns obtained with Sukkula, it could be seen that a prominent ~1800bp IRAP fragment is missing in a few saffron accessions (see * in Figure 3.4C). However, later the same accessions were re-PCR amplified and an identical IRAP banding pattern was generated. Thus the missing bands were most probably due to adding a less than optimum amount of template DNA during the initial round of PCR. Moreover, high levels of IRAP polymorphism were evident within and among the sexually reproducing *Crocus* species (Figures 3.1, 3.2). Along with 17 saffron accessions, '*C. sativus cartwrightianus*' was used and its amplification pattern was different from typical saffron *C. sativus* (lane 18 Figure 3.4). This accession has morphological similarities to *C. cartwrightianus* cv. *albus* (Figure 1.2A3 and A4). And the IRAP profile for *C. sativus cartwrightianus* based on IRAP *Crocus* Nikita primer was very similar to that of *C. cartwrightianus* cv. *albus* (Figure 3.3D). The overall results indicate that IRAPs constitute a suitable marker system for the detection of genetic variability among the *Crocus* species.

3.3.3 Genome diversity and phylogenetic relationships among *Crocus* species

In order to evaluate the genetic relationships among members of the genus *Crocus*, the binary data (0, 1) obtained from IRAP fragments was pooled for the construction of phylogeny using Neighbor Joining (NJ) method in PowerMarker (Saitou and Nei, 1987). The analysis involved 20 *Crocus* species and a total of 39 accessions (Table 2.1). The phylogenetic analysis of IRAP polymorphisms separated the 20 different *Crocus* species into groups and sub-groups (also referred to as clusters or clades below Figure 3.5). Further, multiple accessions of *Crocus* species were also used for phylogenetic reconstruction, and irrespective of the species, most accessions clustered together (see Figure 3.6). Most groups had strong nodal support as indicated by the high bootstrap support values, often close to 100%, and thus IRAPs provided good discrimination both at the species as well as at the accession level.

Based on the IRAP tree, the 20 *Crocus* species were resolved both section *Crocus* and *Nudiscapus* into three main clades A, B and C (Figure 3.5), although clade A and B may be merged together into one clade. The clade C was further divided into two sub-clades C1 and C2. Clade A consisted of three diploid *Crocus* species and included *C. oreoreticus*, *C. hadriaticus* and *C. cartwrightianus* radiating out as single

branches from the tree (see clade A Figure 3.5). While, clade B comprised six species, including the diploid *C. asumaniae*, *C. thomasi*, *C. mathewii*, *C. pallasii* as well as the triploid *C. sativus*. Similarly, “*C. sativus cartwrightianus*”, is an unrecognized species for which ploidy level is not clear but based on the preliminary cytogenetic investigation carried out here, it is most likely to be a diploid species and is grouped with *C. sativus* (clade B Figure 3.5). Clade C is constituted by 11 out of the 20 *Crocus* species and contained representative species of both section *Crocus* and *Nudiscapus*. This clade consisted of *C. vernus*, *C. tommasinianus*, *C. kotschyanus*, *C. versicolor*, *C. goulimyi*, *C. niveus*, *C. speciosus*, *C. angustifolius*, *C. flavus*, *C. laevigatus* and *C. boryi*. Notably, *C. sativus* (saffron) is placed in between the recognized diploid species *C. pallasii* and “*C. sativus cartwrightianus*”, while *C. mathewii* (series *Crocus*) is the second closest species of *C. pallasii* (Figure 3.5).

For convenience, the IRAP tree that included all 39 accessions of the 20 species was divided into three main clades D, E and F (Figure 3.6), although clades E and F may be united into one mega clade each. Clade D has 94% nodal support included 23 accessions belonging to 9 *Crocus* species while one accession of *C. cartwrightianus* (CcwBD09) radiating out on a separate branch. This clade is divided into two sub-clades, D1, and D2 and included accessions of *C. asumaniae*, *C. cartwrightianus*, *C. oreoreticus*, *C. hadriaticus*, *C. thomasii*, *C. mathewii*, *C. pallasii*, *C. sativus* and *C. sativus cartwrightianus*. Notably, the clade D2 (with 99% nodal support) is consisting of *C. pallasii* subsp. *pallasii* clustering with 5 accessions of *C. sativus* and the unrecognized species ‘*C. sativus cartwrightianus*’ (sub-clade D2 Figure 3.6). Similarly, clade E with 50% nodal support comprises of 10 *Crocus* accessions of 6 species and is constituted by *C. vernus*, *C. tommasinianus*, *C. goulimyi*, *C. versicolor*, *C. niveus* and *C. kotschyanus*. The *C. vernus* and *C. tommasinianus* from *Crocus* series *Verni* and have 100% nodal support. *C. goulimyi*, *C. versicolor* and *C. niveus* from related series and have 58% nodal support, while two accessions of *C. kotschyanus* lay on a sister branch with 99% nodal support (clade E Figure 3.6). Further, clade F comprised of 5 *Crocus* species with 57% nodal support. Only one accession per each species was used and this clade consisted of *C. speciosus*, *C. angustifolius*, *C. flavus*, *C. laevigatus* and *C. boryi* from section *Nudiscapus* (clade F Figure 3.6).

The IRAPs spliced all of the used species and accessions into their respective series. All accessions of the clade D and E exclusively belonged to section *Crocus*, while accessions in clade F are confined to section *Nudiscapus* only. These results

indicate the utility of the IRAP marker system for discriminating taxa and its potential role in analysing phylogeny (Figure 3.6).

By and large, the tree topology for both species and accessions (Figures 3.5, 3.6) is in accordance to that to Mathew (1982) and Petersen et al. (2008) and the position of most (but not all) species is satisfied even at the series level (see below and discussion). Further, the order and clustering of species and accessions in both trees obtained for the 20 species and 39 accessions is very much identical (compare Figures 3.5 and 3.6). At section level there is no discrepancy and species belonging to one series clustered together clades D, E, F (Figure 3.6).

Although, few accessions for example *C. hadriaticus Alepohori* (ChdARD09) clustered with *C. oreocreticus* instead of grouping with other *C. hadriaticus* accessions. *C. cartwrightianus* (CcwBD09) remained separate from other *C. cartwrightianus* accessions, and one out of the three *C. asumaniae* accessions ('white' or CasWD09) did not clustered with *C. asumaniae* accessions, 'alba' and 'S9104' (see sub-clade D1). Both accessions of *C. asumaniae* 'white' and 'alba' were obtained from the Netherlands, while 'S9104' originated from Aseki, Turkey. Similarly, *C. cartwrightianus* (CcwBD09) also came from Netherlands and the other *C. cartwrightianus* accessions were obtained from Rare plants. These accessions are maintained in nurseries but the original area of collection for all accession is not known and this variation in accession may be related to different geographical origin (see discussion).

Table 3.2: IRAP primer combinations with optimum annealing temperatures. ∅ indicates unsuccessful primer combinations for amplification

Primer Name	IRAP <i>Crocus</i> Sukkula	Sukkula	IRAP <i>Crocus</i> Nikita	Nikita	LTR6149	LTR6150	3'LTR	5' LTR1	IRAP <i>Crocus</i> 5'LTR	RTY1	RTY2
IRAP <i>Crocus</i> Sukkula	52°C										
Sukkula	60 °C	62°C									
IRAP <i>Crocus</i> Nikita	54°C	58°C	46 °C								
Nikita	54°C	60°C	48 °C	50 °C							
LTR6149	50°C	60°C	46°C	45°C	40°C						
LTR6150	50°C	60°C	50°C	48°C	45°C	40°C					
3'LTR	54°C	58°C	52°C	54°C	50°C	48°C	50°C				
5' LTR1	52°C	∅	46 °C	45°C	42°C	48.5°C	43°C	42°C			
IRAP <i>Crocus</i> 5'LTR	56°C	58°C	56°C	54°C	54°C	54°C	58°C	56 °C	58°C		
RTY1	52°C	56°C	50°C	48°C	∅	45°C	48°C	45°C	58°C	48°C	
RTY2	54°C	62 °C	50°C	52°C	42°C	42°C	56°C	∅	58°C	48°C	50°C

Table 3.3: Primer combinations, total and polymorphic number of bands, percentage of IRAP polymorphism, band size range and PIC values.

No.	Primer combination	Total bands	Polymorphic bands	Degree of polymorphism (%)	Band size	Maximum band number		Minimum band number		PIC			
										Max	Min.	Avg.	Stdv.
1	Sukkula	110	110	100	150 - 4000	At 390	23	Several accession	1	0.4	0.05	0.2	0.1
2	Sukkula + IRAP <i>Crocus</i> Sukkula	78	78	100	100 - 1800	At 290	27	Several accession	1	0.4	0.05	0.2	0.1
3	Sukkula + Nikita	113	113	100	100 - 4000	At 600	24	Several accession	1	0.4	0.05	0.2	0.1
4	Sukkula + IRAP <i>Crocus</i> Nikita	74	74	100	100 - 2200	At 1000	15	Several accession	1	0.4	0.05	0.2	0.1
5	Sukkula + LTR6149	58	58	100	200 - 4000	At 1200	21	Several accession	1	0.4	0.05	0.2	0.1
6	Sukkula + LTR6150	48	48	100	200 - 3000	At 650, 1500	14	Several accession	1	0.4	0.05	0.2	0.1
7	Sukkula + 3'LTR	72	72	100	100 - 1800	At 230	20	Several accession	1	0.4	0.05	0.2	0.1
8	Sukkula + IRAP <i>Crocus</i> 5'LTR	53	53	100	100 - 2000	Several sizes	10	Several accession	1	0.3	0.05	0.2	0.1
9	Sukkula + 5'LTR1	Nil	Nil	Nil	Nil	Nil		Nil					
10	Sukkula + Reverse TY1	87	87	100	100 - 3300	At 1200	20	Several accession	1	0.4	0.05	0.1	0.1
11	Sukkula + Reverse TY2	95	95	100	130-3000	At 1050	19	Several accession	1	0.4	0.05	0.2	0.1
12	Nikita	61	61	100	110 - 3500	At 200	24	Several accession	1	0.4	0.05	0.2	0.1
13	Nikita+ IRAP <i>Crocus</i> Sukkula	84	84	100	100 - 2500	At 520	34	Several accession	1	0.4	0.05	0.2	0.1
14	Nikita + IRAP <i>Crocus</i> Nikita	71	71	100	100 - 4500	Several sizes	10	Several accession	1	0.3	0.05	0.2	0.1
15	Nikita + LTR6149	89	89	100	100 - 3000	At 2500	20	Several accession	1	0.4	0.05	0.2	0.1
16	Nikita+ LTR6150	61	61	100	110 - 3000	At 790	17	Several accession	1	0.4	0.05	0.2	0.1
17	Nikita + 3'LTR	77	77	100	100 - 3200	At 280, 290	16	Several accession	1	0.4	0.05	0.2	0.1
18	Nikita+ IRAP <i>Crocus</i> 5'LTR	80	80	100	110 - 2400	At 190-320	21	Several accession	1	0.4	0.05	0.1	0.1
19	Nikita+ 5'LTR1	87	87	100	100 - 4000	Several sizes	15	Several accession	1	0.4	0.05	0.2	0.1
20	Nikita + Reverse TY1	64	64	100	120 - 2200	At 1250	16	Several accession	1	0.4	0.05	0.1	0.1
21	Nikita+ Reverse TY2	47	47	100	100 - 1700	At 650	9	Several accession	1	0.3	0.05	0.1	0.1
22	IRAP <i>Crocus</i> Sukkula	110	110	100	100 - 3800	At 1600	25	Several accession	1	0.4	0.05	0.2	0.1
23	IRAP <i>Crocus</i> Sukkula + IRAP <i>Crocus</i> Nikita	104	104	100	120 - 3200	At 800	16	Several accession	1	0.4	0.05	0.2	0.1
24	IRAP <i>Crocus</i> Sukkula + LTR6149	77	77	100	100 - 4200	At 100	21	Several accession	1	0.4	0.05	0.2	0.1
25	IRAP <i>Crocus</i> Sukkula+ LTR6150	72	72	100	100 - 3000	At 840	17	Several accession	1	0.4	0.05	0.2	0.1
26	IRAP <i>Crocus</i> Sukkula + 3'LTR	81	81	100	100 - 2000	At 240, 400	17	Several accession	1	0.4	0.05	0.2	0.1
27	IRAP <i>Crocus</i> Sukkula + IRAP <i>Crocus</i> 5'LTR	77	77	100	100 - 2400	At 190	17	Several accession	1	0.4	0.05	0.2	0.1
28	IRAP <i>Crocus</i> Sukkula + 5'LTR1	84	84	100	100 - 3000	At 700	15	Several accession	1	0.4	0.05	0.2	0.1
29	IRAP <i>Crocus</i> Sukkula + Reverses TY1	91	91	100	100 - 3100	At 1200	27	Several accession	1	0.4	0.05	0.2	0.1
30	IRAP <i>Crocus</i> Sukkula + Reverses TY2	99	99	100	120 - 4000	At 240	17	Several accession	1	0.4	0.05	0.2	0.1
31	IRAP <i>Crocus</i> Nikita	84	84	100	150 - 4000	At 1200	20	Several accession	1	0.4	0.05	0.2	0.1
32	IRAP <i>Crocus</i> Nikita + LTR6149	71	71	100	100 - 3000	At 120	18	Several accession	1	0.4	0.05	0.2	0.1
33	IRAP <i>Crocus</i> Nikita + LTR6150	71	71	100	100 - 4000	At 100	22	Several accession	1	0.4	0.05	0.2	0.1
34	IRAP <i>Crocus</i> Nikita + 3'LTR	75	75	100	100 - 2900	At 230	21	Several accession	1	0.4	0.05	0.2	0.1

Table 3.3: continued

No.	Primer combination	Total bands	Polymorphic bands	Degree of polymorphism (%)	Band size	Maximum number of accessions sharing one band (size indicated)		Minimum band number		PIC			
										Max.	Min.	Avg.	Stdv.
35	IRAP <i>Crocus</i> Nikita + IRAP <i>Crocus</i> 5'LTR	78	78	100	120 - 3100	At 200	22	Several accession	1	0.4	0.05	0.2	0.1
36	IRAP <i>Crocus</i> Nikita + 5'LTR1	81	81	100	110 -2700	At 280	21	Several accession	1	0.4	0.05	0.1	0.1
37	IRAP <i>Crocus</i> Nikita + Reverse TY1	61	61	100	100 - 3000	At 1500	20	Several accession	1	0.4	0.05	0.2	0.1
38	IRAP <i>Crocus</i> Nikita + Reverse TY2	101	101	100	100 - 4100	At 590, 1700	22	Several accession	1	0.4	0.05	0.2	0.1
39	LTR6149	57	57	100	300 - 3300	At 1300	13	Several accession	1	0.3	0.05	0.2	0.1
40	LTR6149 + LTR6150	59	59	100	120 - 4000	Several sizes	15	Several accession	1	0.4	0.05	0.1	0.1
41	LTR6149 + 3'LTR	72	72	100	100 - 2800	At 200	25	Several accession	1	0.4	0.05	0.2	0.1
42	LTR6149 + IRAP <i>Crocus</i> 5'LTR	72	72	100	140 - 2700	At 700	17	Several accession	1	0.4	0.05	0.2	0.1
43	LTR6149 + 5'LTR1	80	80	100	160 - 2700	At 1950	17	Several accession	1	0.4	0.05	0.2	0.1
44	LTR6149 + Reverse TY1	Nil	Nil	Nil	Nil								
45	LTR6149 + Reverse TY2	57	57	100	120 - 2500	At 900	20	Several accession	1	0.4	0.05	0.1	0.1
46	LTR6150	67	67	100	100 - 3500	At 600	24	Several accession	1	0.4	0.05	0.2	0.1
47	LTR6150 + 3'LTR	70	70	100	150 - 2750	At 700, 2750	16	Several accession	1	0.4	0.05	0.2	0.1
48	LTR6150 + IRAP <i>Crocus</i> 5'LTR	70	70	100	100 - 2400	At 300	19	Several accession	1	0.4	0.05	0.2	0.1
49	LTR6150 + 5'LTR1	70	70	100	100 - 3000	At 1000	26	Several accession	1	0.4	0.05	0.1	0.1
50	LTR6150 + Reverse TY1	56	56	100	120 - 2500	At 600	17	Several accession	1	0.4	0.05	0.2	0.1
51	LTR6150 + Reverse TY2	66	66	100	110 -2000	At 720	22	Several accession	1	0.4	0.05	0.2	0.1
52	3'LTR	51	51	100	180 - 4000	At220	25	Several accession	1	0.4	0.05	0.2	0.1
53	3'LTR + IRAP <i>Crocus</i> 5'LTR	76	76	100	110 -2100	At 330	16	Several accession	1	0.4	0.05	0.2	0.1
54	3'LTR + 5'LTR1	69	69	100	110 - 2100	At 800	32	Several accession	1	0.4	0.05	0.2	0.1
55	3'LTR + Reverse TY1	70	70	100	100 - 2900	At 570	32	Several accession	1	0.4	0.05	0.2	0.1
56	3'LTR + Reverse TY2	68	68	100	110-2700	At 220	18	Several accession	1	0.4	0.05	0.2	0.1
57	IRAP <i>Crocus</i> 5'LTR	73	73	100	100 - 3800	At 650	19	Several accession	1	0.4	0.05	0.2	0.1
58	IRAP <i>Crocus</i> 5'LTR + 5'LTR1	84	84	100	100 - 2900	At 250	17	Several accession	1	0.4	0.05	0.2	0.1
59	IRAP <i>Crocus</i> 5'LTR + Reverse TY1	79	79	100	120 - 4500	At 250	17	Several accession	1	0.4	0.05	0.2	0.1
60	IRAP <i>Crocus</i> 5'LTR + Reverse TY2	82	82	100	120 - 3600	At 120	24	Several accession	1	0.4	0.05	0.2	0.1
61	5'LTR1	92	92	100	120 - 4000	At780	19	Several accession	1	0.4	0.05	0.2	0.1
62	5'LTR1 + Reverse TY1	40	40	100	180 - 2400	At 810	20	Several accession	1	0.4	0.05	0.1	0.1
63	5'LTR1 + Reverse TY2	Nil	Nil	Nil	Nil								
64	Reverse TY1	44	44	100	100 - 2700	At 1000	29	Several accession	1	0.4	0.05	0.2	0.1
65	Reverse TY1 + Reverse TY2	108	108	100	100 - 4200	At 410	29	Several accession	1	0.4	0.05	0.2	0.1
66	Reverse TY2	88	88	100	120 -2400	At 800	22	Several accession	1	0.4	0.05	0.2	0.1
Total number of bands		4745											
Maximum number of bands		113											
Minimum number of bands		40											

Nil: indicates failure of IRAP primer to produce PCR amplification; Reverse TY1= RY1; Reverse TY2= RY2; several accessions: refers to the presence of an observation (DNA band) in more than one accession
 Values given under PIC (polymorphic information content) are after rounding off

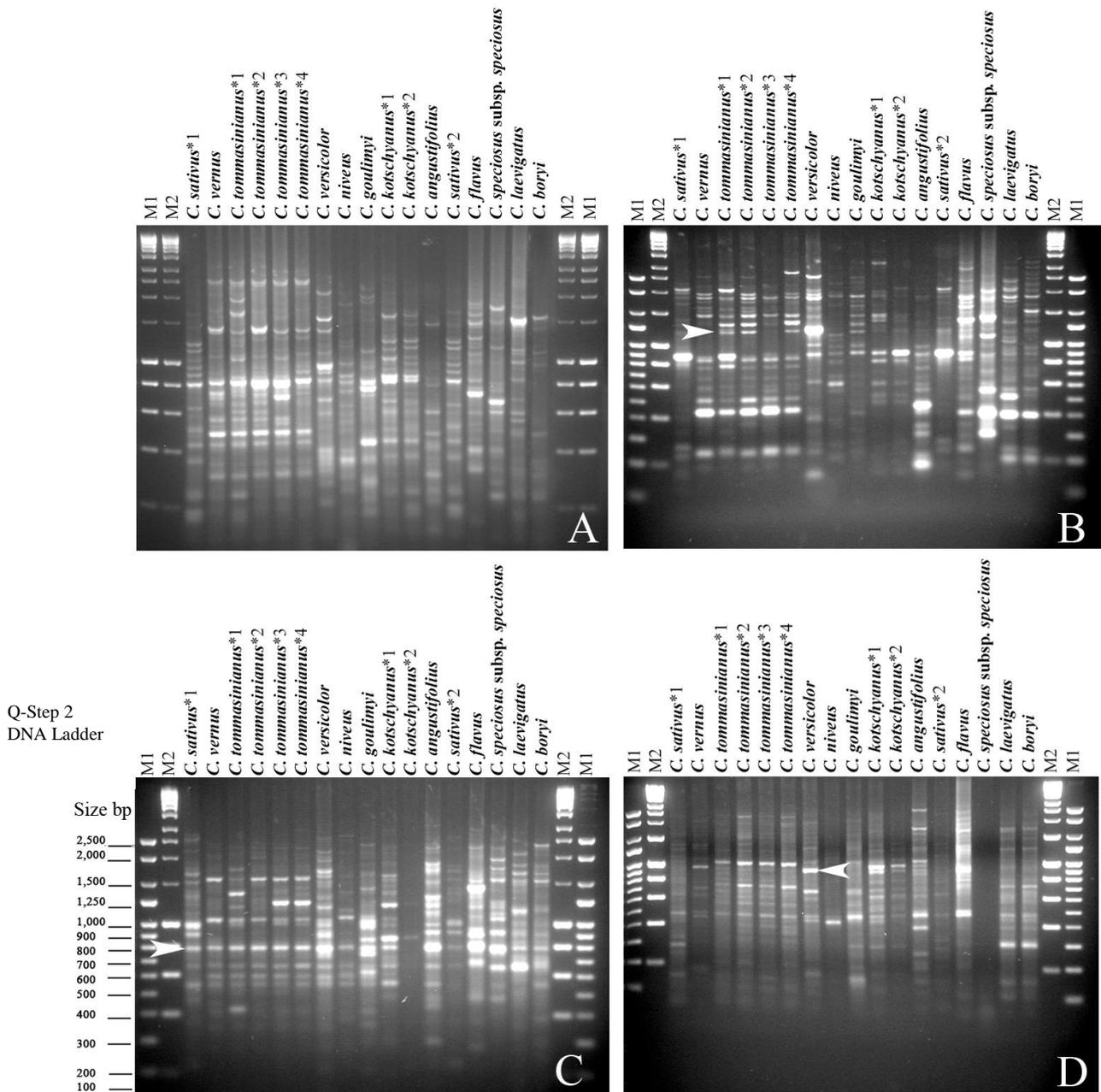


Figure 3.1: IRAP amplification pattern from 12 *Crocus* species using A) Nikita and IRAPs *Crocus* Sukkula combination, B) IRAP_S *Crocus* 5'LTR, C) LTR6150, D) RTY1 and RTY2 combination. Species name is given on the top of every lane, species name followed by asterisks (*) with digits indicating to sub-species or accession. The accessions included *C. tommasinianus**1 (lilac beauty, CtmLD09), *C. tommasinianus**2 (barr purple, CtmBD09), *C. tommasinianus**3 (rubinetta, CtmTD09), *C. tommasinianus**4 (albus, CtmAD09) and *C. kotschyanus**1 (subsp. *kotschyanus*, CkotP09) and *C. kotschyanus**2 (var. *Zonatus*, Ckot/z08). Arrow ahead indicating to common IRAPs bands present in most species (C), as well as to unique bands present in a species or two (B, D). On either side of the agarose gel (2%) is 5µl DNA length markers Q-Step 2 (M1) and HyperLadder I (M2).

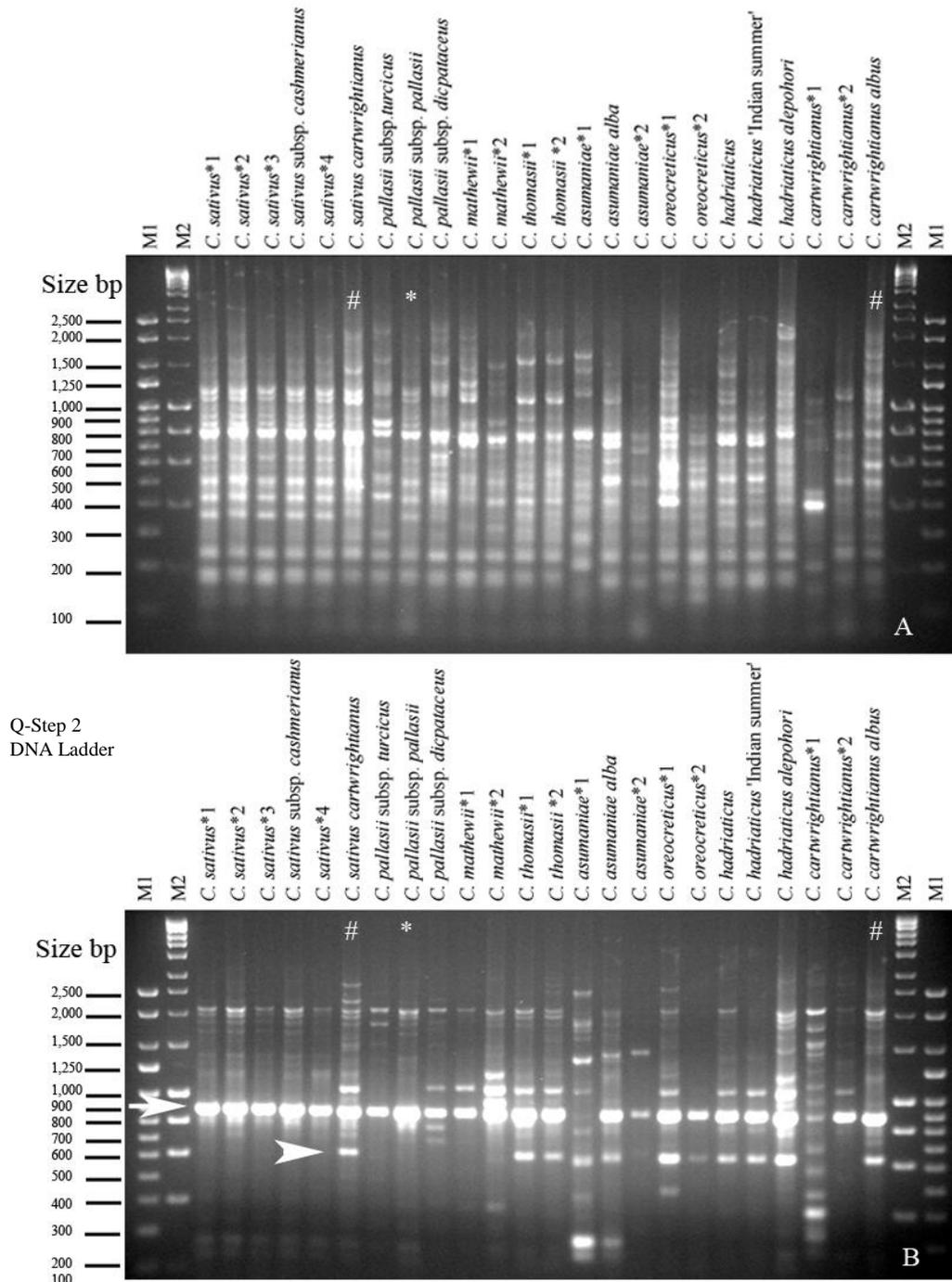


Figure 3.2: IRAP amplification pattern from 24 *Crocus* accessions (9 *Crocus* species) from *Crocus* series *Crocus* using A) Nikita and IRAP_S *Crocus* Sukkula combination, and B) IRAP_S *Crocus* 5'LTR. Species name is given on the top of every lane, species name followed by asterisks (*) with digits indicating to sub-species or accession. The accessions included *C. sativus**1 (CsatP09), *C. sativus**2 (Cstkf09), *C. sativus**3 (CstCD09), *C. sativus**4 (CsatP09), *C. mathewii**1 (CmatD08), *C. mathewii**2 (HKEP.9291, CmtHR09), *C. thomasi**1 (CtmVD09), *C. thomasi**2 (MS978, CtomI09), *C. asumaniae**1 (CasWD09), *C. asumaniae**2 (CasAD09), *C. asumaniae**3 (CasAT09), *C. oreoreticus**1 (CorVR09), *C. oreoreticus**2 (CorVD09), *C. cartwrightianus**1 (CcwBD09), *C. cartwrightianus**2 (CcrCR09), while *C. hadriaticus* (lane 21) is a white flowering accession (ChdWD09). Arrow indicates presence of a common band present in most species, while arrow head indicate to unique band absent in *C. sativus*. On either side of the agarose gel (2%) is 5µl DNA length markers Q-Step 2 (M1) and HyperLadder I (M2).

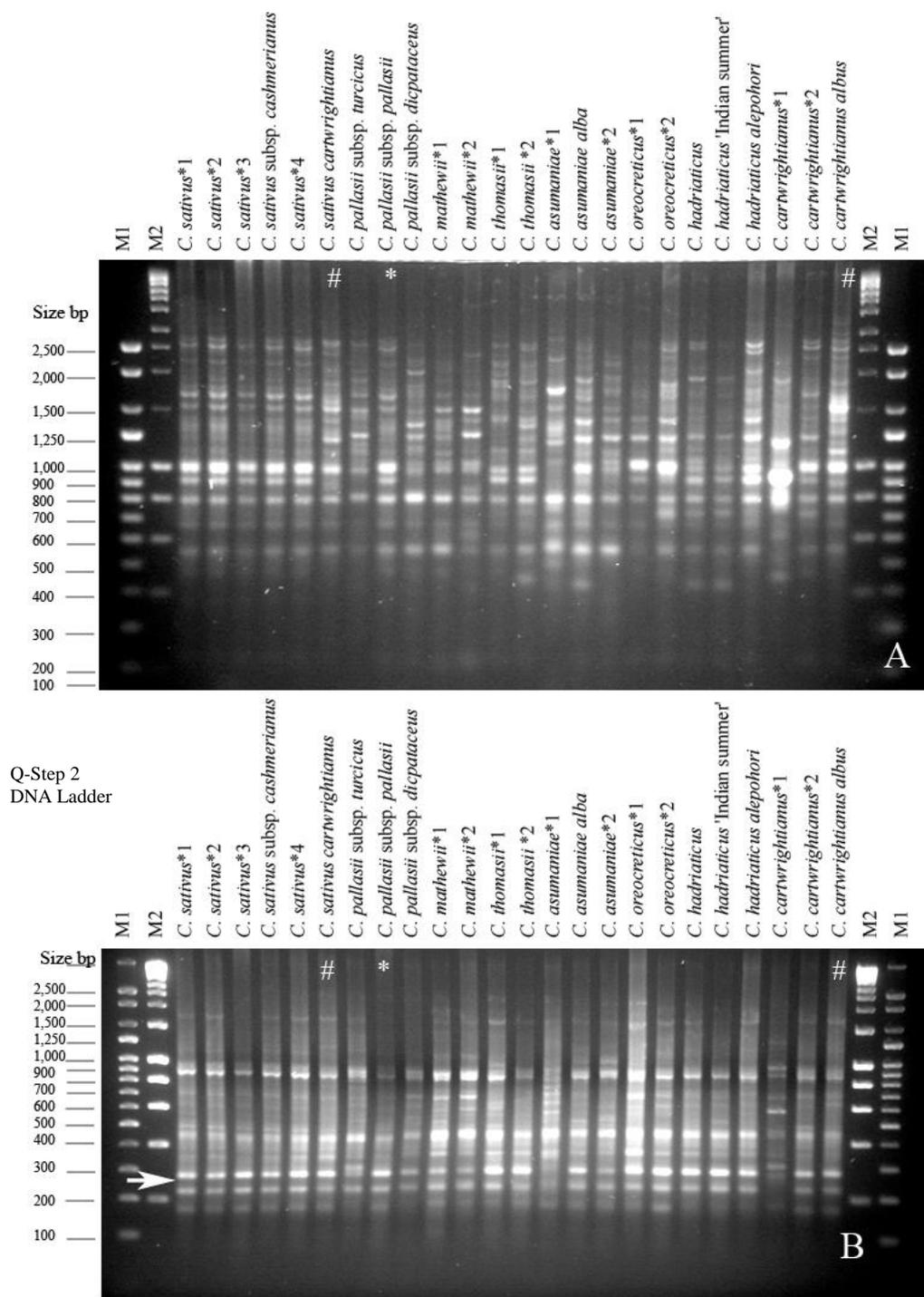


Figure 3.3: IRAP amplification pattern from 24 *Crocus* accessions (9 *Crocus* species) from *Crocus* series *Crocus* using A) LTR6150 and B) RTY1 and RTY2. Combination Species name is given on the top of every lane, species name followed by asterisks (*) with digits indicating to sub-species or accession. The accessions included *C. sativus**1 (CsatP09), *C. sativus**2 (Cstkf09), *C. sativus**3 (CstCD09), *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 (CasAD09), *C. asumaniae**3 (CasAT09), *C. oreoreticus**1 (CorVR09), *C. oreoreticus**2 (CorVD09), *C. cartwrightianus**1 (CcwBD09), *C. cartwrightianus**2 (CcrCR09), while *C. hadriaticus* (lane 21) is a white flowering accession (ChdWD09). Arrow indicates presence of a similar band present in most species. On either side of the agarose gel (2%) is 5 μ l DNA length markers Q-Step 2 (M1) and HyperLadder I (M2).

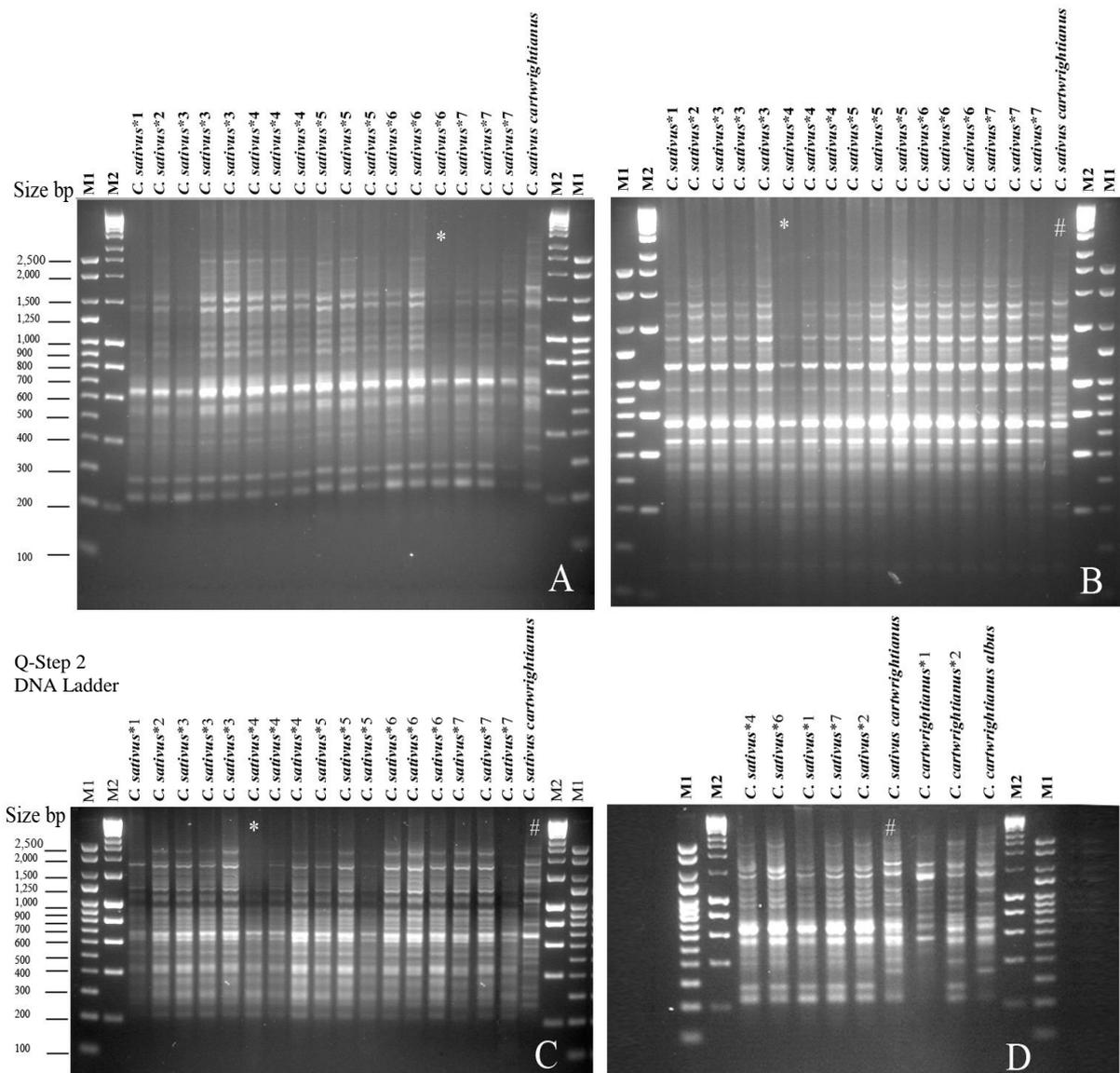


Figure 3.4: IRAP amplification pattern from 17 saffron accessions using A) Nikita, B) 5'LTR1, C) Sukkula LTR primer and D) IRAP_S *Crocus* Nikita. Saffron accessions were obtained from several sources and included *1: JW Dix Export, Netherland (2007), *2: Pottertons Nursery, UK, *3: JW Dix Export, Netherland (2010), *4: Crocus1Bank, Spain, *5: Suttons Nursery, UK, *6: Kashmir, India, *7: Var. *cashmeriensis* from JW Dix Export, Netherland (2009). Hash symbol (#) is *C. sativus cartwrightianus* having similarities with *C. cartwrightianus albus* (D). Asterisk indicates the 1500bp and above IRAP fragments missing in this accession of saffron. On either side of the agarose gel (2%) is 5µl DNA length markers Q-Step 2 (M1) and HyperLadder I (M2).

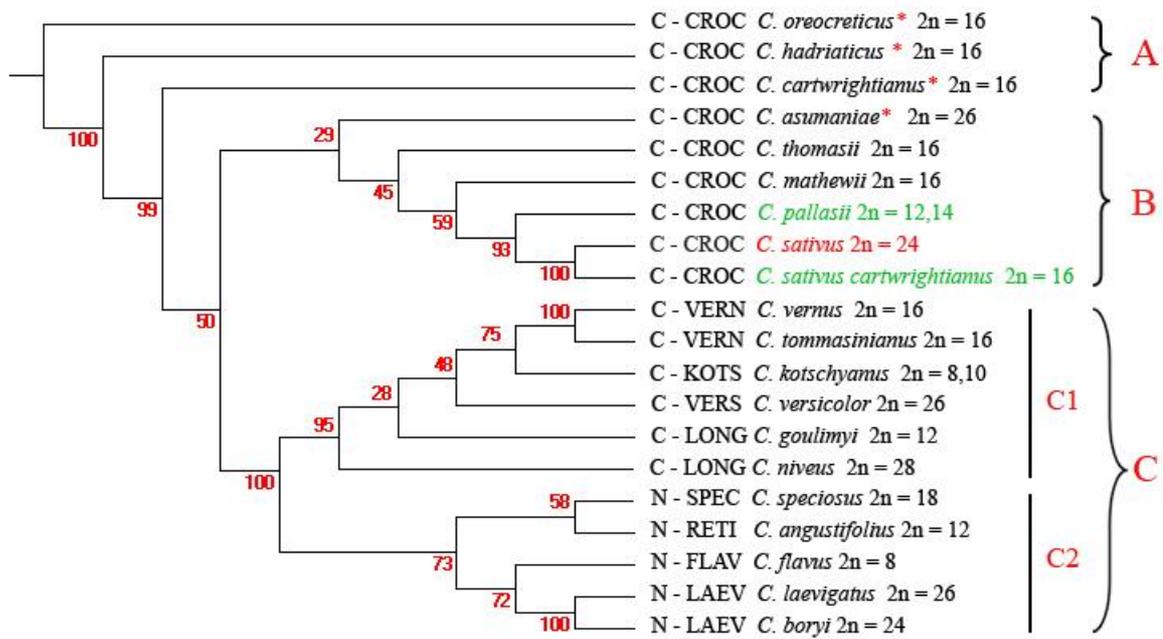


Figure 3.5: Consensus phylogenetic model of IRAPs data for 20 *Crocus* species by Neighbour Joining algorithm. Bootstrap consensus tree is inferred from 1000 replicates computed by PowerMarker software and represented in percentage on nodes. The phylogenetic analysis included 11 primers in 63 combinations and suggests that *C. sativus* is more closely related to *C. pallasii* and *C. sativus cartwrightianus*. *Crocus* species are divided into three main clades A, B and C. Clade A, B and sub-clade C1 are representatives of section *Crocus* while sub-clade C2 are representative of section *Nudiscapus* (see section 3.3.3 and below).

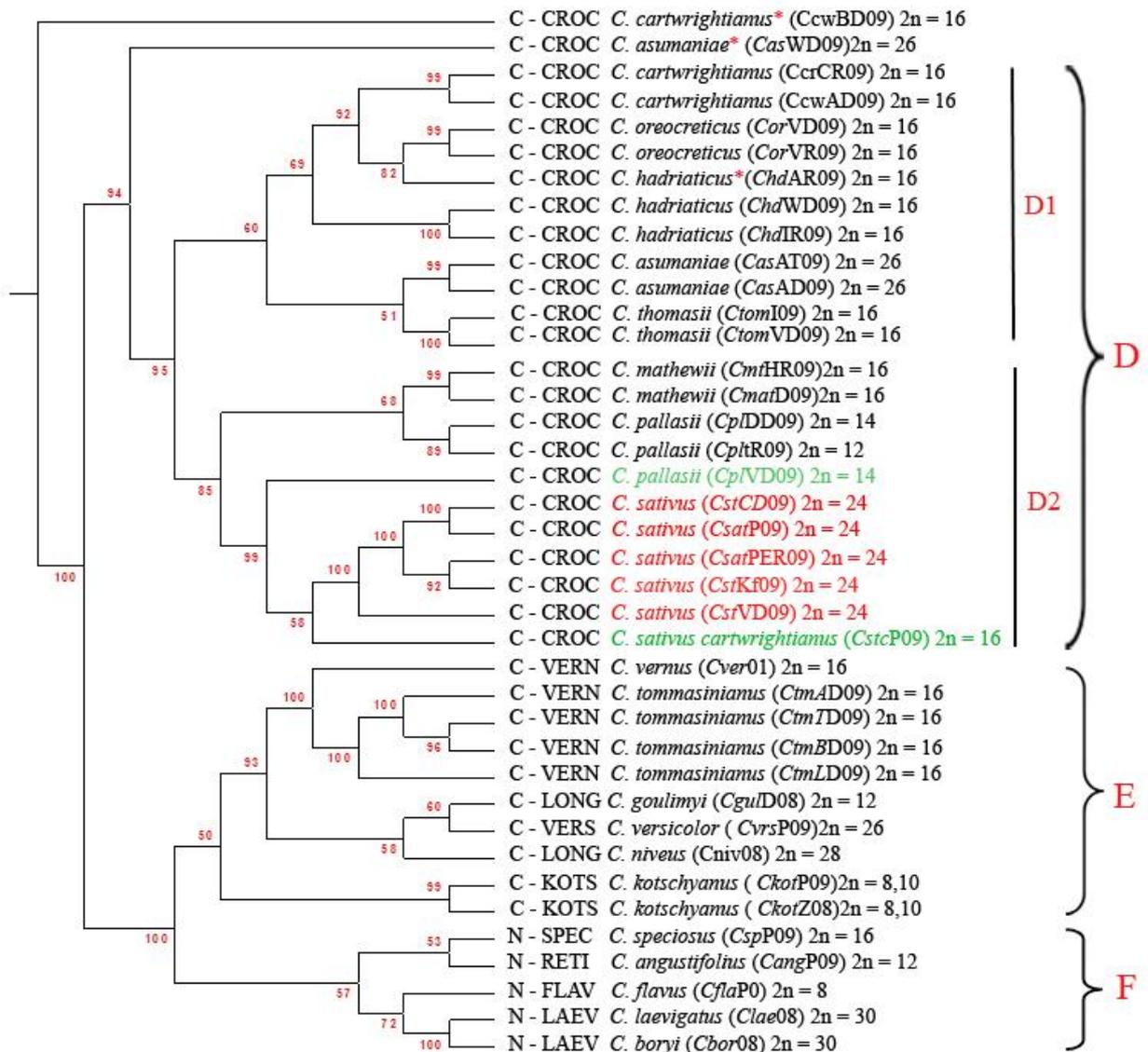


Figure 3.6: Consensus phylogenetic model of IRAPs data for 39 *Crocus* accessions (20 species) by Neighbour Joining algorithm. Bootstrap consensus tree is inferred from 1000 replicates computed by PowerMarker software and represented in percentage on nodes. The phylogenetic analysis included 11 primers in 63 combinations and suggests that *C. sativus* is more closely related to *C. pallasii* subsp. *pallasii*, and *C. sativus cartwrightianus*. *Crocus* accessions are divided into three main clades D, E and F. Clade D and E are representatives of section *Crocus* while F are representative of section *Nudiscapus* (see section 3.3.3 and below).

3.4 Discussion

Crop improvement is hardly possible without the presence of genetic variation that must be found between parents of a chosen cross, to be selected and then transmitted into subsequent generations (Heslop-Harrison and Schwarzacher, 2012; McCouch *et al.*, 2013). DNA markers have been playing a key role in the understanding of genetic variability and diversity as well as in tracing of individuals or lines carrying genes of interest (Todorovska *et al.*, 2005). The current study tested IRAP markers based on published sources as single primers and combinations of pairs (Table 3.1). Those giving amplification with *Crocus* DNA are shown, with experimentally determined optimum annealing temperatures (Table 3.2), and were applied to understand the phylogenetic relationship and genomic structures of the 20 species of *Crocus*.

All 63 IRAP primer combinations resulted in the amplification of 9490 clear and unambiguous alleles from 40 assays and the polymorphic information content (PIC) varied from 0.4 to 0.05 (Table 3.3 and section 3.2.4). The annealing temperature for IRAP markers were set up lower than the theoretical annealing temperatures (compare Table 3.1 and 3.2), and in all cases the reproducibility of amplification was confirmed by re-running the same primer combination. The banding pattern was reproducible and only rare discrepancies in banding patterns were observed. For example, when the DNA concentration was increased even up to 10-fold (from 20-200ng) minor differences were found, especially for DNA fragments of larger size and minor intensity (asterisks in Figure 3.4). However, the dominant strong bands used for polymorphism scoring remained virtually identical (arrows in Figure 3.1-3.3), indicating the robustness of the method. This contrasts with RAPD methods, generating smaller numbers of bands from each primer combination, where extreme variability has been found depending on DNA extraction and PCR conditions (Kalendar *et al.*, 1999, 2011).

Different sequences within plant genomes develop polymorphisms at a different rate (Todorovska *et al.*, 2005; Saeidi *et al.*, 2008). SSRs or microsatellites are highly polymorphic even within accessions and retroelements amplify and reinsert in the genome at a slower rate, while chromosome number variation (other than through polyploidy) usually occurs only between species (see Heslop-Harrison and Schmidt, 2012). But as bulk of the genomic DNA comprises of retroelements. Therefore, IRAPs is a preferred choice and are appropriate markers for both wild and domesticated species, and have also been applied to detect diversity and relationships within crops species (Flavell *et al.*, 1998; Nair *et al.*, 2005; Saeidi *et al.*, 2008; Carvalho *et al.*, 2010). Although, in larger

evolutionary distances IRAPs may prove less useful because similarly sized bands are not identical-by-descent (Teo *et al.*, 2005; Kalendar *et al.*, 2011). However, in this study knowledge of polymorphisms and genomic relationships in species of *Crocus* section *Crocus* and some more distant relatives, and particularly the diversity and origin of *C. sativus* (saffron) was aimed. In most cases maximum information was gained when two IRAP primers were applied, but in certain instances additional DNA fragments were amplified with a single primer as well (see Table 3.3). The Sukkula primer alone as well as in combination with other primers produced the maximum number of IRAP bands; other primers gave either only a small number and/or poorly resolved smearing bands (Figure 3.3B). The low number of IRAPs fragments indicates that there are fewer copies of the respective retrotransposon family in the *Crocus* genome than other plants genomes, but other possibilities such as nesting or mismatches in priming sites also exist or the elements are distantly spaced (Alavi-Kia *et al.*, 2008). Whereas, the presence of numerous IRAP bands indicating abundance of a retrotransposon family (Saeidi *et al.*, 2008; Kalendar *et al.*, 2011). No significant differences were observed as far as the average number of bands amplified by one or combination of primers was concerned. The average number of bands produced by the LTR primer alone was 75, while 76 bands were obtained by the combination of LTR primers (Table 3.3). These results can be compared to those of Teo *et al.* (2005) in banana where, the authors obtained 12 bands on average for a single LTR primer and 13 bands from two LTR primers. The *Crocus* genome size is some 10x larger than banana, so the increased number of fragments might be expected.

There have been considerable recent interests in understanding the relationships and diversity within the genus *Crocus*, and the existence of saffron within the genus makes it more important (see Chapter I). IRAPs revealed extensive polymorphisms between all species and high levels of polymorphism were evident between species as well as within accessions of the sexually reproducing *Crocus* (Figures 3.1-3.3). Improvements are possible by improving genetics of species but to date, only rare genetic variation has been reported within saffron accessions (Alvarez-Orti *et al.*, 2004; Nehvi *et al.*, 2007; Agayev *et al.*, 2009; Fernandez *et al.*, 2011). Among the 17 saffron accessions analysed here, no consistent differences were detected and all IRAP primer pairs showed homogeneous banding patterns (Figure 3.4). Further, the number of accessions did not relate to the level of polymorphism as described in previous studies, where the number of allele's detected and genetic diversity was strongly correlated with the number of accessions (Cui *et al.*, 2010). At the moment quality of saffron (stigma) is defined by the

local landraces, growing conditions (e.g. soil, water, temperature and altitude), collection and processing techniques and not by true genotypic variation (Alvarez-Orti *et al.*, 2004; Agayev *et al.*, 2009; Grilli Caiola and Canini, 2010).

There were no bands unique to *C. sativus*; however IRAPs generated unique bands from the non-saffron species (arrow head in Figure 3.2). Further, the relatedness of *C. sativus* with other members of the series *Crocus* was evident (arrow in Figure 3.1C). The number of IRAP bands in saffron was not greater than the number of bands amplified in the diploids, and the chances of selective (biased) and variable PCR amplification was overcome by using at least two rounds of PCR and sometimes DNA extracted in different years had no significant variation in banding pattern. The sterile triploid nature of saffron is well known, and in the current analysis IRAPs could not generate any unique banding pattern whereby origin of bands may be related to any one of the diploid species, and thus the concept of autopolyploid saffron was not supported (Figure 3.1-3.3). Although, the overall PCR profile of *C. sativus* was different from all species analysed here, but among the analysed species, it was most similar to that of *C. pallasii* subsp. *pallasii* (see Figures 3.2, 3.3 and below) and hence this subspecies (albeit with $2n=2x=14$ chromosomes) is the most likely an ancestor of *C. sativus* while *C. cartwrightianus* (see below) is supported as the other ancestral species

The phylogenetic analysis of IRAPs placed the different species and accessions into their respective sections *i.e.* *Crocus* or *Nudiscapus* and multiple accessions of the same species as a cluster within a clade (Figures 3.5, 3.6). The position of all species is in accordance to previous reports (Mathew, 1982; Petersen *et al.*, 2008). Further, sub-branching at accession level indicates the genomic diversity within accessions and high bootstrap support values show confidence and usefulness of IRAPs in discriminating between species and accessions (Figure 3.6). The *C. hadriaticus* *Alepohori* (ChdARD09), *C. asumaniae* accession 'white' (CasWD09) and *C. cartwrightianus* (CcwBD09) did not cluster with the accessions of the respective species (sub-clades D1, Figure 3.6). However, as the exact geographical origin or collection point is not known, the variation in accessions may be attributed partly to their different origin or hybridization history imposed under human agricultural practices. Evidence of a wide range of genetic and epigenetic alterations, including deletion events and elimination of non-coding, low and high copy sequences, has been well documented in both natural and synthetic polyploids (Gaeta *et al.*, 2007). Further, it has also been demonstrated that stress or unusual environmental stimuli like hybridization and tissue culture may induce

heritable DNA changes and in plants this is often associated with the accumulation and rise in the activities of transposable elements (Kubis *et al.*, 2003; Ågren and Wright, 2011; Heslop-Harrison and Schwarzacher, 2011). Still, as most of the species were purchased commercially from nurseries, this discrepancy may also be attributed to accidental mixing or inaccurate labelling and highlights the immense significance of a worldwide germplasm resource for *Crocus*. At the moment, the CrocusBank collection of the genus is perhaps the most important and precise collection (see www.crocusbank.org). The subgenus *Crocus* comprises of two sections (Mathew, 1982) and both the sections *Crocus* and *Nudiscapus* are probably monophyletic (Harpke *et al.*, 2013). In the current analysis species of both sections are clearly separated within the tree (Figure 3.5). Series *Crocus* is described as a strongly supported monophyletic group (Petersen *et al.*, 2008 and Figure 1.8) and the IRAP analysis does not contradict this inferred monophyletic origin as all members of the series are grouped together in clades A, B or sub-clade C1 (Figure 3.5).

Different species of *Crocus* series *Crocus* have been found related to *C. sativus* and have been considered as potential ancestors of *C. sativus* (see Chapter 1). For example, *C. cartwrightianus* shows morphological similarity to *C. sativus* (Figure 1.2A1&A2) and even today, *C. cartwrightianus* is used as “wild” saffron. Several studies that used morphology as well as karyotype analysis of the allied *C. sativus* species demonstrated that *C. cartwrightianus* is one of the progenitors of *C. sativus* (Maw, 1886; Mathew, 1982; Grilli Caiola *et al.*, 2004). Further, the diploid *C. oreoreticus* is similar to *C. cartwrightianus* and has also been considered as a possible ancestor of the *C. sativus* (Burt, 1948). Similarly the occurrences of repetitive DNA sequences have also been employed in phylogenetic analysis of the genus, but their contribution to the understanding of *Crocus* phylogeny was limited (Frello and Heslop-Harrison, 2000a; Frello *et al.*, 2004). However, their results did not support Mathew’s classification and led them to discuss the possibility of far-reaching hybridization and fast speciation within the genus. In case of allotriploid saffron, *C. cartwrightianus*, *C. hadriaticus*, *C. oreoreticus* (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, where each one could contribute the basic set of $x=8$ chromosomes. The AFLP method has provided further insights and revealed *C. cartwrightianus* and *C. thomasii* to be the closest relatives of *C. sativus* (Zubor *et al.*, 2004). Further, flow cytometry analysis that involved the diploid species of series *Crocus* including *C. cartwrightianus* and *C. thomasii* revealed *C. cartwrightianus* to be the most likely

ancestor of *C. sativus* (Brandizzi and Caiola, 1998). The results obtained here do not contradict these results as *C. sativus* is flanked by these diploid species of the series *Crocus* on the sister branches (Figures 3.5, 3.6). Furthermore, based on IRAP markers, *C. almehensis* and *C. michelosnii* were found closer to *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* (Figure 1.8). The results here show considerable variation between accessions of *C. cartwrightianus* (Figure 3.4D; 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* genomes (Figure 3.4 & Figure 4.13).

In the current analysis *C. sativus* is placed in between the recognized diploid species *C. pallasii* subsp. *pallasii* and '*C. sativus cartwrightianus*' (Figure 3.5). These results are contrasting to the RADP data analysis of Grilli Caiola *et al.* (2004), where the authors found *C. cartwrightianus* to be the closest relatives of saffron, followed by *C. thomasii*. However, the authors also mentioned *C. pallasii* and *C. asumaniae* to be the more distantly related to *C. sativus*. Still our results are in agreement with Erol *et al.*, (2013), where the authors found maximum similarity between one accession of *C. pallasii* subsp. *pallasii* and *C. sativus*. Similarly, chloroplast, ribosomal and nuclear single copy genes sequences focused on phylogeny and suggested *C. cartwrightianus* and *C. pallasii* as ancestral species of *C. sativus* (Harpke *et al.*, 2013) results that parallel the findings here. However, based on Inter Simple Sequence Repeat (ISSR) marker analysis *C. cartwrightianus* cv. *albus* was more closely related to *C. sativus* than to *C. cartwrightianus* (Rubio-Moraga *et al.*, 2009). In this analysis too, *C. sativus* clustered with '*C. sativus cartwrightianus*' with 100% nodal support (clade B, Figure 3.5), and the possibility of this species being *C. cartwrightianus* cv. *albus* cannot be rule out (clade D, Figure 3.6).

The existence of polyploidy in various taxa of plants is associated with certain advantages (compared to their diploid ancestors) and it may be responsible for their success (Heslop-Harrison, 2012). Different copies of the same alleles allow a greater chance of adaptation and leads to neo-functionalization (Adams and Wendel, 2005). However, major disadvantage associated with polyploidization is the disruption of meiotic cell cycle and often leading to the formation of aneuploids that may be sterile. *C. sativus* is

a sterile triploid plant, propagated exclusively by vegetative means, and studies have shown the existence of genetic diversity on a limited scale only (Sik *et al.*, 2008; Nemati *et al.*, 2012). Domestication of plants and animals some 10000 years ago has resulted in both genetic and phenotypic changes and since then our major crops have evolved and spread due to their adaptability to diverse environments (Eckardt, 2010; Matsuoka, 2011). Several of our important crops species, such as rice, have been domesticated several times with limited introgression that transferred key domestication alleles between divergent rice gene pools (Kovach *et al.*, 2007).

It is difficult to address how or when saffron originated and naturalized but a number of studies carried out over the recent years have suggested clonal origin of world's saffron. These studies hypothesised an ancient spontaneous hybridization event in nature that resulted in a unique triploid clone of *C. sativus* or saffron (Mathew, 1977; Caiola *et al.*, 2004; Rubio-Moraga *et al.*, 2009; Fluch *et al.*, 2010). Due to its sterility it relied solely on multiplication by corms and still continues to be propagated vegetatively (Dhar *et al.*, 1988, Piqueras *et al.*, 1999); thus, saffron growing around the world today may be just one clone (Jacobsen and Ørgaard, 2004 and section 1.3.2). Molecular studies regarding the clonal origin of *C. sativus* have applied RAPDs, SSRs, ISSR, AFLPs, IRAPs, ESTs and chloroplast DNA markers, and their conclusion indicated that there is just one saffron cultivar grown worldwide (Caiola *et al.*, 2004; Alavi-Kia *et al.*, 2008; Rubio-Moraga *et al.*, 2009; Fluch *et al.*, 2010). For most crops, domestication is seen as a bottleneck reducing genetic variation, and further artificial selection has advantages in maintaining its genetic characteristics but causes reduction in genetic diversity; in the case of saffron there is no robust evidence for genetic variation (e.g. Table 3.3). Phenotypic differences stated above could be attributed to differences in climate and cultivation practices (Caiola *et al.*, 2004; Rubio-Moraga *et al.*, 2009; Fluch *et al.*, 2010). (Figure 1.4), While Variation in saffron quality is mainly due to the methodology followed in processing the stigmas or may arise from adulterants to it, and is independent of the origin of the corms cultivated by farmers (Alvarez-Orti *et al.* 2004; Maggi *et al.*, 2011; Torelli *et al.*, 2014). Given the high levels of polymorphism between the wild species and even individual accessions, minimal if any variation was evident in *C. sativus*, despite accessions from a broad geographical range being included (Figure 3.4). Thus the IRAP data most likely indicate a single origin and naturalization of the triploid *C. sativus* and support its clonal propagation.

Evolutionary history of the genus *Crocus* is very complex as indicated by intensive species hybridization and explosive speciation in the evolution of *Crocus* (Frello *et al.*, 2004) and that could be one of the selective pressures in the origin of saffron (Fernandez, 2007). Nevertheless, high variation between accessions within each of the species (other than *C. sativus*) was evident, and it is clear that much more extensive collections will be required to circumscribe the taxonomic units. Many of the wild *Crocus* species, although locally abundant in their native range, are difficult to maintain in cultivation. Further, *Crocus* species are perennial and flowers many times during its lifespan. Thus in theory, it offers the possibility to hybridise with a number of coexisting genotypes and overlapping generations within the population (Larsen *et al.*, 2015). Further, the hybrid origin of cultivars such as Golden Yellow (3x) or Stellaris (2x) in genus *Crocus* is well-documented (Ørgaard *et al.*, 1995). These evidences suggest interspecific hybridization may occur occasionally, with consequences allowing gene-flow and homogenization, through to hybrid speciation, and leading to uncertain delimitation of species. Similarly, microspores of saffron can germinate on the stigma of other species of *Crocus* series *Crocus* and the vice versa. Viable seeds have been obtained from the all diploids of the *C. sativus* aggregate through self and cross-pollination as well as from a cross between *C. sativus* with *C. cartwrightianus* (Grilli Caiola *et al.*, 2010; 2011). These results further support hybridization and introgression within species of the *Crocus* series *Crocus*. Such outbreeding with relatively long-distance gene exchange may be associated with the present diversity in *Crocus* that keeps the genus flexible and species are able to thrive in diverse habitats and hence given rise to rich speciation (Larsen *et al.*, 2015).

For almost 80% for their food intake, humans depend on fewer than a dozen of the approximately 300,000 flowering plants species. Thus only a fraction of the available genetic diversity within species is overly exploited (McCouch *et al.*, 2013). The current study emphasizes the importance of assessing genetic diversity in germplasm characterization and conservation. A wide genetic base is of great importance for the development of improved varieties and it will be of great importance in the improvement of saffron cultivars, exploitation of their genetic diversity and conservation of the *Crocus* germplasm.

4 CHAPTER IV: PHYLOGENETIC RELATIONSHIPS OF *CROCUS* SPECIES AND POLYPLOID NATURE OF SAFFRON INFERRED FROM EST-SSR, SNPS AND BARCODING GENES

4.1 Introduction

In spite of the historical role of morphological traits in understanding the affinities of plant taxa (Linnaeus, 1753). Subsequently, the availability of DNA-based markers has become increasingly widespread and important in resolving controversies related to phylogeny by estimating more objectively and precisely genetic variation between taxa (APGIII: Bremer *et al.*, 2009). The bulk of genetic variation at nucleotide level is often not visible at a phenotypic level. In contrast phenotype, which is the result of complex interactions between genotype and environment, is influenced by environmental conditions and displays variation under different sets of environmental stimuli that may be the result of divergent or convergent evolution. Therefore, relationships deriving from morphological traits alone may not be completely accurate (see Varshney and Dubey, 2009; Yang *et al.*, 2013). In the recent years, studies on genetic diversity in several crop species, such as rice, wheat, barley, brassica and banana have added much to our understanding and the number of phenotypic markers in these species has increased to potentially hundreds of DNA-based useful polymorphisms (Todorovska *et al.*, 2005; Heslop-Harrison and Schwarzacher, 2012).

From the late 1990s to present PCR-based approaches are perhaps the most powerful and important. The robustness, stability and reproducible nature make them ideal tools for marker assisted selection (MAS) breeding programs, allowing hundreds of genotypes to be screened in minimal time at low costs (Varshney and Dubey, 2009). There are several powerful PCR-based marker systems in use for understanding the phylogenetic relationships and diversity levels within both plants and animals (Lu *et al.*, 2013) and some of the most frequently applied markers are described in (Chapter I, section 1.8). However, because of the higher levels of polymorphism, low cost and known map-based locations, EST-SSRs and SNPs are considered the most efficient molecular markers which have a wide range of applications in genetic mapping, gene tagging, and studies of genetic diversity and evolution (Gao *et al.*, 2004; Gadaleta *et al.*, 2011; Heslop-Harrison and Schwarzacher, 2012).

ESTs represent parts of the expressed genes and SSR markers generated from EST sequences are potential candidates for genes with known or putative functions and may be applied to comparative studies in related species. EST-SSRs are generally less polymorphic than random SSRs but both types of markers perform similarly in estimating genetic diversity (Gao *et al.*, 2004; Ramu *et al.*, 2013). Further, the development of EST-SSR is relatively easy because of the public availability of the thousands of ESTs in the GenBank database (<http://www.ncbi.nlm.nih.gov/dbEST/>) that allow users to investigate sequence qualities and EST structural features.

SNPs refer to genomic variation caused by a single nucleotide mutation at a specific locus (see Chapter I). Higher density, genetic stability and the amenable nature of SNPs to high-throughput automated analysis make them the preferred choice for detailed analysis of genome-wide association mapping and precision mapping (Isobe *et al.*, 2013). To date, the degree of genetic variation and phylogenetic relationships in several crop important species such as, *Asparagus* (Caruso *et al.*, 2008), *Capsicum* (Jeong *et al.*, 2010), English walnut (Ciarmiello *et al.*, 2011), cucumber (Hu *et al.*, 2011), wheat (Gao *et al.*, 2004; Gadaleta *et al.*, 2011), red clover (Isobe *et al.*, 2013), maize (Frascaroli *et al.*, 2013) and *Sorghum* (Ramu *et al.*, 2013) have been successfully assessed with EST-SSR and SNP markers.

Recently DNA barcoding which employs small, standardized portions of the genome (maturase K, *matK*; ribulose 1,5-bisphosphate carboxylase/oxygenase large subunit, *rbcL*; intragenic spacer between tRNAH is GUG gene and photosystem II thylakoid membrane protein of Mr 32.000 gene, *trnHpsbA*; internal transcribed spacer of nuclear ribosomal DNA, ITS) as substitutes for morphology have been widely applied to species identification as well as in phylogeny (Hebert *et al.*, 2003; Kress and Erickson, 2007; Seberg and Petersen, 2009; Gismondi *et al.*, 2013). These genes are universally present in the target lineages and the unique sequence diversity allows discrimination among species (Kress *et al.*, 2005; Yang *et al.*, 2013). The nuclear ribosomal DNA internal transcribed spacer 2 (ITS2) region and MaturaseK gene (*matK*) of the chloroplast of are regarded as universal DNA barcodes (Hollingsworth *et al.*, 2009; Yao *et al.*, 2010).

