Acta fytotechn zootechn, 28, 2025(4): 381–393
http://www.acta.fapz.uniag.sk

Original Paper

PCR-Based Profiling of Genetic Variability in Domesticated *Capsicum* spp.: Implications for Biodiversity

Silvia Farkasová[1], Lucia Urbanová*[1], Jana Žiarovská[2]

[1]*Slovak University of Agriculture in Nitra, Research Centre Agrobiotech, Slovakia*
[2]*Slovak University of Agriculture in Nitra, Faculty of Agrobiology and Food Resources, Institute of Plant and Environmental Sciences, Slovakia*

Article Details: Received: 2025-09-30 | Accepted: 2025-11-11 | Available online: 2025-12-31

https://doi.org/10.15414/afz.2025.28.04.381-393

The original tropical genus *Capsicum* spp. is an important crop worldwide, particularly the species *Capsicum chinense* Jacq and *Capsicum annuum* L. Like many crops subjected to long-term domestication and breeding, and under the continuous influence of human activity on the environment, the *Capsicum* genus faces a significant loss of biodiversity and a gradual reduction in the availability of its genetic resources. This study aimed to evaluate the applicability of existing PCR marker techniques for both interspecific and intraspecific analysis, and to examine the extent of genetic similarity or diversity among the genomes of the cultivated *Capsicum* spp. Five varieties of *C. chinense* Jacq and *C. annuum* L. were used in PCRs to detect genetic variability by four DNA-based dominant marker techniques: CDDP (Conserved DNA-derived polymorphism), iPBS (Inter-primer binding site amplification), PBA (Cytochrome P450-based polymorphism) and BBAP (Bet v 1-based amplicon polymorphism). The number of *loci* for all methods was 310 and it varied from 34 (PBA), 62 (BBAP) and 96 (iPBS) to 117 (CDDP), P% for all methods was 79.68% (45.71% PBA, 82.26% BBAP, 84.38% iPBS and 86.32% CDDP). The CDDP primers had the maximum discriminating power at 0.718, whilst the PBA primers displayed the lowest value at 0.379 among the applied approaches. The pair Ohnivec – Jalahot had the highest genetic similarity across all approaches, evaluated at 0.335. The lowest genetic similarity was detected between the pairings H. chocolate and Ohnivec, reported at 0.629. All used DNA-based marker techniques were evaluated as applicable and informative in the genus *Capsicum* spp., from moderately informative PBA, through highly informative iPBS and BBAP to very highly informative CDDP.

Keywords: *Capsicum chinense* Jacq, *Capsicum annuum* L., DNA marker technique, biodiversity

1 Introduction

1.1 Capsicum spp.

The genus *Capsicum* (family Solanaceae) originates from the tropical regions of Central and South America, from where it gradually spread to the tropics of Asia and Africa, and later to the temperate zones of Europe and North America. Of the thirty-eight known species, six have been domesticated, the most important of which are *Capsicum annuum* L., *Capsicum chinense* Jacq, and *Capsicum frutescens* L. These species are valued for their pungent taste, which is caused by alkaloid complexes (capsaicinoids), also used as extracts in alternative medicine (Ramchiary and Kole 2019). The most intensive breeding has occurred in *C. annuum* L., which includes various pungent varieties used in Mexican cuisine – where it was originally domesticated – as well as sweet "vegetable" varieties popular in temperate regions. The fruits of *C. chinense* Jacq are primarily used as a spice due to their distinctive flavour and high pungency.

These two species exhibit high morphological variability; however, like other domesticated crops their molecular variability has been naturally reduced by local selective breeding (Pickersgill 1997). On the other hand, genetic diversity can be increased through outcrossing under field conditions, depending on the environment, location, and population diversity.

***Corresponding Author:** Lucia Urbanová. Research Centre Agrobiotech, Slovak University of Agriculture, Trieda Andreja Hlinku 2, 949 76 Nitra, Slovakia
e-mail: lucia.urbanova@uniag.sk ORCID: https://orcid.org/0000-0001-7434-2677

1.2 Genetic Improvement and Environmental Adaptation in Capsicum Cultivation

In addition to targeted breeding, natural hybridization between species or variants within a species has also occurred, resulting in varieties such as Bhut Jolokia (a probable hybrid of *C. chinense* and *C. frutescens*) (Bosland and Baral 2007), or one of the most pungent cultivar, Trinidad Moruga Scorpion, a natural hybrid originating from the Trinidad and Tobago region in the Caribbean (Bosland et al., 2012). Breeding has focused on genes responsible for pungency, resistance to biotic and abiotic stress, and the ability to adapt to different climate zones. These efforts are related to the main domestication syndromes in the genus *Capsicum* including non-deciduous and pendent fruits (providing protection from sun and birds), as well as varying degrees of pungency based on consumer preferences (Kumar et al., 2018). *Capsicum* spp. is affected by a wide range of biotic stresses and is sensitive to growing conditions. The optimal temperatures for growing are 25–30 °C (day) and 18–20 °C (night) with more than 80% relative humidity. In temperate climates, indoor cultivation (polyhouse, greenhouse, or net shades) is more suitable with essential supplementing of nitrogen, phosphorus and potassium and water ranging between 600 to 1,250 mm. High light intensity in the open field is also not favourable for fruit development (Pramanik et al., 2020; FAO). In 2023, total production of *Capsicum* spp. (with *Pimenta* spp. as a group "chillies and peppers") reached 44.1 mil. t, with China being the largest producer, accounting for more than 17.4 mil. t of pepper (FAOSTAT, 2023). Droughts and flooded regions, generated by climate change, pose challenges for modern breeding, which increasingly relies on information from the fields of molecular biology and genetics.

1.3 DNA Marker Methodologies

DNA markers can identify genotypic variations, exposing polymorphisms within the same or among distinct species. The genetic variability can benefit from DNA mutations as a point mutation, DNA segment duplications, translocations, and mistakes during DNA replication (Amiteye, 2021) that enrich the genome and support genetic biodiversity. Arbitrarily amplified DNA markers may present drawbacks, including the co-migration of paralogous bands instead of orthologous ones, nested priming, and heteroduplex formation, among others. The rapid advancement of public genomic databases has facilitated the development of gene-targeted markers, which can be valuable for investigating genetic variability and diversity (Poczai et al., 2013).

1.3.1 CDDP (Conserved DNA-Derived Polymorphism)

The technique utilises gene sequences from conserved gene families that are present in multiple copies within the genomes of plants. Primers were designed to target well-characterised plant genes, including *MYB* (involved in secondary metabolism, abiotic and biotic stress, and cell morphology), *ERF* (transcription factors that play a role in plant disease resistance pathways), *KNOX* (homeobox genes), *MADS* (genes responsible for the initiation and development of floral organs), and *ABP1* (gene encoding auxin binding protein). Although the reproducibility of this technique has been demonstrated to be high, it is not entirely consistent (Poczai et al., 2013).

1.3.2 iPBS (Inter-Primer Binding Site Amplification)

Retrotransposons provide great enrichment to genome through transport ability. They constitute a very diverse population of genetic elements, demonstrating insertional polymorphism both intra- and inter-plant taxa. Their prevalence across plant species is correlated with genomic complexity, since species with smaller genomes often have a reduced fraction of transposable elements (Poczai et al., 2013). Kalendar et al. (2010) proposed the iPBS approach that functions as an independent marker system and as a fast separation technique for retrotransposons. iPBS amplification depends on the ubiquitous existence of the tRNA complement, serving as a binding site for reverse transcriptase (PBS) in LTR retrotransposons.

1.3.3 PBA (Cytochrome P450-Based Polymorphism)

The *CYP* (Cytochrome P450) gene family is present in most of eukaryotes which contain a diverse array of enzymes, comprising approximately 35,000 members. *CYP* genes are essential for detoxifying xenobiotics, protecting plants from biotic and abiotic stresses. They are organised into clans, with each clan originating from a single ancestral gene, land plants possess eleven clans (Nelson and Werck-Reichhart, 2011). Authors Yamanaka et al. (2003) developed DNA markers for assessing genome-wide diversity in higher plant-species based on mammalian *P450* genes.

