

Article

New EST-SSR Markers for Individual Genotyping of Opium Poppy Cultivars (*Papaver somniferum* L.)

Jakub Vašek ^{1,*} , Daniela Čílová ¹, Martina Melounová ¹, Pavel Svoboda ², Pavel Vejl ¹, Radka Štikarová ¹, Luboš Vostrý ¹, Perla Kuchtová ³ and Jaroslava Ovesná ²

¹ Czech University of Life Sciences Prague, FAFNR, Department of Genetics and Breeding, Kamýčká 129, 6 Suchdol, 16500 Prague, Czech Republic; cilova@af.czu.cz (D.Č.); melounova@af.czu.cz (M.M.); vejl@af.czu.cz (P.V.); stikarova@af.czu.cz (R.Š.); vostry@af.czu.cz (L.V.)

² Crop Research Institute, Division of Crop Genetics and Breeding, Drnovská 507/73, 6 Ruzyně, 16106 Prague, Czech Republic; pavel.svoboda@vurv.cz (P.S.); ovesna@vurv.cz (J.O.)

³ Czech University of Life Sciences, FAFNR, Department of Agroecology and Crop Production, Kamýčká 129, 6 Suchdol, 16500 Prague, Czech Republic; kuchtova@af.czu.cz

* Correspondence: vasek@af.czu.cz; Tel.: +420-22438-2562

Received: 31 October 2019; Accepted: 18 December 2019; Published: 19 December 2019



Abstract: High-quality simple sequence repeat (SSR) markers are invaluable tools for revealing genetic variability which could be utilized for many purposes, such as breeding new varieties or the identifying current ones, among other applications. Based on the analysis of 3.7 million EST sequences and 15 genomic sequences from bacterial artificial chromosome (BAC) libraries, 200 trinucleotide genic (EST)-SSR and three genomic (gSSR) markers were tested, where 17 of them fulfilled all criteria for quality markers. Moreover, the reproducibility of these new markers was verified by two genetics laboratories, with a mean error rate per allele and per locus equal to 0.17%. These markers were tested on 38 accessions of *Papaver somniferum* and nine accessions of another five species of the *Papaver* and *Argemone* genera. In total, 118 alleles were detected for all accessions (median = 7; three to ten alleles per locus) and 88 alleles (median = 5; three to nine alleles per locus) within *P. somniferum* alone. Multivariate methods and identity analysis revealed high resolution capabilities of the new markers, where all but three pair accessions (41 out of 47) had a unique profile and opium poppy was distinguished from other species.

Keywords: *Papaver somniferum* L.; Papaveraceae; EST-SSR; microsatellites; individual genotyping

1. Introduction

Papaver somniferum L. (opium poppy) is a well-known and extensively studied representative of the *Papaver* genus, numbering over 100 species [1]. As a traditional plant, it has accompanied man since at least the early Neolithic age [2,3]. Today, *P. somniferum* is an economically important plant grown throughout the world for two main reasons. Firstly, *P. somniferum* serves as a source of secondary metabolites such as morphine, thebaine, codeine, or noscapine, which are utilized by pharmaceutical companies for producing medicines with analgesic, antitussive, sedative, or anti-tumor effects [4,5]. Its dark side is infamously the abuse of its sap (“opium latex”) for the production of heroin [6]. Secondly, poppy seeds do not contain a high amount of morphinan alkaloids when they are properly harvested and/or treated [7,8]; thus they are also used in the food industry, typically in Central and Eastern Europe, given their historical and cultural background.

The opium poppy is a diploid ($2n = 14$) annual plant with a prevailing self-pollinating mode of cross-breeding [9,10]. This characteristic simplifies the whole breeding process, which starts from heterogeneous material of different origin and usually continues by the selection of individual “mother”

plants with the required properties. Then follows several generations of self-pollination, necessary for genetic stabilization, homozygotization, and obtaining a sufficient amount of seeds. New varieties should fulfill the UPOV requirements (The International Union for the Protection of New Varieties of Plants) abbreviated DUS (Distinctness, Uniformity, Stability; see more details on UPOV webpages) and could be defined as a pure line, which is characterized by a very high degree of homozygosity and homogeneity. It should be mentioned that other breeding strategies are possible (e.g., creation of hybrid varieties) but are rarely performed.

Modern poppy varieties are classified into three broad categories, as their respective utilization reflects the breeding purpose [11]. The industrial category represents varieties with generally high alkaloid content or varieties producing high amounts of specific chemical compounds like thebaine or noscapine [5]. In contrast, the culinary category includes varieties where the breeding was aimed at low alkaloid content and enhanced flavor. Some varieties combine features of the previously mentioned categories and thus are classified as dual purpose. An extra category could represent varieties for ornamental purposes. Irrespective of category, there is a high demand for suitable markers which allow one to evaluate the diversity among the germplasm collection, accelerate the breeding process, or identify individual varieties. There are typically two approaches, which are quite often intertwined; the first one is based on chemoclassification according to type and number of chemical compounds, and the second one exploits different types of DNA markers.

The application of highly sensitive methods such as gas (GC) or liquid chromatography (HPLC) with mass spectrometry (MS) naturally allows for the discrimination between the *Papaver* species [12], determining high and low alkaloid content varieties [1,13], the geographical origin of poppy varieties [14], or promising breeding lines [15,16]. Nevertheless, it seems that these methods are not sufficiently reliable at the fine scale necessary for the discrimination of individual varieties. This is probably due to the complex influence of many environmental conditions changing the alkaloid content [6,17,18] leading to a great intra-cultivar variability. Here, DNA markers, which are not influenced by these factors, have proven to be helpful. The first attempts to evaluate genetic variability and diversity among varieties were done by random amplified polymorphic DNA (RAPD), inter-simple sequence repeat (ISSR), or amplified fragment length polymorphism (AFLP) methods [19–22] because these methods do not require a priori knowledge about the genome. Simultaneously, with an accumulation of sequencing data, microsatellite markers were also developed [23–26].

Microsatellites, or simple sequence repeats (SSR), belong to broad family of tandemly repetitive DNA. They are characterized by a short nucleotide motif of 1–6 bp (also called repeat unit) with a total length up to 100 bp [27,28], where their further classification depends on the number of the repeat units, the specific nucleotide motif, the composition of the whole repetition and the localization within the cell or the genome. They have many desirable attributes to make them suitable as molecular markers. They are highly abundant, (nonrandomly) distributed throughout the whole genomes of both prokaryotic and eukaryotic organisms, hypervariable, multiallelic in nature, locus specific, and codominant [29,30]. Moreover, SSR markers are also amenable for multiplexing and high-throughput genotyping [31,32].

When establishing microsatellites as a marker, it is necessary to obtain information about the sequences flanking their own repetitive region. The sequences could be retrieved directly from genomic DNA, cDNA (complementary DNA) or only a short part of cDNA called EST (expressed sequence tag), both of which are created by the reverse transcription of RNA. Thus, one can classify SSR markers as genomic (gSSR) or genic (EST-SSR). Genic SSR markers possess several advantages over genomic SSR markers, as they are considered more reliable and robust, could be potentially directly connected to associated traits because they refer to polymorphism in the expressed portion of their genome and are more transferable across to closely relative species [29,33]. On the other hand, they are usually less polymorphic than their genomic counterparts and their amplicon size can differ from expectation given the presence of introns in flanking regions [29].

The usefulness of SSR markers in the field of plant genetics is proven in their application in the creation of linkage maps, the characterization of genetic diversity in germplasm collections, marker-assisted selection, QTL identification, gene mapping, DNA fingerprinting, paternity analysis, and evolutionary or populations studies, among others [29,32,34,35]. These analyses were made on many plant species, such as *Ambrosia artemisiifolia* [36], *Vigna unguiculata* L. Walp. [37], *Brassica oleracea* [38], *Solanum tuberosum* [39], and *Oryza sativa* L. [40], and demonstrated the versatility of microsatellite markers. The number of SSR markers for the opium poppy is increasing, but they are poorly characterized in most cases; therefore, only a few of them are suitable for variety identification. Thus, our aim was to develop and characterize a new set of polymorphic markers resembling Mendelian inheritance in diploid organisms, amenable for multiplexing and automatic processing of large quantities of samples.

