

## Article

# Disentangling *Crocus Series Verni* and Its Polyploids

Irena Raca<sup>1</sup>, Frank R. Blattner<sup>2,3</sup>, Nomar Espinosa Waminal<sup>2</sup>, Helmut Kerndorff<sup>4</sup>, Vladimir Randelović<sup>1</sup>  
and Dörte Harpke<sup>2,\*</sup>

<sup>1</sup> Department of Biology and Ecology, University of Niš, 18000 Niš, Serbia

<sup>2</sup> Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), D-06466 Seeland-Gatersleben, Germany

<sup>3</sup> German Centre of Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, D-04103 Leipzig, Germany

<sup>4</sup> São Romão, Cx 548 A, 8150-058 São Brás de Alportel, Portugal

\* Correspondence: harpke@ipk-gatersleben.de

**Simple Summary:** In plants, the occurrence of polyploid lineages, which are plants with multiple instead of two sets of chromosomes, is quite common. Polyploids can originate as autopolyploids within a species or by combining the genomes of different species resulting in allopolyploids. Within the group of spring crocuses, a polyploid complex exists where it is unclear how it evolved and which species eventually contributed to polyploid formation. Among *Crocus* species, evolutionary analyses are further complicated by widely varying chromosome numbers that do not clearly correlate with di- or polyploidy. To reconstruct the evolution of these polyploids, we combine chromosome counts, genome size estimations, phylogenetic analyses based on maternally and bi-parentally inherited genomes, co-ancestry analysis, and morphometric data for all species potentially involved in polyploid formation. Through this approach, we show that polyploids in the *Crocus heuffelianus* group are allopolyploids that originated multiple times involving different parental genotypes and reciprocal crosses. Chromosome numbers partly changed after polyploidization. Numbers found in polyploids are therefore no longer in all cases additive values of their parents' chromosomes. We conclude that in crocuses, only an approach combining evidence from different analysis methods can uncover the evolutionary history of species if polyploidization is involved.



**Citation:** Raca, I.; Blattner, F.R.;

Waminal, N.E.; Kerndorff, H.;

Randelović, V.; Harpke, D.

Disentangling *Crocus Series Verni* and Its Polyploids. *Biology* **2023**, *12*, 303.

[https://doi.org/10.3390/](https://doi.org/10.3390/biology12020303)

[biology12020303](https://doi.org/10.3390/biology12020303)

Academic Editor: Lorenzo Peruzzi

Received: 30 December 2022

Revised: 6 February 2023

Accepted: 7 February 2023

Published: 14 February 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Spring crocuses, the eleven species within *Crocus series Verni* (Iridaceae), consist of di- and tetraploid cytotypes. Among them is a group of polyploids from southeastern Europe with yet-unclear taxonomic affiliation. Crocuses are generally characterized by complex dysploid chromosome number changes, preventing a clear correlation between these numbers and ploidy levels. To reconstruct the evolutionary history of series *Verni* and particularly its polyploid lineages associated with *C. heuffelianus*, we used an approach combining phylogenetic analyses of two chloroplast regions, 14 nuclear single-copy genes plus rDNA spacers, genome-wide genotyping-by-sequencing (GBS) data, and morphometry with ploidy estimations through genome size measurements, analysis of genomic heterozygosity frequencies and co-ancestry, and chromosome number counts. Chromosome numbers varied widely in diploids with  $2n = 8, 10, 12, 14, 16,$  and  $28$  and tetraploid species or cytotypes with  $2n = 16, 18, 20,$  and  $22$  chromosomes. *Crocus longiflorus*, the diploid with the highest chromosome number, possesses the smallest genome ( $2C = 3.21$  pg), while the largest diploid genomes are in a range of  $2C = 7–8$  pg. Tetraploid genomes have  $2C$  values between  $10.88$  pg and  $12.84$  pg. Heterozygosity distribution correlates strongly with genome size classes and allows discernment of di- and tetraploid cytotypes. Our phylogenetic analyses showed that polyploids in the *C. heuffelianus* group are allotetraploids derived from multiple and partly reciprocal crosses involving different genotypes of diploid *C. heuffelianus* ( $2n = 10$ ) and *C. vernus* ( $2n = 8$ ). Dysploid karyotype changes after polyploidization resulted in the tetraploid cytotypes with 20 and 22 chromosomes. The multi-data approach we used here for series *Verni*, combining evidence from nuclear and chloroplast phylogenies, genome sizes, chromosome numbers, and genomic heterozygosity for ploidy estimations, provides a way to disentangle the evolution of plant taxa with complex karyotype changes that can be used for the analysis of other groups within *Crocus* and beyond. Comparing these results with morphometric analysis results in characters that can discern the different taxa currently subsumed under *C. heuffelianus*.

**Keywords:** chromosome numbers; *Crocus heuffelianus* group; *Crocus* series *Verni*; dysploidy; genome size; genotyping-by-sequencing (GBS); morphometry; phylogeny; polyploidy

## 1. Introduction

Polyploidization or whole-genome duplication (WGD) is a common process in plants, resulting in individuals with different ploidy levels. Two major mechanisms are discerned: if WGD happens within a species, the resulting polyploid is termed autopolyploid, whereas hybridization between two different species followed by WGD is termed allopolyploidy [1]. Autopolyploids might suffer at least initially from reduced fertility due to distorted chromosome distribution during meiosis, while allopolyploids usually undergo normal meiosis. However, one has to understand auto- and allopolyploids as the endpoints of a continuum. Chromosome pairing and distribution among daughter cells depends on the overall similarity of the chromosomes. Thus, even within a species, chromosomes can be relatively different or rather similar in an allopolyploid if closely related species are involved. Polyploidization is often a driver of evolution, as doubling of genes releases one of the homeologous copies from purifying selection, allowing it to obtain new functions [2]. This event should be advantageous, particularly in stressful habitats or during changing environmental conditions [3]. Through time, polyploid genomes will accumulate differences not only at the level of allelic differences but, through karyotype changes, also regarding the overall structure of the genomes [1]. This results in diploidization, i.e., former homeologous chromosomes are no longer recognizable as such. Ancient polyploidization events are therefore not easy to detect and most often need in-depth genome analysis to be revealed [2,4]. For more recent WGD events, chromosome numbers and genome sizes can be indicative. While there are taxa where chromosome numbers correlate clearly with ploidy levels and genome size [5,6], dysploid chromosome number changes can blur such correlations through breaking and/or fusion of chromosomes [7]. In addition, downsizing or enlarging of genomes [4,8–10] through the loss of DNA or activation of transposable elements can hinder ploidy level recognition. However, these latter processes are normally acting at a slower pace than changes in chromosome numbers [11,12].

*Crocus* series *Verni* B.Mathew is a group of mostly spring-flowering crocuses from Central and South Europe, some of them being important ornamentals. The series consists of eleven species with unclear phylogenetic relationships, as earlier approaches with molecular markers arrived only at badly resolved species groups and species identification was also partly uncertain [13–15] or sampling incomplete [16]. Chromosome numbers range from  $2n = 8$  to  $2n = 28$  [16–23], often with uncertain ploidy levels, as high chromosome numbers do not necessarily correlate with larger genome sizes and, therefore, higher ploidy level [16]. Particular populations from the Balkan Peninsula, thought to taxonomically belong to *C. heuffelianus* Herb., exhibit highly diverse chromosome numbers with  $2n = 10, 18, 19, 20, 22,$  and  $23$  chromosomes [16,18]. Harpke et al. [15], in their account of series *Verni*, concluded from karyotype analysis that certain populations of *C. heuffelianus* (as well as *C. neglectus* Peruzzi and Carta) might have resulted from polyploidization but were not able to confirm this further. We refer to these potential polyploids throughout this article as “*C. cf. heuffelianus*”, “*C. cf. tommasinianus*”, and “*C. cf. vernus*”, as their taxonomic status is unclear and name changes seem still premature to us. *Crocus neglectus*, in contrast, was already recognized as a separate species [15].

Here we intend to understand the evolution of *Crocus* series *Verni* with particular reference to the origin of the polyploid species and cytotypes. To arrive at this goal, we use molecular phylogenetic approaches based on (a) genotyping-by-sequencing (GBS [24]) to obtain highly informative genome-wide single-nucleotide polymorphisms (SNPs) for a robust phylogeny, (b) nuclear rDNA internal transcribed spacers (ITS) plus 14 single-copy gene sequences for tracing bi-parentally inherited genome parts and chloroplast DNA sequences for inferring maternal lineages within the study group, and (c) morpho-

anatomical analyses to find traits that can discern the taxa, and combine these data with (d) chromosome counts and (e) genome size estimations of diverse populations in order to infer di- and polyploid taxa and cytotypes and reveal their parental contributors and geographic distribution.

## 2. Materials and Methods

### 2.1. Plant Materials

Our study includes plants from 63 populations: 24 *C. heuffelianus*/*C. cf. heuffelianus*, 13 *C. vernus* (L.) Hill/*C. cf. vernus*, nine *C. tommasinianus* Herb., three *C. bertiscensis* Raca, Harpke, Shuka, and V.Randjel., three *C. neapolitanus* (Ker Gawl.) Loisel., two *C. neglectus*, two *C. kosaninii* Pulević, one of each of the other series *Verni* species, two populations of outgroup *C. malyi* Vis., and one ornamental cultivar (Table 1, Supplementary Table S1). To differentiate the  $2n = 18$  karyotypes of *C. cf. heuffelianus*, we labeled them Western Carpathian clade (WCC), Pannonian-Illyric clade (PIC), and Southern Carpathian clade (SCC). The sampling covers the whole distribution area of the series except for the westernmost *C. vernus* populations from the Pyrenees, which could only be included in the chloroplast and nuclear single-copy marker dataset (Table 1). To encumber poaching on the wild populations we here provide only rather general locations for the studied materials instead of populations' GPS coordinates. Chloroplast and nuclear single-copy markers were investigated for a smaller number of representatives of series *Verni* while the GBS analyses were based on the most exhaustive number of individuals (Table 1).

**Table 1.** Materials from *Crocus* ser. *Verni* used in this study (for an extended list see Supplementary Table S1).

Species	Individuals in Chloroplast/ GBS/NSCG <sup>1</sup> /Morpho-Anatomical Analyses	Origin	x/Chromosome Number/Average 2C Genome Size
<i>Crocus bertiscensis</i> Raca, Harpke, Shuka, and V.Randjel.	4/6/1/0	Albania (ALB), Montenegro (MNE)	2x/12/6.66 pg
<i>Crocus etruscus</i> Parl.	2/3/0/0	Italy (ITA)	2x/8/7.58 pg
<i>Crocus ilvensis</i> Peruzzi and Carta	4/5/1/0	Italy (ITA)	2x/8/7.88 pg
<i>Crocus kosaninii</i> Pulević	2/3/2/0	Kosovo (XKX), Serbia (SRB)	2x/14/7.95 pg
<i>Crocus longiflorus</i> Raf.	2/1/2/0	Italy (ITA)	2x/28/3.21 pg
<i>Crocus neglectus</i> Peruzzi and Carta	1/1/1/0	Germany (GER)	4x/16/12.24 pg
<i>Crocus neapolitanus</i> (Ker Gawl.) Loisel.	3/2/1/0	Italy (ITA)	2x/8/-
<i>Crocus siculus</i> Tineo	0/1/0/0	Italy (ITA)	2x/8/-
<i>Crocus tommasinianus</i> Herb.	5/25/2/0	Bosnia and Herzegovina (BIH), Italy (ITA), Montenegro (MNE), Serbia (SRB)	2x/16/5.53 pg
<i>Crocus vernus</i> (L.) Hill	14/27/4/100	Albania (ALB), Bosnia and Herzegovina (BIH), France (FRA), Montenegro (MNE), Slovenia (SLO), Switzerland (CHE)	2x/8/5.78 pg
<i>Crocus heuffelianus</i> Herb.	8/22/4/70	Romania (ROU), Slovakia (SVK), Ukraine (UKR)	2x/10/7.73 pg
<i>Crocus cf. heuffelianus</i> (SCC)	7/19/1/70	Romania (ROU)	4x/18/12.84 pg
<i>Crocus cf. heuffelianus</i> (PIC)	5/26/1/40	Bosnia and Herzegovina (BIH), Slovenia (SLO)	4x/18/10.88 pg
<i>Crocus cf. heuffelianus</i> (WCC)	3/9/2/35	Slovakia (SVK)	4x/18/12.75 pg
<i>Crocus cf. heuffelianus</i>	12/23/0/80	Montenegro (MNE), Serbia (SRB)	4x/20/11.82 pg
<i>Crocus cf. heuffelianus</i>	3/14/1/40	Albania (ALB), Kosovo (XKX)	4x/22/11.95 pg
<i>Crocus cf. vernus</i>	5/9/1/0	Albania (ALB)	4x/16/12.38 pg
<i>Crocus cf. tommasinianus</i>	0/1/0/0	Montenegro (MNE)	4x/-/-
<i>Crocus malyi</i> Vis. (outgroup)	2/2/2/0	Croatia (HRV)	2x/30/-

<sup>1</sup> NSCG = nuclear single-copy genes plus rDNA ITS.