1.3.4 BBAP (Bet v 1-Based Amplicon Polymorphism)

Among other things, the main function of pathogen-related (PR) proteins is the response to abiotic or biotic stress. They are divided into 17 protein families including the PR-10 protein family, encoded by the *ypr10* genes (Fernandes et al., 2013). *ypr10* genes have significant

sequence similarity and constitute a uniform group, with this uniformity thought to be preserved by coordinated evolution (Schenk et al., 2009); rendering them attractive candidates for the development of marker techniques. Bet v 1 is a notable member of the PR-10 protein family due to its allergic ability. Family membership is based on homology in 3D structure and highly conserved amino acid sequence regions. Such conservation has enabled the advancement of the BBAP technique by Žiarovská and Urbanová (2022). The approach was developed to characterise genomic variability of length polymorphism based on the sequences of homologous genes within the PR-10.

Molecular markers have been utilised in plant research for several decades, with applications including marker-assisted selection of genetic resources for breeding new varieties regarding conservation strategies (Poczai et al., 2013). Domestication and plant breeding form domestication syndromes or the ability to survive only in narrowly specified conditions. Preserving biodiversity is vital for future generations, as original plant genetic resources frequently offer solutions for plant survival in challenging growth conditions and under stress. These topics are intensively addressed by international organizations such as the Crop Trust, which published a report on "A Global Strategy for the Conservation and Use of *Capsicum* Genetic Resources" (Barchenger and Khoury, 2022), as well as by the United Nations and the European Commission, which implement biodiversity conservation strategies into international or local law (Strategic Plan for Biodiversity 2011–2020).

The objectives of this study were to assess the feasibility of selected PCR marker techniques at both interspecific and intraspecific levels and to observe the degree of similarity or capture the genetic diversity of the genomes of the most widely cultivated species and their popular varieties. The selected marker techniques were designed to target specific genes, all of which are involved in plant stress responses. Additionally, the study includes a bioinformatic analysis that identifies general differences between the genomes of the two *Capsicum* species examined, alongside a summary of currently available data from the NCBI database.

2 Material and Methods

2.1 Biological Material and DNA Extraction

The biological material consisted of species *Capsicum chinense* Jacq and *Capsicum annuum* L. and its 5 chilli pepper varieties (Table 1) grown in the experimental plots of the Botanical Garden of the Slovak University of Agriculture in Nitra and collected in autumn 2024. Genomic DNA isolation was done by commercial

Table 1 Description of biological material

Species	Cultivar	Origin	SHU	Parentage
Capsicum annuum L.	Jalahot	Mexico	15,000	Jalapeňo hybrid
	Ohnivec	Czech Republic	2,500	related to Palivec
Capsicum chinense Jacq	Habanero chocolate	Jamaica (Caribean region)	300–450 tis	naturally selected variant of Habanero group
	TMS caramel	Moruga, Trinidad and Tobago region in Caribean	1.2–1.8 mil	natural variant of Trinidad Moruga Scorpion
	TMS yellow	Moruga, Trinidad and Tobago region in Caribean	0.8–1.5 mil	colour variant of Trinidad Moruga Scorpion

SHU – scoville heat units

Table 2 PCR reaction components and conditions

Marker technique	DNA (ng)	Primer (nM)	Annealing temperature (°C)	Polymerisation time (s)	PCR cycles (times)
Conserved DNA-derived polymorphism (CDDP)	10	400	54	60	35
Inter-primer binding site amplification (iPBS)	10	1200	55	120	35
Cytochrome P450-based polymorphism (PBA)	10	500	50	120	45
Bet v 1-based amplicon polymorphism (BBAP)	10	500	63	45	40

kit GeneJET Genomic DNA Purification Kit (Thermo Scientific), the manufacturer's steps were followed. The concentration and quality of the isolated DNA were measured using a NanoPhotometerTM P360 (Implen).

2.2 PCR Amplification and Statistical Evaluation of Markers

All PCR reactions were performed on a SureCycler 8800 thermal cycler (Agilent). The temperature profiles and reaction components are displayed in Table 2 and primer sequences in Table 3. All reactions were performed in 10 µl of volume using EliZyme Robust PCR Master Mix (2X) (Elisabeth Pharmacon®).

PCR reaction products were separated on 3% agarose gels stained with the intercalating dye GelRed® Nucleic Acid Gel Stain (Biotinum). The fragments were separated at 110 V for 120 minutes. Subsequently, they were visualized using a BDA digital system 30 transilluminator (Analytik Jena). The electropherogram images obtained were processed using free software GelAnalyzer. The amplicons were evaluated using a standard HyperLadder™ 50bp (Meridian BioScience®). Based on the data obtained, binary matrices were made, based on the presence or absence of the given fragments. Distance matrices were subsequently created based on the Jaccard coefficient using R studio with vegan package (2.7.1) (Oksanen et al., 2025). Varieties were clustered using "ward.D" method into individual dendrograms. For statistical calculations online software iMEC (Amiryousefi et al., 2018) (Online Marker Efficiency Calculator) was used. The assessment of the level of informativeness based on the PIC-value was assessed according to the scale of Serrote et al. (2020).

2.3 In silico Analyses

Bioinformatic analyses were conducted using the NCBI database (https://www.ncbi.nlm.nih.gov/), where information regarding the availability of gene sequences and whole genome records was examined. Subsequently, the genomes of *Capsicum annuum* L. (NCBI: GCF_002878395.1) and *Capsicum chinense* Jacq (NCBI: GCA_002271895.2) were compared using the tool Comparative Genome Viewer (https://www.ncbi.nlm.nih.gov/cgv) with a minimum

Table 3 Overview of used DNA-marker techniques with primer sequences and source references

Marker technique	Name	Primer sequence (5´-3´)	Reference
Conserved DNA-derived polymorphism (CDDP)	WRKY F1	TGGCGSAAGTACGGCCAG	Collard and Mackill (2009)
	WRKY R1	GTGGTTGTGCTTGCC	
	WRKY R2	GCCCTCGTASGTSGT	
	WRKY R3	GCASGTGTGCTCGCC	
	WRKY R2b	TGSTGSATGCTCCCG	
	WRKY R3b	CCGCTCGTGTGSACG	
Bet v 1-based am-plicon polymorphism (BBAP)	F1	CCTGGAACCATCAAGAAG	Žiarovská and Ur-banová (2022)
	R1	TTGGTGTGGTAGTTGCTG	
	R2	TTGGTGTGGTAGTGGCTG	
	R3	TTGGTGTGGTACTGGCTG	
	R4	TTGGTGTGGTACTTGCTG	
Cytochrome P450-based polymorphism (PBA)	CYP A1 F	GCCAACCTTTCTAACAATGC	Yamanaka et al. (2003)
	CYP A1 R	AAGGACATGCTCTGACCATT	
	CYP B6 F	GACTCTTGCTACTCCTGGTT	
	CYP B6 R	CGAATACAGAGCTGATGAGT	
	CYP C19 F	TCCTTGTGCTCTGTCTCTCA	
	CYP C19 R	CCATCGATTCTTGGTGTTCT	
Inter-primer binding site amplification (iPBS)	1846	CTGGCATTTCCATTGTCGTCGATGC	Kalendar et al. (2010)
	1886	ATTCTCGTCCGCTGCGCCCCTACA	
	1897	AGTTTGGCATAGAAAAATCGAGCCAAC	
	1899	TGAGTTGCAGGTCCAGGGATCA	
	2079	AGGTGGGCGCCA	

S – C/G

alignment size set at 1,000 bp. Primer BLAST tool was employed to assess the potential for primer binding (CDDP, BBAP, PBA) to the available sequences in the NCBI database for both *C. annuum* L. and *C. chinense* Jacq. The analyses were conducted using the default parameters offered by the tool (https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome).

