agronomy

MDPI

*Article*

# A Reliable Molecular Diagnostic Tool for CA90 (*Castanea sativa* × *Castanea crenata*) Hybrid Identification Through SSR

**Toufiq Soale Yussif [1,2], Nadine Evora da Cruz [1], Valentim Coelho [1,2,3], Eugénia Gouveia [1,2,3,*] and Altino Branco Choupina [1,2,3]**

[1] Escola Superior Agrária de Bragança, Instituto Politécnico de Bragança, Campus de Santa Apolónia, Apartado 1172, 5301-855 Bragança, Portugal; toufiq@ipb.pt (T.S.Y.); nadinecruzzy@gmail.com (N.E.d.C.); valentimcoelho@ipb.pt (V.C.); albracho@ipb.pt (A.B.C.)
[2] Centro de Investigação de Montanha (CIMO), Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal
[3] Laboratório para a Sustentabilidade e Tecnologia em Regiões de Montanha, Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal
* Correspondence: egouveia@ipb.pt; Tel.: +351-273303276

**Abstract:** Chestnut trees are an essential source of both food and timber. However, the severe threats from invasive pests and diseases compromise their existence and productivity. In Europe, chestnut hybridization programs have been initiated to produce resilient rootstocks in response to ink disease. However, the gap in the identification of these hybrid plants is typically based on field observations and morphological features and remains a challenge. Our study presents a marker set for distinguishing between chestnut hybrid CA90 (*Castanea sativa* × *Castanea crenata*), a hybrid with demonstrated resistance to *Phytophthora cinnamomi*, and other varieties using microsatellite (SSR) markers and bioinformatics tools. We used 35 chestnut samples, including three CA90 controls, hybrids sampled within Portugal, with an aim to define the profiles of the chestnut hybrids and varieties in this study based on band patterns and SSR motifs. We selected and modified nine distinct SSR primers with null allelic features from 43 already developed simple sequence repeat (SSR) markers. PCR amplification and agarose gel electrophoresis were used to amplify and visualize the DNA bands. To confirm genetic variations, 27 amplified bands were sequenced by Sanger sequencing. This analysis identified 31 SSRs across 22 SSR-containing sequences, with trinucleotide (67.74%) repeats being the most common, followed by repeats of dinucleotide (22.58%), mononucleotide (6.45%), and hexanucleotide (3.23%). A total of 18 alleles were observed for the nine loci. The alleles ranged from one to three per locus for the 35 samples. The novel locus CP4 could only be found in CA90 hybrids. This tool can aid in identifying and selecting disease-resistant hybrids, thereby contributing to chestnut production and management strategies.

**Keywords:** chestnut hybrids; *Castanea sativa*; microsatellites markers (SSR); CA90



Academic Editor: Ai-Sheng Xiong

Received: 21 January 2025
Revised: 17 February 2025
Accepted: 20 February 2025
Published: 23 February 2025

**Citation:** Yussif, T.S.; Evora da Cruz, N.; Coelho, V.; Gouveia, E.; Choupina, A.B. A Reliable Molecular Diagnostic Tool for CA90 (*Castanea sativa* × *Castanea crenata*) Hybrid Identification Through SSR. *Agronomy* **2025**, *15*, 543. https://doi.org/10.3390/agronomy15030543



## 1. Introduction

The chestnut tree (*Castanea*: Fagaceae) has been valued for its economic, cultural, and ecological significance in Europe, Asia, and North America. Their family includes other important tree species, such as oaks (*Quercus* spp.) [1]. In Portugal, chestnut trees (mainly *Castanea sativa*) have significant cultural and economic importance. The regions of Trás-os-Montes and Beiras are known for their extensive chestnut orchards and high-quality production, accounting for more than 80% of production (34,000 to 38,000 tons) [2],

particularly varieties like "Martainha", "Longal", and "Judia". Chestnuts are celebrated annually at the *Magusto* festival in early November, where locals gather to roast and celebrate the harvest [3,4].