## 2.2. DNA Extraction and Sanger Sequencing

Total genomic DNA was extracted from silica-gel-dried leaf tissue with the DNeasy Plant Mini Kit (Qiagen) according to the instructions of the manufacturer. After DNA extraction, we checked DNA quality and concentration on 1% agarose gels. For the amplification of the two chloroplast regions, we used the primers *matKf*, *rpS16in1\_r*, *rpS16in1\_f*, *trnQr*, *ycf1bF*, and *ycf1bR* [25]. PCR amplification protocols for all markers followed Harpke and Kerndorff [25]. Forward and reverse strands of both regions were directly Sanger sequenced on an ABI 3730 XL using the amplification primers, edited where necessary, and assembled into single sequences in GENEIOUS Prime 2022.1.1 [26]. Afterward, the sequences were aligned using MAFFT version 1.5.0 [27] within GENEIOUS and manually corrected.

## 2.3. Genotyping-by-Sequencing

To obtain genome-wide SNPs, GBS analyses [24] were conducted for 91 di- and 102 tetraploid individuals, with one of the latter included twice as a replicate. For the library preparation, 200 ng of genomic DNA was used and cut with the two restriction enzymes *Pst*I-HF (NEB) and *Msp*I (NEB). Library preparation, individual barcoding, and single-end sequencing on the Illumina NovaSeq were performed following Wendler et al. [28].

Barcoded reads from the 194 samples were de-multiplexed using the CASAVA pipeline 1.8 (Illumina). Adapter trimming of GBS sequence reads was performed with CUTADAPT [29] within IPYRAD v.0.9.58 [30] and reads shorter than 60 bp after adapter removal were discarded. GBS reads were clustered using the IPYRAD 0.7.5 [30] pipeline with a clustering threshold of 0.85. We tested diverse IPYRAD settings but at the end the default settings of parameter files generated with IPYRAD were optimal for the other parameters. We generated one output that included the outgroup *C. malyi*, which was used for phylogenetic analyses, and a second output without *C. malyi*, which was used for principal component analysis (PCA) and population assignment analyses.

## 2.4. Analyses of Population Structure

Principal component analysis (PCA) was conducted in IPYRAD. For model-based Bayesian population assignment analysis we used the R package LEA [31]. Population assignment was performed for  $K = 1-15$  with 20 repetitions each and ploidy set to four. Additionally, FASTSTRUCTURE [32] was used with “simple” as prior for  $K = 1-15$  with 20 repetitions each for cross-validation. The optimal  $K$  was then determined with the function “chooseK” in FASTSTRUCTURE. The Q-matrices obtained with LEA (for  $K = 5$ ,  $K = 8$ ) and FASTSTRUCTURE (for  $K = 4$ ), which include the ancestral assignment frequencies, were sorted using the R package tidyverse [33] and plotted with ggplot2 [34], discerning different ancestral clusters with color-coding. The ggplot2 package was also used for plotting the PCA results.

## 2.5. Heterozygosity and *Fst* Determination

The allelic constitution of SNP positions was checked in the *vcf* file obtained with IPYRAD to infer if and to what extent at these positions more than two alleles for a heterozygous SNP were present in the diploid individuals. As overall only less than 3% of SNPs were not biallelic, DNASP v. 6 [35] was used to infer the heterozygosity as well as the fixation index (*Fst*) of the data based on the *vcf* files generated with IPYRAD for di- and polyploid individuals. The DNASP output was used to calculate the ratio of heterozygous sites to the total number of sites of the samples.

## 2.6. Next-Generation Sequencing of Nuclear Markers

Fourteen nuclear single-copy markers were amplified (see Supplementary Table S2) using the Phusion High Fidelity DNA polymerase (ThermoFisher Scientific). To obtain sequences of the nuclear rDNA internal transcribed spacer region (ITS), we used the primers ITS-A and ITS-B [36] following the protocol of Blattner [37]. PCR products of each amplicon of one sample were pooled together regardless of their concentration, purified

using a NucleoFast plate (Marcherey-Nagel), and finally diluted in 34  $\mu$ L triple-distilled water. Sixteen microliters of the purified amplicon pool were digested using the NEBNext™ dsDNA Fragmentase kit (NEB) for an incubation time of 1.5 min and another 16  $\mu$ L of the amplicon pool was digested for 4.5 min following the manufacturer's instructions (NEB). Both were pooled, 100  $\mu$ L triple distilled water was added, and a NucleoFast plate (Marcherey-Nagel) was used to remove too-small fragments and contaminants. Fifty microliters were subjected to size-selection targeting fragment sizes of 400–600 bp using BluePippin (Sage Science), blunt-end repaired, and used for sequencing library preparation according to the Illumina TruSeq DNA library protocol. Adaptors and barcodes were ligated to the samples. The libraries were size-selected with BluePippin. Fragment size distribution and DNA concentration were evaluated on an Agilent BioAnalyzer High Sensitivity DNA Chip and using the Qubit DNA Assay Kit in a Qubit 2.0 Fluorometer (Life Technologies). Finally, the DNA concentration of the libraries was checked with a quantitative PCR run. Cluster generation on Illumina cBot and paired-end sequencing  $2 \times 250$  bp on the Illumina HiSeq 2000/2500 and NovaSeq6000 platform, respectively, followed Illumina's recommendation and included 1% Illumina PhiX library as internal control. The targeted output per sample was 300,000 reads. Reads were initially iteratively mapped against the forward primer as reference or already existing sequences of the marker in GENEIOUS Prime 2022.1.1 using the GENEIOUS mapping tool. In a second step, the pre-defined reads were assembled into haplotypes, i.e., representing the different alleles present in the sequenced marker region. Finally, three out of 14 nuclear single-copy markers (*orcf*, *rcf2*, *topo6*) were informative and were used for the investigation of haplotype differences of di- and polyploids to determine the parental species involved in allotetraploid formation.

### 2.7. Chromosome Counts

Chromosome counts were either obtained from the literature, obtained from direct observations in this study, or extrapolated for a few individuals based on the genome size data together with published chromosome counts of nearby populations. For direct observations, roots were cut about 2 cm from the tips, pretreated with 2 mM 8-hydroxyquinoline for 5 h at room temperature, and then kept in cold water overnight in a refrigerator. The roots were fixed in Carnoy's solution (3:1 ethanol:acetic acid) for 24 h and stored in 70% ethanol until use. Slide preparation was carried out according to Waminal et al. [38] and Rodríguez-Domínguez et al. [39]. Slides were fixed in 2% formaldehyde solution (47608, Sigma-Aldrich) for 3 min and dehydrated in an ethanol series (70%, 90%, 99%). Chromosomes were stained with 1  $\mu$ g/ $\mu$ L DAPI in  $2 \times$  SSC. Images were captured using a  $100 \times$  objective of an Olympus BX61 fluorescence microscope (Olympus).

### 2.8. Genome Size Measurements

Genome sizes were measured for 134 individuals. Due to a lack of material, we could not measure the genome sizes of *C. neapolitanus* and *C. siculus*. Genome sizes for *C. bertiscensis* were partially taken from Raca et al. [16].

Genome size was determined using propidium iodide (PI) as a stain in flow cytometry with a Cyflow Space (Sysmex Partec) flow cytometer, following essentially the procedure described in Jakob et al. [6]. We mainly used rye (*Secale cereale*; 16.01 pg/2C) or pea (*Pisum sativum*; 9.09 pg/2C) as internal size standards and the buffer CyStain PI Absolute P (Sysmex Partec). Genome size measurements aimed at identifying diploids and polyploids. To link the genome sizes with the molecular data, we used silica-gel-dried leaves from the same individual used for DNA extraction whenever possible. Initial tests showed that fresh and silica-gel-dried materials arrived at the same genome size estimations in *Crocus*. However, the quality of data obtained is slightly lower for dried leaves. A detailed overview of the material measured and standards used is given in Supplementary Table S3.

### 2.9. Phylogenetic Inferences and Origin of Allopolyploids

Maximum parsimony (MP) analyses were conducted in PAUP\* 4.0a169 [40] using a two-step heuristic search as described in Blattner [37] with 1000 initial random addition sequences (RAS) restricting the search to 25 trees per replicate. The resulting trees were afterwards used as starting trees in a search with maxtree set to 10,000. To test clade support, bootstrap analyses were run on all datasets with re-sampling 1000 times with the same settings as before, except that we did not use the initial RAS step. PAUP\* was also used to infer the best-fitting model of sequence evolution for the sequence datasets using the Bayesian information criterion (Table 2). Analyses were run for (a) a dataset consisting of the combined sequences of the two chloroplast marker regions *matK-trnQ* and *ycf1*; for the GBS-derived sequences including (b) the diploid species and (c) the diploid plus all tetraploid individuals of the *C. heuffelianus/C. vernus* complex plus *C. neglectus*, the only tetraploid species formally recognized to date; (d) for three nuclear single-copy genes; and (e) for the rDNA ITS region. The nuclear single-copy and ITS datasets were used to identify homeotic alleles in tetraploids and their diploid parents instead of deriving detailed phylogenetic relationships. The MP analysis of GBS data derived from di- and tetraploid individuals was used to infer the closest diploid parent of allopolyploids and to see if allopolyploids are monophyletic or originated multiple times.

**Table 2.** Characteristics of the analyzed molecular datasets for *Crocus* ser. *Verni*.

	Chloroplast ( <i>matK-trnQ</i> + <i>ycf1</i> )	GBS Data 2x/2x + 4x	Nuclear Gene Regions <i>orcp/rcf2/topo6/ITS</i>
Number of sequences	81	93/194	37/42/39/22
Alignment lengths	4123	187,846	1160/568/819/646
Constant characters	3989	178,748/176,958	982/490/737/599
Variable characters	134	9098/10,888	178/78/82/47
Parsimony-informative characters	114	5899/6954	125/47/38/33
Number of MP trees	4	45/5400	142/10k/10k/10k <sup>1</sup>
MP tree length	156	16,298/25,300	333/99/89/56
Consistency index (CI)	0.87	0.58/0.45	0.57/0.83/0.96/0.88
Retention index (RI)	0.98	0.83/0.77	0.66/0.88/0.97/0.93
Model of sequence evolution (BIC)	GTR + I	GTR + I + $\Gamma$	HKY + I + $\Gamma$ /HKY + I + $\Gamma$ /TrN + I/HKY + I

<sup>1</sup> 10,000 trees were set as maxtree in MP analyses.

For the GBS-derived data, we also calculated SVDquartets in PAUP\*, evaluating all quartets for the dataset consisting of 92 diploid individuals running 500 bootstrap re-samples. Individuals were partitioned according to their species affiliation, and trees were selected using QFM quartet assembly and the multispecies coalescent (MSC) as tree model.

Bayesian phylogenetic inference (BI) was conducted in MRBAYES 3.2.7 [41] for the chloroplast, the nuclear single-copy genes, and the ITS datasets. In BI, two times four chains were run for 5 million generations specifying the respective model of sequence evolution. A tree was sampled every 500 generations. Converging log-likelihoods, potential scale reduction factors for each parameter, and inspection of tabulated model parameters in MRBAYES suggested that stationary had been reached in all cases. The first 25% of trees of each run were discarded as burn-in.

In addition to the phylogeny-based inference of parental species of allotetraploids we also used STACKS v2.55 [42] to generate an input file for RADPAINTER [43]. A locus needed to be present in 80% of the individuals of a population and in 50% of all individuals to be processed. Population structure was inferred using 10,000 burn-in steps in the Monte Carlo Markov Chain (MCMC) analysis with 100,000 further iterations, keeping every 1000th sample. This run was continued, adding an additional 100,000 steps, treating the original run as burn-in. To obtain the best posterior state for the tree, 1000 attempts were used. Results were visualized in R with the functions provided with RADPAINTER.

The GBS-related Supplementary Datasets S1–S4 are available online through the e!DAL PGP data repository (<https://doi.org/10.5447/ipk/2023/5>, accessed on 10 February 2022).

### 2.10. Morpho-Anatomical Analyses

The morphological analysis was performed on fresh material, including 435 individuals in total (*C. vernus*: five populations; *C. heuffelianus*: two populations; mixed *C. heuffelianus*/*C. cf. heuffelianus*  $2n = 18$  SCC: two populations; *C. cf. heuffelianus*  $2n = 18$  SCC: three populations; *C. cf. heuffelianus*  $2n = 18$  WCC: two populations; *C. cf. heuffelianus*  $2n = 18$  PIC: two populations; *C. cf. heuffelianus*  $2n = 20$ : four populations; *C. cf. heuffelianus*  $2n = 22$ : two populations). Leaf cross-sections were made using a manual microtome [44]. The leaf sections of all 435 individuals were double-stained with safranin (1 g of safranin dissolved in 100 mL of 50% ethanol) and alcian blue (1 g of dye dissolved in 100 mL of distilled water, with a couple of crystals of phenol and three drops of glacial acetic acid). Stained sections were then dehydrated through an alcohol series (50%, 70%, 96%, 100%), examined, and photographed with a Leica DM 1000 microscope (Leica Microsystems) [16,45,46]. Anatomical features were measured in IMAGEJ [47]. A list of 42 characters from the literature relevant for this group [46,48,49] related to morphology and leaf anatomy was taken into consideration (see Supplementary Table S4). The qualitative characters were standardized as states represented by numbers (see Supplementary Table S4).