3 Results and Discussion

The aim of the study, to assess the application of marker techniques on two species of the genus *Capsicum*, was achieved using four DNA-based marker techniques (PBA, BBAP, CDDP, and iPBS) to determine both intraspecific and interspecific variability of their genomes. Two *Capsicum annuum* L. varieties (Jalahot, Ohnivec) and three *Capsicum chinense* Jacq varieties (Habanero chocolate, TMS caramel, and TMS yellow) were selected for the study as the species most frequently encountered in the breeding of modern chili pepper varieties. *In vitro* analyses were preceded by extraction of genomic DNA, the concentration of which ranged between 16–29 ng.µl^{-1}, with a purity of 1.8 ±0.2 at absorbance A260/280. All PCR series contained no-template controls, which were negative for the presence of amplicons.

The CDDP marker technique was evaluated as the most suitable for studying the genome variability of varieties and selected species of the genus *Capsicum* from the point of view of several statistical indexes. It achieved the highest PIC value approaching the maximum for codominant genes can be 0.5 and stood out with the total number of *loci* and amplicons, with relatively high discriminatory power (Table 4). The iPBS and BBAP techniques were evaluated as average, and PBA as poorly suitable.

3.1 CDDP

The PIC value of 0.498 indicates a very high level of informativeness. The percentage of polymorphic *loci* (P%) was 86.32%, and the discriminatory power (D) was 0.718. Five primer combinations (WRKY) were employed for the CDDP, which proved sufficient to distinguish the samples and only 15 of the 117 *loci* were monomorphic. All individual primer pairs produced polymorphic profiles. The primer pair WRKY F + R1 yielded a total of 66 amplicons, with an average of 13.2 amplicons per sample (ranging from 11 to 16). The amplicon sizes ranged from 184 to 1,550 bp. All cultivars shared sizes of 251, 302, and 466 bp, while unique amplicons occurred in Jalahot (420 bp) and TMS caramel (912 bp). The primer pair WRKY F + R2 generated a total of 76 amplicons, the highest yield among the five primer combinations, resulting in an average of 15.2 amplicons per sample (ranging from 14 to 16). The amplicon sizes varied from 161 to 1649 bp. All samples exhibited shared sizes of 344, 597, 933, and 1,185 bp. Unique amplicons were identified in H. chocolate (1,649 bp), TMS caramel (501, 529, 901, 1,476 bp), TMS yellow (802 bp), and Ohnivec (239 bp). Primers WRKY F + R3 generated a total of 55 amplicons, averaging 11 per sample (ranging from 7 to 12). The sizes varied from 185 to 1,802 bp. All cultivars exhibited shared amplicons of 395 and 1,260 bp. Unique products were identified in H. chocolate (653, 1,471, 1,702 bp) and Ohnivec (920 bp). With the primer combination WRKY F + R2b, a total of 38 amplicons were generated, averaging 7.6 per sample (ranging from 6 to 9). All cultivars exhibited shared amplicons of 228 and 1,143 bp. Unique amplicons were observed in Ohnivec (193 and 706 bp), H. chocolate (604 and 880 bp), TMS caramel (645 bp), and TMS yellow (1,711 bp). The final primer combination, WRKY F + R3b, generated a total of 73 amplicons, averaging 14.6 per

Table 4 Basic statistical information proceeded for all varieties of *Capsicum* spp. by used DNA marker techniques

Marker technique	PIC	Amplicon range (bp)	Total number of amplicons	*loci*	P %	D
Conserved DNA-derived polymorphism (CDDP)	0.498	122–1,802	308	117	86.32	0.718
Inter-primer binding site amplification (iPBS)	0.388	127–2,190	256	96	84.38	0.435
Cytochrome P450-based polymorphism (PBA)	0.278	81–1,753	133	34	45.71	0.379
Inter-primer binding site amplification (iPBS)	0.365	68–1,806	186	62	82.26	0.641
all methods	0.369	–	895	310	79.68	0.671

PIC – polymorphism information content; P – polymorphism; D – discriminatory power

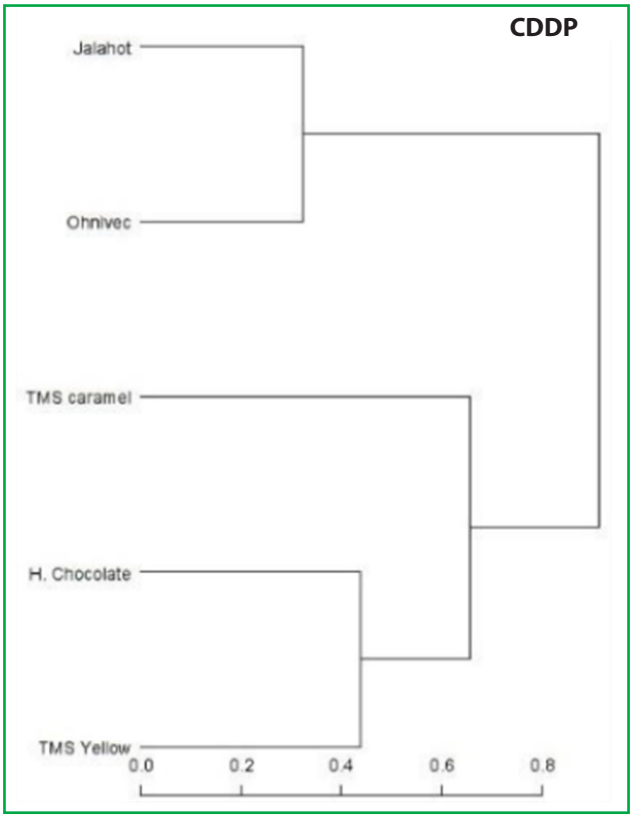

Figure 1 Dendrogram describing the genetic distance of five cultivars of the genus *Capsicum* based on the CDDP technique according to Jaccard coefficient
The legend shows the values of Jaccard distance indices

sample (ranging from 12 to 17). The amplicon sizes ranged from 135 to 1,448 bp. All cultivars shared the sizes of 125, 187, 240, 293, 419, and 543 bp. Unique amplicons were identified in the profiles of Jalahot (559 and 1,260 bp), Ohnivec (268 and 809 bp), and TMS caramel (260 bp).

The CDDP technique showed excellent interspecific discrimination, distinguishing species with Jaccard indices ranging from 0.583 (TMS caramel – Jalahot) to 0.716 (H. chocolate – Jalahot). The most similar profiles were found for Jahalot – Ohnivec (0.324), while the highest index was observed for the mentioned pair H. chocolate – Jalahot. According to Figure 1, the samples were divided into two groups based on species classification, as expected. On the other hand, the varieties TMS caramel and TMS yellow were unexpectedly separated, with a Jaccard dissimilarity index of 0.563, which was higher than that of the pair TMS yellow – H. chocolate (0.439).