2. Results

2.1. Bioinformatic Analysis

Searching within ~382,000 cDNA sequences for loci with a trinucleotide motif and 10 or more repeats led to 585 loci being found, where for 308 loci (52.6%) primers were designed and 299 (51.1%) of them were unique according to GMATA software (Genome-wide Microsatellite Analyzing Tool Package). Similar searching within BAC (bacterial artificial chromosome) genomic sequences revealed 12 gSSR loci. Primers were designed for all of them, but subsequent analysis left only three (25%) primers fulfilling the quality requirements.

2.2. Marker Testing

As a first step, a total of 200 EST-SSR and three gSSR markers were tested by agarose gel electrophoresis, where 166 (~82%) of them produced a visible band(s). This amplification control was done in two rounds on a subset of eight accessions. The first round of amplification was done at $T_a = 60$ °C and the non-amplifying loci were tested again at $T_a = 57$ °C.

The second step comprised of one (preliminary test with four accessions only) or two (all accessions) rounds of fragment analysis by capillary electrophoresis for 38 promising markers (~18.7%), but only 17 (~8.4%) of them fulfilled all the requirements for high-quality markers. The potential markers were mostly excluded due to multiple peak profiles, suggesting amplification of two or more loci at once, even after several rounds of primer redesign. The results of studies involved in SSR markers development are summarized for comparison purposes in Table 1.

Table 1. Comparison of studies reporting the development of simple sequence repeat (SSR) markers in opium poppy.

Nr. of Tested Markers	Amplifying Markers	Useful Markers	Total Nr. of Alleles*/Fragments**	Mean per Locus (Papaveroidae)	Mean per Locus (<i>P. somniferum</i>)	Transferability ^a (%)	Study
100	96	53	207 **	3.9	1.7	97	[24]
93	76	67	562 **	8.4	3.77	88.7	[25]
22	17	6	17 *	-	2.83	33	[26]
12	12	8	20 *	-	2.5	-	[23]
203	166	17	118 *	6.94	5.18	29.41	this study

^a cross-species transferability.

2.3. Marker Polymorphism, Transferability, and Reproducibility

The analyzed samples were divided into three groups in order to gain a better understanding of the distribution of genetic variability (Table 2). The first group include all accessions and represents the total variability detected by our microsatellite panel; the second group comprises of the accessions

of *P. somniferum*, irrespective of purpose class; and the last group contains only the cultivars of *P. somniferum* amenable for culinary usage, although some of them are classified as dual purpose or even industrial. In total, 118 alleles were detected with a median of seven alleles and three to ten alleles per locus. As expected, the number of alleles in other groups decreased to 88 (*P. somniferum* only) or 77 (*P. somniferum* culinary cultivars) alleles with a median of five and four alleles with three to nine and two to nine alleles per locus, respectively. The observed heterozygosity was generally very low and ranged between 0.048–0.190 with a median of 0.110 for the first group, between 0.026–0.158 with a median of 0.050 for the second group, and between 0.000–0.176 with median 0.060 for the third group (Table 2). We also calculated the PIC (polymorphic information content) to get a different kind of estimation of marker polymorphism. Its values varied between 0.245–0.757 with a median of 0.575 for the first group, between 0.164–0.764 with a median of 0.517 for the second group, and between 0.081–0.779 with median of 0.468 (Table 2).

Table 2. Number of alleles per locus (k), observed heterozygosity (Ho) and polymorphic information content (PIC) for all three groups of accessions.

Marker	All Accessions			<i>P. somniferum</i>			<i>P. somniferum</i> -Culinary		
	k	H _O	PIC	k	H _O	PIC	k	H _O	PIC
OPEST026	10	0.190	0.724	8	0.158	0.682	7	0.176	0.662
OPEST048c	9	0.114	0.599	7	0.053	0.536	6	0.029	0.450
OPEST051c	5	0.077	0.538	5	0.079	0.517	5	0.088	0.485
OPEST053c	10	0.109	0.674	6	0.053	0.598	6	0.059	0.551
OPEST061	7	0.119	0.694	6	0.132	0.672	6	0.118	0.654
OPEST081c	8	0.190	0.559	6	0.158	0.526	5	0.147	0.503
OPEST086d	3	0.068	0.245	3	0.026	0.164	2	0.029	0.081
OPEST099	8	0.109	0.590	3	0.053	0.428	3	0.029	0.425
OPEST102b	9	0.087	0.505	5	0.053	0.371	4	0.029	0.344
OPEST106	6	0.093	0.703	4	0.079	0.631	4	0.059	0.615
OPEST120b	7	0.159	0.605	3	0.105	0.478	3	0.088	0.432
OPEST126b	6	0.049	0.501	5	0.053	0.422	3	0.029	0.381
OPEST131	6	0.111	0.518	4	0.105	0.399	4	0.088	0.418
OPEST156	3	0.053	0.366	3	0.053	0.366	3	0.029	0.368
OPEST169	6	0.075	0.531	5	0.079	0.517	2	0.088	0.485
OPEST177b	6	0.048	0.349	6	0.053	0.228	5	0.000	0.105
OPGSSR001	9	0.154	0.693	9	0.158	0.699	9	0.088	0.654
mean	6.94	0.104	0.564	5.18	0.084	0.500	4.53	0.068	0.466
median	7	0.101	0.575	5	0.066	0.517	4	0.059	0.468

Cross species transferability was tested on nine accessions of five species of the genus *Papaver* and *Argemone*, both belonging to the Papaveraceae family (Table S1). The highest transferability was reached for *P. nudicaule*, with 13 amplifying markers, followed by *P. orientale* (10), *P. glaucum* (9), *P. rhoeas* (8), and *A. mexicana* (2). Only 1 out of the 17 markers, namely OPEST099, amplified for each tested species and another five (OPEST48c, OPEST53c, OPEST102b, OPEST120b, and OPEST131) amplified for each species but one. On the other hand, only the marker OPEST156 was strictly specific for *P. somniferum* and did not produce any detectable signal for the other species in both labs. However, five out of seventeen markers (29.41%) amplified for all species within the *Papaver* genus, seven markers (41.18%) amplified for all species but one within the *Papaver* genus, 12 markers (70.60%) amplified for two out of four *Papaver* species, 15 markers amplified (88.24%) for at least one species within the *Papaver* genus, and only two markers (11.76%) produced a signal peak across the genera *Papaver* and *Argemone*.

For the *P. somniferum* dataset, the mean e_a (mean error per allele) and e_l (mean error per locus) reached the same value equal to 0.17%. In other species, reproducibility was substantially lower with mean $e_a = 27.18\%$ and mean $e_l = 32\%$. The detailed information about the error rate per allele and per locus for the other species dataset is presented in Table S2.

2.4. Assessment of Genetic Diversity and Discrimination Power of Markers

Resolving the similarities/dissimilarities between the accessions was done with the help of two exploratory methods, namely hierarchical cluster analysis (CLU) and principal component analysis (PCA). As the best clustering algorithm, UPGMA (unweighted pair group method with arithmetic mean) was selected according to the highest value of CC (cophenetic correlation coefficient) and lowest *delta* parameter (Table S3). As can be seen in Figure 1a,b, the new SSR markers were able to discriminate *P. somniferum* accessions (green, blue color) from other species (red color) with the exception of one *P. nudicaule* sample. Moreover, within *P. somniferum* accessions, the ornamental varieties (blue color) were found to be most similar to each other, although they did not create one enclosed cluster. Furthermore, according to CLU (Figure 1a) it is visible that new markers discriminate all but two pairs of accessions—that is, Aplaus vs Orfeus and Orel vs Sokol (all *P. somniferum* varieties). There is another pair of accessions (Gerlach-2 and Gerlach-34) with no differences, but these accessions belong to the same variety. Interestingly, the same situation was not observed for the Bergam, Maraton, Opal, or Orel varieties (Figure 1a).