Crocus, as described in Chapter I, is a genus with some 88 recognized species, divided into two sections *Crocus* and *Nudiscapus*. Section *Crocus* is divided into six series each including between 2 and 10 species (Figure 1.9) and saffron is the most economically important species in the genus *Crocus*. Despite the genus having been the subject of many investigations, the genetic base of saffron and phylogeny of the genus remain topics of

major interest (Caiola and Canini, 2010; Erol *et al.*, 2013). However, only rare diversity has been reported within saffron grown worldwide, and the phylogeny of the species within the genus is poorly understood (see Chapter I and below). Peterson *et al.* (2008) analysed DNA sequences from five plastid regions in the genus *Crocus*. Their results were largely successful in resolving the different series (with a few ambiguities), but within the series, resolution was poor. Thus, the goal of this research was to apply SNP and barcoding markers as well as to develop novel EST-SSR markers for *Crocus* by screening of ESTs and to examine 1) heterozygosity within accessions; 2) variation within species; 3) similarities and phylogeny of sequences between species; and 4) the ancestral species present in the triploid saffron crocus, *C. sativus*.

4.2 Material and methods

4.2.1 Plant materials and DNA extraction

The study included 43 accessions from 23 different species of the genus *Crocus*. A full list of accessions used is given in the materials and methods (Chapter II, Table 2.2). Plants were grown and total genomic DNA extracted from young leaves and floral buds of the *Crocus* species and accessions as described in materials and methods. A few additional DNA sequences for species not available here, for analysis their sequences were downloaded from GenBank (see also below).

4.2.2 PCR markers and primer design

A total of 20 PCR markers, including 4 SNPs, 11 EST-SSRs and 5 previously known barcoding markers were applied in the current study (Table 4.1). SNP markers were obtained from M. Santaella and J.A. Fernández (University of Castilla-La Mancha, Spain, partners in the Crocusbank project), while the EST-SSR markers were developed during this study. A total of 6,603 unique *Crocus* ESTs (available from Saffron Genes database <http://www.saffrongenes.org> mostly from the research groups of Giovanni Giuliano and Jose-Antonio Fernández) were screened for presence of SSRs using the online tandem repeat finder application (<http://tandem.bu.edu/trf/trf.html>). A total of only 15 ESTs containing SSRs were obtained, or which it was possible to design primers from 11 with a repeat size between 2 and 4bp, and a minimum length of 17 repeat units for a dinucleotide and 11 repeats for trinucleotides were used. Out of the above acquired ESTs, 11 primer

pairs flanking bi, tri or tetra nucleotide repeats were designed using Primer 3 (http://www.frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). Preferred product size, primer length and GC content were as described in materials and methods (Chapter II) and primer pairs were synthesized from Sigma (www.sigmaaldrich.com/).

4.2.3 PCR amplification and gel electrophoresis

Annealing temperature for all primer pairs (Table 4.1) was optimized using a Tprofessional Gradient Thermocycler (Biometra). PCR conditions and gel electrophoresis was carried out as described in materials and methods (Chapter II). Primers amplified products of the expected size for all microsatellites except for the tetranucleotide spanning 14 repeats.

4.2.4 Cloning and sequencing of PCR products

The products of two primers (one SNP and EST-SSR each) were selected to assess genetic diversity and relationship of the *Crocus* species (Table 2.1). In order to verify the presence of SSRs, SNPs and determine the source of allelic sequence variation, eluted PCR products of TC₂₅ and ATPs from 43 *Crocus* accessions of 23 species were cloned in pGEM[®]-T Easy vector. While for the barcoding genes, purified PCR products from accessions from the *Crocus* series *Crocus* was sequenced directly (without cloning) with custom primers (see Table 4.1 and below). All sequencing reactions were carried out commercially as described in the materials and methods (Chapter II and Appendix 4).

4.2.5 Sequence variability and phylogenetic analysis

Multiple sequence alignment of the TC₂₅, ATPs and barcoding gene sequences was performed using Geneious multiple alignment tool imbedded in Geneious R6 (Biomatters Ltd. available at <http://www.geneious.com>) and improved by eye where necessary. Phylogenetic reconstruction and estimation of nucleotide variability were carried out using Geneious R6. The evolutionary history was inferred by using Neighbour Joining (NJ) method based on the Tamura 3-parameter model (Tamura, 1992). Nodal support was assessed via bootstrapping, and the bootstrap consensus tree was inferred from 1000 replicates (Felsenstein, 2005) using default settings of the software.

Table 4.1: List of PCR primers used to amplify DNA sequences from *Crocus* species given along GenBank accession number, repeat region, primer sequence, GC content, annealing temperature and expected PCR produce size.

No.	Marker name	GenBank accession	Repeat region/ amplified gene	Primers Sequence (5'-3')	GC content (%)	Annealing temperature (°C)	Expected product (bp)
1	EST-SSR	BM 027735	(AG) ₂₃	F:TGCAAAAATCTGTGTCGTGTC R: TTGATCAACAACATAGACTGG	42.9 38.1	60	234
2	EST-SSR	BM 027734	(CT) ₁₇	F:GAGAGGTGTTAGTGGAGGTTGTC R: TCCATGGACAAATGGAAAGAG	52.2 42.9	60	210
3	EST-SSR	BM027660	(GTA) ₁₁	F: GGGCGTAAAGAACCAACATC R: CTCCACCTCCAGGAAATAACAG	50 50	60	329
4	EST-SSR	EX142902	(GCA) ₁₁	F: GAGGCGATGGACGTCTTG R: CAAGATCAGCCCTAACAAATATCC	61.1 41.7	60	239
5	EST-SSR	EU424137.1	(ATGT) ₁₄	F: GATTCATGTACGTGTGAGTTGC R: TGAACCTAACAAATATAGCACACA	45.5 34.6	64	171
6	EST-SSR	GU372953.1	(GA) ₃₀	F: GCAACGGTGCTAAAGAGGTC R: CAACTCCCACATGTGTTTCG	55 50	60	242
7	EST-SSR	GU372952.1	(AC) ₃₄	F: GGAATCTTTGCCGAGTGTC R: GCGCTAATGCTTTACCAACC	50 50	60	222
8	EST-SSR	GQ414769.1	(CT) ₂₂	F: GAGGTCCAAGGTGCTGACAT R: CCAGTGCAGGTGTTCTCTCA	55 55	60	193
9	EST-SSR	EF535584.1	(TC) ₂₅	F: TCCCTACACCAACAAAACC R: CCTGAAACCTGGAGGAAGTG	50 55	60	190
10	EST-SSR	GU372958.1	(GAA) ₁₆	F: CCGGTCAGTACAAACCACT R: ACCCCTAGATCCCCAGACAC	52 60	60	239
11	EST-SSR	GU372955.1	(AAG) ₂₅	F: TGGCCGTTATACCACTACCC R: CTCTCCTTGTGCCTATTCCA	55 47	60	191
12	ATP _s *1	EX145466	-	F:CTGGACTCTTCTTC GCATCTTT R:GCAGAAAACAGAGTTCAAACAAT	45.4 36.3	58	205
13	A111*1	HO045228	-	F:TCGATGATCCAACCAGGAAG R: GAGCTCTTG CCTCCATTCC	50 57.8	58	152
14	A38*1	*2	-	F:ACTTCCAGGAACCACTTCTC R:GGGAGAAACTCAAATCCACTG	50 45.5	58	300
15	A82*1	HO045224	-	F:CCTATTGAGGAGCCACAACC R:CGCGTGTGTGCTGATCTCTG	55 60	58	116

Table 4.1: continued.

No.	Marker name	GenBank accession	Repeat region/ amplified region	Primers Sequence (5'-3')	GC content (%)	Annealing temperature (°C)	Expected product (bp)
16	<i>rbcL</i>	*3	<i>rbcL</i>	1F: ATGTCACCACAAACAGAAAC 724R: TCGCATGTACCTGCAGTAGC	-	50	726
17	ITS	*4	ITS	ITS 4: TCCTCCGCTTATTGATATGC ITS 5: CAAGATCAGCCCTAACAAATATCC	-	52.9	329
18	<i>matK</i>	*3	<i>matK</i>	XF: TAATTTACGATCAATTCATTC 5R: GTTCTAGCACAAAGAAAGTCG	-	50	802
19	<i>matK</i>	*3	<i>matK</i>	390F: CGATCTATTCATTCAATATTTTC 1326R: TCTAGCACACGAAAGTCGAAGT	-	50	966
20	<i>trnH-psbA</i>	*5	<i>trnH-psbA</i>	F: GTTATGCATGAACGTAATGCTC R: CGCGCATGGTGGATTCCACAATCC	-	53.9	651

*1: unpublished oligos were obtained from M. Santaella, J.A. Fernandez (personal communication)

*2: primer sequences were provided by Santaella and Fernandez (personal communication)

*3: chloroplast gene (ribulose-bisphosphate carboxylase and maturase K)

*4: nuclear genes

*5: ribosomal genes

4.3 Results

4.3.1 PCR markers analysis

PCR products of ATPs, TC₂₅ and barcoding genes were sequenced and included *C. pallasii* subsp. *pallasii* (CplR09), *C. veneris*, *C. cancellatus* and *C. biflorus* too, raising the total number of accessions to 44. Further, DNA from four accessions *C. cartwrightianus* (Figure 4.1), *C. vernus*, *C. versicolor*, *C. speciosus* (Figure 4.1) gave no amplification or inconsistent results across all primers, presumably because of its poor DNA quality. However, these accessions have given products later except *C. vernus* and are included in sequence analysis.

Amplification of four nuclear DNA sequences (SNP markers) from 40 accessions of 20 *Crocus* species, including five *C. sativus* accessions is given (Figures 4.1, 4.2). The ATPs marker (Figures 4.1A, 4.2A) showed two distinct bands of 205bp with *C. sativus* and some wild accessions having both bands, some wild species having only the lower band (Figure 4.1A). In some lanes, the lower band was broader, suggesting two products of similar size. Primers for A111 showed two clear bands in *C. sativus* band of 152bp, or only the lower band, or gave no amplification in most of the wild species (Figures 4.1B, 4.2B). It was interesting that most members of the series *Crocus* amplified the lower band, while the larger band was obtained in *C. pallasii* subsp. *pallasii* and *C. asumaniae* only among the wild species (Figures 4.2B). Primers for A38 produced a monomorphic band of 300bp across all accessions of the series *Crocus* except in *C. cartwrightianus* *1 (Figure 4.2C). A few wild species of the genus *Crocus* did not amplify this band (Figure 4.2C). Primers for A82 marker also amplified the expected product of 116bp product from most accessions of the series *Crocus* (Figure 4.2D), while from the wild species of the genus it produced multiple bands and mostly of different sizes than expected (Figure 4.2D). In several accessions including *C. sativus* the band was broader, and is perhaps made of two products of similar sizes.

Figures 4.3 and 4.4 shows the amplification pattern of the 11 EST-SSR markers developed in this study. These EST-SSR markers have given products within the expected range from *Crocus* accessions and species (Table 4.1). The AG₂₃ marker showed multiple bands along the expected 234bp with *C. sativus* and several wild accessions of the series *Crocus* (Figures 4.4A). The wild accessions from other series of the genus lack this band (Figure 4.3A) and *C. pallasii* subsp. *dispathaceus* produced four bands (Figure 4.4A). In *C. sativus* the band was broader, and when the amplified PCR product was run later on a 4% agarose gel, the band clearly separated into 3 distinct bands. Primers for CT₁₇ produced

210bp amplicons and showed three clear bands in *C. sativus*, gave two bands or only the lower band in the wild species of the series *Crocus* (Figures 4.4B). Except *C. tommasinianus**1, *C. niveus* and *C. laevigatus*, the wild accessions from other series of the genus did not show this band (Figure 4.3B). Furthermore, the lower two bands produced by *C. sativus* showed resemblance to *C. pallasii* subsp. *pallasii* (Figure 4.4B). The GTA₁₁ marker showed distinct multiple bands of approximately 329bp with *C. sativus* and several wild accessions of the series *Crocus*, and two species showing only the clear lower band. In most accessions the higher band was broader, suggesting two products of similar size (Figure 4.4C). Only one accession of *C. tommasinianus* did not amplify this band, all other accessions of the wild species produced a clear single band (Figure 4.3C). Primers for the GCA₁₁ marker (Figures 4.3D, 4.4D) produced a thick strong band of approximately 239bp from all accessions of *C. sativus* and multiple bands from other accessions in series *Crocus* except in *C. pallasii* subsp. *pallasii*, *C. thomasii*, *C. cartwrightianus**2 and *C. cartwrightianus albus*. Further, the thickness of the band in *C. sativus* is roughly three times that of the single band seen in the three accessions above, and may be related to the ploidy level in *C. sativus* (Figures 4.4D). Similarly, the wild species from other series of the genus showed a single band of higher molecular weight than seen in *C. sativus*. While *C. tommasinianus**1, *C. flavus*, *C. laevigatus* and *C. boryi* produced double bands, in few accessions no amplification was seen (Figure 4.3D). Primers for ATGT₁₄ show a single band of approximately 171bp in *C. sativus* and produced similar bands in *C. pallasii* subsp. *pallasii*, *C. hadriaticus* (one accession) and *C. cartwrightianus* (Figures 4.4E). Other accessions of the series *Crocus* that produced a single band, were polymorphic, or produced double bands, While in *C. thomasii* no amplification was seen (Figure 4.4E). Wild accessions from series other than *Crocus* did not show this band, or amplified a band of lower molecular weight than that of *C. sativus* except *C. tommasinianus**1 (Figure 4.3E). Primers for the GA₃₀ marker produced a monomorphic band of approximately 242bp from all accessions of the genus *Crocus* except in *C. vernus* (Figures 4.3F, 4.3F). The AC₃₄ marker produced a ladder like banding pattern characteristic for the tandemly repeated DNA sequences. This marker produced multiple bands along the expected 222bp product from all *Crocus* accessions analysed except one accession of *C. thomasii* (Figures 4.3G, 4.4G). Primers for CT₂₂ markers resulted in a single band of approximately 193bp from all accessions of the series *Crocus* except, *C. pallasii* subsp. *pallasii*, *C. thomasii*, *C. oreoreticus* and *C. hadriaticus* ‘Indian summer’ produced double bands (Figure 4.4H). Wild accessions from series other than *Crocus* resulted in multiple bands or a single broader band, suggesting two products of

similar size, except in *C. angustifolius* which showed a single band (Figure 4.3H). The TC₂₅ primers produced multiple bands from all accessions of the genus except a single band in *C. kotschyanus* (Figures 4.3I, 4.4I). In most lanes, the band was broader, suggesting multiple products of similar size. In *C. sativus*, as well as in several members of the series *Crocus* the PCR amplified product of TC₂₅ was run on a 4% agarose gel and the product clearly separated into 3 distinct bands, ranging from ~190bp to 226bp in size (Figure 4.10). Furthermore, primers for CAA₁₆ marker produced a monomorphic band of approximately 239bp from all analysed *Crocus* accessions (Figures 4.3J, 4.4J). All five accessions of *C. sativus* produced multiple bands and interestingly, *C. pallasii* subsp. *pallasii*, *C. cartwrightianus albus* and *C. asumaniae* produced similar marker profile suggesting the close relationship of these species to *C. sativus* at least at the CAA₁₆ marker level (Figure 4.4J). Further, the size of CAA₁₆ marker amplified in *C. sativus* is broader and more intense than any other wild species (Figure 4.3J). The AAG₂₅ marker showed three bands in *C. sativus* and some wild accessions such as *C. sativus cartwrightianus* produced two bands, while few wild accessions showed only one polymorphic band of approximately 191bp (Figure 4.4K). With the exception of *C. tommasinianus* and *C. speciosus* subsp. *speciosus* no other wild accession of the series other than *Crocus* show this band (Figure 4.3K). This marker could show clear differences among the members of the series *Crocus* and even between accessions of the same species (compare accession in Figure 4.4K).

The ATPs markers, A111, A38, A82, AG₂₃, CT₁₇, GTA₁₁ and GCA₁₁ were also applied to assess genetic diversity within the 17 *C. sativus* accessions. These accessions were obtained from several sources having wide geographic distribution (see Table 2.2 and Chapter II). All these markers indiscriminately amplified monomorphic bands, and all samples produced identical marker profiles, suggesting the scarcity of genetic diversity within *C. sativus* genome (Figure 4.5).

Figure 4.6 shows PCR amplification of five barcoding genes from members of the *Crocus* series *Crocus*. All five markers, *rbcL* (Figure 4.6A), *matK* XF+5R (Figure 4.6B), *matK* 390F+1326R (Figure 4.6C), *trnH-psbA* (Figure 4.6D) and ITS 4+ITS 5 (Figure 4.6E) successfully produced PCR amplicons from 15 *Crocus* series *Crocus* accessions (9 species). The ITS 4+ITS 5 combination produced multiple polymorphic bands, perhaps due to multiple ITS copies from the accessions, all other markers produced a single band within the range of expected sizes (Table 4.1). Therefore, the ITS 4+ITS 5 combination were not included in further sequence analysis.

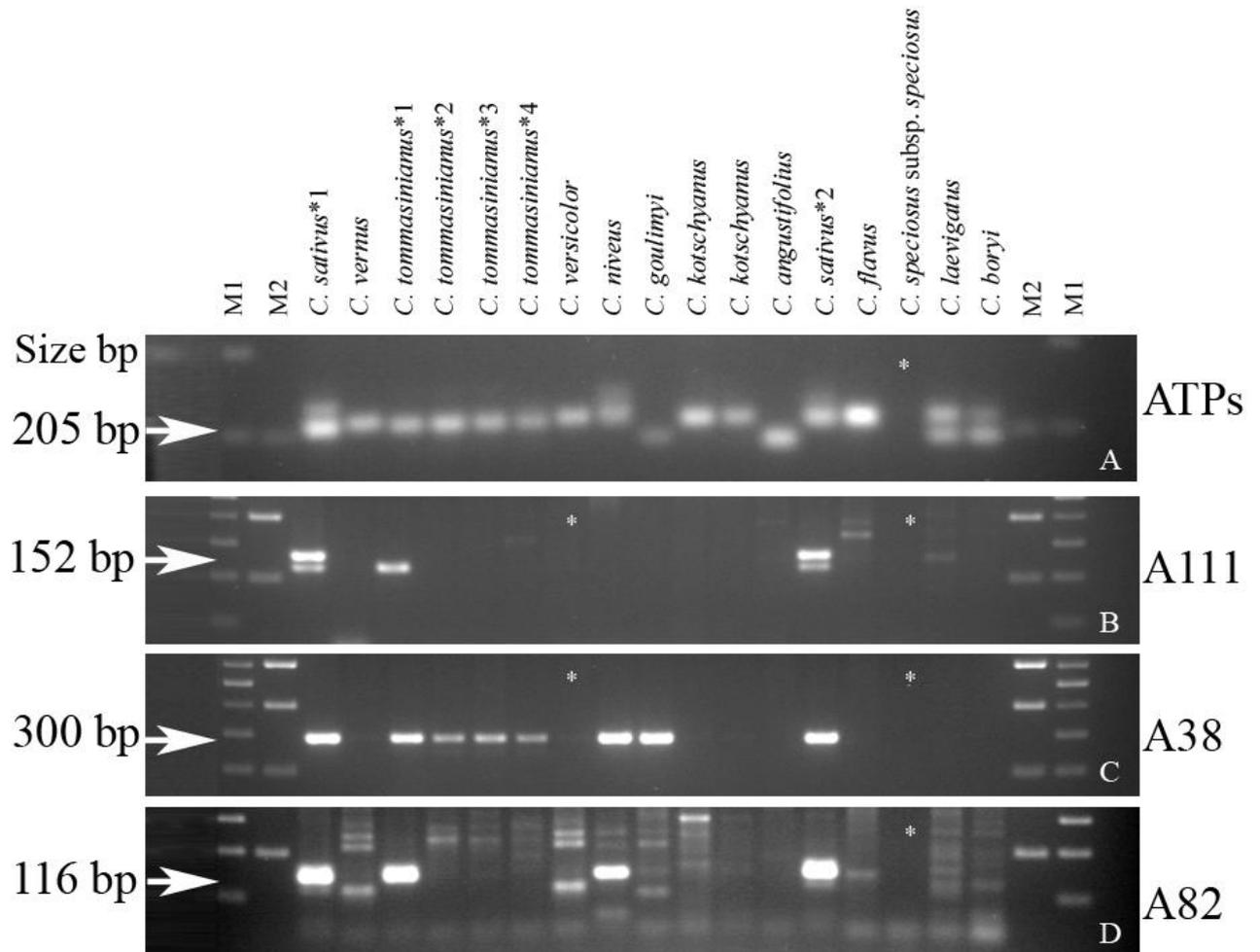


Figure 4.1: PCR amplification pattern of the SNP markers A) ATPs, B) A111, C) A38, D) A82 from 17 *Crocus* accessions of 12 *Crocus* species from other *Crocus* series. Name of species is given on the top of every lane, species name followed by asterisks (*) with digits indicating to sub-species or accession. The accessions included *C. sativus**1 (CsatP09), *C. sativus**2 (Cstkf09), *C. tommasinianus**1 (lilac beauty, CtmLD09), *C. tommasinianus**2 (barr purple, CtmBD09), *C. tommasinianus**3 (rubinetta, CtmTD09), *C. tommasinianus**4 (albus, CtmAD09) and *C. kotschyanus**1 (subsp. *kotschyanus*, CkotP09) and *C. kotschyanus**2 (var. *Zonatus*, Ckot/z08). Arrow indicates the amplified products of expected sizes. On either side of the agarose gel (2%) is 5µl DNA length markers Q-Step 2 (M1) and HyperLadder I (M2).

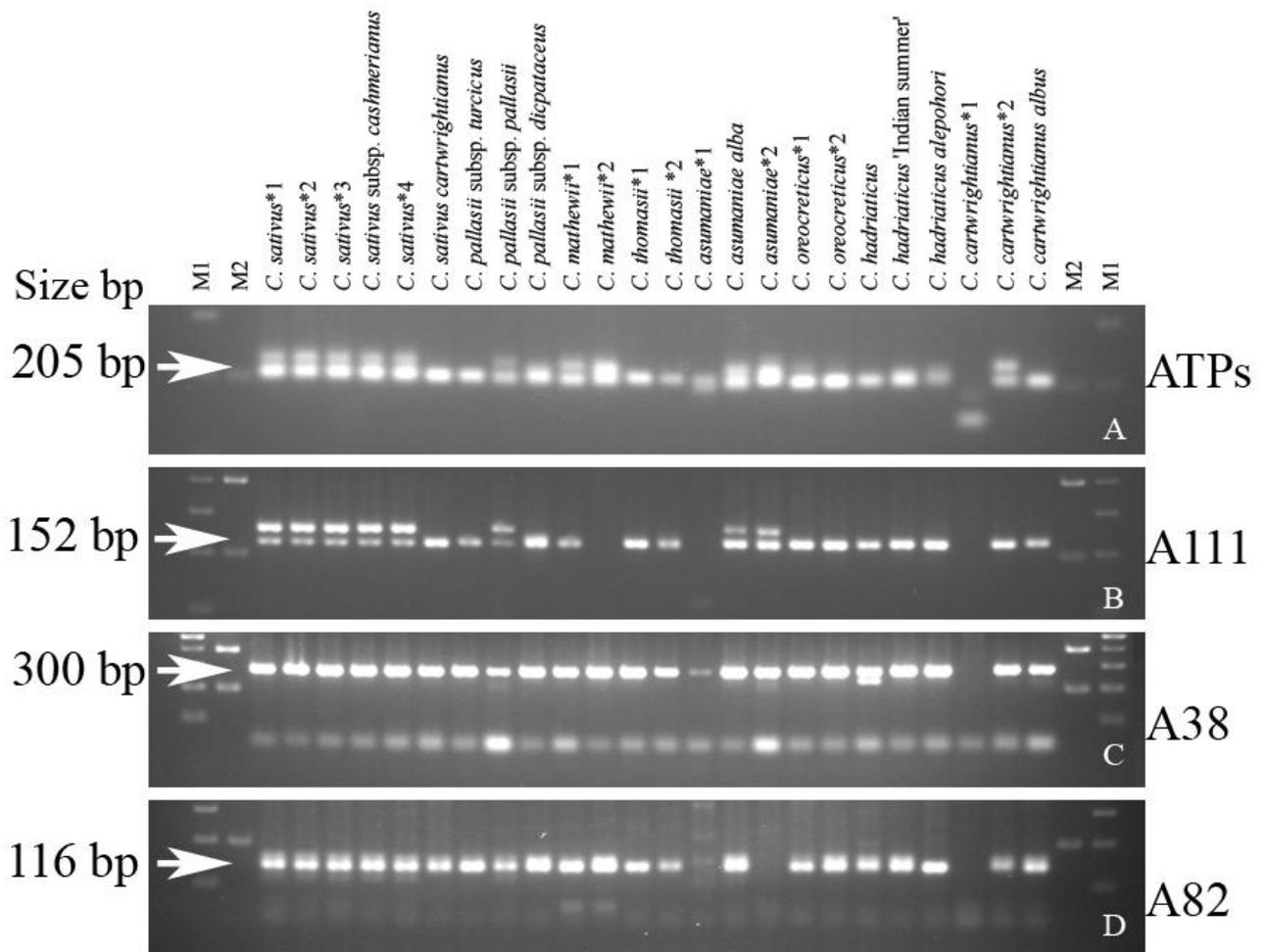


Figure 4.2: PCR amplification pattern of the SNP markers, A) ATPs, B) A111, C) A38, D) A82, from 24 *Crocus* accessions of 9 *Crocus* species series *Crocus*. (Species name is given on the top of every lane, species name followed by asterisks (*) with digits indicating to subspecies or accession. The accessions included *C. sativus**1 (CsatP09), *C. sativus**2 (Cstkf09), *C. sativus**3 (CstCD09), *C. sativus**4 (CsatP09), *C. mathewii**1 (CmatD08), *C. mathewii**2 (HKEP.9291, CmtHR09), *C. thomasii**1 (CtmVD09), *C. thomasii**2 (MS978, CtomI09), *C. asumaniae**1 (CasWD09), *C. asumaniae**2 (CasAD09), *C. asumaniae**3 (CasAT09), *C. oreoreticus**1 (CorVR09), *C. oreoreticus**2 (CorVD09), *C. cartwrightianus**1 (CcwBD09), *C. cartwrightianus**2 (CcrCR09), while *C. hadriaticus* (lane 21) is a white flowering accession (ChdWD09). Arrow indicates the amplified products of expected sizes. On either side of the agarose gel (2%) is 5µl DNA length markers Q-Step 2 (M1) and HyperLadder I (M2).

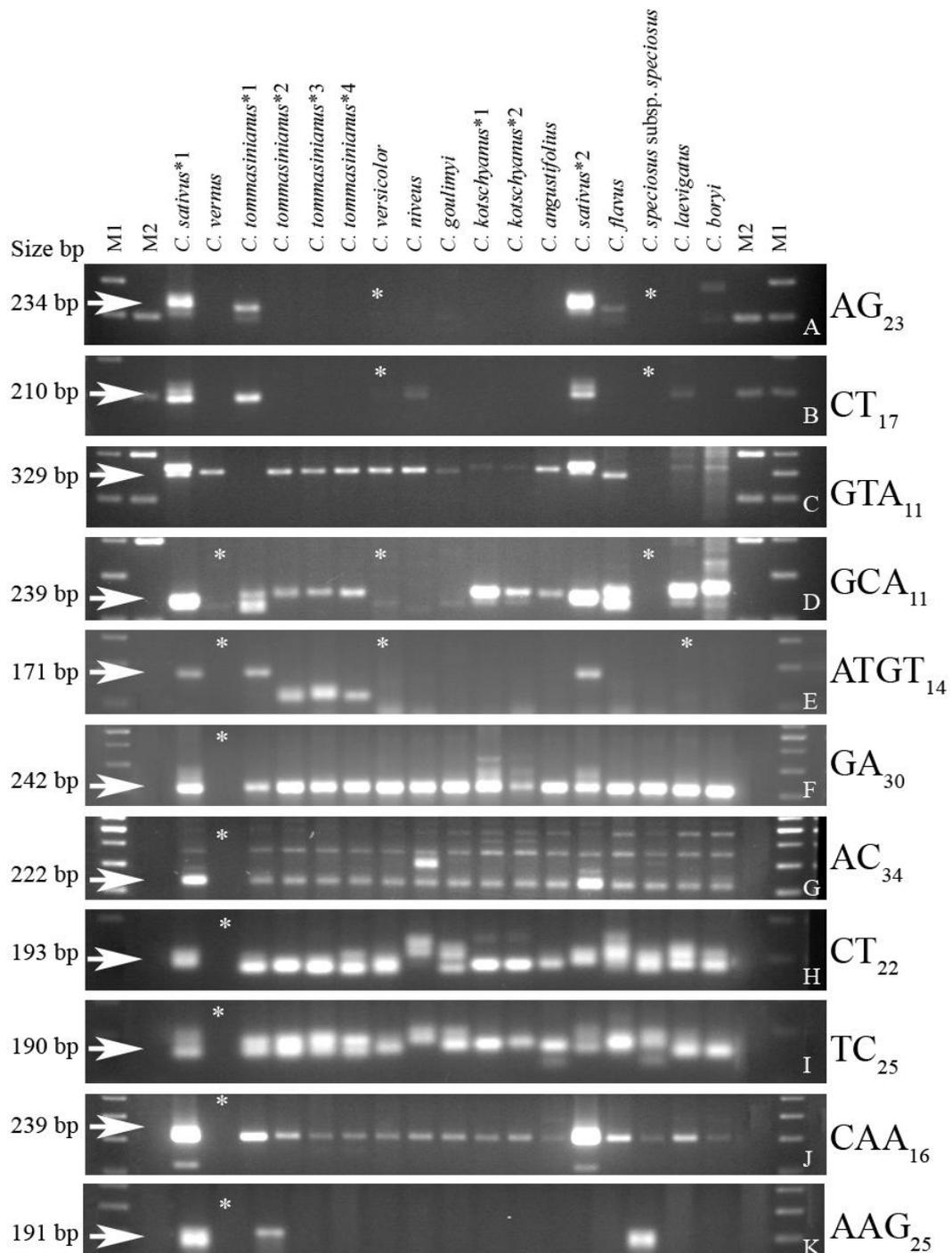


Figure 4.3: PCR amplification pattern of EST-SSR markers, A) AG₂₃, B) CT₁₇, C) GTA₁₁, D) GCA₁₁, E) ATGT₁₄, F) GA₃₀, G) AC₃₄, H) CT₂₂, I) TC₂₅, J) CAA₁₆, and K) AAG₂₅, from 17 *Crocus* accessions of 12 *Crocus* species from other *Crocus* series. Name of species is given on the top of every lane, species name followed by asterisks (*) with digits indicating to sub-species or accession. The accessions included *C. sativus**1 (CsatP09), *C. sativus**2 (Cstkf09), *C. tommasinianus**1 (lilac beauty, CtmLD09), *C. tommasinianus**2 (barr purple, CtmBD09), *C. tommasinianus**3 (rubinetta, CtmTD09), *C. tommasinianus**4 (albus, CtmAD09) and *C. kotschyanus**1 (subsp. *kotschyanus*, CkotP09) and *C. kotschyanus**2 (var. *Zonatus*, Ckot/z08). Arrow indicates the amplified product of expected sizes. On either side of the agarose gel (2%) is 5µl DNA length markers Q-Step 2 (M1) and HyperLadder I (M2).

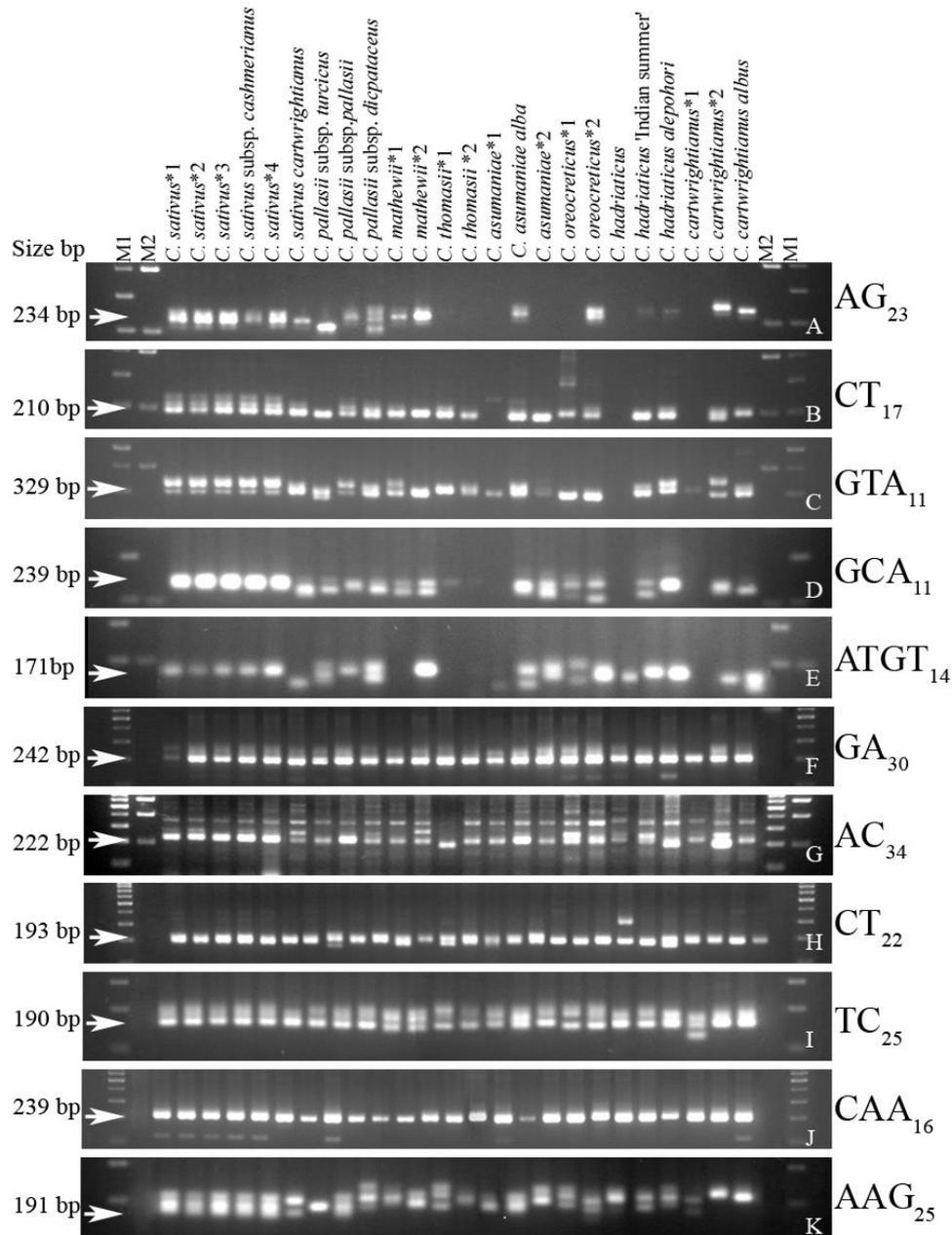


Figure 4.4: PCR amplification pattern of the EST-SSR markers A) AG₂₃, B) CT₁₇, C) GTA₁₁, D) GCA₁₁, E) ATGT₁₄, F) GA₃₀, G) AC₃₄, H) CT₂₂, I) TC₂₅, J) CAA₁₆, and K) AAG₂₅ from 24 *Crocus* accessions of 9 *Crocus* species series *Crocus*. Species name is given on the top of every lane, species name followed by asterisks (*) with digits indicates sub-species or accession. The accessions included *C. sativus**1 (CsatP09), *C. sativus**2 (Cstkf09), *C. sativus**3 (CstCD09), *C. sativus**4 (CsatP09), *C. mathewii**1 (CmatD08), *C. mathewii**2 (HKEP.9291, CmtHR09), *C. thomasii**1 (CtmVD09), *C. thomasii**2 (MS978, CtomI09), *C. asumaniae**1 (CasWD09), *C. asumaniae**2 (CasAD09), *C. asumaniae**3 (CasAT09), *C. oreocreticus**1 (CorVR09), *C. oreocreticus**2 (CorVD09), *C. cartwrightianus**1 (CcwBD09), *C. cartwrightianus**2 (CcrCR09), while *C. hadriaticus* (lane 21) is a white flowering accession (ChdWD09). Arrow indicates the amplified products of expected sizes. On either side of the agarose gel (2%) is 5µl DNA length markers Q-Step 2 (M1) and HyperLadder I (M2).

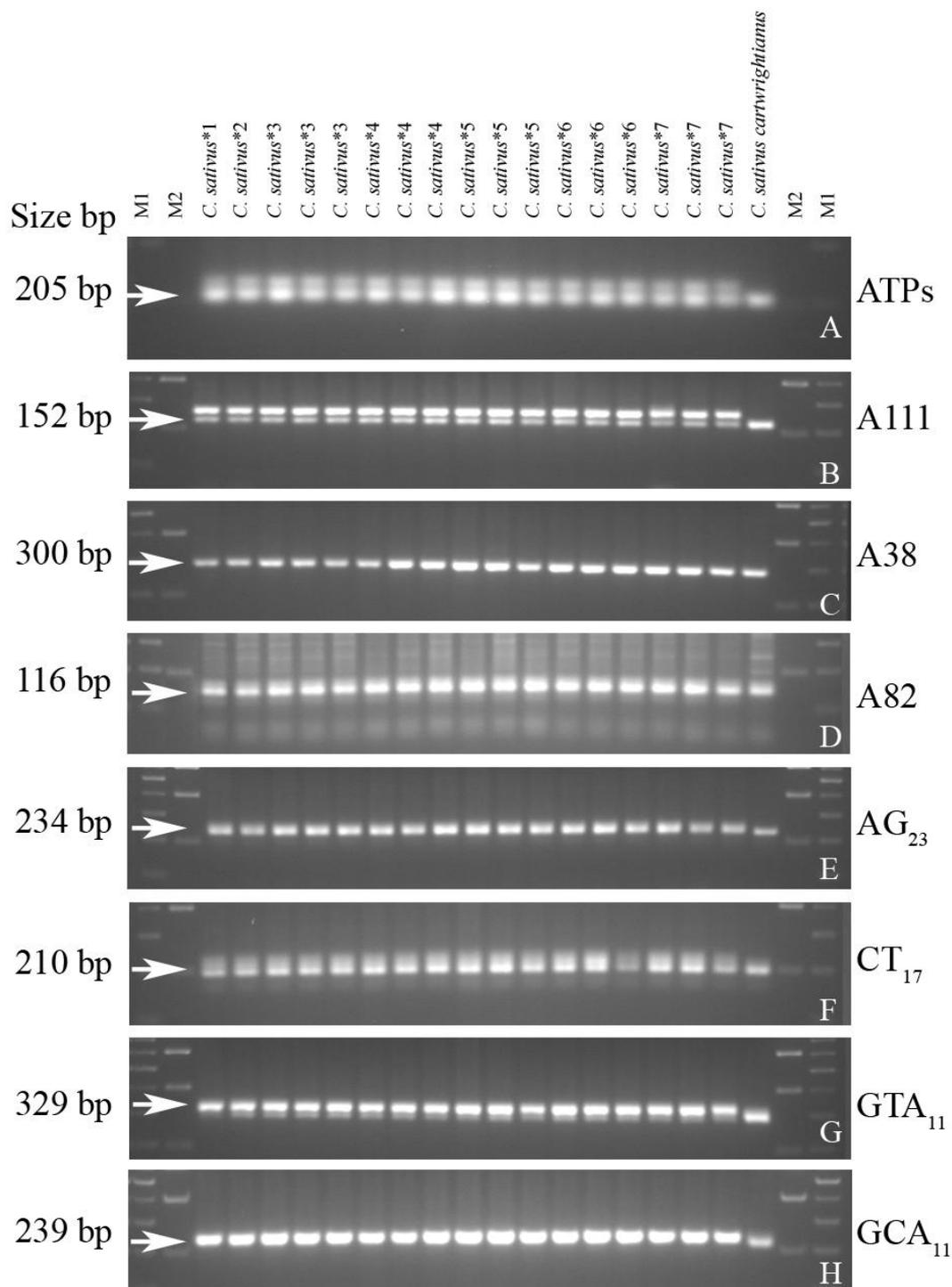


Figure 4.5: PCR amplification pattern of the SNP markers ATPs (A), A111 (B), A38 (C), A82 (D) and EST-SSR markers AG₂₃ (E), CT₁₇ (F), GTA₁₁ (G) and GCA₁₁ (H) from 17 saffron accessions. *1: JW Dix Export, Netherland (2007), *2: Pottertons Nursery, UK, *3: JW Dix Export, Netherland (2010), *4: CrocusBank, Spain, *5: Suttons Nursery, UK, *6: Kashmir, India, *7: Var. *cashmeriensis* from JW Dix Export, Netherland (2009). Hash symbol (#) is *C. sativus cartwrightianus* this accession was purchased under this unrecognized name and was run as a control; Arrow indicates the amplified product of expected sizes. On either side of the agarose gel (2%) is 5µl DNA length markers Q-Step 2 (M1) and HyperLadder I (M2).

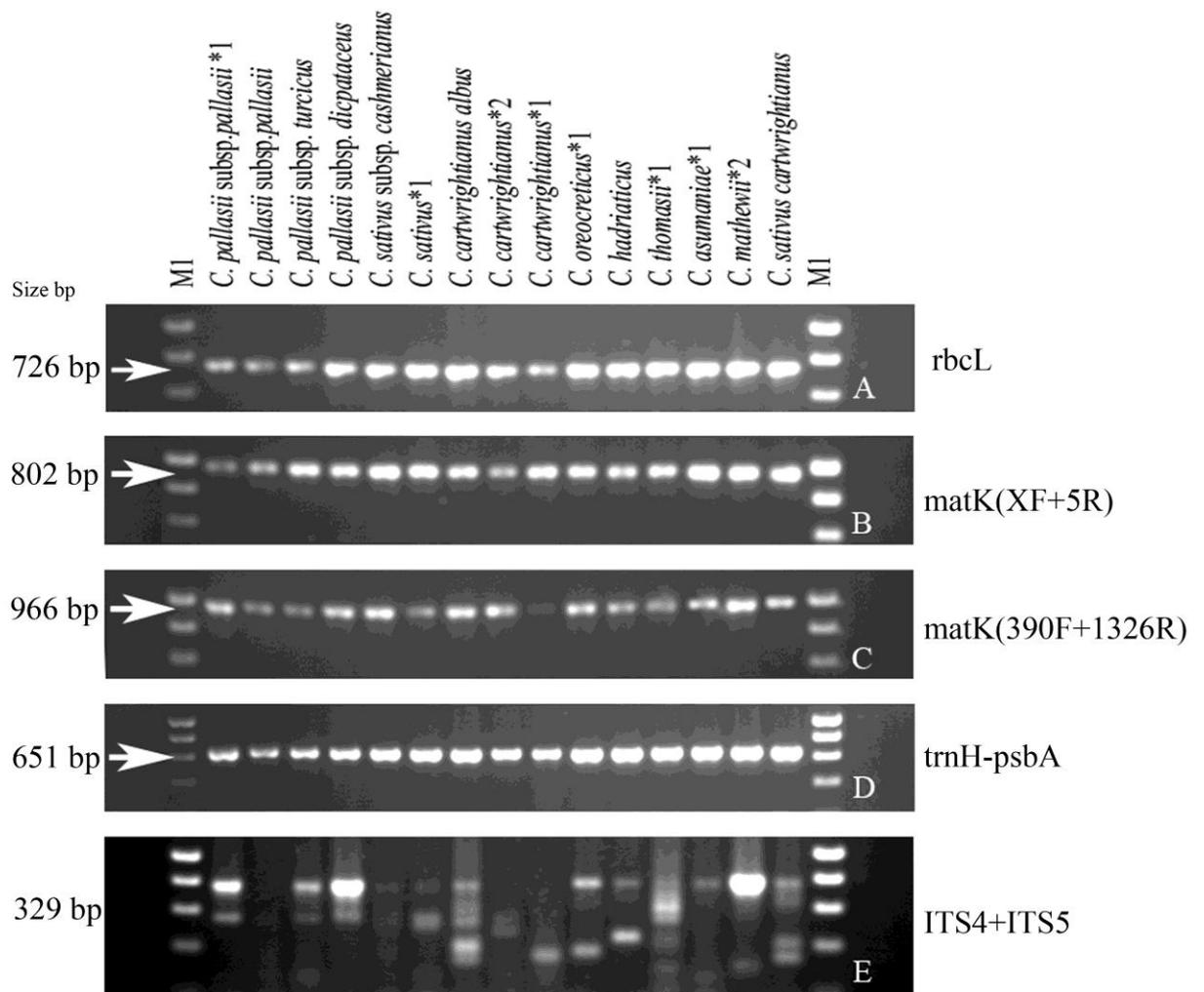


Figure 4.6: PCR amplification pattern of the five barcoding gene A) *rbcL*, B) *matK* XF+5R, C) *matK* 390F+1326R, D) *trnH-psbA* and E) ITS 4+ITS 5 markers from 15 *Crocus* accessions (9 species) series *Crocus*. *1: *C. pallasii* subsp. *pallasii* (CplR09) is an accession from Rareplants nursery, UK and used here only. On either side of the agarose gel (2%) is 5µl DNA length markers Q-Step 2 (M1).

4.3.2 Nucleotide sequence variation and phylogenetic tree of the ATPs region from saffron accession and *Crocus* species

Clear bands of ATPs were cloned and sequenced from 43 *Crocus* accessions except *C. vernus* (Table 2.2). For both primers at least 5-10 colonies were sequenced per accession. A total of 230 clones were sequenced for ATPs, out of 230 clones, 107 were not included in the analysis because they were either too short or having 100% identity to another clone of the same accession already included in the analysis (see Appendix 4). Thus the final ATPs analysis included 123 sequences.

Sequencing of the ATPs marker (arrow in Figure 4.1A, 4.2A) resulted in sequences of 3 lengths that ranged from 202bp to 226bp in size. Overall, 11 sequences were of 202bp length, 90 of 205bp and 21 of 226bp. The aligned sequence has insertions or deletions (indels) of 1bp - 6bp and conserved pattern of SNPs, dividing all accessions into three major alleles and eight sequence groups (Figure 4.8B). The consensus ATPs sequence is an AT rich having ~44.4% GC content and when BLASTN searched, it show 84% and 83% homology to the ATPs sequence of vacuolar ATP synthase subunit B family protein coding DNA sequences (CDS) as annotated in *Brachypodium distachyon* (XM_003569618.1) and *Populus trichocarpa* (XM_006379232.1) mRNA complete CDS respectively.

To assess further internal organization of the *Crocus* ATP synthase fragment, these 123 sequences were assembled *de novo* using Geneious R6 software (parameters including gaps up to 5 and word length 10 were defaults for “Highest Sensitivity/Slow” mode) Although an unconventional use of the assembly algorithm, this proved valuable in grouping the most similar sequences and discounting random nucleotide variations. The groups were closely similar to those identified by inspection of the sequences (Figure 4.8B). Based on sequence alignment, the 123 ATPs could be classified into five groups, while the *de novo* assembly made 8 contigs of these, seven are major variants with 6 to 41 sequences, one represented twice and one, a hybrid of two other sequences, was ungrouped (Table 4.2). With 30 cycles of amplification, Taq polymerase errors are typically 1 in 1000 (*i.e.* one occurrence in 5 of the 200bp sequences here, or one in every 8 nucleotide positions of the 123 sequences), so single occurrences of polymorphisms were largely discounted in the analysis of results below.

The top eight sequences (serial 1 to 8) of the group 1 and 2 (Figure 4.8B) are 205bp long and do not include any *C. sativus* sequences. These eight sequences share six SNPs with group 3 (C at 76th and 127th, T at 89th and 170th, A at 107th, G at 208th) and can be divided

into two groups based on their SNPs. The first group, from *C. pallasii* to *C. asumaniae* (serial 1 to 4) have T at 28th, G at 88th and 110th, T at 124th and 130th, G at 150rd, and C at 203th base pair position. While, the second group *C. oreoreticus* to *C. hadriaticus* (serial 5 to 8), has unique SNPs such as A at 55th and 96th, C at 130rd and 167th base pair position (Figure 4.8B). These eight sequences are clustering close to one another (clade A, Figure 4.8A) and in the *de novo* assembly they are placed in contig 7 and 8 (Table 4.2). Sequences of the group 3 (serial 9 to 29) are 226bp long and have 34 SNPs across the length of the sequence. Further, this group is represented by sequences of *C. sativus* as well as accessions of both section *Crocus* and *Nudiscapus* and (Figure 4.8B). These 21 sequences (clade B) are divided into contig 4 and 5 by *de novo* assembly each having 10 ATPs sequences each, while the *C. sativus cartwrightianus* sequence is separate, due to a single nucleotide mutation G at position 162 (Figure 4.8B). Interestingly, in the phylogenetic tree too, all these group 3 sequences are together, but *C. sativus cartwrightianus* sequence is out of the group (clade B, Figure 4.8A).

The 4th group contains 53 sequences; this group starts with *C. flavus* and ends with *C. sativus* sequences (serial 30 to 82). Sequences in this group are of three types: most of the sequences contain both T and C SNPs at the 28th and 52nd base pair position (Figure 4.8B), while six sequences contain only the C nucleotide at the 52nd base pair position. Similarly, 11 sequences within this group are 202bp and have 3bp deletion, while *C. mathewii*, *C. thomasii* and *C. pallasii* sequences are 209bp with 4bp addition (Figure 4.8B). All these sequences are grouped into clade C (Figure 4.8A), and *de novo* assembly placed sequences of this group in contigs 2, 3 and 6 (Table 4.2). The last two sequences of *C. niveus* and *C. sativus* in this group (serial 81 to 82) are recombinant, having similarity to both group 4 and 5; Such hybrid or recombinant sequences may be attributed to chromosomal recombination events (Figure 4.8B). Further, the *C. flavus* sequence is out of the clade C (Figure 4.8A).

The last 41 sequences of group 5 (serial 83 to 121) have C, C, A, G and T SNPs at 46th, 135th, 145th, 151th and 182nd base pair positions respectively. The *C. asumaniae* (serial 83) sequence is hybrid between group 4 and 5. While the last two sequences of *C. thomasii* and *C. kotschianus* (serial 122 to 123) are missing the last T nucleotide at the 185th base pair position (Figure 4.8B). All these sequences are assembled into contig 1 by the *de novo* assembly analysis (Table 4.2) and clade D in the phylogenetic tree (Figure 4.8A).

In summary, the results from the ATP synthase sequence analysis show that each of the four groups A to D included sequences from both series *Crocus* and series *Nudiscapus*; many sequence variants were shared between multiple species. The occurrence of *C. sativus*

sequences in groups B to D (Figures 4.8A), indicated that the ATP synthase gene is present in more than one copy.

The ATPs sequences of saffron were aligned separately and it revealed three alleles based on unique SNPs, which defines each allele type (Figure 4.7A). Among the 30 ATPs sequences of saffron accessions, 13 sequences (serial 1-13) had two SNPs T and C nucleotide at 28th and 52nd base pair position respectively. While the second allele type (serial 14-21) has five SNPs C, C, A, G and T nucleotide at 46th, 134th, 144th, 151st, 178th bp position respectively. The third allele type (serial 22-30) has both SNPs as well as of 1 to 6bp insertions/addition across the sequence length. The twelve SNPs in the third allele type are C, T, C, A, A, C, C, C, T, A, C and G at 76th, 89th, 102nd, 104th, 107th, 112th, 114th, 127th, 170th, 180th, 183rd and 204th bp position respectively (Figure 4.7). It is interesting to note, that the third allele type has the maximum number of SNPs as well as all insertions are only confined to the this allele (Figure 4.7B)

Table 4.2: *de novo* sequence assembly statistics of the 123 *Crocus* ATPs sequences.

Statistics	Unused reads	All contigs	Contigs >=100bp
Number of	1	8	8
Min Length (bp)	220	205	205
Median Length (bp)		208	208
Mean Length (bp)	220	211	211
Max Length (bp)	220	226	226
N50 Length (bp)		209	209
Number of contigs >= N50		4	4
Length Sum (bp)	220	1,692	1,692

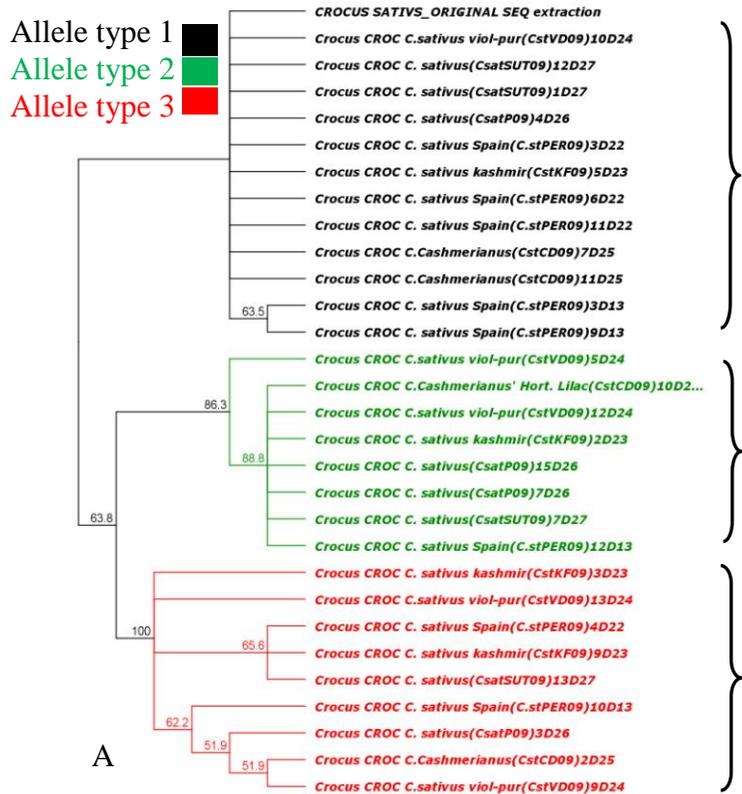
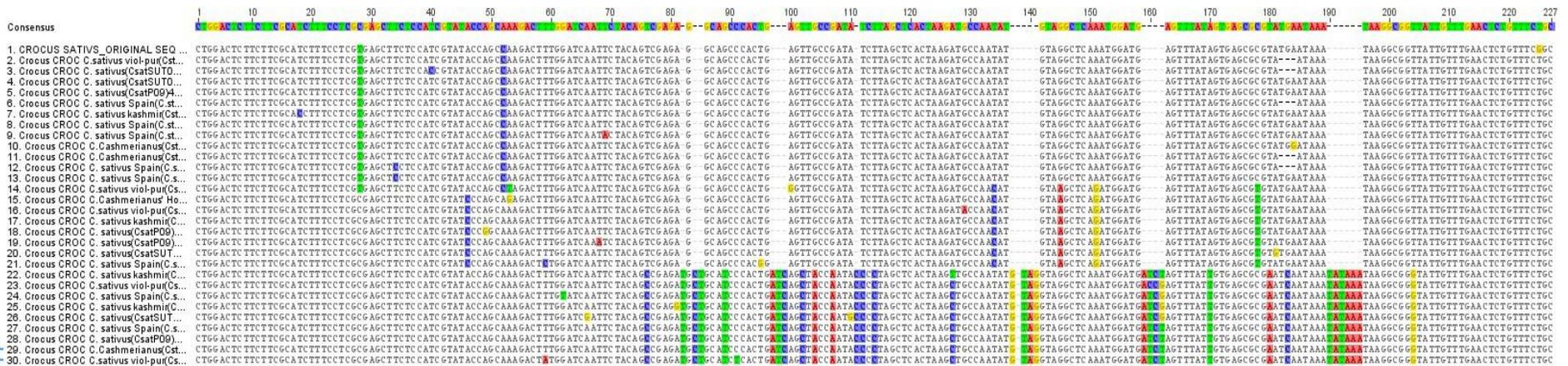


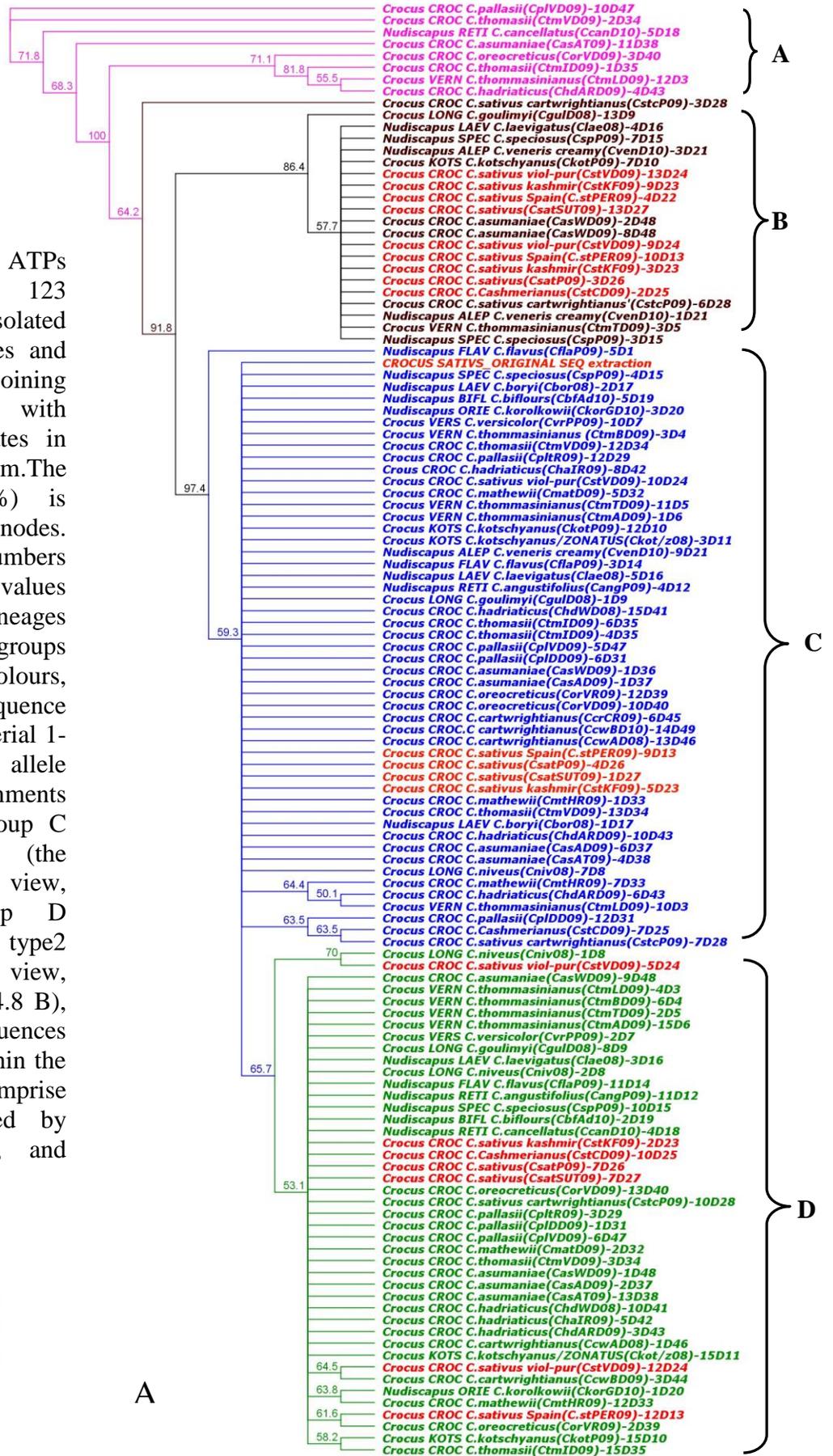
Figure 4.7: A) Molecular phylogenetic analysis of ATPs sequences by Neighbour Joining (NJ) method with 1000 bootstrap replicates using the Geneious R6 program. *Crocus sativus* accessions show clear separation into three clades given in colours, allele type1: contains 13 *C. sativus* accession, allele type2: contains 8 *C. sativus* accession, allele type3: contains 9 *C. sativus* accession.

B) Multiple sequence alignment of the ATPs sequences isolated from 30 accession of *C. sativus* showing three distinct allele types based on SNPs position, allele type1(serial 1-13) has two SNPs T and C nucleotide at 28th and 52nd base pair position respectively, allele type2 (serial 14-21) has five SNPs C, C, A, G and T nucleotide at 46th, 134th, 145th, 151st, 178th bp position respectively, allele type3 (serial 22-30) has twelve SNPs C, T, C, A, A, C, C, T, A, C and G at 76th. 89th, 102nd, 104th, 107th, 112th, 114th, 127th, 170th, 180th, 183rd and 204thbp position respectively. Consensus sequence length is 227bp. Highlighted nucleotides “ATCG” indicating to the single nucleotide polymorphism detected in the species; dashes indicate deletions/ insertion or gaps in the sequence alignments.

B

Figure 4.8: A) ATPs phylogenetic tree of 123 nucleotide sequences isolated from 43 *Crocus* species and accessions. Neighbour-Joining tree was constructed with 1000 bootstrap replicates in Geneious R6 program. The bootstrap support (%) is shown near the nodes. Branches without numbers received bootstrap values smaller than 50%. Lineages divided into four main groups shown in different colours, Group A unresolved sequence alignments (tree view serial 1-8), Group B represent allele type3 (sequence alignments view, serial 9-29), Group C represent allele type1 (the sequence alignments view, serial 30-80), Group D represent allele type2 (sequence alignments view, serial 81-123) (Figure 4.8 B), and the *C. sativus* sequences are indicated in red within the groups. Names comprise section name followed by series, species name, and accession number.

B, Allele type 3
 C, Allele type 1
 D, Allele type 2



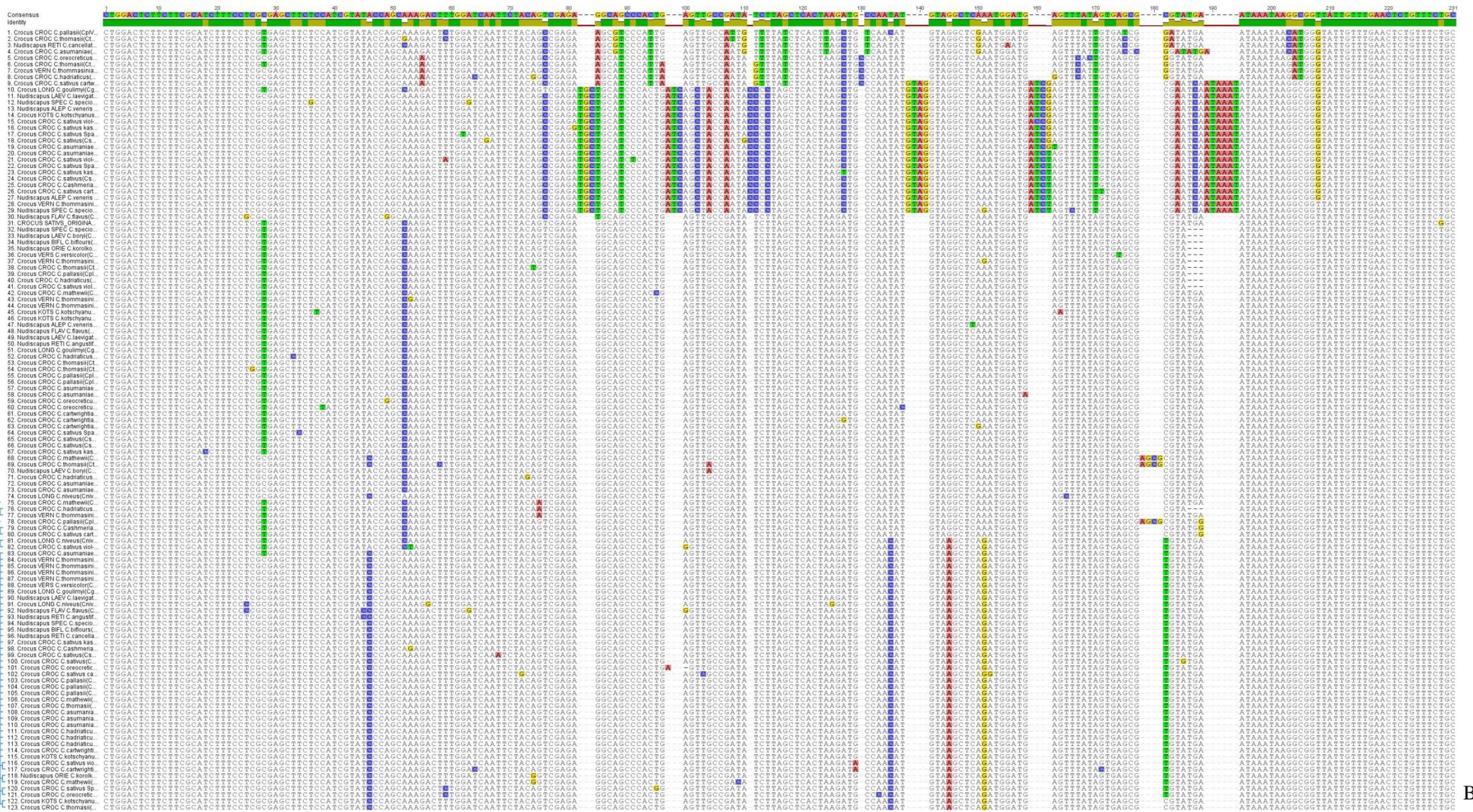


Figure 4.8: B) Multiple sequence alignment view of the ATPs sequences isolated from 30 accessions of *C. sativus*. The nucleotide sequences were aligned using the Geneious R6 program, Showing three sequence allele types, allele type1(serial 30-80) had two SNPs T and C nucleotide at 28th and 52nd base pair position respectively, allele type2 (serial 81-123) has five SNPs C, C, A, G and T nucleotide at 46th, 135th, 145th, 151th and 182nd bp position respectively, allele type3 (serial 9-29) has twelve SNPs C, T, C, A, A, C, C, T, A, C and G at 76th, 89th, 102nd, 104th, 107th, 112th, 114th, 127th, 170th, 180th, 183rd and 204th bp position respectively. The sequences presented in (A, serial 1-8) do not group with any of the three type of alleles see (Figure 4.8 A). The consensus sequence length is 227bp. Highlighted nucleotides “ATCG” indicating to the single nucleotide polymorphism detected in the species; dashes indicate deletions/ insertion or gaps in the sequence alignments.