3.2 iPBS

Initial screening was conducted using 17 iPBS primers, of which 5 primers exhibiting the best amplification profiles were selected for further analysis. The PIC value of 0.388 indicates a high level of informativeness, while the P% was recorded at 84.38%, and the D value was 0.435. Five primers were employed, which proved sufficient to distinguish the samples, with only 15 of the 96 *loci* being monomorphic. The primer pair 1,886 yielded a total of 63 amplicons, with an average of 12.6 amplicons per sample (ranging from 9 to 15). The amplicon sizes ranged from 284 to 2,190 bp. All cultivars shared amplicon sizes of 334, 585, and 1,607 bp, while unique amplicons were identified in Jalahot (519 bp) and TMS caramel (1,098 bp). Primer 1,897 generated a total of 42 amplicons, averaging 8.4 amplicons per sample (ranging from 7 to 11). The amplicon sizes ranged from 163 to 1,490 bp, with all samples exhibiting shared sizes of 163, 278, 504, and 1,490 bp. Unique amplicons were identified only in Jalahot (431 bp). Primer 1,899 produced 34 amplicons, averaging 6.8 amplicons per sample (ranging from 6 to 8), with sizes varying from 261 to 939 bp. All cultivars exhibited shared amplicons at 261, 371, 788, and 939 bp, while unique products were found in TMS caramel (596 and 693 bp). Using primer 2,079, a total of 50 amplicons were generated, averaging 10 per sample (ranging from 9 to 12). All cultivars exhibited shared amplicons at

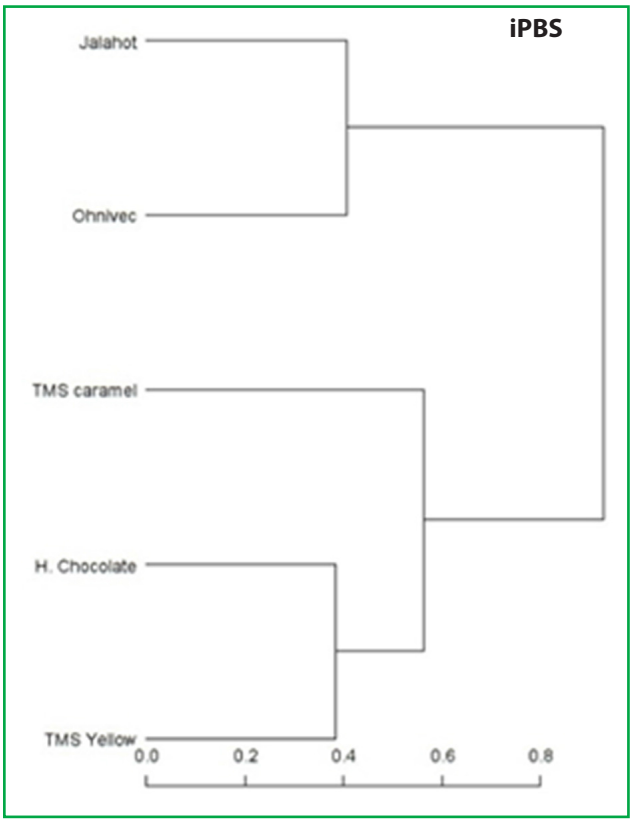

Figure 2 Dendrogram describing the genetic distance of five cultivars of the genus *Capsicum* based on the iPBS technique according to Jaccard coefficient
The legend shows the values of Jaccard distance indices

sizes of 334, 582, and 1,343 bp, with unique amplicons observed in Jalahot (455 bp), Ohnivec (477 bp), and TMS caramel (893 bp). Lastly, primer 1,846 generated 67 amplicons, averaging 13.4 per sample (ranging from 11 to 15), with sizes ranging from 127 to 1,517 bp. Unique amplicons were identified in the profiles of Jalahot (248 bp), Ohnivec (523, 739, and 1,231 bp), H. chocolate (698 bp), and TMS caramel (209 and 318 bp). No shared amplicons were found.

The dendrogram of genetic distances based on the results of the iPBS technique (Figure 2) was very similar to that of the CDDP technique. The same clusters were formed on the dendrogram, but they differ in the values of the Jaccard indices. The lowest value was observed for the pair TMS yellow – H. chocolate (0.383), and the highest for Ohnivec – H. chocolate (0.680), both threshold values being lower than in the case of CDDP. The profiles of the TMS yellow and caramel varieties were again distinguished by more than 50% dissimilarity. Six out of ten pairs reported values higher than 0.600.

3.3 PBA

Among the applied methodologies, PBA proved to be the least effective at the varietal level. Primer combinations CYP A1 F + R and CYP B6 F + R were able to distinguish between the two studied species. PIC had a value of 0.278, the lowest among the techniques used; the primers were moderately informative, P% was 45.71%, and D was 0.379. The primer combination CYP A1 F + R generated a total of 37 amplicons, with an average of 7.4 per sample (ranging from 5 to 10), in a size range from 112 to 1,753 bp. All cultivars shared amplicons at sizes 275, 350, 409, and 526 bp. The primers CYP B6 F + R produced 51 amplicons, with an average of 10.2 per sample (ranging from 9 to 11), in a size range from 90 to 1,095 bp. All samples shared amplicon sizes of 90, 122, 433, 571, 868, and 1,095 bp. The primers CYP C19 F + R generated a monomorphic profile in the size range from 81 to 1,131 bp. No cultivar had a unique amplicon in their profile using the PBA technique.

The Jaccard index values were highest for TMS yellow – Jalahot/Ohnivec (0.429) and lowest for Ohnivec – Jalahot and TMS caramel – H. chocolate (both 0.000) (Figure 3). Unlike in the CDDP and iPBS techniques, the H. chocolate – TMS caramel pair shared an identical profile, while the TMS yellow profile lacked the 146 bp band amplified by the CYP A1 primer pair.

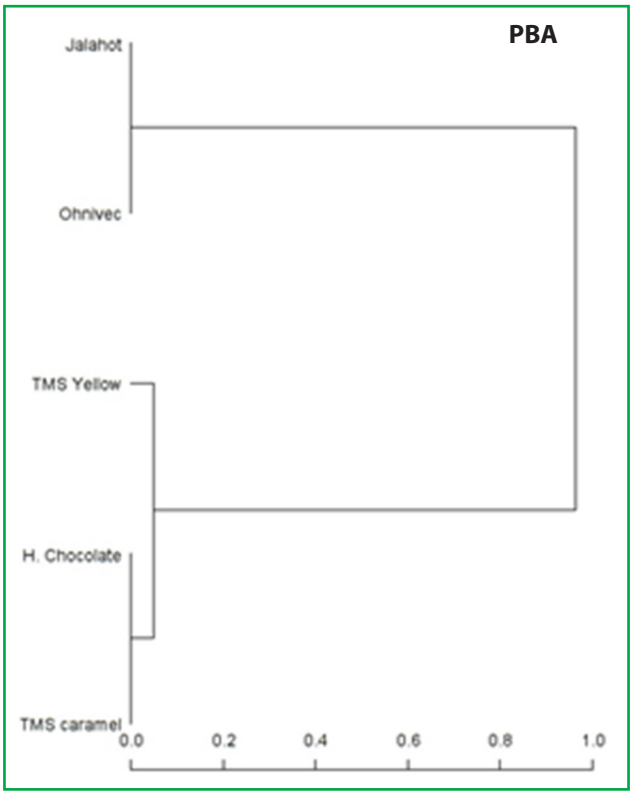

Figure 3 Dendrogram describing the genetic distance of five cultivars of the genus *Capsicum* based on the PBA technique according to Jaccard coefficient
The legend shows the values of Jaccard distance indices

3.4 BBAP

BBAP technique performed as a highly informative marker according to PIC value (0.365), P% was 82.26 and D was 0.641. The primer combination F + R1 generated a total of 41 amplicons, averaging 8.2 amplicons per sample, with a range of 7 to 10. The sizes of the products varied from 92 to 1,086 bp. All samples shared an amplicon measuring 1,086 bp. Unique amplicons were identified in the H. chocolate profile (1,096 bp, 274 bp, 394 bp), as well as in TMS yellow (91 bp). Primers F + R2 generated amplicons ranging in size from 76 to 1,695 bp; in total, 40 amplicons were produced, with an average of 8 per sample (ranging from 6 to 12). This primer combination did not yield any amplicons for TMS yellow; however, the other cultivars exhibited shared amplicons of sizes 136 and 495 bp. Jalahot displayed unique amplicons at sizes 76 and 357 bp. Primers F + R3 generated a total of 58 amplicons, ranging from 75 to 1,806 bp, resulting in an average of 11.6 amplicons per sample (ranging from 10 to 15). All cultivars exhibited amplicons of sizes 296, 544, 624, and 1,806 bp. Unique amplicons were identified in the profiles of Jalahot (118 bp), TMS caramel (125 bp), and TMS yellow (177 and 257 bp).