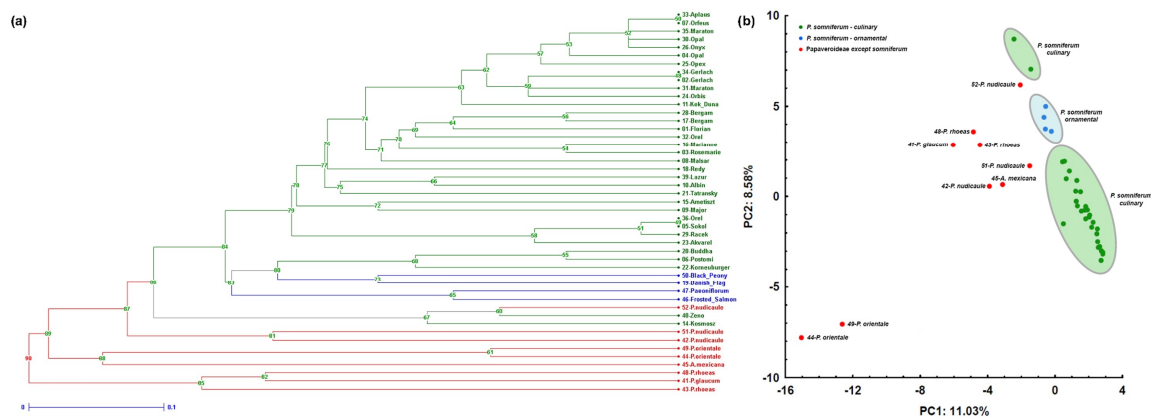


Figure 1. (a) Hierarchical cluster analysis (CLU) and (b) principal component analysis (PCA) analysis where *P. somniferum* accessions (green) with distinguished ornamental varieties (blue) and other species (red) are shown.

A similar result was obtained by PCA analysis, where the projection onto two principal axes is shown in Figure 1b. Two synthetic variables explained 19.61% of the total amount of variability, where the first axis divided *P. somniferum* accessions from the other species and the second axis further separated accessions both within *P. somniferum* and the other species. We also observed an analogous result during CLU for one *P. nudicaule* sample, showing high similarity to some *P. somniferum* accessions, and the existence of two *P. somniferum* “outliers” represented by the Kozmosz and Zeno varieties.

Identity analysis revealed that an exact match in all tested loci was found for the same pairs of accessions as mentioned for CLU analysis (Table S4). It also means that 41 unique profiles out of 47 accessions were found. When allelic mismatch starting from one to five alleles was allowed, the number of same pairs of accessions rose to 9, 15, 25, 37, and 52, respectively (Tables S4–S9a,b).

3. Discussion

Data mining allows for the fast designing of tens to hundreds of thousands of SSR markers in silico, but their quality and usefulness can only be proven by empirical analysis in the “wet” lab. We tried to raise our chances of obtaining a set of quality markers with sufficient polymorphism and potentially single locus specificity by setting up several criteria. At the very beginning, we limited our search to only microsatellites with a trinucleotide motif and no less than 10 repeats. Besides SSRs with a trinucleotide motif, we could also use hexanucleotide microsatellites, but microsatellite abundance with longer motifs typically decreases within the genome [41,42], and information about their polymorphism

is lacking in the literature. The reason for using microsatellites with a motif of a multiple of three is obvious, because potential mutations preserving the reading frame are less severe than frameshift mutations. Moreover, it is well known that microsatellite polymorphism increases with total length, due to the higher chance of slippage of DNA polymerase during replication [30,43].

As we found, the second goal (single locus specificity) imposed a limit on the markers of choice. *P. somniferum* has a complex, not yet fully understood evolutionary history, probably including genomic introgression of two or even three species [9]. Thus, *P. somniferum* ($2n = 22$) could be classified as an alloaneuploid, where its basic chromosome number $x = 11$ is derived from the more ancestral $x = 7$, given the existence of a hypothetical triploid hybrid that may have preceded the speciation of *P. somniferum* [9]. Although Marciano et al. [6] mentioned that the diploid nature of *P. somniferum* simplified the workflow for genetic identification and compared it with the development of markers for human identification, we worried that they probably underestimated the evolutionary complexity of the genome development typical for flowering plants and the consequences of allopolyploidy [44,45]. Namely, for *P. somniferum*, analysis of a 401 kb long genomic sequence of gene cluster for noscapine synthesis was performed, revealing gene duplicities, structural rearrangement, and many DNA or retrotransposable elements [5]. This implies that similar structures should not be limited to only one genomic region and that analogous situations could occur throughout the whole poppy genome. Despite such unfavorable conditions, our goal was to develop markers resembling typical Mendelian inheritance in a diploid organism.

The empirical verification of locus specificity was ensured by two complementary approaches. The first idea was that most of our accessions were varieties and, by definition, should fulfil the DUS criterion. Thus, for varieties of autogamous plants, like *P. somniferum* is, we expected a very high degree of homozygosity, i.e., a low heterozygosity (Table 2). This was verified after the analysis of all the final markers, where 24 out of 34 (~71%) *P. somniferum* varieties (including the varieties for ornamental purposes) were homozygotes for all markers and most of the rest (six out of ten) were heterozygotes for only one or two markers (i.e., homozygotes for 15 or 16 markers). It helped to eliminate markers where a clear profile and a maximum of two peaks per sample were detected, but all or almost all plants seemed to be heterozygotes, which was improbable given our knowledge about the tested material. Only Lee et al. [26] also tried to exclude nonspecific markers by the elimination of markers with three or more peaks per sample, but they were unable to recognize and discard seemingly locus-specific markers with a heterozygous-like profile because they tested unknown genetic material obtained through narcotic seizures. The second idea relied on the verification of locus specificity through the direct sequencing of PCR products, followed by comparison against reference sequences obtained during data mining, where only sequences with little or no noise were accepted. This noise in several markers suggested weak co-amplification from other locus/loci, but the low amount of such sequences was not detectable, even with highly sensitive capillary electrophoresis, and thus this seems to be unimportant for methods based on fragment analysis.

At the beginning of marker testing, we obtained the same percentage of successful amplification (82%) as Selale et al. [25], but the number of retained markers was lower. Selale et al. [25] tested 93 EST-SSR markers and retained 67 markers (72%), versus 203 tested markers and 17 (8.4%) retained markers in this study (Table 1). Moreover, when we take into account the total number of tested primer pairs (272) for 203 markers, then the percentage of retained markers drops to 6.25%, a ratio of 1:15 (accepted: discarded markers). Celik et al. [24] also reported a higher yield of markers in opium poppies when they tested 100 genomic SSR and 53 (53%) were found to be useful (Table 1). Such discrepancy is probably caused by a subtle yet important difference in the methodology of the mentioned studies. They analyzed mixtures of 10 or 20 plants per sample and each allele was counted by the presence (1) or absence (0) of a peak signal of appropriate size, with no restriction on how many peaks (i.e., alleles) were allowed per marker. Such an approach could be very useful when one wants to quickly retrieve many markers, obtain a relatively large amount of data on a genome-wide level, and overcome the problem with polyploidy, but everything comes at a price. There is a loss of

the codominant nature of the SSR markers and the information about the source of polymorphism. This could lead to bias in the analyses where a higher level of precision is needed (e.g., mapping or individual genotyping). This approach is also technically challenging for routine screening because such markers are not easily amenable for multiplexing without an elevated risk of error, given the higher order of interactions between the primers with unknown specificity, and the individual analysis of tens of such markers is costly and time consuming.