Principal component (PCA) and discriminant (CDA) analyses for morpho-anatomical characters were computed using STATISTICA 7.0 [50]. Due to unbalanced sample numbers (two populations of diploid vs. 13 populations of polyploid *C. cf. heuffelianus*), the set of differential characters was defined based on representative populations for each taxon/cytotype derived from type localities (140 individuals; Supplementary Table S4), computing a PCA. PCA served as a tool to point out the significant traits. The characters highlighted as important with the PCA were furthermore used for CDA. Moreover, eight a priori-defined groups of parental species and polyploids were included in CDA: *C. vernus* (100 individuals); *C. heuffelianus* (40); mixed *C. heuffelianus*/*C. cf. heuffelianus*  $2n = 18$  SCC (40); *C. cf. heuffelianus*  $2n = 18$  SCC (60); *C. cf. heuffelianus*  $2n = 18$  WCC (35); *C. cf. heuffelianus*  $2n = 18$  PIC (40); *C. cf. heuffelianus*  $2n = 20$  (80); *C. cf. heuffelianus*  $2n = 22$  (40) (435 individuals; Supplementary Table S4). The CDA results were plotted with ggplot2 [34].

## 3. Results

### 3.1. Determination of Ploidy Levels

**Chromosome counts**—Most chromosome counts obtained in this study coincided with previous reports (Supplementary Table S3). The lowest chromosome count was in *C. etruscus*, *C. ilvensis*, and *C. vernus* ( $2n = 8$ ) followed by *C. heuffelianus* ( $2n = 10$ ); *C. bertiscensis* ( $2n = 12$ ); *C. kosaninii* ( $2n = 14$ ); *C. neglectus*, *C. tommasinianus*, and *C. cf. vernus* ( $2n = 16$ ); PIC, SCC, and WCC populations of *C. cf. heuffelianus* ( $2n = 18$ ); *C. cf. heuffelianus* from Montenegro and Serbia ( $2n = 20$ ); *C. cf. heuffelianus* from Albania and Kosovo ( $2n = 22$ ); and *C. longiflorus* ( $2n = 28$ ). The chromosome count of a *C. vernus*-derived polyploid population from Central Albania, herein referred to as *C. cf. vernus* ( $2n = 16$ ), is reported here for the first time (Supplementary Table S3, Supplementary Figure S1).

**Genome sizes**—The smallest genome size was observed in *C. longiflorus* ( $2C = 3.21$  pg), followed by *C. tommasinianus* ( $2C = 5.53$  pg) and *C. vernus* ( $2C = 5.78$  pg). *Crocus bertiscensis* was observed with an average genome size of  $2C = 6.66$  pg. *Crocus etruscus* ( $2C = 7.58$  pg), *C. ilvensis* ( $2C = 7.88$  pg), *C. heuffelianus* ( $2C = 7.73$  pg), and *C. kosaninii* ( $2C = 7.95$  pg) had similar genome sizes (Supplementary Table S3). Populations of *C. heuffelianus* with higher chromosome counts ( $2n = 18, 20, 22$ ) were measured with average genome sizes per cytotype ranging between  $2C = 10.88$  and  $2C = 12.84$  pg (Table 1). Similar genome sizes were measured for *C. neglectus* ( $2C = 12.24$  pg) and *C. cf. vernus* ( $2C = 12.23$  pg).

Plotting of genome sizes against the chromosome counts showed a general negative relationship between genome size and chromosome number in both diploid and tetraploid taxa. This relationship is weaker when *C. longiflorus* is excluded (Supplementary Figure S1).

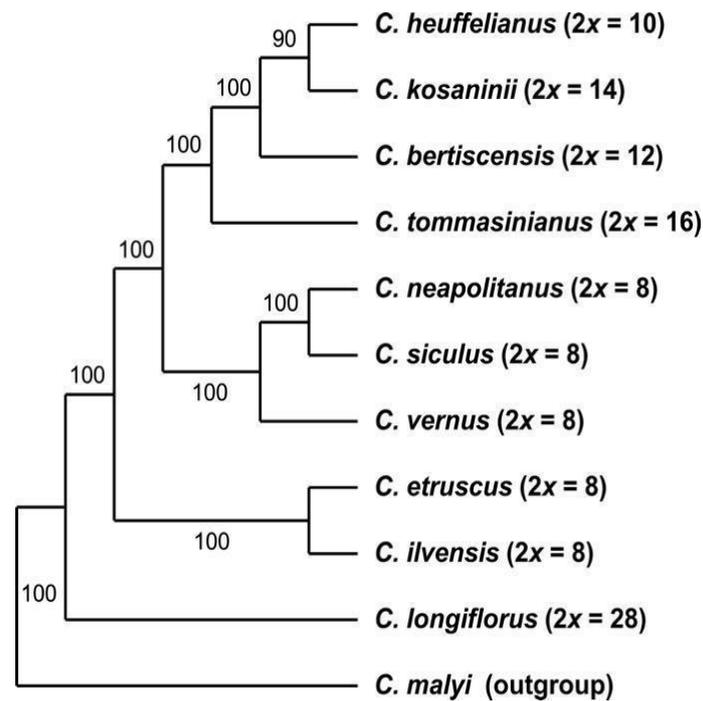
*Crocus longiflorus* is sister to the core series *Verni* taxa and has a seemingly polyploid chromosome count ( $2n = 28$ ), although genome size indicates a diploid genome (Supplementary Table S3).

**GBS-derived heterozygosity to ploidy**—Most SNPs in the datasets were bi-allelic even in polyploids. The heterozygosity  $H_0$  ranged between 0.0115 and 0.0628 for the GBS data set used (Supplementary Table S5). Individuals can be divided into two groups: samples with a  $H_0$  of 0.0115 to 0.0278 and samples with a  $H_0$  of 0.0359 to 0.0628 (Supplementary Table S5). If heterozygosity is considered in the context of known genome sizes, the group with the higher  $H_0$  possesses the larger genome sizes ( $2C = >10$  pg). The group with lower  $H_0$  also has smaller genome sizes ( $2C = <8$  pg) and generally lower chromosome numbers (except for *C. longiflorus*) and comprises *C. vernus* ( $2C = 5.78$  pg,  $2n = 8$ ), *C. etruscus* ( $2C = 7.58$  pg,  $2n = 8$ ), *C. ilvensis* ( $2C = 7.88$  pg,  $2n = 8$ ), *C. heuffelianus* ( $2C = 7.73$  pg,  $2n = 10$ ), *C. bertiscensis* ( $2C = 6.66$  pg,  $2n = 12$ ), *C. kosaninii* ( $2C = 7.95$  pg,  $2n = 14$ ), *C. tommasinianus* ( $2C = 5.53$  pg,  $2n = 16$ ), and *C. longiflorus* ( $2C = 3.21$  pg,  $2n = 28$ ). As a consequence, all samples with a  $H_0$  below 0.03 are considered to represent diploids, while all samples with a  $H_0$  above 0.035 are considered polyploids. The latter group includes *C. cf. heuffelianus* with higher chromosome counts and genome sizes ( $2C = 10.88$  to  $12.84$  pg,  $2n = 18, 20, 22$ ), *C. neglectus* ( $2C = 12.24$  pg,  $2n = 16$ ), and *C. cf. vernus* ( $2C = 12.23$  pg,  $2n = 16$ ).

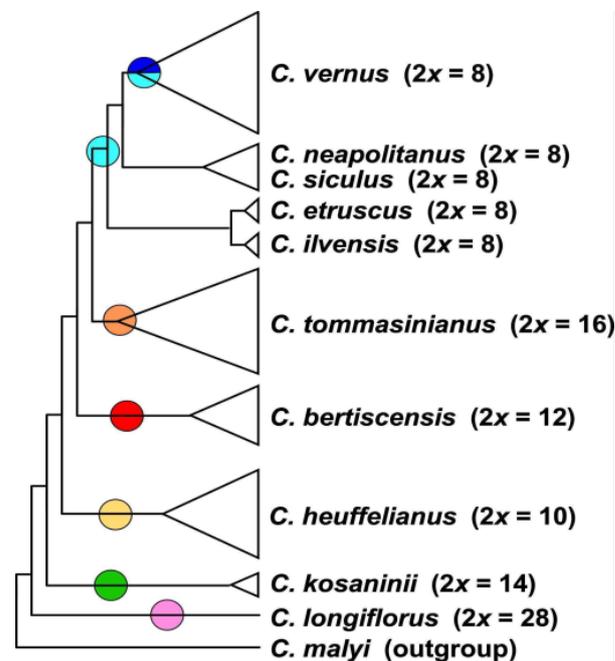
### 3.2. Phylogenetic Inference and Origin of Allopolyploids

**GBS-derived data**—Initially, we created a single dataset for all GBS-derived analyses in IPYRAD, i.e., including all 194 sequences in one alignment (2009 loci, 187,846 bp alignment length, 20.88% missing sites). From this, we derived the dataset that includes only the diploid accessions by excluding all polyploid individuals from the data matrix (Table 2). The SVDquartets analysis of the diploid species (Figure 1) using the multispecies coalescent was used to infer the phylogenetic relationships among the diploid species, which are the basic taxonomic units in this group. Here *C. longiflorus* ( $2n = 2x = 28$ ) from Sicily is the sister species to all other taxa within series *Verni*. The next two consecutively branching clades consist of the species with  $2n = 2x = 8$  with *C. ilvensis* grouping together with *C. etruscus* followed by the clade of *C. neapolitanus*, *C. siculus*, and *C. vernus*. These species all occur in Italy or the Alps. The latter group is sister to a clade that harbors the eastern species *C. heuffelianus*, *C. bertiscensis*, *C. kosaninii*, and *C. tommasinianus* with the higher chromosome numbers of  $2n = 2x = 10, 12, 14$ , and  $16$ , respectively.

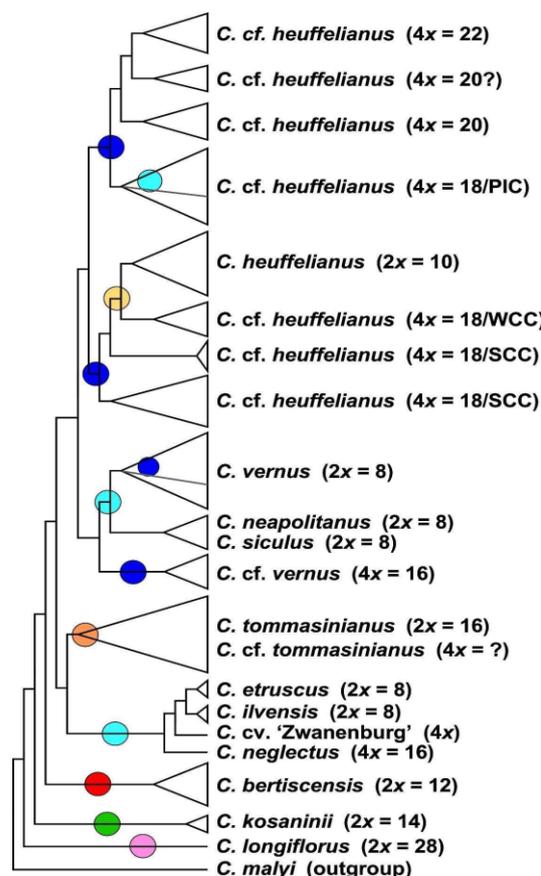
The MP analysis of diploid taxa (Figure 2, Supplementary Figure S2) differs strongly from the diploid's SVDquartets tree topology. *Crocus longiflorus* is in both analyses sister to all other series *Verni* species. In the next clade the positions of the  $2n = 8$  and  $2n \geq 10$  taxa are, however, reversed. Although the species in the tree received very high bootstrap support, for the clades along the backbone of the tree, support values are low (Supplementary Figure S2), so that the topology of this tree has no strong support. We used MP mainly to infer the topology of the polyploids in relation to their diploid progenitors. In MP allopolyploids mostly group within or as sister to the parental species where they share higher genetic similarity and, as MP is sensitive to reticulate data structure, indicates if a polyploid is monophyletic or might consist of different subgroups. In the analysis of the combined di- and tetraploid GBS data (Figure 3, Supplementary Figure S3), *C. longiflorus*, *C. kosaninii*, and *C. bertiscensis* are the first species branching off, with the latter being sister to a large clade consisting of three subclades: (a) *C. etruscus*, *C. ilvensis*, and *C. neglectus* being sister to *C. tommasinianus*, (b) *C. neapolitanus*, *C. siculus*, and *C. vernus* as sister group of (c) *C. heuffelianus*. The polyploids were grouped in between the diploid taxa (Figure 3, Supplementary Figure S3).



**Figure 1.** Phylogenetic tree of diploid species and accessions of *Crocus* ser. *Verni* based on the GBS dataset analyzed with SVDquartets analysis using the multispecies coalescent as tree model. Numbers along branches indicate bootstrap support values; chromosome numbers are provided in brackets.