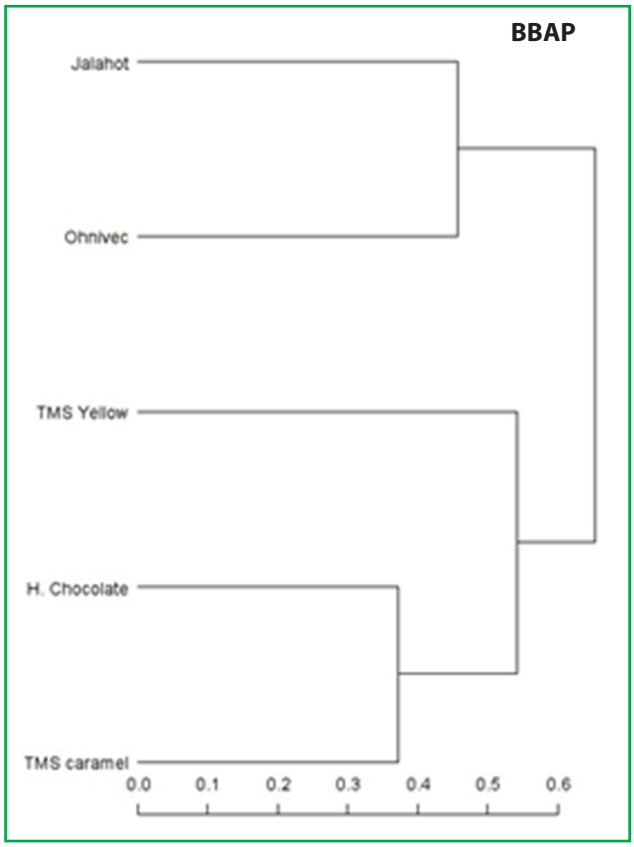

Figure 4 Dendrogram describing the genetic distance of five cultivars of the genus *Capsicum* based on the BBAP technique according to Jaccard coefficient
The legend shows the values of Jaccard distance indices

The final primer combination, F + R4, generated a total of 47 amplicons, averaging 9.4 per sample (ranging from 8 to 10). The size range of these amplicons was 68–835 bp. The amplification profiles demonstrated greater similarity among the cultivars, with all sharing amplicon sizes of 85, 120, 175, 220, 668, and 835 bp. Unique amplicons were noted in the profiles of Jalahot (355 bp), Ohnivec (229 bp), and TMS caramel (68 bp).

All values of the Jaccard index of genetic distances were very similar with the BBAP technique (Figure 4) (median 0.500 ±0.074). The lowest value was observed for TMS caramel – H. chocolate (0.373), and the highest for TMS yellow – Jalahot (0.608), followed closely by the value for the pair H. chocolate – Ohnivec (0.604). The profile of TMS caramel was more like H. chocolate than to TMS yellow (0.510), as was also the case with the iPBS technique.

3.5 Combined Application of PBA, CDDP, BBAP and iPBS Marker Systems

Apart from the identical profiles observed in the PBA technique, the intraspecific genetic diversity of *C. annuum* L. ranged from 0.324 (CDDP) to 0.457 (BBAP). The genetic diversity of *C. chinense* Jacq ranged from 0.373 (BBAP) to 0.641 (CDDP). Across the combination all used techniques (Figure 5), the pair Ohnivec – Jalahot (0.335) showed the highest genetic similarity, while the lowest genetic similarity was observed in the pair H. chocolate – Ohnivec (0.629).

The total number of *loci* was 310, with a polymorphism percentage of 79.68%. The PIC value reached 0.370, indicating high informativeness, and the discriminatory power value was 0.671. Overall, 80% of the pairs had a genetic distance greater than 50%, and 60% had a genetic distance greater than 60%. In conclusion, the genetic diversity between the yellow and caramel colour types of the TMS cultivar was evaluated at 0.479, which was higher than the genetic distances between H. chocolate – TMS yellow (0.385) and H. chocolate – TMS caramel (0.472).

Biodiversity assessments and monitoring programs may require information at various levels such as morphological, biochemical and molecular. DNA marker systems can identify differences that are often

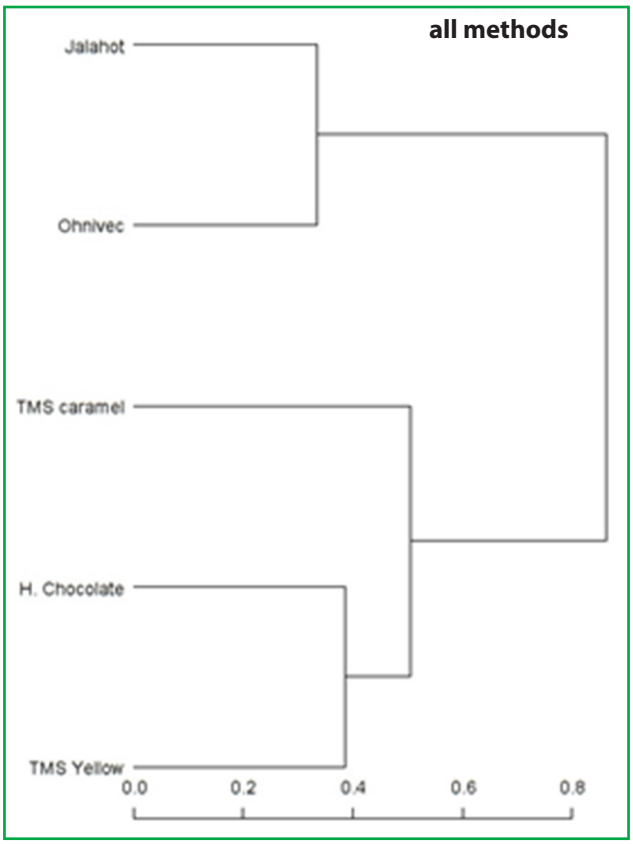

Figure 5 Dendrogram describing the genetic distance of five cultivars of the genus *Capsicum* based on resulting profiles of all used techniques according to Jaccard coefficient
The legend shows the values of Jaccard distance indices

undetectable, as in certain crops, genetic variations result in observable phenotypic distinctions, whereas in other instances, they remain obscure (Swarup et al., 2021). DNA marker systems contribute significantly to a comprehensive understanding of the diversity within available plant genetic resources. This information is crucial for modern plant breeding (Nadeem et al., 2017), which must consider not only the needs of consumers and producers, but also the preservation of food safety, biodiversity, and the protection of the environment and ecological balance. For genome mapping without prior sequence data, several DNA marker systems have been developed, such as AFLP, CAPS, CDDP, CpSSR, DArT, EST, iPBS, IRAP, ISSR, PBA, RAPD, RAMP, RBIP, REMAP, RFLP, SAMPL, SCAR, SNP, SRAP, SSCP, SSR, STS, and others (Nadeem et al., 2017). According to Aziz et al. (2023), increasing the number of markers and expanding genome coverage is expected to enhance the reliability of the data. These DNA marker methodologies offer both efficacy and cost-effectiveness. Yamanaka et al. (2003) tested the PBA technique on 51 plants across 28 taxonomic families. The authors recommend the technique as a useful tool to assess and detect diversity in plant species that do not have relevant genetic markers. The BBAP technique was employed for several vegetable species (Urbanová and Žiarovská 2021).

The decline in *Capsicum* spp. biodiversity is evident, as shown by ongoing efforts to preserve origin, traditional, local or valuable cultivars for both biological and cultural reasons (Barchenger and Khoury, 2022). However, the level of current risk remains under-discussed. According to Liu et al. (2023), domestication and selective breeding have reduced genetic diversity by favouring traits like fruit size, pungency, and disease resistance.

3.6 In Silico Analysis

Extensively utilised database NCBI, was employed to analyse the genomic information of the two species under investigation. Presently (July 2025), *C. annuum* L. comprises 25 assembly/genome records, 77,082 genes, and 395,129 protein entries. The species *C. chinense* Jacq is less studied, with only 3 assembly/genome records present in the database. In terms of gene records, 135 are available, along with 44,606 protein entries. Both species possess 12 chromosomes, their comparison is illustrated in Figure 6, where the alignments are displayed with a minimum alignment threshold of 1,000 bp.