A total of 17 selected markers identified 118 alleles for all accessions, with a median of seven alleles per locus, which is comparable even with the variability of the genomic SSR markers in populations of the allogamous species *P. rhoeas* [46]. Nonetheless, variability within *P. somniferum* itself is more important for our purpose, where the number of alleles decreased by about 25% (88 alleles, median = 5) for all *P. somniferum* accessions and about 35% when we take account only culinary varieties (77 alleles, median = 4). Unfortunately, the direct comparison of variability within *P. somniferum* species is only possible with the study of Lee et al. [26], where they found at six loci 17 alleles with a mean of 2.83 alleles per locus, and Mičianová et al. [23] who reported 20 alleles at 8 loci with a mean of 2.5 alleles per locus (Table 1). More extensive studies [24,25] with a similar number of samples adopted a different strategy of marker development and analysis (see the previous part of discussion) and we can only perform an indirect and thus biased comparison when we accept the premise that one fragment is equal to one allele. In the former study [24], 207 (all accessions) and 90 (*P. somniferum* accessions) fragments per 53 gSSR markers were detected, with a mean of 3.9 and 1.7 fragments per marker, respectively. The latter study [25] reported 562 (all accessions) and 253 (*P. somniferum*) fragments per 67 EST-SSR markers, with a mean of 8.4 and 3.77 fragments per markers, respectively (Table 1). We also performed an empirical evaluation of 21 published markers (Table S10) where all but two (psom 12 and psom 17) of the markers were discarded due to multiple peak profiles or insufficient polymorphism (our internal limit was at least 3 alleles per marker for *P. somniferum* accessions) and thus we were unable to compare it with our SSR panel. All this information showed the same or higher levels of performance in newly developed markers and a congruence about relatively low genetic variability within varieties of *P. somniferum*.

Cross-species transferability of new markers within the genus seems to be rather low (29.41%) when we accept the published range from 23% to 96% [33], and this probably depends on the sequence conservation of the involved genes given their functional importance and the phylogenetic distances of compared species [29]. Surprisingly, a much higher level of transferability within the *Papaver* genus was reported not only for the EST-SSR markers (97%) [25], but also for genomic SSR (at least 88.7%) [24]. It seems that these markers targeted more conservative regions of the genome, both inter and intragenic, or lacked specificity. In contrast, Lee et al. [26] mentioned two out of six markers (33%) successfully amplifying for at least one species of the *Papaver* and *Escholzia* genera, and one marker (16.67%) amplifying both for the genera *Papaver* and *Escholzia*, which is closer to our results. Although the number of amplifying markers in other species is lower, they still represent a valuable source of markers for the less studied *Papaver* species, where a sufficient amount of sequences and markers is still lacking.

As far as we know, this is the first study reporting marker reproducibility within or between papaver species and thus we are unable to directly compare our results with other articles dealing with *P. somniferum* species. It seems that our mean error rate per locus (in our case $e_l = e_a = 0\text{--}2.86\%$, mean = 0.17%) is lower than the usual mean error rate, which ranges from 0.5% to 1% [47]. This result is also in concordance with newer studies in olive (*Olea europaea*) where e_a ranged from 0.7% to 6.2% with a mean of 2.26% [48] or *Alnus incana* ($e_a = 0\text{--}8\%$, mean = 1.4%) and *Alnus glutinosa* ($e_a = 0\text{--}6.3\%$, mean = 1.3%) [49]. Unfortunately, none of these studies mentioned e_l .

A substantially different situation was revealed by comparison with other species' datasets, where all metrics showed a very high error rate (Table S2). The main reason probably lies in the evolutionary distances of the assessed species, causing a lower specificity and thus reduces the amplification capability of markers primarily developed for *P. somniferum*. This corresponds with the

most frequent source of error, which was amplification failure. Other explanations include human factors, cross-contamination, or different criteria for accepting a signal as an allele; but these kinds of problems might not be restricted to other species datasets, and it was not noticed for *P. somniferum* datasets. It is also necessary take into account that few samples per marker were evaluated and any mismatch led to a high error rate. As a result of these findings, we advise any usage of our markers in species other than *P. somniferum* to be taken with caution. This is also the reason why a more conservative approach was chosen and a lower transferability was reported, because only markers amplifying in both labs are labelled by a + sign (Table S1).

The new microsatellite system distinguished all accessions except for two pairs of varieties (Orel vs Sokol and Aplaus vs Orfeus), as proved by CLU and PCA analyses (Figure 1a,b). Unfortunately, little is known about the breeding history or pedigree of the mentioned cultivars, but both Orel and Sokol are white-seeded poppies. Furthermore, there is information available which shows that Orel and another white-seeded variety, Racek, were selected as individual plants from an unstable population of a local (unnamed) variety [50], so they should share some genetic similarities (Figure 1a). These cultivars have the same breeder (Oseva Pro Ltd., Opava, Czech Republic), thus we can speculate that Sokol also has its origin in the same local variety; we can then talk about isolines, which could be difficult to resolve, even with additional markers. The second pair of cultivars belong to a category of blue-seeded poppies, but they have different breeders, and no information about the breeding history of Aplaus has been published. This result suggests that the gene pool actually used for new varieties is quite narrow and/or that our system lacks robustness, as it was implicated by the identity analysis (ID) paired comparison of varieties. Although we could argue that the probability of a random match is much lower (e.g., $pID = 1.09 \times 10^{-16}$; Table S4) than the probability ($pID = 1.04 \times 10^{-3}$) published by Mičianová et al. [23], one should treat such values carefully and also look behind the numbers to see their meaning. Such estimators were developed for application in forensic genetics or studies of wild populations and calculate the probability of a random match within an idealized population according to the Hardy–Weinberg law and other assumptions (see more about the topic in [51,52]). In contrast, our study and others analyzed varieties bred by man and they cannot be considered as a population of randomly mating individuals by any means and many varieties could share a substantial portion of their genomes due to their origin, breeding history, and pedigree. This is the reason why we disagree with the argumentation in [23] and, despite the very low p-values for a random match in our case, we regard such values as uninformative and, thus, the obtained result leads us to have doubts about the system's robustness and to call for a higher number of markers. The last interesting outcome from multivariate analysis showed that there were differences between accessions of the same variety, such as Bergam, Maraton, Opal, or Orel (Figure 1a). At the present time, it is not possible to conclude what the source of these differences is, but probable explanations include some degree of internal variability within varieties, human error (e.g., wrong label, unintended crossing in gene bank), or various degrees of seed stock influencing the purity of the seeds (C1 vs E degree).

4. Materials and Methods

4.1. Plant Material

In total, 29 varieties of *Papaver somniferum* (opium poppy), which included all three types of variety classified according to purpose (culinary, dual purpose, pharmaceutical/industrial), 4 varieties of *P. somniferum* typically used for ornamental purpose, 4 other species of the genus *Papaver* (*P. glaucum*, *P. nudicaule*, *P. orientale*, and *P. rhoeas*) and 1 distant relative species of the genus *Argemone* (*Argemonemexicana*), were collected. In total, 47 accessions were analyzed, because the same varieties from multiple sources were received in several cases (Table 3). All accessions were obtained in the form of seeds. These seeds were put into separated flowerpots and planted in a greenhouse under regulated conditions. Later, a few leaves of each plant were cut in the plant's phenological phase 13–15 according to the BBCH (Biologische Bundesanstalt, Bundessortenamt und Chemische Industrie)

scale [53] and were utilized as source material for the isolation of DNA. Thus, each sample contained DNA from one plant.

Table 3. Analyzed accessions of *P. somniferum* (opium poppy) and related species.