4.3.3 Nucleotide sequence variation and phylogenetic tree of the TC₂₅ region from saffron accession and *Crocus* species

PCR products of the EST-SSR primer TC₂₅ (arrow in Figure 4.3, 4.4) were cloned and sequenced with universal M13 primers for all 43 *Crocus* accessions except *C. vernus* (Table 2.2). TC₂₅ is a compound microsatellite comprising (CA)₇(TC)_{25.5} (-*CTTTAT* CACA CACACACACA TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTCT AGAAGAT-) motif (italics indicate immediately flanking sequence). For each accession, 5-10 clones were sequenced, and a total of 300 sequences ranging from 134bp to 226bp in size were obtained. However, the final analysis included only 107 representative sequences of all *Crocus* accessions, while the remaining 193 sequences were discarded for being too short, too diverse, or because they had 100% homology to another clone of the same accession already included in the analysis. The consensus TC₂₅ sequence is AT rich and is 190bp long. Sequences that are included in the phylogenetic analysis ranged from 155bp to 223bp in size (Figures 4.10B).

The TC₂₅ sequences displayed significant polymorphism, particularly within the microsatellite region. However, for convenience all the 107 sequences were divided into clades E and F (Figures 4.10A, 4.10B), while at the nucleotide level the sequence can be separated into three main alleles based on the presence/absence of CA or TC repeat region (see below and Figure 4.10B and Appendix 4).

Group E (serial 1 to 4) contains 4 sequences of 2 accessions of *C. kotschianus* divided into two sister branches and one sub branch with (bootstrap support 87.1). Sequences of 175bp length remained out of group (group E, Figure 4.10A). While, Group F contain all the remaining species (serial 5 to 107) well supported (bootstrap 100%) and rooted with the four *C. kotschianus* sequences (group E) and divided into four sub-clades. Group F1 contains 41 sequences (serial 5 to 45) of 151-200bp long. The sequences, with poor bootstrap resolution contain accessions from both section *Nudiscapus* and *Crocus*, showing the relationships and presence of shared alleles within the genus. For example *C. cancellatus* sequences are grouping with *C. cartwrightianus* (serial 22, 23 and serial 24, 25). Similarly, *C. asumaniae* shares a branch with *C. versicolor* at serials (20 and 21) and another at serials 35 and 36 (see Figure 4.10A); and *C. laevigatus* with *C. boryi* (*C. pallasii* subsp. *dispathaceus* (serial 26 to 30). Although several accessions of *C. asumaniae* and *C. tommasinianus* share

one type of sequence and more than one allele, sequences from these accessions were not resolved into clusters (group F1, Figure 4.10A). Moreover, some 45 sequences from F1 had T-rich stretches without a clear SSR motif. The well-supported clade F2 contains 12 sequences of 191-198bp long from 9 different species (serial 46 to 57). These sequences do not contain the CA motif (Figure 4.10B). The sequences from five species (serial 51 to 57), were nearly identical at the nucleotide level (bootstrap support 79.5; *C. thomasi*, *C. cartwrightianus*, *C. asumaniae*, *C. sativus* and *C. hadriaticus*) with additional variation in *C. biflorus*, *C. korolkowii*, *C. oreoreticus* and *C. goulimyi*. While, 10 sequences had a TC SSR motif with no significant CA repeats, and most included some 3' T and A bases (Figure 4.10B).

Sub-clade F2 includes sequences from 9 species (*C. mathewii*, *C. pallasii*, *C. oreoreticus*, *C. sativus*, *C. cartwrightianus*, *C. asumaniae*, *C. hadriaticus* and *C. thomasi* from Series Crocus, *C. tommasinianus* from series Verni, as well as “*C. sativus cartwrightianus*” (the garden-named non-species). Sub-clade F3 contains 24 sequences of 161-163bp long, having a single base pair insertion or deletion (serial 58 to 81). These sequences contain the SSR sequence related to the reference sequence, *-CTTTAT CACACCAC TCTTTTCTT TTCTCTCTC CTTTTGCTAT AGAAGAT-*, but are without conspicuous SSR motifs (Figure 4.12, 4.10A; bootstrap support 68.1). Sub-clade F4 (serial 82 to 107) contains 25 sequences along with the reference TC₂₅ sequence for which the primer pairs were designed (Figure 4.10A, bootstrap support 73.8). These sequences are 168-202bp long containing both CA and TC repeat region with between 10 and 30 TC repeats including some degeneracy. The reference TC₂₅ sequence (of *C. sativus*) was 190bp length size, and an identical sequence was amplified from *C. mathewii* and *C. pallasii* subsp. *pallasii* only (Figure 4.12, 4.10B).

A second round of analysis of the flanking sequence with the microsatellite domain deleted in the nucleotide alignment was carried out. The analysis included SNPs in the 95bp 5' region and the 50bp at the 3' region flanking the SSR domain; the grouping of accessions and species in the phylogenetic tree was broadly similar to that including the microsatellite region, and showed no well-supported groups with sequences from only a few species. Unlike, ATPs and the published but undescribed *Crocus* ESTs (see discussion below), contig assembly was not useful because of the microsatellite region, and did not relate to any natural

groupings (Figure 4.10B). Interestingly, the TC₂₅ results also revealed the existence of 3 distinct alleles of 163bp; 190bp and 198bp in *C. sativus* (see Figures 4.9B, 4.10B).

A large study of ESTs from *C. sativus* (Giuliano *et al.*, 2008) includes 6,603 ESTs, which his algorithm assembled into 1,893 clusters, “each corresponding to a different expressed gene”. I used the total of 6,908 ESTs present in GenBank/EBI database and carried out a similar assembly using Geneious but with some relaxed parameters so homologous sequences were grouped. It also had the effect of making more extended contigs with more sequences aligned, a total of 6,180 out of 6,908 reads were assembled to produce 767 contigs and 728 reads were not assembled (Table 4.3). Giuliano *et al.* (2008) makes no mention of evidence of multiple copies or whole genome duplication, but in about 10% of these, there was evidence for three different alleles of the same gene (Figure 4.11).

Table 4.3: *De novo* sequence assembly statistics of the 6,908 *Crocus* EST sequences.

Statistics	Unused reads	All contigs	Contigs >=100bp	Contigs >=1000bp
Number of	728	767	767	48
Min Length (bp)	100	100	100	1,002
Median Length (bp)		532	532	1,196
Mean Length (bp)	425	564	564	1,351
Max Length (bp)	2,149	3,141	3,141	3,141
N50 Length (bp)		631	631	1,284
Number of contigs >= N50		252	252	19
Length Sum (bp)	309,803	432,837	432,837	64,886

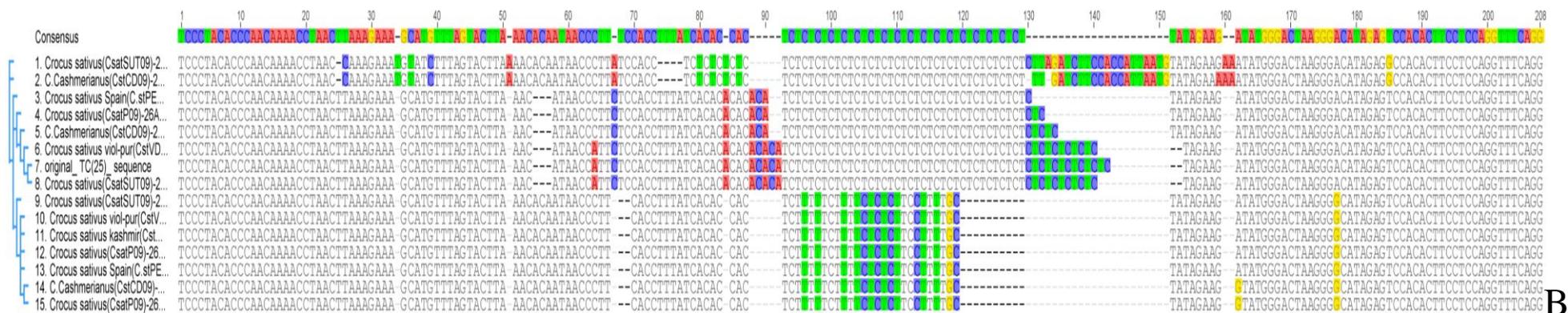
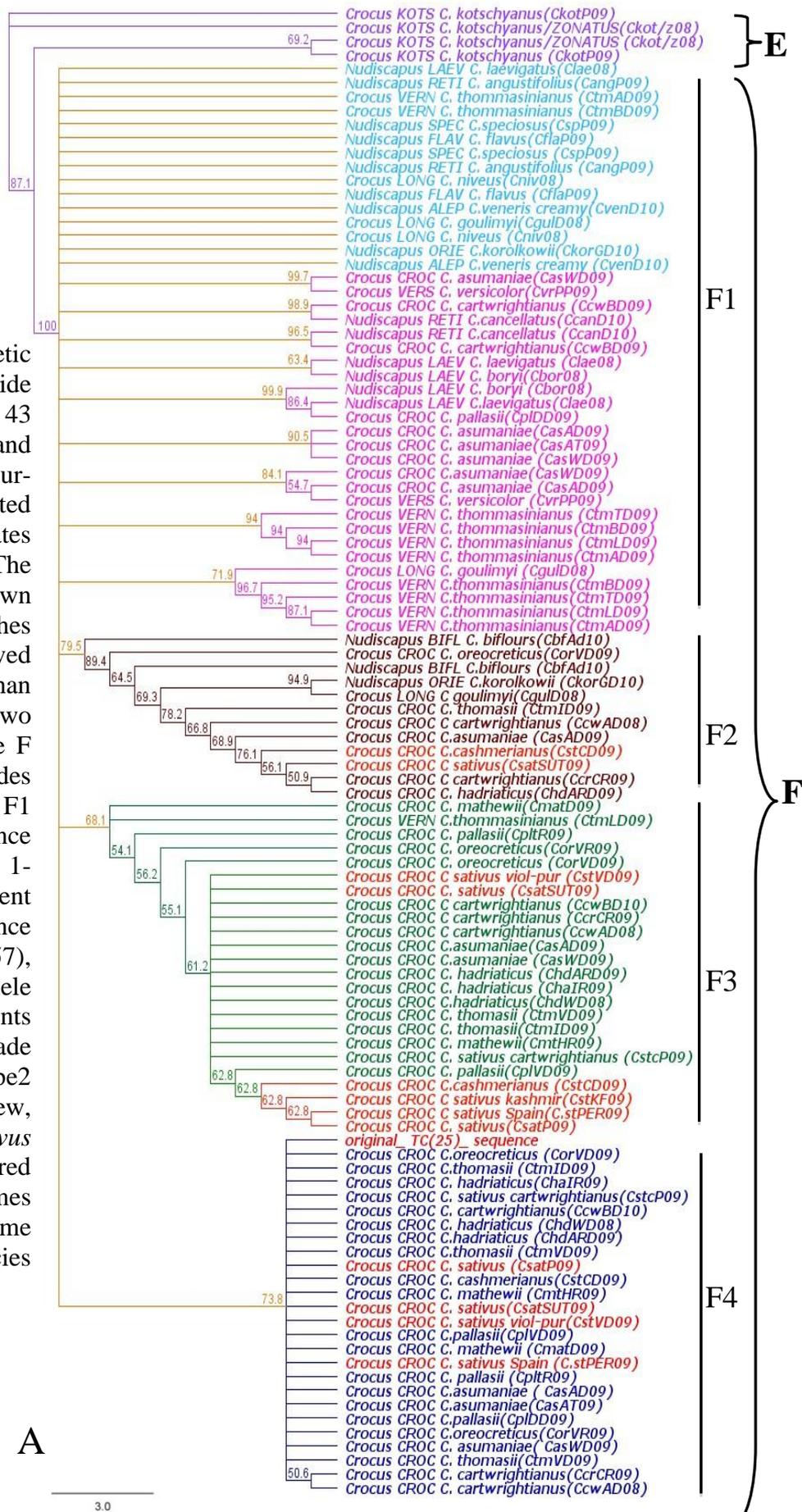


Figure 4.9: A) Molecular phylogenetic analysis of the TC₂₅ sequence of the *C. sativus* accession. Neighbour Joining (NJ) method with 1000 bootstrap replicates using the Geneious R6 program. *Crocus sativus* accessions show clear separation into three clades based on the TC₂ microsatellite motifs, allele type1: contains two accessions of *C. sativus*, allele type2: contains six accessions of *C. sativus*, allele type3: contains seven accessions of *C. sativus*.

B) Multiple sequence alignment view of the TC₂₅ sequences isolated from fifteen *C. sativus* accessions showing three sequence allele types, allele type1 had two sequences (serial 1-2) these sequences do not contain the CA motif, allele type2 (serial 3-8) had the CA motif and SSR (TC₂₅) repeats, allele type3 (serial 9-15) had the CA motif and T-rich stretches without a clear SSR motif. The consensus sequence length is 208bp. Highlighted nucleotides “ATCG” indicating to the single nucleotide polymorphism detected in the species and the microsatellite region; dashes indicate deletions/ insertion or gaps in the sequence alignment.

Figure 4.10: A) Phylogenetic tree of 107 TC₂₅ nucleotide sequences isolated from 43 accessions of saffron and *Crocus* species. Neighbour-Joining tree was constructed with 1000 bootstrap replicates in Geneious R6 program. The bootstrap support (%) is shown near the nodes. Branches without numbers received bootstrap values smaller than 50%. Lineages divided into two main groups E and F, clade F divided in three sub-clades shown in different colours, F1 unresolved sequence alignments (tree view series 1-49), sub-clade F2 represent allele type1 (sequence alignment view, serial 49-57), sub-clade F3 represent allele type3 (the sequence alignments view, serial 58-81), sub-clade F4 represent allele type2 (sequence alignment view, serial 82-107), the *C. sativus* sequences are indicated in red within the groups. Names comprise section name followed by series, species name, accession number.



4.3.4 Nucleotide sequence variation and phylogenetic tree of the barcoding genes

Species and accessions of the *Crocus* series *Crocus* were used to test universality of the *matK*, *trnH*, *rbcL* and ITS primers (Table 4.1). The success levels for *matK* XF, *matK* 390F, *trnH*, *rbcL* and ITS primers were 100% and all five primers amplified regions of the expected sizes (see below and Figure 4.6). The ITS primers amplified multiple copies within individuals (Figure 4.6) and products of the ITS were not sequenced. Examination of sequence quality and coverage indicated that *matK* 390F, *matK* XF, *trnH* and *rbcL* generated high quality sequences and are described below.

4.3.4.1 *matK* 390F+1326R plastid gene sequence

Seventeen (17) sequences ranging from 802-965bp length of were aligned and used in the analysis. *C. moabiticus* and *C. banaticus* sequences (EU497045 and EU496995) were downloaded from NCBI. These sequences contained 929 (96.2%) identical sites and pairwise identity was 99.4%. The entire sequence contains 6 SNPs dispersed between 220-839bp regions at 220, 252, 356, 531, 540 and 839th base pair position. Also *C. cartwrightianus* sequence has a 6bp insertion from 673-679bp (see Appendix 4, Figure A4.7).

4.3.4.2 *matK* XF+5R plastid gene sequence

Seventeen sequences ranging from 802-942bp along with *C. moabiticus* and *C. banaticus* (EU497045 and EU496995) sequences downloaded from NCBI were applied in the analysis. The sequences contained 903 (95.9%) identical sites while pairwise identity was 99.3%. The region 210-2550bp is SNP rich region containing 11 SNPs at 214, 246, 308, 350, 412, 426, 455, 525 and 534,827,862th base pair position (see Appendix 4, Figure A4.8).

4.3.4.3 *trnH* plastid gene sequence

The analysis involved seventeen sequences ranging from 605-650bp along two sequences of *C. moabiticus* and *C. banaticus* (EU110227 and EU110175) from NCBI. The sequences contained 635 (97.5%) identical sites and pairwise identity was 99.3%. the SNP

rich region is located between 500-520bp containing 6 SNPs at 505, 506, 511, 512, 517 and 518th base pair position. *C. moabiticus* also has a 6bp insertion between 560-565bp (see Appendix 4, Figure A4.10).

4.3.4.4 *rbcL* plastid gene sequence

The *rbcL* analysis included sixteen sequences ranging from 700-726bp. The sequence of *C. banaticus* (JX903213) was downloaded from NCBI, while *C. moabiticus* sequence has not been submitted to the database yet. The *rbcL* region contained 6 SNPs between 1-670bp at 5, 35, 446, 479, 575, 608 and 666th base pair position (see Appendix 4, Figure A4.9).

4.3.4.5 Phylogenetic analysis

By and large, all sequences either amplified or downloaded from the NCBI were very much identical except for the single nucleotide polymorphism, discriminating different species and accession (Figure 4.13). *C. banaticus*, a member of the subgenus *Crociris* was used as the outgroup here. Levels of species discrimination based on the sequences from *matK*, *trnH* and *rbcL* were carried out alone, as well as by combining sequences of all four regions in head to tail orientation (*matK* 390F, *matK* XF, *trnH* and *rbcL* respectively). Species and accession discriminating power in the phylogenetic trees constructed using *matK* 390F, *matK* XF, *trnH*, *rbcL* sequences alone, were not very different from that of the combined sequence. Rather, various combinations of all four loci in the composite tree were more powerful in resolving and differentiating between species and accessions than either of the loci individually (Figure 4.13).

The phylogenetic tree is divided into two clades G and H (each sister to two unresolved single-sequence branches); Clade G contains one species, *C. mathweii*, along three accessions of *C. pallasii*. Similarly, *C. banaticus* is present as a sister branch with the purple flowering accession of *C. cartwrightianus* (CcWBD10), while *C. asumaniae* and *C. moabiticus* group with clade G and H, having 76.5% nodal support. However, they are more closely related to clade G than to clade H (see Clade G, Figure 4.13). The chromatogram of *C. asumaniae* with *matK* 390F+1326R and *rbcL* revealed diversity in nucleotide sequence (see Appendix 4, Figure A4.6) and could be the potential ancestral species of the whole group or it itself may be a polyploid. In clade H, six closely related

species of *C. sativus* are grouped together. It was interesting that *C. sativus* accessions *i.e.* the Spanish and *Cashmirianus* (*C. sativus* accession) group separately and the *C. cartwrightianus* accession (CcrCR09) appeared to be the closest species to the *C. sativus* accession from Spain. Further, this accession could be the potential maternal parent of the Spanish saffron (Clade H, Figure 4.13). Morphologically *C. cartwrightianus* var. *albus* and *C. sativus cartwrightianus* (the garden-named non-species) are similar (Figure 1.2) and in the phylogenetic tree of the barcoding genes are grouped together in one sub-clade with 82.1% nodal support. The ATPs and TC₂₅ results were also similar, and both species might be one, or *C. sativus cartwrightianus* might be a hybrid between *C. cartwrightianus* var. *albus* and *C. sativus*. Further, *C. hadriaticus* and *C. thomasi* group together having weak support with 50.3% nodal support. Similarly, all *C. pallasii* accessions clustered together with 98.2% nodal support, but *C. pallasii* (cplVD09) grouped with *C. sativus* (Clade H, Figure 4.13). Nevertheless, the barcoding genes do not contradict the allotriploid nature of *C. sativus* and the results of ATPs and TC₂₅, and indicated that one allele in *C. sativus* is most probably originating from *C. pallasii* (Clade G, Figure 4.13) and a second allele is from *C. cartwrightianus* (Clade H, Figure 4.13).



Figure 4.13: Molecular phylogenetic analysis of four barcoding genes (combined), *matK* XF, *matK* 390F, *trnH*, *rbcL*. Neighbour-Joining tree was constructed with 1000 bootstrap replicates in Geneious R 6. Consensus percentages are given at each node using Tamura 3-parameter model to calculate genetic distance. The analysis included 17 nucleotide sequences from *Crocus* series *Crocus* separated into two main lineages shown by different colours, clade G and clade H cluster with three sister branches and *C. banaticus* used as an outgroup from subgenus *Crociris*. *C. moabiticus* and *C. banaticus* sequences were downloaded from NCBI. The *C. sativus* sequences are indicated in red within the groups. Names above comprise section name followed by series, species name and unique accession identification number (Table 2.2).

4.4 Discussion

The PCR primers developed here, as well as ‘universal primers’ used for phylogenetic studies, were successful in amplifying genomic DNA fragments of the expected sizes from the species of *Crocus* tested from series of the section *Crocus* and *Nudiscapus* (Figures 4.1 to 4.6). Indeed, all but one of the 20 primer pairs (Table 4.1) tested amplified sequences of the correct length from at least one accession of every species. The conservation of the primers and amplification confirms the close phylogenetic relationships of these species, despite of their relatively diverse morphology (Figure 1.4& Figure1.2) and the taxonomic definition of the individual species (Jacobsen and Ørgaard, 2004; Petersen *et al.*, 2008; Seberg and Petersen, 2009).

Genetic diversity in *C. sativus* (saffron) is limited or largely unknown. Further, lack of pollination coupled with homoeologous recombination adds further to the existing genetic bottlenecks (Nehvi *et al.* 2007; Rubio-Moraga *et al.*, 2009; Fernandez *et al.*, 2011). Therefore, the current study was designed to capture interspecific variation particularly within saffron accessions, and 10 species included more than one accession (from 2 to 5). A number of studies based on morphological traits and molecular markers have been carried out over recent years to understand diversity in saffron and clarify *Crocus* phylogeny (see Chapters I, III and below). However, little insight has been gained so far on this front. Perhaps, the first extensive study was carried out by Petersen *et al.* (2008, also see Figure 1.8), who sequenced five plastid genes from 86 recognized species of the genus *Crocus*. To add further to existing knowledge, here 20 primer pairs including 5 universal barcodes and 15 other EST-SSR and SNP markers were applied and products of 6 primers were sequenced with multiple accessions of many of the species used.

SSR markers are highly polymorphic and transferable across species and even genera, while ESTs are very informative in gene tracking, but the conserved nature of genes in related species results in low levels of polymorphism (Gao *et al.*, 2004; Xue *et al.*, 2008). To see maximum polymorphism, over six thousand publicly available ESTs of *C. sativus* (D’Agostino *et al.*, 2007) were *in silico* screened for SSRs and applied along with SNP markers (Table 4.1). The overall results highlight the possibilities of using molecular approaches to find ancestors of saffron and to determine relationships

between *C. sativus*, and other species within and outside of the section *Crocus* of the genus. The results with the EST-SSR, SNPs and universally proposed barcode primers in this chapter complement the results obtained with IRAP markers (see Chapter III) in discriminating and establishing genetic relationships in *Crocus* species. Further, the current work adds information to identify polymorphisms, define genome relationships, and interpret phylogeny and ancestry of the genus (Figures 4.7-4.13 and below).

4.4.1 ATPs sequence variation

PCR results indicate that ATPs (ATP synthase, EC 3.6.3.14, making ATP from ADP and inorganic phosphate) sequences are well-conserved and ancient components of the genus *Crocus* as they were amplified from all member of the genus (Figures 4.1A, 4.2A). A total of 123 ATPs sequences were applied to understand the relationships of *C. sativus* with related species, and the phylogenetic analysis indicated the presence of three types of alleles, showing unique SNPs in each allele (Figures 4.7, 4.8 and below). Sequences in clade A, and a few sequences in the clades B (*'C. sativus cartwrightianus'* serial 9), C (*C. flavus*, serial 30) and D (*C. niveus*, *C. sativus* and *C. asumaniae*, serial 81, 82 and 83) were apparently recombinant (Figure 4.8B). These variants were not detected in any of the *C. sativus* clones, but the evolutionary pattern cannot be reconstructed. Sequences in clade B (20 sequences), C (50 sequences) and D (40 sequences) contain members of both section *Crocus* and *Nudiscapus* and have a specific type of allele and sequences of *C. sativus* are represented in these three clades (Figures 4.8A). Sequences in clade B were amplified from nine species, *C. goulimyi*, *C. laevigatus*, *C. speciosus*, *C. veneris*, *C. kotschianus*, *C. sativus*, *C. asumaniae*, *'C. sativus cartwrightianus'* and *C. tommasinianus*. Except *C. asumaniae* and *C. sativus*, belonging to *Crocus* series *Crocus*, the other six species belong to different series of sections *Crocus* and *Nudiscapus* and *'C. sativus cartwrightianus'* is an un-recognized species presumably of garden origin (see Mathew, 1982 and Figure 1.9). The *C. asumaniae* is present as a sub-group with *C. sativus* in Petersen *et al.* (2008) in a clade with strong 96% nodal support. Although *Crocus asumaniae* has 26 chromosomes, still the universal barcode sequence data of *C. asumaniae* also indicated heterozygosity in sequence chromatogram (see Appendix 4 Figure A 4.6 and above).

Nevertheless, grouping of these species in clade B is indicating the origin or shared origin of this allele in *C. sativus* from *C. asumaniae* or another related common ancestor (see Figures 4.8A). Sequences in clade C and D are predominantly (but not exclusively) from section *Crocus*, and the sequence of *C. sativus* reveals intraspecific diversity (see clades C, D). Four out of five *C. sativus* sequences in clade C grouped between *C. cartwrightianus* (CcwAD08) and *C. mathewii* (CmtHR09) and four out of seven *C. sativus* sequences in clade D grouped with *C. cancellatus* (CcanD10) and *C. oreocreticus* (CorVD09). The *C. cartwrightianus*, *C. mathewii* and *C. oreocreticus* all have 16 somatic chromosomes and are among the most likely parents of *C. sativus* (Fernandez, 2006; Petersen *et al.*, 2008; Seberg and Petersen 2009; Harpke *et al.*, 2013), as is supported by the results of this chapter.

Monocots have undergone one to several rounds of whole genome duplication (WGD) events. For example at least two WGDs have taken place prior to the divergence of cereals and other grasses (see Stein, 2007; Jiao *et al.*, 2011). The genome size in *C. sativus* is estimated to be greater than 30,000 Mbp (based on *C. vernus*; $2n=8$; 11,000Mbp), which is about 80 times larger than *Arabidopsis thaliana* or twice the size of the barley genome per haploid genome (Frello *et al.*, 2004; Candan *et al.*, 2009). However, nothing is known about WGD events in the genus *Crocus*, which are now known to have occurred widely several times in the evolution of plant lineages (see D'Hont *et al.*, 2012). The basic chromosome number of around $x=8$ does not give enough evidence for the recent polyploidy in saffron (Harpke *et al.*, 2013). Although these ancient WGD events are nowadays detected by sensitive analysis of predicted protein sequences from whole genome DNA analysis, they in general would not be detected by PCR primer amplification or high-stringency hybridization of probes. Moreover, the completed sequenced genomes have revealed a considerable amount of redundant genes that are attributed to these WGDs (see Soltis *et al.*, 2009). However, the multiple variants of all the nuclear sequence found in many species, and particularly *C. sativus*, suggests that the nuclear sequences and perhaps the ATP synthase gene (Figures 4.8A, 4.8B) is present in more than one copy. It was also important that most species were found on several branches, indicating that the duplication of the sequence has occurred before the separation of sections *Crocus* and *Nudiscapus*, and also that there has been no deletion of the sequence during subsequent speciation.

It is widely believed that diversity is scarce in *C. sativus* due to its autotriploid nature (Brighton 1977; Mathew 1982; Fluch *et al.*, 2009), but the ATP synthase gene analysis was extremely helpful in identifying not only diversity within *C. sativus* accessions, as well as it suggesting the probability of *C. sativus* being an allotriploid rather than to be autotriploid having a single ancestral species. Thus the results shown here go parallel to the assumptions published in few earlier reports (Fernandez, 2006; Grilli Caiola and Canini, 2010). The nuclear ATP synthase gene is well annotated in the poplar genome (Tusken *et al.*, 2006) and is a vacuolar ATP synthase subunit B family protein CDS. In poplar, it appears on chromosomes IV and IX, two chromosomes shown to have extensive duplicated segments. As genome sequencing and mapping is advancing in *Crocus*, this may be an interesting ‘anchor’ to examine genome evolution and duplication events, and test assembly results.

4.4.2 TC₂₅ EST-SSR sequence variation

Microsatellites are a ubiquitous class of simple repetitive DNA present in all eukaryotes and development of SSR markers has provided a key resource for investigating genetic diversity (Heslop-Harrison and Schwarzacher, 2012). Like ATPs, the sequences (-*CTTTAT CACA CACACACACA TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTCT AGAAGAT*-) that are flanking the (CA)₇ (TC)₂₅ repeat motif in the reference *C. sativus* sequence are well conserved across all the 44 species analysed (Figures 4.3I, 4.4I). The nucleotide sequence revealed the presence of several variant copies of (TC)₂₅ within each genome and, the microsatellite motif itself is highly variable, with little correlation to the known species relationships (Figure 4.10B, 4.12). The results here are in agreement with those of Derakhshan *et al.* (2014), where the authors reported amplification of 2-8 alleles per locus using different SSR markers developed from safflower (see Derakhshan *et al.*, 2014).

Phylogenetic analysis of the genus *Crocus* has led to the suggestion that it is not monophyletic (Petersen *et al.*, 2008), although others believe in its monophyletic origin (Harpke *et al.*, 2013). Here too, individual species showed high levels of polymorphism in sequence, and most species were included in several major branches and no well resolved phylogeny related to major branches was evident (Figure 4.10A). The current analysis reveals that (TC)₂₅ sequences are present in more than one copy in each *Crocus*

genome, and the PCR is amplifying homoeologous sites. Notably, all species from series *Crocus* form some reasonably well-resolved branches, and in particular, the *C. sativus* sequences are always related closely to those from *C. pallasii* ssp. *pallasii* and *C. cartwrightianus* (Figure 4.10A). In the phylogenetic reconstruction of the genus *Crocus*, all the 107 (TC)₂₅ sequences were divided into two main groups (clades E, F). In clade F, *C. sativus* sequences are distributed in three sub-clades F2, F3 and F4 with strong nodal support (Figure 4.10A). Although section *Nudiscapus* is inferred as monophyletic, in my analysis of ATP synthetase gene (see above) and TC₂₅, the sequences of section *Nudiscapus* are distributed in more than one clade. However, these results do not contradict previous results, where homoeologous copies of pCOSAt103 gene were identified in about one third the *Nudiscapus* taxa (Harpke *et al.*, 2013) and the authors suggested allotetraploid origin for the section *Nudiscapus*.

In most angiosperms, compound microsatellites are found such as CA-TC. In plants AC/GT satellites are scarce compared to mammalian genomes (Morgante 2002), while the tetranucleotide SSRs are much less frequent in coding than in the non-coding regions (Scotti *et al.*, 2000). But in this analysis, only one tetranucleotide (ATGT)₁₄ tested, and amplified multiple bands from most species, indicating the existence of multiple copies of (ATGT)₁₄ (Figure 4.3E, 4.4E). Further, microsatellites evolve rapidly and variation in number of units of the repeats is common; their cross taxon utility is also well documented (Varshaney *et al.*, 2005). Given the range of accessions studied here, both within and between *Crocus* species (see Table 2.2), it is expected that length polymorphism will be high, certainly compared to that in related varieties of crops. For example, Saeidi *et al.* (2006) found microsatellite variation between *Aegilops tauschii* accessions gave no useful taxonomic or phylogenetic conclusions because it was so high, although the same microsatellites were known to be valuable in studying wheat pedigrees including the *Ae. tauschii* D genome.

In *A. thaliana*, rice, soybean, maize and wheat SSR frequency was reported higher in the ESTs compared to the non-coding genomic DNA. While, abundance of AG/CT repeats and lower frequency of AT microsatellite has been reported in ESTs (Morgante, 2002). However, in the current analysis out 6,603 *Crocus* ESTs, only 15 ESTs containing 2-4bp repeat regions were found, and the results here contradict the above but go parallel to those suggesting SSRs are rare in the protein-coding regions

(Wang *et al.*, 1994; Tóth *et al.*, 2000). In the *Crocus* microsatellite analysis here, there was an unexpected type of variation: 40 sequences had the CT motif, while the remaining 67 had variants with multiple poly T (2 to 7) interspersed mostly with C; 28 sequences had the CA repeat (mostly those also with the CT microsatellite; ten sequences in two groups had only the CT and no CA repeat) while the remaining sequences had 0 to 3 CA repeats (Figure 4.10B). Microsatellites are widely used as rapidly evolving markers to discriminate related germplasm, and their evolution is normally considered to occur through replication slippage and recombination resulting from unequal crossing over, or gene conversion (Tóth *et al.*, 2000; Li *et al.*, 2002). Although there are numerous recent publications exploiting variation in SSRs as molecular markers, there has been almost no work on the mechanisms of evolutionary changes in the last decade, and indeed few studies involving sequencing of polymorphisms between accessions. The results here suggest that a better understanding of the mechanisms of SSR evolution is needed in both plant and animals, for a better understanding of both inter and intraspecific SSR polymorphism. These mechanisms, leading to length variability within the SSR motifs (although giving variation that does not follow a single-stepwise model), do not account for variation of the nature detected here in the SSR region between the *Crocus* accessions. *C. cartwrightianus*, *C. oreocreticus* and *C. hadriaticus* are among the potential progenitor species of *C. sativus* (Mathew *et al.*, 1982; Zubor *et al.*, 2004; Petersen *et al.*, 2008; Seberg and Petersen, 2009). Based on a single copy nuclear gene *pCOSAt103* *C. pallasii* has also been proposed as a candidate ancestor (Harpke *et al.*, 2013), although the authors do not mention the subspecies of *C. pallasii*; but from the results here (e.g. Fig. 4.7). It is evident that the three *C. pallasii* sub-species show as many differences at the DNA level as the recognized species. In analysis of the (TC)₂₅ sequences, *C. sativus* sequences were found associated with *C. cartwrightianus*, *C. oreocreticus*, *C. pallasii* subsp. *pallasii* and *C. hadriaticus* in sub-clades F2, F3 and F4, that would support the allotriploid origin of *C. sativus* as well as the results of ATPs (see above).

4.4.3 Variation in *matK*, *rbcL* and *trnH* chloroplast and mitochondrial (barcoding) genes

Universal DNA barcoding strategies have been employed for diverse groups of both plants and animals and have aided to our understanding and the course of recognizing new species. Barcoding systems in land plants seem to be more puzzling, as in plant genome the substitution rates are considerably lower than those observed in animals (see Newmaster *et al.*, 2006). In the current study, to further clarify *C. sativus* putative ancestors, universal barcoding genes i.e. *matK*, *trnH*, *rbcL* and ITS (Table 4.1) were sequenced from *Crocus* series *Crocus* and assessed. Nucleotide sequence analysis of these genes revealed high levels of polymorphism with few clear patterns of inheritance (Figures 4.6, 4.13) consistent with data from of Peterson *et al.* (2008), variation within wild accessions was so high that no clear phylogeny could be inferred. Furthermore, both the unconventional use of the assembly algorithm to group the most similar sequences (discounting random nucleotide variations) and the phylogenetic trees built using the Neighbour Joining algorithm revealed some important patterns of relationships. In addition, recombinant sequences appeared to be identified by the assembly algorithm as ungrouped sequences (where the sequence had fragments of two other sequences). This feature of sequence evolution is notable and it would be interesting to confirm it with additional primers.

Out of the five marker barcodes, ITS is known to be the most polymorphic and have the highest discriminatory power (Li *et al.*, 2011). No bacterial or fungal contamination was detected in any sample; still several polymorphic copies of ITS region were amplified (Figure 4.6). Perhaps, the multiple copies of ITS have been amplified from different homoeologous groups as *C. sativus* alone has 13 rDNA sites (see Chapter V). As the origin of the multiple ITS copies was not known, the products of ITS barcode were not sequenced. DNA sequences of *matK*, *trnH* and *rbcL* regions were analysed separately as well as assorted in head to tail. In both cases the tree topology remained almost identical (Figure 4.13).

In the current study, *C. banaticus*, the only member of subgenus *Crociris* (Mathew, 1982) was applied as outgroup, and it is located robustly as a sister group to clade G along with *C. cartwrightianus*, *C. asumaniae* and *C. moabiticus*. Previous

mitochondrial and nuclear DNA based studies have placed *C. banaticus* with other members of section *Crocus* (Petersen *et al.*, 2008; Harpke *et al.*, 2013). Furthermore, based analyses of barcoding genes, *C. hadriaticus* and *C. thomasii* have been reported to be closely related to one another (Gismondi *et al.*, 2013), although the authors also proposed an early separation, and suggested very ancestral origin for the two species, rejecting the possibility that either of the two could be the progenitor species for *C. sativus*. The current work too, indicated the two species to be closely related to one another, but contradict their early separation from *C. sativus* or *C. cartwrightianus* group (clade H, Figure 4.13). Both species exist as a subgroup to *C. sativus*, and my these results are in agreement to the published data of Petersen *et al.* (2008). Comparative karyotype analysis has also revealed *C. cartwrightianus* as one of the ancestors of *C. sativus* in case of autotriploidy (Mathew, 1982; Grilli Caiola *et al.*, 2004). DNA sequencing- based studies have also suggested *C. cartwrightianus* as a potential ancestor for *C. sativus* (Petersen *et al.* 2008; Grilli Caiola and Caniani, 2010; Gismondi *et al.*, 2013) and the results here, for the maternally inherited sequences, were consistent with *C. cartwrightianus* as providing the female ancestor of *C. sativus*, and the nuclear gene analysis supported *C. pallasii* subsp. *pallasii* as the second ancestor (Figure 4.13). The latter, contradict the RAPD markers analysis of Grilli Caiola *et al.* (2004) where the authors rule out the hypothesis of close relationships between *C. sativus* and *C. pallasii*.

Nevertheless, intraspecific variation comprises the core of modern evolutionary biology, and its ever-increasing importance is well documented (see Funk and Omland, 2003). Besides identifying the potential ancestors of *C. sativus*, the barcoding genes revealed intraspecific variation in different accessions of *C. pallasii*, *C. cartwrightianus* and *C. sativus*. Possible reasons for such variation could be related to the difference in geographical distribution or nursery practices that may cause heritable epigenetic changes (Slatkin, 1987; Hyten *et al.*, 2006 and above). In the analysis here, substantial variation was revealed between different accessions of *C. cartwrightianus*, and in the future it will be important to survey the full range of variation within this species to see if any accessions more closely match the alleles in the less diverse *C. sativus*. Knowledge about different alleles across the *Crocus* genus has potential for making new hybrids and increasing the genetic base: identification of such variation will also be

useful to allow us to discriminate and identify the authenticity of saffron sample from possible contaminants.

4.4.4 Variation in *C. sativus*

Phenotypic variation, such as flower size, tepals shape, colour intensity, pistil weight and pollen viability suggest the possibility of saffron improvement through selection (Grilli Caiola *et al.*, 2001; Macchia *et al.*, 2013 and Figure 1.4). Nevertheless, clone selection of saffron is one of the major issues addressed today by many research groups (Agayev *et al.*, 2009). Saffron multiplies by vegetative means, which does not induce genomic variations except for the rare mutation, and these too are not easily detectable in a triploid genome (see Chapters I, III). Evolutionary history of the genus *Crocus* is very complex as indicated by intensive species hybridization and explosive speciation in the evolution of *Crocus* (Frello *et al.*, 2004) and that could be one of the selective pressures in the origin of saffron (Fernandez, 2007).

At the nucleotide level different accessions of *C. sativus* grown around the world have shown rare diversity, while differences in saffron quality are mainly attributed to the post-harvest processing of stigmas, and to some extent agronomy, and are independent of the saffron is genetic origin (Grilli Caiola *et al.*, 2004; Ordoudi *et al.*, 2004; Fluch *et al.*, 2009); recent report has revealed the effects of geographical origin, cultivation and environment in saffron (Macchia *et al.*, 2013). Earlier, eleven IRAP markers (in all possible combinations) were applied to gain insights of the genetic variation between saffron accessions grown worldwide, and the data showed minimum diversity within *C. sativus* accessions (Chapter III). Therefore, it is very likely that saffron has originated once and then has undergone artificial selection. Such practice offer advantages for many of the traits required for domestication and in maintaining its genetic characteristics while compromising on most of the variability (Rubio-Moraga *et al.*, 2009).

Genetic heterogeneity provides vigour and is a possible solution to the vulnerability of monocultured crops (Zhu *et al.*, 2000). A solid understanding of the genetic variability and population structure of the wild and cultivated plant populations is necessary for sustainable management, conservation of genetic resources and broadening of the genetic base of plant species (Heslop-Harrison and Schwarzacher,

2012). Thus, one of the chief aims of the current project was to ascertain genetic diversity in *C. sativus*. There are several reports that support the allotriploid nature of *C. sativus* (see Fernandez, 2006; Grilli Caiola and Canini, 2010). The ATPS and TC₂₅ sequences has revealed intraspecific variation in *C. sativus* as well as supported the new concept of allotriploid *C. sativus* (Figures 4.8B, 4.10B). Though, there was no evidence from the nuclear genes that any accession of *C. sativus* included unique sequences. However, for the chloroplast genes analysed here, there was some evidence that the accessions identified as Kashmir/Cashmeriensis had a different range of polymorphisms from those of the European *C. sativus* accessions suggesting multiple maternal origins (Figure 4.13). These results are partly contradictory to my previous IRAP amplified polymorphism results, where minimum to no diversity was detected (Chapter III). However, the sequence based diversity is not known, as the IRAP bands were not sequenced. The results obtained with ATPS, TC₂₅ and barcoding genes indicate that different *C. sativus* species might have evolved through independent events, or genetic differences, found between *C. sativus* (CstCD09, *Cashmirianus*) and *C. sativus* (CstPER09, Spanish) might be due to different habitat selections (see Figure 4.13). Such intraspecific variation has been reported in Italian and Spanish saffron too (Gismondi *et al.*, 2013). Not formally characterized, there is little evident morphological variation between the diverse saffron collections once grown in a common nursery at the Crocusbank collection in Cuenca, Spain (Figures 1.4, 1.12). Recent RAPD and SSR analysis of different Iranian saffron accessions have also indicated the existence of genetic variability among saffron accessions (Namayandeh *et al.*, 2012; Izadpanah *et al.*, 2014). Moreover, the results revealed that the DNA barcoding approach can be used not only for molecular intraspecific discrimination but is equally effectively in tracing the authenticity and geographical origin of saffron, along may be applied along other known DNA or biochemical markers (Maggi *et al.*, 2011; Torelli *et al.*, 2014).

Saffron reproducing asexually, so it is also interesting to consider if new genetic variation is occurring via somatic mutation (Grilli Caiola *et al.*, 2004; Fernandez *et al.*, 2011). This might involve targeting candidate genes where variation is noted in field-grown material, or alternatively genetic mapping of loci in wild species before examination of the equivalent loci in saffron. Furthermore, given this result, it will now be interesting to sequence whole chloroplast genomes from European and Asian saffron

accessions, and to compare these with the sequences from other *Crocus* species. It is possible that organellar genome fragments have been transferred to the nuclear genome (Huang *et al.*, 2003), so it is important that this is ruled out, and the use of whole-genome sequencing with modern methods is likely to be most effective so that we can look more closely at the way in which saffron has evolved.

5 CHAPTER V: RELATIONSHIPS OF *CROCUS* SPECIES BASED ON CYTOGENETIC INVESTIGATION AND ORGANIZATION OF REPETITIVE DNA SEQUENCES

5.1 Introduction

5.1.1 Cytogenetic structure of *Crocus sativus*

Crocus series *Crocus* is a heterogeneous group of largely autumn-flowering crocuses, widespread in the Mediterranean and Asia Minor. Most of the species occur in the wild and a few are grown as ornamentals while the most economically important *Crocus sativus* is grown for the production of saffron and not known in the wild (see Chapter I). The genus *Crocus* shows a wide range of chromosomal variation *i.e.* $2n=6$ to $2n=70$ (*C. candidis* $2n=6$, *C. mathewii* $2n=70$), even at the species level chromosome numbers may vary (Mather, 1932; Karasawa, 1935, 1942; Brighton *et al.*, 1973; Goldblatt and Takei, 1997; Schneider *et al.*, 2012). The tremendous economic importance and the presence of large chromosomes have always inspired cytogenetic investigation of the genus (Karasawa, 1942; Agayev, 2002 and Chapter I). The pioneering cytological work of Himmerbaur, Sugiura, Mather and Karasawa in the early 20th century revealed the variability in chromosome number for *Crocus*: $2n=24$ and $2n=15$ or $2n=14$ (Himmerbaur 1926; Sugiura, 1931; Mather, 1932; Karasawa, 1935). Later, Pathak (1940), Feinbrun (1958), Brighton *et al.*, (1973), Brighton (1977), Mathew (1977), Ghaffari (1986), Goldblatt and Takei, (1997), Ebrahimzadeh *et al.* (1998), Frello and Heslop-Harrison, (2000b), Frello *et al.*, (2004), Fernandez *et al.* (2009) Schneider *et al.*, (2012) have studied the chromosome number and karyotype evolution within the genus *Crocus*. The authors highlighted the extreme complexity of the genus at the karyotype level and most describe *C. sativus* as a triploid species with $2n=3x=24$, $x=8$ (see Table 5.1 and discussion below).

Since cytogenetic analysis provides direct insight into karyotype evolution, it plays a critical role in reconstructing phylogenies. Therefore, today cytogenetics is an integral part of phylogenetic reconstruction and genome mapping projects (Schmidt and

Heslop-Harrison, 1998; Schwarzacher, 2003a; Markova and Vyskot, 2009). Phylogeny of the genus and the parental species involved in *C. sativus* speciation are not explicitly known (see Fernandez *et al.*, 2011). Comparison of the primitive members of the genus *Crocus* (Mathew, 1982) with the closely related genus *Syringodea* ($x=6$) suggest that the basic chromosome number for the genus *Crocus* could be $x=6$, that might have undergone subsequent reduction to $x=4$ by descending dysploidy in several lineages of *Crocus* and polyploidization events may be based on the basic numbers of $x=3-6$ (Goldblatt and Takei, 1997). Comparative morphological and molecular approaches have led to the hypothesis that *C. cartwrightianus*, *C. hadriaticus*, *C. oreoreticus*, *C. thomasi*, *C. pallasii* or *C. cartwrightianus* var. *albus* maybe potential ancestors, of *C. sativus* through hybridization, polyploidy or mutation (Chapters III, IV). Further, it is a widely accepted theory that *C. sativus* has been propagated from a sterile autotriploid clone (Mathew, 1982; Ghaffari, 1986; D'Agostino *et al.*, 2007). The EST-SSR and SNP data clearly demonstrated the existence of some diversity and presence of more than one and up to three copies of genes, indicating that *C. sativus* is most likely an allotriploid (previewed in Chapter IV). With no comparison, saffron is a high value, sustainable crop where improvement is potentially possible through exploitation of diversity within the genus *Crocus*. Several potential candidate species already pinpointed as ancestral species of *C. sativus*: confirmation of it will provide an opportunity to attempt to make new hybrids, to resynthesize saffron, introduce novel diversity and compare the different forms (see below).

Table 5.1: Chromosome number in members of the series *Crocus* (modified from Grilli Caiola and Canini, 2010).

Serial	Species	Subspecies	Authority	Chromosome (2n)	Reference
1	<i>C. asumaniae</i>	-	B. Mathew & T. Baytop	26	Mathew 1999
2	<i>C. cartwrightianus</i>	-	W. Herbert	16	Brighton 1977
3	<i>C. hadriaticus</i>	-	W. Herbert	16	Brighton <i>et al.</i> 1973
4	<i>C. mathewii</i>	-	H. Kerndorff & E. Pasche	16	Mathew 1999
5	<i>C. moabiticus</i>	-	F. Bornmüller & J.E. Dinsmore	14	Kerndorff 1988
6	<i>C. naqabensis</i>	-	D. Al-Eisawi	14	Al-Eisawi 2001
7	<i>C. oreoreticus</i>	-	B. L. Burtt	16	Brighton <i>et al.</i> 1973
8	<i>C. pallasii</i>	<i>pallasii</i>	K. L. Goldbach	14	Mathew 1999; Candan <i>et al.</i> , 2009
		<i>turcicus</i>	B. Mathew	12	Mathew 1999
		<i>dispathaceus</i>	E.A. Bowles	14	Mathew, 1999
9	<i>C. thomasi</i>	-	M. Tenore	16	Brighton <i>et al.</i> 1973
10	<i>C. sativus</i>	-	C. Linnaeus	14, 15 24 24 16, 20, 24, 40 24	Mather, 1932 Karasawa, 1935 Pathak, 1940 Karasawa, 1942 Agayev, 2002

5.1.2 Repetitive DNA organization in *C. sativus*

Polyploid formation has been a major force in the evolution of many plants and animals (Mable, 2013; Madlung, 2013). However, our understanding of the subsequent evolution of DNA sequences that become united in a common nucleus is limited (McClintock, 1983; Heslop-Harrison and Schwarzacher, 2012). The recent advancements in sequencing technologies at affordable costs have given a direct access to the genomic architecture of both plants and animals (see <http://genomesonline.org>). It is now known the gene content and order among angiosperms is relatively uniform and that the remarkable diversity in genome sizes is due to the presence of various classes of repetitive DNA elements (Bennett and Leitch, 2011; Heslop-Harrison and Schwarzacher, 2011).

In plants, repetitive DNA may account for up to 70-80% or even more of their whole genome and may be found dispersed in the form of TEs, that are free to propagate in the genomes (see Chapter III), or as tandem repeats in discrete clusters (Orgel and Crick, 1980; Kubis *et al.*, 2003). The rapid evolution of both tandemly arranged and dispersed repetitive DNA often leads to changes in sequence composition and abundance and it could possibly be the main force responsible in speciation (Schmidt and Heslop-Harrison, 1998; Shapiro and Sternberg, 2005). Thus, understanding the role and nature of repeated DNA elements is extremely important in investigating organizational and phylogenetic relationships among the genomes (Schwarzacher 2003a; Kalendar *et al.*, 2011; Heslop-Harrison and Schmidt, 2012; Estep *et al.*, 2013).

Tandemly repeated DNA comprises sequences of various motifs and lengths that are tandemly organised in the form of long arrays extending from few to tens of kilobases, concentrating at one or more distinct genomic locations and are referred to as satellite DNAs (Frello *et al.*, 2004). Many different satDNA families have been described in plants, showing species or genome specific diversity in their DNA sequence and chromosomal distribution (Schweizer *et al.*, 1988; Vershinin *et al.*, 1994, Contento *et al.*, 2005). Often long arrays of different satellites (referred to as library of satDNAs) may coexist in the same genome (Kuhn *et al.*, 2007), mainly concentrated in the heterochromatic regions around the centromere, or at interstitial or subtelomeric regions (Heslop-Harrison and Schwarzacher, 2011). Despite their abundance, the biological

significance of most repetitive DNA still remains uncertain and most is considered as selfish DNA (Orgel and Crick, 1980; Shapiro and Sternberg, 2005). Nevertheless, regions of chromosomes rich in repetitive DNA families have lower susceptibility to recombination events (Kuhn *et al.*, 2007; Estep *et al.*, 2013).

Fluorescent *in situ* hybridization (FISH) and genomic *in situ* hybridization (GISH) are powerful techniques for identifying chromosomes and following their alteration during evolution over long time-scales associated with speciation, as well as over shorter periods associated with plant breeding (Schwarzacher *et al.*, 1992; Heslop-Harrison *et al.*, 2003; Frello *et al.*, 2004). Within the genome, tandemly repeated DNA sequences are convenient landmarks for many aspects of genome analysis. Some repeats show chromosome, genome or species specificity and can be combined with total genomic DNA probe as a tool for the recognition chromosomal arm, chromosome or a genome in case of polyploids, thus providing evidence of rare gene flow in natural populations (Zhao *et al.*, 1998; Carvalho *et al.*, 2009). Interestingly, except for *C. vernus* where repetitive DNA sequences have been isolated and tested across wild species of the genus (Frello and Heslop-Harrison, 2000a, 2000b; Frello *et al.*, 2004), to the best of my knowledge no major satellite DNA sequences (abundant non-rRNA-related and non-telomere tandem repeats) have been isolated from *Crocus*.

The aim of the current work was to 1) exploit FISH or GISH using total labelled genomic DNA as a probe to chromosome spreads of *C. sativus* to ascertain the potential ancestral species; 2) isolate and characterize some repetitive DNA sequences from *C. sativus*; 3) investigate physical organization of the tandemly organized repetitive DNA sequences along the chromosomes of *C. sativus* and series *Crocus*; and 4) use any isolated sequences of repetitive DNA, and microsatellite, 5S and 45S rDNA sequences as chromosome markers to identify homologous chromosomal pairs and to establish the karyotype for *C. sativus*.

5.2 Materials and Methods

The list the species used in the current study is given as (Table 5.2). Except for *C. pallasii* for which accessions of sub species were included in the current study, for all other species only one accession was used. While intra-specific variation in tandem

repeat array size and localization is well known, width of sampling was given priority over testing multiple accessions here. Sources of all the material are given in (Table 5.2).

5.2.1 DNA extraction, restriction enzyme digestion and gel electrophoresis

Total genomic DNA was extracted from the *Crocus* species (Table 5.2) following the standard CTAB method described in Materials & Methods (Chapter II). Genomic DNA was digested with restriction endonucleases; *HaeIII*, *HindIII*, *BamHI*, *Sau3AI*, *DraI* and *EcoRI* (New England BioLabs) in the presence of appropriate buffers. Restriction digestion and gel electrophoresis conditions were as described earlier in (Chapter II).

5.2.2 Isolation of repetitive DNA sequences from *C. sativus*

Appropriate restricted genomic DNA products of *HaeIII* and *DraI* of 200-3000bp were eluted and cloned into pGEM®-T Easy vectors (Promega). About 50-100 white colonies were chosen, grown on LB agar plates, transferred onto positively charged nylon membrane and hybridized for 16hrs with digoxigenin labelled *C. sativus* genomic DNA probe. The process of cloning, plasmid selection and dot blot hybridization was as described in materials and methods (Chapter II and appendix 5).

5.2.3 *In situ* hybridization

For *in situ* hybridization both meiotic and mitotic chromosomes were prepared from fixed material on clean glass slides. Chromosomal preparation, probe labelling and *in situ* hybridization followed the protocol of Schwarzacher and Heslop-Harrison (2000) and described in materials and methods (Chapter 2, section 2.2.11-2.2.12).

5.2.4 Probes used

For *in situ* hybridization probes used included:

pTa71 contains a 9kb *EcoRI* fragment of the repeat unit of 25S-5.8S-18S rDNA isolated from *T. aestivum* (Gerlach and Bedbrook, 1979) and was linearised with *EcoRI* before labelling.

Total genomic DNA from *Crocus* species (Table 5.2) was sheared to 3-5kb fragments by autoclaving before labelling. GISH was carried out with and without autoclaved

genomic DNA from *C. sativus* (20-30x of the probe concentration) that was added to the mixture as blocking DNA.

Table 5.2: List of *Crocus* species used in the study given along University of Leicester identification code and source of origin.

Serial	Species	Subspecies	University of Leicester Number	Source
1	<i>C. asumaniae</i>	-	CasWD09	JW Dix Export (The Netherlands)
2	<i>C. cartwrightianus</i>	-	CcwBD09	JW Dix Export (The Netherlands)
3	<i>C. hadriaticus</i>	-	ChdWD09	JW Dix Export (The Netherlands)
4	<i>C. mathewii</i>	-	CmatD08	JW Dix Export (The Netherlands)
5	<i>C. oreoreticus</i>	-	CorVD09	JW Dix Export (The Netherlands)
6	<i>C. pallasii</i>	<i>pallasii</i>	CplVD09	JW Dix Export (The Netherlands)
7	<i>C. thomasii</i>	-	CtomVD09	JW Dix Export (The Netherlands)
8	<i>C. sativus</i>	-	CstPER09	Spain

5.3 Results

Multi target *in situ* hybridization with genomic and repetitive DNA probes was applied simultaneously to the spread mitotic and meiotic chromosomes of *C. sativus*. Genomic probes from *Crocus* series *Crocus* were used with an aim to identify the chromosomal complement from the potential donor parents, while unique banding patterns of repetitive DNA was helpful in identify and designating the chromosomal pairs for karyotyping *C. sativus*. The same strategy was applied to test some more distantly related species of *C. sativus* (outside *Crocus* series *Crocus*) as controls, and to rule out the possibility of any cross hybridization.

5.3.1 Karyotype of *C. sativus*

Figures 5.1 and 5.2 illustrate the karyotype and ideogram of *C. sativus* following *in situ* hybridization with total genomic DNA from '*C. sativus cartwrightianus*', *C. pallasii* subsp. *pallasii* and 45S rDNA. The genomic probes showed disperse weak hybridization signals and relatively stronger uniform signals around the centromeric region of all chromosomes (Figures 5.1, 5.2C). Interestingly, both the genomic probe from '*C. sativus cartwrightianus*' and *C. pallasii* subsp. *pallasii* labeled with biotin (red) and 45S rDNA (pTa71) probe labelled with digoxigenin (green) strongly labelled the rDNA regions.

Further, both probes localized 45S rDNA sites at discrete sites on 13 chromosomes in the triploid *C. sativus*. All 45S rDNA sites were terminal, and none was intercalary (Figures 5.1, 5.2). Several cells from different accessions of *C. sativus* (Spanish and Kashmirian) were analysed and there were no polymorphisms between accessions as far as the number or locations of 45S rDNA sites was concerned (see discussion).

Here the karyotype of *C. sativus* is drawn and position of the 45S rDNA on each chromosome is shown (Figure 5.1A). Chromosome types are numbered according to their sizes and presence of satellites, while length of chromosome is taken as per Agayev (2002) and Fernández *et al.* (2009); although length measurements can be made on chromosomes prepared for *in situ* hybridization using 8-hydroxyquinoline pretreatment and enzymatic digestion, the length variation is larger than with protocols using more extreme pretreatment and acid digestion, so the published results were used here; no substantial inconsistencies were noted. The first triplet consists of sub-acrocentric chromosomes with large but relatively polymorphic satellite regions. These are the largest chromosomes in the karyotype of *C. sativus* having strong DAPI bands on the long arm at the sub-centromere and this triplet can be easily distinguished. The second triplet of chromosomes is also sub-acrocentric, characterized by the presence of satellites on the long arm of all the three chromosomes. Further, the intensity of 45S rDNA fluorescence in this triplet was different, and this variation is most likely due to the different sizes of satellite region (compare chromosomes in Figure 5.1B). The third pair is sub-metacentric while the fourth triplet is comprised by metacentric chromosomes, and all contain 45S rDNA sites. The 45S rDNA signals are comparatively weaker on the fourth triplet as compared to the third triplet of chromosomes (Figure 5.1A, B). Similarly, the fifth triplet comprises of heteromorphic chromosomes, one of the chromosomes 5(1) is metacentric and has a prominent 45S rDNA site and a strong DAPI band near the centromere (arrows in Figures 5.1B, 5.2) while the other two chromosomes 5(2,3) are sub-acrocentric and have no satellites nor the strong DAPI bands. Sub-metacentric chromosomes comprise the sixth and seventh triplet, while the eighth triplet are metacentric chromosomes. No rDNA sites were detected in the latter three

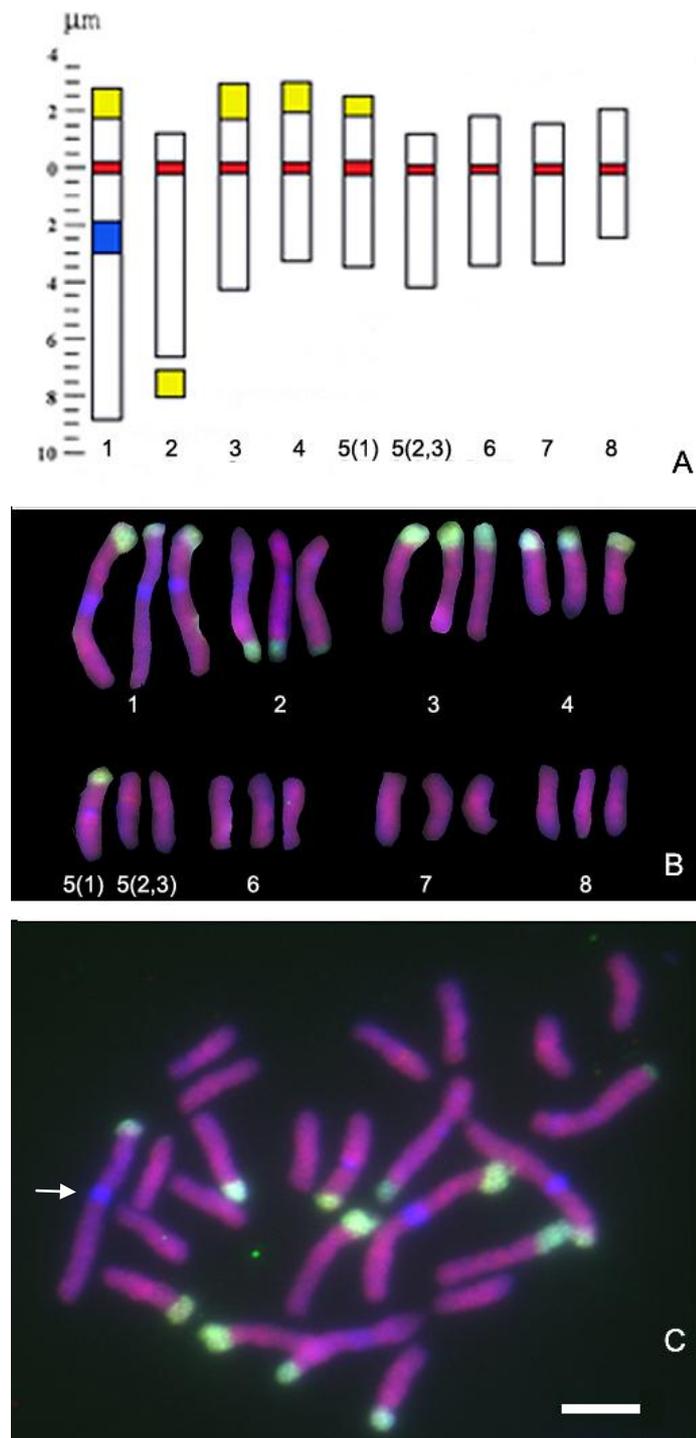


Figure 5.1: **A)** An ideogram of *C. sativus* ($2n=3x=24$, $x=8$) based on length of chromosome, position of centromere and presence of 45S rDNA sites, 1-8 indicating to morphologically similar chromosome (triplet8), chromosomes 1,2,5(2,3): Subacrocentric, chromosomes 3,4,5(1),8:Metacentric, chromosomes 6,7:Submetacentric (from Agayev *et al.*, 2002). **B)** Root-tip metaphase chromosomes of *C. sativus* after fluorescent *in situ* hybridization with 45S rDNA clone, pTa71 labelled with digoxigenin 11-dUTP (detected in green) and total genomic DNA from '*C. sativus cartwrightianus*' labelled with biotin 16-dUTP (detected in red). For karyotype, **C)** metaphase plate shows individual chromosomes from were cut and paired together (**B**)⁴Arrows indicate DAPI bands. Bar represents 5μm.

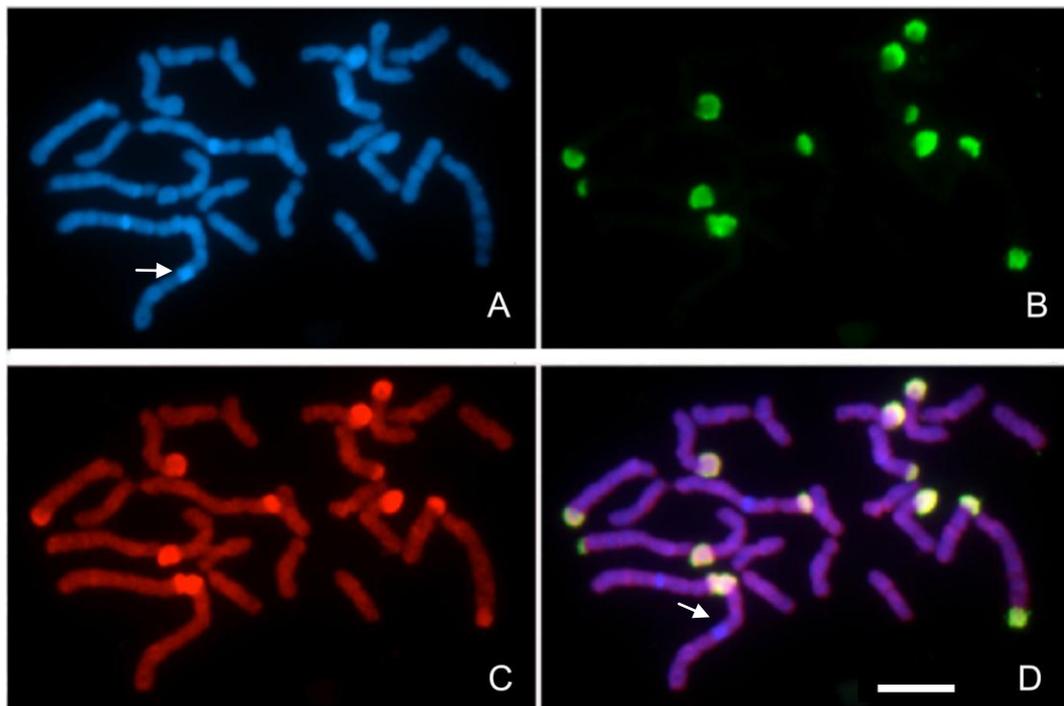


Figure 5.2: Root-tip metaphase chromosomes of *C. sativus* ($2n=3x=24$) after fluorescent *in situ* hybridization (FISH). **A)** *C. sativus* chromosomes fluorescent blue with DAPI staining. **B)** Hybridization pattern of the 45S rDNA clone, pTa71 labelled with digoxigenin 11-dUTP (detected in green) showing 13 45S rDNA sites. **C)** *In situ* hybridization of the total genomic DNA from *C. pallasii* subsp. *pallasii* labelled with biotin 16-dUTP (detected in red) showing hybridization signals with the 13 45S rDNA sites and on single chromosomal arms and centromeric region of few smaller chromosomes. **D)** Overlay of A, B and C images. Arrows indicate DAPI bands. Bar represents 5 μ m.

5.3.2 Meiotic chromosomes pairing in *C. sativus*

Figures 5.3 to 5.5 illustrate the meiotic pairing behaviour of chromosomes in the pollen mother cells (PMC) of *C. sativus*. Heterotypic meiotic division was observed in the PMC of *C. sativus*, where the conjugations of chromosomes were variable and non-uniform. In several instances all 24 chromosomes paired into eight trivalents or even possible quadrivalents (Figure 5.3A). In other PMC trivalents and bivalents were observed, while in still some other cases a mixture of trivalents, bivalents and univalent were observed. In these heterotypic meiotic divisions lagging *C. sativus* chromosomes were seen (Figure 5.3B, C). Furthermore, variation in the number of paired chromosomes ranged from five to eight for the trivalent, one to three for bivalents and from zero to three for the univalent (see Figures 5.3A, B, C, and D).