A search for probable targets of the utilized primers was performed using the primer-BLAST program, depending on the data accessible in the database. The findings were provided for *C. annuum* L. due to inadequate genetic data for *C. chinense* Jacq to conduct a similar search. Primers for all techniques were aligned with the genome of *C. annuum* L. The results are summarised to reflect all primer matches in accordance with the specified technique. It is important to note that there was not a 100% match to all the targets. The analysis was limited to the CDDP, BBAP, and PBA methodologies, as iPBS amplifies non-specific regions.

In total, 49 primer target matches were identified in NCBI for *C. annuum* L., of which 10% were represented by *wrky* genes, while 21% matched uncharacterized sequences (Figure 7). Additionally, up to 21% of the predicted products were associated with the flowering time control protein FPA.

The BLAST primer analysis using CYP primer combinations identified only 14 (Figure 8) matches within the *C. annuum* L. nucleotide records. One match was associated with a retrotransposon (LC434371.1), while 7 matches corresponded to *beta-amyrin synthase* mRNA sequences, and 5 matches were linked to serine/threonine protein phosphatase 6 regulatory subunit mRNA records.

Regarding the BBAP technique, 159 potential target products were identified across the four primer combinations for *C. annuum* L. The highest percentage

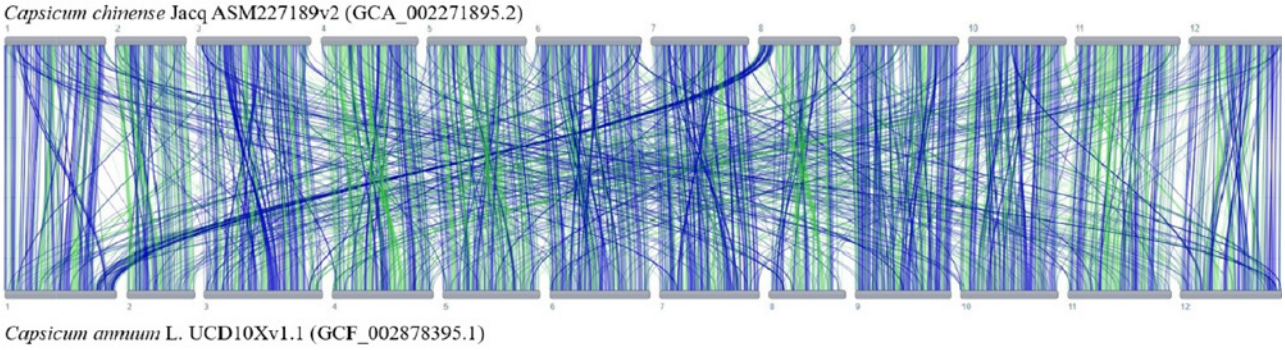

Figure 6 Whole genome alignment between *C. chinense* Jacq (NCBI: GCA_002271895.2) and *C. annuum* L. (NCBI: GCF_002878395.1) with minimum nucleotide alignment at 1000 bp (NCBI-Comparative Genome Viewer tool)

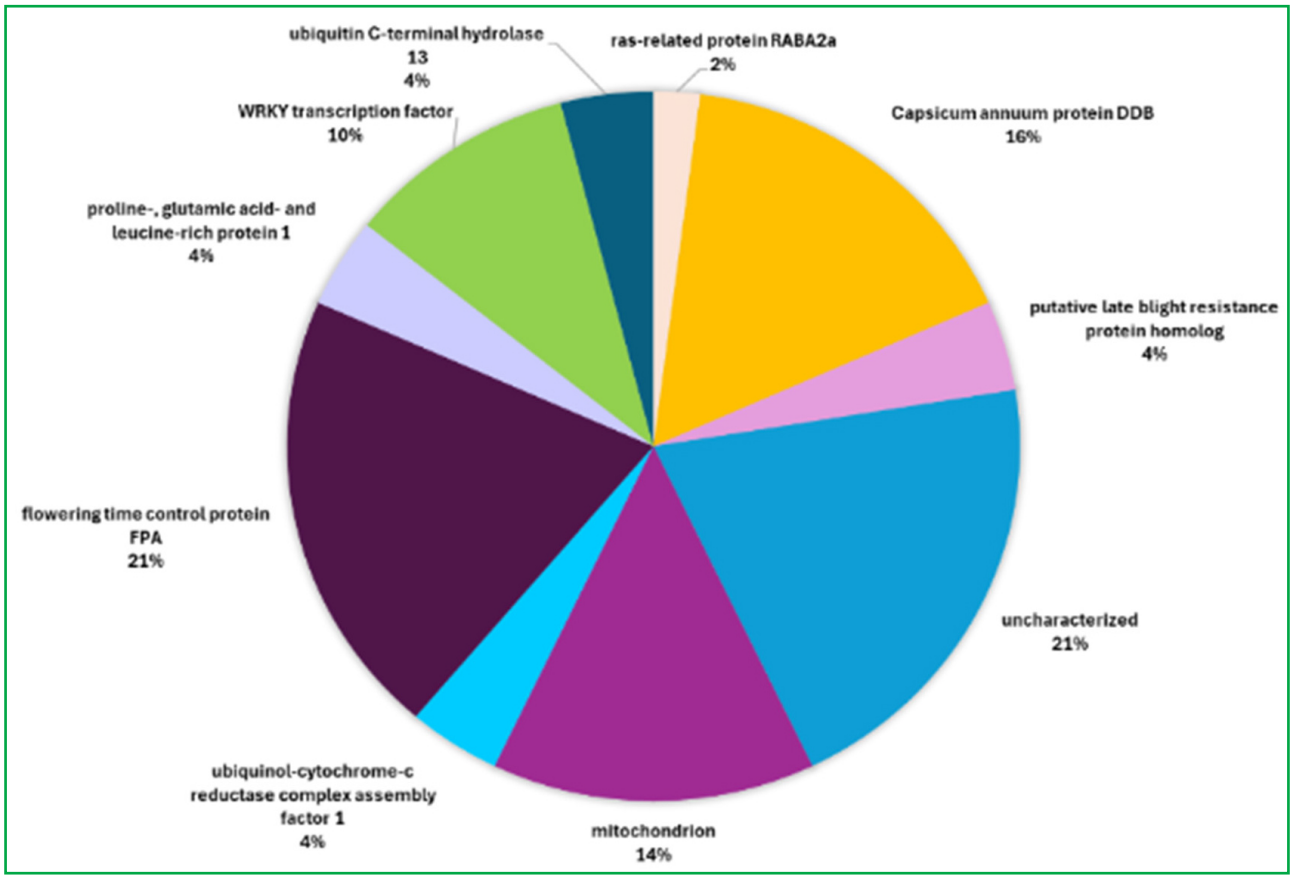

Figure 7 Predicted binding sites of WRKY primer set in the *Capsicum annuum* L. genome
NCBI Primer-BLAST Analysis