Accession Number	Assigned Code	Species	Genebank Evidence Number	Variety	Source
1	FLO	<i>P. somniferum</i>	15O0800171	Florian	genebank Opava ^a
2	GER	<i>P. somniferum</i>	15O0800148	Gerlach	genebank Opava ^a
3	ROS	<i>P. somniferum</i>	15O0800164	Rosemarie	genebank Opava ^a
4	OPA	<i>P. somniferum</i>	15O0800169	Opal	genebank Opava ^a
5	SOK	<i>P. somniferum</i>	15O0800179	Sokol	genebank Opava ^a
6	POS	<i>P. somniferum</i>	15O0800195	Postomi	genebank Opava ^a
7	ORF	<i>P. somniferum</i>	15O0800190	Orfeus	genebank Opava ^a
8	MAL	<i>P. somniferum</i>	15O0800183	Malsar	genebank Opava ^a
9	MAJ	<i>P. somniferum</i>	15O0800182	Major	genebank Opava ^a
10	ALB	<i>P. somniferum</i>	15O0800159	Albín	genebank Opava ^a
11	KEK	<i>P. somniferum</i>	15O0800093	Kék Duna	genebank Opava ^a
14	KOS	<i>P. somniferum</i>	15O0800173	Kozmosz	genebank Opava ^a
15	AME	<i>P. somniferum</i>	15O0800198	Ametiszt	genebank Opava ^a
16	MAE	<i>P. somniferum</i>	15O0800031	Marianne	genebank Opava ^a
17	BER	<i>P. somniferum</i>	15O0800184	Bergam	genebank Opava ^a
18	RED	<i>P. somniferum</i>	15O0800189	Redy	genebank Opava ^a
19	DAN	<i>P. somniferum</i>		Danish Flag	market
20	BUD	<i>P. somniferum</i>	15O0800185	Buddha	genebank Opava ^a
21	TAT	<i>P. somniferum</i>		Tatranský	Červený Újezd ^b
22	KOR	<i>P. somniferum</i>		Korneuburger	Červený Újezd ^b
23	AKV	<i>P. somniferum</i>		Akvarel	Červený Újezd ^b
24	ORB_II	<i>P. somniferum</i>		Orbis	CISTA ^c
25	OPE_II	<i>P. somniferum</i>		Opex	CISTA ^c
26	ONY_II	<i>P. somniferum</i>		Onyx	CISTA ^c
28	BER_II	<i>P. somniferum</i>		Bergam	CISTA ^c
29	RAC_II	<i>P. somniferum</i>		Racek	CISTA ^c
30	OPA_II	<i>P. somniferum</i>		Opal	CISTA ^c
31	MAR	<i>P. somniferum</i>	15O0800181	Maraton	genebank Opava ^a
32	ORE	<i>P. somniferum</i>	15O0800187	Orel	genebank Opava ^a
33	APL_II	<i>P. somniferum</i>		Aplaus	CISTA ^c
34	GER_II	<i>P. somniferum</i>		Gerlach	CISTA ^c
35	MAR_II	<i>P. somniferum</i>		Maraton	CISTA ^c
36	ORE_II	<i>P. somniferum</i>		Orel	CISTA ^c
39	LAZ	<i>P. somniferum</i>		Lazur	Červený Újezd ^b
40	ZEN	<i>P. somniferum</i>		Zeno	Červený Újezd ^b
41	GLA	<i>P. glaucum</i>		seed mixture	market
42	GNO	<i>P. nudicaule</i>		Gnome	market
43	RHO	<i>P. rhoeas</i>		seed mixture	market
44	PIZ	<i>P. orientale</i>		Pizziato	market
45	MEX	<i>A. mexicana</i>		seed mixture	market
46	FRO	<i>P. somniferum</i>		Frosted Salmon	market
47	PAE	<i>P. somniferum</i>		Paeoniflorum	market
48	DAW	<i>P. rhoeas</i>		Dawn Chorus	market
49	SCHA	<i>P. orientale</i>		Scharlach	market
50	PEO	<i>P. somniferum</i>		Black Peony	market
51	NUD_I	<i>P. nudicaule</i>		seed mixture	market
52	NUD_II	<i>P. nudicaule</i>		seed mixture	market

^a part of Oseva Ltd company and Plant oil research institute, ^b research station Červený Újezd, ^c Central Institute for Supervising and Testing in Agriculture.

DNA isolation was done by the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's recommendations. The quality and quantity of the isolated DNA was verified by

UV spectrometry (S-111107AW nanophotometer, Implen) and electrophoretic separation in 1% (*w/v*) agarose gel.

4.2. Bioinformatic Analysis and Primers Design

Filtered and clipped cDNA sequences were downloaded from the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (NCBI). Only the sequences produced by pyrosequencing were chosen, because 454 sequencing typically produces longer reads than Illumina and thus there is a higher chance of obtaining a sufficiently long stretch of simple repeats. About 3.7 million cDNA sequences, with a length of 1.89 Gb in total, and 15 genomic sequences from the BAC library, with a total length 2.7 Mb, were analyzed. This was followed by several rounds of data processing and assembling with the help of the online bioinformatics tool EGassembler [54], resulting in 382531 so-called contigs and singletons, with a total length 187 Mb.

SSR mining was done by GMATA software [55], where microsatellites with a trinucleotide motif and 10 or more repeats were utilized for marker development. The same software was used for the preliminary primer design for amplicons within the range 100–300 bp and T_m 60 °C. When necessary, other primers were designed using Primer3 plus [56,57] and OligoEvaluator software (an online web tool provided by Sigma-Aldrich, St. Luis, MO, USA).

Moreover, 21 genomic SSR (gSSR) or EST-SSR markers (Table S10) published by Celik et al. [24] and Lee et al. [26] for *P. somniferum* and Kati et al. [46] for *P. rhoeas* were chosen and analyzed for comparison purposes.

4.3. SSR Loci Amplification

A PCR reaction mixture, with a total volume of 10 µl, included 10 ng of DNA, 0.2 µM F and R primer (Table 4, Tables S11 and S12), and 1× Multiplex PCR Master Mix Plus (Qiagen, Hilden, Germany). The temperature profile was as follows: 1 cycle of pre-denaturation at 95 °C for 600 s, 35 cycles of denaturation at 94 °C for 30 s, annealing at 57 °C or 60 °C for 90 s and elongation at 72 °C for 60 s, and 1 cycle of final elongation at 60 °C for either 600 s (for agarose gel electrophoresis) or 4800 s (for capillary electrophoresis). The amplification was done by a C-1000 (Bio-Rad, Hercules, CA, USA) or T-gradient Thermo (Analytic Jena, Jena, Germany) thermocycler. The presence of amplicons and the expected size control was done by electrophoretic separation in 2% (*w/v*) agarose gel with a size standard GeneRuler 100 bp DNA Ladder (Thermo Fisher Scientific, Waltham, MA, USA). Only the amplicons with one band and in the expected size range were chosen for the next step of analysis by capillary electrophoresis. There were also several cases when one band was observed but the fragment size was two to eight times higher than expected. Such fragments were sequenced, and new primers were eventually designed prior to capillary electrophoresis.

4.4. Fragment Analysis

The chosen markers were analyzed by capillary electrophoresis and thus F primer was fluorescently labelled by 6FAM, VIC, NED or PET dye. The amplification condition and temperature profile of PCR was the same as for agarose gel electrophoresis. The verification of the results and the reproducibility control was ensured by analysis of all accessions, both in the molecular laboratory of the Department of Genetics and Breeding (DGB) and in the molecular laboratory of the Crop Research Institute (CRI) on different capillary instruments. In the former laboratory, PCR products were diluted with deionized water at a ratio of 1:99–199 depending on tested marker and 1 µl of diluted PCR product was mixed with 12 µL Hi-Di formamide (Thermo Fisher Scientific, Waltham, MA, USA) and 0.2 µl of GeneScan LIZ600 size standard (Thermo Fisher Scientific, Waltham, MA, USA). The samples were separated by ABI PRISM 310 (Thermo Fisher Scientific, Waltham, MA, USA) with a 47 cm-long capillary filled with POP4 polymer (Thermo Fisher Scientific, Waltham, USA). For fragment data analysis and allele identification, the GeneMapper v4.1 program (Thermo Fisher Scientific, Waltham, MA, USA) was utilized. In the latter laboratory, the PCR products were diluted with deionized water in a ratio of

1:65–199 and 1 µL of appropriate product was mixed with 10 µL of Hi-Di formamide (Thermo Fisher Scientific, Waltham, MA, USA) and 0.2 µL of GeneScan LIZ500 size standard (Thermo Fisher Scientific, Waltham, MA, USA). Separation was done by ABI 3500 (Thermo Fisher Scientific, Waltham, USA) with a 61 cm-long capillary filled with POP7 polymer (Thermo Fisher Scientific, Waltham, MA, USA). Blind analysis irrespective of the DGB lab results was done in GeneMapper v5.0 (Thermo Fisher Scientific, Waltham, MA, USA).

Table 4. Information about primers, fluorescent dyes, and size range of new genic (EST)-SSR markers.