The early meiotic pachytene chromosomes analysed here did not pair completely: rather incomplete pairing was evident in most cases, and three partially paired chromosomes were detected. Here only representative images and enlarged partially paired triplet chromosomes are shown (Figure 5.3E, 5.4). The meiotic pachytene chromosomes in (Figure 5.4A) probed with total genomic DNA from *C. tomasii* (green) and *C. asumaniae* (red) revealed eight bright signals of 45S rDNA sites after hybridization with *C. asumaniae* genomic DNA. The majority of the chromosomes are intertwining, and the univalent, bivalent or trivalent nature of the chromosomes is not very clear. However, one trivalent chromosome enlarged on the right hand side (RHS) clearly shows incompletely paired chromosomes (see arrow Figure 5.4A). Further, the pachytene chromosomes of *C. sativus* hybridized with labelled total genomic DNA from *C. hadriaticus* (red) and *C. mathewii* (green). This metaphase also revealed seven to eight 45S rDNA sites, that are labelled by *C. mathewii* genomic DNA (Figure 5.4B). Most chromosomes are present as a complex network of 'fibres', however, in the proximity of interphase nucleus three fibres corresponding to a triplet of incompletely paired meiotic chromosomes are seen and enlarged on the RHS of the same plate (see arrow in Figure 5.4B). Similarly, (Figure 5.4C) is an early pachytene stage probed with *C. pallasii* subsp. *pallasii* (red) and pTa71 (green), where clustering of 45S rDNA may reveal potential bivalents and trivalents. Along the 45S rDNA sites, three chromosomes in the form of thin fibres are seen, indicating the incompletely paired triplet chromosomes (arrow in Figure 5.4C). The appearance of the univalent, lagging

chromosomes or incomplete pairing in the first meiotic division may be due to weak affinity between the chromosomes. Furthermore, in several PMC the reduction divisions were abnormal, where the chromosomes making non-disjunctions and the final division of PMC resulted into 3-7 or 8 nuclei (see A-H, Figure 5.5). Also in different PMC, significant variation in the size of individual cells was observed (compare Figure 5.5A, B). Although, the actual number of chromosomes in the daughter nuclei could not be counted, but as a result of the failure of chromosomes to pair and segregate faithfully, the daughter cells clearly contained uneven number of chromosomes (Figure 5.5I). On account of the abnormal meiotic division in PMC, the pollen is often somewhat deformed.

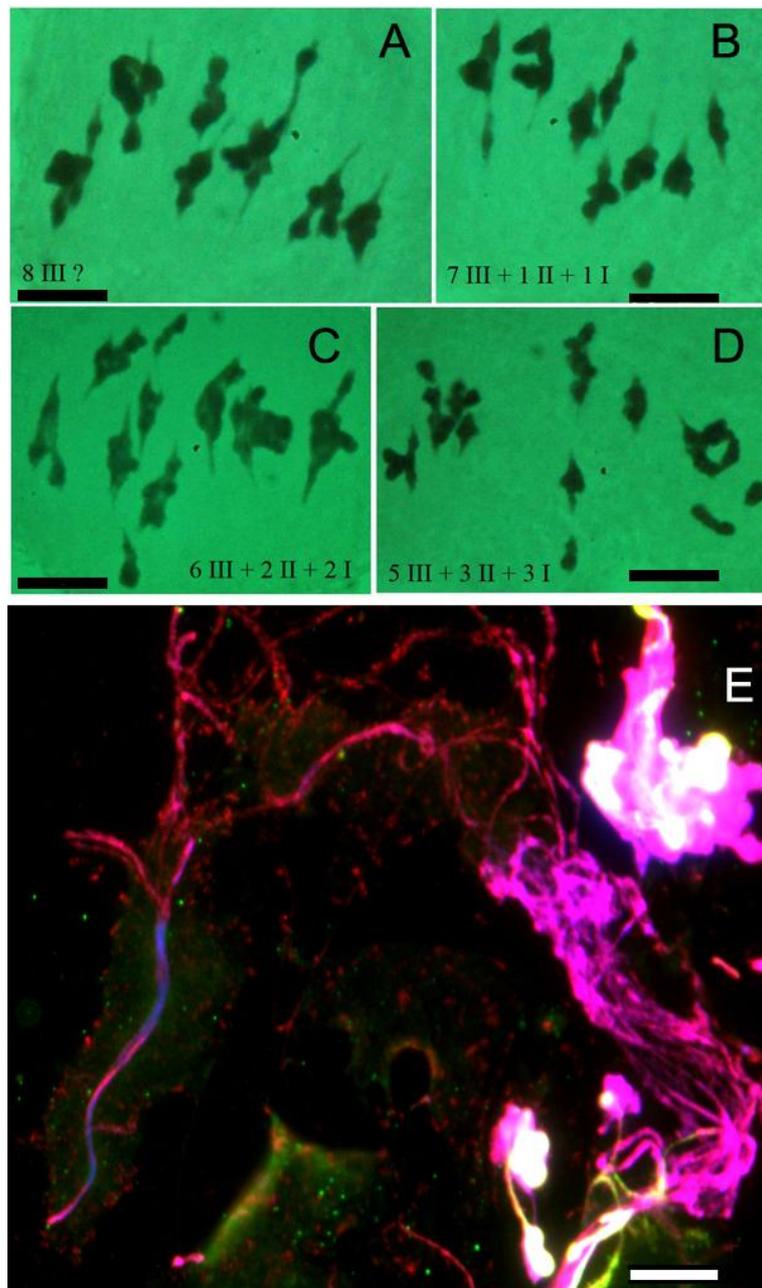


Figure 5.3: **A-D)** Meiotic chromosomes stained with acetocarmine at metaphase I in *C. sativus* ($2n=3x=24$). The cells show as few as 7-8 structures of paired chromosomes, representing trivalents, bivalents and univalent chromosomes. The “?” is indicating ambiguity. **E)** Chromosomes of *C. sativus* at meiotic prophase probed with *C. pallasii* subsp. *pallasii* labelled with biotin 16-dUTP (detected in red) and 45S rDNA clone, pTa71 labelled with digoxigenin 11-dUTP (detected in green). Arrow indicates the incompletely paired triplet chromosome. Bar represents $10\mu\text{m}$ in A, B, C, D and $5\mu\text{m}$ in E. (Stained meiotic figures jointly with Drs John Bailey and Farah Badakshi.)

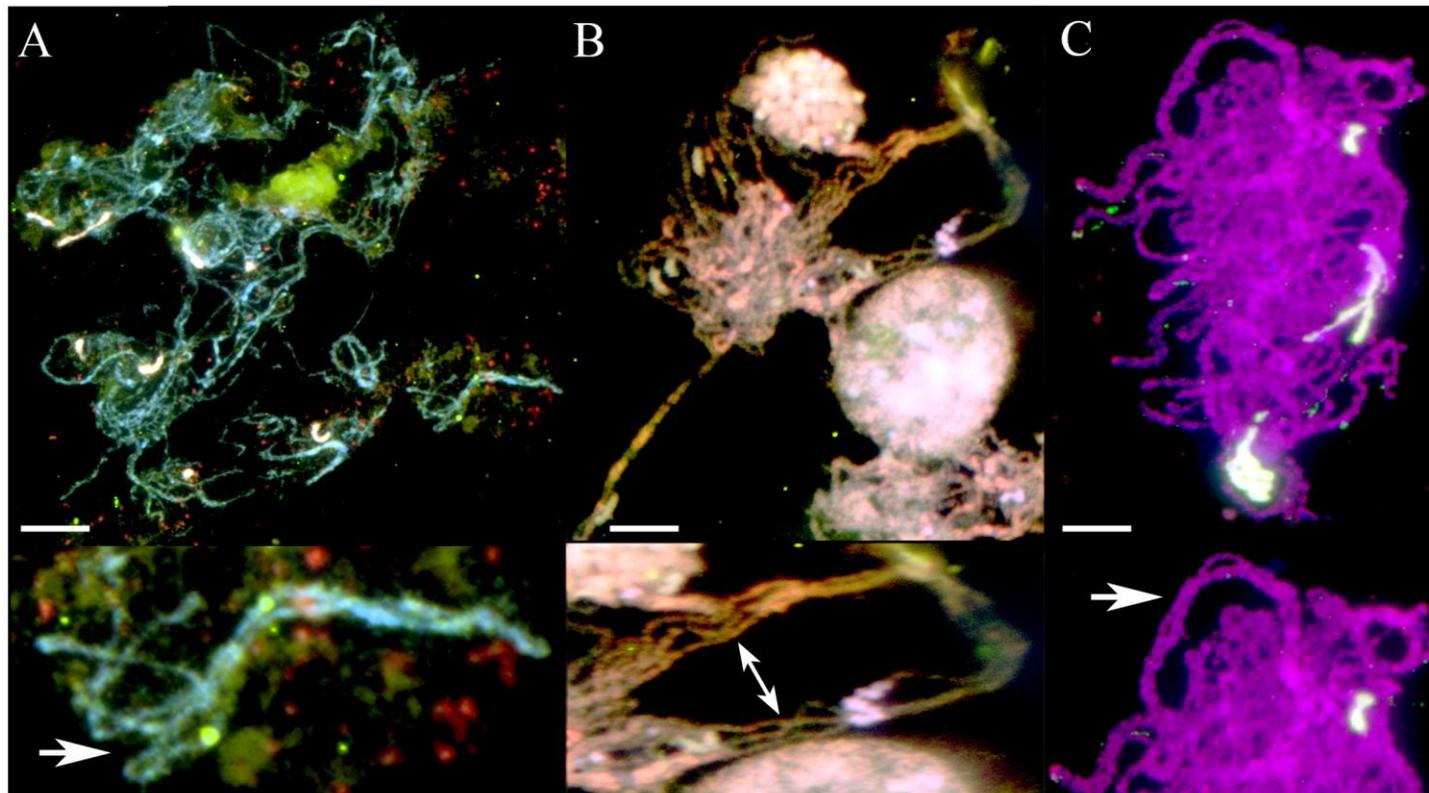


Figure 5.4: Meiotic pachytene chromosomes of *C. sativus* ($2n=3x=24$) after fluorescent *in situ* hybridization (FISH). **A)** *C. sativus* chromosomes probed with *C. tomasii* labelled with digoxigenin 11-dUTP (detected in green) and *C. asumaniae* labelled with biotin 16-dUTP (detected in red). **B)** *C. sativus* chromosomes probed with *C. hadriaticus* labelled with biotin 16-dUTP (detected in red) and *C. mathewii* labelled with digoxigenin 11-dUTP (detected in green). **C)** *C. sativus* chromosomes probed with *C. pallasii* subsp. *pallasii* labelled with biotin 16-dUTP (detected in red) and 45S rDNA clone, pTa71 labelled with digoxigenin 11-dUTP (detected in green). Arrows indicate multiple paired chromosome axes seen in enlargements (right). Bar represents 5µm in main images



Figure 5.5: Heterotypic pollen mother cells (PMCs) following aberrant meiotic division in *C. sativus* ($2n=3x=24$) most likely with three A, B) three, C, D) four, E, F) five, G, H) six or up to eight cells. Non-disjunction of chromosomes resulted in daughter cells with uneven numbers of chromosomes (I), forming restitution nuclei and becoming cellularized. Bar represents 10 μ m.

5.3.3 *In situ* hybridization and characterization of potential donor parents

In situ hybridization gave discrimination along chromosomes when the preparations were extended in length and methods were optimized to show the best discrimination of signal. However, under these pretreatment and preparation conditions, there was some variability in extension of individual chromosomes. Centromeres were sometimes difficult to localize, particularly when there was strong *in situ* hybridization signal. For each probe two representative images are given and where possible chromosomes were aligned into triplet groups. Although it was not possible to align chromosomes, as consistent groups of three in all metaphases and prometaphases, the results were informative given the weak discrimination of chromosomes by the *in situ* hybridization signal. Summary of the *in situ* hybridization is given as (Table 5.3).

5.3.3.1 *In situ* hybridization with genomic DNA from *C. cartwrightianus* and *C. pallasii* subsp. *pallasii*

The *in situ* hybridization in both metaphase and prometaphase chromosomes of *C. sativus* probed with labelled genomic DNA from *C. cartwrightianus* (in B, detected in red) and *C. pallasii* subsp. *pallasii* (in C, detected in green) was fairly uniform (Figures 5.6, 5.7). Further, the hybridization patterns indicated that DNA sequences from genomic DNA of both *C. cartwrightianus* and *C. pallasii* subsp. *pallasii* show high homology to major regions of all 24 chromosomes in *C. sativus*. Both genomic probes showed weak hybridization signals to most chromosomes, while some regions particularly the whole arms of largest chromosomes and centromeric, pericentromeric regions of the short chromosomes showed preferential hybridization (see chromosome 4, 5 and 8 Figures 5.6, 5.7). Much of these dispersed *in situ* hybridization signals are most probably due to different classes of transposable and other repetitive DNA elements.

Based on the hybridization patterns of the three colours (DAPI and *in situ* hybridization signal), as well as karyotype shown (Figure 5.1), sizes and (sometimes indistinct) centromere positions, chromosomes were matched, and in general three similar chromosomes were placed in each group, 1 to 8, in the karyotypes (see Figures 5.6, 5.7). Chromosome 1 is the largest and all three chromosomes in the two metaphases are similar. A prominent DAPI-positive sub-centromeric band collocates with a green

band with minimal red hybridization. The distal part of the long arm shows only weak hybridization with the probes, while both probes show strong hybridization to the region of the short arm, slightly stronger with green than red (seen as a more yellow short arm in the overlay Figures 5.6A, 5.7A). Chromosome 2 is also large and has a satellite (45S rDNA site) on the long arm; the third member of the group is slightly smaller and shows a slightly different hybridization pattern. The short arm has strong intercalary hybridization to both probes while the long arm has relatively uniform hybridization of both probes except for exclusion from the terminal satellite region. Chromosomes 3 and 4 are middle-sized and the end of the long arm shows little hybridization, while the broad centromeric region, particularly in chromosome 4 (like chromosome 5(1) too), shows stronger hybridization with the green probe and red. The left-hand chromosome 5 has more centromeric *in situ* and DAPI signal than the other two. The three smallest chromosomes show distinctive morphology (with 8 being more metacentric than 6 and 7), and with the hybridization patterns that form three fairly well defined groups 6 and 7 (with strongest signal intercalary on the short arm and 8 with more centromeric *in situ* signal). Nevertheless, chromosomes in further complete and partial metaphases or prometaphases seen elsewhere in the slides fitted the patterns seen in these two metaphases and in general could be assigned to the same groups (Figures 5.6, 5.7).

5.3.3.2 *In situ* hybridization with genomic DNA from *C. thomasi* and *C. asumaniae*

In situ hybridization of *C. sativus* metaphase chromosomes is given in (figures 5.8 and 5.9), labelled with total genomic of *C. thomasi* (detected in red) and *C. asumaniae* (detected in green). The genomic probe from *C. thomasi* indicated a higher degree of homology to *C. sativus* chromosomes than *C. asumaniae*, and hybridized to all major regions of the 24 chromosomes (see B in Figures 5.8, 5.9). Variation in signal intensity and hybridization pattern on the same as well as within the triplet chromosome was evident (see Figure 5.4 and below)

The genomic probe from *C. asumaniae* showed strong hybridization to the 45S rDNA sites, and relatively weak, but uniform hybridization, along chromosome arms with some exclusion from centromeric regions; particularly where there were DAPI-positive bands such as in chromosome 1 and 5(1) (see D in Figures 5.8, 5.9). As with the

45S probe, major and minor hybridization sites were seen, and sometimes sites had been lost but comparable to the 45S rDNA karyotype (compare Figures 5.1 and 5.8).

As in the other *C. sativus* metaphases, chromosomes could be placed into groups of three based on their morphology, DAPI staining, rDNA sites and *in situ* hybridization patterns. However, the *C. thomasii* probe often discriminated one of each group of three chromosomes by a somewhat different hybridization pattern (see C, Figures 5.8, 5.9). Two members of the chromosome 1 showed four distinct bands, while the third showed more diffuse but still strong hybridization. Among the chromosome 2 pair, one chromosome showed stronger overall hybridization than the other members of the pair, and the hybridization pattern on the chromosome 3 was more or less identical, where the small arm showing more intense but uniform *C. thomasii* signals. Similarly, variation in signals of chromosome 4 was evident; the red signals are strong in the centromeric region of one chromosome, but present on more or less uniform on the other two chromosomes. The chromosome 5 is heteromorphic (see karyotype in Figure 5.1) and so was the *in situ* hybridization pattern with *C. thomasii* labelled DNA. On chromosome 6 and 8, the signals are centromeric and relatively uniform, but on chromosome 7, the signals are more intense on one chromosome as compared to the other two chromosomes (see Figures 5.8, 5.9). It is also clear from the results, that there is less homology of *C. asumaniae* probe to the *C. sativus* chromosomes (compare Figures 5.2, 5.8) suggesting less relationship to the ancestral species. The *C. thomasii* genomic DNA probe (see B, in Figures 5.8, 5.9) like the other genomic probes did not show strong differential labelling of any genome or group of 8 chromosomes. However, within the karyotype, about half of the groups showed a pair of closely similar chromosomes, and one which was more distinct. This is supportive firstly of a close relationship of *C. thomasii* to one of the ancestral genomes, and secondly to a $2x + 1x$, rather than an autotriploid or $x + x + x$ amphitriploid origin of *C. sativus*.

5.3.3.3 *In situ* hybridization with genomic DNA from *C. hadriaticus* and *C. mathewii*

Figure 5.10 illustrates the *in situ* pattern of *C. sativus* metaphase chromosomes hybridized with labelled total genomic DNA from *C. hadriaticus* (detected in red) and *C. mathewii* (detected in green). Two representative metaphases are given, where the results are comparable. Good quality spread metaphase chromosomes were not

available, and the current cells did not allow a cut-out of individual chromosomes into pairs of three, as previously carried out for the other genomic probes (see Figures 5.6-5.9). The labelled *C. hadriaticus* genomic DNA produced dispersed weak hybridization signals on all chromosomes, except for a few chromosomes (less than eight) where the signals are quite strong and uniform along the whole chromosomes (compare B and F in Figure 5.10). Similar to other genomic probes, the *C. hadriaticus* probe also labelled the centromeric and sub-centromeric regions of most chromosomes (see B, Figure 5.10). Whereas the *C. mathewii* total genomic DNA weakly labelled the entire genome of *C. sativus* but specifically labelled few chromosomes (most probably 3) and the 45S rDNA regions (compare C and G in Figure 5.10). The prominent DAPI bands that most probably representing the sub-metacentric repetitive DNA region of the large chromosome is also labelled by the *C. mathewii* genomic DNA (see C in Figure 5.10). Previous labelling of total genomic DNA from several species generated probes which specifically label the 45S rDNA regions (see for example Figures 5.2, 5.9), so it is not surprising that the *C. mathewii* label shows this hybridization pattern (see C in Figure 5.10). The 45S rDNA may represent almost 5% of the entire DNA present in *C. sativus* (see Figure 5.1) and the current results further revealed the overall affinity of *C. hadriaticus* to be much higher with the *C. sativus* as compared to the *C. mathewii*.

5.3.3.4 *In situ* hybridization with genomic DNA from *C. oreocreticus* and ‘*C. sativus cartwrightianus*’

Comparable *in situ* hybridization patterns were obtained in both metaphases of *C. sativus* probed with labelled genomic DNA from *C. oreocreticus* (B, F detected in red) and the garden origin ‘*C. sativus cartwrightianus*’ (C, G detected in green Figure 5.11). The overall hybridization of both *C. oreocreticus* and ‘*C. sativus cartwrightianus*’ indicated that DNA sequences from genomic DNA of both show similarity to major regions of all 24 chromosomes in *C. sativus*. The *C. oreocreticus* probe weakly labelled the entire chromosomes and specifically labelled the centromeric, sub-centromeric regions (see B, F in Figure 5.11). Similarly, the DAPI-positive bands on large chromosomes as well as the rDNA regions are also strongly labelled. Interestingly, some interstitial regions among *C. sativus* chromosomal arms showed strong hybridization.

Although, no experimental evidence is available at the moment but such interstitial high intensity signals may be compared to chromosomal translocations (see C in Figure 5.11).

Except for some minor discrepancies, the *in situ* hybridization results obtained with '*C. sativus cartwrightianus*' is largely uniform in both metaphases (C, G in Figure 5.11). The unevenness could be due to the loss of some repetitive DNA during labelling, or denaturation steps involved in GISH, or larger regions of single copy DNA that are not labelled strongly. However, in both metaphases, the genomic hybridization signals were seen on most chromosomes, and some regions particularly the centromeric, pericentromeric regions and rDNA sties showed strong hybridization (C and G in Figure 5.11). Similar to other genomic probes, the prominent DAPI positive sub-centromeric bands collate with '*C. sativus cartwrightianus*' and indicate the presence of related families of repetitive DNA within the members of the genus *Crocus*. My previous molecular marker results revealed a close relationship of '*C. sativus cartwrightianus*' with *C. sativus* (see chapters III, IV); indeed the *in situ* results too confirmed this and indicated that '*C. sativus cartwrightianus*' under high stringency conditions labelled most of the chromosomes uniformly.

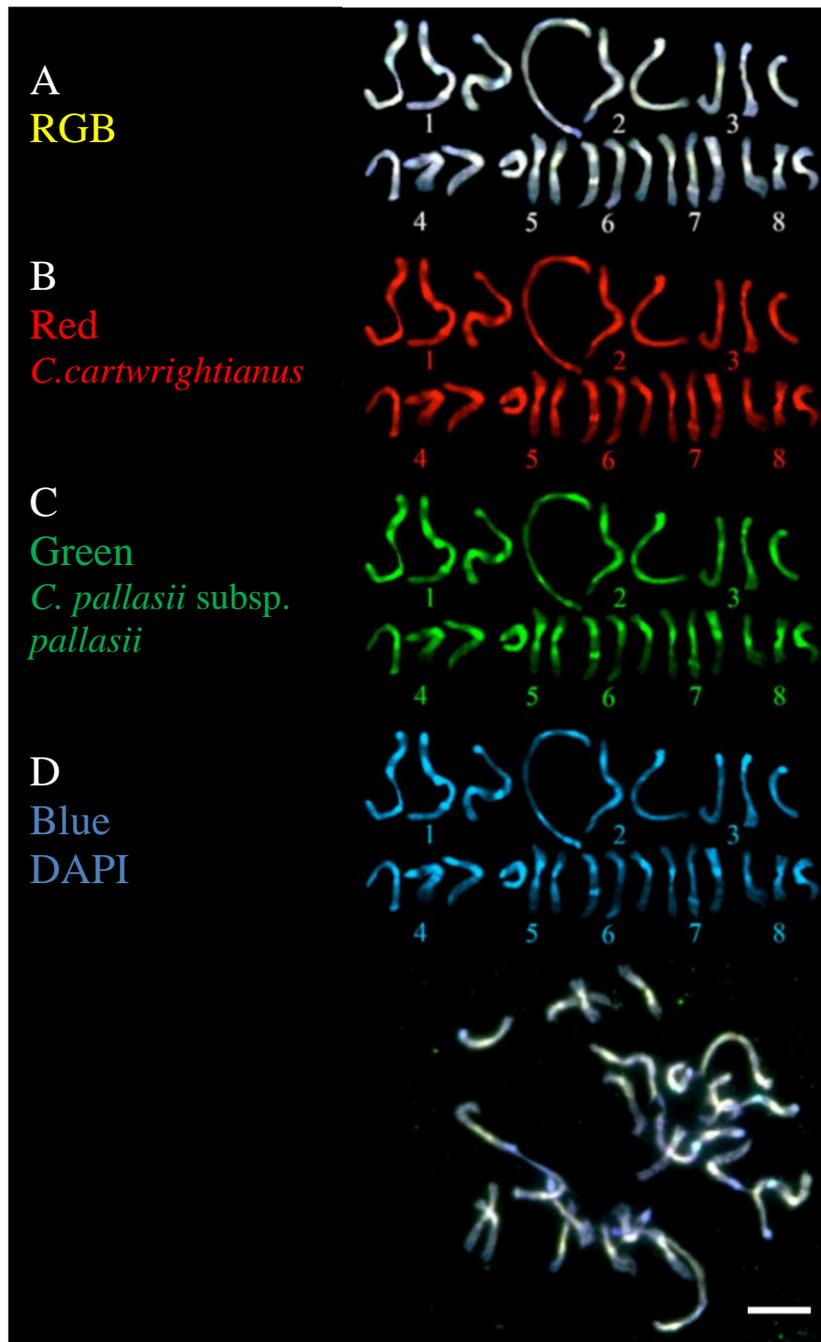


Figure 5.6: Prometaphase karyotype of *C. sativus* labelled with genomic DNA from *C. cartwrightianus* (in B, biotin label, detected in red) and *C. pallasii* subsp. *pallasii* (in C, digoxigenin label, detected in green), and the DNA stained with DAPI (in D), and with channels as overlay in the cut-out karyotype (A) and complete chromosome figure (bottom). The cut out is presented conventionally with long arm uppermost and in size order from longest to shortest and three similar chromosomes were placed in each group 1 to 8 in the karyotype. Bar represents 5 μ m.

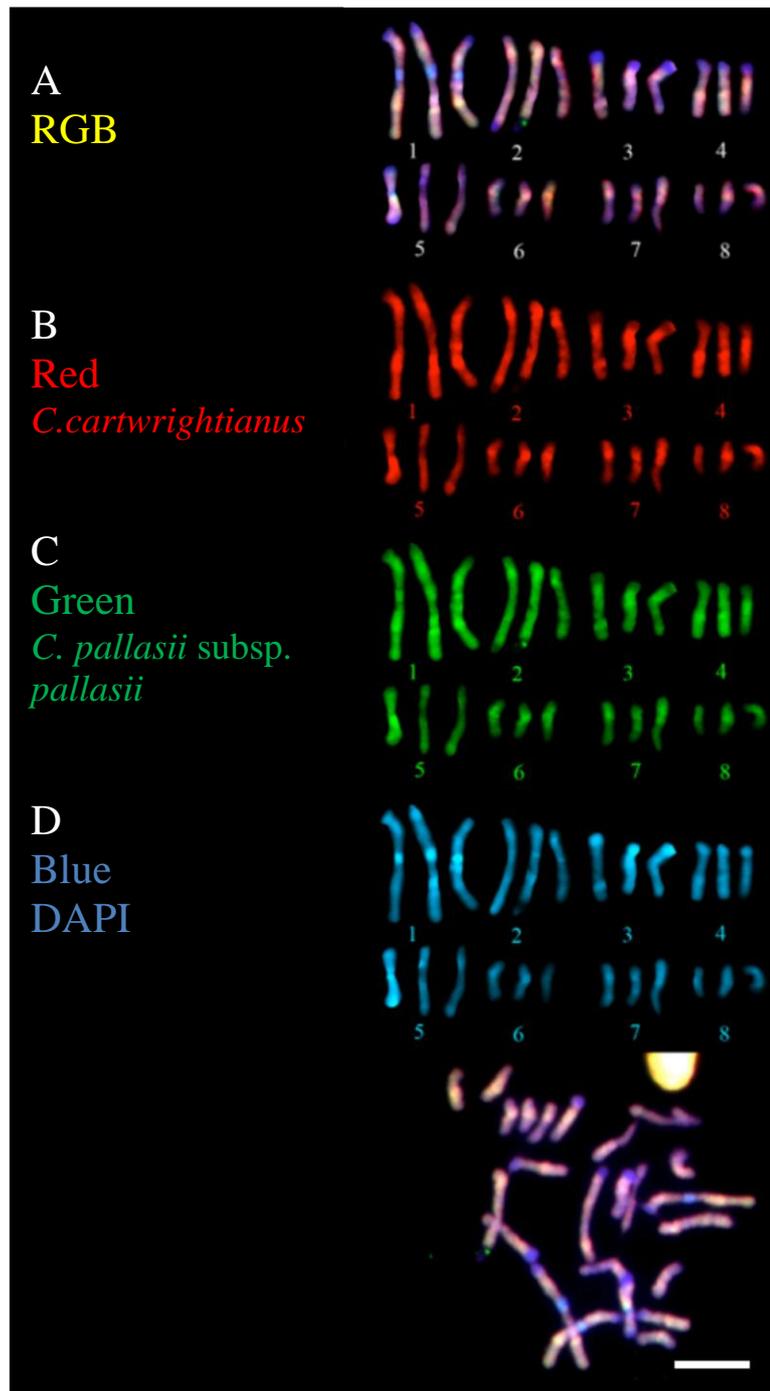


Figure 5.7: Metaphase karyotype of *C. sativus* labelled with genomic DNA from *C. cartwrightianus* (in B, biotin label, detected in red) and *C. pallasii* subsp. *pallasii* (in C, digoxigenin label, detected in green), and the DNA stained with DAPI (in D), and with channels as overlay in the cut-out karyotype (A) and complete chromosome figure (bottom). The cut out is presented conventionally with long arm uppermost and in size order from longest to shortest and three similar chromosomes were placed in each group 1 to 8 in the karyotype. Bar represents 5 μ m.

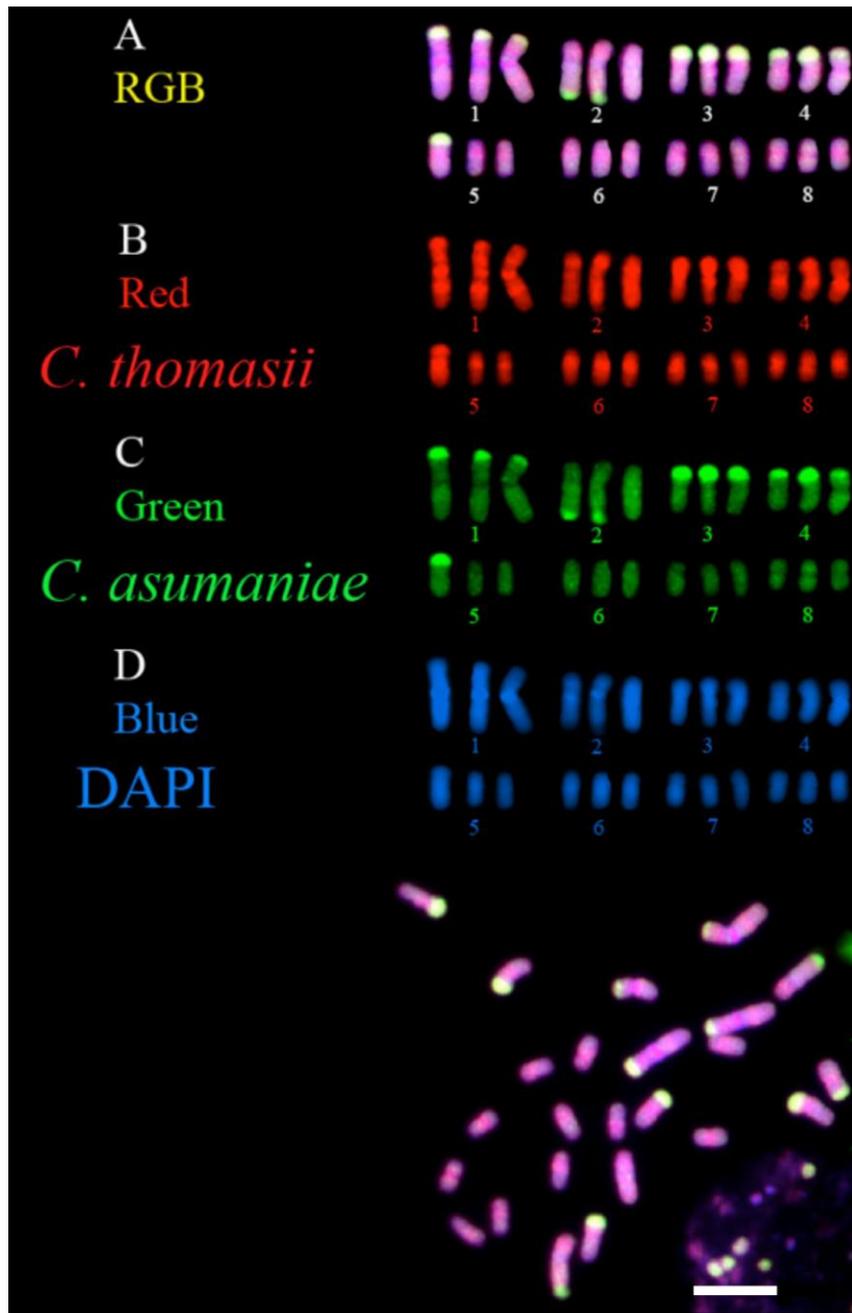


Figure 5.8: Metaphase karyotype of *C. sativus* labelled with genomic DNA from *C. thomasi* (in B, biotin label, detected in red) and *C. asumaniae* (in C, digoxigenin label, detected in green), and the DNA stained with DAPI (in D), and with channels as overlay in the cut-out karyotype (A) and complete chromosome figure (bottom). The cut out is presented conventionally with long arm uppermost and in size order from longest to shortest and three similar chromosomes were placed in each group 1 to 8 in the karyotype. Bar represents 5 μ m.

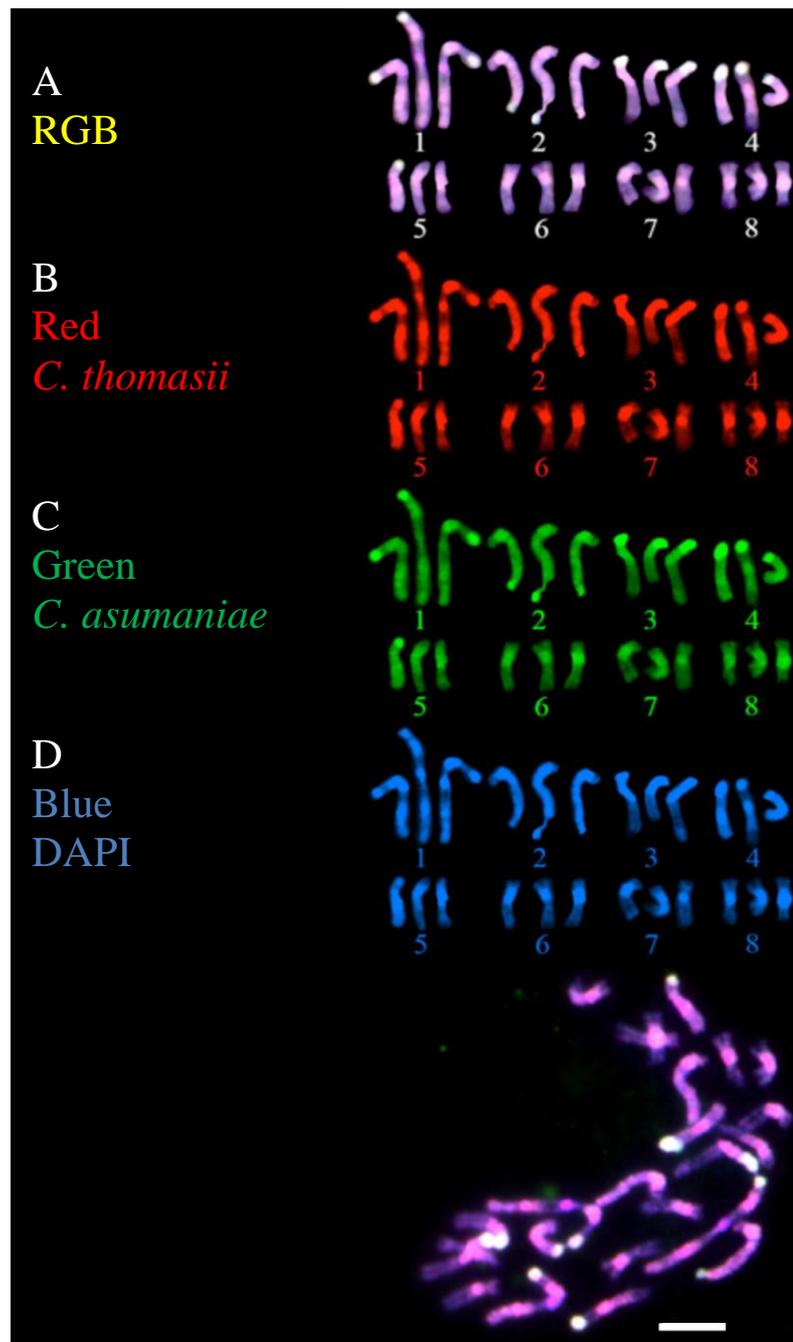


Figure 5.9: Metaphase karyotype of *C. sativus* labelled with genomic DNA from *C. thomasi* (in B, biotin label, detected in red), *C. asumaniae* (in C, digoxigenin label, detected in green), DNA stained with DAPI (in D), channels as overlay in the cut-out karyotype (A) and complete chromosome figure (bottom). The cut out is presented conventionally with long arm uppermost and in size order from longest to shortest. Three similar chromosomes were placed in groups 1-8 in the karyotype. [Bar represents 5µm.](#)

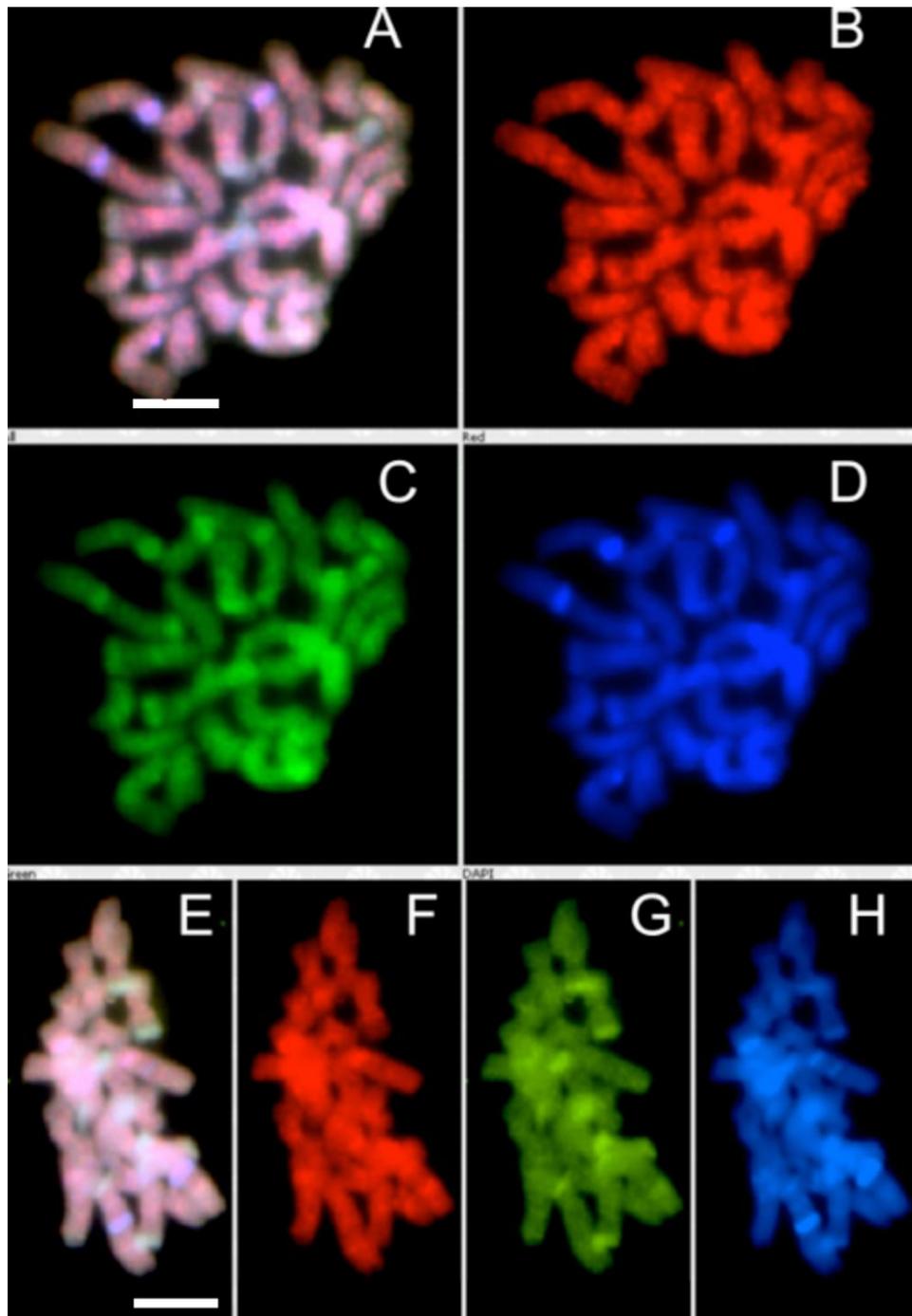


Figure 5.10: Root-tip metaphase chromosomes of *C. sativus* ($2n=24$) after fluorescent *in situ* hybridization (GISH) of the total genomic DNA from *C. hadriaticus* (**B, F**) labelled with biotin 16-dUTP (detected in red) and *C. mathewii* (**C, G**) labelled with digoxigenin 11-dUTP (detected in green). *C. sativus* chromosomes fluorescent blue with DAPI (**D, H**). Overlay of red, green and blue filters (**A, E**). Bar represents 5 μ m.

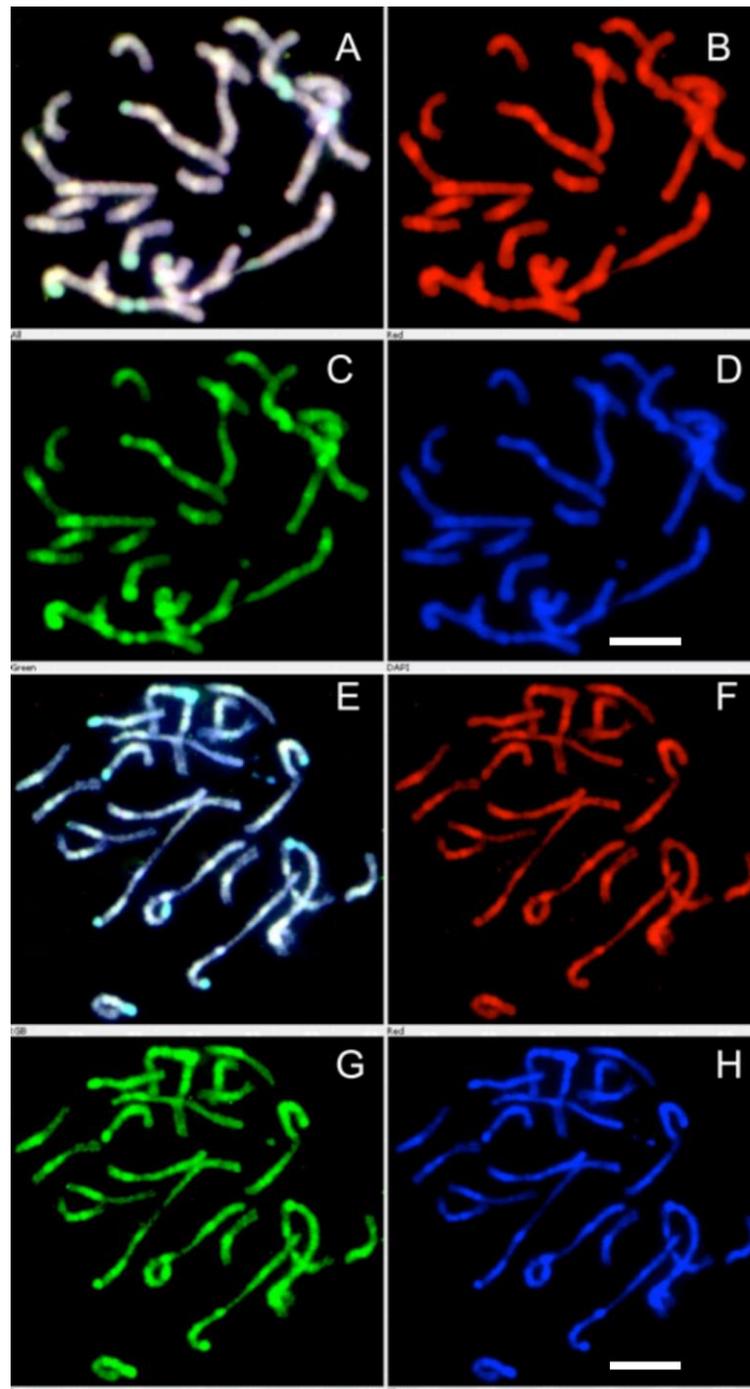


Figure 5.11: Root-tip metaphase chromosomes of *C. sativus* ($2n=24$) after fluorescent *in situ* hybridization (GISH) of the total genomic DNA from *C. oreocreticus* (**B**, **F**) labelled with biotin 16-dUTP (detected in red) and '*C. sativus cartwrightianus*' (**C**, **G**) labelled with digoxigenin 11-dUTP (detected in green). *C. sativus* chromosomes fluorescent blue with DAPI (**D**, **H**). Overlay of red, green and blue filters (**A**, **E**). Bar represents $5\mu\text{m}$.

5.4 Discussion

5.4.1 Ancestral species of *C. sativus*

The current phylogeny of the genus *Crocus* is based on an extensive set of published morphological and molecular data that consistently put *C. sativus* closely associated with other members of the genus, in *Crocus* series *Crocus* (see Mathew, 1982; Petersen *et al.*, 2008; Seberg and Petersen, 2009; Harpke *et al.*, 2013). Earlier, potential ancestral species of *C. sativus* were identified using IRAPs, EST-SSRs, barcoding genes and SNP markers (see Chapters III and IV). Here, genomic *in situ* hybridization was used on both the meiotic as well as the mitotic chromosomes to infer the parental genome during the cell cycle to therefore identify the ancestral genome of *C. sativus*. The molecular cytogenetics approach is based on the phylogenetic information gained from previous result chapters, augmented with published reports where *C. sativus* is placed closely to related species in series *Crocus*: *C. cartwrightianus*, *C. thomasi*, *C. pallasii* (different subsp. *pallasii* accession) *C. hadriaticus*, *C. oreocreticus* and *C. mathewii* (also see Table 5.1). The relatively strong genomic *in situ* hybridization seen with all the species used here (Figures 5.6 to 5.11) confirms the close relationship of the repetitive DNA sequences present in all the genomes used. Within the overall hybridization patterns, I was able to interpret differences between the species used as genomic probes, although the results will need additional confirmation.

Some variation was observed in the genomic hybridization pattern among different batches of probes prepared from the same species, as has been found with genomic *in situ* hybridization with many species where the repetitive DNA component is not consistently labelled. The results shown were obtained after optimizing *in situ* hybridization conditions and changing the standard protocol (see Schwarzacher and Heslop-Harrison, 2000) that included GISH with and without blocking DNA, high and low stringency washing conditions, denaturation of the chromosomes and probe at different temperatures etc. By and large, consistent results were obtained using blocking DNA (20-30x of probe concentration) and then subsequent high stringency washing condition of 75-80%. The *in situ* results obtained with genomic probes labelled with different nucleotides (biotin and digoxigenin) were comparable not only on mitotic

spreads, but also in between mitotic and meiotic chromosomes (compare Figures 5.4A with 5.8, 5.9 and 5.4B with 5.10).

Taken together, the *in situ* hybridization patterns (Figures 5.6, 5.7) indicate that DNA sequences from genomic DNA of *C. cartwrightianus*, *C. thomasi*, *C. pallasii* subsp. *pallasii* and '*C. sativus cartwrightianus*' indicate the highest homology to major regions of all 24 chromosomes in *C. sativus* (Figures 5.6, 5.7, 5.11). Among the species belonging to the series *Crocus*, several studies have indicated *C. cartwrightianus* to be the most probable candidate for the origin of saffron (Brighton, 1977; Mathew, 1977; Petersen *et al.*, 2008). Similarly, several recent studies have also shown the affinity of *C. pallasii* subsp. *pallasii* with *C. sativus* (Sanei *et al.*, 2007; Erol *et al.*, 2013). The results obtained here support to previous findings, but partially contradict the findings of Grilli Caiola *et al.* (2004), where the authors consider *C. pallasii* to be more distantly related to *C. sativus* (see also Chapters III, IV). In fact both *C. cartwrightianus* (2n= 16) and *C. pallasii* subsp. *pallasii* (2n=14) are fertile, diploid, autumn flowering plants. Further, both occur in the wild and are found in areas overlapping with *C. sativus* or saffron cultivation (Brighton, 1977; Mathew, 1992; Grilli Caiola and Canini, 2010). Thus the possibility of their recent or past hybridization resulting in triploid saffron seems very likely.

The '*C. sativus cartwrightianus*' is an unrecognized species (see Table 2.2) presumably a garden-origin variant, or hybrid, which has both morphological and DNA based similarities with *C. cartwrightianus* cv. *albus* (see A4, A4 Figures 1.2 and Chapter III). In spite of this, *C. cartwrightianus* and *C. cartwrightianus* cv. *albus* have 16 chromosomes, and still both have visible differences not only in morphology but also in pollen grain structure and germination (Karasawa, 1956; Grilli Caiola, 1995). Nevertheless, both *C. cartwrightianus* cv. *albus* and *C. sativus* are infertile, and anomalous pollen grain percentage, pollen size and *in vitro* percentage pollen germination together with ISSR profile, suggest that *C. cartwrightianus* cv. *albus* is more similar to *C. sativus* than some accessions of the purple form of *C. cartwrightianus* (Rubio-Moraga *et al.*, 2009; Grilli Caiola and Canini, 2010). There is substantial intraspecific variation within *C. cartwrightianus* accessions in morphology (e.g. stigma length; personal observation see A1, A2 Figures 1.2) and DNA markers (Chapters III and IV). The GISH results also indicated the DNA sequence-based similarities between

C. cartwrightianus and '*C. sativus cartwrightianus*' (Figure 5.11). Because of the close relationships of the two *Crocus* species and the *albus* variant to each other and the hybrid species (saffron), there is no clear discrimination of one or two sets of 8 chromosomes, contrasting with, for example, the triploid hybrid *Crocus* 'Golden yellow' (*C. flavus* x *C. angustifolius*, species in different sections of the genus), and many other hybrids from different plant families, where genomic *in situ* hybridization clearly discriminates the ancestral origin of chromosomes in hybrids (see Ørgaard *et al.*, 1995; Schwarzacher *et al.*, 2003a).

Much of the dispersed *in situ* hybridization signal is likely to originate from transposable elements, because they are dispersely distributed along the chromosomes. It will be interesting to analyse sequences of larger numbers of TE from both the species used as probes: these can now be obtained and analysed with high-throughput DNA sequencing approaches. It may be that, as in *Brassica oleracea* with a CACTA transposon, regions show species-specific amplification (Alix *et al.*, 2008) and may be usable for identification of the ancestral origins of the different chromosomes. The stronger, more band-like signals seen in both *in situ* hybridization and DAPI staining suggest the presence of distinctive tandemly repeated DNA sequences (see Heslop-Harrison and Schwarzacher, 2011). However, cloning experiments during this work and also by Frello and Heslop-Harrison (2000a) did not identify abundant tandem repeat families. Although the authors (Frello *et al.*, 2004) isolated some centromeric tandem repeats but these did not show discrimination between species or chromosomes, in contrast to the genomic DNA probe here where differences were detected. Again, it will be valuable to analyse large amounts of genomic DNA sequence from *C. sativus* and other *Crocus* species to examine the nature and evolution of the repetitive DNA families present.

A small-scale cloning and screening experiment in the course of the present work has identified some candidate repetitive sequences (see Appendix 5). Although time constraints did not allow either Southern hybridization or *in situ* hybridization, to explore the long range and physical organization of these isolated repetitive DNA elements within the genus *Crocus*; these experiments will form important tools of future work (Appendix 5). Repetitive DNA sequences located around the centromeres of the largest chromosome in *C. sativus* will be of particular interest (Figure 5.10). It is likely

that high throughput sequencing will reveal large numbers of tandem repeats and using bioinformatics strategy we may be able to see if certain repeats are chromosome specific. Such investigation will have several implications in future introgression, genomic enrichment of saffron.

Table 5.3: Brief summary of the GISH analysis of *C. sativus*.

Labelled genomic probe	Metaphase chromosomes							
	<i>C. sativus</i>	<i>C. sativus</i>	<i>C. sativus</i>	<i>C. sativus</i>	<i>C. sativus</i>	<i>C. sativus</i>	<i>C. sativus</i>	<i>C. sativus</i>
<i>C. thomasi</i>	Eu +/+ Het +/-							
<i>C. asumaniae</i>		Eu -/- Het +/+						
<i>C. hadriaticus</i>			Eu +/+ Het +/+ Het +/-					
<i>C. mathewii</i>				Eu +/+ Het +/+				
<i>C. cartwrightianus</i>					Eu +/+ Eu +/- Het +/-			
<i>C. oreoreticus</i>						Eu +/+ Eu +/- Het +/-		
<i>C. pallasii</i> subsp. <i>pallasii</i>							Eu +/+ Eu +/- Het +/-	
<i>C. sativus cartwrightianus</i>								Eu +/+ Eu -/- Het +/-
Labelled genomic probe	Euchromatin signals				Heterochromatin signals			
<i>C. thomasi</i>	Uniform, 8-12 strong, others weak				Large strong centromeric (particularly on small chromosomes) and terminal bands, some are NORs			
<i>C. asumaniae</i>	Uniform weak signals on all 24 chromosomes				Strong NORs and centromeric to whole small arm, often only localized to NORs			
<i>C. hadriaticus</i>	Up to 8 strong, remaining uniform weak signals				Centromeric or terminal to whole arm signal on 12-16 chromosome, some are NORs			
<i>C. mathewii</i>	Signals similar to <i>C. hadriaticus</i> , 8 strong, others uniform weak signals				Similar to <i>C. asumaniae</i> , strong labelling of all NORs, DAPI bands			
<i>C. cartwrightianus</i>	16 strong, the intensity of signals is often not the same on both arms, 8 weak, sometimes no big difference				Strong centromeric or terminal to whole arm signal on 21 chromosome, some are NORs			
<i>C. oreoreticus</i>	Uniform to all 24 chromosomes, some have sub-terminal strong signals.				Large sub-telomeric, centromeric signals, some NORs, sometime interstitial regions only			
<i>C. pallasii</i> subsp. <i>pallasii</i>	Very similar to <i>C. cartwrightianus</i> , but often localized strong interstitial signals				Similar to <i>C. cartwrightianus</i> , sometimes strong hybridization to NORs			
<i>C. sativus cartwrightianus</i>	16-20 strong, sometime all uniform signals				Strong NORs and centromeric to whole arm, hybridization to DAPI bands			

5.4.2 Karyotype and allopolyploid nature of *C. sativus*

Karyological complexity of the genus *Crocus* is known and previous studies have revealed variability in *C. sativus* chromosome number from $2n=14$, 15, 20, 24 and 40 (Sugiura, 1931; Mather, 1932; Karasawa, 1935, 1942; Ghaffari, 1986; Fernandez *et al.*, 2009). Cytological examinations reported here, from 2010 to 2014 however, revealed no less than $2n=24$ mitotic chromosomes. Therefore, authors who counted less or more than 24 chromosomes in *C. sativus* are most likely mistaken, either through wrong identification of material or the cytological method, including counting of satellites as complete chromosomes. Similar results to these reported here were obtained by Agayev (2002) and Fernandez *et al.* (2009), the authors investigated *C. sativus* karyotype by analyzing material of different geographical regions and discovered no differences in chromosomal count. Agayev (2002) and Agayev *et al.* (2010) present karyotype of *C. sativus* based on alkali-hydrolysis to make chromosome preparations and aceto-iron-hematoxylin staining. The chromosome morphology from their *C. sativus* accession was consistent with that reported here; however, they found a difference in the satellite (45S rDNA) in the Kashmirian accession and assumed the Spanish saffron is a cultivar genetically different to the Kashmiri saffron. However, in the current analysis, no such differences were observed (see Figure 5.1) and the variation may have occurred in somatic material. Thus it may be argued that *C. sativus* originated only once in history, and then subsequently spread, undergoing certain genetic or phenotypic changes in the subsequent evolution. I could also observe slight morphological differences amongst the different accessions as well as at the genetic level based on EST-SSR and SNP markers (Figure 1.4 and Chapter IV).

Karyotyping is an effective tool for the physical comparison of chromosomes and may provide direct insight into the ancestral species involved in hybrid formation (Frello *et al.*, 2004). The karyotype of *C. sativus* based on the length of chromosomes, rDNA sites, centromeric position and DAPI bands reveal e.g. eight groups of three chromosomes (Figures 5.1, 5.6, 5.7). The majority of previous studies have suggested *C. sativus* to be autotriploid (Chapter I, III and IV) and its sterility has been attributed to its aberrant pollens, because the female gametophyte is less frequently sterile and some introgression is possible (see Grilli Caiola and Canini, 2010). Karasawa (1935), considered saffron is an autotriploid species whose chromosomes at metaphase form

eight trivalents, and proposed its probable origin from a diploid *Crocus*. Similarly, Ghafarri (1986) interpreted his meiotic figures with frequent trivalents as showing *C. sativus* is autotriploid, although, he also observed other meiotic figures of different pairing configurations at low frequencies. He also shows chromosomes lagging at anaphase and possible bridges, features that could lead to formation of restitution nuclei and multiple pollen mother cell formation as illustrated in (figure 5.5) above. All the species used for genomic probing, showed some degree of labelling on to the chromosomes of *C. sativus*, which does indicate that the genomes are indeed closely related. However, both the meiotic preparations (Figures 5.3 to 5.4) and mitotic *in situ* hybridization (Figures 5.6 to 5.11) do not support autotriploidy, since one of the three genomes was often different in hybridization with genomic DNA probes, and meiotic figures showed variable pairing (Figures 5.3, 5.4 and below).

Meiotic pachytene analysis of *C. sativus* revealed the presence of univalents, bivalents and trivalents. Presence of trivalent chromosomes supports autotriploidy, while univalents, bivalents and quadravalents suggest an allotriploid origin of *C. sativus*. By and large, the *in situ* results indicated that the three genomes are similar to each other and at least two and maybe all the three genomes possibly arose from different ancestral species (Figures 5.4, 5.7). Both bivalents and trivalents chromosomes indicate that no two genomes are exactly similar but that two are more closely related than the third one (Figure 5.3). However, more work needs to be done, and detailed meiotic analysis will be extremely helpful in resolving the phylogeny of *C. sativus*. Unlike hexaploid *Triticum aestivum*, where a single major locus, the *Ph1* is responsible for its true diploid-like behaviour (Hao *et al.*, 2011), in *Crocus* genes controlling meiotic pairing are unknown. Mechanisms controlling pairing of chromosomes could be more complex. Varying degrees of pairing between chromosomes, heterotypic meiosis and restitution nuclei were seen (Figures 5.2 - 5.5).

Allotriploids are believed to form a higher frequency of crossovers than their diploid progenitors (Leflon *et al.*, 2010). Further, two diploid species, *C. cartwrightianus* and *C. pallasii* subsp. *pallasii* was found to be closely related to *C. sativus* (Figures 5.3, 5.4). It could be assumed from the analysis that if the two genomes are highly similar in *C. sativus*, then possible evidence of homoeologous pairing amongst these genomes may be seen. Indeed, the *in situ* hybridization with *C. pallasii*

subsp. *pallasii* indicated possible translocation within the chromosomes (see Figure 5.7). Further, the karyotypes of both species show similarities with *C. sativus* (Sanei *et al.*, 2007; Agayev *et al.*, 2009) and particularly the presence of odd chromosome in the karyotype of *C. pallasii* subsp. *pallasii* and *C. sativus* provides further evidence of their shared ancestry (see Figure 5.1 and Table 5.4). A high degree of variation in the karyotype of *C. pallasii* subsp. *pallasii* (from Western and Central Turkey) as well as cytological similarities between *C. pallasii* subsp. *pallasii* with *C. dispathaceus* have been reported (Brighton, 1977). Being closely related, growing in adjacent areas and having the same number of chromosomes the author proposed possible hybridization between the two subspecies. It is very likely such a hybridization event may result in a series of chromosomal interchanges and/or pericentric inversion that may result in the origin of odd chromosome present in the karyotype of *C. sativus* today (Figure 5.1 and Table 5.4).

There is no doubt that a better understanding of the *Crocus* phylogeny and chromosome evolution in *Crocus* could be gained if genomic DNA for PCR-based results (see Chapters III, IV), probes (labell DNA) for *in situ* hybridization and chromosome numbers (of the potential ancestors) were checked from plants coming from natural populations. However, to date no study that addressed *C. sativus* ancestry, and where the material is obtained directly from natural populations. In virtually all cases, *Crocus* corms are supplied by commercial suppliers and rarely come from collections of scientific organizations, so the true origin of the plants cannot be ascertained (Table 2.2). Further, the commercial suppliers (nurseries) grow *Crocuses* for ornamental and gardening purposes and therefore, the identity of plants in which chromosomes were counted remains doubtful. Future work, using all members of the *Crocus* series *Crocus* needs to be done from natural populations, as accessing the karyotypes from species collected in wild will make a direct comparison and can be used to support the molecular markers results.

Table 5.4: Karyotype of the members of *Crocus* series *Crocus*

No	<i>Crocus</i> species	Chromosome Number (2n)	Chromosome 1	Chromosome 2	Chromosome 3	Chromosome 4	Chromosome 5	Chromosome 6	Chromosome 7	Chromosome 8
1	<i>Crocus sativus</i>	24	3 large acrocentric with satellite on the short arms	3 large acrocentric with small satellite on the long arms	Medium size sub-metacentric	Medium size Metacentric	1 metacentric 2,3 acrocentric	Small acrocentric	Small acrocentric	Small metacentric
2	<i>C. cartwrightianus</i>	16	Large acrocentric chromosomes	Large acrocentric with small satellite on the long arms	Small metacentric	Small metacentric	Small metacentric	Small metacentric	Small metacentric	Small acrocentric
3	<i>C. hadriaticus</i>	16	Large acrocentric chromosomes	Large acrocentric with small satellite on the long arms	Sub-metacentric or acrocentric chromosomes	Sub-metacentric or acrocentric chromosomes	Sub-metacentric or acrocentric chromosomes	Sub-metacentric or acrocentric chromosomes	Sub-metacentric or acrocentric chromosomes	Sub-metacentric or acrocentric chromosomes
4	<i>C. thomasii</i>	16	Large acrocentric chromosomes	Large acrocentric with small satellite on the long arms	Metacentric or sub-metacentric	Metacentric or sub metacentric	Metacentric or sub metacentric	Metacentric or sub-metacentric	Small acrocentric	Small acrocentric
5	<i>C. pallasii</i> subsp. <i>pallasii</i>	14	Large acrocentric, Sub-metacentric	Large acrocentric, Sub-metacentric	One smaller pair of acrocentric with satellites on the long arms sub-metacentric	Metacentric and sub-metacentric	Metacentric and sub metacentric	Metacentric and sub-metacentric	Metacentric and sub-metacentric	
6	<i>C. pallasii</i> subsp. <i>dispathaceus</i>	14	Acrocentric	Acrocentric with small satellite on the long arms	Acrocentric	Metacentric	Metacentric	Smaller sub-metacentric	Acrocentric	
7	<i>C. pallasii</i> subsp. <i>turcicus</i>	12	Large metacentric with small satellites on one arm	Acrocentric	Acrocentric	Acrocentric	Smaller metacentric	Smaller metacentric		

6 CHAPTER VI: GENERAL DISCUSSION

The conclusions from individual results chapters have been discussed in the appropriate sections above. The general discussion here aims to show the overall progress towards the objectives set out in (Chapter I Introduction) and highlight a few broader implications of the work for both academic and applied areas. The current work also discusses future research opportunities as well as identifying prospects of the overall project.

6.1 Genome studies and breeding in orphan crops “Cultigens”

Over the years, plant breeders have been remarkably successful in developing new and highly productive cultivars of all major crops with desired traits. There has been a tremendous increase in crop productivity, and it may be attributed for a good reason to the application of Mendel’s principles in breeding as well as to a better understanding of the crop genomics (see Chapter I). All that has been achieved is inspite of the rapid emergence of more virulent races of pathogens, and in more disturbed and changing global environmental conditions (Chakraborty and Newton, 2011). So far the applications of traditional plant breeding practices, coupled with more recent genetic engineering approaches, have succeeded in steadily increasing and maintaining food for the ever-growing population (see Borlaug, 1983). However, when the global population reaches nearly 9.4 billion people in 2050, the challenges for agriculture will be overwhelming (<http://www.fao.org>). Perhaps, key to the success stories in agriculture are all related directly or indirectly to the identification and maintenance of useful biodiversity within species.

The dried stigmas of *C. sativus* (saffron) are highly valued since antiquity, and consumed as a spice or drug of immense significance, with hardly any sector of life where saffron has no applications. Today is a great time for scientists interested in whole genome sequences and large-scale genomics. Still, inspite of all the recent advances in modern agriculture systems, and in generating molecular data, reconstructing species-level phylogenies and identification of useful diversity in non-model crops such as *C. sativus* remains a challenge. The earliest farmers knowingly or unknowingly maintained useful genetic variation in nearly every species that could be chosen from the wild. These ancient farmers planted, harvested and reselected these species in order to gradually develop

improved populations with a range of desirable traits. However, after the sudden genetic selection leading to the first domesticated species and the birth of agriculture, perhaps marked initiation of the loss of genetic diversity. Even today, the desirable genes, which were either selected by man (the first breeder) or nature itself, are dispersed within both the domesticated and wild plant populations and can be reselected if required (see Vaughan *et al.*, 2007). In saffron too, variation may be confined to natural resources. In the first place identifying and preserving any diversity in *C. sativus* is of utmost importance: a project to secure vital diversity in *Crocus* was initiated in the form of “the CrocusBank project” (see www.crocusbank.org/). The failure to secure local land races that carry genetic variation coupled with destruction of natural habitats in the Mediterranean possibly added to the reduction of genetic diversity in saffron. Further, sterility and exclusive vegetative propagation in saffron result in offspring’s that are genetically identical to the mother corms. Further, lack of mechanization and people no longer interested to labour in fields for saffron cultivation, and urbanization, along with the global climatic changes had also some role in loss of locally grown saffron materials that cannot be neglected (Fernandez *et al.*, 2011 and Chapter I).