Chestnut trees (*Castanea* spp.) are susceptible to invasive pests and diseases, which can seriously jeopardize their existence and productivity due to their genetic adaptation to different environmental conditions. Chestnut ecosystems are currently threatened by different stress factors (natural or anthropogenic), such as climate change, the abandonment of traditional orchards, wildfire, and an increased incidence of pests and diseases [5]. The pathogens of ink disease, the oomycetes *Phytophthora cinnamomi* and *Phytophthora cambivora*, and the chestnut blight fungi *Cryphonectria parasitica* are the most severe chestnut disease pathogens [6,7]. At the end of the nineteenth century, ink disease drastically threatened European chestnuts. The first record of this disease in Portugal was in 1838, although it is believed to have been present in Spain since 1726 [8]. The repercussions of biotic and abiotic stress factors and predictions of the implications of climate change on pathogens shift distribution, coupled with the consequent manifestation of distinct disease phenotypes, led to the development of novel chestnut hybrids that are better suited to prevailing climatic conditions and biotic stress [9].

To counter these severe diseases, mitigating research methodologies have been diligently employed, ultimately culminating in identifying a promising approach: developing hybridized plant varieties fortified with resistance mechanisms [7,10]. Hybrid chestnuts derived from Asian and European sources have proven notably resistant to ink diseases [11]. Since the 1950s, breeding programs with European and Japanese chestnuts (*C. sativa* × *C. crenata*) have been established in Portugal, France, and Spain to obtain hybrids that are tolerant to ink disease while maintaining fruit production and quality traits to satisfy commercial demands [12,13]. More common control approaches are correctly managing nurseries/orchards, using resistant rootstocks for propagation, or planting resistant hybrids for production [14]. However, clear strategies for distinguishing and maintaining established hybrid plants are yet to emerge, resulting in farmers inadvertently cultivating non-hybrid varieties. The selection of cultivars is an essential step in establishing and maintaining orchards. Helping nurseries and growers maintain good genetic records will help answer questions about specific cultivars' performance. In the Mediterranean countries, chestnut cultivar identification and genetic resources assessment are significant issues leading to research by groups in countries such as Italy [15], Spain [16], and Portugal [4,17].

In Portugal, chestnut genetic diversity was analyzed through a preliminary study using single sequence repeat (SSR), and it was successfully used to type traditional Portuguese chestnut varieties [4]. The results suggested low genetic variability among the Portuguese chestnut varieties. Santos et al. [18] revealed that *C. crenata* and *C. mollissima* have been the foundation of the Portuguese chestnut breeding programs. Another more recent breeding program established by Costa et al. [19] focused on the development of *C. sativa* × *C. crenata* and *C. sativa* × *C. mollissima* hybrids, where four F1 hybrids were chosen for extensive propagation according to their potential to proliferate and root in vitro, their ink disease resistance levels, and field development [20,21]. A clonal collection of Japanese chestnut hybrids was crossed with European chestnut (*C. sativa*) to produce nuts, timber, and rootstocks for local cultivars [11]. Moreover, a group of hybrid genotypes selected as the most resistant to *P. cinnamomi* is being propagated as improved genetic materials for new rootstocks released to the market [12].

Chestnut varietal characterization has been carried out in recent decades using different approaches, and more recently, the development of DNA markers has provided a direct study of genotypes that enable the identification of chestnut cultivars [22]. In the last 20 years, microsatellite markers (SSRs) consisting of tandem repeats of polymorphic short-

sequence motifs (2–8 bp), co-dominant, inherited, highly polymorphic, and reproducible, are the most suitable to detect genetic variability [11,23]. Hybrid identification plays a crucial role in plant breeding, ensuring genetic purity and enhancing crop improvement programs. SSR markers provide an efficient molecular approach for distinguishing parental lines from hybrids, facilitating breeding programs and genetic conservation strategies in chestnut [24]. Alessandri et al. [25] monitored gene flow between wild and cultivated chestnut trees, revealing the significant role of hybridization in shaping genetic diversity through SSR markers. Tumpa et al. [26] also confirmed the high effectiveness of both genomic SSR (gSSR) and expressed sequence tag SSR (EST-SSR) markers in describing the genetic diversity of sweet chestnut, highlighting their utility in hybrid identification. A comprehensive analysis involving 216 chestnut accessions, from Japanese, Chinese, European, and American species and Japanese–Chinese hybrids, utilized SSR markers to assess genetic diversity and relationships. The study identified 21 synonym groups and 189 distinct genotypes, demonstrating the effectiveness of SSR markers in distinguishing between chestnut species and their hybrids [27]. More recently, SSRs has been used, in maize genomic and evolutionary breeding [28], the validation of candidate genes in sugarcane [29], olive gene validation [30,31], and *Quercus* genetic diversity [32]