Marker	Dye		Primer sequence 5' - 3'	Size range [bp]	GenBank AC Number
OPEST026	6FAM	F-	GTGAGGAGGACGAGCTTTTG	105–148	MK744101
		R-	gtttcttCCGTTGTAAAATACCGACTGC		
OPEST048c	PET	F-	CGTGAGAAGCTAGAACAGAAAGA	174–204	MK744102
		R-	TCGTTCACTGAGTTCTGATATGA		
OPEST051c	VIC	F-	GGGTTCTTTTGTCTACTTCTTTCTT	147–162	MK744103
		R-	AAGGTGTCGGTGCCCAGC		
OPEST053c	NED	F-	TCAATACCCACAAAAGGAGGA	172–203	MK744104
		R-	gtttcttTCAAGACAAAAGAAACCAAGCCA		
OPEST061	6FAM	F-	GGCTGCTGCTTCTTTTCATC	191–236	MK744105
		R-	ATAGGGCAAACCTGCCTGCTA		
OPEST081c	PET	F-	AGTAAAACGATCCGTACCTACCTGA	133–176	MK744106
		R-	CGTTTTTCTACAGGGTTGATTTCTGA		
OPEST086d	PET	F-	ACTTTTCCCCCTCTTCAGTAGC	223–244	MK744107
		R-	TCCAGTCCACATCAGGATCA		
OPEST099	6FAM	F-	TTAACAGATCCGCATTCCA	262–288	MK744108
		R-	CACCGATTGTACCACGAAGA		
OPEST102b	VIC	F-	CGCCACCACATATTTCTCTG	184–206	MK744109
		R-	GGTTGTCGGCATAGAAGGAA		
OPEST106	6FAM	F-	CACCAAATCTCATTGCCTGA	166–191	MK744110
		R-	CCCTAATCGGATGGATCAAA		
OPEST120b	6FAM	F-	TAGTGGTTGCTCGTAGCGTC	138–156	MK744111
		R-	TCACGGTTCTTCTATCATGGTG		
OPEST126b	6FAM	F-	GTTTCTCACGGAGGGATTG	206–228	MK744112
		R-	CCGTTTCCCAACTTCGTAGA		
OPEST131	VIC	F-	GTTCCAAACCACCAACCACA	224–250	MK744113
		R-	TTGTGAGGCCCTAGAGAGGA		
OPEST156	6FAM	F-	TTTAGCTTACAATGGTGGGAGA	264–270	MK744114
		R-	GAAACCGTAGCCAGGTGAAA		
OPEST169	VIC	F-	TCCAACGCAAGCAATTACAA	165–205	MK744115
		R-	GCCACTTCGTAACCCAGGTA		
OPEST177b	VIC	F-	TCTCCGTAACCTGAAGAACAGA	96–112	MK744116
		R-	TGGTGGCAGTGAATTTGAT		
OPGSSR001	VIC	F-	TGCGGCTTCTAATCATCCTT	218–244	MK744117
		R-	CCATCAACTTCGCACAGCTA		

Due to the different conditions in both labs, which influenced fragment mobility and thus allele size estimation, the allelic size standards were developed, and the nomenclature defined for each allele in the DGB lab. Allelic ladders (ALs) containing all or almost all of the detected alleles depending on the marker were created by mixing the PCR products of the appropriate genotypes with a ratio of 2:1 of heterozygotes and homozygotes (see example ladder in Figure S1). The pooled PCR products were further diluted in a ratio of 1:99–350 with deionized water and 1 µL of appropriate ALs was prepared and separated simultaneously with the tested samples under the same conditions as mentioned earlier.

4.5. Sequencing

The locus specificity of each marker tested by capillary electrophoresis was checked by the Sanger sequencing method. All markers were amplified in multiple copies, separated in 1% (*w/v*) agarose gel, and excised by clean scalpel after 1 h separation. DNA purification was performed using the GenJet Gel Extraction kit (Thermo Fisher Scientific, Waltham, MA, USA) according to the manual's recommendations. The clean amplicons, together with the specific primers, were sent to the Eurofins Genomics (Eurofins Genomics Germany GmbH, Ebersberg, Germany) sequencing service and were bidirectionally sequenced. The succeeding quality evaluation, assembly, alignment with reference sequence, and necessary manual correction of all sequences was done by Sequencing Analysis Software v5.2 (Thermo Fisher Scientific, Waltham, MA, USA), BioEdit Software v7.0.9.0 [58] and an online version of MUSCLE software [59] available on webpages of The European Bioinformatics Institute (EMBL-EBI).

4.6. Statistical Data Analysis

Descriptive statistics, including the number of alleles with frequencies estimation and identity analysis (ID), was performed in Cervus v3.0.3 software [60]. ID was done by a series of paired comparisons of each accession with a 0–5 allowed allele mismatch. STATISTICA software v13.3 [61] was utilized for exploratory data analysis (EDA) by principal component analysis (PCA) and DARwin5 software [62] for EDA by the hierarchical clustering (CLU) method. For the purpose of CLU analysis, the microsatellite data were transformed to binary data with presence (1) and absence (0) coding for alleles. The created distance matrix was based on the Dice coefficient [63] with 10,000 bootstrap steps. The best clustering method was selected according to the highest value of the cophenetic correlation coefficient (CC) and lowest *delta* parameter in the NCSS program [64]. As an input for PCA analysis, a trinary matrix of individual allelic frequencies was created, where the frequency had a value of 1 or 0 for presence or absence of the appropriate allele for homozygotes, whereas heterozygotes had an allele presence value of 0.5.

For the reproducibility analysis, two metrics quantifying error rates were utilized—the error rate per allele (e_a) and the error rate per locus (e_l) according to Pompanon et al. [47]. Each error rate was calculated as the number of mismatches between reference genotypes (here DGB dataset) and replicates (here CRI dataset) where n genotypes had been genotyped t times. The analysis was done separately for *P. somniferum* and other species.

The mean error rate per allele (e_a) is expressed as a ratio between the number of allelic mismatches (m_a) and the number of replicated alleles ($2nt$ for diploid organisms) [47].

$$e_a = m_a/2nt \quad (1)$$

The mean error rate per locus (e_l) is expressed as a ratio between the single-locus genotypes, including at least one allelic mismatch (m_l), and the number of replicated single-locus genotypes (nt) [47].

$$e_l = m_l/nt \quad (2)$$

Supplementary Materials: The following are available online at <http://www.mdpi.com/2223-7747/9/1/10/s1>, Figure S1: Example of allelic ladder for marker OPEST48c, Table S1: Transferability of EST-SSR markers among other species of Papaveraceae family, Table S2: Error rate per allele and per locus for other species datasets, Table S3: CC and delta values for each tested clustering method, Table S4: Identity analysis with 0 allelic mismatches allowed, Table S5: Identity analysis with one allelic mismatch allowed, Table S6: Identity analysis with two allelic mismatches allowed, Table S7: Identity analysis with three allelic mismatches allowed, Table S8a: Identity analysis with four allelic mismatches allowed, Table S8b: Identity analysis with four allelic mismatches allowed, Table S9a: Identity analysis with five allelic mismatches allowed, Table S9b: Identity analysis with five allelic mismatches allowed, Table S10: Empirical evaluation of 21 published SSR markers, Table S11: Tested EST-SSR markers, Table S12: Tested gSSR markers.

Author Contributions: Conceptualization, J.V., P.V. and J.O.; Data curation, J.V.; Formal analysis, J.V. and L.V.; Investigation, D.Č., M.M. and R.Š.; Resources, D.Č., M.M. and P.K.; Software, P.S.; Supervision, J.O.; Validation,

P.S.; Visualization, J.V.; Writing—original draft, J.V.; Writing—review and editing, J.V., P.S., P.V. and J.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Ministry of Agriculture of the Czech Republic, grant number QK1720263 “Diagnostic method for laboratory control of Papaver L. authenticity”.

Acknowledgments: We would like to greatly thank Opava gene bank, namely Andrea Rychlá, and Central Institute for Supervising and Testing in Agriculture, namely Petr Zehnálek, for providing the poppy varieties.