Hybrids may deliver higher yields and better quality than their parents alone (Madlung, 2013). Identification of the ancestral parent species for *C. sativus* holds much promise for advances in *Crocus* agriculture. Identification of the potential donor parents will not only help in reintroducing novel biodiversity into saffron, by exploiting the wild ancestors, and if certain traits could be transferred into saffron, it would allow the fixation of heterosis and a set of given genetic combinations. Further, understanding of the genetic mechanisms underpinning sterility may allow it to overcome so has potential implications for saffron growers as well as in modern agriculture and may open gateways for a technology that could exploit and facilitate triploid hybrid vigor. Thus this may pave a way to hybrid seed or other propagation systems in saffron as well as for the exploitation of sterile plants with higher ploidies. To date, the potential ancestry of *C. sativus* is not resolved and the current work, identified candidate ancestral species using molecular cytogenetic approaches (see Chapters III, IV, V). The availability of whole genome sequences for *C. sativus* and potential diploid ancestors as well as the generation of complementary large-scale transcriptomic and protein interaction data will opens novel avenues of research that would allow exploitation and may facilitate the indefinite use of hybrid vigor. Thus in future we may possibly obtain true seed

from a crop that is currently propagated vegetatively (also see Van de Peer and Pires, 2012; Madlung, 2013).

6.2 Origin and genetic diversity in *Crocus sativus*

The results here show that, outside *C. sativus*, there was a high level of diversity, evident at morphological (varietal for cultivated species) and genetic level. Individual accessions of *Crocus* species obtained from different sources and often individual corms, while clearly related, showed extensive variation at the DNA level in nearly every assay. Interestingly, accessions of the same species purchased from different nurseries showed significant differences, for example comparing the purple flowering *C. cartwrightianus* where the *C. cartwrightianus* accession CcrCR09 (Rare Plants) which is grouping very closely with *C. sativus* while CcwBD09 (JW Dix Export, The Netherlands) is on the outer periphery of the group (see sub-clade D1, Figure 3.6). No doubt, all species within *Crocus* series *Crocus* are closely related and except for *C. sativus*, all others are fertile and hence diploid (or possibly tetraploid and acting as pseudo-diploid). Indeed, variation within species for sequences and IRAP patterns was often as high as between different *Crocus* species. The results obtained here are generally consistent with other molecular studies of wild species that are sexually reproducing, although the new results found higher variation than in many other groups: the molecular analysis of the DNA also showed considerable differences between all species within the *Crocus* series *Crocus*. In the IRAPs, this variation was typically greater than that within species (see Chapter III). However, the sequence analysis here identified almost no well-supported branches that included only a single species. Thus, in agreement with other studies, it is likely that speciation has occurred relatively recent (see Chapter IV). Unlike the situation in many other genera with unknown relationships, the markers used here were not able to resolve any major evolutionary lineages within *Crocus* series *Crocus*, although they did support the monophyletic origin of the natural series.

Exclusive asexual multiplication in *C. sativus* (saffron), that lacks recombination, have revealed the existence of limited genetic variability within saffron grown worldwide, consistent with most of the previous studies. Despite the 6000 high quality *Crocus* ESTs (see D'Agostino *et al.*, 2007) there were a limited number of suitable ESTs to develop useable EST-SSR markers for *Crocus* (Chapter IV). The overall published results taken together

suggest that *C. sativus* is a single clone that has been propagated vegetatively and distributed over much of the world. The results showed here too with the IRAPs strongly support the clonal origin and subsequent vegetative spread of saffron germplasm.

There is minimal variation in *C. sativus* at the morphological (Figure 1.4) as well as at the DNA level, although existence of single-locus (gene) somatic mutations cannot be ruled out (see Chapter IV). Except for a minor discrepancy, the *in situ* pattern was consistent and the *in situ* hybridization with genomic probes was informative and gave discrimination along the chromosomes of *C. sativus*. However, the shared ancestry and DNA sequence based similarities between members of *Crocus* series *Crocus* did not allow clear discrimination of one or two sets of 8 chromosomes, as for example previously carried out for the triploid hybrid *Crocus* 'Golden yellow' *C. flavus* x *C. angustifolius* (Ørgaard *et al.*, 1995). It is noteworthy, that *C. sativus* is most probably an allopolyploid species, with ancestors arising within the *Crocus* section of the genus, and *C. cartwrightianus*, *C. thomasi* and *C. pallasii* subsp. *pallasii* being the best candidate ancestors (see below). The IRAPs, sequence and cytological evidence presented in (Chapters III, IV and V) respectively support this conclusion.

One reason why GISH did not allow clear discrimination of one or more sets of 8 chromosomes, could be the ancestral form does not exist anymore or its ancestral species has evolved rapidly in the triploid condition. A similar condition is described for bread wheat (genome AABBDD) that originated ~10,000 years ago, where the A (*Triticum urartu*) and D (*Aegilops tauschii*) genome donors are known, but the origin of the B genome is still unclear. Although *Ae. speltoides* (from the Sitopsis section) is seen to be the most likely B genome donor (Feldman and Levy, 2005). It is likely that whole genome approaches such as Genotyping by Sequencing (GBS) developed by Liu *et al.* (2014) will give additional information about the relationships of the genomes. As discussed above with respect to SSRs and IRAPs, the choice of the most appropriate marker system relies not only on the availability of markers, but also on the genetic structure of a species to be examined (Heslop-Harrison and Schwarzacher, 2012), and the results here indicate that whole genome surveys, separated for nuclear and organellar genomes, will be essential to confirm ancestral relationships in saffron.

6.3 Molecular markers, genome diversity and evolution

The CrocusBank project for the first time collected saffron from around the world and provided a baseline for the identification of useful diversity among saffron accessions and to identify the diploid ancestral species. Molecular markers for identification of chromosomes and genomes are important in wide hybridization and alien introgression programmes, which have enabled plant breeders to exploit variation from diverse germplasm. However, given the large genome sizes in *Crocus*, and unknown levels of variation at the start of this study, it was essential to use a complementary approach and a range of different markers from both heterochromatic and euchromatic regions. The results revealed possibilities to use molecular approaches and infer ancestry of *C. sativus* (saffron) as well as to determine relationships of *C. sativus* with other species.

Based on IRAPs profiling *C. sativus* showed maximum similarity to the purple flowering *C. cartwrightianus* (accession CcrCR09), *C. pallasii* subsp. *pallasii* and ‘*C. sativus cartwrightianus*’ (see Figure 3.6 Chapter III). SNPs data indicated similarity of *C. thomasii* with *C. hadriaticus* and suggested *C. oreoreticus* to be one of the potential ancestor species for the two (clade A Figure 4.8A). Furthermore, the sequences from *C. sativus*, *C. asumaniae*, ‘*C. sativus cartwrightianus*’, *C. hadriaticus* and *C. kotschyanus* are found in three clades (clades B, C, D Figure 4.8A) and this not only indicates the relationship of *C. sativus* with these species, but also highlights the potential hybrid nature and reticulate evolution of *Crocus* species. Beside this, the EST data also revealed affinities of *C. asumaniae* with *C. versicolor* and the purple flowering *C. cartwrightianus* (JW Dix Export, The Netherlands) with *C. cancellatus*. The diversity in sexual *C. pallasii* subsp. *dispathaceus* was high grouped with *C. laevigatus* and *C. boryi* instead of clustering with *C. pallasii* subsp. *pallasii* (see Figure 4.10). The EST data also suggested the relationship of *C. asumaniae*, *C. cartwrightianus* cv. *albus* and *C. thomasii* (see Figure 4.10 and Chapter IV for details) and the barcoding genes data complemented the IRAP, SNP and ESTs data, showing *C. cartwrightianus* cv. *albus* grouping with ‘*C. sativus cartwrightianus*’ and we speculated both to be the same species (clade H, Figure 4.13). More importantly, the purple flowering *C. cartwrightianus* (accession CcrCR09) grouped with the Spanish accession of *C. sativus*. Thus it is reasoned that *C. cartwrightianus* (accession CcrCR09) could be one of the top candidate ancestors for *C. sativus* (see Figure 4.13 and Chapter IV for details). The cumulative results

reveals, that *C. sativus* is most similar to *C. pallasii* subsp. *pallasii*, *C. cartwrightianus*, *C. thomasii*, as well as the garden origin '*C. sativus cartwrightianus*'. All of them have karyotype structures similar to *C. sativus*, and except for *C. pallasii* subsp. *pallasii*, all have $2n = 16$ (see Figure 6.1). Further, the results shown here are reproducible and are well supported by published data, where the authors used different approaches to answer the puzzle of *C. sativus* ancestry. Thus the current results in addition to indicating the allotriploid nature, also illustrated this species may be considered as a possible donor parents.

Based on the diversity analysed here, it is clear that further collections of wild *Crocus* species are required from across their ranges. Given the difficulty of maintaining the material in Gene banks or gardens, collection for DNA and attempts at propagation should be linked to *in situ* conservation of the areas where *Crocus* is wild.

Although the current analysis is perhaps one of the most extensive on *C. sativus* there are still options for further work. Other types of markers for *in situ* hybridization using natural diploid species collected from the wild as a source of DNA and chromosomes is extremely important. Further, the integration of several molecular approaches has been important and in the present study the conclusions are not derived from only one type of marker, rather IRAP, ESTs, SNPs and barcoding genes were analysed, and the conclusions deriving from the PCR and sequencing was confirmed by cytological observation. Given the complexity of genome, even deeper sampling of sequence would not have been certain to increase the strength of the sequence results alone, for example.

In the future, with the decreasing costs of whole genome and RNA sequencing, it might be hoped that extensive genome sequencing and more transcriptome sequence will enable better characterization of the relationships between the different species in *Crocus*. In particular, given the results with *in situ* hybridization present in (Chapter V), it is likely to be important to elucidate the nature and extent of whole genome duplication events that have happened during *Crocus* evolution in relatively recent periods. It is also to be hoped that the nature of the chromosome fission, fusion and duplication events can be better characterized, as done, for example, in Brassica by Cheng *et al.* (2013).

Crocus sativus origin data can be used for resynthesis of the species with potentially better characteristics. Also there is the potential for transformation with specific genes: if a flowering modification consisting of homoeotic transformation of anthers into stigmas could be achieved, production of saffron would double, and quality would increase due to lack of

pollen contamination. Thus a better understanding of the genetics of *C. sativus* and the complex genome structure along the phylogenetic relationships of the hybrid and wild types must be gained. This will allow dissecting the loci that influence tolerance, high yield and domestication-like traits in *C. sativus*.

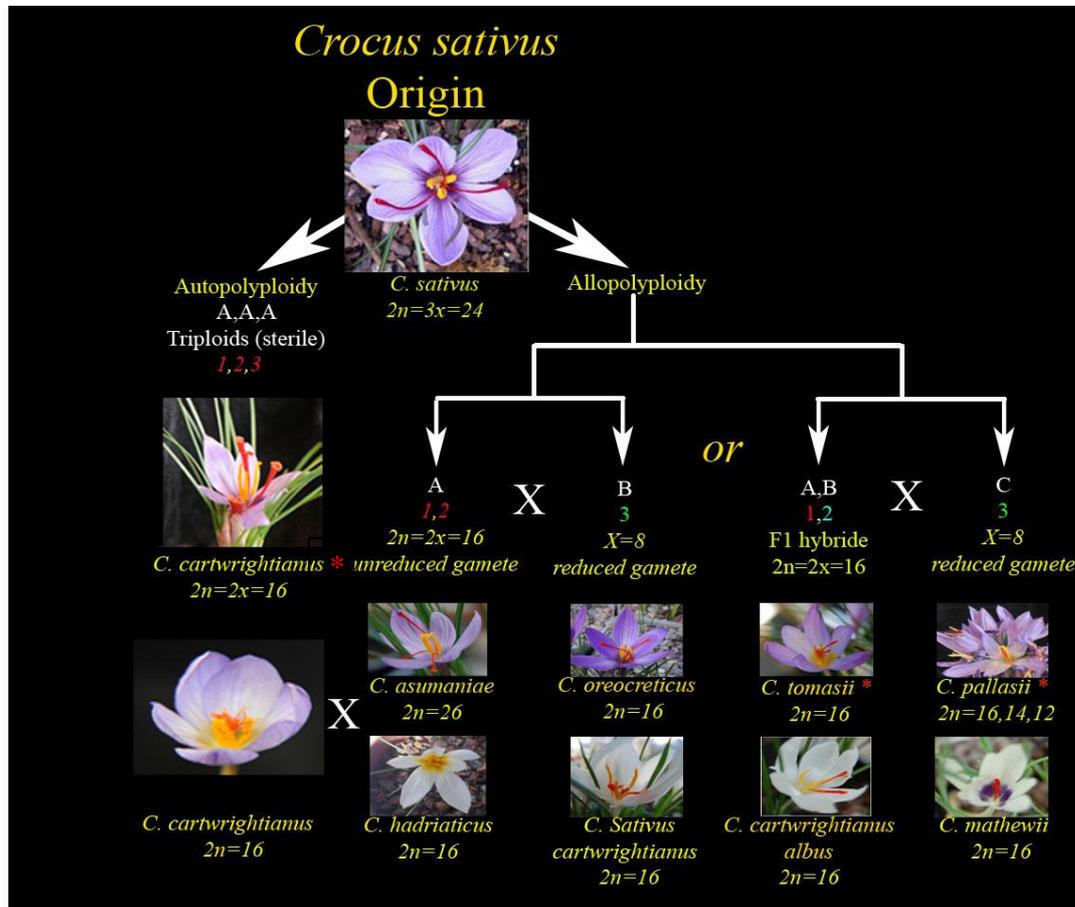


Figure 6.1: Floral morphology and chromosome numbers shows the potential hybrid nature of *C. sativus* (autotriploid or the allotriploid origin) and genetic relationship of saffron with other diploid *Crocus* species series *Crocus*. On the top *C. sativus* ($2n=3x=24$, $x=8$) and asterisk indicates to the flowers of the most potential three ancestor species closer to *C. sativus*. 1) *C. cartwrightianus*, accession (CcrCR09), showing maximum morphological and genetically similarity to saffron suggesting the maternal ancestor (see result in chapter IV, phylogeny tree of barcoding genes), 2) *C. pallasii* subsp. *pallasii* being the best candidate ancestors from nuclear DNA markers results (see Chapters III & IV, phylogeny tree of IRAPs, SNP and an EST-SSR markers), 3) *C. thomasii*, (there is a cytological evidence presented in Chapter V) all Chapters are respectively support this conclusion.

7 REFERENCES

- 1000 Genomes Project Consortium (2010). A map of human genome variation from population-scale sequencing. *Nature*, 467(7319), 1061-1073.
- Abdullaev FI, Espinosa-Aguirre JJ (2004). Biomedical properties of saffron and its potential use in cancer therapy and chemoprevention trials. *Cancer Detection and Prevention*, 28: 426-32.
- Adams MD, Kelley JM, Gocayne JD, Dubnick M, Polymeropoulos MH, Xiao H, Merril CR, Wu A, Olde B, Moreno RF, Kerlavage AR, McCombie WR, Venter JC (1991). Complementary DNA sequencing: Expressed sequence tags and human genome project. *Science*, 252: 1651-1656.
- Agayev YM (2002). New features in karyotype structure and origin of saffron, *Crocus sativus* L. *Cytologia*, 67: 245-252.
- Agayev YM, Fernandez JA, Zarifi E (2009). Clonal selection of saffron (*Crocus sativus* L.): The first optimistic experimental results. *Euphytica*, 169: 81-99.
- Agayev YM, Shakib AM, Soheilvand S, Fathi M (2006, October). Breeding of saffron (*Crocus sativus*): possibilities and problems. In *II International Symposium on Saffron Biology and Technology*, 739 (pp. 203-207).
- Agayew YM, Zarifi E, Fernández J-A (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.
- Akhunov ED, Akhunova AR, Anderson OD, Anderson JA, Blake N, Clegg MT, Coleman-Derr D, Conley EJ, Crossman CC, Deal KR, Dubcovsky J, Gill BS, Gu YQ, Hadam J, Heo H, Huo N, Lazo GR, Luo M, Ma YQ, Matthews DE, McGuire PE, Morrell PL, Qualset CO, Renfro J, Tabanao D, Talbert LE, Tian C, Toleno DM, Warburton ML, You FM, Zhang W, Dvorak J (2010). Nucleotide diversity maps reveal variation in diversity among wheat genomes and chromosomes. *Bmc Genomics*, 11.
- 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 & Biotechnological Equipment*, 22: 795-800.
- Al-Eisawi D (2001). Two new species of *Iridaceae*, *Crocus naqabensis* and *Romulea petraea* from Jordan. *Arab Gulf Journal of Scientific Research*, 19:167-169
- Ali SI, Mathew B (2000). *Iridaceae*. In: Ali SI & Qaiser M (eds.), *Flora of Pakistan*, St. Louis: Press of University of Karachi and Missouri Botanical Garden, 202: 1-35.
- Alix K, Joets J, Ryder CD, Moore J, Barker GC, Bailey JP, King GJ (2008). The CACTA transposon Bot1 played a major role in Brassica genome divergence and gene proliferation. *The Plant Journal*, 56: 1030-1044.

- Álvarez-Ortí M, Gómez LG, Rubio A, Escribano J, Pardo J, Jiménez F, Fernández JA (2004). Development and gene expression in saffron corms. *Acta Horticulturae*, (ISHS) 650:141-153.
- Anastasaki E, Kanakis C, Pappas C, Maggi L, del Campo CP, Carmona M, Polissiou MG (2009). Geographical differentiation of saffron by GC–MS/FID and chemometrics. *European Food Research and Technology*, 229: 899-905.
- Andrabi SMH, Maxwell WMC (2007). A review on reproductive biotechnologies for conservation of endangered mammalian species. *Animal Reproduction Science*, 99: 223-243.
- Arrigo N, Barker MS (2012). Rarely successful polyploids and their legacy in plant genomes. *Current Opinion in Plant Biology* 15:140-146.
- Ayad WG, Hodgkin T, Jaradat A and Rao VR (1997). Molecular Genetic Techniques for Plant Genetic Resources. Report of an IPGRI Workshop, 9–11 October 1995, Rome, Italy. Rome: International Plant Genetic Resources Institute.
- Aytekin A, Acikgoz, AO (2008). Hormone and microorganism treatments in the cultivation of saffron (*Crocus sativus* L.) plants. *Molecules*, 13: 1135-1147.
- Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, et al. (2008) Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *PLoS ONE* 3: e3376.
- Bardil A, de Almeida JD, Combes MC, Lashermes P, Bertrand B (2011). Genomic expression dominance in the natural allopolyploid *Coffea arabica* is massively affected by growth temperature. *New Phytologist*, 192: 760–774.
- Barrett SC, Harder LD (1996). Ecology and evolution of plant mating. *Trends in Ecology and Evolution*, 11(2), 73-79.
- Basirnia ASHKAN, Darvishzadeh R, Mandoulakani BA (2014). Retrotransposon insertional polymorphism in sunflower (*Helianthus annuus* L.) lines revealed by IRAP and REMAP markers. *Plant Biosystems-An International Journal Dealing with all Aspects of Plant Biology*, (2014): 1-12.
- Bedbrook JR, Jones J, O'Dell M, Thompson RD, Flavell RB (1980). A molecular description of telomeric heterochromatin in *Secale* species. *Cell*, 19: 545-60.
- Begum R, Zakrzewski F, Menzel G, Weber B, Alam SS, Schmidt, T. (2013). Comparative molecular cytogenetic analyses of a major tandemly repeated DNA family and retrotransposon sequences in cultivated jute *Corchorus* species (*Malvaceae*). *Annals of Botany*, 112: 123-134.
- Bennett MD (1984). The genome, the natural karyotype and biosystematics. In: Grant WF. Ed. *Plant biosystematics*. Toronto, Canada: Academic Press, pp 41–66.
- Bennett MD, Leitch IJ (2011). Nuclear DNA amounts in angiosperms: Targets, trends and tomorrow. *Annals of Botany*, 107: 467-590.
- Biswas MK, Xu Q, Deng XX (2010). Utility of RAPD, ISSR, IRAP and REMAP markers for the genetic analysis of *Citrus* spp. *Scientia Horticulturae*, 124: 254-261.
- Boff T, Schifino-Wittmann MT (2003). Paleopolyploidy and segmental allopolyploidy in species of *Leucaena* Benth: evidence from meiotic behaviour analysis. *Hereditas*, 138: 27–35.

- Borlaug NE (1983). Contributions of conventional plant breeding to food production. *Science*, 219: 689-693.
- Börner A, Neumann K, Kobiljski B (2011, December). Wheat Genetic Resources—How to Exploit. In *Czech Journal of Genetics and Plant Breeding* (Vol. 47, No. Special Issue, pp. S43-S48). Institute of Agricultural Economics and Information for the Czech Academy of Agricultural Sciences.
- Brandizzi F, Caiola MG (1998). Flow cytometric analysis of nuclear DNA in *Crocus sativus* and allies (*Iridaceae*). *Plant systematics and evolution*, 211: 149-154.
- Bremer B, Bremer K, Chase M, Fay M, Reveal J, Soltis D, Stevens P. (2009). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society*, 161: 105-121.
- Brighton CA (1977). Cytology of *Crocus sativus* and its allies (*Iridaceae*). *Plant Systematics and Evolution*, 128: 137-157.
- Brighton CA, Mathew B, Marchant CJ (1973). Chromosome counts in the genus *Crocus* (*Iridaceae*). *Kew Bulliten*, 28: 451–464.
- Buggs RJA (2013). The consequences of polyploidy and hybridisation for transcriptome dynamics Unravelling gene expression of complex crop genomes. *Heredity*, 110: 97-98.
- Burgess DJ (2013). Chromatin: Retrotransposons and heterochromatin spreading. *Nature Reviews Genetics*, 14: 79-79.
- Burt RL (1948). *Crocus oreocreticus*. *Phyton* 1: 224-225
- Cardle L, Ramsay L, Milbourne D, Macaulay M, Marshall D, Waugh R (2000). Computational and experimental characterization of physically clustered simple sequence repeats in plants. *Genetics*, 156: 847-854.
- Carmona M, Zalacain A, Alonso GL (2006). El color, sabor y aroma del azafrán especia. Albacete : Altabán Ediciones; Spain.
- Carrió E, Güemes J (2013). The role of a mixed mating system in the reproduction of a Mediterranean subshrub (*Fumana hispidula*, *Cistaceae*). *Journal of Plant Research*, 126(1), 33-40.
- Caruso M, Federici CT, Roose ML (2008). EST-SSR markers for asparagus genetic diversity evaluation and cultivar identification. *Molecular Breeding*, 21: 195-204.
- Carvalho A, Guedes- Pinto H, Martins- Lopes P, Lima- Brito J (2010). Genetic variability of Old Portuguese bread wheat cultivars assayed by IRAP and REMAP markers. *Annals of Applied Biology*, 156: 337-345.
- Carvalho A, MartÃn A, Heslop-Harrison JS, Guedes-Pinto H, Lima-Brito J (2009). Identification of the spontaneous 7BS/7RL intergenomic translocation in one F1 multigeneric hybrid from the *Triticeae* tribe. *Plant Breeding*, 128: 105-108.
- Castilho A, Heslop-Harrison JS (1995). Physical mapping of 5S and 18S-25S rDNA and repetitive DNA sequences in *Aegilops umbellulata*. *Genome*, 38: 91-96.
- Chakraborty S, Newton AC (2011). Climate change, plant diseases and food security: An overview. *Plant Pathology*, 60: 2-14.

- Charles M, Belcram H, Just J, Huneau C, Viollet A, Couloux A, Segurens B, Carter M, Huteau V, Coriton O, Appels R, Samain S, Chalhou B (2008). Dynamics and Differential Proliferation of Transposable Elements During the Evolution of the B and A Genomes of Wheat. *Genetics*, 180: 1071-1086.
- Charlesworth D (2006). Evolution of plant breeding systems. *Current Biology*, 16(17), R726-R735.
- Cheng F, Mandáková T, Wu J, Xie Q, Lysak MA, Wang X (2013). Deciphering the diploid ancestral genome of the mesohexaploid *Brassica rapa*. *The Plant Cell Online*, 25: 1541-1554.
- Chichiricco G, Grilli Caiola M (1986). *Crocus sativus* pollen germination and pollen tube growth in vitro and after intraspecific and interspecific pollination. *Canadian Journal of Botany*, 64(11), 2774-2777.
- Ciarmiello LF, Piccirillo P, Pontecorvo G, De Luca A, Kafantaris I, Woodrow P (2011). A PCR based SNPs marker for specific characterization of English walnut (*Juglans regia* L.) cultivars. *Molecular Biology Reports*, 38: 1237-1249.
- Contento A, Heslop-Harrison JS, Schwarzacher T (2005). Diversity of a major repetitive DNA sequence in diploid and polyploid *Triticeae*. *Cytogenetic and Genome Research*, 109: 34-42.
- Coşkun F, Selvi S, Satil F (2010). Phylogenetic relationships of some Turkish *Crocus* (*Iridaceae*) taxa based on morphological and anatomical characters. *Turkish Journal of Botany*, 34: 171-178.
- Cuadrado A, Jouve N (1995). Fluorescent in situ hybridization and C-banding analyses of highly repetitive DNA sequences in the heterochromatin of rye (*Secale montanum* Guss.) and wheat incorporating *S. montanum* chromosome segments. *Genome*, 38: 795-802.
- Cui H, Moe KT, Chung JW, Cho YI, Lee Gapark YJ (2010). Genetic diversity and population structure of rice accessions from South Asia using SSR markers. *Korean Journal of Breeding Science*, 42: 11-22.
- D'Agostino N, Pizzichini D, Chiusano ML, Giuliano G (2007). An EST database from saffron stigmas. *BMC Plant Biology*, 7: 53
- D'Hont A *et al.*, (2012). The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature*, 488: 213–217.
- D'Onofrio C, De Lorenzis G, Giordani T, Natali L, Cavallini A, Scalabrelli G (2010). Retrotransposon-based molecular markers for grapevine species and cultivars identification. *Tree Genetics & Genomes*, 6: 451-466.
- de Juan JA, López CH, Muñoz RM, Picornell MR (2009). Yield and yield components of saffron under different cropping systems. *Industrial Crops and Products*, 30: 212-219.
- Deo B (2003). Growing saffron - the world's most expensive spice. *Crop Food Research*, 20: 1-4.

- Derakhshan E., Majidi MM, Sharafi Y, Mirlohi A (2014). Discrimination and genetic diversity of cultivated and wild safflowers (*Carthamus* spp.) using EST-microsatellites markers. *Biochemical Systematics and Ecology*, 54: 130-136.
- Devaux C, Lepers C, Porcher E (2014). Constraints imposed by pollinator behaviour on the ecology and evolution of plant mating systems. *Journal of Evolutionary Biology*, 27(7), 1413-1430.
- Devos KM (2010). Grass genome organization and evolution. *Current Opinion in Plant Biology*, 13: 139-145.
- Devos KM, Gale MD (1992). The use of random amplified polymorphic DNA markers in wheat. *Theoretical and Applied Genetics*. 84: 567-572.
- Dewannieux M, Esnault C, Heidmann T (2003). LINE-mediated retrotransposition of marked Alu sequences. *Nature Genetics*, 35: 41-48.
- Dhar AK, Sapru R, Rekha K (1988). Studies on saffron in Kashmir. Variation in natural population and its cytological behavior. *Crop Improvement*, 15: 48-52.
- Doležel J, Vrána J, Šafář J, Bartoš J, Kubaláková M, Šimková H (2012). Chromosomes in the flow to simplify genome analysis. *Functional & Integrative Genomics*, 12: 397-416.
- Doyle JJ, Doyle JL (1990). Isolation of plant DNA from fresh tissue. *Focus* 12: 13-15.
- Duc G, Bao S, Baum M, Redden B, Sadiki M, Suso MJ, Zong X (2010). Diversity maintenance and use of *Vicia faba* L. genetic resources. *Field Crops Research*, 115: 270-278.
- Ebrahimzadeh H, Saboora A, Noori-Dalooi MR, Ghaffari SM (1998). Chromosomal studies on four Iranian *Crocus* species (Iridaceae). *Iranian Journal of Botany* 7: 180-192
- Eckardt NA (2010). Evolution of domesticated bread wheat. *Plant Cell*, 22: 993.
- Ellis THN, Poyser SJ, Knox MR, Vershinin AV, Ambrose MJ (1998). Tyl–copia class retrotransposon insertion site polymorphism for linkage and diversity analysis in pea. *Molecular Genetics and Genomics*, 260: 9–19.
- Ellstrand NC, Hoffman CA (1990). Hybridization as an avenue of escape for engineered genes. *BioScience* 40: 438–442.
- Emerson KJ, Merz CR, Catchen JM, Hohenlohe PA, Cresko WA, et al. (2010). Resolving postglacial phylogeography using high-throughput sequencing. *Proceedings of the National Academy of Sciences* 107: 16196–16200.
- Engelmann F, Engels JMM (2002). Technologies and strategies for ex situ conservation. In: Managing Plant Genetic Diversity. Ramanatha RV, Brown AHD, Jackson MT, eds., Wallingford, Rome, CAB International, IPGRI, pp.89-104.
- Eric SM, Mohammadin S, Edger PP (2012). Ancient whole genome duplications, novelty and diversification: the WGD Radiation Lag-Time Model. *Current Opinion in Plant Biology*, 15: 147-153.
- 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.
- Erol O, Kucuker O, Uzen E (2008). Corm tunic morphology of Turkish *Crocoideae* (Iridaceae) and their systematic significance. *Nordic Journal of Botany*, 26: 66-73.

- 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.
- European Saffron White Book (2006). Problems and Strategies for improving the quality and strengthen competitiveness. In: López Córcoles, H., Carmona, M., López Fuster, P., Alonso, G., *et al.* (Ed.). INTERREG IIIC. Project co-financed by UE. This book has been translated into Spanish, French, Italian and Greek, p. 213.
- Feinbrun N (1958). Chromosome numbers in *Crocus*. *Genetica* 29: 172-192.
- Feldman M, Levy AA (2005). Allopolyploidy - A shaping force in the evolution of wheat genomes. *Cytogenetic and Genome Research*, 109: 250-258.
- Felsenstein J. (2005). PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.
- Ferello 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.
- Fernandez JA (2004). Biology, biotechnology and biomedicine of saffron. *Recent Research of Development in Plant Science* 2:127–159.
- Fernandez JA (2007). Genetic resources of saffron and allies (*Crocus* spp.). *Acta Horticulturae* 739:167–185.
- Fernández JA, Santana O, Guardiola JL, Molina RV, Heslop-Harrison JS, Borbely G, De-Los-Mozos-Pascual M (2011). The World Saffron and *Crocus* collection: strategies for establishment, management, characterisation and utilisation. *Genetic Resources and Crop Evolution*, 58: 125-137.
- Fernández, JA, Agayev, YM, Zarifi, E. (2009, May). A study of karyotypes in the *Crocus sativus* L. aggregate and origin of cultivated saffron. In *III International Symposium on Saffron: Forthcoming Challenges in Cultivation, Research and Economics* 850: (pp. 47-54).
- Finnegan DJ (1989). Eukaryotic transposable elements and genome evolution. *Trends in Genetics*, 5: 103-107.
- Flavell AJ, Dunbar E, Anderson R, Pearce SR, Hartley R, Kumar A (1992). Tyl-copia group retrotransposons are ubiquitous and heterogeneous in higher plants. *Nucleic Acids Research*, 20: 3639-3644.
- Flavell AJ, Knox MR, Pearce SR, Ellis THN (1998). Retrotransposon-based insertion polymorphisms (RBIP) for high throughput marker analysis. *Plant Journal*, 16: 643-650.
- Fluch S, Hohl K, Stierschneider M, Kopecky D (2010). *Crocus sativus* L. Molecular Evidence on its Clonal Origin. *Acta Horticulturae*, 850: 41-46.
- Fluch S, Hohl K, Stierschneider M, Kopecky D, Kaar B. (2009). *Crocus sativus* L.-molecular evidence on its clonal origin. In *III International Symposium on Saffron: Forthcoming Challenges in Cultivation, Research and Economics*, 850: (pp. 41-46).
- Frascaroli E, Schrag TA, Melchinger AE. (2013). Genetic diversity analysis of elite European maize (*Zea mays* L.) inbred lines using AFLP, SSR, and SNP markers reveals

- ascertainment bias for a subset of SNPs. *Theoretical and Applied Genetics*, 126: 133-141.
- Frello S, Heslop-Harrison JS (2000a). Repetitive DNA sequences in *Crocus vernus* Hill (*Iridaceae*): the genomic organization and distribution of dispersed elements in the genus *Crocus* and its allies. *Genome*, 43: 902-909.
- Frello S, Heslop-Harrison JS (2000b). 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.
- Fu YB (2003). Applications of bulking in molecular characterization of plant germplasm: a critical review. *Plant Genetic Resources*, 1: 161-167.
- Fu YB, Somers DJ (2009). Genome-wide reduction of genetic diversity in wheat breeding. *Crop Science*, 49: 161-168.
- Fuchs J, Demidov D, Houben A, Schubert I (2006). Chromosomal histone modification patterns - from conservation to diversity. *Trends in Plant Science*, 11: 199-208.
- Funk DJ, Omland KE. (2003). Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics*, 34: 397-423.
- Gadaleta A, Giancaspro A, Zacheo S, Nigro D, Giove SL, Colasuonno P, Blanco A (2011). Comparison of genomic and EST-derived SSR markers in phylogenetic analysis of wheat. *Plant Genetic Resources*, 9: 243-246
- Gaeta RT, Pires JC, Iniguez-Luy F, Leon E, Osborn TC (2007). Genomic changes in resynthesized *Brassica napus* and their effect on gene expression and phenotype. *Plant Cell*, 19: 3403-3417.
- Gall JG, Pardue ML (1969). Formation and detection of RNA-DNA hybrid molecules in cytological preparations. *Proceedings of the National Academy of Sciences of the United States of America*, 69: 378-383.
- Gao LF, Jing RL, Huo NX, Li Y, Li XP, Zhou RH, Chang XP, Tang JF, Ma ZY, Jia JZ (2004). One hundred and one new microsatellite loci derived from ESTs (EST-SSRs) in bread wheat. *Theoretical and Applied Genetics*, 108: 1392-1400.
- Gaut B, Ross Ibarra J (2008). Selection on major components of angiosperm genomes. *Science*, 320: 484-486.
- Geetha SB, Subbareddy S, Satishchandra P, Moses R, Chandra S (2013). Morphological and molecular screening of French bean (*Phaseolus vulgaris* L.) germplasm using SCAR markers for *Colletotrichum lindemuthianum* (Sacc. and Magn.) Scrib. causing anthracnose resistance. *Archives Of Phytopathology And Plant Protection*, 46: 84-97.
- Gerlach WL, Bedbrook JR. (1979). Cloning and characterization of ribosomal RNA genes from wheat and barley. *Nucleic Acids Research*, 7: 1869-1885.
- Ghaffari SM (1986). Cytogenetic studies of cultivated *Crocus sativus* (*Iridaceae*). *Plant systematics and evolution*, 153: 199-204.

- Ghaffari SM, Bagheri A (2009). Stigma variability in saffron (*Crocus sativus* L.). *African Journal of Biotechnology*, 8: 601-604.
- Ghorbani M (2007). The economics of saffron in Iran. *Acta Horticulturae*, 739:321–331.
- Giaccio M (2004). Crocetin from saffron: an active component of an ancient spice. *Crit Reviews in Food Science and Nutrition*, 44: 155-72.
- Gill BS, Friebe BR (1998). Plant cytogenetics at the dawn of the 21st century. *Current Opinion in Plant Biology*, 1: 109-115
- Gill BS, Friebe BR, White FF (2011). Alien introgressions represent a rich source of genes for crop improvement. *Proceedings of the National Academy of Sciences of the United States of America*, 108: 7657-7658.
- Gismondi A, Fanali F, Labarga JMM, Caiola MG, Canini A (2013). *Crocus sativus* L. genomics and different DNA barcode applications. *Plant Systematics and Evolution*, 299: 1859-1863.
- Goldblatt P, Davies TJ, Manning JC, van der BM, Savolainen V (2006). Phylogeny of *Iridaceae* subfamily *Crocoideae* based on a combined multigene plastid DNA analysis. *Aliso*, 22: 399-411.
- Goldblatt P, Rodriguez A, Powell MP, Davies TJ, Manning JC, Van der Bank M and Savolainen V (2008). *Iridaceae* ‘Out of Australasia’? Phylogeny, biogeography, and divergence time based on plastid DNA sequences. *Systematic Botany*, 33:495-508.
- Goldblatt P, Takei M (1997). Chromosome cytology of *Iridaceae* – patterns of variation, determination of base number, and modes of karyotype change. *Annals of the Missouri Botanic Garden*, 84: 285–304.
- Goodwillie C, Kalisz S, Eckert CG (2005). The evolutionary enigma of mixed mating systems in plants: occurrence, theoretical explanations, and empirical evidence. *Annual Review of Ecology, Evolution and Systematics*, 36:47–79
- Goodwillie C, Partis KL, West JW (2004). Transient self-incompatibility confers delayed selfing in *Leptosiphon jepsonii* (*Polemoniaceae*). *International Journal of Plant Sciences*, 165:387–94
- Gresta F, Avola G, Lombardo GM, Siracusa L, Ruberto G (2009). Analysis of flowering, stigmas yield and qualitative traits of saffron (*Crocus sativus* L.) as affected by environmental conditions. *Scientia Horticulturae*, 119: 320-324.
- Gresta F, Lombardo GM, Siracusa L, Ruberto G (2008). Saffron, an alternative crop for sustainable agricultural systems. A review. *Agronomy for Sustainable Development*, 28: 95-112.
- Grilli Caiola GM, Castagnola M, Chichiriccò G (1985). Ultrastructural study of Saffron pollen. *Plant Biosystem*, 119: 61-66.
- Grilli Caiola GM, Di Somma D, Lauretti P (2001). Comparative study on pollen and pistil of *Crocus sativus* L. (*Iridaceae*) and its allied species. *Annali di Botanica Roma*, 1: 93–103
- Grilli Caiola M, Di Somma D, Lauretti P (2011). Comparative study of pollen and pistil in *Crocus sativus* L. (*Iridaceae*) and allied species. *Annali di Botanica*, 58.

- Grilli Caiola M, Leonardi D, Canini A (2010). Seed structure in *Crocus sativus* L., *C. cartwrightianus* Herb., *C. thomasi* Ten., and *C. hadriaticus* Herb. at SEM. *Plant Systematics and Evolution*, 285(1-2), 111-120.
- Grilli Caiola MG (1995). A study on pollen grains of *Crocus cartwrightianus* (Iridaceae). *Plant Systematics and Evolution*, 198: 155-166.
- Grilli Caiola MG (2005). Embryo origin and development in *Crocus sativus* L. (Iridaceae). *Plant Biosystemistry*, 139: 335–343
- Grilli Caiola MG, Canini A (2010) Looking for saffron's (*Crocus sativus* L.) parents. In: Husaini AM (Ed) Saffron. *Functional Plant Science and Biotechnology*, 4 (Special Issue 2), 1-14
- Grilli Caiola MG, Caputto P, Zanier R (2004). RAPD analysis in *Crocus sativus* L. accessions and related *Crocus* species. *Biologia Plantarum*, 48: 375–380.
- Grilli Candan F, Sik L, Kesercioglu T (2009). Cytotaxonomical studies on some *Crocus* L. taxa in Turkey. *African Journal of Biotechnology*, 8: 4374-4377.
- Grodzicker T, Williams J, Sharp P (1974). Physical mapping of temperature sensitive mutations of adenovirus. *Cold Spring Harbor Symposia on Quantitative Biology* 39:439-446.
- Hao M, Luo J, Yang M, Zhang L, Yan Z, Yuan Z, Zheng Y, Zhang H, Liu D, Gustafson P (2011). Comparison of homoeologous chromosome pairing between hybrids of wheat genotypes Chinese Spring *ph1b* and Kaixian-luohanmai with rye. *Genome*, 54: 959-964.
- Harper AL, Trick M, Higgins J, Fraser F, Clissold L, Wells R *et al.* (2012). Associative transcriptomics of traits in the polyploid crop species *Brassica napus*. *Nature Biotechnology*, 30: 798–802.
- 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.
- Haworth AH (1800). On the cultivation of crocuses with a short account of the different species known at present. *Transactions of the Horticulture Society of London*, 1: 122–139.
- Hebert PDN, Cywinska A, Ball SL, de Waard JR (2003). Biological identifications through DNA barcodes. *Proceedings Royal Society of London: Biological Sciences*, 270:313–321.
- Heitkam T, Schmidt T. (2009). BNR - A line family from *Beta vulgaris*- contains a RRM domain in open reading frame 1 and defines a L1 sub-clade present in diverse plant genomes. *Plant Journal*, 59: 872-882.
- Herbert W (1847). A history of the species of *Crocus*. *Journal of Horticulture Society Lond*, 2: 249–293.
- Heslop-Harrison JP (2010). Genes in evolution: The control of diversity and speciation. *Annals of Botany*, 106: 437-438.
- Heslop-Harrison JS (1991). The molecular cytogenetics of plants. *Journal of Cell Science*, 100: 15-21.

- Heslop-Harrison JS (2000a). RNA, genes, genomes and chromosomes: repetitive DNA sequences in plants. In: Olmo E, Redi CA eds. *Chromosomes Today, Vol 13: 45-56*.
- Heslop-Harrison JS (2000b). Comparative genome organization in plants: From sequence and markers to chromatin and chromosomes. *Plant Cell*, 12: 617-635.
- Heslop-Harrison JS, Schmidt T (2012). Plant Nuclear Genome Composition. eLS, John Wiley & Sons Ltd: Chichester. <http://dx.doi.org/10.1002/9780470015902.a0002014.pub2>
- Heslop-Harrison JS, Schwarzacher T (2011). Organisation of the plant genome in chromosomes. *Plant Journal*, 66:18-33.
- Heslop-Harrison JS, Schwarzacher T (2012). Genetics and genomics of crop domestication: prospects for the 21st century. *Plant Biotechnology and Agriculture*, 3-18.
- Heslop-Harrison JS, Schwarzacher T (2013). Nucleosomes and centromeric DNA packaging. *Proceedings of the National Academy of Sciences of the United States of America*, 110: 19974-19975
- Heslop-Harrison Y, Shivanna KR (1977). The receptive surface of the angiosperm stigma. *Annals of Botany*, 41: 1233–1258.
- Heslop-Harrison, J. S. (2012). Genome evolution: extinction, continuation or explosion?. *Current Opinion in Plant Biology*, 15: 115-121.
- Himmerbaur W (1926). Zur Entwicklungsgeschichte von *Crocus sativus*. *Tschirch- Festschrift, Leipzig*. 335-350.
- Hipp AL, Eaton DA, Cavender-Bares J, Fitzek E, Nipper R, Manos PS (2014). A framework phylogeny of the American oak clade based on sequenced RAD data. *PLoS ONE* 9(4): e93975.
- Hiscock SJ, Allen AM (2008). Diverse cell signalling pathways regulate pollen–stigma interactions: the search for consensus. *New Phytologist*, 179: 286–317.
- Hiscock SJ, McInnis SM (2003). The diversity of self-incompatibility systems in flowering plants. *Plant Biology*, 5: 23–32.
- Hodgkin T and Rao VR (2002). People, plants and DNA: perspectives on the scientific and technical aspects of conserving and using plant genetic resources. In: Engels JMM, Rao VR, Brown AHD and Jackson MT (eds) *Managing Plant Genetic Diversity*. Rome: International Plant Genetic Resources Institute, pp. 469–480.
- Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, et al. (2010). Population Genomics of Parallel Adaptation in Threespine Stickleback using Sequenced RAD Tags. *PLoS Genet* 6: e1000862.
- Hollingsworth PM, Forrest LL, Spouge JL, Hajibabaei M, Ratnasingham S, et al. (2009). A DNA barcode for land plants. *Proceedings of the National Academy of Sciences of the United States of America*, 106:12794–12797.
- Holt WV, Pickard AR (1999). Role of reproductive technologies in genetic resource banks in animal conservation. *Reviews of Reproduction*, 4: 143–150.
- Hosseinzadeh H, Nassiri- Asl M (2012). Avicenna's (Ibn Sina) the Canon of Medicine and Saffron (*Crocus sativus*): A Review. *Phytotherapy Research*, 27: 475-483.

- Hu J, Wang L, Li J (2011). Comparison of genomic SSR and EST-SSR markers for estimating genetic diversity in cucumber. *Biologia Plantarum*, 55:577-580.
- Huang CY, Ayliffe MA, Timmis JN (2003). Direct measurement of the transfer rate of chloroplast DNA into the nucleus. *Nature*, 422: 72-76.
- Hyten DL, Song Q, Zhu Y, Choi IY, Nelson RL, Costa JM, Cregan PB (2006). Impacts of genetic bottlenecks on soybean genome diversity. *Proceedings of the National Academy of Sciences of the United States of America*, 103: 16666-16671.
- Isobe S, Boller B, Klimenko I, Kölliker S, Rana JC, Sharma TR, Shirasawa K, Hirakawa H, Sato S, Tabata S (2013). Genome-wide SNP marker development and QTL identification for genomic selection in red Clover. In *Breeding strategies for sustainable forage and turf grass improvement* (pp. 29-36). Springer Netherlands.
- 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(8): 1941-1944.
- Jacobsen N, Orgaard M (2004). *Crocus cartwrightianus* on the Attica Peninsula *Acta Horticulturae*, 650: 65-69.
- Jeong HJ, Jo YD, Park SW, Kang BC (2010). Identification of Capsicum species using SNP markers based on high resolution melting analysis. *Genome*, 53: 1029-1040.
- Jiang J, Birchler JA, Parrott WA, Dawe RK (2003). A molecular view of plant centromeres. *Trends Plant Science*, 8: 570-575.
- Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, Tomsho LP, Hu Y, Liang H, Soltis PS, Soltis DE, Clifton SW, Schlarbaum SE, Schuster SC, Ma H, Leebens-Mack J, Depamphilis CW (2011). Ancestral polyploidy in seed plants and angiosperms. *Nature*, 473: 97-100.
- John HA, Birnstiel ML, Jones KW (1969). RNA-DNA hybrids at the cytological level. *Nature*, 223: 582-587.
- Kafi M, Koocheki A, Rashed MH, Nassiri M (Eds.). (2006). *Saffron (Crocus sativus) production and processing* (Vol. 8). Enfield, NH: Science Publishers.
- 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.
- Kalendar R, Grob T, Regina M, Suoniemi A, Schulman A (1999). IRAP and REMAP: two retrotransposon-based DNA fingerprinting techniques. *Theoretical and Applied Genetics*, 98: 704-711.
- Kalendar R, Schulman AH (2006). IRAP and REMAP for retrotransposon-based genotyping and fingerprinting. *Nature Protocols* 1(5): 2478-2484.
- Kalendar R, Tanskanen J, Immonen S, Nevo E, Schulman A (2000). Genome evolution of wild barley (*Hordeum spontaneum*) by BARE-1 retrotransposon dynamics in response to sharp microclimatic divergence. *Proceedings of the National Academy of Sciences of the United States of America*, 97: 6603-6607.
- Kapitonov VV, Jurka J (2008). A universal classification of eukaryotic transposable elements implemented in Repbase. *Nature Reviews Genetics*, 9: 411-412.

- Karasawa K (1935). Cytological studies on the cultivated *Crocus*. *Japanese Journal of Botany*, 11: 162-168.
- Karasawa K (1942). 25. List of the Chromosome Numbers in the Genus *Crocus*. *Proceedings of the Imperial Academy*, 18: 117-122.
- Karp A, Kresovich S, Bhat KV, Ayad WG and Hodgkin T (1997). Molecular Tools in Plant Genetic Resources Conservation: A Guide to the Technologies. IPGRI Technical Bulletin No. 2. Rome: International Plant Genetic Resources Institute.
- Kearse M, Moir R ETC, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28: 1647-1649.
- Kerndorff H (1988). Observations on *Crocus (Iridaceae)* in Jordan with special reference to *Crocus moabiticus*. *Herbertia (USA)*, 41: 33-53.
- King J, Armstead I, Harper J, Ramsey L, Snape J, Waugh R, King I (2013). Exploitation of interspecific diversity for monocot crop improvement. *Heredity*, 110: 475-483.
- Korzun V (2002). Use of molecular markers in cereal breeding. *Cellular and Molecular Biology Letters*, 7: 811-820.
- Kovach MJ, Sweeney MT, McCouch SR (2007). New insights into the history of rice domestication. *Trends in Genetics*, 23, 578-587.
- Kress WJ, Erickson DL (2007). A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. *PLoS one*, 2: e508.
- Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH (2005). Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences of the United States of America*, 102: 8369-8374.
- 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.
- Kuhn G, Schwarzacher T, Heslop-Harrison JS (2007). Long-range organisation of two non-homologous satellite DNAs from the *Drosophila buzzatii* cluster: Mutual relative arrangements and molecular drive. *Chromosome Research*, 15: 35-36.
- Kumar A, Bennetzen JL (1999). Plant retrotransposons. *Annual Review Genetics*, 33: 479-532.
- Kumar R, Singh V, Devi K, Sharma M, Singh MK, Ahuja PS (2008). State of art of saffron (*Crocus sativus* L.) agronomy: a comprehensive review. *Food Reviews International*, 25: 44-85.
- Lander ES (1996). The new genomics: global views of biology. *Science*, 274:536-539.
- Lapitan VC, Brar DS, Abe T, Redoña ED (2007). Assessment of genetic diversity of Philippine rice cultivars carrying good quality traits using SSR markers. *Breeding science*, 57: 263-270.
- 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(1), 425-437.

- Leflon, M., L. Grandont, F. Eber, V. Huteau, O. Coriton, L. Chelysheva, E. Jenczewski and A. M. Chevre (2010). Crossovers get a boost in *Brassica* allotriploid and allotetraploid hybrids. *Plant Cell*, 22: 2253-2264.
- Leigh F, Kalendar R, Lea V, Lee D, Donini P, Schulman AH (2003). Comparison of the utility of barley retrotransposon families for genetic analysis by molecular marker techniques. *Molecular Genetics and Genomics*, 269: 464-474.
- Leitch IJ, Bennet MD (1997). Polyploidy in angiosperms. *Trends in Plant Science*, 12: 470-476.
- Li DZ, Lian MG, Hong TL, Hong W, Xue JG, Jian QL, Zhi DC (2011). Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. *Proceedings of the National Academy of Sciences of the United State of America*, 108: 19641-19646.
- Li HJ, Conner RL, Chen Q, Graf RJ, Laroche A, Ahmad F, Kuzyk AD (2004). Promising genetic resources for resistance to *wheat streak mosaic virus* and the wheat curl mite in wheat-*Thinopyrum* partial amphiploids and their derivatives. *Genetic Resources and Crop Evolution*, 51: 827-835.
- Linnaeus C (1753). *Species plantarum*, 1957 facsimile. *Bernard Quaritch, London*.
- Liu H, Bayer M, Druka A, Russell JR, Hackett CA, Poland J, Waugh R (2014). An evaluation of genotyping by sequencing (GBS) to map the Breviaristatum-e (ari-e) locus in cultivated barley. *BMC genomics*, 15: 104.
- Liu K, Muse SV (2005). PowerMarker: integrated analysis environment for genetic marker data. *Bioinformatics*, 21: 2128–2129.
- Lovo J, Winkworth RC, Mello-Silva R (2012). New insights into *Trimezieae* (*Iridaceae*) phylogeny: what do molecular data tell us?. *Annals of Botany*, 110: 689-702.
- Lu JJ, Kang JY, Feng SG, Zhao HY, Liu JJ, Wang HZ (2013). Transferability of SSR markers derived from *Dendrobium nobile* expressed sequence tags (ESTs) and their utilization in *Dendrobium* phylogeny analysis. *Scientia Horticulturae*, 158: 8-15.
- Ma J, Wing RA, Bennetzen JL, Jackson SA (2007). Plant centromere organization: a dynamic structure with conserved functions. *Trends in Genetics*, 23: 134-139.
- Ma XQ, Zhu DY, Li CSP, Dong TTK, Tsim KWK (2001). Authentic identification of stigma croci (stigmas of *Crocus sativus*) from its adulterants by molecular genetics analysis. *Planta Medica*, 67: 183–186.
- Mable BK (2013). Polyploids and hybrids in changing environments: winners or losers in the struggle for adaptation? *Heredity*, 110(2), 95-96.
- Macchia M, Ceccarini L, Molfetta I, Cioni PL, Flamini G (2013). Studies on saffron (*Crocus sativus* L.) from Tuscan Maremma (Italy): effects of geographical origin, cultivation environment and drying method on volatile emission. *International Journal of Food Science & Technology*, 48: 2370-2375.
- Madlung A (2013). Polyploidy and its effect on evolutionary success: old questions revisited with new tools. *Heredity*, 110: 99-104.
- 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.
- Makri OE, Ferlemi AV, Lamari FN, Georgakopoulos CD (2013). Saffron administration prevents selenite-induced cataractogenesis. *Molecular Vision*, 19: 1188-1197.
- Mandoulakani BA, Piri Y, Darvishzadeh R, Bernoosi I, Jafari M (2012). Retroelement insertional polymorphism and genetic diversity in *Medicago sativa* populations revealed by IRAP and REMAP markers. *Plant Molecular Biology Reporter*, 30: 286-296.
- Manninen O, Kalendar R, Robinson J, Schulman AH (2000). Application of BAR/M retrotransposon markers to the mapping of a major resistance gene for net blotch in barley. *Molecular and General Genetics*, 264: 325-334
- Markova M, Vyskot B (2010). New horizons of Genomic *in situ* hybridization. *Cytogenetic and Genome Research*, 126: 368-375.
- Mather K (1932). Chromosome variation in *Crocus*. I. *Journal of Genetics*, 26: 129–142.
- Mathew B (1977). *Crocus sativus* and its allies (*Iridaceae*). *Plant Systematics and Evolution*, 128: 89-103.
- Mathew B (1982). The Crocus. A Revision of the Genus *Crocus* (*Iridaceae*). Timber Press, Portland.
- Mathew B (1984). *Crocus* L. In: Davis PH [ed.], Flora of Turkey and the East Aegean Islands, vol 8: pp 413–438. Edinburgh University Press, Edinburgh.
- Mathew B (1999). Botany, taxonomy and cytology of *C. sativus* L. and its allies. *Saffron: Crocus sativus L. Amsterdam, Netherlands: Harwood Academic*, 19-30.
- Mathew B (2000). *Crocus* L. In: Guner A *et al.* (eds) Flora of Turkey and the East Aegean Islands, vol 11. Edinburgh University Press, Edinburgh, pp. 271–274.
- Matsuoka Y (2011). Evolution of polyploid triticum wheats under cultivation: The role of domestication, natural hybridization and allopolyploid speciation in their diversification. *Plant and Cell Physiology*, 52: 750-764.
- Maw G (1886). A Monograph of the Genus *Crocus*, Dulau & Co., London, pp. 238.
- Maxted N, Scholten M, Codd R, Ford-Lloyd B (2007). Creation and use of national inventory of crop wild relatives. *Biological Conservation*, 140: 142–159.
- McClintock B (1983). The significance of responses of the genome to challenge. *Physiology Or Medicine Literature Peace Economic Sciences*, 180.
- McCouch S, Baute GJ, Bradeen J, Bramel P, Bretting PK, Buckler E, Zamir D (2013). Agriculture: Feeding the future. *Nature*: 499: 23-24.
- McGregor CE, van Treuren R, Hoekstra R and van Hintum TJJ (2002). Analysis of the wild potato germplasm of the series *Acaulia* with AFLPs: implications for *ex situ* conservation. *Theoretical and Applied Genetics*, 104: 146–156.
- Molina RV, Valero M, Navarro Y, García-Luis A, Guardiola JL (2004). The effect of time of corm lifting and duration of incubation at inductive temperature on flowering in the saffron plant *Crocus sativus* L.). *Scientia Horticulturae*, 103: 79-91.

- Molina RV, Valero M, Navarro Y, Guardiola JL, García-Luis A (2005). Temperature effects on flower formation in saffron (*Crocus sativus* L.). *Scientia Horticulturae*, 103: 361-379.
- Mollafilabi A, Aslami MH (2010). Replacement of Saffron (*Crocus sativus* L.) with Poppy (*Papaver somniferum* L.) and Its Socio-Economic Results in Afghanistan. *Acta Horticulturae* 850: 299-302.
- Morgante M, Hanafey M, Powell W (2002). Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nature Genetics*, 30: 194-200.
- Murat F, Xu JH, Tannier E, Abrouk M, Guilhot N, Pont C, Messing J, Salse J (2010). Ancestral grass karyotype reconstruction unravels new mechanisms of genome shuffling as a source of plant evolution. *Genome Research*, 20: 1545-1557.
- Nair AS, Teo CH, Schwarzacher T, Heslop-Harrison JS (2005). Genome classification of banana cultivars from South India using IRAP markers. *Euphytica*, 144: 85-290.
- Namayandeh A, Nemati Z, Kamelmanesh MM, Mokhtari M, Mardi M (2013). Genetic relationships among species of Iranian *crocus* (*Crocus* spp.). *Crop Breeding Journal*, 3: 61-67.
- Negbi M (1999). Saffron cultivation: past, present and future prospects. In: Negbi M (ed.): Saffron *Crocus sativus* L. *Harwood Academy Publ. Amsterdam*. pp. 1-19.
- Nehvi FA, Wani SA, Dar SA, Makhdoomi MI, Allie BA, Mir ZA (2007). Biological interventions for enhancing saffron productivity in Kashmir. *Acta Horticulturae* (ISHS), 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.
- Newmaster SG, Fazekas AJ, Ragupathy S (2006). DNA barcoding in land plants: evaluation of *rbcL* in a multigene tiered approach. *Botany*, 84: 335-341.
- Ohshima, K. (2013). RNA-mediated gene duplication and retroposons: retrogenes, LINEs, SINEs, and sequence specificity. *International Journal of Evolutionary Biology*, 2013.
- Ollerton J, Winfree R, Tarrant S (2011). How many flowering plants are pollinated by animals? *Oikos*, 120: 321–326.
- Ordoudi SA, Befani CD, Nenadis N, Koliakos GG, Tsimidou MZ (2009). Further examination of antiradical properties of *Crocus sativus* stigmas extract rich in crocins. *Journal of Agriculture Food Chemistry*, 57:3080-6.
- Ordoudi SA, Tsimidou MZ (2004). Saffron quality: Effect of agricultural practices, processing and storage. In *Production Practices and Quality Assessment of Food Crops Volume 1*: pp. 209-260. Springer Netherlands.
- Ørgaard, M., N. Jacobsen and J. S. Heslop-Harrison (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.
- Orgel, L. E., & Crick, F. H. (1980). Selfish DNA: the ultimate parasite. *Nature*, 284: 604-607.

- Ortiz R, Simon P, Jansky S, Stelly D (2009). Ploidy manipulation of the gametophyte, endosperm and sporophyte in nature and for crop improvement: a tribute to Professor Stanley J. Peloquin (1921-2008). *Annals of Botany* 104: 795-807.
- Özhatay N (2002). Diversity of bulbous monocots in Turkey with special reference. Chromosome numbers. *Pure Applied Chemistry*, 74: 547–555.
- Ozkan H, Kafkas S, Ozer MS, Brondolini A (2005). Genetic relationships among South-East Turkey wild barley populations and sampling strategies of *Hordeum spontaneum*. *Theoretical and Applied Genetics*, 112: 12-20.
- Page RD (1996). TreeView: an application to display phylogenetic trees on personal computers. *Computational Molecular Biology*, 12: 357-358.
- Patel D, Power JB, Anthony P, Badakshi F, Heslop-Harrison JP, Davey MR (2011). Somatic hybrid plants of *Nicotiana x sanderae* (+) *N. debneyi* with fungal resistance to *Peronospora tabacina*. *Annals of Botany*, 108: 809-819.
- Pathak GN (1940). Studies in cytology of *Crocus*. *Annals of Botany*, 4: 227-256.
- Paunescu A (2006). “Strategii de conservare a biodiversității”. Biodiversitatea de la Concepte Fundamentale la Aplicații Biotehnologice, Ed. Academiei. Romane Publisher, Bucharest, pp.120-126.
- Perrino P, Robertson LD, Solh MB (1991). Maintenance, evaluation and use of faba bean germplasm collections: problems and prospects. *Sé'rie Sé'minaires Options Mediterrane'ennes*, 10: 21-31.
- 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.
- Phillips R, Rix EM (1989). Bulbs. The Pan garden plants series. London: Pan.
- Piqueras A, Han BH, Escribano J, Rubio C, Hellin E, Fernandez JA (1999). Development of cormogenic nodules and microcorms by tissue culture, a new tool for the multiplication and genetic improvement of saffron. *Agronomie* 19: 603-610.
- Pires JC, Hertweck KL (2008). A Renaissance of Cytogenetics: Studies in Polyploidy and Chromosomal Evolution 1. *Annals of the Missouri Botanical Garden*, 95: 275-281.
- Pitman NCA, Jorgensen PM (2002). Estimating the size of the world's threatened flora. *Science*, 298: 989.
- Plessner O, Negbi M, Ziv M, Basker D (1989). Effects of temperature on the flowering of the saffron *Crocus* (*Crocus sativus* L.): induction of hysteranthly. *Isreal Journal of Botany*, 38: 1–7.
- Poma A, Fontecchio G, Carlucci G, Chichiricò G (2012). Anti-inflammatory Properties of Drugs from Saffron *Crocus*. *Anti-Inflammatory and Anti-Allergy Agents Medicinal Chemistry*, 11: 37-51.
- Premkumar K, Abraham SK, Santhiya ST, Ramesh A (2003). Protective effects of saffron (*Crocus sativus* L.) on genotoxin-induced oxidative stress in Swiss albini mice. *Phytotherapy Research*, 17: 614-7.

- Qi L, Echalié B, Friebe B, Gill BS (2003). Molecular characterization of a set of wheat deletion stocks for use in chromosome bin mapping of ESTs. *Functional and Integrative Genomics*, 3: 39-55.
- Raina BL, Agarwal SG, Bhatia AK, Gaur GS. (1996). Changes in Pigments and Volatiles of Saffron (*Crocus sativus* L) During Processing and Storage. *Journal of the Science of Food and Agriculture*, 71: 27-32.
- Ramu P, Billot C, Rami JF, Senthilvel S, Upadhyaya HD, Reddy LA, Hash CT (2013). Assessment of genetic diversity in the sorghum reference set using EST-SSR markers. *Theoretical and Applied Genetics*, 126: 2051-2064.
- Rios JL, Recio MC, Giner RM, Manez S (1996). An update review of saffron and its active constituents. *Phytotherapy Research*, 10: 189-193.
- Rosato M, Galián JA, Rosselló JA (2012). Amplification, contraction and genomic spread of a satellite DNA family (E180) in *Medicago* (*Fabaceae*) and allied genera. *Annals of Botany*, 109: 773-782.
- Ross-Ibarra J, Morrell PL, Gaut BS (2007). Plant domestication, a unique opportunity to identify the genetic basis of adaptation. *Proceedings of the National Academy of Sciences of the United States of America*, 104: 8641-8648.
- 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.
- Rudall P, Mathew B (1990). Leaf anatomy in *Crocus* (*Iridaceae*). *Kew Bulliten* 45(3): 535-544.
- Sabine J (1829). An account of the crocuses cultivated in the Society's Garden. *Transactions of the Horticulture Society of London*, 7(8): 419-432.
- 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.
- Saeidi H, Rahiminejad MR, Vallian S, Heslop-Harrison JS (2006). Biodiversity of diploid D-genome *Aegilops tauschii* Coss. in Iran measured using microsatellites. *Genetic Resources and Crop Evolution* 53: 1477-1484.
- Saitou N, Nei M (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4: 406-425.
- Salina EA, Sergeeva EM, Adonina IG, Shcherban AB, Belcram H, Huneau C, Chalhoub B (2011). The impact of Ty3-gypsy group LTR retrotransposons Fatima on B genome specificity of polyploid wheats. *BMC Plant Biology*, 11: 99.
- Sambrook J, Fritsch EF, Maniatis T (1989). *Molecular cloning* (Vol. 1, No. 7.58). New York: Cold spring harbor laboratory press.
- Sanei M, Rahimyan H, Agayev YM, Soheilvand S (2007). New Cytotype of *Crocus pallasii* subsp. *haussknechtii* from West of Iran. *Acta Horticulturae*, 739: 107-111.
- Santana MF, Batista AD, Ribeiro LE, de Araújo EF, de Queiroz MV (2013). Terminal repeat retrotransposons as DNA markers in fungi. *Journal of Basic Microbiology*, 53: 823-827

- Saxena RB (2010). Botany, Taxonomy and Cytology of *Crocus sativus* series *Pharmacognosy Review AYU*, 31: 374-380.
- Scalone R, Kolf M, Albach DC (2013). Mating system variation in Veronica (Plantaginaceae): inferences from pollen/ovule ratios and other reproductive traits. *Nordic Journal of Botany*, 31(3), 372-384.
- Schmidt M, Betti G, Hensel A (2007). Saffron in phytotherapy: pharmacology and clinical uses. *Wiener Medizinische Wochenschrift*, 157: 315-319.
- Schmidt T (1999). LINEs, SINEs and repetitive DNA: non-LTR retrotransposons in plant genomes. *Plant Molecular Biology*, 40: 903-910.
- Schmidt T, Heslop-Harrison JS (1998). Genomes, genes and junk: The large-scale organization of plant chromosomes. *Trends in Plant Science*, 3: 195-199.
- Schmidt T, Kubis S, Heslop-Harrison JS (1995). Analysis and chromosomal localization of retrotransposons in sugar beet (*Beta vulgaris* L.): LINEs and Ty1-copia-like elements as major components of the genome. *Chromosome Research*, 3: 335-345.
- Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S et al. (2009). The B73 maize genome: complexity, diversity, and dynamics. *Science*, 326: 1112-1115.
- Schneider I, Kerndorff H, Pasche E (2012). Chromosome numbers of Turkish *Crocus* (*Liliiflorae*, *Iridaceae*) and their geographical distribution. *Feddes Repertorium*, 123: 73-79.
- Schwarzacher T (2003a). Meiosis, recombination and chromosomes: a review of gene isolation and fluorescent *in situ* hybridization data in plants. *Journal of Experimental Botany*, 54: 11-23.
- Schwarzacher T (2003b). DNA, chromosomes, and *in situ* hybridization. *Genome*, 46: 953-962.
- Schwarzacher T (2008). Chromosomes, recombination and proteins at meiosis - A tribute to Peter Moens (1931-2008). *Chromosome Research*, 16: 679-682.
- Schwarzacher T, Anamthawatjansson K, Harrison GE, Islam A, Jia JZ, King IP, Leitch AR, Miller TE, Reader SM, Rogers WJ, Shi M, Heslop-Harrison JS (1992). Genomic *in situ* hybridization to identify alien chromosomes and chromosome segments in wheat. *Theoretical and Applied Genetics*, 84: 778-786.
- Schwarzacher T, Heslop-Harrison JS (2000). *Practical in situ hybridization*. BIOS Scientific Publishers Ltd.
- Schweizer, G., Ganal, M., Ninnemann, H., & Hemleben, V. (1988). Species-specific DNA sequences for identification of somatic hybrids between *Lycopersicon esculentum* and *Solanum acaule*. *Theoretical and applied genetics*, 75: 679-684.
- Scotti I, Magni F, Fink R, Powell W, Binelli G, Hedley PE. (2000). Microsatellite repeats are not randomly distributed within Norway spruce (*Picea abies* K.) expressed sequences. *Genome*, 43: 41-46.
- Seberg, O., Peterson, G., 2009. How many loci does it take to DNA barcode a *Crocus*? *PLoS ONE*, 4: e4598.