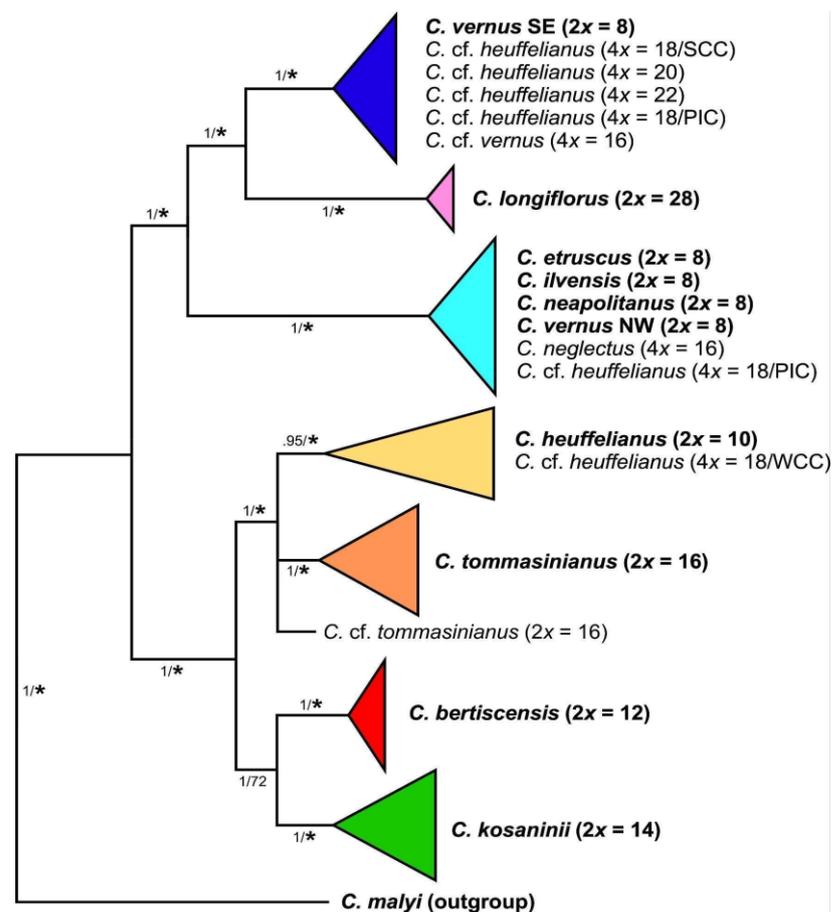
**Figure 2.** Schematic representation of the strict consensus tree topology of 45 most parsimonious trees derived from an MP analysis of the GBS dataset including only diploid accessions of *Crocus* ser. *Verni* (for original data see Supplementary Figure S2). In brackets, the ploidy levels and chromosome numbers are given. Colored circles refer to the chloroplast types present in the respective clades (below).



**Figure 3.** Schematic representation of the strict consensus tree topology of 5400 most parsimonious trees derived from an MP analysis of the GBS dataset including di- and tetraploid accessions of *Crocus* ser. *Verni* (for original data see Supplementary Figure S3). In brackets the ploidy levels and chromosome numbers are given. For tetraploids with  $4x = 18$  the geographic affiliation is also provided, as Pannonian-Illyric Clade (PIC), Southern Carpathian Clade (SCC), and Western Carpathian Clade (WCC). Colored circles refer to the chloroplast types present in the respective clades (below).

**Chloroplast-derived data**—Series *Verni* diploids are split into two major groups in the phylogenetic tree of combined chloroplast data. The first group comprises *C. heuffelianus* ( $2n = 2x = 10$ ), *C. bertiscensis* ( $2n = 2x = 12$ ), *C. kosaninii* ( $2n = 2x = 14$ ), and *C. tommasinianus* ( $2n = 2x = 16$ ) (Figure 4, Supplementary Figure S4). The second group consists of two subgroups: (a) a group comprising the  $2n = 2x = 8$  species *C. etruscus*, *C. ilvensis*, *C. neapolitanus*, *C. siculus*, and *C. vernus* from its northern and western distribution range (Alps to Pyrenees, NW type), and (b) *C. vernus* from the southeastern distribution range (Dinaric Alps; SE type), which has *C. longiflorus* ( $2n = 2x = 28$ ) as its sister group.

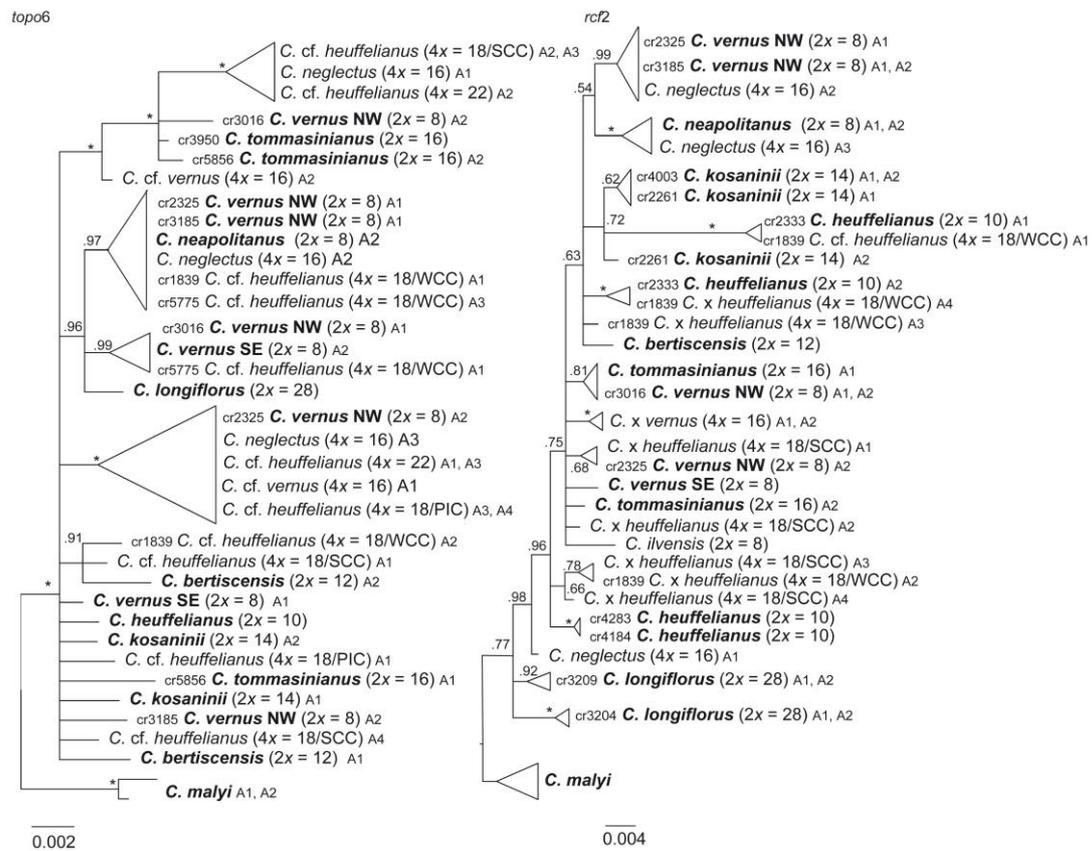
The allotetraploids were grouped with their maternal parents. *Crocus neglectus* ( $2n = 4x = 16$ ) possesses a chloroplast similar to *C. ilvensis*. *Crocus* cf. *heuffelianus* (WCC;  $2n = 4x = 18$ ) from Slovakia groups with *C. heuffelianus*. Most of the *C. cf. heuffelianus* (PIC;  $2n = 4x = 18$ ) individuals from Bosnia and Herzegovina as well as Slovenia grouped in a clade with the western  $2n = 2x = 8$  species and are partly identical with the NW chloroplast type of *C. vernus* (e.g., *C. vernus* from Slovenian Alps is identical with *C. cf. heuffelianus* from Bosnia and Herzegovina). However, we also found *C. cf. heuffelianus* PIC grouping with the SE chloroplast type of *C. vernus*, indicating an independent origin. *Crocus* cf. *heuffelianus* (SCC;  $2n = 4x = 18$ ) from Romania,  $2n = 4x = 20$  from Montenegro and Serbia,  $2n = 4x = 22$  from Northern Albania and Kosovo, as well as *C. cf. vernus* ( $2n = 4x = 16$ ) from Central Albania were found in the clade with the southeastern *C. vernus* populations (SE type).



**Figure 4.** Schematic representation of relationships of species and cytotypes of *Crocus* ser. *Verni* obtained through Bayesian phylogenetic inference of sequences from two chloroplast regions (for original data see Supplementary Figure S4). Diploid accessions, as basic units in the series, are given in boldface. Numbers along branches indicate Bayesian posterior probabilities/maximum parsimony bootstrap values ( $\geq 50\%$ ), with asterisks for bootstrap support values  $> 80\%$ .

**Nuclear single-copy markers**—Three variable nuclear single-copy regions (*orcp*, *rcf2*, *topo6*; Supplementary Table S2) were chosen to identify the position of alleles of the allopolyploids by the criterion that the marker regions showed differences between the potential diploid parental species. Up to four alleles could be found within the allopolyploids (Figure 5, Supplementary Figures S5–S7). *Crocus* cf. *heuffelianus*  $2n = 4x = 18$  (WCC) and *C. cf. heuffelianus*  $2n = 4x = 18$  (PIC) were found to be grouping with *C. vernus* from its western distribution range and/or with *C. heuffelianus*. *Crocus* cf. *heuffelianus*  $2n = 4x = 18$  (SCC) shared similar alleles with *C. vernus* from its eastern distribution range and/or *C. heuffelianus*. The same applies for *C. cf. heuffelianus*  $2n = 4x = 20$  and  $2n = 4x = 22$ . *Crocus neglectus* ( $2n = 4x = 16$ ) grouped with *C. ilvensis* or within a clade comprising *C. neapolitanus* as well as other allotetraploids in two of the selected nuclear single-copy genes. Generally, the gene-tree topologies were different from the topologies in chloroplast or GBS-derived datasets and differences also occurred among the single-copy genes. For example, in *topo6*, *C. longiflorus* occupies a similar position as in the chloroplast marker tree where it groups in one clade with other Italian taxa such as *C. neapolitanus*, *C. ilvensis*, and *C. vernus* (Figure 5). In the *rcf2*-derived tree its position is similar to the GBS results, where it groups as sister to the core series *Verni* taxa. *Crocus tommasinianus*, or one of its alleles, was found in one clade with *C. vernus* (NW) in *topo6* and *rcf2* (Figure 5), while it groups with *C. bertiscensis*, *C. kosaninii*, and *C. heuffelianus* in the SVDquartets of GBS and the chloroplast trees. The *orcp* data set had the highest number of parsimony-informative sites of the three closely

examined nuclear single-copy genes, but also the highest homoplasmy and was mostly characterized by polytomies (Table 2, Supplementary Figure S7).



**Figure 5.** Phylogenetic trees obtained through Bayesian phylogenetic inference based on haplotype sequences of nuclear single-copy genes in *Crocus* ser. *Verni*. Numbers along branches indicate BI posterior probabilities (pp). Support values of 1.0 are indicated by asterisks. Allelic differences (A1–A4) in these markers were used to track the bi-parental contributions of diploids to allotetraploids. If more than one individual per species or cytotype was included, their DNA identifier (cr number) is provided. For detailed information, see Supplementary Figures S5 and S6.

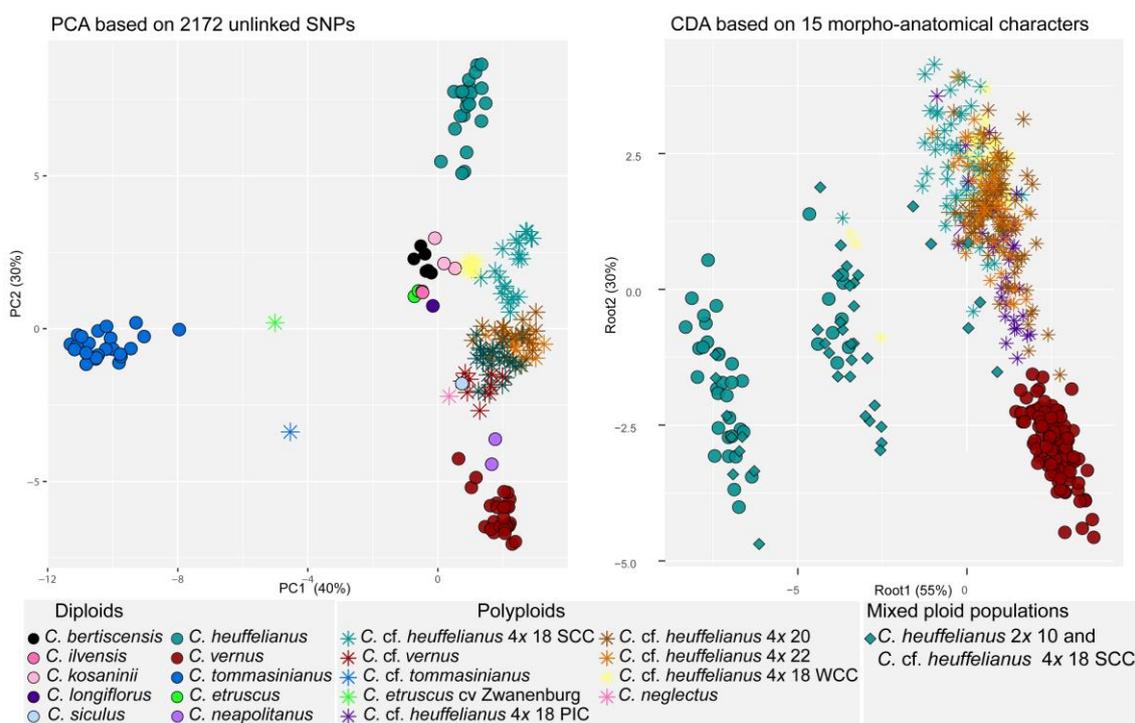
**ITS**—In the ITS tree (Supplementary Figure S8) *C. longiflorus* is sister to the other series *Verni* taxa (pp 1.0). In the sister clade of *C. longiflorus*, *C. kosaninii* is separated from the remaining species of the series but with very low support (pp 0.61). The majority of series *Verni* species is found in a large polytomy with only a few subclades. In one of the subclades, *C. ilvensis* groups with *C. etruscus*, while most of the other subclades are formed by samples of the same species, with the exception of the subclade comprising *C. vernus*. Here, *C. neapolitanus*, *C. siculus*, *C. vernus*, and all allotetraploids included in the dataset can be found (pp 1.0; Supplementary Figure S8), albeit their relationships remained unresolved.