Conflicts of Interest: The authors declare that they have no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Stranska, I.; Skalicky, M.; Novak, J.; Matyasova, E.; Hejnak, V. Analysis of selected poppy (*Papaver somniferum* L.) cultivars: Pharmaceutically important alkaloids. *Ind. Crop. Prod.* **2013**, *41*, 120–126. [[CrossRef](#)]
2. Coward, F.; Shennan, S.; Colledge, S.; Conolly, J.; Collard, M. The spread of Neolithic plant economies from the Near East to northwest Europe: A phylogenetic analysis. *J. Archaeol. Sci.* **2008**, *35*, 42–56. [[CrossRef](#)]
3. Merlin, M.D. Archaeological evidence for the tradition of psychoactive plant use in the old world. *Econ. Bot.* **2003**, *57*, 295–323. [[CrossRef](#)]
4. Kabera, J.N.; Semana, E.; Mussa, A.R.; He, X. Plant secondary metabolites: Biosynthesis, classification, function and pharmacological properties. *J. Pharm. Pharm.* **2014**, *2*, 377–392.
5. Winzer, T.; Gazda, V.; He, Z.; Kaminski, F.; Kern, M.; Larson, T.R.; Li, Y.; Meade, F.; Teodor, R.; Vaistij, F.E.; et al. A *Papaver somniferum* 10-gene cluster for synthesis of the anticancer alkaloid noscapine. *Science* **2012**, *336*, 1704–1708. [[CrossRef](#)]
6. Marciano, M.A.; Panicker, S.X.; Liddil, G.D.; Lindgren, D.; Sweder, K.S. Development of a method to extract opium poppy (*Papaver somniferum* L.) DNA from heroin. *Sci. Rep.* **2018**, *8*, 2590. [[CrossRef](#)] [[PubMed](#)]
7. Lachenmeier, D.W.; Sproll, C.; Musshoff, F. Poppy seed foods and opiate drug testing—where are we today? *Ther. Drug. Monit.* **2010**, *32*, 11–18. [[CrossRef](#)] [[PubMed](#)]
8. Jankovičová, K.; Ulbrich, P.; Fuknová, M. Effect of poppy seed consumption on the positive results of opiates screening in biological samples. *Leg. Med.* **2009**, *11*, S416–S418. [[CrossRef](#)] [[PubMed](#)]
9. Lavania, U.C.; Srivastava, S. Quantitative delineation of karyotype variation in *Papaver* as a measure of phylogenetic differentiation and origin. *Curr. Sci.* **1999**, *77*, 7.
10. Tétényi, P. Opium Poppy (*Papaver somniferum*): Botany and Horticulture. In *Horticultural Reviews*; Janick, J., Ed.; John Wiley & Sons, Inc.: Oxford, UK, 2010; pp. 373–408. ISBN 978-0-470-65062-2.
11. Labanca, F.; Ovesnà, J.; Milella, L. *Papaver somniferum* L. taxonomy, uses and new insight in poppy alkaloid pathways. *Phytochem. Rev.* **2018**, *17*, 853–871. [[CrossRef](#)]
12. Choe, S.; Kim, S.; Lee, C.; Yang, W.; Park, Y.; Choi, H.; Chung, H.; Lee, D.; Hwang, B.Y. Species identification of *Papaver* by metabolite profiling. *Forensic Sci. Int.* **2011**, *211*, 51–60. [[CrossRef](#)] [[PubMed](#)]
13. Schulz, H.; Baranska, M.; Quilitzsch, R.; Schütze, W. Determination of alkaloids in capsules, milk and ethanolic extracts of poppy (*Papaver somniferum* L.) by ATR-FT-IR and FT-Raman spectroscopy. *Analyst* **2004**, *129*, 917–920. [[CrossRef](#)] [[PubMed](#)]
14. Odell, L.R.; Skopec, J.; McCluskey, A. Isolation and identification of unique marker compounds from the Tasmanian poppy *Papaver somniferum* N. *Forensic Sci. Int.* **2008**, *175*, 202–208. [[CrossRef](#)] [[PubMed](#)]
15. Németh-Zámbori, É.; Jászberényi, C.; Rajhárt, P.; Bernáth, J. Evaluation of alkaloid profiles in hybrid generations of different poppy (*Papaver somniferum* L.) genotypes. *Ind. Crop. Prod.* **2011**, *33*, 690–696. [[CrossRef](#)]
16. Gümüşçü, A.; Arslan, N.; Sarihan, E.O. Evaluation of selected poppy (*Papaver somniferum* L.) lines by their morphine and other alkaloids contents. *Eur. Food Res. Technol.* **2008**, *226*, 1213–1220. [[CrossRef](#)]
17. Dittbrenner, A.; Mock, H.P.; Börner, A.; Lohwasser, U. Variability of alkaloid content in *Papaver somniferum* L. *J. Appl. Bot. Food Qual.* **2009**, *82*, 103–107.
18. Ziegler, J.; Facchini, P.J.; Geißler, R.; Schmidt, J.; Ammer, C.; Kramell, R.; Voigtländer, S.; Gesell, A.; Pienkny, S.; Brandt, W. Evolution of morphine biosynthesis in opium poppy. *Phytochemistry* **2009**, *70*, 1696–1707. [[CrossRef](#)]

19. Güçlü, G.B.; Gürkök, T.; Koyuncu, M.; Arslan, N.; Parmaksız, İ. Genetic characterization of Turkish commercial opium poppy (*Papaver somniferum* L.) cultivars using ISSR and SSR markers. *J. New Results Sci.* **2014**, *3*, 48–57.
20. Parmaksız, İ.; Özcan, S. Morphological, chemical, and molecular analyses of Turkish papaver accessions (Sect. *Oxytona*). *Turk. J. Bot.* **2011**, *35*, 1–16.
21. Acharya, H.S.; Sharma, V. Molecular characterization of opium poppy (*Papaver somniferum*) germplasm. *Am. J. Infect. Dis.* **2009**, *5*, 155–160. [[CrossRef](#)]
22. Saunders, J.A.; Pedroni, M.J.; Penrose, L.D.J.; Fist, A.J. AFLP analysis of opium poppy. *Crop. Sci.* **2001**, *41*, 1596. [[CrossRef](#)]
23. Mičianová, V.; Ondreičková, K.; Muchová, D.; Klčová, L.; Hudcovicová, M.; Havrlentová, M.; Mihálik, D.; Kraic, J. Forensic application of EST-derived STR markers in opium poppy. *Biologia* **2017**, *72*, 587–594. [[CrossRef](#)]
24. Celik, I.; Gultekin, V.; Allmer, J.; Doganlar, S.; Frary, A. Development of genomic simple sequence repeat markers in opium poppy by next-generation sequencing. *Mol. Breed.* **2014**, *34*, 323–334. [[CrossRef](#)]
25. Şelale, H.; Çelik, I.; Gültekin, V.; Allmer, J.; Doğanlar, S.; Frary, A. Development of EST-SSR markers for diversity and breeding studies in opium poppy. *Plant. Breed.* **2013**, *132*, 344–351. [[CrossRef](#)]
26. Lee, E.J.; Jin, G.N.; Lee, K.L.; Han, M.S.; Lee, Y.H.; Yang, M.S. Exploiting expressed sequence tag databases for the development and characterization of gene-derived simple sequence repeat markers in the opium poppy (*Papaver somniferum* L.) for forensic applications. *J. Forensic Sci.* **2011**, *56*, 1131–1135. [[CrossRef](#)]
27. Richard, G.-F.; Kerrest, A.; Dujon, B. Comparative Genomics and Molecular Dynamics of DNA Repeats in Eukaryotes. *Microbiol. Mol. Biol. Rev.* **2008**, *72*, 686–727. [[CrossRef](#)]
28. Ellegren, H. Microsatellites: Simple sequences with complex evolution. *Nat. Rev. Genet.* **2004**, *5*, 435–445. [[CrossRef](#)]
29. Kalia, R.K.; Rai, M.K.; Kalia, S.; Singh, R.; Dhawan, A.K. Microsatellite markers: An overview of the recent progress in plants. *Euphytica* **2011**, *177*, 309–334. [[CrossRef](#)]
30. Buschiazzo, E.; Gemmell, N.J. The rise, fall and renaissance of microsatellites in eukaryotic genomes. *Bioessays* **2006**, *28*, 1040–1050. [[CrossRef](#)]
31. Guichoux, E.; Lagache, L.; Wagner, S.; Chaumeil, P.; Léger, P.; Lepais, O.; Lepoittevin, C.; Malausa, T.; Revardel, E.; Salin, F.; et al. Current trends in microsatellite genotyping. *Mol. Ecol. Resour.* **2011**, *11*, 591–611. [[CrossRef](#)]
32. Parida, S.K.; Kalia, S.K.; Kaul, S.; Dalal, V.; Hemaprabha, G.; Selvi, A.; Pandit, A.; Singh, A.; Gaikwad, K.; Sharma, T.R.; et al. Informative genomic microsatellite markers for efficient genotyping applications in sugarcane. *Theor. Appl. Genet.* **2009**, *118*, 327–338. [[CrossRef](#)] [[PubMed](#)]
33. Varshney, R.K.; Graner, A.; Sorrells, M.E. Genic microsatellite markers in plants: Features and applications. *Trends Biotechnol.* **2005**, *23*, 48–55. [[CrossRef](#)] [[PubMed](#)]
34. Vieira, M.L.C.; Santini, L.; Diniz, A.L.; Munhoz, C.D.F. Microsatellite markers: What they mean and why they are so useful. *Genet. Mol. Biol.* **2016**, *39*, 312–328. [[CrossRef](#)] [[PubMed](#)]
35. Hayward, A.C.; Tollenaere, R.; Dalton-Morgan, J.; Batley, J. Molecular Marker Applications in Plants. In *Plant Genotyping*; Batley, J., Ed.; Springer: New York, NY, USA, 2015; Volume 1245, pp. 13–27. ISBN 978-1-4939-1965-9.
36. Meyer, L.; Causse, R.; Pernin, F.; Scalone, R.; Bailly, G.; Chauvel, B.; Délye, C.; Le Corre, V. New gSSR and EST-SSR markers reveal high genetic diversity in the invasive plant *Ambrosia artemisiifolia* L. and can be transferred to other invasive *Ambrosia* species. *PLoS ONE* **2017**, *12*, e0176197. [[CrossRef](#)] [[PubMed](#)]
37. Chen, H.; Wang, L.; Liu, X.; Hu, L.; Wang, S.; Cheng, X. De novo transcriptomic analysis of cowpea (*Vigna unguiculata* L. Walp.) for genic SSR marker development. *BMC Genet.* **2017**, *18*, 65. [[CrossRef](#)] [[PubMed](#)]
38. El-Esawi, M.A.; Germaine, K.; Bourke, P.; Malone, R. Genetic diversity and population structure of *Brassica oleracea* germplasm in Ireland using SSR markers. *C. R. Biol.* **2016**, *339*, 133–140. [[CrossRef](#)] [[PubMed](#)]
39. Kishine, M.; Tsutsumi, K.; Kitta, K. A set of tetra-nucleotide core motif SSR markers for efficient identification of potato (*Solanum tuberosum*) cultivars. *Breed. Sci.* **2017**, *67*, 544–547. [[CrossRef](#)] [[PubMed](#)]
40. Bora, A.; Choudhury, P.R.; Pande, V.; Mandal, A.B. Assessment of genetic purity in rice (*Oryza sativa* L.) hybrids using microsatellite markers. *3 Biotech* **2016**, *6*, 50. [[CrossRef](#)]
41. Subirana, J.A.; Messeguer, X. Structural families of genomic microsatellites. *Gene* **2008**, *408*, 124–132. [[CrossRef](#)]