- Senerchia N, Wicker T, Felber F, Parisod C (2013). Evolutionary Dynamics of Retrotransposons Assessed by High-Throughput Sequencing in Wild Relatives of Wheat. *Genome Biology and Evolution*, 5: 1010-1020.
- Sergeeva EM, Salina EA, Adonina IG, Chalhoub B. (2010). Evolutionary analysis of the CACTA DNA-transposon Caspar across wheat species using sequence comparison and in situ hybridization. *Molecular Genetics and Genomics*, 284:11-23.
- Shapiro, J. A., & Sternberg, R. (2005). Why repetitive DNA is essential to genome function. *Biological Reviews*, 80: 227-250.
- Sharaf-Eldin M, Elkholy S, Fernandez JA, Junge H, Cheetham R, Guardiola J, Weathers P (2008). *Bacillus subtilis* FZB24 (R) Affects flower quantity and quality of Saffron (*Crocus sativus*). *Planta Medica*, 74: 1316.
- 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 Biological Science*, 2: 73–78.
- Siracusa L, Gresta F, Avola G, Albertini E, Raggi L, Marconi G, Ruberto G (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.
- Slatkin M (1987). Gene flow and the geographic structure of natural. *Science*, 236: 787-792.
- Soltis DE, Albert VA, Leebens-Mack J, Bell CD, Paterson AH, Zheng C, Sankoff D, DePamphilis CW, Wall PK, Soltis PS. (2009). Polyploidy and angiosperm diversification. *American Journal of Botany*, 96: 336-348.
- Soltis DE, Buggs RJA, Doyle JJ, Soltis PS (2010). What we still don't know about polyploidy. *Taxon* 59: 1387–1403.
- Soltis DE, Soltis PS (1999). Polyploidy: recurrent formation and genome evolution. *Trends in Ecology and Evolution*, 14: 348-352.
- Soltis PS, Soltis DE (2000). The role of genetic and genomic attributes in the success of polyploids. *Proceedings of the National Academy of Sciences of the United States of America*, 97: 7051-7.
- Southern EM (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology*, 98: 503–517.
- Souza-Chies, Tatiana T., Eliane Kaltchuk dos Santos, Lilian Eggers, Alice Mainieri Flores, Eudes M. Stiehl Alves, Juliana Fachineto, Juliana Lustosa *et al.* (2012). Studies on diversity and evolution of Iridaceae species in southern Brazil. *Genetics and Molecular Biology*, 35: 1027-1035.
- Srivastava R, Ahmed H., Dixit RK. (2010). *Crocus sativus* L.: A comprehensive review. *Pharmacognosy reviews*, 4: 200.
- Staub JE, Serquen FC, Gupta M (1996). Genetic markers, map construction, and their application in plant breeding. *Hort Science*, 31: 729-741.
- Stein N (2007). *Triticeae* genomics: Advances in sequence analysis of large genome cereal crops. *Chromosome Research*, 15: 21-31.

- Sugiura T (1931). A list of chromosome number in Angiosperms. *Botanical Magazine Tokyo*, 43, 353.
- Syomin BV, Ilyin YV (2005). Diversity of LTR retrotransposons and their role in genome reorganization. *Russian Journal of Genetics*, 41: 430-435.
- Takeda S, Matsuoka M (2008). Genetic approaches to crop improvement: Responding to environmental and population changes. *Nature Reviews Genetics*, 9: 444-457.
- Talbert LE, Bruckner PL, Smith LY, Sears R, Martin TJ (1996). Development of PCR markers linked to resistance to *Wheat streak mosaic virus* in wheat. *Theoretical and Applied Genetics*, 93: 463-467.
- 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, L'Aquila, Italy, pp 47-96
- Tamura K (1992). Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C-content biases. *Molecular Biology and Evolution*, 9: 678-687.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011). MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, 28: 2731-2739.
- Tang ZX, Fu SL, Ren ZL, Zhang T, Zou YT, Yang ZJ, Li GR, Zhou JP, Zhang HQ, Yan BJ, Zhang HY, Tan FQ, Gustafson P (2011). Diversity and evolution of four dispersed repetitive DNA sequences in the genus *Secale*. *Genome*, 54: 285-300.
- Tautz D (1989). Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research*, 17:6463-6471.
- Teo CH, Tan SH, Ho CL, FaridahQZ, Othman YR, Heslop-Harrison JS, Kalendar R, Schulman AH (2005). Genome constitution and classification using retrotransposon-based markers in the orphan crop banana. *Journal of Plant Biology*, 48: 96-105.
- The International Standardization for Organization 3632-1 and 2. Saffron (*Crocus sativus* L.) ISO, (2003). Part 1: Specifications. Part 2: Test Methods. International Standards Organisation, Geneva, Switzerland, ISO 3632-1/2.
- Todorovska E, Abumhadi N, Kamenarova K, Jeleva D, Kostova A, Christov N, Alexandrova N, Jacquemin JM, Anzai H, Nakamura C, Atanassov A. (2005). Biotechnological approaches for cereal crops improvement part II: Use of molecular markers in cereal breeding. *Biotechnology and Biotechnological Equipment*, 19: 91-104.
- 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.
- Tóth G, Gáspári Z, & Jurka J (2000). Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Research*, 10: 967-981.
- Tsoktouridis G, Krigas N, Karamplianis T, Constantinidis T, Maloupa E (2009). Genetic differences among wild Greek *Crocus* taxa and cultivated saffron (*Crocus sativus* L.). In: 3rd International symposium on saffron Forthcoming challenges in cultivation research and economics. Krokos, Kozani, Greece Book of Abstracts, pp. 37.

- Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Henrissat, B (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science*, 313: 1596-1604.
- Van de Peer Y, Pires JC (2012). Getting up to speed. *Current Opinion in Plant Biology*, 15: 111-114.
- Varshney RK, Dubey A. (2009). Novel genomic tools and modern genetic and breeding approaches for crop improvement. *Journal of Plant Biochemistry and Biotechnology*, 18: 127-138.
- Varshney RK, Graner A, Sorrells ME. (2005). Genic microsatellite markers in plants: features and applications. *Trends in Biotechnology*, 23: 48-55.
- Vaughan DA, Balázs E, Heslop-Harrison JS (2007). From crop domestication to superdomestication. *Annals of Botany*, 100: 893-901.
- Vencovsky R, Crossa J (1999). Variance effective population size under mixed self and random mating with applications to genetic conservation of species. *Crop Science*, 39: 1282–1294.
- Vershinin A, Svitashv S, Gummesson PO, Salomon B, von Bothmer R, Bryngelsson T (1994). Characterization of a family of tandemly repeated DNA sequence in *Triticeae*. *Theoretical and Applied Genetics*, 89: 217-225.
- Vershinin AV, Alkhimova EG, Heslop-Harrison JS (1996). Molecular diversification of tandemly organized DNA sequences and heterochromatic chromosome regions in some *Triticeae* species. *Chromosome Research*, 4: 517-25.
- Vicient CM, Jaaskelainen M, Kalendar R, Schulman AH (2001) Active retrotransposons are a common feature of grass genomes. *Plant Physiology*, 125: 1283-1292
- Vignal A, Milan D, San Cristobal M (2002). A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics Selection Evolution*, 34: 275-305.
- Villalobos VM, Engelmann F (1995). *Ex situ* conservation of plant germplasm using biotechnology. *World Journal of Microbiology and Biotechnology*, 11: 375-382.
- Vos P, Hogers R, Bleeker M, Reijans M, Lee TVD, Hornes M, Friters A, Pot J, Paleman J, Kuiper M, Zabeau M (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, 23: 4407–4414.
- Wang Z, Weber JL, Zhong G, Tanksley SD (1994). Survey of plant short tandem DNA repeats. *Theoretical and Applied Genetics*, 88: 1-6.
- Wessler SR (2006). Transposable elements and the evolution of eukaryotic genomes. *Proceedings of the National Academy of Sciences of the United States of America*, 103: 17600-17601.
- Williams JGK, AR Kubelik, K.J. Livak, J.A. Rafalski, and S.V. Tingey. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research*, 18: 6531–6535.
- Winterhalter P, Straubinger M (2000). Saffron renewed interest in an ancient spice. *Food Reviews International*, 16: 39-59.

- Wollrab C, Heitkam T, Holtgräwe D, Weisshaar B, Minoche AE, Dohm JC, Schmidt T (2012). Evolutionary reshuffling in the Errantivirus lineage Elbe within the *Beta vulgaris* genome. *The Plant Journal*, 72: 636-651.
- Xue S, Zhang Z, Lin F, Kong Z, Cao Y, Li C, Yi H, Mei M, Zhu H, Wu J, Xu H, Zhao D, Tian D, Zhang C, Ma Z (2008). A high-density intervarietal map of the wheat genome enriched with markers derived from expressed sequence tags. *Theoretical and Applied Genetics*, 117: 181-189.
- Yang W, Kang X, Yang Q, Lin Y, Fang M (2013). Review on the development of genotyping methods for assessing farm animal diversity. *Journal of Animal Science and Biotechnology*, 4: 2.
- Yao H, Song J, Liu C, Luo K, Han J (2010). Use of ITS2 region as the universal DNA barcode for plants and animals. *PLoS ONE* 5: e13102.
- Yu M, Chu J, Ma R, Shen Z, Fang J (2013). A novel strategy for the identification of 73 *Prunus domestica* cultivars using random amplified polymorphic DNA (RAPD) markers. *African Journal of Agricultural Research*, 8: 243-250.
- Zanier R, Grilli Caiola M (2011). Self incompatibility mechanisms in the *Crocus sativus* aggregate (*Iridaceae*): a preliminary investigation. *Annali di Botanica*, 58.
- Zhao, XP, Si Y, Hanson RE, Crane CF, Price HJ, Stelly DM, Paterson AH (1998). Dispersed repetitive DNA has spread to new genomes since polyploid formation in cotton. *Genome Research*, 8: 479-492.
- Zhu, Youyong, Hairu Chen, Jinghua Fan, Yunyue Wang, Yan Li, Jianbing Chen, JinXiang Fan et al. (2000). Genetic diversity and disease control in rice. *Nature*, 406: 718-722.
- Zhuang D, Aoki Y, Kageyama K, Fukui H (2013). Development of Species-specific SCAR markers for identification of rose species, *Rosa multiflora*. *Journal of the Japanese society for Horticultural Science*, 82: 78-82.
- 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.

Important websites referred in the thesis

<http://www.crocusbank.org/>

<http://www.saffronomics.org/>

<http://www.saffrongenes.org>

<http://www.fao.org/>

<http://www.ncbi.nlm.nih.gov/dbEST/>

<http://tandem.bu.edu/trf/trf.html>

http://www.frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi

<http://www.sigmaaldrich.com/>

<http://genomesonline.org>

<http://www.iucnredlist.org/>

8 Appendix 4

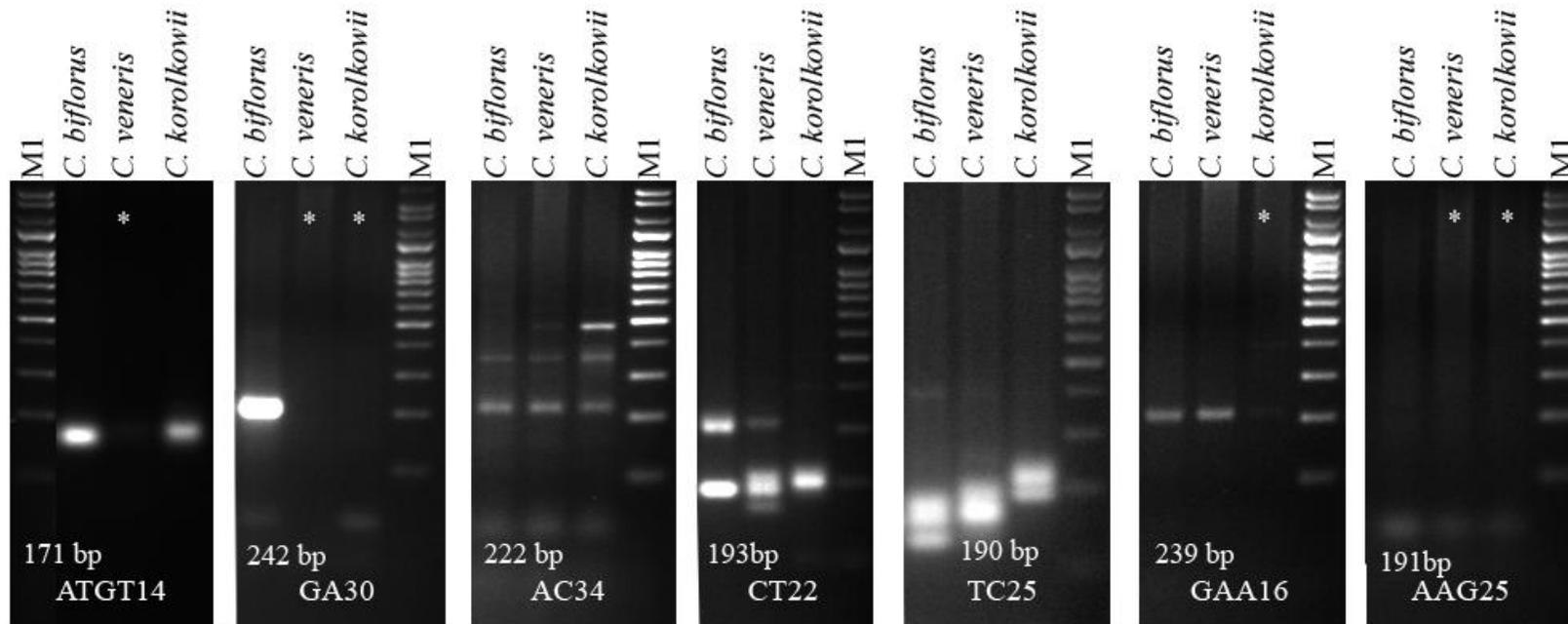


Figure A4.1: PCR amplification pattern of the EST-SSR markers from *C. biflorus*, *C. veneris* and *C. korolkowii*. Markers used from left to right include: ATGT₁₄, GA₃₀, AC₃₄, CT₂₂, TC₂₅, CAA₁₆ and AAG₂₅. Name of species is given on the top of every lane. Markers name along the expected product size is given at the base. On side of the agarose gel (2%) is a DNA length marker Q-Step 2 (M1).

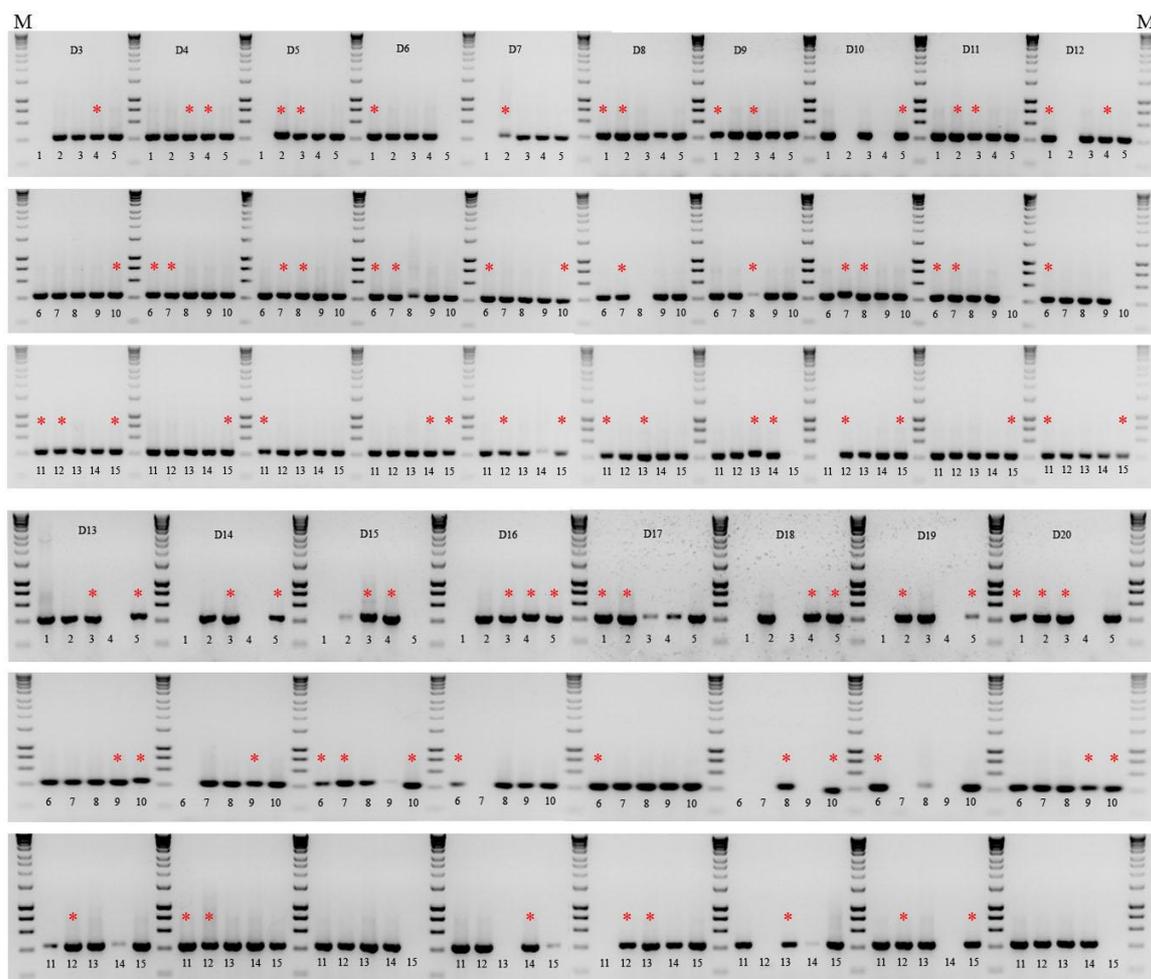


Figure A4.2: Colony PCR for the ATPs gene sequences. For transformation three different concentrations of the bacterial cells were used and placed here one above the other as a column. “M” is DNA size marker HyperLadder I. Asterisks indicating to the sequenced bacterial colony, and values with letters (D3-D20) indicating to the *Crocus* accessions that included: D3 *C. tommasinianus* (CtmLD09), D4 *C. tommasinianus* (CtmBD09), D5 *C. tommasinianus* (CtmTD09), D6 *C. tommasinianus* (CtmAD09), D7 *C. versicolor* (CvrPP09), D8 *C. niveus* (Cniv08), D9 *C. goulimyi* (CgulD08), D10 *C. kotschyanus* (CkotP09), D11 *C. kotschyanus* (Ckot/z08), D12 *C. angustifolius* (CangP09), D13 *C. korolkowii* (Ckor08), D14 *C. flavus* (CflaP09), D15 *C. speciosus* (CspP09), D16 *C. laevigatus* (Clae08), D17 *C. boryi* (Cbor08), D18 *C. cancellatus* (CcanD10), D19 *C. biflorus* (CbfAD10), D20 *C. verneris* (CvenD10). See also Table A4.1 & A4.2.

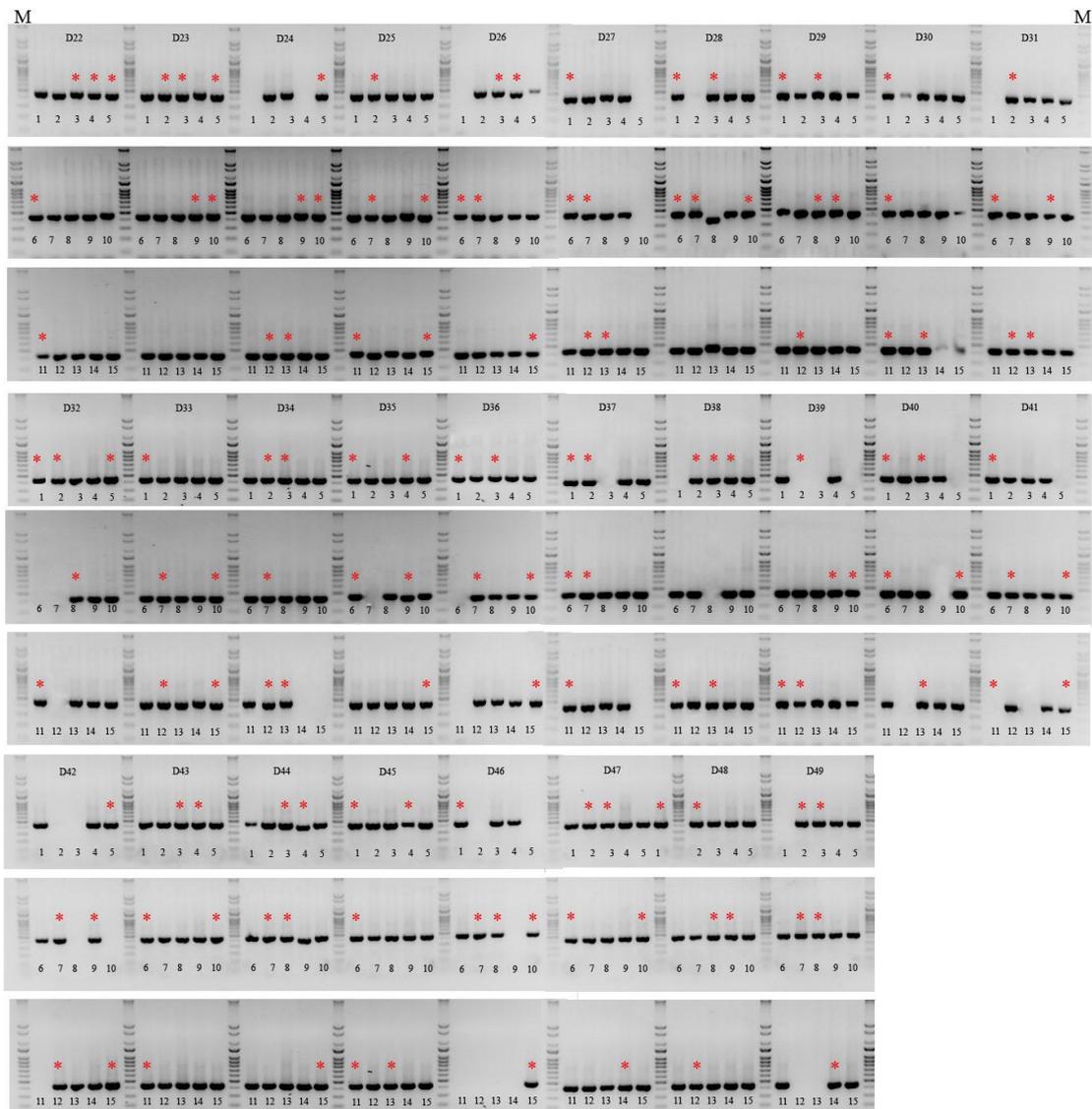


Figure A4.3: Colony PCR for the ATPs gene sequences. For transformation three different concentrations of the bacterial cells was used and placed here one above the other as a column. “M” is DNA size marker Q-step 2. Asterisks indicating to the sequenced bacterial colony, and values with letters (D22-D49) indicating to the *Crocus* accessions that included: D22 *C. sativus*, J.Perez, Spain (CstPER09), D23 *C. sativus*, (Kashmir, Cstkf09), D24 *C. sativus* violet purple, dark striped (CstVD09), D25 *C. sativus* Cashmirianus Hort Lilac (CstCD09), D26 *Crocus sativus* (CsatP09), D27 *Crocus sativus* (Cstst09), D28 *Crocus sativus cartwrightianus* cv. *albus* (CstcP09), D29 *C. pallasii* (CpltR09), D30 *C. pallasii* (CplVD09), D31 *C. pallasii* (CplDD09), D32 *C. mathewii* (CmatD09jean), D33 *C. mathewii* (CmtHR09), D34 *C. thomasii* (CtmVD09), D35 *C. thomasii* (CtomI09john), D36 *C. asumaniae* (CasWD09), D37 *C. asumaniae* (CasAD09), D38 *C. asumaniae* (CasAT09jhon), D39 *C. oreoreticus* (CorVR09), D40 *C. oreoreticus* (CorVD09), D41 *C. hadriaticus* (ChdWD08), D42 *C. hadriaticus* (ChaIR09), D43 *C. hadriaticus* (ChdARD09), D44 *C. cartwrightianus* (CcwBD09), D45 *C. cartwrightianus* (CcrCR09), D46 *C. cartwrightianus* (CcwAD08), D47 *C. pallasii* (CplVD09), D48 *C. asumaniae* (CasWD09), D49 *C. cartwrightianus* (CcwBD10).

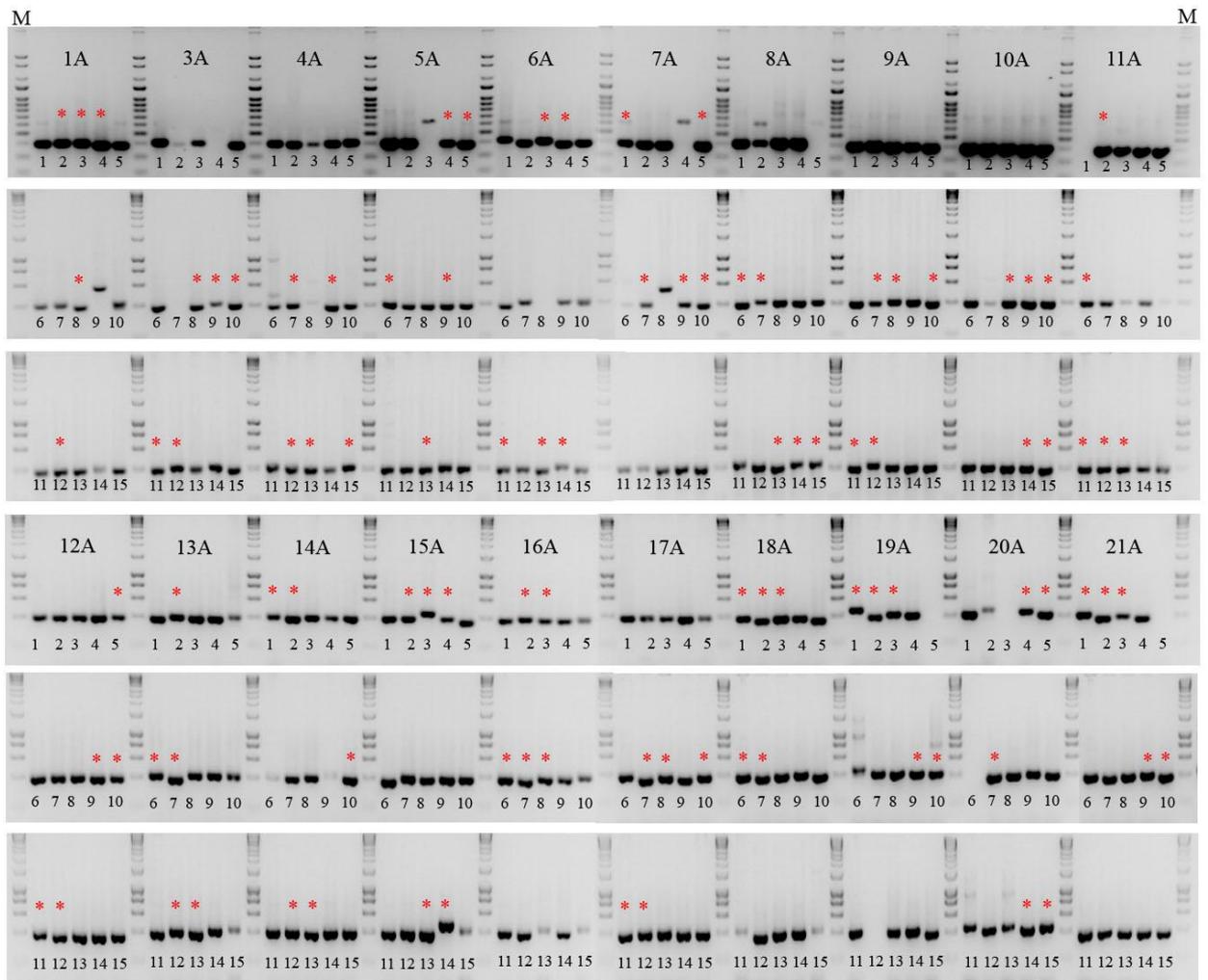


Figure A4.4: Colony PCR for the TC₂₅ sequences. For transformation three different concentrations of the bacterial cells was used and placed here one above the other as a column. “M” is DNA size marker Q-step 2. Asterisks indicating to the sequenced bacterial colony, and values with letters (1A-21A) indicating to the *Crocus* accessions that included: 1A *C. sativus* (J.Perz, Spain, CstPER09), 3A *C. tommasinianus* (CtmLD09), 4A *C. tommasinianus* (CtmBD09), 5A *C. tommasinianus* (CtmTD09), 6A *C. tommasinianus* (CtmAD09), 7A *C. versicolor* (CvrPP09), 8A *C. niveus* (Cniv08), 9A *C. goulimyi* (CgulD08), 10A *C. kotschyanus* (CkotP09), 11A *C. kotschyanus* (Ckot/z08), 12A *C. angustifolius* (CangP09), 13A *C. korolkowii* (Ckor08), 14A *C. flavus* (CflaP09), 15A *C. speciosus* (CspP09), 16A *C. laevigatus* (Clae08), 17A *C. boryi* (Cbor08), 18A *C. cancellatus* (CcanD10), 19A *C. biflorus* (CbfAD10), 21A *C. verneris* creamy (CvenD10).

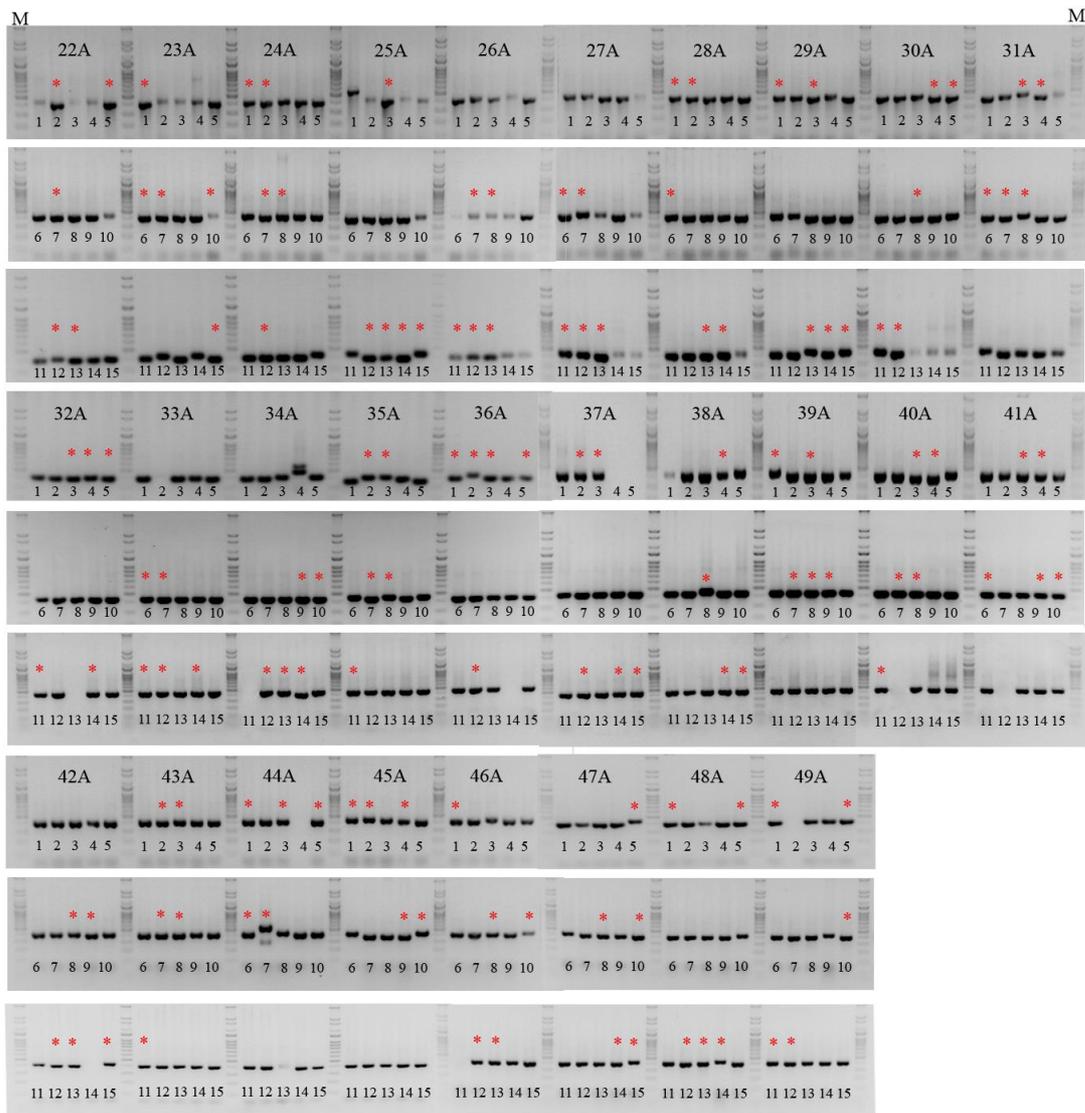


Figure A4.5: Colony PCR for the TC₂₅ sequences. For transformation three different concentrations of the bacterial cells was used and placed here one above the other as a column. “M” is DNA size marker Q-step 2. Asterisks are indicating the sequenced bacterial colony, and values with letters (22A-49A) indicating to the *Crocus* accessions that included: 22A *C. sativus*, J.Perez, Spain (CstPER09), 23A *C. sativus*, (Kashmir, Cstkf09), 24A *C. sativus*, violet purple, dark striped (CstVD09), 25A *C. sativus* Cashmirianus Hort Lilac (CstCD09), 26A *Crocus sativus* (CsatP09), 27A *Crocus sativus* (Cstsut09), 28A *Crocus sativus cartwrightianus* cv. *albus* (CstcP09), 29A *C. pallasii* (CpltR09), 30A *C. pallasii* (CplVD09), 31A *C. pallasii* (CplDD09), 32A *C. mathewii* (CmatD09jean), 33A *C. mathewii* (CmtHR09), 34A *C. thomasii* (CtmVD09), 35A *C. thomasii* (CtomI09john), 36A *C. asumaniae* (CasWD09), 37A *C. asumaniae* (CasAD09), 38A *C. asumaniae* (CasAT09jhon), 39A *C. oreocreticus* (CorVR09), 40A *C. oreocreticus* (CorVD09), 41A *C. hadriaticus* (ChdWD08), 42A *C. hadriaticus* (ChaIR09), 43A *C. hadriaticus* (ChdARD09), 44A *C. cartwrightianus* (CcwBD09), 45A *C. cartwrightianus* (CcrCR09), 46A *C. cartwrightianus* (CcwAD08), 47A *C. pallasii* (CplVD09), 48A *C. asumaniae* (CasWD09), 49A *C. cartwrightianus* (CcwBD10).

Table A4.1: List of *Crocus* species and accessions along information about transformed colonies of the ATPs and TC₂₅ gene sequences. Colony sequence identification code and sequence length in bp.

No	Section	Series	Species	Accession	TC ₂₅		ATPs	
					No of seq.	Seq. Length (bp)	No of seq.	Seq. Length (bp)
1	<i>Crocus</i>	<i>Crocus</i>	<i>C.sativus (J.Perz, Spain)</i>	CstPER09	1 A 2	178	-	-
					1 A 3	Not good	-	-
					1 A 4	164	-	-
					1 A 8	164	-	-
					1 A 12	163	-	-
3	<i>Crocus</i>	Verni	<i>C.tommasinianus</i>	CtmLD09	3 A 8	163	4D3	205
					3 A 9	Not good	15D3	202
					3 A 10	178	12D3	206
					3 A 11	166	11D3	205
					3 A 12	188	10D3	205
4	<i>Crocus</i>	Verni	<i>C.tommasinianus</i>	CtmBD09	4 A 7	Identical to 4 A 12	3D4	202
					4 A 9	162	4D4	202
					4 A 12	175	6D4	205
					4 A 13	166	7D4	205
					4 A 15	Not good	15D4	202
5	<i>Crocus</i>	Verni	<i>C.tommasinianus</i>	CtmTD09	5 A 4	175	2D5	205
					5 A 5	166	3D5	226
					5 A 6	188	7D5	205
					5 A 9	Identical to 5 A 5	8D5	205
					5 A 13	166	11D5	205
6	<i>Crocus</i>	Verni	<i>C.tommasinianus</i>	CtmAD09	6 A 3	Identical to 6 A 14	1D6	205
					6 A 4	Identical to 6 A 11	6D6	205
					6 A 11	166	7D6	209
					6 A 13	156	14D6	205
					6 A 14	183	15D6	205
7	<i>Crocus</i>	Versicolores	<i>C.versicolor</i>	CvrPP09	7 A 1	173	2D7	205
					7 A 5	Identical to 7 A 10	6D7	201
					7 A 7	Identical to 7 A 1	10D7	202
					7 A 9	173	15D7	205
					7 A 10	165	12 D7	205
8	<i>Crocus</i>	Longiflori	<i>C.niveus</i>	Cniv08	8 A 6	206	1D8	205
					8 A 7	190	2D8	205
					8 A 13	174	7D8	205
					8 A 14	201	11D8	205
					8 A 15	Identical to A11	13D8	205
9	<i>Crocus</i>	Longiflori	<i>C.goulimy</i>	CgulD08	9 A 7	Identical to 9 A 11	1D9	205
					9 A 8	186	3D9	205
					9 A 10	171	8D9	205
					9 A 11	171	13D9	226
					9 A 12	192	14D9	205
10	<i>Crocus</i>	Kotschyani	<i>C.kotschyanus</i>	CkotP09	10 A 8	Identical to 10 A 10	5D10	202
					10 A 9	175	7D10	226
					10 A 10	175	8D10	202
					10 A 14	175	12D10	205
					10 A 15	161	15D10	205
11	<i>Crocus</i>	Kotschyani	<i>C.kotschyanus / ZONATUS</i>	Ckot/z08	11 A 2	175	2D11	202
					11 A 6	175	3D11	205
					11 A 11	Identical to 11 A 6	6D11	205
					11 A 12	Identical to 11 A 6	7D11	205
					11 A 13	Identical to 11 A 6	15D11	205
12	Nudiscapus	Reticulati	<i>C.angustifolius</i>	CangP09	12 A 5	Identical to 12 A 12	1D12	205
					12 A 9	169	4D12	205
					12 A 10	Identical to 12a11	6D12	205
					12 A 11	163	11D12	205
					12 A 12	134	15D12	205

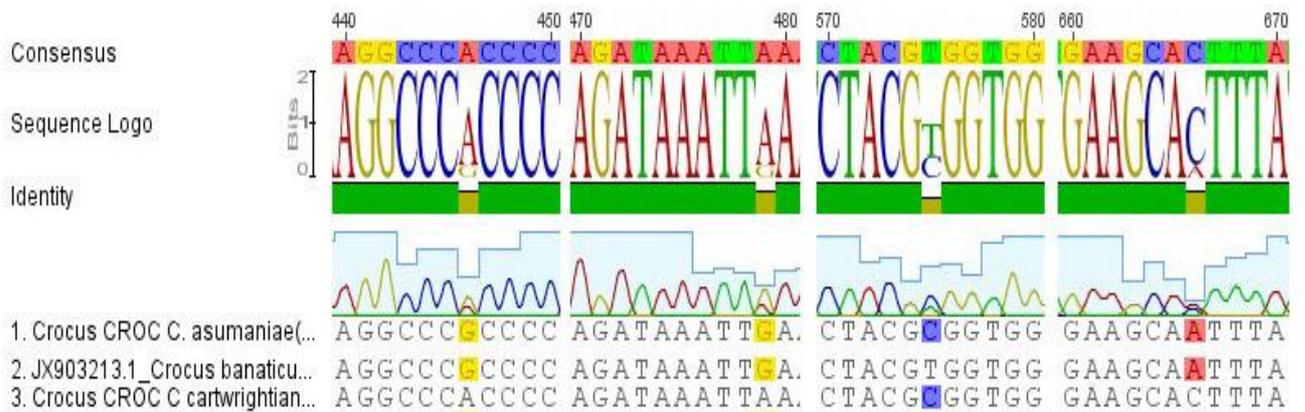
13	Nudiscapus	Orientales	<i>C.korolkowii</i>	Ckor08	13 A 2	172	3D13	202
					13 A 6	190?	5D13	226
					13 A 7	163	9D13	205
					13 A 12	178	10D13	226
					13 A 13	163	12D13	205
14	Nudiscapus	Flavi	<i>C.flavus</i>	CflaP09	14 A 1	178	3D14	205
					14 A 2	170	5D14	205
					14 A 10	170	9D14	205
					14 A 12	176	11D14	205
					14 A 13	156	12D14	202
15	Nudiscapus	Speciosi	<i>C.speciosus</i>	CspP09	15 A 2	172	3D15	226
					15 A 3	221-97%	4D15	202
					15 A 4	165	6D15	205
					15 A 13	167-97%	7D15	226
					15 A 14	264	10D15	205
16	Nudiscapus	Laevigatae	<i>C.laevigatus</i>	Clae08	16 A 2	169	3D16	205
					16 A 3	155	4D16	226
					16 A 6	151	5D16	205
					16 A 7	Not good	6D16	205
					16 A 8	163	14D16	205
17	Nudiscapus	Laevigatae	<i>C.boryi</i>	Cbor08	17 A 7	151	1D17	205
					17 A 8	Identical to 17 A 10	2D17	202
					17 A 10	163	6D17	202
					17 A 11	164	12D17	202
					17 A 12	155	13D17	202
18	Nudiscapus	Reticulati	<i>C.cancellatus ssp</i>	CcanD10	18 A 1	161	4D18	205
					18 A 2	134	5D18	205
					18 A 3	Identical to 18 A 1	8D18	205
					18 A 6	Identical to 18 A 1	10D18	153
					18 A 7	134	13D18	205
19	Nudiscapus	Biflori	<i>C.biflorus</i>	CbfAD10	19 A 1	227	2D19	205
					19 A 2	154	5D19	202
					19 A 3	186	6D19	202
					19 A 9	186	12 D19	205
					19 A 10	167	15D19	202
20	Nudiscapus	Orientales	<i>C.korolkowii</i>	CkrGD10	20 A 4	215	-	-
					20 A 5	Identical to 20 A 14	-	-
					20 A 6	224	-	-
					20 A 14	193	-	-
					20 A 15	217	-	-
21	Nudiscapus	Aleppici	<i>C.verneris creamy</i>	CvenD10	21 A 1	191	1D20	226
					21 A 2	157	2D20	202
					21 A 3	168	3D20	226
					21 A 9	175	9D20	205
					21 A 10	170	10D20	226

Table A4.2: List of *Crocus* accessions from *Crocus* series *Crocus* along with information about transformed colonies of the ATPs and TC₂₅ gene sequences. Colony sequence identification code and sequence length in bp.

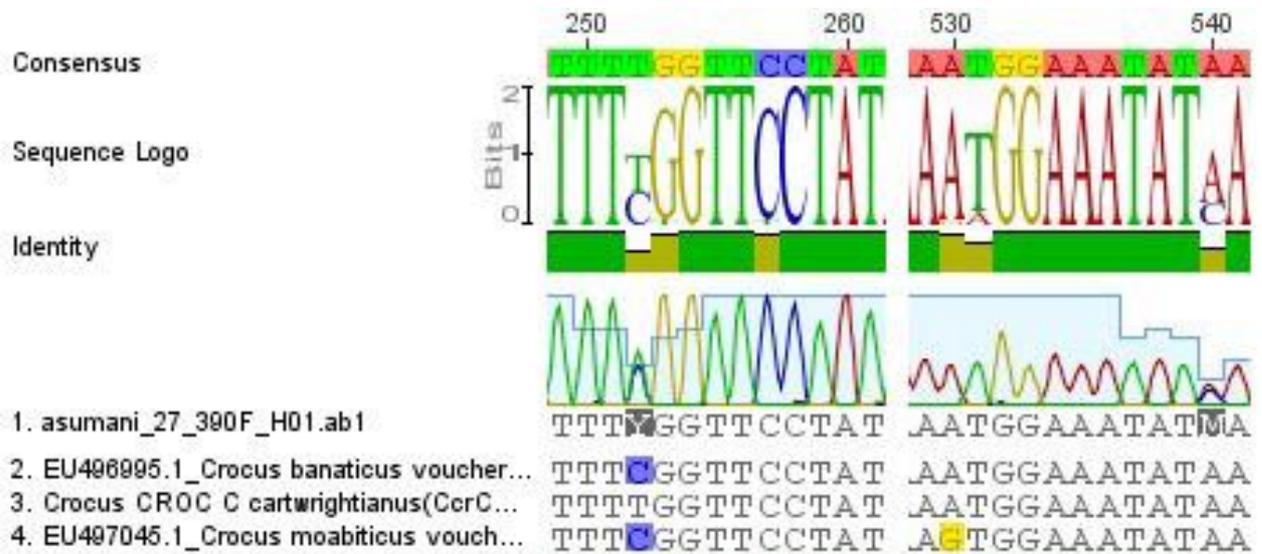
No	Section	Series	Species	Accession	TC ₂₅		ATPs	
					No of seq	Seq length (bp)	No of seq	Seq length (bp)
22	<i>Crocus</i>	<i>Crocus</i>	<i>C.sativus, J.Perez, Spain</i>	CstPER09	22 A 2	Identical to 22 A 12	3D22	202
					22 A 5	166	4D22	226
					22 A 7	163	5D22	202
					22 A 12	162	6D22	205
					22 A 13	Not good	11D22	205
23	<i>Crocus</i>	<i>Crocus</i>	<i>C.sativus, Farah</i>	Cstkf09	23 A 1	163	2D23	205
					23 A 6	Identical to 23 A 10	3D23	226
					23 A 7	Not good	5D23	205
					23 A 10	162	9D23	226
					23 A 15	163	10D23	213
24	<i>Crocus</i>	<i>Crocus</i>	<i>C.sativus, violet purple, dark striped</i>	CstVD09	24 A 1	188	5D24	205
					24 A 2	174	9D24	226
					24 A 7	163	10D24	202
					24 A 8	178	12D24	205
					24 A 12	162	13D24	226
25	<i>Crocus</i>	<i>Crocus</i>	<i>C.sativus Cashmirianus Hort Lilac</i>	CstCD09	25 A 3	162	2D25	226
					25 A 12	163	7D25	205
					25 A 13	196	10D25	205
					25 A 14	182	11D25	202
					25 A 15	174	15D25	202
26	<i>Crocus</i>	<i>Crocus</i>	<i>Crocus sativus</i>	CsatP09	26 A 7	163	3D26	226
					26 A 8	180	4D26	205
					26 A 11	166	6D26	226
					26 A 12	163-52%	7D26	205
					26 A 13	180	15D26	205
27	<i>Crocus</i>	<i>Crocus</i>	<i>Crocus sativus</i>	Cstsut09	27 A 6	188	1D27	205
					27 A 7	162	6D27	205
					27 A 11	198	7D27	205
					27 A 12	186	12D27	202
					27 A 13	166	13D27	226
28	<i>Crocus</i>	<i>Crocus</i>	<i>Crocus sativus cartwrightianus Albus</i>	CstcP09	28 A 1	Identical to 28 A 6	1D28	202
					28 A 2	163	3D28	220
					28 A 6	166	6D28	226
					28 A 13	163	7D28	205
					28 A 14	170	10D28	205
29	<i>Crocus</i>	<i>Crocus</i>	<i>C.pallasii</i>	CpltR09	29 A 1	190	1D29	205
					29 A 3	Identical to 29 A 13	3D29	205
					29 A 13	163	8D29	205
					29 A 14	196	9D29	205
					29 A 15	163	12D29	202
30	<i>Crocus</i>	<i>Crocus</i>	<i>C.pallasii</i>	CplVD09	30 A 4	163	1D30	205
					30 A 5	186	6D30	205
					30 A 8	180	10D30	205
					30 A 11	168	11D30	205
					30 A 12	162	13D30	202
31	<i>Crocus</i>	<i>Crocus</i>	<i>C.pallasii</i>	CplIDD09	31 A 3	Identical to 31 A 4	2D31	205
					31 A 4	194	6D31	205
					31 A 6	151	9D31	205
					31 A 7	168	12D31	209
					31 A 8	196	13D31	202
32	<i>Crocus</i>	<i>Crocus</i>	<i>C.Mathewii</i>	CmatD09jean	32 A 3	164	1D32	201
					32 A 4	184	2D32	205
					32 A 5	164	5D32	205
					32 A 11	188	11D32	205
					32 A 14	Not good	8D32	202

33	<i>Crocus</i>	<i>Crocus</i>	<i>C.Mathewii</i>	CmtHR09	33 A 6	182	1D33	209
					33 A 7	163	7D33	205
					33 A 11	156	10D33	202
					33 A 12	163	12D33	205
					33 A 14	184	15D33	205
34	<i>Crocus</i>	<i>Crocus</i>	<i>C.thomasii</i>	CtmVD09	34 A 9	184	2D34	205
					34 A 10	163	3D34	205
					34 A 12	202	7D34	205
					34 A 13	163	12D34	202
					34 A 14	182	13D34	209
35	<i>Crocus</i>	<i>Crocus</i>	<i>C.thomasii</i>	CtomI09john	35 A 2	163	1D35	206
					35 A 3	191	4D35	205
					35 A 7	168	6D35	205
					35 A 8	Not good 167	9D35	205
					35 A 11	168	15D35	205
36	<i>Crocus</i>	<i>Crocus</i>	<i>C.asumaniae</i>	CasWD09	36 A 1	165	1D36	205
					36 A 2	174	3D36	205
					36 A 3	165	7D36	205
					36 A 5	173	10D36	205
					36 A 8	165	15D36	205
37	<i>Crocus</i>	<i>Crocus</i>	<i>C.asumaniae</i>	CasAD09	37 A 2	192	1D37	205
					37 A 3	165	2D37	205
					37 A 12	171	6D37	205
					37 A 14	192	7D37	205
					37 A 15	163	11D37	205
38	<i>Crocus</i>	<i>Crocus</i>	<i>C.asumaniae</i>	CasAT09jhon	38 A 4	171	2D38	205
					38 A 5	194	3D38	205
					38 A 8	Identical to 38 A 4	4D38	205
					38 A 14	164	11D38	206
					38 A 15	200	13D38	205
39	<i>Crocus</i>	<i>Crocus</i>	<i>C.oreoreticus</i>	CorVR09	39 A 1	162	2D39	205
					39 A 3	Identical to 39 A 9	9D39	205
					39 A 7	198	10D39	202
					39 A 8	162	11D39	205
					39 A 9	162	12D39	205
40	<i>Crocus</i>	<i>Crocus</i>	<i>C.oreoreticus</i>	CorVD09	40 A 3	Not good	1D40	206
					40 A 4	171 53%	3D40	206
					40 A 7	162	6D40	206
					40 A 8	174	10D40	205
					40 A 11	168	13D40	205
41	<i>Crocus</i>	<i>Crocus</i>	<i>C.hadriaticus</i>	ChdWD08	41 A 3	Identical to 41 A 9	1D41	202
					41 A 4	168	7D41	205
					41 A 6	163	10D41	205
					41 A 9	163	11D41	205
					41 A 10	172	15D41	205
42	<i>Crocus</i>	<i>Crocus</i>	<i>C.hadriaticus</i>	ChaIR09	42 A 8	Identical to 42 A 15	5D42	205
					42 A 9	168	7D42	202
					42 A 12	HQ<85%	9D42	205
					42 A 13	168	12D42	205
					42 A 15	163	15D42	202
43	<i>Crocus</i>	<i>Crocus</i>	<i>C.hadriaticus</i>	ChdARD09	43 A 2	163	3D43	205
					43 A 3	178	4D43	206
					43 A 8	196	6D43	202
					43 A 9	163	10D43	205
					43 A 11	198	11D43	202
44	<i>Crocus</i>	<i>Crocus</i>	<i>C.cartwrightianus</i>	CcwBD09	44 A 1	177	3D44	205
					44 A 3	161	4D44	202
					44 A 5	Identical to 44 A 7	7D44	205
					44 A 6	178	8D44	205
					44 A 7	143	15D44	205
45	<i>Crocus</i>	<i>Crocus</i>	<i>C.cartwrightianus</i>	CcrCR09	45 A 1	Identical to 45 A 9	1D45	202
					45 A 2	198	4D45	202
					45 A 4	163	6D45	205
					45 A 9	163	11D45	202
					45 A 10	170	13D45	205
46	<i>Crocus</i>	<i>Crocus</i>	<i>C.cartwrightianus</i>	CcwAD08	46 A 1	Identical to 46 A 13	1D46	205
					46 A 8	162	7D46	205
					46 A 10	192	8D46	202

					46 A 12	178	10 D46	205
					46 A 13	163	15D46	205
47	<i>Crocus</i>	<i>Crocus</i>	<i>C.pallasii</i>	CplVD09	47 A 5	166	2D47	202
					47 A 8	Identical to 47 A 5	5D47	205
					47 A 10	163	6D47	205
					47 A 14	163	10D47	205
					47 A 15	188	14D47	205
					48 A 1	Identical to 48 A 12	1D48	205
48	<i>Crocus</i>	<i>Crocus</i>	<i>C.asumaniae</i>	CasWD09	48 A 5	171	2D48	226
					48 A 12	163	8D48	226
					48 A 13	171	9D48	205
					48 A 14	198	12D48	205
					49 A 1	172	2D49	202
49	<i>Crocus</i>	<i>Crocus</i>	<i>C.cartwrightianus</i>	CcwBD10	49 A 5	170	4D49	202
					49 A 10	168	7D49	205
					49 A 11	163	8D49	205
					49 A 12	Not good	14D49	205



A



B

Figure A4.6: **A)** Cut out sequence logo of the *matK* 390F+1326R plastid gene. Highlighted nucleotides “G, C, A” showing heterozygosity in *C. asumaniae* (2n=26, x=8). **B)** Cut out sequence logo of the *rbcL* plastid gene. Highlighted nucleotides (Y, M) showing heterozygosity in *C. asumaniae* (2n=26, x=8).

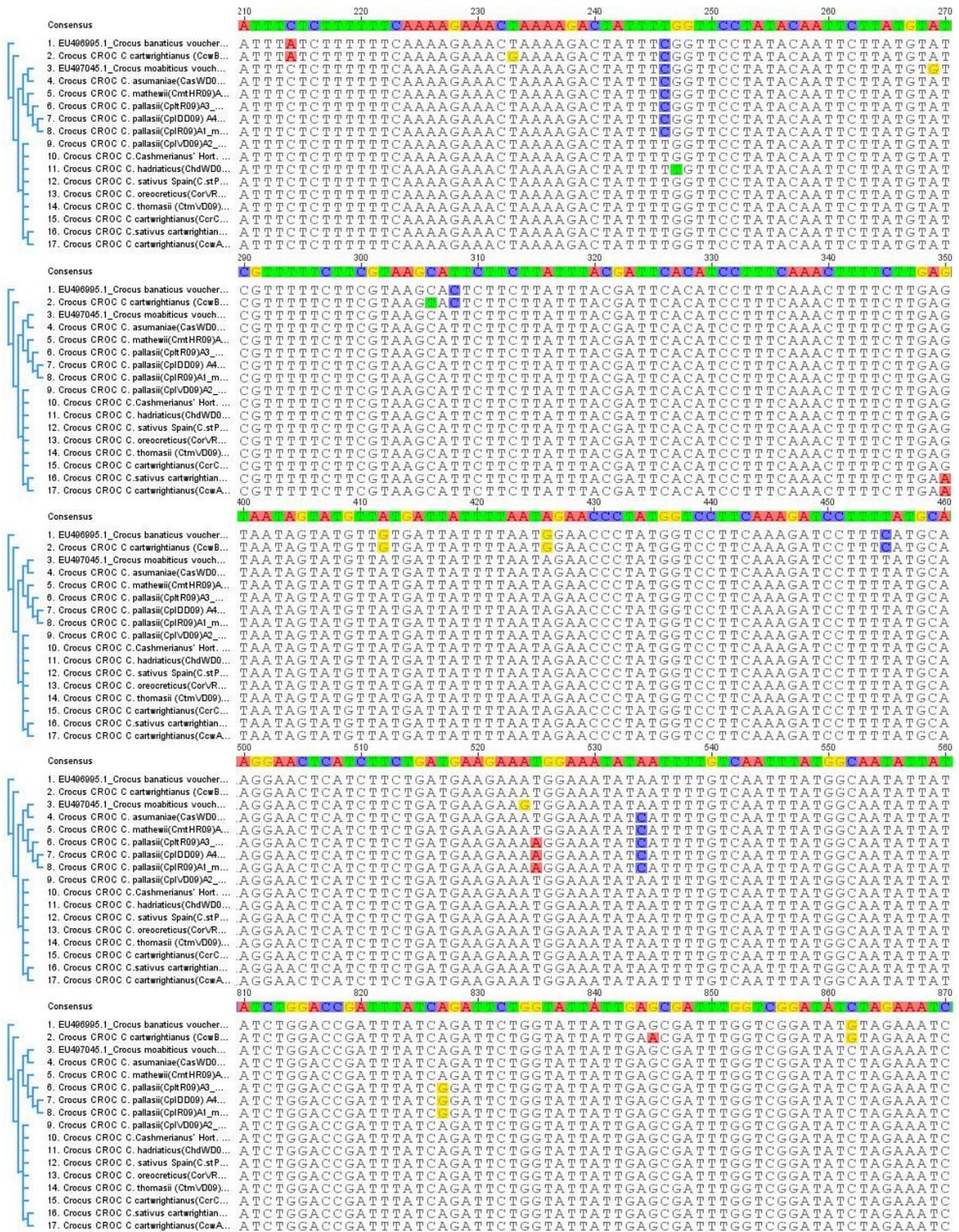


Figure A4.8: Multiple sequence alignment view of the *matK* XF+5R -plastid gene sequences isolated from 17 accessions of *Crocus* species series *Crocus* along with *C. moabiticus* and *C. banaticus* (EU497045 and EU496995) sequences downloaded from NCBI. Sequence length is 802-942bp. Highlighted nucleotides "ATCG" indicating the single nucleotide polymorphism detected in the species.

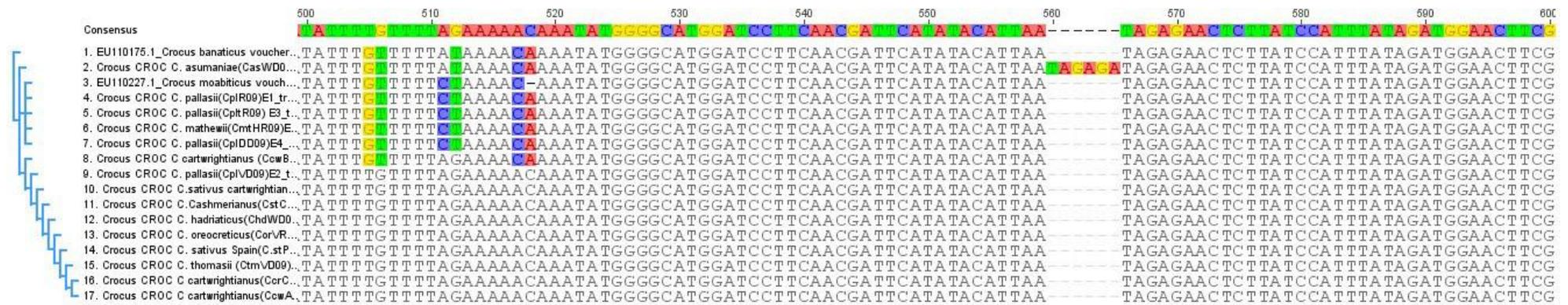


Figure A4.10: Multiple sequence alignment view of the *trnH* -plastid gene sequences isolated from 17 accessions of *Crocus* species series *Crocus* along with *C. moabiticus* and *C. banaticus* (EU110227 and EU110175) sequences downloaded from NCBI, Sequence length is 605-650bp . Highlighted nucleotides “ATCG” indicating to the single nucleotide polymorphism detected in the species and dashes indicate deletions.

1: Consensus *matK* 390F+1326R plastid gene sequences of *Crocus* series *Crocus*, obtained by aligning both forward and reverse sequences for each species.

>EU497045.1_ *Crocus. moabiticus_voucher_C1927_maturase_K_ (matK)_ gene A new nucleotide sequence entered manually.*

AAATCCTTCAATGCTGGATTCAAGATGTTCCCTCTTTTGCATTCTTTCGCGATTCTTCTTCACGAATATCATAATTGGAAT
AGTTTTTAAATTAAGTACTCAGAATAAATCTATTTCTCTTTTTCAAAAGAACTAAAAGACTATTTTCGGTTCCTATACAATTC
TTATGTGTATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCATTCTTCTTATTTACGATTTCACATCCTTTCAAACCTTT
TCTTGAGCGAAGATATTTCTATGGAATAATGGAACATCTTCAAACGGAACATCTTATAATAGTATGTTATGATTATTTT
AATAGAACCCTATGGTCTTTCAAAGATCCTTTTATGCTTATTTGACTATGTCGATGTC AAGGAAAAGCAATTCGGCTTCAAAA
GGAACCTCATCTTCTGATGAAGAAGTGGAAATATAATTTTGTCAATTTATGGCAATATTATTTTCACCTTTTGGTATCAAT
CATACAGGATCCATATAAACCAACTATCAAACCATTCTTCCATTTCTGGGTATCTTTCAAGTTTATTAATAAATAA
TTCGACGGTAAGGAATCAAATGTTAGATATTCATTTCTAATAGATACTTACTACGAAATTTGATACCGCAGTCCCAG
TTATTTTCTTATTAGATCTTTATCTAAAGCTCATTTTGTACTGATCGGGACATCCTATTAGTAAGCCCATCTGGACCG
ATTTATCAGATTCTGGTATTATTGAGCGATTTGGTTCGGATATCTAGAAATCTTTCTCATTATCATAGTGGGTCCTCAA
AAAAA-----

>EU496995.1_ *Crocus. banaticus_voucher_C1821_maturase_K_ (matK)_ gene A new nucleotide sequence entered manually.*

AAATCCTTCAATGCTGGATTCAAGATGTTCCCTCTTTTGCATTCTTTCGCGATT
CTTCTTCACGAATATCATAATTGGAATAGTTTTTAAATTAAGTACTCAGAATAAATCTATTTATCTTTTTTCAAAGAACTA
AAAGACTATTTTCGGTTCCTATACAATCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCACTCTTCT
TATTTACGATTTCACATCCTTTCAAACCTTTTCTTGAGCGAAGATATTTCTATGGAATAATGGAACATCTTCAAACGAAAC
ATCTTATAATAGTATGTTGTGATTATTTAATGGAACCCTATGGTCTTCAAAGATCCTTTTCATGCATTATGCTCGATGT
CAAGGAAAAGCAATTCGGCTTCAAAGGAACTCATCTTCTGATGAAGAAATGGAATATAATTTTGTCAATTTATGG
CAATATTATTTTCACCTTTTGGTATCAATCATACAGGATCCATATAAACCAACTATCAAACCATTCTTTCCATTTTCTGGG
TTATCTTCAAAGTTTACTAAAAAATTCTTTCGACGGTAAGGAATCAAATGTTAGATAATTCATTTCTAATAGATACTCTT
ACTACGAAATTTGATACCGCAGTCCCAGTTATTTTCTTATTAGATCTTTATCAAAGCTCAATTTTGTACCGTATCGGG
ACATCCCATTAGTAAGCCCATCTGGACCGATTTATCAGATTCTGGTATTATTGAGCGATTTGGTTCGGATATGTAGAAAT
CTTCTCATTATCATAGTGGGTCCTCAAAAAACA-----

>*Crocus_CROC_C. cartwrightianus_ (CcwBD10)9D_matK_ (390F_1326R).*

TCGATCTATTCATTCAATATTTCCCTTTTTAGAGGACAAATTATTACATTTACATTATGTATCAGATATACTAATACCCC
ATCCCATTATGGAATCTTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTCTTTTGCATTCTTTCGCGATTCT
TTCTTCACGAATATCATAATTGGAATAGTTTTTAAATTAAGTACTCAGAATAAATCTATTTATCTTTTTTCAAAGAACTAA
AAGACTATTTTCGGTTTTCTATACAATCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCACTCTTCT
ATTTACGATTTCACATCCTTTCAAACCTTTTCTTGAGCGAAGATATTTCTATGGAATAATGGAACATCTTCAAACGAAAC
TCTTATAATAGTATGTTGTGATTATTTAATGGAACCCTATGGTCTTCAAAGATCCTTTTCATGCATTATGCTCGATGTC
AAGGAAAAGCAATTCGGCTTCAAAGGAACTCATCTTCTGATGAAGAAATGGAATATAATTTTGTCAATTTATGGC
AATATTATTTTCACCTTTTGGTATCAATCATACAGGATCCATATAAACCAACTATCAAACCATTCTTTCCATTTTCTGGGT
TATCTTCAAAGTTTACTAAAAAATTCTTTCGACGGTAGGAATCAAATGTTAGATAATTCATTTCTAATAGATACTCTTA
CTTACTACGAAATTTGATACCGCAGTCCCAGTTATTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACCGGTA
TCGGGACATCCTATTAGTAAGCCCATCTGGACCGATTTATCAGATTCTGGTATTATTGAGCGATTTGGTTCGGATATGTA
GAAATCTTTCTCATTATCATAGTGGGTCCTCAAAAAAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTTTCGTT
TTTGCTAAAAA

>*Crocus_CROC_C. cartwrightianus (CcrCR09) D8matK_ (390F_1326R).*

TCGATCTATTCATTCAATATTTCCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCCC
CATCCCATTATGGAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTCTTTTGCATTCTTTCGCGATT
CTTCTTCACGAATATCATAATTGGAATAGTTTTTAAATTAAGTACTCAGAATAAATCTATTTCTCTTTTTTCAAAGAACTA
AAAGACTATTTTGGTTCCTATACAATCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCACTCTTCT
TATTTACGATTTCACATCCTTTCAAACCTTTTCTTGAACGAAAGATATTTCTATGGAATAATGGAACATCTTCAAACGGAAC
ATCTTATAATAGTATGTTATGATTATTTAATAGAACCCTATGGTCTTCAAAGATCCTTTTATGCATTATGCTCGATGT
CAAGGAAAAGCAATTCGGCTTCAAAGGAACTCATCTTCTGATGAAGAAATGGAATATAATTTTGTCAATTTATGG
CAATATTATTTTCACCTTTTGGTATCAATCATACAGGATCCATATAAACCAACTATCAAACCATTCTTTCCATTTTCTGGG
TTATCTTCAAAGTTTACTAAAAAATTCTTTCGACGGTAGGAATCAAATGTTAGATAATTCATTTCTAATAGATACTCTTA
CTACGAAATTTGATACCGCAGTCCCAGTTATTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACTGTATCGGGAC
ATCCTATTAGTAAGCCCATCTGGACCGATTTATCAGATTCTGGTATTATTGAGCGATTTGGTTCGGATATCTAGAAATCT
TTCTCATTATCATAGTGGGTCCTCAAAAAAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTTTCGTTTTTGTCTA
AAAA

>*Crocus_CROC_C_cartwrightianus* (CcwAD08) D7__matK_ (390F_1326R).

TCGATCTATTCATTCAATATTTCCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCC
CATCCCATTCCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTCTTTTGCATTTCTTGGCATT
CTTCTTCACGAATATCATAATTGGAATAGTTTTTAATTACTCAGAATAAACTATTTCTCTTTTTTCAAAGAAACTA
AAAGACTATTTTGGTTCCCTATACAATTTCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCATTCTTCT
TATTACGATTCACATCCTTTCAAACCTTTCTTGAACGAAGATATTTCTATGGAAAAATGGAACATCTTCAAACGGAAC
ATCTTATAATAGTATGTTATGATTATTTTAATAGAACCCTATGGTCCTTCAAAGATCCTTTTATGCATTATGCTCGATGT
CAAGGAAAAGCAATTTCTGGCTTCAAAGGAACACTATCTTCTGATGAAGAAATGGAATAAATTTTGTCAATTTATGG
CAATATTATTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAAACCAATTTTCCATTTTCTGGG
TTATCTTCAAGTTTACTAAAAAATTTCTTCGACGGTAAGGAATCAAATGTTAGATAAATTCATTTCTAATAGATACTCTT
ACTACGAAATTTGATACCGCAGTCCCAGTTATTTTCTTATTAGATCTTTATCAAAGCTCAATTTTGTACTGTATCGGG
ACATCTATTAGTAAGCCCATCTGGACCGATTATCAGATTCTGGTATTATTGAGCGATTTGGTCGGATATCTAGAAATC
CTTTCTCATTATCATAGTGGTCTCTCAAAAAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTTCGTTTTTGC
TAAAAA

>*Crocus_CROC_C. Kashmirianus'* Hort. Lilac (CstCD09)5D_matK_ (390F_1326R).

TCGATCTATTCATTCAATATTTCCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCC
ATCCCATTCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTCTTTTGCATTTCTTGGCATT
TTTCTTCACGAATATCATAATTGGAATAGTTTTTAATTACTCAGAATAAACTATTTCTCTTTTTTCAAAGAAACTAA
AAGACTATTTTGGTTCCCTATACAATTTCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCATTCTTCT
ATTTACGATTCACATCCTTTCAAACCTTTCTTGAACGAAGATATTTCTATGGAAAAATGGAACATCTTCAAACGGAACA
TCTTATAATAGTATGTTATGATTATTTTAATAGAACCCTATGGTCCTTCAAAGATCCTTTTATGCATTATGCTCGATGTC
AAGGAAAAGCAATTTCTGGCTTCAAAGGAACACTATCTTCTGATGAAGAAATGGAATAAATTTTGTCAATTTATGGC
AATATTATTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAAACCAATTTTCCATTTTCTGGGT
TATCTTTCAAGTTTACTAAAAAATTTCTTCGACGGTAAGGAATCAAATGTTAGATAAATTCATTTCTAATAGATACTCTTA
CTACGAAATTTGATACCGCAGTCCCAGTTATTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACTGTATCGGGA
CATCCTATTAGTAAGCCCATCTGGACCGATTATCAGATTCTGGTATTATTGAGCGATTTGGTCGGATATCTAGAAATC
TTTCTCATTATCATAGTGGTCTCTCAAAAAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTTCGTTTTTGC
AAAAA

>*Crocus_CROC_C. hadriaticus* (ChdWD08) D11_matK_ (390F_1326R).

TCGATCTAATTCATTCAATATTTCCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCC
CATCCCATTCCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTCTTTTGCATTTCTTACGATT
CTTCTTCACGAATATCATAATTGGAATAGTTTTTAATTACTCAGAATAAACTATTTCTCTTTTTTCAAAGAAACTA
AAAGACTATTTTTGGTTCCCTATACAATTTCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCATTCTTCT
ATTTACGATTCACATCCTTTCAAACCTTTCTTGAACGAAGATATTTCTATGGAAAAATGGAACATCTTCAAACGGAACA
TCTTATAATAGTATGTTATGATTATTTTAATAGAACCCTATGGTCCTTCAAAGATCCTTTTATGCATTATGCTCGATGTC
AAGGAAAAGCAATTTCTGGCTTCAAAGGAACACTATCTTCTGATGAAGAAATGGAATAAATTTTGTCAATTTATGGC
AATATTATTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAAACCAATTTTCCATTTTCTGGGT
TATCTTTCAAGTTTACTAAAAAATTTCTTCGACGGTAAGGAATCAAATGTTAGATAAATTCATTTCTAATAGATACTCTTA
CTACGAAATTTGATACCGCAGTCCCAGTTATTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACTGTATCGGGA
CATCCTATTAGTAAGCCCATCTGGACCGATTATCAGATTCTGGTATTATTGAGCGATTTGGTCGGATATCTAGAAATC
TTTCTCATTATCATAGTGGTCTCTCAAAAAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTTCGTTTTTGC
AAAAA

>*Crocus_CROC_C. oreoreticus* (CorVR09)10D_matK_ (390F_1326R).

TCGATCTATTCATTCAATATTTCCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCC
ATCCCATTCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTCTTTTGCATTTCTTGGCATT
TTTCTTCACGAATATCATAATTGGAATAGTTTTTAATTACTCAGAATAAACTATTTCTCTTTTTTCAAAGAAACTAA
AAGACTATTTTGGTTCCCTATACAATTTCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCATTCTTCT
ATTTACGATTCACATCCTTTCAAACCTTTCTTGAACGAAGATATTTCTATGGAAAAATGGAACATCTTCAAACGGAACA
TCTTATAATAGTATGTTATGATTATTTTAATAGAACCCTATGGTCCTTCAAAGATCCTTTTATGCATTATGCTCGATGTC
AAGGAAAAGCAATTTCTGGCTTCAAAGGAACACTATCTTCTGATGAAGAAATGGAATAAATTTTGTCAATTTATGGC
AATATTATTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAAACCAATTTTCCATTTTCTGGGT
TATCTTTCAAGTTTACTAAAAAATTTCTTCGACGGTGAATCAAATGTTAGATAAATTCATTTCTAATAGATACTCTTACT
ACGAAATTTGATACCGCAGTCCCAGTTATTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACTGTATCGGGAC
ATCCTATTAGTAAGCCCATCTGGACCGATTATCAGATTCTGGTATTATTGAGCGATTTGGTCGGATATCTAGAAATCT
TTCTCATTATCATAGTGGTCTCTCAAAAAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTTCGTTTTTGTCTA
AAAA

>*Crocus_CROC_C. asumaniae* (CasWD09) D13_matK_ (390F_1326R)

TCGATCTATTCATTCAATATTTCCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCCC
ATCCCATTCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTCTTTTGCATTTCTTGGCATT
TTCTTTCACGAATATCATAATTGGAATAGTTTTTAATTACTCAGAATAAAATCTATTTCTCTTTTTTCAAAGAAACTAA
AAGACTATTTCCGGTTCCTATACAATTCCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCATTCTTCTT
ATTTACGATTCACATCCTTTCAAACCTTTTCTTGAGCGAAGATATTTCTATGGAAAAATGGAACATCTTCAAACCGGAACA
TCTTATAATAGTATGTTATGATTATTTAATAGAACCCTATGGTCCTCAAAGATCCTTTTATGCATTATGCTCGATGTC
AAGGAAAAGCAATTCTGGCTTCAAAGGAACCTCATCTTCTGATGAAGAAAATGGAATATCATTTTGTCAATTTATGGC
AATATTATTTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAACCATTCTTCCATTTTCTGGGT
ATCTTTCAAGTTTACTAAAAAATCTTTCGACGGTAAGGAATCAAATGTTAGATAATTCATTTCTAATAGATACTCTTAC
TACGAAATTTGATACCGCAGTCCCAGTTATTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACTGTATCGGGAC
ATCTTATTAGTAAGCCCATCTGGACCGATTTATCAGATTCTGGTATTATTGAGCGATTTGGTCGGATATCTAGAAATCT
TTCTCATTATCATAGTGGTCTCAAACAAAACAGGGGTGTATCGAATAAAGTATATACTTCGACTTTCGGTTTTTGCTA
AAAA

>*Crocus_CROC_C. mathweii* (CmtHR09) D14_matK_ (390F_1326R)

TCGATCTATTCATTCAATATTTCCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCCC
ATCCCATTCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTCTTTTGCATTTCTTGGCATT
TTTCTTTCACGAATATCATAATTGGAATAGTTTTTAATTACTCAGAATAAAATCTATTTCTCTTTTTTCAAAGAAACTAA
AAGACTATTTCCGGTTCCTATACAATTCCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCATTCTTCTTA
TTACGATTCACATCCTTTCAAACCTTTTCTTGAGCGAAGATATTTCTATGGAAAAATGGAACATCTTCAAACCGGAACAT
CTTATAATAGTATGTTATGATTATTTAATAGAACCCTATGGTCCTCAAAGATCCTTTTATGCATTATGCTCGATGTC
AGGAAAAGCAATTCTGGCTTCAAAGGAACCTCATCTTCTGATGAAGAAAATGGAATATCATTTTGTCAATTTATGGCA
ATATTATTTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAACCATTCTTCCATTTTCTGGGT
ATCTTTCAAGTTTACTAAAAAATCTTTCGACGGTAAGGAATCAAATGTTAGATAATTCATTTCTAATAGATACTCTTAC
TACGAAATTTGATACCGCAGTCCCAGTTATTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACTGTATCGGGAC
ATCTTATTAGTAAGCCCATCTGGACCGATTTATCAGATTCTGGTATTATTGAGCGATTTGGTCGGATATCTAGAAATCT
TTCTCATTATCATAGTGGTCTCAAACAAAACAGGGGTGTATCGAATAAAGTATATACTTCGACTTTCGGTTTTTGCTA
AAAA

>*Crocus_CROC_C. pallasii* (CpIDD09) 4D_matK_ (390F_1326R)

TCGATCTATTCATTCAATATTTCCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCCC
ATCCCATTCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTCTTTTGCATTTCTTGGCATT
TTTCTTTCACGAATATCATAATTGGAATAGTTTTTAATTACTCAGAATAAAATCTATTTCTCTTTTTTCAAAGAAACTAA
AAGACTATTTCCGGTTCCTATACAATTCCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCATTCTTCTT
ATTTACGATTCACATCCTTTCAAACCTTTTCTTGAGCGAAGATATTTCTATGGAAAAATGGAACATCTTCAAACCGGAACA
TCTTATAATAGTATGTTATGATTATTTAATAGAACCCTATGGTCCTCAAAGATCCTTTTATGCATTATGCTCGATGTC
AAGGAAAAGCAATTCTGGCTTCAAAGGAACCTCATCTTCTGATGAAGAAAATGGAATATCATTTTGTCAATTTATGGC
AATATTATTTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAACCATTCTTCCATTTTCTGGGT
TATCTTTCAAGTTTACTAAAAAATCTTTCGACGGTAAGGAATCAAATGTTAGATAATTCATTTCTAATAGATACTCTTA
CTACGAAATTTGATACCGCAGTCCCAGTTATTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACTGTATCGGGA
CATCCTATTAGTAAGCCCATCTGGACCGATTTATCGGATTTCTGGTATTATTGAGCGATTTGGTCGGATATCTAGAAATC
TTTCTCATTATCATAGTGGTCTCAAACAAAACAGGGGTGTATCGAATAAAGTATATACTTCGACTTTCGGTTTTTGCT
AAAAA

>*Crocus_CROC_C. pallasii* (CpIR09) 1D_matK_ (390F_1326R)

TCGATCTATTCATTCAATATTTCCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCCC
ATCCCATTCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTCTTTTGCATTTCTTGGCATT
TTTCTTTCACGAATATCATAATTGGAATAGTTTTTAATTACTCAGAATAAAATCTATTTCTCTTTTTTCAAAGAAACTAA
AAGACTATTTCCGGTTCCTATACAATTCCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCATTCTTCTT
ATTTACGATTCACATCCTTTCAAACCTTTTCTTGAGCGAAGATATTTCTATGGAAAAATGGAACATCTTCAAACCGGAACA
TCTTATAATAGTATGTTATGATTATTTAATAGAACCCTATGGTCCTCAAAGATCCTTTTATGCATTATGCTCGATGTC
AAGGAAAAGCAATTCTGGCTTCAAAGGAACCTCATCTTCTGATGAAGAAAATGGAATATCATTTTGTCAATTTATGGC
AATATTATTTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAACCATTCTTCCATTTTCTGGGT
TATCTTTCAAGTTTACTAAAAAATCTTTCGACGGTAAGGAATCAAATGTTAGATAATTCATTTCTAATAGATACTCTTA
CTACGAAATTTGATACCGCAGTCCCAGTTATTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACTGTATCGGGA
CATCCTATTAGTAAGCCCATCTGGACCGATTTATCGGATTTCTGGTATTATTGAGCGATTTGGTCGGATATCTAGAAATC
TTTCTCATTATCATAGTGGTCTCAAACAAAACAGGGGTGTATCGAATAAAGTATATACTTCGACTTTCGGTTTTTGCT
AAAAA

>*Crocus_CROC_C. pallasii* (CplR09) D3_matK_ (390F_1326R).

TCGATCTATTCATTCAATATTTCCCTTTTTAGAGGACAAATTTTACATTTCCATTATGTATCAGATATACTAATACCCC
ATCCCATTCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTCTTTTGCATTTCTTGGCATT
TTCTTTCACGAATATCATAATTGGAATAGTTTTTAATTACTCAGAATAAAATCTATTTCTCTTTTTTCAAAGAAACTAA
AAGACTATTTTCGGTTCCTATACAATTCCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCATTCTTCTT
ATTTACGATTCACATCCTTTCAAACCTTTCTTGAGCGAAGATATTTCTATGGAAAAATGGAACATCTTCAAACCGGAACA
TCTTATAATAGTATGTTATGATTATTTAATAGAACCCTATGGTCCTTCAAAGATCCTTTTATGCATTATGCTCGATGTC
AAGGAAAAGCAATTCTGGCTTCAAAGGAACATCTTCTGATGAAGAAAAGGAAATATCATTTTGTCAATTTATGGC
AATATTATTTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAAACCATTTCTCCATTTTCTGGGT
TATCTTTCAAGTTTACTAAAAAATTTCTCGACGGTAAGGAATCAAATGTTAGATAATTCATTTCTAATAGATACTCTTA
CTACGAAATTTGATACCGCAGTCCCAGTTATTTTCTTATAGATCTTTATCTAAAGCTCAATTTTGTACTGTATCGGGA
CATCCTATTAGTAAGCCCATCTGGACCGATTTATCGGATTCTGGTATTATTGAGCGATTTGGTCGGATATCTAGAAATC
TTTCTCATTATCATAGTGGTCTCAAAGAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTTCGTTTTTGTCT
AAAA

>*Crocus_CROC_C. pallasii* (CplVD09) _ D2_matK_ (390F_1326R).

TCGATCTATTTTCATTCAATATTTCCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCC
CATCCCATTCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTCTTTTGCATTTCTTGGCATT
CTTTCTTACGAATATCATAATTGGAATAGTTTTTAATTACTCAGAATAAAATCTATTTCTCTTTTTTCAAAGAAACTAA
AAAGACTATTTTGGTTCCTATACAATTCCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCATTCTTCT
TATTTACGATTCACATCCTTTCAAACCTTTCTTGAGCGAAGATATTTCTATGGAAAAATGGAACATCTTCAAACCGGAAC
ATCTTATAATAGTATGTTATGATTATTTAATAGAACCCTATGGTCCTTCAAAGATCCTTTTATGCATTATGCTCGATGT
CAAGGAAAAGCAATTCTGGCTTCAAAGGAACATCTTCTGATGAAGAAAATGGAATATAATTTTGTCAATTTATGG
CAATATTATTTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAAACCATTTCTTTCCATTTTCTGGG
TTATCTTTCAAGTTTACTAAAAAATTTCTCGAGTAAGGAATCAAATGTTAGATAATTCATTTCTAATAGATACTCTTAC
TACGAAATTTGATACCGCAGTCCCAGTTATTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACTGTATCGGGAC
ATCCTATTAGTAAGCCCATCTGGACCGATTTATCAGATTCTGGTATTATTGAGCGATTTGGTCGGATATCTAGAAATCT
TTCTCATTATCATAGTGGTCTCAAAGAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTTCGTTTTTGTCTA
AAAA

>*Crocus_CROC_C. sativuscartwrightianus* 'Albus'(CstcP09)15D_matK_(390F_1326R).

TCGATCTATTCATTCAATATTTCCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCCC
ATCCCATTCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTCTTTTGCATTTCTTGGCATT
TTTCTTTCACGAATATCATAATTGGAATAGTTTTTAATTACTCAGAATAAAATCTATTTCTCTTTTTTCAAAGAAACTAA
AAGACTATTTTGGTTCCTATACAATTCCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCATTCTTCTT
ATTTACGATTCACATCCTTTCAAACCTTTCTTGAGCGAAGATATTTCTATGGAAAAATGGAACATCTTCAAACCGGAACA
TCTTATAATAGTATGTTATGATTATTTAATAGAACCCTATGGTCCTTCAAAGATCCTTTTATGCATTATGCTCGATGTC
AAGGAAAAGCAATTTCTGGCTTCAAAGGAACATCTTCTGATGAAGAAAATGGAATATAATTTTGTCAATTTATGGC
AATATTATTTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAAACCATTTCTTTCCATTTTCTGGGT
TATCTTTCAAGTTTACTAAAAAATTTCTCGACGGTAAGGAATCAAATGTTAGATAATTCATTTCTAATAGATACTCTTA
CTACGAAATTTGATACCGCAGTCCCAGTTATTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACTGTATCGGGA
CATCCTATTAGTAAGCCCATCTGGACCGATTTATCAGATTCTGGTATTATTGAGCGATTTGGTCGGATATCTAGAAATC
TTTCTCATTATCATAGTGGTCTCAAAGAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTTCGTTTTTGTCT
AAAA

>*Crocus_CROC_C. sativus_Spain* (C.stPER09) A6_matK_ (390F_1326R).

TCGATCTATTCATTCAATATTTCCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCCC
ATCCCATTCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTCTTTTGCATTTCTTGGCATT
TTTCTTTCACGAATATCATAATTGGAATAGTTTTTAATTACTCAGAATAAAATCTATTTCTCTTTTTTCAAAGAAACTAA
AAGACTATTTTGGTTCCTATACAATTCCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCATTCTTCTT
ATTTACGATTCACATCCTTTCAAACCTTTCTTGAGCGAAGATATTTCTATGGAAAAATGGAACATCTTCAAACCGGAACA
TCTTATAATAGTATGTTATGATTATTTAATAGAACCCTATGGTCCTTCAAAGATCCTTTTATGCATTATGCTCGATGTC
AAGGAAAAGCAATTTCTGGCTTCAAAGGAACATCTTCTGATGAAGAAAATGGAATATAATTTTGTCAATTTATGGC
AATATTATTTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAAACCATTTCTTTCCATTTTCTGGGT
TATCTTTCAAGTTTACTAAAAAATTTCTCGACGGTAAGGAATCAAATGTTAGATAATTCATTTCTAATAGATACTCTTA
CTACGAAATTTGATACCGCAGTCCCAGTTATTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACTGTATCGGGA
CATCCTATTAGTAAGCCCATCTGGACCGATTTATCAGATTCTGGTATTATTGAGCGATTTGGTCGGATATCTAGAAATC
TTTCTCATTATCATAGTGGTCTCAAAGAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTTCGTTTTTGTCT
AAAA

>*Crocus_CROC_C. thomasi* (CtmVD09) D12_matK_ (390F_1326R).

TCGATCTATTCATTCAATATTCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCCC
ATCCCATTCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCTCTTTTGCATTTCTTGCAATTC
TTCTTCACGAATATCATAATTGGAATAGTTTTTAATTACTCAGAATAAATCTATTTCTCTTTTTTCAAAAAGAACTAA
AAGACTATTTTGGTTCCTATACAATTCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCATTCTTCTT
ATTACGATTCACATCCTTTCAAACCTTTCTTGAGCGAAGATATTTCTATGGAAAAATGGAACATCTTCAAACGGAACA
TCTTATAATAGTATGTTATGATTATTTAATAGAACCCATGGTCCTTCAAAGATCCTTTTATGCATTATGCTCGATGTC
AAGGAAAAGCAATTCTGGCTTCAAAAAGGAAGCTATCTTCTGATGAAGAAATGGAAATATAATTTTGTCAATTTATGGC
AATATTATTTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAAACCATCTTCCATTTTCTGGGT
TATCTTTCAAGTTTACTAAAAAATTCTTCGACGGTAAGGAATCAAATGTTAGATAATTCATTTCTAATAGATACTCTTA
CTACGAAATTTGATACCGCAGTCCAGTTATTTTCTTATTAGATCTTATCTAAAGCTCAATTTTGTACTGTATCGGGA
CATCCTATTAGTAAGCCATCTGGACCGATTTATCAGATTCTGGTATTATTGAGCGATTTGGTCGGATATCTAGAAATC
TTTCTATTATCATAGTGGGTCTCAAAAAAACAGGGTTGTATCGAATAAAGTATATACTTCGACTTTCGTTTTTGCT
AAAAA

2: Consensus *matK* XF+5R plastid gene sequences of *Crocus* series *Crocus*, obtained by aligning both forward and reverse sequences for each species.

>EU496995.1_ *Crocus banaticus*_voucher_C1821_maturase_K_ (*matK*)_ gene A new nucleotide sequence entered manually.

AAATCCTTCAATGCTGGATTCAAGATGTTCCCTCTTTTGCATTCTTTCGCGATTCTTTCTTCACGAATATCATAATTGGAAT
AGTTTTTTAATTACTCAGAATAAATCTATTATCTTTTTTCAAAGAACTAAAAGACTATTCGGTTCCTATACAATTC
TTATGTATATGAATGTGAATTTGTATTCTTTTCTTCGTAAGCACTCTTCTTATTACGATTACATCCTTTCAAACCTTT
TCTTGAGCGAAGATATTTCTATGGAAAAATGGAACATCTTCAAACAGAACATCTTATAATAGTATGTTGTGATTATTTT
AATGGAACCTATGGTCCTTCAAAGATCCTTTTCATGCATTATGCTCGATGTCAAGGAAAAAGCAATTCTGGCTTCAAAA
GGAACCTCATCTTCTGATGAAGAAATGGAATATAATTTTGTCAATTTATGGCAATATTATTTTCACTTTTGGTATCAAT
CATAACAGGATCCATATAAACCCTATCAAACCTATTCTTCCATTTTCTGGGTTATCTTTCAAGTTTACTAAAAAATTC
TTTGACGGTAAGGAATCAAATGTTAGATAATTCATTCTAATAGATACTCTTACTACGAAATTTGATACCGCAGTCCCA
GTTATTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACCGTATCGGGACATCCCATAGTAAGCCCATCTGGAC
CGATTTATCAGATTCTGGTATTATTGAGCGATTGGTCCGATATGTAGAAATCTTCTCATTATCATAGTGGTCTCTCA
AAAAACA-----

>*Crocus_CROC_C. cartwrightianus* (CcwBD10) A9_ *matK* _ (XF_5R).

CAATTCATTCAACTTTTCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCCCATCCC
ATTCATATGGAATCTTGGTCAAATCCTTCAATGTGGATTCAAGATGTTCCCTCTTTTGCATTCTTTCGCGATTCTTTCT
TCACGAATATCATAAATGGAATAGTTTTTAACTACTCAGAATAAATCTATTATCTTTTTTCAAAGAACTAAAAGAA
CTATTTTCGGTTCCTATACAATCTTATGTATATGAATGTGAATTTGTATTCGTTTTTCTTCGTAAGCACTCTTCTTATTA
CGATTCACATCCTTTCAAACCTTTCTTGAGCGAAGATATTTCTATGGAAAAATGGAACATCTTCAAACGGAACATCTTA
TAATAGTATGTTGTGATTATTTAATGGAACCTATGGTCTTCAAAGATCCTTTTCATGCATTATGCTCGATGCCAAGG
AAAAGCAATTCTGGCTTCAAAGGAACCTCATCTTCTGATGAAGAAATGGAATATAATTTTGTCAATTTATGGCAATA
TTATTTTCACTTTTGGTATCAATCATAACAGGATACATATAAACCCTATCAAACCTATTCTTCCATTTTCTGGGTTATC
TTTCAAGTTTACTAAAAAATCTTTCGACGGTAAGGAATCAAATGTTAGATAGTTCAATTTCTAGTAGATACTCTTACTAC
GAAATTTGATACCCAGTCCAGTATTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACCGTATCGGGACAT
CCTATTAGTAAGCCCATCTGGACCGATTATCAGATTCTGGTATTATTGAACGATTGGTCCGATATGTAGAAATCTTT
TCATTATCATAGTGGTCTCAAAAAAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTTCTG

>EU497045.1_ *Crocus moabiticus*_voucher_C1927_maturase_K_ (*matK*)_ gene A new nucleotide sequence entered manually.

AAATCCTTCAATGCTGGATTCAAGATGTTCCCTCTTTTGCATTCTTTCGCGATTCTTTCTTCACGAATATCATAATTGGAAT
AGTTTTTTAATTACTCAGAATAAATCTATTCTTTTTTCAAAGAACTAAAAGACTATTCGGTTCCTATACAATTC
TTATGTATATGAATGTGAATTTGTATTCTTTTCTTCGTAAGCACTCTTCTTATTACGATTACATCCTTTCAAACCTTT
TCTTGAGCGAAGATATTTCTATGGAAAAATGGAACATCTTCAAACGGAACATCTTATAATAGTATGTTATGATTATTTT
AATAGAACCCTATGGTCCTTCAAAGATCCTTTTATGCATTATGCTCGATGTCAAGGAAAAAGCAATTCTGGCTTCAAAA
GGAACCTCATCTTCTGATGAAGAAGTGGAAATATAATTTTGTCAATTTATGGCAATATTATTTTCACTTTTGGTATCAAT
CATAACAGGATCCATATAAACCCTATCAAACCTATTCTTCCATTTTCTGGGTTATCTTTCAAGTTTATAAAAAATTA
TTTCGACGGTAAGGAATCAAATGTTAGATAATTCATTTCTAATAGATACTCTTACTACGAAATTTGATACCGCAGTCCCA
GTTATTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACTGTATCGGGACATCCTATTAGTAAGCCCATCTGGAC
CGATTTATCAGATTCTGGTATTATTGAGCGATTGGTCCGATATCTAGAAATCTTTCTCATT
TCATAGTGGTCTCAAAAAACA-----

>*Crocus_CROC_C. asumaniae* (CasWD09) A13_ *matK* _ (XF_5R).

CTATTCATTCAATATTTCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCCCATCCC
ATTCATATGGAATATTGGTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTCTTTTGCATTCTTTCGCGATTCTTTCT
TCACGAATATCATAATTGGAATAGTTTTTAACTACTCAGAATAAATCTATTCTTTTTTCAAAGAACTAAAAGAA
CTATTTTCGGTTCCTATACAATCTTATGTATATGAATGTGAATTTGTATTCGTTTTTCTTCGTAAGCACTCTTCTTATTA
CGATTCACATCCTTTCAAACCTTTCTTGAGCGAAGATATTTCTATGGAAAAATGGAACATCTTCAAACGGAACATCTTA
TAATAGTATGTTATGATTATTTAATAGAACCCTATGGTCTTCAAAGATCCTTTTATGCATTATGCTCGATGTCAAGG
AAAAGCAATTCTGGCTTCAAAGGAACCTCATCTTCTGATGAAGAAATGGAATATCATTTTTGTCAATTTATGGCAATA
TTATTTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCCTATCAAACCTATTCTTCCATTTTCTGGGTTATC
TTTCAAGTTTACTAAAAAATCTTTCGACGGTAAGGAATCAAATGTTAGATAATTCATTTCTAATAGATACTCTTACTAC
GAAATTTGATACCGCAGTCCAGTATTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACTGTATCGGGACATC
CTATTAGTAAGCCCATCTGGACCGATTATCAGATTCTGGTATTATTGAGCGATTGGTCCGATATCTAGAAATCTTTCT
TCATTATCATAGTGGTCTCAAAAAAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTTCTG

>*Crocus_CROC_C. mathewii* (CmtHR09) A14_matK_ (XF_5R).

CAATTCATTCAACTTTTCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCCCATCCCA
TTCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTCTTTTGCATTTCTTGCATTCTTTCTT
CAGGAATATCATAAATGGAAATAGTTTTTAATTAATCTCAGAATAAATCTATTTCTCTTTTTTCAAAGAAAACATAAAGAC
TATTTCCGGTTCCTATACAATTTCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCATTCTTCTTATTTAC
GATTCACATCCTTTCAAACCTTTCTTGAGCGAAGATATTTCTATGGAAAAATGGAACATCTTCAAACGGAACATCTTAT
AATAGTATGTTATGATTATTTAATAGAACCCTATGGTCCTTCAAAGATCCTTTTATGCATTATGCTCGATGTCAAGGA
AAAGCAATTCTGGCTTCAAAGGAACATCTTCTGATGAAGAAAATGGAATATCATTTTGTCAATTTATGGCAATAT
TATTTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAAACCATTCTTTCCATTTTCTGGGTTATCT
TTCAAGTTTACTAAAAAATTCTTCGACGGTAAGGAATCAAATGTTAGATAATTCATTCTAATAGATACTCTTACTACG
AAATTTGATACCGCAGTCCCAGTTATTTTCTTATTAGATCTTTATCTAAAAGCTCAATTTTGTACTGTATCGGGACATCCT
ATTAGTAAGCCCATCTGGACCGATTTATCAGATTCTGGTATTATTGAGCGATTGGTCCGGATATCTAGAAAATCTTTCTC
ATTATCATAGTGGGTCTCAAACAAAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTCTG

>*Crocus_CROC_C. pallasii* (CpltR09) A3_matK_ (XF_5R).

CAATTCATTCAACTTTTCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCCCATCCC
ATTCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTCTTTTGCATTTCTTGCATTCTTTCT
TCACGAATATCATAAATGGAAATAGTTTTTAATTAATCTCAGAATAAATCTATTTCTCTTTTTTCAAAGAAAACATAAAGAC
CTATTTCCGGTTCCTATACAATTTCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCATTCTTCTTATTTA
CGATTCACATCCTTTCAAACCTTTTCTTGAGCGAAGATATTTCTATGGAAAAATGGAACATCTTCAAACGGAACATCTTA
TAATAGTATGTTATGATTATTTAATAGAACCCTATGGTCCTTCAAAGATCCTTTTATGCATTATGCTCGATGTCAAGG
AAAAGCAATTCTGGCTTCAAAGGAACATCTTCTGATGAAGAAAATGGAATATCATTTTGTCAATTTATGGCAATA
TTATTTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAAACCATTCTTTCCATTTTCTGGGTTATC
TTCAAGTTTACTAAAAAATTCTTCGACGGTAAGGAATCAAATGTTAGATAATTCATTCTAATAGATACTCTTACTAC
GAAATTTGATACCGCAGTCCCAGTTATTTTCTTATTAGATCTTTATCTAAAAGCTCAATTTTGTACTGTATCGGGACATC
CTATTAGTAAGCCCATCTGGACCGATTTATCGGATCTGGTATTATTGAGCGATTGGTCCGGATATCTAGAAAATCTTTCT
TCATTATCATAGTGGGTCTCAAACAAAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTCTG

>*Crocus_CROC_C. pallasii* (CplDD09) _A4_matK_ (XF_5R).

CAATTCATTCAACTTTTCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTAGATATACTAATACCCCATCCCATT
CATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTCTTTTGCATTTCTTGCATTCTTTCTTAC
GAAATCATAAATGGAAATAGTTTTTAATTAATCTCAGAATAAATCTATTTCTCTTTTTTCAAAGAAAACATAAAGACTATT
TCGGTTCCTATACAATTTCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCATTCTTCTTATTTACGATT
CACATCCTTTCAAACCTTTTCTTGAGCGAAGATATTTCTATGGAAAAATGACATCTTCAAACGGAACATCTTATAATAGT
ATGTTATGATTATTTAATAGAACCCTATGGTCCTTCAAAGATCCTTTTATGCATTATGCTCGATGTCAAGGAAAAGCA
ATTTCTGGCTTCAAAGGAACATCTTCTGATGAAGAAAATGGAATATCATTTTGTCAATTTATGGCAATATTATTTTC
ACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAAACCATTCTTTCCATTTTCTGGGTTATCTTTCAAGT
TTACTAAAAAATTCTTCGACGGTAAGGAATCAAATGTTAGATAATTCATTCTAATAGATACTCTTACTACGAAATTTG
ATACCGCAGTCCCAGTTATTTTCTTATTAGATCTTTATCTAAAAGCTCAATTTTGTACTGTATCGGGACATCCTATTAGT
AAGCCCATCTGGACCGATTTATCGGATCTGGTATTATTGAGCGATTGGTCCGGATATCTAGAAAATCTTTCTCATTATC
ATAGTGGGTCTCAAACAAAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTCTG

>*Crocus_CROC_C. pallasii* (CplR09) A1_matK_ (XF_5R).

CAATTCATTCAACTTTTCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCCCATCCCA
TTCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTCTTTTGCATTTCTTGCATTCTTTCTT
CAGGAATATCATAAATGGAAATAGTTTTTAATTAATCTCAGAATAAATCTATTTCTCTTTTTTCAAAGAAAACATAAAGAC
TATTTCCGGTTCCTATACAATTTCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCATTCTTCTTATTTAC
GATTCACATCCTTTCAAACCTTTTCTTGAGCGAAGATATTTCTATGGAAAAATGGAACATCTTCAAACGGAACATCTTAT
AATAGTATGTTATGATTATTTAATAGAACCCTATGGTCCTTCAAAGATCCTTTTATGCATTATGCTCGATGTCAAGGA
AAAGCAATTCTGGCTTCAAAGGAACATCTTCTGATGAAGAAAATGGAATATCATTTTGTCAATTTATGGCAATAT
TATTTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAAACCATTCTTTCCATTTTCTGGGTTATCT
TTCAAGTTTACTAAAAAATTCTTCGACGGTAAGGAATCAAATGTTAGATAATTCATTCTAATAGATACTCTTACTACG
AAATTTGATACCGCAGTCCCAGTTATTTTCTTATTAGATCTTTATCTAAAAGCTCAATTTTGTACTGTATCGGGACATCC
TATTAGTAAGCCCATCTGGACCGATTTATCGGATCTGGTATTATTGAGCGATTGGTCCGGATATCTAGAAAATCTTTCT
CATTATCATAGTGGGTCTCAAACAAAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTCTG

> *Crocus _CROC_C pallasii* (CplVD09) A2_matK_ (XF_5R).

CAATTCATTCAACTTTTCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCCCATCCC
ATTCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTTTTTGCAATTTCTTGGCATTCTTTCT
TCACGAATATCATAAATTGGAATAGTTTTTAAATTAATCTCAGAATAAATCTATTTCTCTTTTTTCAAAGAAACTAAAAAGA
CTATTTTGGTTCTATACAATTTCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCATTCTTCTTATTTA
CGATTCACATCCTTTCAAACCTTTCTTGAGCGAAGATATTTCTATGGAAAAATGGAACATCTTCAAACGGAACATCTTA
TAATAGTATGTTATGATTATTTAATAGAACCCTATGGTCTTCAAAGATCCTTTTATGCATTATGCTCGATGTCAAGG
AAAAGCAATTCTGGCTTCAAAGGAACATCTTCTGATGAAGAAATGGAATATAATTTTGTCAATTTATGGCAATA
TTATTTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAAACCATTCTTCCATTTTCTGGGTTATC
TTTCAAGTTTACTAAAAAATCTTTCGACGGTAAGGAATCAAATGTTAGATAATTCATTTCTAATAGATACTCTTACTAC
GAAATTTGATACCGCAGTCCCAGTTATTTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACTGTATCGGGACATC
CTATTAGTAAGCCCATCTGGACCGATTTATCAGATTCTGGTATTATTGAGCGATTGGTCCGGATATCTAGAAATCTTTC
TCATTATCATAGTGGTCTCAAAAAAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTTCTG

> *Crocus _CROC_C. Kashmirianus'* Hort. Lilac (CstCD09) A5_matK_ (XF_5R).

CAATTCATTCAACTTTTCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCCCATCCC
ATTCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTTTTTGCAATTTCTTGGCATTCTTTCT
TCACGAATATCATAAATTGGAATAGTTTTTAAATTAATCTCAGAATAAATCTATTTCTCTTTTTTCAAAGAAACTAAAAAGA
CTATTTTGGTTCTATACAATTTCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCATTCTTCTTATTTA
CGATTCACATCCTTTCAAACCTTTCTTGAGCGAAGATATTTCTATGGAAAAATGGAACATCTTCAAACGGAACATCTTA
TAATAGTATGTTATGATTATTTAATAGAACCCTATGGTCTTCAAAGATCCTTTTATGCATTATGCTCGATGTCAAGG
AAAAGCAATTCTGGCTTCAAAGGAACATCTTCTGATGAAGAAATGGAATATAATTTTGTCAATTTATGGCAATA
TTATTTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAAACCATTCTTCCATTTTCTGGGTTATC
TTTCAAGTTTACTAAAAAATCTTTCGACGGTAAGGAATCAAATGTTAGATAATTCATTTCTAATAGATACTCTTACTAC
GAAATTTGATACCGCAGTCCCAGTTATTTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACTGTATCGGGACATC
CTATTAGTAAGCCCATCTGGACCGATTTATCAGATTCTGGTATTATTGAGCGATTGGTCCGGATATCTAGAAATCTTTC
TCATTATCATAGTGGTCTCAAAAAAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTTCTG

> *Crocus _CROC_C. hadriaticus* (ChdWD08) A11_matK_ (XF_5R).

CAATTCATTCAACTTTTCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCCCATCCC
ATTCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTTTTTGCAATTTCTTACGATTCTTTCT
TCACGAATATCATAAATTGGAATAGTTTTTAAATTAATCTCAGAATAAATCTATTTCTCTTTTTTCAAAGAAACTAAAAAGA
CTATTTTGGTTCTATACAATTTCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCATTCTTCTTATTTA
CGATTCACATCCTTTCAAACCTTTCTTGAGCGAAGATATTTCTATGGAAAAATGGAACATCTTCAAACGGAACATCTTA
TAATAGTATGTTATGATTATTTAATAGAACCCTATGGTCTTCAAAGATCCTTTTATGCATTATGCTCGATGTCAAGG
AAAAGCAATTCTGGCTTCAAAGGAACATCTTCTGATGAAGAAATGGAATATAATTTTGTCAATTTATGGCAATA
TTATTTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAAACCATTCTTCCATTTTCTGGGTTATC
TTTCAAGTTTACTAAAAAATCTTTCGACGGTAAGGAATCAAATGTTAGATAATTCATTTCTAATAGATACTCTTACTAC
GAAATTTGATACCGCAGTCCCAGTTATTTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACTGTATCGGGACATCC
TATTAGTAAGCCCATCTGGACCGATTTATCAGATTCTGGTATTATTGAGCGATTGGTCCGGATATCTAGAAATCTTTC
CATTATCATAGTGGTCTCAAAAAAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTTCTG

> *Crocus _CROC_C. oreoreticus* (CorVR09) A10_matK_ (XF_5R).

CAATTCATTCAACTTTTCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCCCATCCC
ATTCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTTTTTGCAATTTCTTGGCATTCTTTCT
TCACGAATATCATAAATTGGAATAGTTTTTAAATTAATCTCAGAATAAATCTATTTCTCTTTTTTCAAAGAAACTAAAAAGA
CTATTTTGGTTCTATACAATTTCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCATTCTTCTTATTTA
CGATTCACATCCTTTCAAACCTTTCTTGAGCGAAGATATTTCTATGGAAAAATGGAACATCTTCAAACGGAACATCTTA
TAATAGTATGTTATGATTATTTAATAGAACCCTATGGTCTTCAAAGATCCTTTTATGCATTATGCTCGATGTCAAGG
AAAAGCAATTCTGGCTTCAAAGGAACATCTTCTGATGAAGAAATGGAATATAATTTTGTCAATTTATGGCAATA
TTATTTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAAACCATTCTTCCATTTTCTGGGTTATC
TTTCAAGTTTACTAAAAAATCTTTCGACGGTAAGGAATCAAATGTTAGATAATTCATTTCTAATAGATACTCTTACTAC
GAAATTTGATACCGCAGTCCCAGTTATTTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACTGTATCGGGACATC
CTATTAGTAAGCCCATCTGGACCGATTTATCAGATTCTGGTATTATTGAGCGATTGGTCCGGATATCTAGAAATCTTTC
TCATTATCATAGTGGTCTCAAAAAAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTTCTG

> *Crocus _CROC_C. sativus_Spain (C.stPER09) A6_matK_ (XF_5R).*

CAATTCATTCAACTTTTCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCCCATCCC
ATTCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTTTTTGCATTTCTTGGCATTCTTTCT
TCACGAATATCATAATTGGAATAGTTTTTAATTAATCTCAGAAATAAATCTATTCTCTTTTTTCAAAGAAACTAAAAGA
CTATTTTGGTTCTATAACAATCTTATGTATATGAATGTGAATTTGTATTTCGTTTTCTTCGTAAGCATTCTTCTTATTTA
CGATTCACATCCTTTCAAACCTTTCTTGAACGAAGATATTCTATGGAAAAATGGAACATCTTCAAACGGAACATCTTA
TAATAGTATGTTATGATTATTTAATAGAACCCTATGGTCTTCAAAGATCCTTTTATGCATTATGCTCGATGTCAAGG
AAAAGCAATTCTGGCTTCAAAGGAACATCTTCTGATGAAGAAATGGAATATAATTTGTCAATTTATGGCAATA
TTATTTTCACTTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAAACCATCTTCCATTTCTGGGTTATC
TTTCAAGTTTACTAAAAAATCTTTCGACGGTAAGGAATCAAATGTTAGATAATTCATTTCTAATAGATACTCTTACTAC
GAAATTTGATACCGCAGTCCCAGTTATTTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACTGTATCGGGACATC
CTATTAGTAAGCCATCTGGACCGATTTATCAGATTCTGGTATTATTGAGCGATTGGTCCGATATCTAGAAATCTTTC
TCATTATCATAGTGGTCTCAAACAAAAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTTCTG

> *Crocus _CROC_C. thomasi (CtmVD09) A12_matK_ (XF_5R).*

CAATTCATTCAACTTTTCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCCCATCCC
ATTCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTTTTTGCATTTCTTGGCAATTTCTTCTT
CACGAATATCATAATTGGAATAGTTTTTAATTAATCTCAGAAATAAATCTATTCTCTTTTTTCAAAGAAACTAAAAGAC
TATTTTGGTTCTATAACAATCTTATGTATATGAATGTGAATTTGTATTTCGTTTTCTTCGTAAGCATTCTTCTTATTTAC
GATTCACATCCTTTCAAACCTTTCTTGAACGAAGATATTCTATGGAAAAATGGAACATCTTCAAACGGAACATCTTAT
AATAGTATGTTATGATTATTTAATAGAACCCTATGGTCTTCAAAGATCCTTTTATGCATTATGCTCGATGTCAAGGA
AAAGCAATTCTGGCTTCAAAGGAACATCTTCTGATGAAGAAATGGAATATAATTTTGTCAATTTATGGCAATAT
TATTTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAAACCATCTTCCATTTTCTGGGTTATCT
TTCAAGTTTACTAAAAAATCTTTCGACGGTAAGGAATCAAATGTTAGATAATTCATTTCTAATAGATACTCTTACTACG
AAATTTGATACCGCAGTCCCAGTTATTTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACTGTATCGGGACATCC
TATTAGTAAGCCATCTGGACCGATTTATCAGATTCTGGTATTATTGAGCGATTGGTCCGATATCTAGAAATCTTTC
CATTATCATAGTGGTCTCAAACAAAAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTTCTG

> *Crocus _CROC_C. cartwrightianus (CcrCR09) A8_matK_ (XF_5R).*

CAATTCATTCAACTTTTCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCCCATCCC
ATTCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTTTTTGCATTTCTTGGCATTCTTTCT
TCACGAATATCATAATTGGAATAGTTTTTAATTAATCTCAGAAATAAATCTATTCTCTTTTTTCAAAGAAACTAAAAGA
CTATTTTGGTTCTATAACAATCTTATGTATATGAATGTGAATTTGTATTTCGTTTTCTTCGTAAGCATTCTTCTTATTTA
CGATTCACATCCTTTCAAACCTTTCTTGAACGAAGATATTCTATGGAAAAATGGAACATCTTCAAACGGAACATCTTA
TAATAGTATGTTATGATTATTTAATAGAACCCTATGGTCTTCAAAGATCCTTTTATGCATTATGCTCGATGTCAAGG
AAAAGCAATTCTGGCTTCAAAGGAACATCTTCTGATGAAGAAATGGAATATAATTTTGTCAATTTATGGCAATA
TTATTTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAAACCATCTTCCATTTTCTGGGTTATC
TTTCAAGTTTACTAAAAAATCTTTCGACGGTAAGGAATCAAATGTTAGATAATTCATTTCTAATAGATACTCTTACTAC
GAAATTTGATACCGCAGTCCCAGTTATTTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACTGTATCGGGACATC
CTATTAGTAAGCCATCTGGACCGATTTATCAGATTCTGGTATTATTGAGCGATTGGTCCGATATCTAGAAATCTTTC
TCATTATCATAGTGGTCTCAAACAAAAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTTCTG

> *Crocus _CROC_C.sativus cartwrightianus_'Albus'(CstcP09) A15_matK_ (XF_5R).*

CAATTCATTCAACTTTTCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCCCATCCC
ATTCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTTTTTGCATTTCTTGGCATTCTTTCT
TCACGAATATCATAATTGGAATAGTTTTTAATTAATCTCAGAAATAAATCTATTCTCTTTTTTCAAAGAAACTAAAAGA
CTATTTTGGTTCTATAACAATCTTATGTATATGAATGTGAATTTGTATTTCGTTTTCTTCGTAAGCATTCTTCTTATTTA
CGATTCACATCCTTTCAAACCTTTCTTGAACGAAGATATTCTATGGAAAAATGGAACATCTTCAAACGGAACATCTTA
TAATAGTATGTTATGATTATTTAATAGAACCCTATGGTCTTCAAAGATCCTTTTATGCATTATGCTCGATGTCAAGG
AAAAGCAATTCTGGCTTCAAAGGAACATCTTCTGATGAAGAAATGGAATATAATTTTGTCAATTTATGGCAATA
TTATTTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAAACCATCTTCCATTTTCTGGGTTATC
TTTCAAGTTTACTAAAAAATCTTTCGACGGTAAGGAATCAAATGTTAGATAATTCATTTCTAATAGATACTCTTACTAC
GAAATTTGATACCGCAGTCCCAGTTATTTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACTGTATCGGGACATC
CTATTAGTAAGCCATCTGGACCGATTTATCAGATTCTGGTATTATTGAGCGATTGGTCCGATATCTAGAAATCTTTC
CATTATCATAGTGGTCTCAAACAAAAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTTCTG

> *Crocus_CROC_C. cartwrightianus* (CcwAD08) A7_matK_ (XF_5R).

CAATTCATTCAACTTTTCCTTTTTAGAGGACAAATTATTACATTTCCATTATGTATCAGATATACTAATACCCCATCCC
ATTCATATGGAAATATTGGTTCAAATCCTTCAATGCTGGATTCAAGATGTTCCCTCTTTTGCATTTCTTGGCATTCTTTCT
TCACGAATATCATAATTGGAATAGTTTTTAATTACTCAGAATAAATCTATTCTCTTTTTTCAAAGAAACTAAAAGA
CTATTTTGGTTCTATACAATTCCTTATGTATATGAATGTGAATTTGTATTTCGTTTTTCTTCGTAAGCATTCTTCTTATTTA
CGATTCACATCCTTTCAAACCTTTCTTGAACGAAGATATTCTATGGAAAAATGGAACATCTTCAAACGGAACATCTTA
TAATAGTATGTTATGATTATTTAATAGAACCCTATGGTCCTTCAAAGATCCTTTTATGCATTATGCTCGATGTCAAGG
AAAAGCAATTCTGGCTTCAAAGGAACATCTTCTGATGAAGAAATGGAAATATAATTTGTCAATTTATGGCAATA
TTATTTTCACTTTTGGTATCAATCATAACAGGATCCATATAAACCAACTATCAAACCATTCTTCCATTTTCTGGGTTATC
TTTCAAGTTTACTAAAAAATCTTCGACGGTAAGGAATCAAATGTTAGATAATTCATTTCTAATAGATACTCTTACTAC
GAAATTTGATACCGCAGTCCCAGTTATTTTCTTATTAGATCTTTATCTAAAGCTCAATTTTGTACTGTATCGGGACATC
CTATTAGTAAGCCCATCTGGACCGATTTATCAGATTCTGGTATTATTGAGCGATTGGTCCGATATCTAGAAATCTTTC
TCATTATCATAGTGGTCTCAAAAAACAGGGGTTGTATCGAATAAAGTATATACTTCGACTTTCTT

3: Consensus *rbcl* plastid gene sequences of *Crocus* series *Crocus*, obtained by aligning both forward and reverse sequences for each species.

>JX903213.1_ *Crocus banaticus* voucher_D.K. Kim_09-004_ribulose-1,5-bisphosphate_carboxylase_large_subunit_(*rbcl*)_-----

GGATTTAAAGCTGGTGTAAAGATTACAGATTGACTTATTATACTCCGGAGTACGAAACCAAAGATACTGATATCTTG
GCAGCATTCCGAGTAACTCCTCAACCCGGAGTTCCCTGCTGAAGAGGCGGGGGCAGCGGTAGCTGCGGAATCTTCTACT
GGTACATGGACAACAGTGTGGACTGATGGACTTACCAGTCTTGATCGTTACAAAGGACGATGCTATCATATCGAGGCC
GTTGTTGGGGAGGAAAAATCAATATATTGCTTATGTAGCTTATCCTTTAGACCTTTTTGAAGAAGGTTCTGTTACTAATA
TGTTTACTTCCATTGTGGTAATGTATTTGGTTTCAAAGCCCTACGAGCTCTACGTCTGGAGGATTTGCGAATCCCCC
TGCTTATCCAAAACCTTTCCAAGGCCCGCCCCATGGCATCCAGGTTGAAAAGAGATAAAATGAAACAAGTATGGTCGTCC
CCTATTGGGATGTACTATTAACAAAATTTGGGATTATCTGCAAAAAACTACGGTAGAGCGGTTTATGAATGTCTACG
TGGTGGGCTTGATTTTACCAAGGATGATGAAAACGTGAACTCACAACTTTTATGCGTTGGAGAGACCGTTTCTTATTT
TGTGCTGAAGCAATTTATAAAGCGCAGGCCGAAACAGGTGAAATCAAAGGACATTACTTGAATGCAACTGCAG

> *Crocus* _CROC_ *C. cartwrightianus* (CwBD10)9B_ *rbcl*__ (1F_724R).

CCCCAACAGAACTAAAGCAAGTGTGGATTAAAGCTGGTGTAAAGATTACAGATTGACTTATTATACTCCGGAGT
ACGAAACCAAAGATACTGATATCTTGGCAGCATTCCGAGTAACTCCTCAACCCGGAGTTCCCTGCTGAAGAAGCGGGG
GCAGCGGTAGCTGCGGAATCTTCTACTGGTACATGGACAACAGTGTGGACTGATGGACTTACCAGTCTTGATCGTTAC
AAAGGACGATGCTATCATATCGAGGCCGTTGTTGGGGAGGAAAAATCAATATATTGCTTATGTAGCTTATCCTTTAGAC
CTTTTTGAAGAAGGTTCTGTTACTAATATGTTTACTTCCATTGTGGGTAATGTATTTGGTTTCAAAGCCCTACGAGCTCT
ACGTCTGGAGGATTTGCGAATTTCCCCTGCTTATTCCAAAACCTTTCAAAGGCCACCCCATGGCATCCAGGTTGAAAG
AGATAAATTAACAAGTATGGTCGTCCCTATTGGGATGTACTATTAACCAAATTTGGGATTATCTGCAAAAAACTA
CGGTAGAGCGGTTTATGAATGTCTACGCGGTGGGCTTGATTTTACCAAGGATGATGAAAATGTGAACTCACAACTTT
TATGCGTTGGAGAGACCGTTTCTTATTTTGTGCTGAAGCACTTTATAAAGCGCAGGCCGAAACAGGTGAAATCAAAGG
ACATTACTTGAATGCTACTGCAG

> *Crocus* _CROC_ *C. asumaniae* (CasWD09)13B_ *rbcl*__ (1F_724R).

CCCCAACAGAACTAAAGCAAGTGTGGATTAAAGCTGGTGTAAAGATTACAGATTGACTTATTATACTCCGGAGT
ACGAAACCAAAGATACTGATATCTTGGCAGCATTCCGAGTAACTCCTCAACCCGGAGTTCCCTGCTGAAGAAGCGGGG
GCAGCGGTAGCTGCGGAATCTTCTACTGGTACATGGACAACAGTGTGGACTGATGGACTTACCAGTCTTGATCGTTAC
AAAGGACGATGCTATCATATCGAGGCCGTTGTTGGGGAGGAAAAATCAATATATTGCTTATGTAGCTTATCCTTTAGAC
CTTTTTGAAGAAGGTTCTGTTACTAATATGTTTACTTCCATTGTGGGTAATGTATTTGGTTTCAAAGCCCTACGAGCTCT
ACGTCTGGAGGATTTGCGAATTTCCCCTGCTTATTCCAAAACCTTTCAAAGGCCCGCCCCATGGCATCCAGGTTGAAAG
AGATAAATTAACAAGTATGGTCGTCCCTATTGGGATGTACTATTAACCAAATTTGGGATTATCTGCAAAAAACTA
CGGTAGAGCGGTTTATGAATGTCTACGCGGTGGGCTTGATTTTACCAAGGATGATGAAAATGTGAACTCACAACTTT
TATGCGTTGGAGAGACCGTTTCTTATTTTGTGCTGAAGCACTTTATAAAGCGCAGGCCGAAACAGGTGAAATCAAAGG
ACATTACTTGAATGCTACTGCAG

> *Crocus* _CROC_ *C. mathweii* (CmtHR09)14B_ *rbcl*__ (1F_724R).

CCCCACAGAACTAAAGCAAGTGTGGATTAAAGCTGGTGTAAAGATTACAGATTGACTTATTATACTCCGGAGT
ACGAAACCAAAGATACTGATATCTTGGCAGCATTCCGAGTAACTCCTCAACCCGGAGTTCCCTGCTGAAGAAGCGGGG
GCAGCGGTAGCTGCGGAATCTTCTACTGGTACATGGACAACAGTGTGGACTGATGGACTTACCAGTCTTGATCGTTAC
AAAGGACGATGCTATCATATCGAGGCCGTTGTTGGGGAGGAAAAATCAATATATTGCTTATGTAGCTTATCCTTTAGAC
CTTTTTGAAGAAGGTTCTGTTACTAATATGTTTACTTCCATTGTGGGTAATGTATTTGGTTTCAAAGCCCTACGAGCTCT
ACGTCTGGAGGATTTGCGAATTTCCCCTGCTTATTCCAAAACCTTTCAAAGGCCACCCCATGGCATCCAGGTTGAAAG
AGATAAATTAACAAGTATGGTCGTCCCTATTGGGATGTACTATTAACCAAATTTGGGATTATCTGCAAAAAACTA
CGGTAGAGCGGTTTATGAATGTCTACGCGGTGGGCTTGATTTTACCAAGGATGATGAAAATGTGAACTCACAACTTT
TATGCGTTGGAGAGACCGTTTCTTATTTTGTGCTGAAGCACTTTATAAAGCGCAGGCCGAAACAGGTGAAATCAAAGG
ACATTACTTGAATGCTACTGCAG

> *Crocus* _CROC_ *C. pallasii* (CplDD09)4B_ *rbcl*__ (1F_724R).