### 3.3. Phylogenomic Analysis

The 192 GBS sequences of series *Verni* (excluding the outgroup *C. malyi*) with a threshold number of 120 samples sharing a locus resulted in a dataset comprising 2207 loci with 18.38% missing sites. The data matrix of unlinked SNPs included 2172 SNPs. The sample with the lowest number of reads (559,922) and loci (866) was *C. siculus*.

The GBS-based PCA (Figure 6) placed *C. tommasinianus* clearly separate from all other samples in the negative part of the PC1 axis. In between *C. tommasinianus* and the majority of all other taxa, in the positive part of the PC1, *C. heuffelianus* and *C. vernus* can be found. However, *C. heuffelianus* in the positive part of the PC2 axis (5 or higher) and *C. vernus*

in the negative part of the PC2 axis (−4 or lower) were distinct from each other. The two individuals of *C. neapolitanus* were close to *C. vernus*. *Crocus etruscus* and *C. ilvensis* were placed together in close proximity to *C. bertiscensis*, *C. kosaninii*, and *C. longiflorus* representatives, all between −1 and 1 on the PC1 axis and 0 to 4 on the PC2 axis. *Crocus siculus* was found in the lower negative part of the PC2 axis (ca. −2), partly overlapping with the polyploids *C. cf. heuffelianus*, *C. neglectus*, and *C. cf. vernus*. They were placed in between *C. heuffelianus* and *C. vernus* but always in the positive part of the PC1 axis.



**Figure 6.** PCA based on 2127 unlinked GBS-derived SNPs (left panel) comprising all species of *Crocus* ser. *Verni*, and CDA based on 15 morpho-anatomical characters (right panel) of *C. vernus*, *C. heuffelianus*, and *C. cf. heuffelianus*.

Co-ancestry analysis with RADPAINTER showed admixture for *C. neglectus* (Supplementary Figure S9) with both *C. etruscus* (co-ancestry 234.2, 241.8) and *C. ilvensis* (co-ancestry 217.9, 221.8) as well as with *C. neapolitanus* (co-ancestry 99.2, 103.7). *Crocus etruscus* cv. ‘Zwanenburg’ shares a relatively high co-ancestry with *C. etruscus* (380.6, 386.3) followed by *C. ilvensis* (301.3, 301.1) and *C. tommasinianus* from Italy (110.5). *Crocus cf. tommasinianus* was found to be admixed with *C. vernus* (181.7) and *C. tommasinianus* (135.1) growing at the same location (Montenegro, Mt. Lovcen). The highest co-ancestry of *C. cf. heuffelianus*  $2n = 4x = 18$  (PIC) and diploid series *Verni* species was found with *C. vernus* (52.8–86.9), while it was lower with *C. heuffelianus* (44.0–51.6). The level of co-ancestry for *C. cf. heuffelianus*  $2n = 4x = 18$  (SCC) was high in the mixed-ploidy populations with *C. heuffelianus* (127.5–128.8), while it was between 50 and 101.5 with other *C. heuffelianus* populations and between 49.2 and 59.4 with *C. vernus*. *Crocus cf. heuffelianus*  $2n = 4x = 18$  from the Western Carpathians (WCC) had its highest co-ancestry level with *C. heuffelianus* (49.4–93.4). Its co-ancestry shared with *C. vernus* ranged between 51.6 and 62.0. The co-ancestry levels of *C. cf. heuffelianus*  $2n = 4x = 20$  and  $2n = 4x = 22$  shared with *C. vernus* were lower than in the other allotetraploid *C. cf. heuffelianus* (48.0–58.2 and 47.8–58.5). The same applies to the shared co-ancestry with *C. heuffelianus* (46.1–54.0 and 45.9–52.2). In addition, they are even lower than the co-ancestry shared between diploid species such as *C. bertiscensis* and *C. kosaninii* (57.1–62.9).

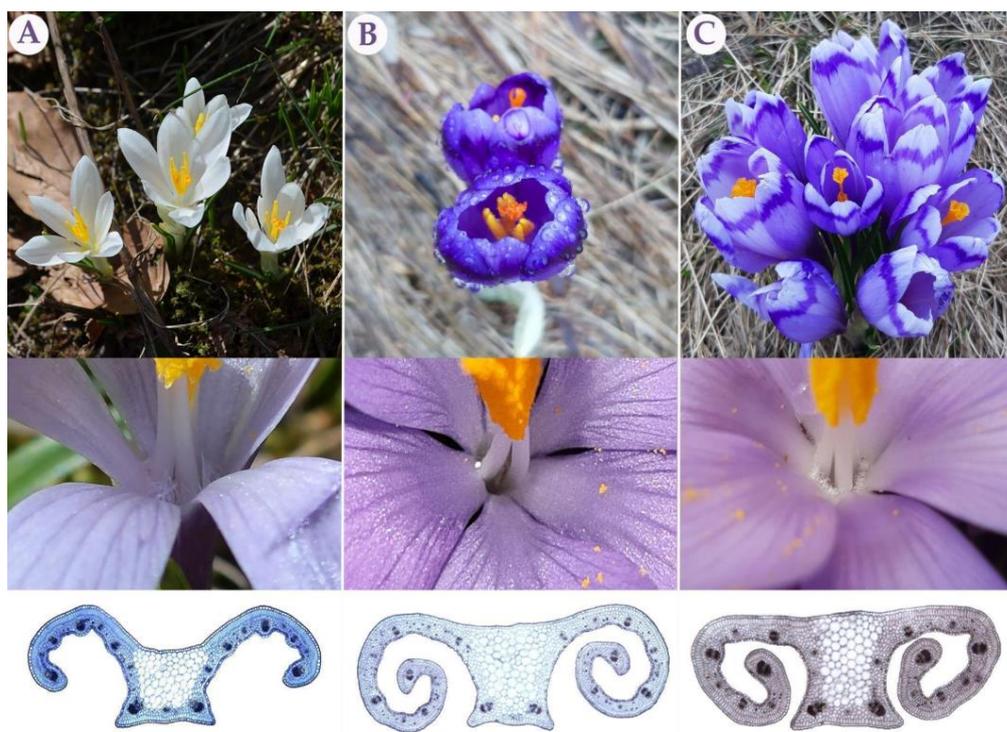
In the population assignment analysis (K with the lowest entropy was  $K = 8$ ), most of the alleles present in *C. tommasinianus* samples were assigned to one ancestral population (Supplementary Figure S10). Moreover, *C. vernus* alleles were mostly assigned to one ancestral population, and representatives of *C. heuffelianus* were mostly assigned to their own ancestral population as well. However, *C. bertiscensis*, *C. etruscus*, *C. ilvensis*, *C. kosaninii*, and *C. longiflorus* appeared admixed, partly sharing *C. heuffelianus* and *C. vernus* (*C. etruscus*, *C. ilvensis*, and *C. longiflorus*) patterns or additionally having alleles assigned to *C. tommasinianus* and/or to *C. cf. heuffelianus*. The different *C. cf. heuffelianus* groups were partly assigned to their own ancestral population showing no admixture with LEA ( $K = 5$ ,  $K = 8$ ) or to *C. vernus* ( $K = 4$ ) with FASTSTRUCTURE. *Crocus cf. vernus* and *C. neglectus* were partly assigned to the *C. vernus* ancestral population but also to different ancestral populations of *C. cf. heuffelianus*. *Crocus etruscus* cv. ‘Zwanenburg’ and *C. cf. tommasinianus* showed admixture, with parts of their alleles derived from *C. tommasinianus*. In the case of *C. cf. tommasinianus*, *C. vernus* contributed genomic materials, and *C. etruscus* cv. ‘Zwanenburg’ was complemented by *C. etruscus*.

Assigning the several polyploid groups mostly to one ancestral population while showing diploids as admixed was also observed for lower K or other SNP subsampling and/or other ancestral assignment methods (Supplementary Figures S9 and S10) regardless of the analysis program used.

*Crocus cf. heuffelianus* had the lowest *Fst* with *C. heuffelianus* and *C. vernus* ( $2n = 4x = 18$  WCC: *Fst* = 0.24 and *Fst* = 0.22;  $2n = 4x = 18$  SCC: *Fst* = 0.14 and *Fst* = 0.17;  $2n = 4x = 18$  PIC: *Fst* = 0.17 and *Fst* = 0.12;  $2n = 4x = 20$ : *Fst* = 0.18 and *Fst* = 0.15;  $2n = 4x = 22$ : *Fst* = 0.22 and *Fst* = 0.19) or, in cases of *C. cf. heuffelianus* ( $2n = 4x = 18$  PIC), with *C. neapolitanus* (*Fst* = 0.15; Supplementary Table S6). The lowest *Fst* for *C. cf. vernus* was observed towards *C. vernus* (*Fst* = 0.15), followed by *C. neapolitanus* (*Fst* = 0.20) and *C. heuffelianus* (*Fst* = 0.23). *Crocus neglectus* had the lowest *Fst* towards *C. etruscus* (*Fst* = 0.07), *C. vernus* (*Fst* = 0.08), *C. neapolitanus* (*Fst* = 0.10), and *C. ilvensis* (*Fst* = 0.10). The cultivar ‘Zwanenburg’ had the lowest *Fst* (0.00) with *C. etruscus*, followed by *C. ilvensis* (*Fst* = 0.09) and *C. tommasinianus* (*Fst* = 0.14). The lowest *Fst* for *C. cf. tommasinianus* was found with *C. tommasinianus* (*Fst* = 0.14) and *C. vernus* (*Fst* = 0.13).

### 3.4. Morpho-Anatomical Analyses

The PCA of the morpho-anatomical dataset highlighted 14 characters with PC scores  $>0.70$ : outer and inner perigone segment length and width (Outer\_ps\_l, Outer\_ps\_w, Inner\_ps\_l, Inner\_ps\_w), anther length (Anther\_l), throat hair (Th\_hair), leaf section height and width (Section\_h, Section\_w), arm length (Arm\_l), central parenchyma area (Parenchyma\_a), palisade cell and tissue height (Pal\_cell\_h, Pal\_tissue\_h), spongy tissue height (Sp\_tissue\_h), and xylem area (Xy\_a) (Supplementary Table S7). This set was extended with stigma/anther ratio (S/a\_r) (PC1 = 0.68) as the most important discriminative feature for *C. vernus* confirmed by previous research [15]. Finally, the CDA of the complete dataset (435 individuals of *C. vernus*, *C. heuffelianus*, mixed *C. heuffelianus*/*C. cf. heuffelianus*  $2n = 18$  SCC, *C. cf. heuffelianus*  $2n = 18$  SCC, *C. cf. heuffelianus*  $2n = 18$  WCC, *C. cf. heuffelianus*  $2n = 18$  PIC, *C. cf. heuffelianus*  $2n = 20$ , and *C. cf. heuffelianus*  $2n = 22$ ) was computed based on 15 previously mentioned characters. The clear separation of *C. heuffelianus* (Figure 6) in the negative part of the CDA of both axes was caused by the absence of throat hair (Figure 7, Supplementary Table S8). The mixed populations of the diploid *C. heuffelianus* and its polyploid  $2n = 18$  SCC cytotype overlapped with these two taxa (Figure 6), while all other polyploid populations were grouped in the positive part of both axes (Figure 6). The characters responsible for differentiation along the second axis were leaf cross-section width and arm length (Supplementary Table S8).