In Portugal, non-hybrid chestnut plants are cheap, whereas hybrids such as "ferosacre" (CA90) command a higher price. However, in certain instances, rootstock vendors may offer *Castanea sativa* plants while asserting that they are hybrid cultivars. To date, some farmers still depend on traditional methods of hybrid identification, such as morphological characteristics and field observations. However, these methods have limitations in distinguishing closely related individuals from genetically similar hybrids [33]. Molecular techniques, particularly microsatellite analysis, have revolutionized the field of hybrid identification by providing a precise means of differentiating between individuals based on their genetic profiles. Meanwhile, more attention has been paid to genetic diversity and not to hybrid identification tools that enable farmers to plant resistant varieties. It is important to fight plant diseases and find answers to cultivar identification challenges. Our main objective in this study is to develop and validate a genetic identification tool using microsatellite (SSR) markers to distinguish ink disease (Phytophthora cinnamomi)-resistant chestnut hybrids (particularly CA90, a *Castanea sativa* × *Castanea crenata* hybrid) from other varieties to enhance the selection and deployment of resilient rootstocks, ultimately supporting sustainable chestnut production and management strategies. This study is essential because the chestnut industry lacks certification and validation programs for cultivars. Genetic marker technology can play a significant role in helping the industry maintain genetic standards.

## 2. Materials and Methods

### *2.1. Plant Materials and DNA Extraction*

In summer 2023, a total of 35 representative samples consisting of hybrids and traditional Portuguese varieties were collected from chestnut trees in some parts of Portugal: 12 from St[a] Comba, 10 from Vila-Boa de Serapicos, and 13 from Deifil Green Biotechnology, a research institution specializing in in vitro plant propagation for species with inefficient natural reproduction systems. Table 1 provides a detailed description of each accession. Three (3) samples were positive controls, CA90 hybrids (*C. sativa* × *C. crenata*), with well-defined genetic profiles among St[a] Comba samples. After sample collection, the leaves were cleaned with 70% ethanol and stored at −20 °C until DNA extraction.

**Table 1.** Details of the 35 chestnut samples, including laboratory designated codes. M = mark; P = place.

| Lab Code | Samples | GPS/Field Tag | Sampling Location |
|---|---|---|---|
| A | CA90 positive control | M111111.2 P222973.9 | St[a] Comba |
| B | CA90 positive control | M110916.0 P2225509 | St[a] Comba |
| C | CA90 positive control | M110932.6 P222509.7 | St[a] Comba |
| D | *C. Sativa* | TD | Vila-Boa de Serapicos |
| 1 | CA90 | T1 | Vila-Boa de Serapicos |
| 2 | CA90 | T2 | Vila-Boa de Serapicos |
| 3 | CA90 | T3 | Vila-Boa de Serapicos |
| 4 | CA90 | T4 | Vila-Boa de Serapicos |
| 5 | CA90 | T5 | Vila-Boa de Serapicos |
| 6 | CA90 | T6 | Vila-Boa de Serapicos |
| 7 | CA90 | T7 | Vila-Boa de Serapicos |
| 8 | CA90 | T8 | Vila-Boa de Serapicos |
| 9 | CA90 | T9 | Vila-Boa de Serapicos |
| 10 | *C. mollissima* 60907 | T10 | Deifil Green Biotechnology |
| 11 | *C. mollissima* E2604 | T11 | Deifil Green Biotechnology |
| 12 | *C. mollissima* Y0204 | T12 | Deifil Green Biotechnology |
| 13 | *C. mollissima* Z1408 | T13 | Deifil Green Biotechnology |
| 14 | CA90 | T14 | Deifil Green Biotechnology |
| 15 | Martaínha 2 | T15 | Deifil Green Biotechnology |
| 16 | Bouche de Betizac 1 | T16 | Deifil Green Biotechnology |
| 17 | Bouche de Betizac 2 | T17 | Deifil Green Biotechnology |
| 18 | Martaínha 1 | T18 | Deifil Green Biotechnology |
| 19 | Cota | T19 | Deifil Green Biotechnology |
| 20 | Judia | T20 | Deifil Green Biotechnology |
| 21 | Marsol | T21 | Deifil Green Biotechnology |
| 22 | Précoce Migoule | T22 | Deifil Green Biotechnology |
| 23 | Putative CA90 | M111478.8 P223081.2 | St[a] Comba |
| 24 | Putative CA90 | M11609.9 P222967.4 | St[a] Comba |
| 25 | Putative CA90 | M111649.4 P222888.7 | St[a] Comba |
| 26 | Putative CA90 | M111494.7 P222815.0 | St[a] Comba |
| 27 | Putative CA90 | M111497.6 P223077.5 | St[a] Comba |
| 28 | Putative CA90 | M111191.6 P222934.5 | St[a] Comba |
| 29 | Putative CA90 | M111512.1 P222808.6 | St[a] Comba |
| 30 | Putative CA90 | M111486.4 P222934.5 | St[a] Comba |
| 31 | Putative CA90 | M111477.6 P222821.3 | St[a] Comba |