CCCCACAGAACTAAAGCAAGTGTGGATTAAAGCTGGTGTAAAGATTACAGATTGACTTATTATACTCCGGAGT
ACGAAACCAAAGATACTGATATCTTGGCAGCATTCCGAGTAACTCCTCAACCCGGAGTTCCCTGCTGAAGAAGCGGGG
GCAGCGGTAGCTGCGGAATCTTCTACTGGTACATGGACAACAGTGTGGACTGATGGACTTACCAGTCTTGATCGTTAC
AAAGGACGATGCTATCATATCGAGGCCGTTGTTGGGGAGGAAAAATCAATATATTGCTTATGTAGCTTATCCTTTAGAC
CTTTTTGAAGAAGGTTCTGTTACTAATATGTTTACTTCCATTGTGGGTAATGTATTTGGTTTCAAAGCCCTACGAGCTCT
ACGTCTGGAGGATTTGCGAATTTCCCCTGCTTATTCCAAAACCTTTCAAAGGCCACCCCATGGCATCCAGGTTGAAAG
AGATAAATTAACAAGTATGGTCGTCCCTATTGGGATGTACTATTAACCAAATTTGGGATTATCTGCAAAAAACTA
CGGTAGAGCGGTTTATGAATGTCTACGCGGTGGGCTTGATTTTACCAAGGATGATGAAAATGTGAACTCACAACTTT
TATGCGTTGGAGAGACCGTTTCTTATTTTGTGCTGAAGCACTTTATAAAGCGCAGGCCGAAACAGGTGAAATCAAAGG
ACATTACTTGAATGCTACTGCAG

> *Crocus _CROC_C. pallasii (CpltR09)3B_rbc1_ (1F_724R).*

CCCCACAGAAACTAAAGCAAGTGCTGGATTAAAGGCTGGTGTAAAGATTACAGATTGACTTATTATACTCCGGAGT
ACGAAACCAAAGATACTGATATCTTGCCAGCATTCCGAGTAACTCCTCAACCCGGAGTTCCTGCTGAAGAAGCGGGG
GCAGCGGTAGCTGCGGAATCTTCTACTGGTACATGGACAACAGTGTGGACTGATGGACTTACCAGTCTTGATCGTTAC
AAAGGACGATGCTATCATATCGAGGCCGTTGTTGGGGAGGAAAAATCAATATATTGCTTATGTAGCTTATCCTTTAGAC
CTTTTTGAAGAAGGTTCTGTACTAATATGTTTACTTCCATTGTGGGTAATGTATTTGGTTTCAAAGCCCTACGAGCTCT
ACGTCTGGAGGATTTGCGAATTCCTGCTTATTCCAAAACCTTCCAAGGCCACCCCATGGCATCCAGGTTGAAAG
AGATAAATTAACAAGTATGGTCTGCCCTATTGGGATGTACTATTAACCAAAAATTGGGATTATCTGCAAAAAACTA
CGGTAGAGCGGTTTATGAATGTCTACGCGGTGGCTTGATTTTACCAAGGATGATGAAAATGTGAACTACAACCTTT
TATGCGTTGGAGAGACCGTTTCTTATTTTGTGCTGAAGCACTTTATAAAGCGCAGGCCGAAACAGGTGAAATCAAAGG
ACATTACTTGAATGCTACTGCAG

> *Crocus _CROC_C. pallasii (CplR09)1B_rbc1_ (1F_724R)*

CCCCAACAGAAACTAAAGCAAGTGCTGGATTAAAGGCTGGTGTAAAGATTACAGATTGACTTATTATACTCCGGAGT
ACGAAACCAAAGATACTGATATCTTGCCAGCATTCCGAGTAACTCCTCAACCCGGAGTTCCTGCTGAAGAAGCGGGG
GCAGCGGTAGCTGCGGAATCTTCTACTGGTACATGGACAACAGTGTGGACTGATGGACTTACCAGTCTTGATCGTTAC
AAAGGACGATGCTATCATATCGAGGCCGTTGTTGGGGAGGAAAAATCAATATATTGCTTATGTAGCTTATCCTTTAGAC
CTTTTTGAAGAAGGTTCTGTACTAATATGTTTACTTCCATTGTGGGTAATGTATTTGGTTTCAAAGCCCTACGAGCTCT
ACGTCTGGAGGATTTGCGAATTCCTGCTTATTCCAAAACCTTCCAAGGCCACCCCATGGCATCCAGGTTGAAAG
AGATAAATTAACAAGTATGGTCTGCCCTATTGGGATGTACTATTAACCAAAAATTGGGATTATCTGCAAAAAACTA
CGGTAGAGCGGTTTATGAATGTCTACGCGGTGGCTTGATTTTACCAAGGATGATGAAAATGTGAACTACAACCTTT
TATGCGTTGGAGAGACCGTTTCTTATTTTGTGCTGAAGCACTTTATAAAGCGCAGGCCGAAACAGGTGAAATCAAAGG
ACATTACTTGAATGCTACTGCAG

> *Crocus _CROC_C. Kashmirianus (CstCD09)5B_rbc1_ (1F_724R).*

CCCCAACAGAAACTAAAGCAAGTGCTGGATTAAAGGCTGGTGTAAAGATTACAGATTGACTTATTATACTCCGGAGT
ACGAAACCAAAGATACTGATATCTTGCCAGCATTCCGAGTAACTCCTCAACCCGGAGTTCCTGCTGAAGAAGCGGGG
GCAGCGGTAGCTGCGGAATCTTCTACTGGTACATGGACAACAGTGTGGACTGATGGACTTACCAGTCTTGATCGTTAC
AAAGGACGATGCTATCATATCGAGGCCGTTGTTGGGGAGGAAAAATCAATATATTGCTTATGTAGCTTATCCTTTAGAC
CTTTTTGAAGAAGGTTCTGTACTAATATGTTTACTTCCATTGTGGGTAATGTATTTGGTTTCAAAGCCCTACGAGCTCT
ACGTCTGGAGGATTTGCGAATTCCTGCTTATTCCAAAACCTTCCAAGGCCACCCCATGGCATCCAGGTTGAAAG
AGATAAATTAACAAGTATGGTCTGCCCTATTGGGATGTACTATTAACCAAAAATTGGGATTATCTGCAAAAAACTA
CGGTAGAGCGGTTTATGAATGTCTACGCGGTGGCTTGATTTTACCAAGGATGATGAAAATGTGAACTACAACCTTT
TATGCGTTGGAGAGACCGTTTCTTATTTTGTGCTGAAGCACTTTATAAAGCGCAGGCCGAAACAGGTGAAATCAAAGG
ACATTACTTGAATGCTACTGCAG

> *Crocus _CROC_C. hadriaticus (ChdWD08)11B_rbc1_ (1F_724R).*

CCCCACAGAAACTAAAGCAAGTGCTGGATTAAAGGCTGGTGTAAAGATTACAGATTGACTTATTATACTCCGGAGT
ACGAAACCAAAGATACTGATATCTTGCCAGCATTCCGAGTAACTCCTCAACCCGGAGTTCCTGCTGAAGAAGCGGGG
GCAGCGGTAGCTGCGGAATCTTCTACTGGTACATGGACAACAGTGTGGACTGATGGACTTACCAGTCTTGATCGTTAC
AAAGGACGATGCTATCATATCGAGGCCGTTGTTGGGGAGGAAAAATCAATATATTGCTTATGTAGCTTATCCTTTAGAC
CTTTTTGAAGAAGGTTCTGTACTAATATGTTTACTTCCATTGTGGGTAATGTATTTGGTTTCAAAGCCCTACGAGCTCT
ACGTCTGGAGGATTTGCGAATTCCTGCTTATTCCAAAACCTTCCAAGGCCACCCCATGGCATCCAGGTTGAAAG
AGATAAATTAACAAGTATGGTCTGCCCTATTGGGATGTACTATTAACCAAAAATTGGGATTATCTGCAAAAAACTA
CGGTAGAGCGGTTTATGAATGTCTACGCGGTGGCTTGATTTTACCAAGGATGATGAAAATGTGAACTACAACCTTT
TATGCGTTGGAGAGACCGTTTCTTATTTTGTGCTGAAGCACTTTATAAAGCGCAGGCCGAAACAGGTGAAATCAAAGG
ACATTACTTGAATGCTACTGCAG

> *Crocus _CROC_C. oreoreticus (CorVR09)10B_rbc1_ (1F_724R).*

CCCCACAGAAACTAAAGCAAGTGCTGGATTAAAGGCTGGTGTAAAGATTACAGATTGACTTATTATACTCCGGAGT
ACGAAACCAAAGATACTGATATCTTGCCAGCATTCCGAGTAACTCCTCAACCCGGAGTTCCTGCTGAAGAAGCGGGG
GCAGCGGTAGCTGCGGAATCTTCTACTGGTACATGGACAACAGTGTGGACTGATGGACTTACCAGTCTTGATCGTTAC
AAAGGACGATGCTATCATATCGAGGCCGTTGTTGGGGAGGAAAAATCAATATATTGCTTATGTAGCTTATCCTTTAGAC
CTTTTTGAAGAAGGTTCTGTACTAATATGTTTACTTCCATTGTGGGTAATGTATTTGGTTTCAAAGCCCTACGAGCTCT
ACGTCTGGAGGATTTGCGAATTCCTGCTTATTCCAAAACCTTCCAAGGCCACCCCATGGCATCCAGGTTGAAAG
AGATAAATTAACAAGTATGGTCTGCCCTATTGGGATGTACTATTAACCAAAAATTGGGATTATCTGCAAAAAACTA
CGGTAGAGCGGTTTATGAATGTCTACGCGGTGGCTTGATTTTACCAAGGATGATGAAAATGTGAACTACAACCTTT
TATGCGTTGGAGAGACCGTTTCTTATTTTGTGCTGAAGCACTTTATAAAGCGCAGGCCGAAACAGGTGAAATCAAAGG
ACATTACTTGAATGCTACTGCAG

> *Crocus _CROC_ C. pallasii (CplVD09)2B_rbc1__ (1F_724R).*

CCCCACAGAAACTAAAGCAAGTGCTGGATTAAAGGCTGGTGTAAAGATTACAGATTGACTTATTATACTCCGGAGT
ACGAAACCAAAGATACTGATATCTTGGCAGCATTCCGAGTAACTCCTCAACCCGGAGTTCCTGCTGAAGAAGCGGGG
GCAGCGGTAGCTGCGGAATCTTCTACTGGTACATGGACAACAGTGTGGACTGATGGACTTACCAGTCTTGATCGTTAC
AAAGGACGATGCTATCATATCGAGGCCGTTGTTGGGGAGGAAAAATCAATATATTGCTTATGTAGCTTATCCTTTAGAC
CTTTTGAAGAAGGTTCTGTTACTAATATGTTTACTTCCATTGTGGGTAATGTATTTGGTTTCAAAGCCCTACGAGCTCT
ACGTCTGGAGGATTTGCGAATTCCCCTGCTTATTCCAAAACTTTCCAAGGCCACCCCATGGCATCCAGGTTGAAAG
AGATAAATTAACAAGTATGGTCTGCCCTATTGGGATGTACTATTAACCAAAAATTGGGATTATCTGCAAAAAACTA
CGGTAGAGCGGTTTATGAATGTCTACGTGGTGGGCTTGATTTTACCAAGGATGATGAAAACGTGAACTCACAACCTTT
TATGCGTTGGAGAGACCGTTTCTTATTTTGTGCTGAAGCACTTTATAAAGCGCAGGCCGAAACAGGTGAAATCAAAGG
ACATTACTTGAATGCTACTGCAG

> *Crocus _CROC_ C. sativus cartwrightianus 'Albus' (CstcP09)15B_rbc1__ (1F_724R).*

CCCCACAGAAACTAAAGCAAGTGCTGGATTAAAGGCTGGTGTAAAGATTACAGATTGACTTATTATACTCCGGAGT
ACGAAACCAAAGATACTGATATCTTGGCAGCATTCCGAGTAACTCCTCAACCCGGAGTTCCTGCTGAAGAAGCGGGG
GCAGCGGTAGCTGCGGAATCTTCTACTGGTACATGGACAACAGTGTGGACTGATGGACTTACCAGTCTTGATCGTTAC
AAAGGACGATGCTATCATATCGAGGCCGTTGTTGGGGAGGAAAAATCAATATATTGCTTATGTAGCTTATCCTTTAGAC
CTTTTGAAGAAGGTTCTGTTACTAATATGTTTACTTCCATTGTGGGTAATGTATTTGGTTTCAAAGCCCTACGAGCTCT
ACGTCTGGAGGATTTGCGAATTCCCCTGCTTATTCCAAAACTTTCCAAGGCCACCCCATGGCATCCAGGTTGAAAG
AGATAAATTAACAAGTATGGTCTGCCCTATTGGGATGTACTATTAACCAAAAATTGGGATTATCTGCAAAAAACTA
CGGTAGAGCGGTTTATGAATGTCTACGTGGTGGGCTTGATTTTACCAAGGATGATGAAAACGTGAACTCACAACCTTT
TATGCGTTGGAGAGACCGTTTCTTATTTTGTGCTGAAGCACTTTATAAAGCGCAGGCCGAAACAGGTGAAATCAAAGG
ACATTACTTGAATGCTACTGCAG

> *Crocus _CROC_ C. thomasi (CtmVD09)12B_rbc1__ (1F_724R).*

CCCCACAGAAACTAAAGCAAGTGCTGGATTAAAGGCTGGTGTAAAGATTACAGATTGACTTATTATACTCCGGAGT
ACGAAACCAAAGATACTGATATCTTGGCAGCATTCCGAGTAACTCCTCAACCCGGAGTTCCTGCTGAAGAAGCGGGG
GCAGCGGTAGCTGCGGAATCTTCTACTGGTACATGGACAACAGTGTGGACTGATGGACTTACCAGTCTTGATCGTTAC
AAAGGACGATGCTATCATATCGAGGCCGTTGTTGGGGAGGAAAAATCAATATATTGCTTATGTAGCTTATCCTTTAGAC
CTTTTGAAGAAGGTTCTGTTACTAATATGTTTACTTCCATTGTGGGTAATGTATTTGGTTTCAAAGCCCTACGAGCTCT
ACGTCTGGAGGATTTGCGAATTCCCCTGCTTATTCCAAAACTTTCCAAGGCCACCCCATGGCATCCAGGTTGAAAG
AGATAAATTAACAAGTATGGTCTGCCCTATTGGGATGTACTATTAACCAAAAATTGGGATTATCTGCAAAAAACTA
CGGTAGAGCGGTTTATGAATGTCTACGTGGTGGGCTTGATTTTACCAAGGATGATGAAAACGTGAACTCACAACCTTT
TATGCGTTGGAGAGACCGTTTCTTATTTTGTGCTGAAGCACTTTATAAAGCGCAGGCCGAAACAGGTGAAATCAAAGG
ACATTACTTGAATGCTACTGCAG

> *Crocus _CROC_ C. cartwrightianus (CcwAD08)7B_rbc1__ (1F_724R).*

CCCCACAGAAACTAAAGCAAGTGCTGGATTAAAGGCTGGTGTAAAGATTACAGATTGACTTATTATACTCCGGAGT
ACGAAACCAAAGATACTGATATCTTGGCAGCATTCCGAGTAACTCCTCAACCCGGAGTTCCTGCTGAAGAAGCGGGG
GCAGCGGTAGCTGCGGAATCTTCTACTGGTACATGGACAACAGTGTGGACTGATGGACTTACCAGTCTTGATCGTTAC
AAAGGACGATGCTATCATATCGAGGCCGTTGTTGGGGAGGAAAAATCAATATATTGCTTATGTAGCTTATCCTTTAGAC
CTTTTGAAGAAGGTTCTGTTACTAATATGTTTACTTCCATTGTGGGTAATGTATTTGGTTTCAAAGCCCTACGAGCTCT
ACGTCTGGAGGATTTGCGAATTCCCCTGCTTATTCCAAAACTTTCCAAGGCCACCCCATGGCATCCAGGTTGAAAG
AGATAAATTAACAAGTATGGTCTGCCCTATTGGGATGTACTATTAACCAAAAATTGGGATTATCTGCAAAAAACTA
CGGTAGAGCGGTTTATGAATGTCTACGTGGTGGGCTTGATTTTACCAAGGATGATGAAAACGTGAACTCACAACCTTT
TATGCGTTGGAGAGACCGTTTCTTATTTTGTGCTGAAGCACTTTATAAAGCGCAGGCCGAAACAGGTGAAATCAAAGG
ACATTACTTGAATGCTACTGCAG

> *Crocus _CROC_ C. sativus Spain (C.stPER09) A6_rbc1__ (1F_724R).*

CCCCAACAGAAACTAAAGCAAGTGCTGGATTAAAGGCTGGTGTAAAGATTACAGATTGACTTATTATACTCCGGAGT
ACGAAACCAAAGATACTGATATCTTGGCAGCATTCCGAGTAACTCCTCAACCCGGAGTTCCTGCTGAAGAAGCGGGG
GCAGCGGTAGCTGCGGAATCTTCTACTGGTACATGGACAACAGTGTGGACTGATGGACTTACCAGTCTTGATCGTTACA
AAGGACGATGCTATCATATCGAGGCCGTTGTTGGGGAGGAAAAATCAATATATTGCTTATGTAGCTTATCCTTTAGACC
TTTTGAAGAAGGTTCTGTTACTAATATGTTTACTTCCATTGTGGGTAATGTATTTGGTTTCAAAGCCCTACGAGCTCTA
CGTCTGGAGGATTTGCGAATTCCCCTGCTTATTCCAAAACTTTCCAAGGCCACCCCATGGCATCCAGGTTGAAAGA
GATAAATTAACAAGTATGGTCTGCCCTATTGGGATGTACTATTAACCAAAAATTGGGATTATCTGCAAAAAACTAC
GGTAGAGCGGTTTATGAATGTCTACGTGGTGGGCTTGATTTTACCAAGGATGATGAAAACGTGAACTCACAACCTTT
ATGCGTTGGAGAGACCGTTTCTTATTTTGTGCTGAAGCACTTTATAAAGCGCAGGCCGAAACAGGTGAAATCAAAGGA
CATTACTTGAATGCTACTGCAG

> *Crocus_CROC_C. cartwrightianus* (CcrCR09)8B_rbcl_ (1F_724R).

CCCAAAACAGAACTAAAGCAAGTGTGGATTAAAGGCTGGTGTAAAAGATTACAGATTGACTTATTATACTCCGGAGT
ACGAAACCAAAGATACTGATATCTTGGCAGCATTCCGAGTAACTCCTCAACCCGGAGTTCCTGCTGAAGAAGCGGGG
GCAGCGGTAGCTGCGGAATCTTCTACTGGTACATGGACAAACAGTGTGGACTGATGGACTTACCAGTCTTGATCGTTAC
AAAGGACGATGCTATCATATCGAGGCCGTTGTTGGGGAGGAAAATCAATATATTGCTTATGTAGCTTATCCTTTAGAC
CTTTTGAAGAAGGTTCTGTACTAATATGTTTACTTCCATTGTGGGTAATGTATTTGGTTTCAAAGCCCTACGAGCTCT
ACGTCTGGAGGATTTGCGAATCCCCCTGCTTATTCCAAAACTTCCAAGGCCACCCCATGGCATCCAGGTTGAAAG
AGATAAATTAACAAGTATGGTCGTCCCTATTGGGATGTACTATTAACCAAAAATTGGGATTATCTGCAAAAAACTA
CGGTAGAGCGGTTTATGAATGTCTACGTGGTGGGCTTGATTTTACCAAGGATGATGAAAACGTGAACTCACAACCTTT
TATGCGTTGGAGAGACCGTTTCTTATTTTGTGCTGAAGCACTTTATAAAGCGCAGGCCGAAACAGGTGAAATCAAAGG
ACATTACTTGAATGCTACTGCAG

4: Consensus *trnH* plastid gene sequences of *Crocus* series *Crocus*, obtained by aligning both forward and reverse sequences for each species.

>EU110227.1 *Crocus moabiticus* voucher_C1927_photosystem_II_protein_D1_(psbA)_gene A new nucleotide sequence entered manually.

ACTGCCTTGATCCACTTGGCTACATCCGCCCTTATCTAGCTAAAGGATTTTCTCTTTTTCCATTCATCATTATTGTAT
TTATTCTTACCTTCATACTTCGATCGAGATATTCTATTGGACATAGAATGCCAATCTTTAAAATGTAAAAAAGGAGTA
ATCCGCTGTGACACGTTCACTAAAAAATCCTTTTGTAGCCAATCATTTATCGGGAAAAATTCAAAACTCAACAT
GAGGGAGGAGAAAGAAATAATAGTAACTTGGTCTCGGGCATCTACCATTATACCCACAATGATTGGCCATACAATCG
CTATTTCATAATGGAAGGAGCATTTCCTATTTATATAACAGATCGTATGGTAGGTCACAAATTGGGAGAATTCGCGC
CTACTCTGACTTTCGCGAGACATGCGAGAAACGATAATAAATCTCGTCGTTAATTTTGGATAGATCAAACAAGGGTGG
GGGATCCATGCCCATATTTGTTTTTCTAAAACAAATATGGGGCATGGATCCTTCAACGATTCATATACATTAATAGAG
AACTCTTATCCATTTATAGATGGAACCTTCGACAGCAGCTAGGTCTAAAGGAAAGTT-----

>EU110175.1 *Crocus banaticus* voucher_C1821_photosystem_II_protein_D1_(psbA)_gene A new nucleotide sequence entered manually.

ACTGCCTTGATCCACTTGGCTACATCCGCCCTTATCTAGCTAAAGGATTTTCTCTTTTTCCATTCATCATTATTGTAT
TTATTCTTACCTTCATACTTCGATCGAGATATTCTATTGGACATAGAATGCCAATCTTTAAAATGTAAAAAAGGAGTA
ATCCGCTGTGACACGTTCACTAAAAAATCCTTTTGTAGCCAATCATTTATCGGGAAAAATTCAAAACTCAACAT
GAGGGAGGAGAAAGAAATAATAGTAACTTGGTCTCGGGCATCTACCATTATACCCACAATGATTGGCCATACAATCG
CTATTTCATAATGGAAGGAGCATTTCCTATTTATATAACAGATCGTATGGTAGGTCACAAATTGGGAGAATTCGCGC
CTACTCTGACTTTCGCGAGACATGCGAGAAACGATAATAAATCTCGTCGTTAATTTTGGATAGATCAAACAAGGGTGG
GGGATCCATGCCCATATTTGTTTTTATAAAACAAAATATGGGGCATGGATCCTTCAACGATTCATATACATTAATAG
AGAACTCTTATCCATTTATAGATGGAACCTTCGACAGCAGCTAGGTCTAAAGGAAAGTT-----

> *Crocus* _CROC_ *C. asumaniae* (CasWD09) E13_trnh_ (trnH-psbA_).

GGATTACAAATCCACTGCCTTGATCCACTTGGCTACATCCGCCCTTATCTAGCTAAAGGATTTTCTCTTTTTCCATTC
ATCATTATTGTATTTATTCTTACCTTCATACTTCGATCGAGATATTCTATTGGACATAGAATGCCAATCTTTAAAATGT
AAAAAAGGAGTAATCCGCTGTGACACGTTCACTAAAAAATCCTTTTGTAGCCAATCATTTATCGGGAAAAATTC
AAAACCTCAACATGAGGGAGGAGAAAGAAATAATAGTAACTTGGTCTCGGGCATCTACCATTATACCCACAATGATTG
GCCATACAATCGCTATTCATAATGGAAGGAGCATTTCCTATTTATATAACAGATCGTATGGTAGGTCACAAATTGG
GAGAATTCGCGCCTACTCTGACTTTCGCGAGACATGCGAGAAACGATAATAAATCTCGTCGTTAATTTTGGATAGATC
AAACAAGGGTGGGGGATCCAT-
GCCCCATATTTGTTTTTATAAAACAAAATATGGGGCATGGATCCTTCAACGATTCATATACATTAATAGAGATAGAGA
ACTCTTATCCATTTATAGATGGAACCTTCGGCAGCAGCTAGGTCTAAAGGAAAGTTGTGAGCATTACGTTTCATGCATAA
CA

> *Crocus* _CROC_ *C. pallasii* (CplR09) E1_trnh_ (trnH-psbA_).

GGATTACAAATCCACTGCCTTGATCCACTTGGCTACATCCGCCCTTATCTAGCTAAAGGATTTTCTCTTTTTCCATTC
ATCATTATTGTATTTATTCTTACCTTCATACTTCGATCGAGATATTCTATTGGACATAGAATGCCAATCTTTAAAATGT
AAAAAAGGAGTAATCCGCTGTGACACGTTCACTAAAAAATCCTTTTGTAGCCAATCATTTATCGGGAAAAATTC
AAAACCTCAACATGAGGGAGGAGAAAGAAATAATAGTAACTTGGTCTCGGGCATCTACCATTATACCCACAATGATTG
GCCATACAATCGCTATTCATAATGGAAGGAGCATTTCCTATTTATATAACAGATCGTATGGTAGGTCACAAATTGG
GAGAATTCGCGCCTACTCTGACTTTCGCGAGACATGCGAGAAACGATAATAAATCTCGTCGTTAATTTTGGATAGATC
AAACAAGGGTGGGGGATCCATGCCCCATATTTGTTTTTCTAAAAACAAAATATGGGGCATGGATCCTTCAACGATTCAT
ATACATTA-----
TAGAGAACTCTTATCCATTTATAGATGGAACCTTCGACAGCAGCTAGGTCTAAAGGAAAGTTGTGAGCATTACGTTTCAT
GCATAACA

> *Crocus* _CROC_ *C. pallasii* (CplR09) E3_trnh_ (trnH-psbA_).

GGATTACAAATCCACTGCCTTGATCCACTTGGCTACATCCGCCCTTATCTAGCTAAAGGATTTTCTCTTTTTCCATTC
ATCATTATTGTATTTATTCTTACCTTCATACTTCGATCGAGATATTCTATTGGACATAGAATGCCAATCTTTAAAATGT
AAAAAAGGAGTAATCCGCTGTGACACGTTCACTAAAAAATCCTTTTGTAGCCAATCATTTATCGGGAAAAATTC
AAAACCTCAACATGAGGGAGGAGAAAGAAATAATAGTAACTTGGTCTCGGGCATCTACCATTATACCCACAATGATTG
GCCATACAATCGCTATTCATAATGGAAGGAGCATTTCCTATTTATATAACAGATCGTATGGTAGGTCACAAATTGG
GAGAATTCGCGCCTACTCTGACTTTCGCGAGACATGCGAGAAACGATAATAAATCTCGTCGTTAATTTTGGATAGATC
AAACAAGGGTGGGGGATCCATGCCCCATATTTGTTTTTCTAAAAACAAAATATGGGGCATGGATCCTTCAACGATTCAT
ATACATTA-----
TAGAGAACTCTTATCCATTTATAGATGGAACCTTCGACAGCAGCTAGGTCTAAAGGAAAGTTGTGAGCATTACGTTTCAT
GCATAACA

> *Crocus _CROC_ C. mathweii* (CmtHR09) E14_trnh_(trnH-psbA_).

GGATTCACAATCCACTGCCTTGATCCACTTGGCTACATCCGCCCTTATCTAGCTAAAGGATTTTCTCTTTTTCCATTC
ATCATTATTGTATTTATTCTTACCTTCATACTTCGATCGAGATATTCTATTGGACATAGAATGCCAATCTTTAAAATGTA
AAAAAAGGAGTAATCCGCTGTGACACGTTCACTAAAAAAAATCCTTTTGTAGCCAATCATTTATCGGGAAAAATTC
AAAACCTCAACATGAGGGAGGAGAAAAGAAATAATAGTAACTTGGTCTCGGGCATCTACCATTATACCCACAATGATTG
GCCATACAATCGCTATTATAATGGAAAAGGAGCATTTCCTATTTATATAACAGATCGTATGGTAGGTCACAAATTGG
GAGAATTCGCGCCTACTCTGACTTTCGCGAGACATGCGAGAAAACGATAATAAATCTCGTCGTTAATTTTGGATAGATC
AAACAAGGGTGGGGGATCCATGCCCCATATTTGTTTTCTAAAACAAAATATGGGGCATGGATCCTTCAACGATTTCAT
ATACATTAA-----
TAGAGAACTCTTATCCATTTATAGATGGAACCTCGACAGCAGCTAGGTCTAAAGGAAAGTTGTGAGCATTACGTTTCAT
GCATAACA

> *Crocus _CROC_ C. pallasii* (CplDD09) E4_trnh_(trnH-psbA_).

GGATTCACAATCCACTGCCTTGATCCACTTGGCTACATCCGCCCTTATCTAGCTAAAGGATTTTCTCTTTTTCCATTC
ATCATTATTGTATTTATTCTTACCTTCATACTTCGATCGAGATATTCTATTGGACATAGAATGCCAATCTTTAAAATGTA
AAAAAAGGAGTAATCCGCTGTGACACGTTCACTAAAAAAAATCCTTTTGTAGCCAATCATTTATCGGGAAAAATTC
AAAACCTCAACATGAGGGAGGAGAAAAGAAATAATAGTAACTTGGTCTCGGGCATCTACCATTATACCCACAATGATTG
GCCATACAATCGCTATTATAATGGAAAAGGAGCATTTCCTATTTATATAACAGATCGTATGGTAGGTCACAAATTGG
GAGAATTCGCGCCTACTCTGACTTTCGCGAGACATGCGAGAAAACGATAATAAATCTCGTCGTTAATTTTGGATAGATC
AAACAAGGGTGGGGGATCCATGCCCCATATTTGTTTTCTAAAACAAAATATGGGGCATGGATCCTTCAACGATTTCAT
ATACATTAATAGAGAACTCTTATCCATTTATAGATGGAACCTCGACAGCAGCTAGGTCTAAAGGAAAGTTGTGAGCATT
TACGTTTCATGCATAACA

> *Crocus _CROC_ C. cartwrightianus* (CcwBD10) E9_trnh_(trnH-psbA_).

GGATTCACAATCCACTGCCTTGATCCACTTGGCTACATCCGCCCTTATCTAGCTAAAGGATTTTCTCTTTTTCCATTC
ATCATTATTGTATTTATTCTTACCTTCATACTTCGATCGAGATATTCTATTGGACATAGAATGCCAATCTTTAAAATGTA
AAAAAAGGAGTAATCCGCTGTGACACGTTCACTAAAAAAAATCCTTTTGTAGCCAATCATTTATCGGGAAAAATTC
AAAACCTCAACATGAGGGAGGAGAAAAGAAATAATAGTAACTTGGTCTCGGGCATCTACCATTATACCCACAATGATTG
GCCATACAATCGCTATTATAATGGAAAAGGAGCATTTCCTATTTATATAACAGATCGTATGGTAGGTCACAAATTGG
GAGAATTCGCGCCTACTCTGACTTTCGCGAGACATGCGAGAAAACGATAATAAATCTCGTCGTTAATTTTGGATAGATC
AAACAAGGGTGGGGGATCCATAGCCCCATATTTGTTTTAGAAAACAAAATATGGGGCATGGATCCTTCAACGATTTCAT
TATACATTAA-----
TAGAGAACTCTTATCCATTTATAGATGGAACCTCGACAGCAGCTAGGTCTAAAGGAAAGTTGTGAGCATTACGTTTCAT
GCATAACA

> *Crocus _CROC_ C. sativuscartwrightianus 'Albus'* (CstcP09)E15_trnh_(trnH-psbA_).

GGATTCACAATCCACTGCCTTGATCCACTTGGCTACATCCGCCCTTATCTAGCTAAAGGATTTTCTCTTTTTCCATTC
ATCATTATTGTATTTATTCTTACCTTCATACTTCGATCGAGATATTCTATTGGACATAGAATGCCAATCTTTAAAATGTA
AAAAAAGGAGTAATCCGCTGTGACACGTTCACTAAAAAAAATCCTTTTGTAGCCAATCATTTATCGGGAAAAATTC
AAAACCTCAACATGAGGGAGGAGAAAAGAAATAATAGTAACTTGGTCTCGGGCATCTACCATTATACCCACAATGATTG
GCCATACAATCGCTATTATAATGGAAAAGGAGCATTTCCTATTTATATAACAGATCGTATGGTAGGTCACAAATTGG
GAGAATTCGCGCCTACTCTGACTTTCGCGAGACATGCGAGAAAACGATAATAAATCTCGTCGTTAATTTTGGATAGATC
AAACAAGGGTGGGGGATCCATAGCCCCATATTTGTTTTAGAAAACAAAATATGGGGCATGGATCCTTCAACGATTTCAT
TATACATTAA-----
TAGAGAACTCTTATCCATTTATAGATGGAACCTCGACAGCAGCTAGGTCTAAAGGAAAGTTGTGAGCATTACGTTTCAT
GCATAACA

> *Crocus _CROC_ C. pallasii* (CplVD09) E2_trnh_(trnH-psbA_).

GGTTACACAATCCACTGCCTTGATCCACTTGGCTACATCCGCCCTTATCTAGCTAAAGGATTTTCTCTTTTTCCATTC
ATCATTATTGTATTTATTCTTACCTTCATACTTCGATCGAGATATTCTATTGGACATAGAATGCCAATCTTTAAAATGTA
AAAAAAGGAGTAATCCGCTGTGACACGTTCACTAAAAAAAATCCTTTTGTAGCCAATCATTTATCGGGAAAAATTC
AAAACCTCAACATGAGGGAGGAGAAAAGAAATAATAGTAACTTGGTCTCGGGCATCTACCATTATACCCACAATGATTG
GCCATACAATCGCTATTATAATGGAAAAGGAGCATTTCCTATTTATATAACAGATCGTATGGTAGGTCACAAATTGG
GAGAATTCGCGCCTACTCTGACTTTCGCGAGACATGCGAGAAAACGATAATAAATCTCGTCGTTAATTTTGGATAGATC
AAACAAGGGTGGGGGATCCATAGCCCCATATTTGTTTTAGAAAACAAAATATGGGGCATGGATCCTTCAACGATTTCAT
TATACATTAATAGAGAACTCTTATCCATTTATAGATGGAACCTCGACAGCAGCTAGGTCTAAAGGAAAGTTGTGAGCA
TTACGTTTCATGCATAACA

> *Crocus _CROC_ C. Kashmirianus (CstCD09) E5_trnh_ (trnH-psbA_).*

GGATTCACAATCCACTGCCTTGATCCACTTGGCTACATCCGCCCTTATCTAGCTAAAGGATTTTCTCTTTTTCCATTC
ATCATTATTGTATTTATTTCTTACCTTCATACTTCGATCGAGATATTTCTATTGGACATAGAATGCCAATCTTTAAAAATGTA
AAAAAAGGAGTAATCCGCTGTGACACGTTCACTAAAAAAAATCCTTTTGTAGCCAATCATTATCGGGAAAAAATCA
AAAACCTCAACATGAGGGAGGAGAAAAGAAATAATAGTAACTTGGTCTCGGGCATCTACCATTATACCCACAATGATTG
GCCATACAATCGCTATTCATAATGGAAAGGAGCATTTCCTATTTATATAACAGATCGTATGGTAGGTCACAAATTGG
GAGAATTCGCGCCTACTCTGACTTTTCGCGAGACATGCGAGAAAACGATAATAAATCTCGTCGTTAATTTTGGATAGATC
AAACAAGGGTGGGGGATCCATAGCCCCATATTTGTTTTAGAAAAACAATATGGGGCATGGATCCTTCAACGATTCA
TATACATTAATAGAGAACTCTTATCCATTTATAGATGGAACCTTCGACAGCAGCTAGGTCTAAAAGGAAAAGTTGTGAGCA
TTACGTTTCATGCATAACA

> *Crocus _CROC_ C. hadriaticus (ChdWD08) E11_trnh_ (trnH-psbA_).*

GGATTCACAATCCACTGCCTTGATCCACTTGGCTACATCCGCCCTTATCTAGCTAAAGGATTTTCTCTTTTTCCATTC
ATCATTATTGTATTTATTTCTTACCTTCATACTTCGATCGAGATATTTCTATTGGACATAGAATGCCAATCTTTAAAAATGTA
AAAAAAGGAGTAATCCGCTGTGACACGTTCACTAAAAAAAATCCTTTTGTAGCCAATCATTATCGGGAAAAAATCA
AAAACCTCAACATGAGGGAGGAGAAAAGAAATAATAGTAACTTGGTCTCGGGCATCTACCATTATACCCACAATGATTG
CCATACAATCGCTATTCATAATGGAAAGGAGCATTTCCTATTTATATAACAGATCGTATGGTAGGTCACAAATTGGG
AGAATTCGCGCCTACTCTGACTTTTCGCGAGACATGCGAGAAAACGATAATAAATCTCGTCGTTAATTTTGGATAGATCA
AACAAAGGGTGGGGGATCCATAGCCCCATATTTGTTTTAGAAAAACAATATGGGGCATGGATCCTTCAACGATTCA
ATACATTAATAGAGAACTCTTATCCATTTATAGATGGAACCTTCGACAGCAGCTAGGTCTAAAAGGAAAAGTTGTGAGCAT
TACGTTTCATGCATAACA

> *Crocus _CROC_ C. oreoreticus (CorVR09) E10_trnh_ (trnH-psbA_).*

GGATTCACAATCCACTGCCTTGATCCACTTGGCTACATCCGCCCTTATCTAGCTAAAGGATTTTCTCTTTTTCCATTC
ATCATTATTGTATTTATTTCTTACCTTCATACTTCGATCGAGATATTTCTATTGGACATAGAATGCCAATCTTTAAAAATGTA
AAAAAAGGAGTAATCCGCTGTGACACGTTCACTAAAAAAAATCCTTTTGTAGCCAATCATTATCGGGAAAAAATCA
AAAACCTCAACATGAGGGAGGAGAAAAGAAATAATAGTAACTTGGTCTCGGGCATCTACCATTATACCCACAATGATTG
GCCATACAATCGCTATTCATAATGGAAAGGAGCATTTCCTATTTATATAACAGATCGTATGGTAGGTCACAAATTGG
GAGAATTCGCGCCTACTCTGACTTTTCGCGAGACATGCGAGAAAACGATAATAAATCTCGTCGTTAATTTTGGATAGATC
AAACAAGGGTGGGGGATCCATAGCCCCATATTTGTTTTAGAAAAACAATATGGGGCATGGATCCTTCAACGATTCA
TATACATTA-----
TAGAGAACTCTTATCCATTTATAGATGGAACCTTCGACAGCAGCTAGGTCTAAAAGGAAAAGTTGTGAGCATTACGTTTCAT
GCATAACA

> *Crocus _CROC_ C. sativus_Spain (C.stPER09) E6_trnh_ (trnH-psbA_).*

GGATTCACAATCCACTGCCTTGATCCACTTGGCTACATCCGCCCTTATCTAGCTAAAGGATTTTCTCTTTTTCCATTC
TCATTATTGTATTTATTTCTTACCTTCATACTTCGATCGAGATATTTCTATTGGACATAGAATGCCAATCTTTAAAAATGTA
AAAAAAGGAGTAATCCGCTGTGACACGTTCACTAAAAAAAATCCTTTTGTAGCCAATCATTATCGGGAAAAAATCA
AAAACCTCAACATGAGGGAGGAGAAAAGAAATAATAGTAACTTGGTCTCGGGCATCTACCATTATACCCACAATGATTG
CCATACAATCGCTATTCATAATGGAAAGGAGCATTTCCTATTTATATAACAGATCGTATGGTAGGTCACAAATTGG
AGAATTCGCGCCTACTCTGACTTTTCGCGAGACATGCGAGAAAACGATAATAAATCTCGTCGTTAATTTTGGATAGATCA
AACAAAGGGTGGGGGATCCATAGCCCCATATTTGTTTTAGAAAAACAATATGGGGCATGGATCCTTCAACGATTCA
ATACATTAATAGAGAACTCTTATCCATTTATAGATGGAACCTTCGACAGCAGCTAGGTCTAAAAGGAAAAGTTGTGAGCAT
TACGTTTCATGCATAACA

> *Crocus _CROC_ C. thomasi (CtmVD09) E12_trnh_ (trnH-psbA_).*

GGATTCACAATCCACTGCCTTGATCCACTTGGCTACATCCGCCCTTATCTAGCTAAAGGATTTTCTCTTTTTCCATTC
ATCATTATTGTATTTATTTCTTACCTTCATACTTCGATCGAGATATTTCTATTGGACATAGAATGCCAATCTTTAAAAATGTA
AAAAAAGGAGTAATCCGCTGTGACACGTTCACTAAAAAAAATCCTTTTGTAGCCAATCATTATCGGGAAAAAATCA
AAAACCTCAACATGAGGGAGGAGAAAAGAAATAATAGTAACTTGGTCTCGGGCATCTACCATTATACCCACAATGATTG
GCCATACAATCGCTATTCATAATGGAAAGGAGCATTTCCTATTTATATAACAGATCGTATGGTAGGTCACAAATTGG
GAGAATTCGCGCCTACTCTGACTTTTCGCGAGACATGCGAGAAAACGATAATAAATCTCGTCGTTAATTTTGGATAGATC
AAACAAGGGTGGGGGATCCATAGCCCCATATTTGTTTTAGAAAAACAATATGGGGCATGGATCCTTCAACGATTCA
TATACATTAATAGAGAACTCTTATCCATTTATAGATGGAACCTTCGACAGCAGCTAGGTCTAAAAGGAAAAGTTGTGAGCA
TTACGTTTCATGCATAACA

> *Crocus _CROC_ C. cartwrightianus (CcrCR09) E8_trnh_ (trnH-psbA_)*.

GGATTACAATCCACTGCCTTGATCCACTTGGCTACATCCGCCCTTATCTAGCTAAAGGATTTTCTCTTTTTCCATT
ATCATTATTGTATTTATTCTTACCTTCATACTTCGATCGAGATATTCTATTGGACATAGAATGCCAATCTTTAAAATGTA
AAAAAAGGAGTAATCCGCTGTGACACGTTCACTAAAAAAAATCCTTTTGTAGCCAATCATTATCGGGAAAAATTCA
AAAACACTCAACATGAGGGAGGAGAAAAGAAATAATAGTAACTTGGTCTCGGGCATCTACCATTATACCCACAATGATTG
GCCATACAATCGCTATTATAATGAAAAGGAGCATTTCCTATTTATATAACAGATCGTATGGTAGGTCACAAATTGG
GAGAATTCGCGCCTACTCTGACTTTCGCGAGACATGCGAGAAAACGATAATAAATCTCGTCGTTAATTTTGGATAGATC
AAACAAGGGTGGGGATCCATAGCCCCATATTTGTTTTAGAAAAACAAATATGGGGCATGGATCCTTCAACGATTCA
TATACATTAATAGAGAACTCTTATCCATTTATAGATGGAACCTTCGACAGCAGCTAGGTCTAAAGGAAAAGTTGTGAGCA
TTACGTTTCATGCATAACA

> *Crocus _CROC_ C. cartwrightianus (CcwAD08) E7_trnh_ (trnH-psbA_)*.

GGATTACAATCCACTGCCTTGATCCACTTGGCTACATCCGCCCTTATCTAGCTAAAGGATTTTCTCTTTTTCCATT
ATCATTATTGTATTTATTCTTACCTTCATACTTCGATCGAGATATTCTATTGGACATAGAATGCCAATCTTTAAAATGTA
AAAAAAGGAGTAATCCGCTGTGACACGTTCACTAAAAAAAATCCTTTTGTAGCCAATCATTATCGGGAAAAATTCA
AAAACACTCAACATGAGGGAGGAGAAAAGAAATAATAGTAACTTGGTCTCGGGCATCTACCATTATACCCACAATGATTG
GCCATACAATCGCTATTATAATGAAAAGGAGCATTTCCTATTTATATAACAGATCGTATGGTAGGTCACAAATTGG
GAGAATTCGCGCCTACTCTGACTTTCGCGAGACATGCGAGAAAACGATAATAAATCTCGTCGTTAATTTTGGATAGATC
AAACAAGGGTGGGGATCCATAGCCCCATATTTGTTTTAGAAAAACAAATATGGGGCATGGATCCTTCAACGATTCA
TATACATTAATAGAGAACTCTTATCCATTTATAGATGGAACCTTCGACAGCAGCTAGGTCTAAAGGAAAAGTTGTGAGCA
TTACGTTTCATGCATAACA

Table A4.3: Nucleotide sequence similarity (%) of the *Crocus* species for *matK* XF+5R plastid gene sequences

	EU496995.1 <i>Crocus banaticus</i>	<i>C.cartwrightianus</i> (CcwBD10)	EU497045.1 <i>Crocus maobiticus</i>	<i>C.asumaniae</i> (CasWD09)	<i>C.mathewii</i> (CmtHR09)	<i>C.pallasii</i> (CpltR09)	<i>C.pallasii</i> (CplDD09)	<i>C.pallasii</i> (CplR09)	<i>C.pallasii</i> (CplVD09)	<i>C.Cashmirianus'</i> <i>Hort. Lilac</i> (CstCD09)	<i>C.hadriaticus</i> (ChdWD08)	<i>C.oreocreticus</i> (CorVR09)	<i>C.sativus</i> (C.stPER09)	<i>C.thomasii</i> (CtmVD09)	<i>C.cartwrightianus</i> (CcrCR09)	<i>C.sativus</i> <i>cartwrightianus</i> 'Albus' '(CstcP09)	<i>C.cartwrightianus</i> (CcwAD08)
EU496995.1 <i>Crocus banaticus</i>		98.5	98.3	98.6	98.6	98.4	98.4	98.4	98.6	98.6	98.4	98.6	98.6	98.5	98.6	98.5	98.5
<i>C.cartwrightianus</i> (CcwBD10)	98.5		97.5	97.8	98	97.9	97.7	97.6	98.1	98.1	97.9	98.1	98.1	98	98.1	98	97.9
EU497045.1 <i>Crocus maobiticus</i>	98.3	97.5		99.4	99.4	99.1	99.1	99.1	99.4	99.4	99.1	99.4	99.4	99.3	99.4	99.3	99.3
<i>C.asumaniae</i> (CasWD09)	98.6	97.8	99.4		99.6	99.5	99.3	99.2	99.5	99.5	99.3	99.5	99.5	99.4	99.5	99.4	99.3
<i>C.mathewii</i> (CmtHR09)	98.6	98	99.4	99.6		99.7	99.5	99.6	99.7	99.7	99.5	99.7	99.7	99.6	99.7	99.6	99.5
<i>C.pallasii</i> (CpltR09)	98.4	97.9	99.1	99.5	99.7		99.8	99.7	99.6	99.6	99.4	99.6	99.6	99.5	99.6	99.5	99.4
<i>C.pallasii</i> (CplDD09)	98.4	97.7	99.1	99.3	99.5	99.8		99.9	99.4	99.4	99.2	99.4	99.4	99.3	99.4	99.3	99.2
<i>C.pallasii</i> (CplR09)	98.4	97.6	99.1	99.2	99.6	99.7	99.9		99.3	99.3	99	99.3	99.3	99.2	99.3	99.2	99
<i>C.pallasii</i> (CplVD09)	98.6	98.1	99.4	99.5	99.7	99.6	99.4	99.3		100	99.8	100	100	99.9	100	99.9	99.8
<i>C.cashmirianus'</i> <i>Hort. Lilac</i> (CstCD09)	98.6	98.1	99.4	99.5	99.7	99.6	99.4	99.3	100		99.8	100	100	99.9	100	99.9	99.8
<i>C.hadriaticus</i> (ChdWD08)	98.4	97.9	99.1	99.3	99.5	99.4	99.2	99	99.8	99.8		99.8	99.8	99.7	99.8	99.7	99.6
<i>C.oreocreticus</i> (CorVR09)	98.6	98.1	99.4	99.5	99.7	99.6	99.4	99.3	100	100	99.8		100	99.9	100	99.9	99.8
<i>C.sativus</i> (C.stPER09)	98.6	98.1	99.4	99.5	99.7	99.6	99.4	99.3	100	100	99.8	100		99.9	100	99.9	99.8
<i>C.thomasii</i> (CtmVD09)	98.5	98	99.3	99.4	99.6	99.5	99.3	99.2	99.9	99.9	99.7	99.9	99.9		99.9	99.8	99.7
<i>C.cartwrightianus</i> (CcrCR09)	98.6	98.1	99.4	99.5	99.7	99.6	99.4	99.3	100	100	99.8	100	100	99.9		99.9	99.8
<i>C.sativus</i> <i>cartwrightianus</i> 'Albus' '(CstcP09)	98.5	98	99.3	99.4	99.6	99.5	99.3	99.2	99.9	99.9	99.7	99.9	99.9	99.8	99.9		99.9
<i>C.cartwrightianus</i> (CcwAD08)	98.5	97.9	99.3	99.3	99.5	99.4	99.2	99	99.8	99.8	99.6	99.8	99.8	99.7	99.8	99.9	

Table A4.4: Nucleotide sequence similarity (%) of the *Crocus* species for *matK* 390F+1326R plastid gene sequences

	EU497045.1 <i>Crocus moabiticus</i>	EU496995.1 <i>Crocus banaticus</i>	<i>C.cartwrightianus</i> (CcwBD10)	<i>C.cartwrightianus</i> (CcrCR09)	<i>C.cartwrightianus</i> (CcwAD08)	<i>C.Cashmirianus'</i> <i>Hort. Lilac</i> (CstCD09)	<i>C.hadriaticus</i> (ChdWD08)	<i>C.oreocreticus</i> (CorVR09)	<i>C.asumaniae</i> (CasWD09)	<i>C. mathewii</i> (CmtHR09)	<i>C. pallasii</i> (CplDD09)	<i>C. pallasii</i> (CplR09)	<i>C. pallasii</i> (CplTR09)	<i>C. pallasii</i> (CplVD09)	<i>C.sativus</i> <i>cartwrightianus</i> 'Albus'(CstcP09)	<i>C. sativus</i> (C.stPER09)	<i>C. thomasii</i> (CtmVD09)
EU497045.1 <i>Crocus moabiticus</i>		98.3	98.4	99.3	99.3	99.4	99.1	99.4	99.4	99.4	99.1	99.1	99	99.4	99.3	99.4	99.3
EU496995.1 <i>Crocus banaticus</i>	98.3		97.9	98.5	98.5	98.6	98.4	98.6	98.6	98.6	98.4	98.4	98.3	98.6	98.5	98.6	98.5
<i>C.cartwrightianus</i> (CcwBD10)	98.4	97.9		98.8	98.8	98.8	98.3	98.8	98.7	98.8	98.4	98.5	98.3	98.7	98.9	98.8	98.7
<i>C.cartwrightianus</i> (CcrCR09)	99.3	98.5	98.8		100	99.8	99.6	99.8	99.5	99.6	99.3	99.4	99.2	99.9	99.9	99.8	99.7
<i>C.cartwrightianus</i> (CcwAD08)	99.3	98.5	98.8	100		99.8	99.6	99.8	99.5	99.6	99.3	99.4	99.2	99.9	99.9	99.8	99.7
<i>C.Cashmirianus'</i> <i>Hort. Lilac</i> (CstCD09)	99.4	98.6	98.8	99.8	99.8		99.6	100	99.7	99.8	99.5	99.6	99.4	99.9	99.9	100	99.9
<i>C.hadriaticus</i> (ChdWD08)	99.1	98.4	98.3	99.6	99.6	99.6		99.6	99.3	99.4	99.1	99.2	99	99.7	99.5	99.6	99.5
<i>C.oreocreticus</i> (CorVR09)	99.4	98.6	98.8	99.8	99.8	100	99.6		99.7	99.8	99.5	99.6	99.4	99.9	99.9	100	99.9
<i>C.asumaniae</i> (CasWD09)	99.4	98.6	98.7	99.5	99.5	99.7	99.3	99.7		99.9	99.6	99.7	99.5	99.6	99.6	99.7	99.6
<i>C.mathewii</i> (CmtHR09)	99.4	98.6	98.8	99.6	99.6	99.8	99.4	99.8	99.9		99.7	99.8	99.6	99.7	99.7	99.8	99.7
<i>C. pallasii</i> (CplDD09)	99.1	98.4	98.4	99.3	99.3	99.5	99.1	99.5	99.6	99.7		99.9	99.7	99.4	99.4	99.5	99.4
<i>C.pallasii</i> (CplR09)	99.1	98.4	98.5	99.4	99.4	99.6	99.2	99.6	99.7	99.8	99.9		99.8	99.5	99.5	99.6	99.5
<i>C.pallasii</i> (CplTR09)	99	98.3	98.3	99.2	99.2	99.4	99	99.4	99.5	99.6	99.7	99.8		99.3	99.3	99.4	99.3
<i>C.pallasii</i> (CplVD09)	99.4	98.6	98.7	99.9	99.9	99.9	99.7	99.9	99.6	99.7	99.4	99.5	99.3		99.8	99.9	99.8
<i>C.sativus</i> <i>cartwrightianus</i> 'Albus' (CstcP09)	99.3	98.5	98.9	99.9	99.9	99.9	99.5	99.9	99.6	99.7	99.4	99.5	99.3	99.8		99.9	99.8
<i>C.sativus</i> (C.stPER09)	99.4	98.6	98.8	99.8	99.8	100	99.6	100	99.7	99.8	99.5	99.6	99.4	99.9	99.9		99.9
<i>C.thomasii</i> (CtmVD09)	99.3	98.5	98.7	99.7	99.7	99.9	99.5	99.9	99.6	99.7	99.4	99.5	99.3	99.8	99.8	99.9	

Table A4.5: Nucleotide sequence similarity (%) of the *Crocus* species for the *rbcl* plastid gene sequences

	JX903213.1 <i>Crocus banaticus</i>	<i>C.asumaniae</i> (CasWD09)	<i>C.mathewii</i> (CmtHR09)	<i>C.pallasii</i> (CplDD09)	<i>C.pallasii</i> (CplR09)	<i>C.pallasii</i> (CplR09)	<i>C.cartwrightianus</i> (CcwBD10)	<i>C.Cashmirianus</i> (CstCD09)	<i>C.hadriaticus</i> (ChdWD08)	<i>C.oreoreticus</i> (CorVR09)	<i>C.pallasii</i> (CplVD09)	<i>C.sativus</i> <i>cartwrightianus</i> 'Albus' (CstcP09)	<i>C.thomasii</i> (CtmVD09)	<i>C.cartwrightianus</i> (CcwAD08)	<i>C.sativus</i> (C.stPER09)	<i>C cartwrightianus</i> (CcrCR09)
JX903213.1 <i>Crocus banaticus</i>		99.6	98.9	98.9	98.9	98.9	98.9	99.1	99.1	99.1	99.1	99.1	99.1	99.1	99.1	99.1
<i>C.asumaniae</i> (CasWD09)	99.6		99.2	99.2	99.2	99.3	99.3	99.3	99.2	99.2	99.2	99.2	99.2	99.2	99.2	99.2
<i>C.mathewii</i> (CmtHR09)	98.9	99.2		100	100	99.9	99.9	99.6	99.7	99.7	99.7	99.7	99.7	99.7	99.4	99.4
<i>C.pallasii</i> (CplDD09)	98.9	99.2	100		100	99.9	99.9	99.6	99.7	99.7	99.7	99.7	99.7	99.7	99.4	99.4
<i>C.pallasii</i> (CplR09)	98.9	99.2	100	100		99.9	99.9	99.6	99.7	99.7	99.7	99.7	99.7	99.7	99.4	99.4
<i>C.pallasii</i> (CplR09)	98.9	99.3	99.9	99.9	99.9		100	99.7	99.6	99.6	99.6	99.6	99.6	99.6	99.6	99.6
<i>C.cartwrightianus</i> (CcwBD10)	98.9	99.3	99.9	99.9	99.9	100		99.7	99.6	99.6	99.6	99.6	99.6	99.6	99.6	99.6
<i>C.cashmirianus</i> (CstCD09)	99.1	99.3	99.6	99.6	99.6	99.7	99.7		99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
<i>C.hadriaticus</i> (ChdWD08)	99.1	99.2	99.7	99.7	99.7	99.6	99.6	99.9		100	100	100	100	100	99.7	99.7
<i>C.oreoreticus</i> (CorVR09)	99.1	99.2	99.7	99.7	99.7	99.6	99.6	99.9	100		100	100	100	100	99.7	99.7
<i>C.pallasii</i> (CplVD09)	99.1	99.2	99.7	99.7	99.7	99.6	99.6	99.9	100	100		100	100	100	99.7	99.7
<i>C.sativus</i> <i>cartwrightianus</i> 'Albus' (CstcP09)	99.1	99.2	99.7	99.7	99.7	99.6	99.6	99.9	100	100	100		100	100	99.7	99.7
<i>C.thomasii</i> (CtmVD09)	99.1	99.2	99.7	99.7	99.7	99.6	99.6	99.9	100	100	100	100		100	99.7	99.7
<i>C.cartwrightianus</i> (CcwAD08)	99.1	99.2	99.7	99.7	99.7	99.6	99.6	99.9	100	100	100	100	100		99.7	99.7
<i>C.sativus</i> (C.stPER09)	99.1	99.2	99.4	99.4	99.4	99.6	99.6	99.9	99.7	99.7	99.7	99.7	99.7	99.7		100
<i>C.cartwrightianus</i> (CcrCR09)	99.1	99.2	99.4	99.4	99.4	99.6	99.6	99.9	99.7	99.7	99.7	99.7	99.7	99.7	100	

Table A4.6: Nucleotide sequence similarity (%) of the *Crocus* species for the *trnH* plastid gene sequences

	EU110227.1 <i>Crocus moabiticus</i>	EU110175.1 <i>Crocus banaticus</i>	<i>C.asumaniae</i> (CasWD09)	<i>C. pallasii</i> (CpIR09)	<i>C. pallasii</i> (CpltR09)	<i>C. mathewii</i> (CmtHR09)	<i>C. pallasii</i> (CplDD09)	<i>C.cartwrightianus</i> (CcwBD10)	<i>C.sativus</i> <i>cartwrightianus</i> 'Albus' (CstcP09)	<i>C. pallasii</i> (CplVD09)	<i>C.Cashmirianus</i> (CstCD09)	<i>C.hadriaticus</i> (ChdWD08)	<i>C.oreoreticus</i> (CorVR09)	<i>C.sativus</i> (C.stPER09)	<i>C.thomasii</i> (CtmVD09)	<i>C.cartwrightianus</i> (CcrCR09)	<i>C.cartwrightianus</i> (CcwAD08)
EU110227.1 <i>Crocus moabiticus</i>		99.7	98.5	99.8	99.8	99.8	99.8	99.3	98.8	98.8	98.8	98.8	98.8	98.8	98.8	98.8	98.8
EU110175.1 <i>Crocus banaticus</i>	99.7		98.9	99.8	99.8	99.8	99.8	99.7	99	99	99	99	99	99	99	99	99
<i>C.asumaniae</i> (CasWD09)	98.5	98.9		98.8	98.8	98.8	98.8	98.6	98	97.7	98	98	98	98	98	98	98
<i>C. pallasii</i> (CpIR09)	99.8	99.8	98.8		100	100	100	99.5	98.9	98.6	98.9	98.9	98.9	98.9	98.9	98.9	98.9
<i>C. pallasii</i> (CpltR09)	99.8	99.8	98.8	100		100	100	99.5	98.9	98.6	98.9	98.9	98.9	98.9	98.9	98.9	98.9
<i>C.mathewii</i> (CmtHR09)	99.8	99.8	98.8	100	100		100	99.5	98.9	98.6	98.9	98.9	98.9	98.9	98.9	98.9	98.9
<i>C. pallasii</i> (CplDD09)	99.8	99.8	98.8	100	100	100		99.5	98.9	98.6	98.9	98.9	98.9	98.9	98.9	98.9	98.9
<i>C. cartwrightianus</i> (CcwBD10)	99.3	99.7	98.6	99.5	99.5	99.5	99.5		99.4	99.1	99.4	99.4	99.4	99.4	99.4	99.4	99.4
<i>C.sativus</i> <i>cartwrightianus</i> 'Albus' (CstcP09)	98.8	99	98	98.9	98.9	98.9	98.9	99.4		99.7	100	100	100	100	100	100	100
<i>C.pallasii</i> (CplVD09)	98.8	99	97.7	98.6	98.6	98.6	98.6	99.1	99.7		99.7	99.7	99.7	99.7	99.7	99.7	99.7
<i>C.Cashmirianus</i> (CstCD09)	98.8	99	98	98.9	98.9	98.9	98.9	99.4	100	99.7		100	100	100	100	100	100
<i>C.hadriaticus</i> (ChdWD08)	98.8	99	98	98.9	98.9	98.9	98.9	99.4	100	99.7	100		100	100	100	100	100
<i>C.oreoreticus</i> (CorVR09)	98.8	99	98	98.9	98.9	98.9	98.9	99.4	100	99.7	100	100		100	100	100	100
<i>C.sativus</i> (C.stPER09)	98.8	99	98	98.9	98.9	98.9	98.9	99.4	100	99.7	100	100	100		100	100	100
<i>C.thomasii</i> (CtmVD09)	98.8	99	98	98.9	98.9	98.9	98.9	99.4	100	99.7	100	100	100	100		100	100
<i>C.cartwrightianus</i> (CcrCR09)	98.8	99	98	98.9	98.9	98.9	98.9	99.4	100	99.7	100	100	100	100	100		100
<i>C.cartwrightianus</i> (CcwAD08)	98.8	99	98	98.9	98.9	98.9	98.9	99.4	100	99.7	100	100	100	100	100	100	

9 Appendix 5

Figure A5.1: Schematic representation of Repetitive DNA isolation from *C. sativus* and colony Dot blot Hybridization procedure .

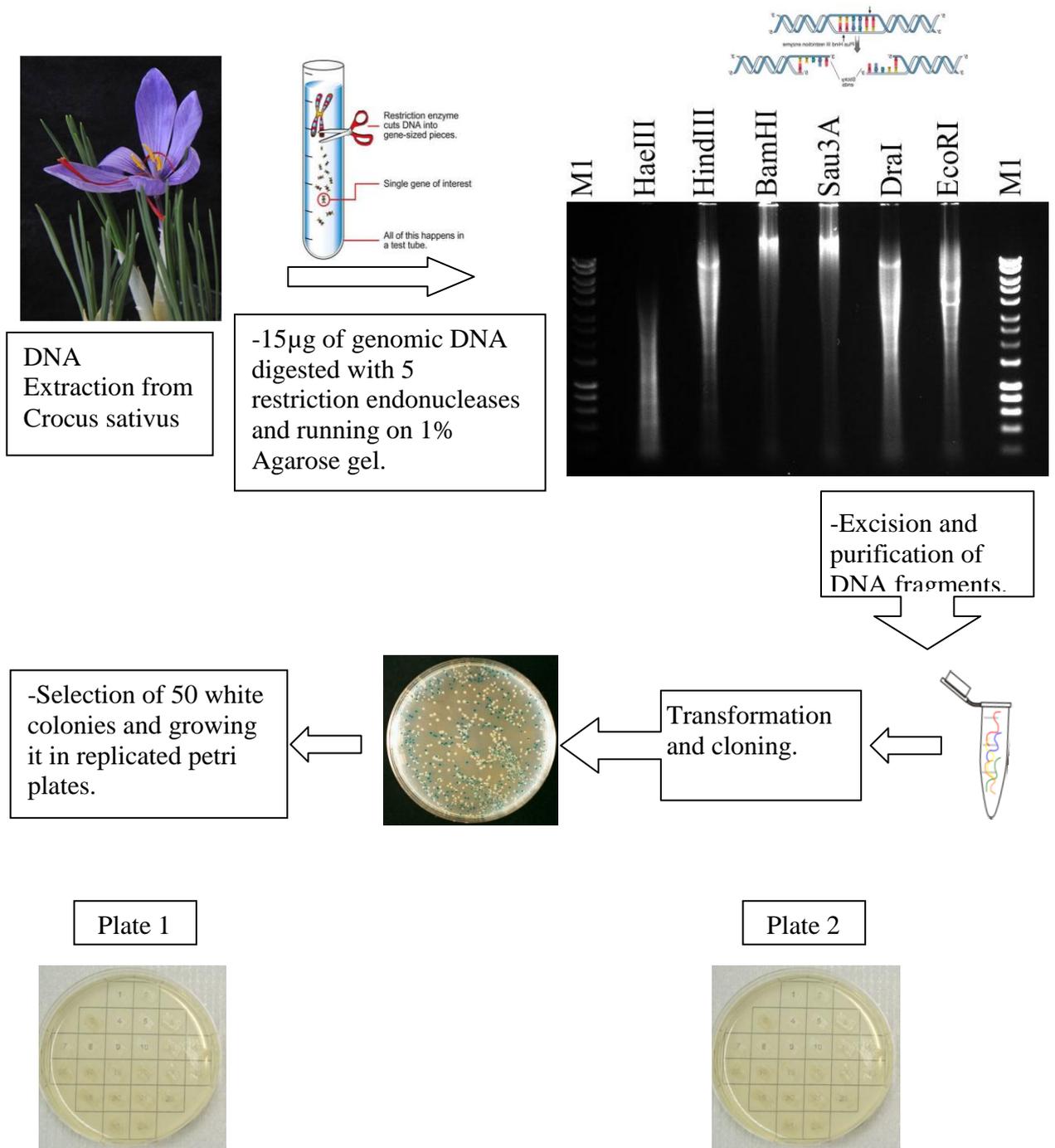
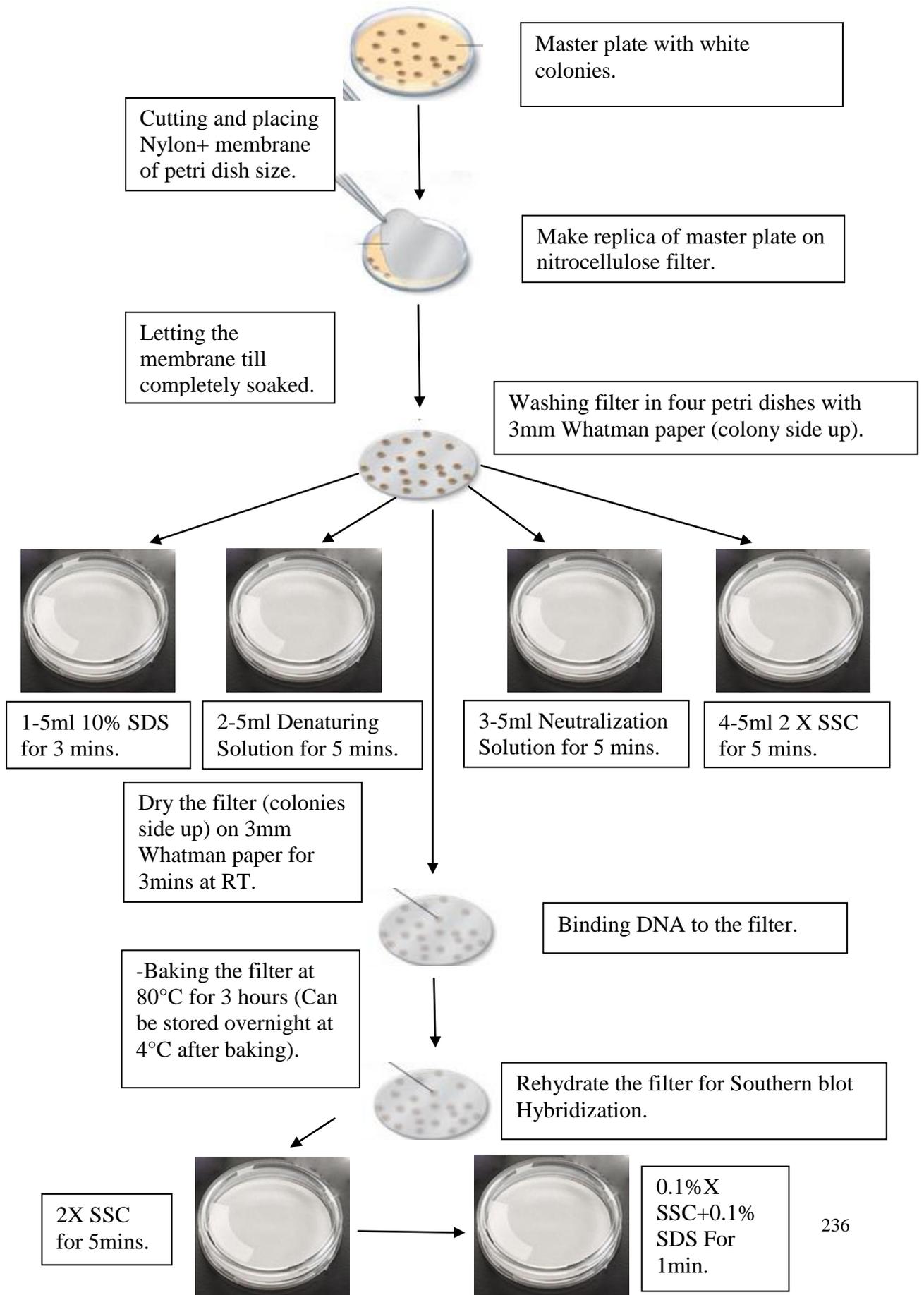


Plate 1:

Colony Dot blot-Hybridization



Southern blot Hybridization

1-Prehybridization

Prehybridization Solution:

- 1ml 20XSSC.
- 50µl Salmon sperm DNA.
- 0.25ml 10% SDS.
- 0.5 ml 50x Denhardt's.
- 3ml Sigma water.

2-Hybridization

Hybridization solution:

- Remove 1ml of prehybridization solution and add.
- 50µl Salmon sperm DNA.
- 0.5ml 50x Dextran sulphate.
- 4 µl DNA probe.

3-Post hybridization

- 50ml of 2XSSC+0.1%SDS (2x) for 5mins at 56°C.
- 50 ml of 0.5XSSC+0.1%SDS (2x) for 15mins at 56°C.

4-Detection

- Wash membrane in Buffer 1+ 0.3% Tween20 for 5min
- Incubate in 50ml Buffer 2 for 30 mins.
- Add anti-digoxigenin-AP (1:5000) diluted in Buffer 2 and put plastic coverslip.
- Incubate in the dark for 30 mins.
- 2x wash in 100ml of Buffer 1 for 15 mins.
- Equilibrate in 20 ml Buffer 3 for 2 mins.
- Add CPD star (1:100) diluted in Buffer3 for 5min in the dark.
- Wrap membrane in cling film (colonies side down).

5-Develop the film

- Placing the membrane in a cassette containing 1-2 intensifier screen.
- Load Kodak/Fuji X-ray film in the dark room.
- Expose the membrane in X-ray films and develop the film.

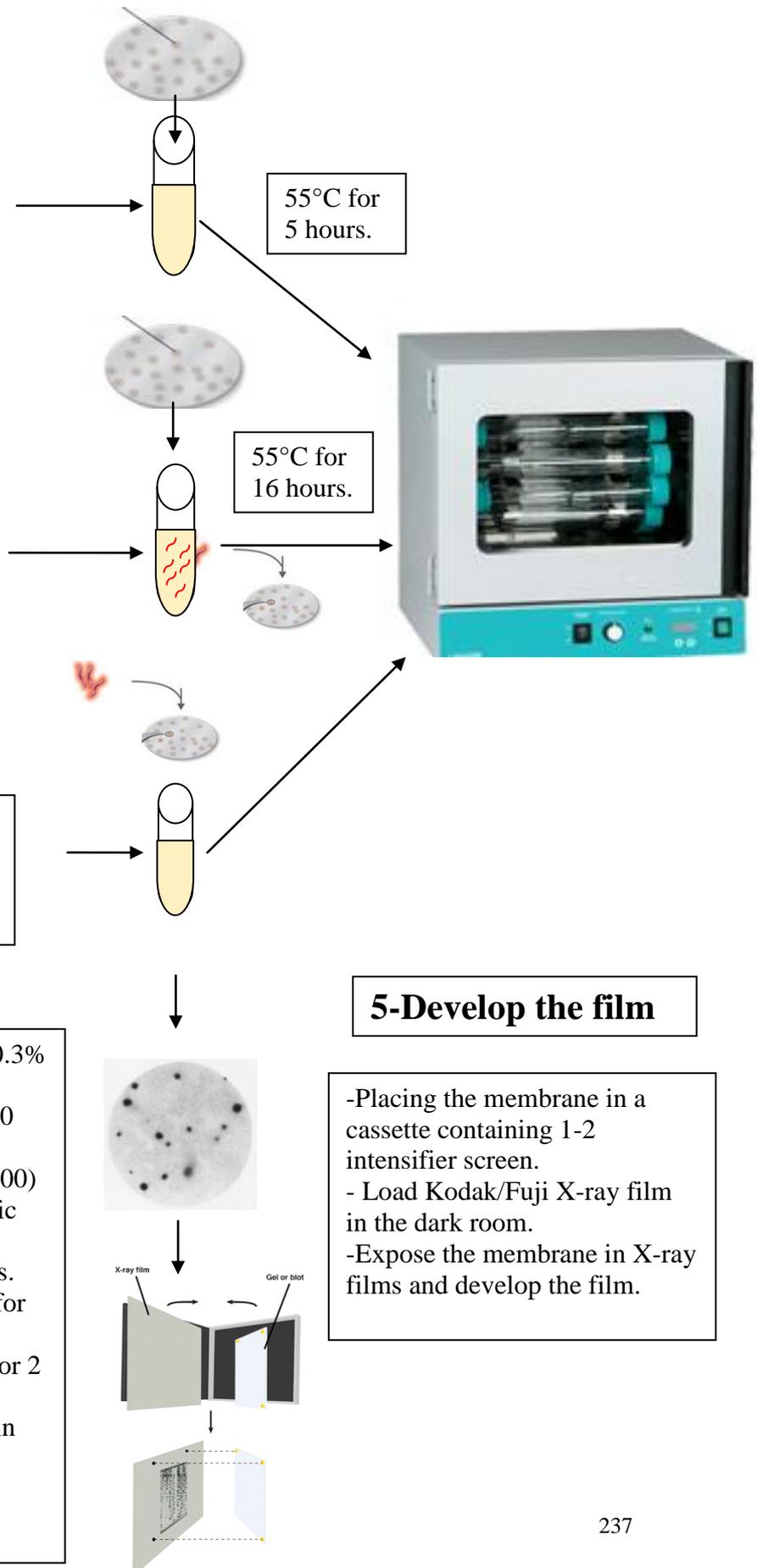
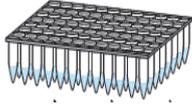
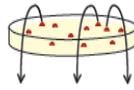
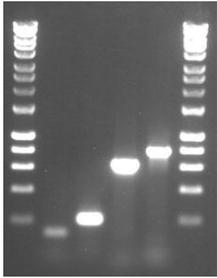


Plate 2:
Colony PCR

-Make master mix for PCR, and use colony as a template DNA.



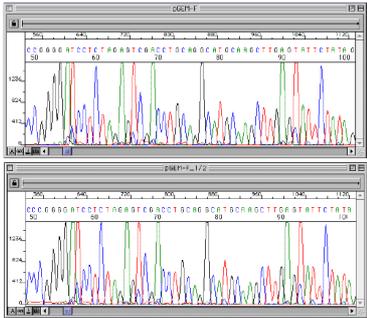
-Gel electrophoresis
-Select colonies for growing on LB media .



- Plasmid extraction
(Confirmation of the insert).



- Sequencing plasmid
-Nucleotide sequence analysis.



-Probe labelling
-Fluorescent *in situ* hybridization.