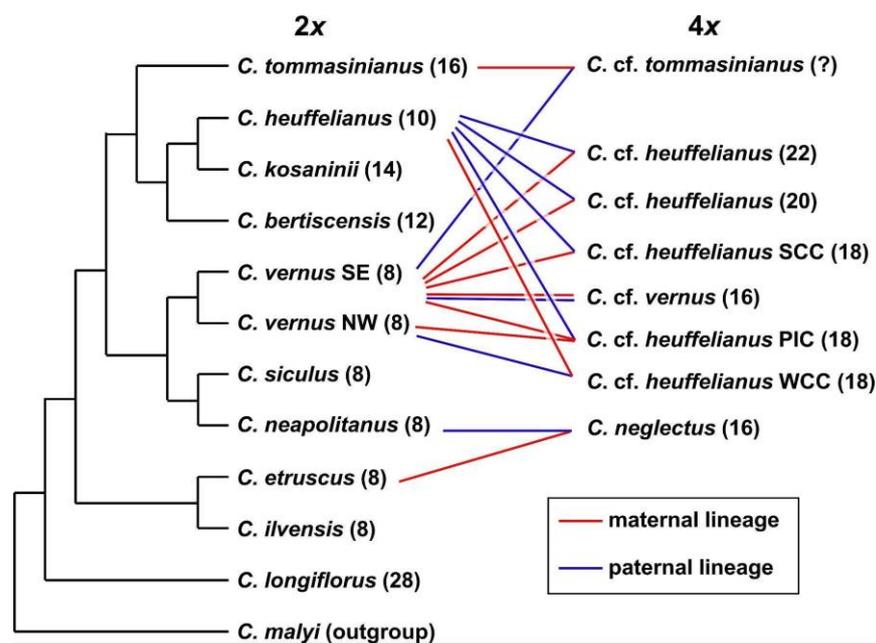


**Figure 7.** Flower, throat, and cross-section of leaf (from top to bottom) of parental species and a *C. cf. heuffelianus* polyploid: *C. vernus* (A), *C. heuffelianus* (B), and polyploid *C. cf. heuffelianus* representatives (C).

#### 4. Discussion

##### 4.1. Recognition of Recent Polyploids and Their Parents

*Crocus heuffelianus* represents one of the biggest taxonomical challenges within *Crocus* due to its high morphological variability. This morphological diversity seems mainly caused by the allotetraploid origin of the karyotypes of taxa possessing higher chromosome counts ( $2n = 18, 20, 22$ ), which became evident in a former molecular study [15]. Through incongruences between the GBS tree (Figures 1–3) and the chloroplast tree (Figure 4) as well as to the ITS tree (Supplementary Figure S8) and/or the affiliation of different alleles of variable nuclear single-copy markers (Figure 5, Supplementary Figures S5–S7) we identified at least seven independent hybridization events involving *C. heuffelianus* ( $2n = 10$ ) and *C. vernus* ( $2n = 8$ ), mostly with *C. vernus* as maternal parent. We summarize these reticulate relationships in Figure 8, a tree based on the diploid's topology derived from the SVDquartets analysis. *Crocus vernus* was found to possess two different chloroplast types depending on the geographical distribution (Eastern Alps to Pyrenees: NW type; Dinaric Alps: SE type). *Crocus vernus* from the Alps hybridized *C. heuffelianus* resulting in allotetraploid Western Carpathian populations (WCC;  $2n = 18$ ) with *C. heuffelianus* as the maternal parent. In the Pannonian-Illyric group of *C. cf. heuffelianus* (PIC;  $2n = 18$ ), we found two different chloroplast haplotypes stemming from the SE and NW *C. vernus* types indicating two crosses involving *C. vernus* as maternal parents (Figure 8). Furthermore, we also observed differences among the chloroplast haplotypes stemming from the NW type of *C. vernus* (Figure 4). One differed by at least two substitutions and two indels from any other NW type (see also branch lengths in Supplementary Figure S4). This indicates multiple hybridization events between *C. vernus* as maternal and *C. heuffelianus* as paternal parent creating the Pannonian-Illyric *C. cf. heuffelianus*. The SE type-carrying *C. vernus* was identified as the maternal species for Southern Carpathian populations of *C. cf. heuffelianus* (SCC;  $2n = 18$ ), as well as for the cytotypes  $2n = 4x = 20$  and  $2n = 4x = 22$  (Figure 8).



**Figure 8.** Schematic representation of species relationships in *Crocus* ser. *Verni* based on GBS, chloroplast, and nuclear single-copy gene data. For the tetraploids the maternal and paternal parents are indicated by lines connecting them to the respective diploid taxa. In brackets, chromosome numbers are provided.

Considering that the Southern Carpathian diploid cytotype of *C. heuffelianus* ( $2n = 10$ ) is ancestral to all polyploid forms and that according to morphological and chorological characteristics it corresponds to the original description, it represents *C. heuffelianus* s. str. [16,49]. It comprises populations with darker perigones and predominantly glabrous throats or hardly visible sparse and short hairs, which makes it distinct from all other series *Verni* species. Several authors reported different distribution ranges for *C. heuffelianus* [51–53]. This confusion is likely caused by confusing *C. heuffelianus* s. str. with its morphologically similar allopolyploids. For instance, some authors reported *C. heuffelianus* s. str. to occur in regions that are, according to our investigations, only inhabited by allopolyploid *C. cf. heuffelianus*. (e.g., throughout Bosnia and Herzegovina [54] or northeastern Italy [55]).

As a result of our study, we confirm a Carpathian distribution of *C. heuffelianus* s. str. in Slovakia, Romania, and Ukraine. *Crocus* cf. *heuffelianus* allopolyploid cytotypes have partly sympatric distributions with one of their parents or are growing in between the parental distribution areas (Supplementary Figure S11).

A new polyploid, *C. cf. vernus* ( $2n = 4x = 16$ ), was found in Central Albania, having *C. vernus* as the maternal parent. Its hybrid origin is indicated by its sister position in the GBS phylogeny, where it did not group within *C. vernus* as it would if it were an autopolyploid [5]. A (segmental) allotetraploid origin is also evident by its position in the SNP-based PCA (Figure 6). In the three closely examined variable nuclear single-copy genes (Figure 5, Supplementary Figures S5–S7), some of *C. cf. vernus*'s alleles were unique or usually grouped close to those of other taxa with  $2n = 8$  chromosomes. Therefore, the paternal parent could either be an extinct *C. vernus*-like genotype, which probably had eight chromosomes, or stems from a *C. vernus* population that we have not yet collected.

Allopolyploids in *C. cf. heuffelianus* and *C. cf. vernus* are also genetically differentiated, as indicated by the  $F_{st}$  values, which were usually higher than 0.15. This genetic differentiation might explain the difficulty in assigning ancestral populations (Supplementary Figure S10) or in the inference of co-ancestry (Supplementary Figure S9), where some allotetraploids were not shown as admixed.

*Crocus neglectus* ( $2n = 4x = 16$ ) could be confirmed here as an allotetraploid with *C. ilvensis* or *C. etruscus* as the maternal parent [15]. The fixation index (*C. etruscus*  $F_{st} = 0.07$ ; *C.*

*ilvensis*  $F_{st} = 0.10$ ) points to a more likely contribution of *C. etruscus*. However, *C. neglectus* shares its chloroplast haplotype with *C. ilvensis*, which was not found in *C. etruscus*. This discrepancy may be explained by the possibility that genetic drift eliminated this type of chloroplast in *C. etruscus* while it persisted in the geographically isolated *C. ilvensis* or that it was just not discovered in the individuals analyzed up to now. Seed size and germination are also more similar to *C. etruscus* [56], while flower bouquet is more similar to *C. ilvensis* [57]. *Crocus neapolitanus* is likely the paternal parent of *C. neglectus*. *Crocus neapolitanus* is genetically very similar to *C. vernus* but has a species-specific allele in one of the nuclear single-copy markers (*rcf2*; Figure 5) found in *C. neglectus*. The relatively high co-ancestry shared between *C. neglectus* and *C. neapolitanus* further supports *C. neapolitanus* as paternal parent.

#### 4.2. General Results Regarding Phylogeny and Systematics

Until the advent of high-throughput sequencing technologies, phylogenetic studies were often restricted to a limited set of markers. Consequently, relationships could not be resolved, as in the case of series *Verni* [15]. In Raca et al. [16], we already successfully increased the resolution by using genome-wide SNP data obtained from GBS. However, since this study aimed to show the phylogenetic affiliation of *C. bertiscensis*, we included neither all the species in *Crocus* ser. *Verni* (*C. siculus* was lacking) nor a higher number of individuals per species and did not analyze species relationships in detail. Here we added additional samples, a chloroplast marker dataset, as well as a dataset comprising nuclear single-copy genes. The latter mainly served to identify the parental origin of the recent allotetraploids. Excluding the allotetraploids in the analysis increased the support values of the tree backbone of the GBS-based MP tree, similar to the BI-based analysis in Raca et al. [16]. However, our SVDquartets species tree (Figure 1) showed a different topology. The relatively high degree of homoplasy in the GBS dataset (Table 2) indicates incomplete lineage sorting (ILS) and/or hybridization that could result in genomic introgression. Incongruences of single-gene trees, as well as the chloroplast tree, support this hypothesis. For example, *C. vernus* today has two different chloroplast haplotypes (SE type and NW type). The NW type is shared with other closely related species such as *C. etruscus*, *C. ilvensis*, *C. neapolitanus*, and *C. siculus*, which all occur in Italy (Supplementary Figure S11). The SE type is found only in the southeastern range of *C. vernus* and groups with *C. longiflorus* as sister to the NW type clade. A possible interpretation could be that these two chloroplast types were once both present in Italy, where introgression of *C. vernus* or its ancestor with *C. longiflorus* occurred. Subsequently, the SE type was sorted out and persisted only in the eastern distribution of *C. vernus*. The position of *C. longiflorus* as sister to *C. vernus* was also observed in one of the nuclear single-copy markers (*topo6*), which supports the hypothesis of an ancient introgression event between *C. vernus* or its ancestor and *C. longiflorus*. There are also other examples where species (or one of their alleles) group with species to which they are only distantly related according to the GBS trees (SVDquartets as well as MP), such as *C. tommasinianus* grouping together with *C. vernus* in *topo6* and *rcf2* (Figure 5).

Single gene trees generally showed a poor resolution, which was the main reason we show only *rcf2* and *topo6*, and even in data sets with a relatively high number of parsimony-informative sites, phylogenetic relationships remained largely unresolved (Supplementary Figure S7) due to the young age of the group and homoplasy in the data. Since multispecies coalescent methods consider ILS and introgression, the SVDquartets tree should more closely reflect the true species relationships although its topology differs from trees derived with concatenation approaches [58].

#### 4.3. Chromosome and Genome Size Evolution

All angiosperms essentially have undergone polyploidization [59]. Dysploidy, chromatin elimination or expansion, nested polyploidizations, introgression, and hybridization can confound the evolutionary dynamics between chromosome number and genome size over time [60–62]. Consequently, these two genomic parameters show independent

evolution and no clear correlation, especially in older polyploids, such as meso- and paleopolyploids [61,63,64]. Nevertheless, some correlation can still be observed in neopolyploids [6,61].

Neopolyploids often have roughly additive genome sizes and chromosome numbers of the progenitors, as in *C. cf. heuffelianus* ( $2n = 18$ ,  $2C = 10.88\text{--}12.84$  pg), *C. neglectus* ( $2n = 16$ ,  $2C = 12.24$  pg), and *C. cf. vernus* ( $2n = 16$ ,  $2C = 12.38$  pg). However, although genome sizes could remain roughly additive in the absence of considerable chromatin loss [65], the additive pattern in chromosome number is often blurred when chromosome fusion (descending dysploidy) or fission (ascending dysploidy) occurs [61,66,67]. Considering the chromosome numbers and genome sizes of the parental genomes of *C. cf. heuffelianus*, *C. heuffelianus* ( $2n = 10$ ), and *C. vernus* ( $2n = 8$ ), it is highly likely that the *C. cf. heuffelianus* cytotypes from Albania, Kosovo, Montenegro, and Serbia ( $2n = 20$  and  $2n = 22$ ,  $2C = 11.95$  pg) have undergone ascending dysploidy. Indeed, a few shorter chromosomes have been observed in these cytotypes, as ascending dysploidy entails chromosome breakage resulting in an increase in chromosome number in general and shorter chromosomes in particular (Supplementary Figure S1).

The cryptic relationship between chromosome number and genome size increases over time in polyploids. This relationship can be observed between *C. longiflorus* and the core series *Verni* species. In general, series *Verni* species with more chromosomes have smaller genomes, thus showing a negative relationship between genome size and chromosome number, with or without *C. longiflorus* (Supplementary Figure S1), which is sister to all other taxa of series *Verni* and was recently moved to this series from the now-defunct series *Longiflori* B.Mathew [15]. *Crocus longiflorus* has the highest chromosome number but the lowest genome size ( $2n = 28$ ,  $2C = 3.21$  pg) in the series. Likewise, *C. tommasinianus* ( $2n = 16$ ), which has twice as many chromosomes as *C. vernus* ( $2n = 8$ ), has a roughly similar genome size to *C. vernus*. Taking into account only chromosome counts, *C. longiflorus* and *C. tommasinianus* would likely be deemed polyploids. However, they show very low heterozygosity scores (Supplementary Table S5), indicating extensive chromatin elimination and essentially diploidized genomes. This chromosome number–genome size relationship between *C. longiflorus* and the core species of series *Verni* can indicate an ancient whole-genome duplication event prior to the divergence of series *Verni*, likely even before the diversification of *Crocus* [68].

A burst of certain repetitive DNA element families may also have promoted genome size expansion in the core series *Verni*. Thus, the genome sizes in the core taxa of series *Verni* have increased despite chromosome number reduction relative to *C. longiflorus*. Nevertheless, this hypothesis can only be supported when fused chromosome blocks [59,69] and expansion of lineage-specific repeat families [70,71] are observed within the core of series *Verni*. This can become possible by comparative genome and repeat analyses between *C. longiflorus* and the other species from series *Verni*. Combined with cytogenetic analyses involving all major groups within the genus, this should allow an understanding of karyotype evolution within the genus.