In the laboratory, all harvested leaves were deprived of veins and reduced to a fine powder in liquid nitrogen using a mortar and pestle. Genomic DNAs were extracted from 0.05 g of leaves using a Thermo Scientific GeneJET Plant Genomic DNA Purification Mini Kit. Vilnius, Lithuania, following the manufacturer's protocol.

### 2.2. SSR Amplification

In order to define a set of microsatellite sequences useful for the fingerprinting and profiling of chestnut samples in this study, we selected and modified nine (9) SSR primer pairs (Table 2) from originally designed primers by Santos et al. [21]. The selection criteria of the primers pairs were based on primers with null alleles and low observed heterozygosity (Ho) values, ranging from 0.000 to 0.552, as our goal was to target specific genotypes in disease-resistant hybrids, thereby enhancing the success of selective breeding. From that study, the primer loci CcPT 0002 and CcPT 0003 were maintained but renamed CP2 and CP3 according to this study (Table 2). The modifications affected CP4 to CP11, including adding one or two nucleotides to the 3′ end of the forward and/or the reverse primers to influence the binding specificity or efficiency of the primers.

**Table 2.** Primer sequences used for SSR analysis and their characteristics.

| NO | Locus | Forward (5′-3′) | Reverse (5′-3′) | Size (bp) | Annealing Tm (°C) |
|---|---|---|---|---|---|
| 1 | CP2 | AGTTCTCCACGAGGCTCAAA | TCCAAGCTGGAGAATCATCA | 220 | 55.3 |
| 2 | CP3 | GGTGCCCAGATTTACGAGAA | ATCGCTTGGAGTCACAGCTT | 240 | 57.3 |
| 3 | CP4 | GCTGCTTCACAACCTTCCTC | GCAAGAGATTCCCTTTGCTG | 220 | 57.3 |
| 4 | CP5 | ACACATGGGGGTGTGAACTT | TTATGGGAAACGGCATCTTC | 125 | 55.3 |
| 5 | CP6 | CCTGTGAGGCTAAGAGAGCG | ACCACGTCGGTGCTTCTAGT | 200 | 59.4 |
| 6 | CP8 | TCGTCCCCTTCTTCATCATC | ATATGGCCAAAAACCCATCA | 250 | 53.2 |
| 7 | CP9 | TTCCACCCAATTGTTACCAC | GATGAAGAAGGGGACGA | 200 | 55.3 |
| 8 | CP10 | ATCCATGAGTGAAAGCCACC | TGGAACAAGAAGCCTCGATT | 250 | 55.6 |
| 9 | CP11 | TCATCCAAGAAGCCCTCAAC | TTCTGCCTCTTTTGTTGCCT | 230 | 55.3 |

Target loci for this study were amplified by performing a PCR reaction in a final volume of 50 µL containing 0.2 mM dNTPs, 1X reaction buffer, 0.8X KCl, 1.5 mM $MgCl_2$, 2.5 units DFS-DNA Taq polymerase, 2.5 µM each of Forward and Reverse primers, 4 µL genomic DNA (20 ng/µL), 0.6 mg/mL BSA, and 29.25 µL deionized water. The following thermocycling conditions were used: initial denaturation, final denaturation, annealing, extension, and final extension at 94 °C-2 min; 94 °C-10 s, 57 °C-30 s, and 72 °C-30 s for 36 cycles and 72–5 min. The internal temperature of the thermocycler was adjusted to 4 °C for the optimal maintenance of amplified DNA fragments. A summary of all the primers used is shown (Table 2). The PCR program was based on Botta et al. [34] with few modifications. After amplification, the fragments were separated using agarose gel electrophoresis of 1.5%. The PCR products were labeled with ethidium bromide, run at 80 Volts, and visualized using the Bio-Rad gel Doc™ Ez Imager (Image Lab application 5.2.1).