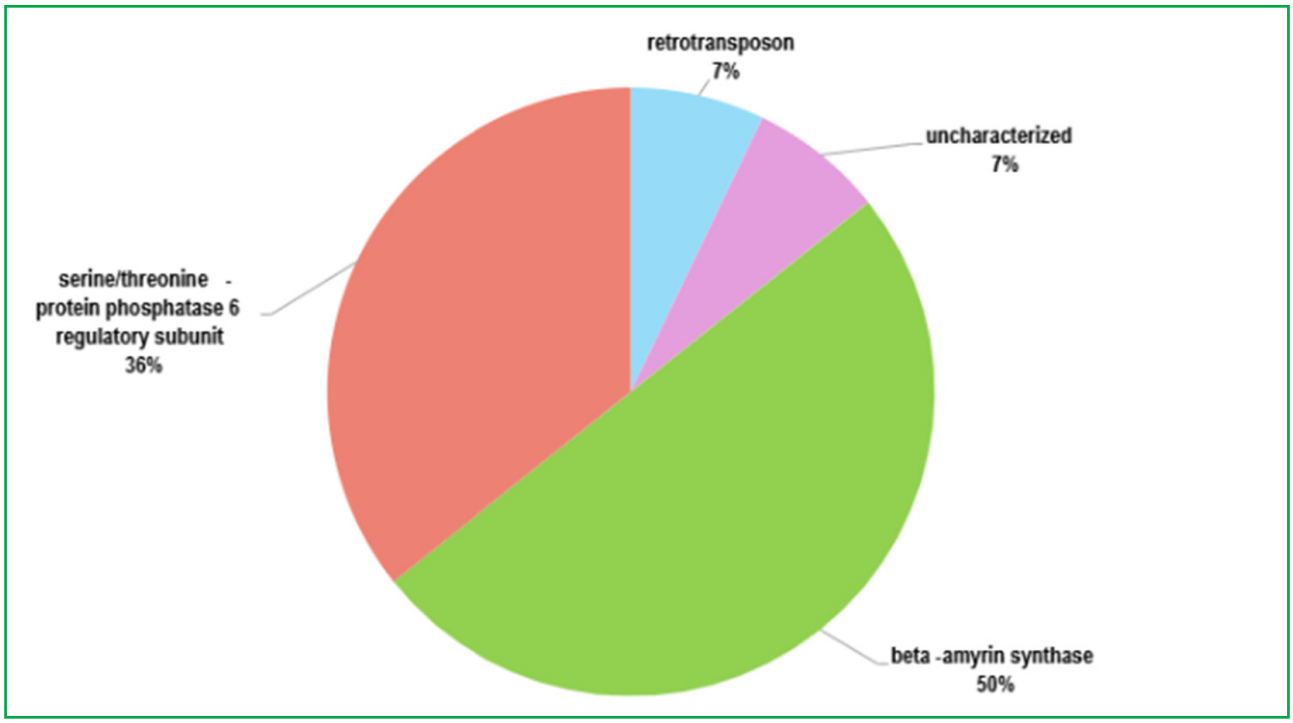

Figure 8 Predicted binding sites of CYP primer set in the *Capsicum annuum* L. genome
NCBI Primer-BLAST Analysis

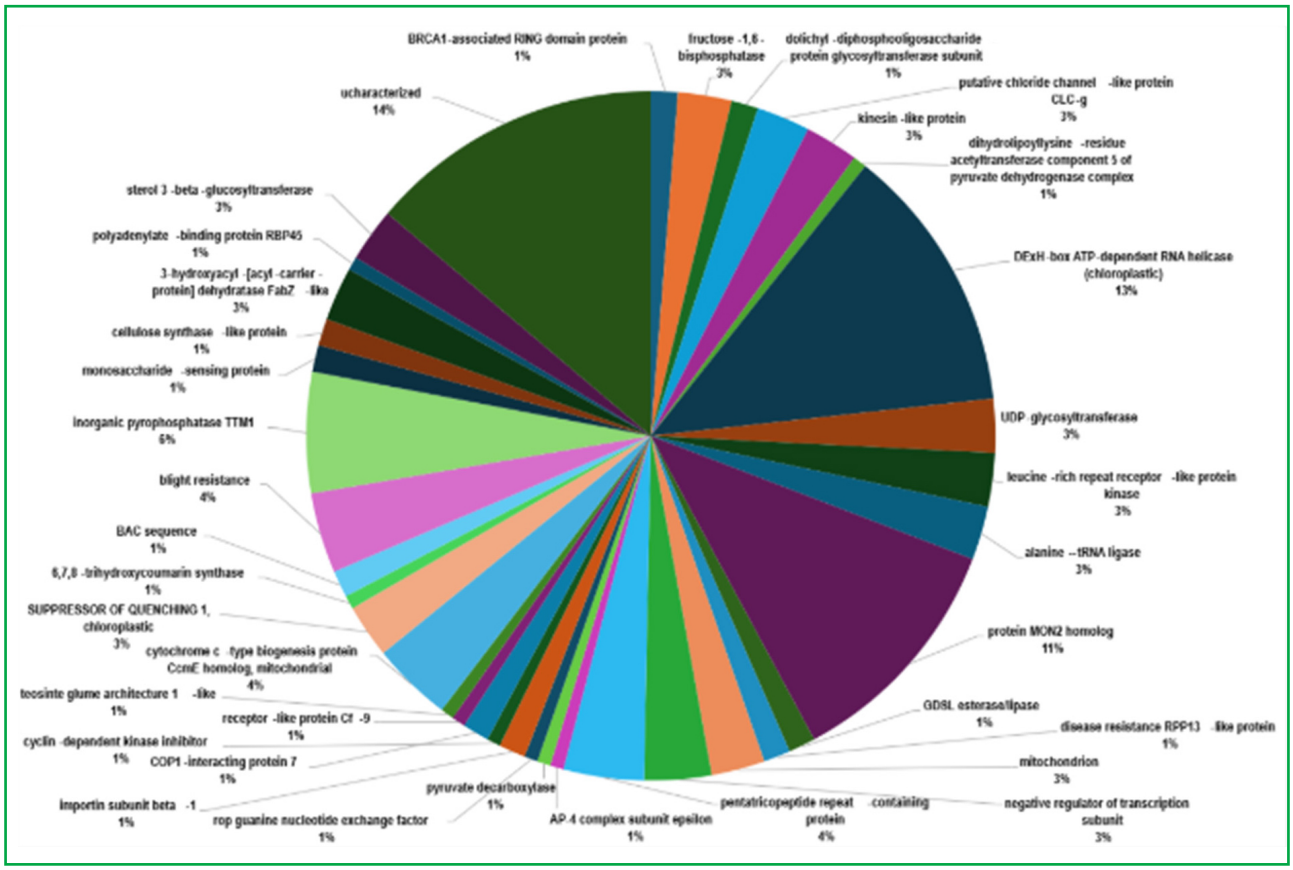

Figure 9 Predicted binding sites of BBAP primer set in the *Capsicum annuum* L. genome NCBI Primer-BLAST Analysis

of matches (14%) was found in uncharacterised sequences, followed by 13% for DexH-box ATP-dependent RNA helicase (chloroplastic) mRNA records and 11% for protein MON2 homolog mRNA sequences. The remaining target sequences are presented in Figure 9.

The identified potential products from the *in silico* analyses were represented by genes with various functions. This indicates that these primer sets are amplifying and identifying polymorphisms linked to different functional areas within plant genomes, as also noted by Poczai et al. (2013) and Yamanaka et al. (2003). The results for the WRKY primer pair matched several sequences related to the WRKY transcription factor. A comprehensive study revealed the existence of 61 *CaWRKY* protein-coding genes. Interestingly, CaWRKY25 is directly linked to the pungency of peppers (Zhang et al., 2023). As for PBA method most matches (36%) were found to serine/threonine protein phosphatase sequences, they play role in plant stress signalling (Yin et al., 2020) and *beta-amyrin synthase* (50%) a key biosynthetic enzyme for terpenoid biosynthesis (Shabani et al., 2010). One potential target product was related to the retrotransposon sequence, which is in correlation with the findings of Yamanaka et al. (2023), where the homology search across plant species using the amplified fragment from the PBA primer sets also found an 8% match to retrotransposon-like fragments. The *in silico* analysis for the BBAP method yielded intriguing results, with the primers used producing numerous potential matches (159) on a diverse array of products. These findings suggest that sequencing some of the products would be beneficial in understanding how the actual PCR products correspond to the *in silico* results. Given that the primers primarily target Bet v 1 homologs and there is currently limited information regarding *ypr10* genes in *Capsicum* spp., the *in silico* results may not accurately reflect reality.

4 Conclusions

Polymorphic marker techniques continue to play a significant role in the study of genetic diversity, offering valuable insights into both intraspecific and interspecific relationships. These methods provide scientists with cost-effective and efficient alternatives to sequencing. Additionally, bioinformatics tools facilitate *in silico* verification of the methodologies' applicability for specific species under investigation.