42. Lai, Y.; Sun, F. The relationship between microsatellite slippage mutation rate and the number of repeat units. *Mol. Biol. Evol.* **2003**, *20*, 2123–2131. [[CrossRef](#)]
43. Levinson, G.; Gutman, G. Slipped-strand mispairing: A major mechanism for DNA sequence evolution. *Mol. Biol. Evol.* **1987**, *4*, 203–221. [[PubMed](#)]
44. 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](#)] [[PubMed](#)]
45. Ramsey, J.; Schemske, D.W. Pathway, mechanisms, and rates of polyploid formation in flowering plants. *Annu. Rev. Ecol. Syst.* **1998**, *29*, 467–501. [[CrossRef](#)]
46. Kati, V.; Le Corre, V.; Michel, S.; Jaffrelo, L.; Poncet, C.; Délye, C. Isolation and characterisation of 11 polymorphic microsatellite markers in *Papaver rhoeas* L. (corn poppy), a major annual plant species from cultivated areas. *Int. J. Mol. Sci.* **2012**, *14*, 470–479. [[CrossRef](#)] [[PubMed](#)]
47. Pompanon, F.; Bonin, A.; Bellemain, E.; Taberlet, P. Genotyping errors: Causes, consequences and solutions. *Nat. Rev. Genet.* **2005**, *6*, 847–859. [[CrossRef](#)]
48. Baldoni, L.; Cultrera, N.G.; Mariotti, R.; Ricciolini, C.; Arcioni, S.; Vendramin, G.G.; Buonamici, A.; Porceddu, A.; Sarri, V.; Ojeda, M.A.; et al. A consensus list of microsatellite markers for olive genotyping. *Mol. Breed.* **2009**, *24*, 213–231. [[CrossRef](#)]
49. Drašnarová, A.; Krak, K.; Vít, P.; Doudová, J.; Douda, J.; Hadincová, V.; Zákavský, P.; Mandák, B. Cross-amplification and multiplexing of SSR markers for *Alnus glutinosa* and *A. incana*. *Tree Genet. Genomes* **2014**, *10*, 865–873. [[CrossRef](#)]
50. Kameníková, L.; Vrbovský, V. White-seeded poppy varieties Orel and Racek. *Czech. J. Genet. Plant. Breed.* **2009**, *45*, 37–38. [[CrossRef](#)]
51. Butler, J.M. *Advanced Topics in Forensic DNA Typing: Interpretation*; Academic Press: Cambridge, MA, USA, 2015.
52. Waits, L.P.; Luikart, G.; Taberlet, P. Estimating the probability of identity among genotypes in natural populations: Cautions and guidelines. *Mol. Ecol.* **2001**, *10*, 249–256. [[CrossRef](#)]
53. Meier, U. *BBCH-Monograph: Growth Stages of Mono- and Dicotyledonous Plants*; Blackwell Wissenschafts-Verlag: Berlin, Germany, 1997.
54. Masoudi-Nejad, A.; Tonomura, K.; Kawashima, S.; Moriya, Y.; Suzuki, M.; Itoh, M.; Kanehisa, M.; Endo, T.; Goto, S. EGASSEMBLER: Online bioinformatics service for large-scale processing, clustering and assembling ESTs and genomic DNA fragments. *Nucleic Acids Res.* **2006**, *34*, W459–W462. [[CrossRef](#)]
55. Wang, X.; Wang, L. GMATA: An integrated software package for genome-scale SSR mining, marker development and viewing. *Front. Plant. Sci.* **2016**, *7*, 1350. [[CrossRef](#)] [[PubMed](#)]
56. Untergasser, A.; Cutcutache, I.; Koressaar, T.; Ye, J.; Faircloth, B.C.; Remm, M.; Rozen, S.G. Primer3—new capabilities and interfaces. *Nucleic Acids Res.* **2012**, *40*, e115. [[CrossRef](#)] [[PubMed](#)]
57. Koressaar, T.; Remm, M. Enhancements and modifications of primer design program Primer3. *Bioinformatics* **2007**, *23*, 1289–1291. [[CrossRef](#)]
58. Hall, T.A. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In *Proceedings of the Nucleic Acids Symposium Series*; Information Retrieval Ltd.: London, UK, 1999; Volume 41, pp. 95–98.
59. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [[CrossRef](#)] [[PubMed](#)]
60. Kalinowski, S.T.; Taper, M.L.; Marshall, T.C. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol. Ecol.* **2007**, *16*, 1099–1106. [[CrossRef](#)] [[PubMed](#)]
61. TIBCO Statistica. Available online: <https://www.tibco.com/resources/product-download/tibco-statistica-trial-download-windows> (accessed on 28 December 2019).
62. Perrier, X.; Jacquemoud-Collet, J.P. DARwin Software. Available online: <http://darwin.cirad.fr/darwin> (accessed on 26 April 2019).
63. Dice, L.R. Measures of the amount of ecologic association between species. *Ecology* **1945**, *26*, 297–302. [[CrossRef](#)]
64. Hintze, J. NCSS; NCSS: Kaysville, UT, USA, 2001.