### 2.3. Sequencing

Twenty-seven DNA amplicons were randomly selected from different sets of nine primers and chestnut samples. For each primer, at least three different bands from different tree samples were selected for sequencing and SSR motif identification to ensure reliability and reproducibility. The selected bands were purified using a DNA Clean & Concentrator kit from ZYMO Research, Irvine, CA, USA, quantified with a nanodrop to assess DNA concentration and purity following the Beer–Lambert law, and sent to the University of Salamanca for Sanger sequencing.

### 2.4. Bioinformatic Analysis

Briefly, (I) sequence processing and quality control of raw DNA sequence reads obtained by Sanger sequencing and trimmed using BioEdit (version 7.2.5) to remove low-quality bases and ensure high-confidence reads was performed. (II) Sequence alignment and similarity search were aligned using ClustalW version 2.1 to check for genetic variations

and ensure accurate sequence comparison. (III) BLASTn (NCBI) was used to compare sequences with reference databases in GenBank, identifying homologous sequences and confirming species identity. (IV) Microsatellite identification (SSR analysis) was based on MISA-web to detect simple sequence repeats (SSRs) in the analyzed sequences, considering the parameters of motif lengths ranging from 1 to 6 bases, a minimum of 5 repeats, and detection of compound microsatellites. Finally, MEGA11 (Molecular Evolutionary Genetics Analysis software, version 11) was used to construct a phylogenetic tree using the maximum likelihood (ML) method and the Tamura-Nei model with a Bootstrap value obtained through 100 replicates, providing a measure of clustering confidence. Strongly supported clades were defined by bootstrap values above 70% indicating robust clustering in several branches suggesting strong statistical support for these groupings. Moderately supported clusters were noted in branches with values between 50 and 70%, while lower bootstrap values (<50) indicate areas of weaker support.

## 3. Results

### *3.1. Effects on Primer Locus Modification*

The modifications of the primers affected the results of this study. Our primer set consistently amplified different regions and could be considered a new variant or loci. We named them locus (CP2, CP3, CP4, CP5, CP6, CP8, CP9, CP10, and CP11). The marker CP2 and CP11 had the lowest number of alleles (one), while CP3 and CP6 both shared the highest number of alleles (three). Regarding motif types, CP3, CP6, CP8, and CP10 exhibited compound motifs, with CP3 showing the highest. See (Table 3).

**Table 3.** Comparison of results for the modified SSR primer loci: number of alleles (Na); *—no motifs found.

| Locus | Sample | Na | Motif | Reference | Loci | Na | Motif | Reference |
|---|---|---|---|---|---|---|---|---|
| CP2 | Marsol | 1 | $(GTG)_9$ | This study | CcPT 0002 | 1 | $(GTG)_9$ | [21] |
| CP2 | CA90 | | * | This study | | | | |
| CP2 | *C. sativa* | | * | This study | | | | |
| CP3 | Marsol | 3 | $(ATC)_6$ | This study | CcPT 0003 | 7 | $(ATC)_8$ | [21] |
| CP3 | *C. sativa* | | $(ATC)_6$ | This study | | | | |
| CP3 | *C. mollissima*. Z1408 | | $(ATC)_6$ | This study | | | | |
| CP3 | CA90 | | $(TGA)_5$tcatgatgaccacaaggattgaa gttagtcacagcatctcggccaccaacgcg ttgggccgcgatgtcgtacgcttttgca (TCT)5 | This study | | | | |
| CP3 | Martainha | | $(TGA)_8$tcatcagtaccacccggcgggttga agttcacaacatcttctccaccacaaaagcgg ggggccatgatgtcatattttgttg (CTT)5 | This study | | | | |
| CP4 | CA90 | 2 | $(GA)_{10}$ | This study | CcPT 0004 | 6 | $(CT)_{10}$ | [21] |
| CP4 | CA90 | | $(GA)_{11}$ | This study | | | | |
| CP5 | CA90 | 3 | $(CT)_8$ | This study | CcPT 0005 | 7 | $(CT)_1$ | [21] |
| CP5 | *C. sativa* | | $(CT)_{12}$ | This study | | | | |
| CP5 | COTA | | $(AG)_7$ | This study | | | | |
| CP6 | Marsol | 2 | $(GAA)_5$ | This study | CcPT 0006 | 4 | (TTC) | [21] |
| CP6 | CA90 | | $(GAA)_5$ gaggaagaagaac$(A)_{12}$ | This study | | | | |
| CP6 | CA90 | | * | This study | | | | |
| CP6 | *C. sativa* | | * | This study | | | | |