## 5. Conclusions

Our study was designed to disentangle the *C. heuffelianus* complex using GBS data and chloroplast markers and combine them with morphology, genome sizes, and chromosome counts. This strategy generally proved to be successful when (a) a broad sampling of allotetraploids and potential parental species are included and (b) the allotetraploids group as sister to their paternal parent in the phylogenetic GBS tree. In cases where our sampling was restricted to just a few samples, such as for some of the Italian species and allotetraploid *C. neglectus*, conclusions remain a bit uncertain. The combination of chloroplast data with only GBS failed to reveal the paternal contributor of *C. cf. heuffelianus* (SCC;  $2n = 18$ ). Here, additional nuclear markers (Figure 5, Supplementary Figure S7) were necessary to identify the NW type of *C. vernus* as the paternal parent. Algorithms specifically designed to detect hybridization signals in GBS data were only partly able to recover the allopolyploids within

series *Verni*, while in several cases co-ancestry values for their parents were lower than even between independent diploids. Thus, such an approach alone is not sufficient to infer polyploidization or indicate the diploid progenitors of polyploids in crocuses.

By linking molecular results and genome sizes with morphology, a clear differentiation of allopolyploids and parental species was possible. In the taxonomically confusing *C. heuffelianus* complex, a circumscription of the diploid taxon and its distinction from the allotetraploids is now possible. *Crocus heuffelianus* s. str. is the diploid cytotype ( $2n = 10$ ) with mostly glabrous throats and darker perigone segments. Together with *C. vernus*, it represents the parental species for all the *C. cf. heuffelianus* allotetraploids. The cytotype  $2n = 18$  of *C. cf. heuffelianus* is split into three groups: Western Carpathian (WCC), Pannonian-Illyric (PIC), and Southern Carpathian (SCC). *Crocus heuffelianus* s. str. is the mother of WCC only, while the NW and SE types of *C. vernus* are maternal lineages of PIC. The SE type of *C. vernus* is the only maternal parent of SCC, as well as for the cytotypes with  $2n = 20$  and  $2n = 22$  chromosomes. All analyzed *C. cf. heuffelianus* polyploids represent morphologically intermediate forms between their parental species, but currently, they cannot be distinguished based on the investigated morphological characters. Instead, chromosome counts are necessary.

While it is possible to unravel more recent polyploidization events, the detection of paleo-polyploidization remains difficult. Our incongruent gene trees indicate past hybridization events, which might have triggered genome size and chromosome number changes. However, while the methods applied here work well in this very young taxon group, they are not satisfactory in uncovering ancient and complex evolutionary histories, particularly those involving highly dynamic genome size and chromosome number changes. In series *Verni*, and in *Crocus* in general, the future availability of genome assemblies will enable comparative cytogenomic analyses to detect potential ancient polyploidization and to trace chromosomal rearrangements resulting in changing karyotypes.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/biology12020303/s1>, Figure S1: Chromosome counts for diploid *C. vernus* and *C. heuffelianus* (A and B), tetraploid *C. cf. vernus* (C), and tetraploid *C. cf. heuffelianus* (D–F). Asterisks show relatively shorter chromosomes. (G) The chromosome number and genome size in diploid and tetraploid *C. ser. Verni* taxa show a negative relationship, which was only significant when *C. longiflorus* was included; Figure S2: Strict consensus MP tree based on 2009 GBS loci including only diploid accessions of *Crocus ser. Verni*; Figure S3: Strict consensus MP tree based on 2009 GBS loci including both diploid and tetraploid accessions and cytotypes of *Crocus ser. Verni*; Figure S4: BI phylogenetic tree of *Crocus ser. Verni* based on chloroplast markers. Numbers at branches provide BI posterior probabilities/bootstrap values from MP analysis. Asterisks indicate bootstrap support values  $\geq 80\%$ ; Figure S5: Strict consensus MP tree of *topo6* in *Crocus ser. Verni*. Allelic differences (A1–A4) in these markers were used to track the bi-parental contributions of diploids to allotetraploids; Figure S6: Strict consensus MP tree of *rcf2* in *Crocus ser. Verni*. Allelic differences (A1–A4) in these markers were used to track the bi-parental contributions of diploids to allotetraploids; Figure S7: Strict consensus MP tree of *orcp* in *Crocus ser. Verni*. Allelic differences (A1–A4) in these markers were used to track the bi-parental contributions of diploids to allotetraploids; Figure S8: Phylogenetic tree of *Crocus ser. Verni* obtained through Bayesian phylogenetic inference based on rDNA ITS sequences. Numbers along branches indicate BI posterior probabilities (pp), pp supports of 1.0 are indicated by asterisks; Figure S9: FINERADSTRUCTURE co-ancestry matrix of the study species based on 10,591 loci (without outgroup species *C. malyi*; *C. siculus* was excluded due to low coverage). Black indicates maximum levels of co-ancestry between two individuals, white the minimum (scale on the right). Numbers below the plots indicate the sample ID; Figure S10: Population structure analysis in *Crocus ser. Verni* based on 2207 GBS loci using FASTSTRUCTURE at  $K = 4$  (30,661 SNPs) and LEA  $K = 5$  and 8 (2172 unlinked SNPs). Each vertical line represents one individual, while each color shows the genetic composition that is assigned into a distinct genetic cluster; Figure S11: Map with the approximate distributions of species in *Crocus ser. Verni*. Distribution areas of different species are indicated by different colors or patterns; Table S1: Detailed information about the individuals included in the chloroplast (CP), genotyping-by-sequencing (GBS), nuclear single-copy genes (NSCG), and morpho-anatomical analyses (MA) of *Crocus ser. Verni* separately (No.) and in total (No. total),

accompanied with voucher numbers; Table S2: Nuclear single-copy marker PCR information in *Crocus* ser. *Verni*; Table S3: Genome size measurements and chromosome counts in *Crocus* ser. *Verni*; Table S4: The dataset for morpho-anatomical analyses (consisting of the parental diploid species *C. vernus* and *C. heuffelianus* and all polyploid cytotypes of *C. cf. heuffelianus*); Table S5: Information on GBS data, loci, and heterozygosity in *Crocus* ser. *Verni*; Table S6: Fixation index (*Fst*) of allotetraploids versus diploids in *Crocus* ser. *Verni*; Table S7: The differential characters based on representative populations dataset of *Crocus* ser. *Verni* highlighted by the principal component analysis (PCA); Table S8: Discriminant analysis (CDA) conducted on the dataset including all populations of diploid accessions of *C. vernus* and *C. heuffelianus* and all polyploid accessions of *C. cf. heuffelianus*; Dataset S1: Alignment of concatenated GBS loci used for phylogenetic analysis; Dataset S2: Variant calling format (vcf) file of the series *Verni* dataset used for FASTSTRUCTURE and LEA; Dataset S3: Unlinked SNP matrix in .geno format used for PCA. Dataset S4: RADPAINTER input file. References [72,73] are cited in Supplementary Materials.

**Author Contributions:** Conceptualization, D.H., V.R. and I.R.; methodology, data analysis, I.R., D.H., F.R.B. and N.E.W.; resources, I.R., V.R., D.H., and H.K.; data curation, D.H.; writing—original draft preparation, F.R.B., I.R. and D.H.; writing—review and editing, all authors; visualization, I.R., F.R.B. and D.H.; funding acquisition, D.H. and I.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by grants of the Ministry of Education, Science and Technological Development of the Republic of Serbia (project No. 173030, contract No. 451-03-68/2020-14/200124) to I.R., Deutscher Akademischer Austauschdienst (DAAD) to I.R., International Association for Plant Taxonomy (IAPT) to I.R., and Deutsche Forschungsgemeinschaft (DFG grant HA 7550/2 and HA 7550/4) to D.H. Costs for open access publishing were partially funded by the Deutsche Forschungsgemeinschaft (DFG grant 491250510) and the Leibniz Open Access Fund.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The analyzed DNA sequences and raw reads are available through ENA study accession number PRJEB57934, Datasets S1–S4 can be downloaded from e!DAL through <https://doi.org/10.5447/ipk/2023/5>, accessed on 10 February 2022.

**Acknowledgments:** We would like to thank I. Faustmann, C. Koch, B. Kraenzlin, and P. Oswald for help with plant cultivation and lab work; A. Himmelbach and S. König for performing Illumina sequencing; and L. Peruzzi, K. Hegedúšová, P. Vantara, T. Huber, S. Milanović, M. Slovák, J. Rukšāns, N. Sarajlić, B. Štulović, M. Mihajlović, D. Jakšić, and L. Shuka for providing *Crocus* materials for molecular and morpho-anatomical analysis and/or assistance during field work. We thank the authorities of the respective countries for supporting our work with collection permits.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Soltis, D.E.; Visger, C.J.; Marchant, D.B.; Soltis, P.S. Polyploidy: Pitfalls and paths to a paradigm. *Am. J. Bot.* **2016**, *103*, 1146–1166. [[CrossRef](#)] [[PubMed](#)]
2. Van de Peer, Y.; Fawcett, J.A.; Proost, S.; Sterck, L.; Vandepoele, K. The flowering world: A tale of duplications. *Trends Plant Sci.* **2009**, *14*, 680–688. [[CrossRef](#)]
3. Van de Peer, Y.; Ashman, T.L.; Soltis, P.S.; Soltis, D.E. Polyploidy: An evolutionary and ecological force in stressful times. *Plant Cell* **2021**, *33*, 11–26. [[CrossRef](#)] [[PubMed](#)]
4. Vu, G.T.H.; Schmutzer, T.; Bull, F.; Cao, H.X.; Fuchs, J.; Tran, T.D.; Jovtchev, G.; Pistrick, K.; Stein, N.; Pecinka, A.; et al. Comparative genome analysis reveals divergent genome size evolution in a carnivorous plant genus. *Plant Genome* **2015**, *8*, plantgenome2015-04. [[CrossRef](#)]
5. Brassac, J.; Blattner, F.R. Species-level phylogeny and polyploid relationships in *Hordeum* (Poaceae) inferred by next-generation sequencing and in silico cloning of multiple nuclear loci. *Syst. Biol.* **2015**, *64*, 792–808. [[CrossRef](#)] [[PubMed](#)]
6. Jakob, S.S.; Meister, A.; Blattner, F.R. The considerable genome size variation of *Hordeum* species (Poaceae) is linked to phylogeny, life form, ecology, and speciation rates. *Mol. Biol. Evol.* **2004**, *21*, 860–869. [[CrossRef](#)] [[PubMed](#)]
7. Albach, D.C.; Greilhuber, J. Genome size variation and evolution in *Veronica*. *Ann. Bot.* **2004**, *94*, 897–911. [[CrossRef](#)] [[PubMed](#)]
8. Leitch, I.J.; Bennett, M.D. Genome downsizing in polyploid plants. *Biol. J. Linn. Soc.* **2004**, *82*, 651–663. [[CrossRef](#)]