Often, these methodologies can reveal differences even with a limited number of samples. All used DNA-based marker methods were evaluated as applicable and informative in the genus *Capsicum* spp., from moderately informative PBA, through highly informative iPBS and BBAP to very highly informative CDDP. All methods described the same interspecific variability; intraspecific variability depended on the technique that was used. Results indicate that employing several DNA-based marker methods is more appropriate for examining genetic diversity or biodiversity rates; nonetheless, it is crucial to use the relevant ones.

Acknowledgments

This study was supported by the Slovak Research and Development Agency, APVV-22–0348.

Conflict of Interest

The authors declare that they have no competing interests.

Author Contributions

S. Farkasová: study design, methodology, collecting of data, data analysis, software, writing original draft preparation, writing review and editing, L. Urbanová: study design, methodology, collecting of data, data analysis, software, writing original draft preparation, writing review and editing, J. Žiarovská: writing original draft preparation, writing review and editing.

AI and AI-Assisted Technologies Use Declaration

English grammar correction – QuillBot and ChatGPT.

References

Amiryousefi, A., Hyvönen, J., & Poczai, P. (2018). iMEC: online marker efficiency calculator. *Appl Plant Sci*, 6(6), e01159. https://doi.org/10.1002/aps3.1159

Amiteye, S. (2021). Basic concepts and methodologies of DNA marker systems in plant molecular breeding. *Heliyon*, 7(10), e08093. https://doi.org/10.1016/j.heliyon.2021.e08093

Aziz, R.R., & Tahir, N.A.R. (2023). Genetic diversity and structure analysis of melon (*Cucumis melo* L.) genotypes using URP, SRAP, and CDDP markers. *Genet Resour Crop Evol*, 70, 799–813. https://doi.org/10.1007/s10722-022-01462-y

Barchenger, D.W., & Khoury, C.K. (2022). *A global strategy for the conservation and use of Capsicum genetic resources.* Global Crop Diversity Trust, Bonn, Germany. https://doi.org/10.5281/zenodo.8367264

Bosland, P., & Baral, J.B. (2007). 'Bhut Jolokia' – the world's hottest known chile pepper is a putative naturally occurring interspecific hybrid. *HortScience,* 42, 222–224. https://doi.org/10.21273/HORTSCI.42.2.222

Bosland, P.W., Coon, D., Reeves, G. (2012). 'Trinidad Moruga Scorpion' pepper is the world's hottest measured chile pepper at more than two million scoville heat units. *HortTechnology,* 22, 535–538. https://doi.org/10.21273/HORTTECH.22.4.534

Collard, B.C.Y., & Mackill, D.J. (2009). Conserved DNA-derived polymorphism (CDDP): a simple and novel method for generating DNA markers in plants. *Plant Mol Biol Rep*, 27(4), 558–562. https://doi.org/10.1007/s11105-009-0118-z

FAO. (2023). *FAOSTAT: Crops and livestock products*. Licence: CC-BY-4.0. (Accessed on 3rd September 2025). https://www.fao.org/faostat/en/#data/QCL/visualize.

FAO. *Land & Water: Pepper.* Licence: CC-BY-4.0. [Accessed on 21 July 2025]. https://www.fao.org/faostat/en/#data/QCL/visualize.

Fernandes, H. et al. (2013). Structural and functional aspects of PR-10 proteins. *The FEBS Journal,* 280(5),1169–99. https://doi.org/10.1111/febs.12114

Kalendar, R. et al. (2010). iPBS: a universal method for DNA fingerprinting and retrotransposon isolation. *Theor Appl Genet*, 121,1419–1430. https://doi.org/10.1007/s00122-010-1398-2

Kumar, S. et al. (2018). Peppers (*Capsicum* spp.): domestication and breeding for global use. In Mandal D, Shukla AC, Siddiqui MW (eds) *Sustainable horticulture, vol 1: diversity, production, and crop improvement, Part III. Crop improvement and biotechnology.* Apple Academic/CRC Press, Waretown (pp. 387–400).

Liu, F., Zhao, J., Sun, H. et al. (2023). Genomes of cultivated and wild *Capsicum* species provide insights into pepper domestication and population differentiation. *Nature Communications*, 14(1), 5487. https://doi.org/10.1038/s41467-023-41251-4

Nadeem, M.A. et al. (2017). DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. *Biotechnol Equip*, 32(2), 261–285. https://doi.org/10.1080/13102818.2017.1400401

Nelson, D., & Werck-Reichhart, D. (2011). A P450-centric view of plant evolution. *Plant J*, 66(1), 194–211. https://doi.org/10.1111/j.1365-313X.2011.04529.x

Oksanen, J. et al. (2025). *Vegan: Community Ecology Package.* R package version 2.7-1. https://CRAN.R-project.org/package=vegan

Pickersgill, B. (1997). Genetic resources and breeding of *Capsicum* spp. *Euphytica,* 96(1),129–133.

Poczai, P. et al. (2013). Advances in plant gene-targeted and functional markers: a review. *Plant Methods*, 9(1), 6. https://doi.org/10.1186/1746-4811-9-6

Pramanik, K. et al. (2020). Factors influencing performance of *Capsicum* under protected cultivation: a review. *International Journal of Environment and Climate Change*, 10(12), 572–588. https://doi.org/10.9734/IJECC/2020/v10i1230339

Ramchiary, N., & Kole, C. (2019). Preface to the volume. *The Capsicum Genome*. Springer International Publishing.

Schenk, M.F. et al. (2009). Characterization of PR-10 genes from eight Betula species and detection of Bet v 1 isoforms in birch pollen. *BMC Plant Biol*, 9(1), 24. https://doi.org/10.1186/1471-2229-9-24

Serrote, C.M.L. et al. (2020). Determining the polymorphism information content of a molecular marker. *Gene*, 726, 144175. https://doi.org/10.1016/j.gene.2019.144175

Shabani, L., Ehsanpour, A.A., & Esmaeili, A. (2010). Assessment of squalene synthase and *beta-amyrin synthase* gene expression in licorice roots treated with methyl jasmonate and salicylic acid using real-time qPCR. *Russ J Plant Physiol*, 57, 480–484. https://doi.org/10.1134/S1021443710040047

Strategic plan for biodiversity 2011–2020. (2011) Secretariat of the Convention on Biological Diversity. https://www.cbd.int

Swarup, S. et al. (2021). Genetic diversity is indispensable for plant breeding to improve crops. *Crop Sci*, 61(2), 839–852. https://doi.org/10.1002/csc2.20377

Urbanová. L., & Žiarovská, J. (2021). Variability of DNA based amplicon profiles generated by Bet v 1 homologous among different vegetable species. *Acta Fytotechnica et Zootechnica*, 24(1), 1–6. https://doi.org/10.15414/afz.2021.24.mi-apa.1-6

Yamanaka, S. et al. (2003). Assessment of cytochrome P450 sequences offers a useful tool for determining genetic diversity in higher plant species. *Theor Appl Genet*, 108(1), 1–9. https://doi.org/10.1007/s00122-003-1403-0

Yin, J. et al. (2020). Expression characteristics and function of CAS and a new *beta-amyrin synthase* in triterpenoid synthesis in birch (*Betula platyphylla* Suk.). *Plant Sci*, 294. https://doi.org/10.1016/j.plantsci.2020.110433

Zhang, L. et al. (2023). Genome-Wide Identification of WRKY Gene Family and Functional Characterization of CcWRKY25 in *Capsicum chinense*. *Int J Mol Sci*., 24(14), 11389. https://doi.org/10.3390/ijms241411389

Žiarovská, J., & Urbanová, L. (2022). Utilization of Bet v 1 homologs based amplified profile (BBAP) variability in allergenic plants fingerprinting. *Biologia*, 77(2), 517–523. https://doi.org/10.1007/s11756-021-00943-2