9. Leitch, I.J.; Fay, M. Plant genome horizons: Michael Bennett's contribution to genome research. *Ann. Bot.* **2008**, *101*, 737–746. [[CrossRef](#)]
10. Meudt, H.M.; Rojas-Andrés, B.M.; Prebble, J.M.; Low, E.; Garnock-Jones, P.J.; Albach, D.C. Is genome downsizing associated with diversification in polyploid lineages of *Veronica*? *Bot. J. Linn. Soc.* **2015**, *178*, 243–266. [[CrossRef](#)]
11. Petrov, D.A. Evolution of genome size: New approaches to an old problem. *Trends Genet.* **2001**, *17*, 23–28. [[CrossRef](#)] [[PubMed](#)]
12. Vu, G.T.H.; Cao, H.X.; Reiss, B.; Schubert, I. Deletion-bias in DNA double-strand break repair differentially contributes to plant genome shrinkage. *New Phytol.* **2017**, *214*, 1712–1721. [[CrossRef](#)] [[PubMed](#)]
13. Petersen, G.; Seberg, O.; Thorsøe, S.; Jørgensen, T.; Mathew, B. A phylogeny of the genus *Crocus* (Iridaceae) based on sequence data from five plastid regions. *Taxon* **2008**, *57*, 487–499. [[CrossRef](#)]
14. Harpke, D.; Meng, S.; Rutten, T.; Kerndorff, H.; Blattner, F.R. Phylogeny of *Crocus* (Iridaceae) based on one chloroplast and two nuclear loci: Ancient hybridization and chromosome number evolution. *Mol. Phylogenet. Evol.* **2013**, *66*, 617–627. [[CrossRef](#)]
15. Harpke, D.; Carta, A.; Tomović, G.; Randelović, V.; Randelović, N.; Blattner, F.R.; Peruzzi, L. Phylogeny, karyotype evolution and taxonomy of *Crocus* series *Verni* (Iridaceae). *Plant Syst. Evol.* **2015**, *301*, 309–325. [[CrossRef](#)]
16. Raca, I.; Harpke, D.; Shuka, L.; Randelović, V. A new species of *Crocus* ser. *Verni* (Iridaceae) with  $2n = 12$  chromosomes from the Balkans. *Plant Biosyst.* **2022**, *156*, 36–42. [[CrossRef](#)]
17. Karasawa, K. Karyological studies in *Crocus*. III. *Jap. J. Bot.* **1943**, *12*, 475–503.
18. Brighton, C.A. Cytological problems in the genus *Crocus* (Iridaceae): I. *Crocus vernus* aggregate. *Kew Bull.* **1976**, *31*, 33–46. [[CrossRef](#)]
19. Baldini, R.M. Karyological observations on two *Crocus* species (Iridaceae) from Tuscany (Italy). *Caryologia* **1990**, *43*, 341–345. [[CrossRef](#)]
20. Micevska, D. Caryological Analysis of the Species of the Genus *Crocus*. Master's Thesis, University of Skopje, Skopje, North Macedonia, 2000.
21. Mitic, D. Caryological Analysis of Some Species of the Genus *Crocus* L. Master's Thesis, University of Skopje, Skopje, North Macedonia, 2001.
22. Peruzzi, L.; Carta, A. *Crocus ilvensis* sp. nov. (sect. *Crocus*, Iridaceae), endemic to Elba Island (Tuscan Archipelago, Italy). *Nord. J. Bot.* **2011**, *29*, 6–13. [[CrossRef](#)]
23. Peruzzi, L.; Carta, A.; Garbari, F. Lectotypification of the name *Crocus sativus* var. *vernus* L. (Iridaceae) and its consequences within *Crocus* ser. *Verni*. *Taxon* **2013**, *62*, 1037–1040. [[CrossRef](#)]
24. Elshire, R.J.; Glaubitz, J.C.; Sun, Q.; Poland, J.A.; Kawamoto, K.; Buckler, E.S.; Mitchell, S.E. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* **2011**, *6*, e19379. [[CrossRef](#)]
25. Kerndorff, H.; Harpke, D. *Crocus cobbii* Kerndorff, Pasche & Harpke species nova (Liliiflorae, Iridaceae) and its relatives. *Stappia* **2022**, *113*, 5–32. [[CrossRef](#)]
26. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649. [[CrossRef](#)]
27. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)] [[PubMed](#)]
28. Wendler, N.; Mascher, M.; Nöh, C.; Himmelbach, A.; Scholz, U.; Ruge-Wehling, B.; Stein, N. Unlocking the secondary gene-pool of barley with next-generation sequencing. *Plant Biotechnol. J.* **2014**, *12*, 1122–1131. [[CrossRef](#)] [[PubMed](#)]
29. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **2011**, *17*, 10. [[CrossRef](#)]
30. Eaton, D.A.R.; Overcast, I. ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics* **2020**, *36*, 2592–2594. [[CrossRef](#)] [[PubMed](#)]
31. Frichot, E.; François, O. LEA: An R package for landscape and ecological association studies. *Methods Ecol. Evol.* **2015**, *6*, 925–929. [[CrossRef](#)]
32. Raj, A.; Stephens, M.; Pritchard, J.K. fastSTRUCTURE: Variational inference of population structure in large SNP data sets. *Genetics* **2014**, *197*, 573–589. [[CrossRef](#)]
33. Wickham, H.; Wickham, M.H. *Package 'Tidyverse'*, 2019; pp. 1–5. Available online: <http://tidyverse.tidyverse.org> (accessed on 20 July 2022).
34. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*; Springer: New York, NY, USA, 2016; pp. 1–189.
35. Rozas, J.; Ferrer-Mata, A.; Sánchez-DelBarrio, J.C.; Guirao-Rico, S.; Librado, P.; Ramos-Onsins, S.E.; Sánchez-Gracia, A. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* **2017**, *34*, 3299–3302. [[CrossRef](#)] [[PubMed](#)]
36. Blattner, F.R. Direct amplification of the entire ITS region from poorly preserved plant material using recombinant PCR. *Biotechniques* **1999**, *27*, 1180–1186. [[CrossRef](#)] [[PubMed](#)]
37. Blattner, F.R. Phylogenetic analysis of *Hordeum* (Poaceae) as inferred by nuclear rDNA ITS sequences. *Mol. Phylogenet. Evol.* **2004**, *33*, 289–299. [[CrossRef](#)]
38. Waminal, N.E.; Park, H.M.; Ryu, K.B.; Kim, J.H.; Yang, T.J.; Kim, H.H. Karyotype analysis of *Panax ginseng* C.A.Meyer, 1843 (Araliaceae) based on rDNA loci and DAPI band distribution. *Comp. Cytogenet.* **2012**, *6*, 425–441. [[CrossRef](#)]
39. Rodríguez-Domínguez, J.M.; Ríos-Lara, L.L.; Tapia-Campos, E.; Barba-Gonzalez, R. An improved technique for obtaining well-spread metaphases from plants with numerous large chromosomes. *Biotech. Histochem.* **2017**, *92*, 159–166. [[CrossRef](#)]

40. Swofford, D.L. *PAUP\*: Phylogenetic Analysis Using Parsimony (\* and Other Methods)*; Sinauer Assoc.: Sunderland, MA, USA, 2002.
41. Ronquist, F.; Teslenko, M.; van der Mark, P.; Ayres, D.L.; Darling, A.; Höhna, S.; Larget, B.; Liu, L.; Suchard, M.A.; Huelsenbeck, J.P. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **2012**, *61*, 539–542. [[CrossRef](#)]
42. Rochette, N.C.; Rivera-Colón, A.G.; Catchen, J.M. Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Mol. Ecol.* **2019**, *28*, 4737–4754. [[CrossRef](#)]
43. Malinsky, M.; Trucchi, E.; Lawson, D.J.; Falush, D. RADpainter and fineRADstructure: Population inference from RADseq data. *Mol. Biol. Evol.* **2018**, *35*, 1284–1290. [[CrossRef](#)] [[PubMed](#)]
44. Gligorijević, S.; Pejčinović, D. Contribution to the methodology of anatomical sections preparation. *Acta Biol. Med. Exp.* **1983**, *8*, 43–45.
45. Raca, I.; Ljubisavljević, I.; Jušković, M.; Randelović, N.; Randelović, V. Comparative anatomical study of the taxa from series *Verni* Mathew (*Crocus* L.) in Serbia. *Biol. Nyssana* **2017**, *8*, 15–22. [[CrossRef](#)]
46. Raca, I.; Jovanovic, M.; Ljubisavljevic, I.; Juskovic, M.; Randelovic, V. Morphological and leaf anatomical variability of *Crocus* cf. *heuffelianus* Herb. (Iridaceae) populations from the different habitats of the Balkan Peninsula. *Turk. J. Bot.* **2019**, *43*, 645–658. [[CrossRef](#)]
47. Schneider, C.A.; Rasband, W.S.; Eliceiri, K.W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **2012**, *9*, 671–675. [[CrossRef](#)] [[PubMed](#)]
48. Mosolygo-L, A.; Sramko, G.; Barabas, S.; Czeglédi, L.; Javor, A.; Molnar, V.A.; Suranyi, G. Molecular genetic evidence for allotetraploid hybrid speciation in the genus *Crocus* L. (Iridaceae). *Phytotaxa* **2016**, *258*, 121–136. [[CrossRef](#)]
49. Raca, I. Taxonomy and Phylogeny of Series *Verni* Mathew (*Crocus* L.) in Southeastern Europe—Morpho-Anatomical, Cytological and Molecular Approach. Ph.D. Thesis, University of Niš, Niš, Serbia, 2021.
50. Statsoft Inc. *STATISTICA (Data Analysis Software System)*, Version 7; Statsoft Inc.: Tulsa, OK, USA, 2004.
51. Herbert, W.; Lindley, J. A history of the species of *Crocus*. *J. R. Hort. Soc.* **1847**, *2*, 249–293.
52. Meusel, H.; Jäger, E.J.; Weinert, E. *Vergleichende Chorologie der zentraleuropäischen Flora*; VEB Fischer: Jena, Germany, 1965.
53. Malynovsky, K.A. *Roslynnisty vysokogirja Ukrainskyykh Karpat*; Naukova dumka: Kyiv, Ukraine, 1980.
54. Beck-Mannagetta, G. *Flora von Bosnien, der Herzegowina und des Sandzaks Novipazar*; C. Gerold: Vienna, Austria, 1904.
55. Peruzzi, L. *Crocus heuffelianus* (Iridaceae), a new record for the Italian flora. *Phytotaxa* **2016**, *261*, 291–294. [[CrossRef](#)]
56. Carta, A.; Probert, R.; Moretti, M.; Peruzzi, L.; Bedini, G. Seed dormancy and germination in three *Crocus* ser. *Verni* species (Iridaceae): Implications for evolution of dormancy within the genus. *Plant Biol.* **2014**, *16*, 1065–1074. [[CrossRef](#)]
57. Carta, A.; Flamini, G.; Cioni, P.L.; Pistelli, L.; Peruzzi, L. Flower bouquet variation in four species of *Crocus* ser. *Verni*. *J. Chem. Ecol.* **2015**, *41*, 105–110. [[CrossRef](#)]
58. Molloy, E.K.; Warnow, T. To include or not to include: The impact of gene filtering on species tree estimation methods. *Syst. Biol.* **2017**, *67*, 285–303. [[CrossRef](#)]
59. Soltis, P.S.; Soltis, D.E. Ancient WGD events as drivers of key innovations in angiosperms. *Curr. Opin. Plant Biol.* **2016**, *30*, 159–165. [[CrossRef](#)]
60. Murat, F.; Armero, A.; Pont, C.; Klopp, C.; Salse, J. Reconstructing the genome of the most recent common ancestor of flowering plants. *Nat. Genet.* **2017**, *49*, 490–496. [[CrossRef](#)]
61. Mandakova, T.; Lysak, M.A. Post-polyploid diploidization and diversification through dysploid changes. *Curr. Opin. Plant Biol.* **2018**, *42*, 55–65. [[CrossRef](#)] [[PubMed](#)]
62. Glombik, M.; Bačovský, V.; Hobza, R.; Kopecký, D. Competition of parental genomes in plant hybrids. *Front. Plant Sci.* **2020**, *11*, 200. [[CrossRef](#)]
63. Lysak, M.A.; Cheung, K.; Kitschke, M.; Bures, P. Ancestral chromosomal blocks are triplicated in *Brassicaceae* species with varying chromosome number and genome size. *Plant Physiol.* **2007**, *145*, 402–410. [[CrossRef](#)] [[PubMed](#)]
64. El-Shehawi, A.M.; Elseehy, M.M. Genome size and chromosome number relationship contradicts the principle of Darwinian evolution from common ancestor. *J. Phylogenet. Evol. Biol.* **2017**, *5*, 179. [[CrossRef](#)]
65. Fedoroff, N.V. Transposable elements, epigenetics, and genome evolution. *Science* **2012**, *338*, 758–767. [[CrossRef](#)] [[PubMed](#)]
66. Guerra, M. Agmatoploidy and symploidy: A critical review. *Genet. Mol. Biol.* **2016**, *39*, 492–496. [[CrossRef](#)]
67. Schubert, I.; Lysak, M.A. Interpretation of karyotype evolution should consider chromosome structural constraints. *Trends Genet.* **2011**, *27*, 207–216. [[CrossRef](#)] [[PubMed](#)]
68. Tank, D.C.; Eastman, J.M.; Pennell, M.W.; Soltis, P.S.; Soltis, D.E.; Hinchliff, C.E.; Brown, J.W.; Sessa, E.B.; Harmon, L.J. Nested radiations and the pulse of angiosperm diversification: Increased diversification rates often follow whole genome duplications. *New Phytol.* **2015**, *207*, 454–467. [[CrossRef](#)]
69. Murat, F.; Xu, J.-H.; Tannier, E.; Abrouk, M.; Guilhot, N.; Pont, C.; Messing, J.; Salse, J. Ancestral grass karyotype reconstruction unravels new mechanisms of genome shuffling as a source of plant evolution. *Genome Res.* **2010**, *20*, 1545–1557. [[CrossRef](#)]
70. Kelly, L.J.; Renny-Byfield, S.; Pellicer, J.; Macas, J.; Novák, P.; Neumann, P.; Lysak, M.A.; Day, P.D.; Berger, M.; Fay, M.F.; et al. Analysis of the giant genomes of *Fritillaria* (Liliaceae) indicates that a lack of DNA removal characterizes extreme expansions in genome size. *New Phytol.* **2015**, *208*, 596–607. [[CrossRef](#)]

71. Macas, J.; Novak, P.; Pellicer, J.; Cizkova, J.; Koblizkova, A.; Neumann, P.; Fukova, I.; Doležel, J.; Kelly, L.J.; Leitch, I.J. In depth characterization of repetitive DNA in 23 plant genomes reveals sources of genome size variation in the legume tribe Fabae. *PLoS ONE* **2015**, *10*, e0143424. [[CrossRef](#)] [[PubMed](#)]
72. Brighton, C.A.; Mathew, B.; Marchant, C.J. Chromosome counts in the genus *Crocus* (Iridaceae). *Kew Bull.* **1973**, *28*, 451–464. [[CrossRef](#)]
73. Thiv, M.; Wörz, A. Die neue Identität des Zavelsteiner Krokus als *Crocus neglectus* nach DNA-Untersuchungen. *Jahresh. Ges. Nat. Württemberg* **2015**, *171*, 163–172.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.