**Table 3.** *Cont.*

| Locus | Sample | Na | Motif | Reference | Loci | Na | Motif | Reference |
|---|---|---|---|---|---|---|---|---|
| CP8 | *C. sativa* | 2 | $(CTCAGA)_5$ gtacaacaaccgacagca $(AAG)_{11}$ | This study | CcPT 0008 | 5 | $(TCT)_1$ | [21] |
| CP8 | CA90 | | $(A)_{12}$ $(TCT)_8$ | This study | | | | |
| CP9 | CA90 | 2 | $(AG)_{10}$ | This study | CcPT 0009 | 8 | (TC) | [21] |
| CP9 | *C. sativa* | | $(TC)_8$ | This study | | | | |
| CP9 | CA90 | | * | This study | | | | |
| CP10 | CA90 | 2 | $(CAC)_6$ $(AAG)_5$ | This study | CcPT 0010 | 5 | (GGT) | [21] |
| CP10 | *C. sativa* | | $(CAC)_6$ $(AAG)_5$ | This study | | | | |
| CP10 | CA90 | | $(GGT)_5$ gggggagccttc $(TCT)_6$ | This study | | | | |
| CP10 | CA90 | | $(GGT)_5$ gggggagccttc $(TCT)_6$ | This study | | | | |
| CP11 | CA90 | 1 | $(GGT)_7$ | This study | CcPT 0011 | 5 | (CAC) | [21] |

### 3.2. Locus Amplifications and Band Pattern Representation

In this study, all primer sets gave good amplifications for all loci in CA90 samples and not in any varieties due to their null allelic nature. The gel images depicted monomorphic bands but high polymorphisms when sequenced, except for CP8, which also showed polymorphic bands on gel. Meanwhile, the accurate detection of SSR primer polymorphism requires high-resolution gels or capillary electrophoresis, which was a limitation of our study since we only relied on the available agarose gel electrophoresis; therefore, we suggest further experimentation with high-resolution gels. A heatmap (Figure 1) was created based on the visual representation of band appearance patterns from the gel images (see Supplementary Materials Figures S1–S3) to create an effective profile to enable the selection of the most effective primer pairs that well differentiated hybrids from varieties.

It was evident from the heatmap that samples 1–9 from Vila-Boa de Serapicos collectively exhibited a remarkable 100% genetic identity with the three controls in all nine primer loci, confirming them as CA90 hybrids. In stark contrast, Sample 14, identified as CA90, from Deifil Green Biotechnology, displayed a questionable genetic match of 55.6% and could not be identified as CA90. The St[a] Comba samples (23–31) could not be identified as CA90 since their genetic status did not match the controls. Additionally, sample D, recognized as "*C. sativa*", exhibited a striking 88.9% genetic resemblance to the controls in all primer loci. This high genetic match with the controls supports CA90 as a natural interspecific hybrid derived from *C. crenata* × *C. sativa*, indicating gene inheritance from both parental lines.

### 3.3. Sequencing Results

The obtained raw DNA sequence reads were subjected to BioEdit (version 7.2.5) for quality control assessment. The resulting sequences were then submitted to ClustalW for sequence alignment and finally to MISA-web following the procedures of [35] for microsatellite prediction (Table 3) to confirm the results of the gel images. The 27 examined sequences collectively comprised 8899 base pairs. Among these, 31 SSRs were identified and distributed across 22 sequences containing SSRs; nine sequences contained more than one SSR, and six SSRs were found to be involved in compound formation. The search parameters for SSRs were set, specifying a motif length ranging from 1 to 6, a repetition minimum of five, and no base pairs between the two microsatellites for compound SSR detection. The analysis revealed that the majority (67.74%) of the identified microsatellites were of the trinucleotide type, with dinucleotide (22.58%), mononucleotide (6.45%), and

hexanucleotide (3.23%) microsatellites also present, albeit in smaller proportions. A total of 18 alleles were obtained for the nine loci. The alleles ranged from one to three per locus for the 35 samples.

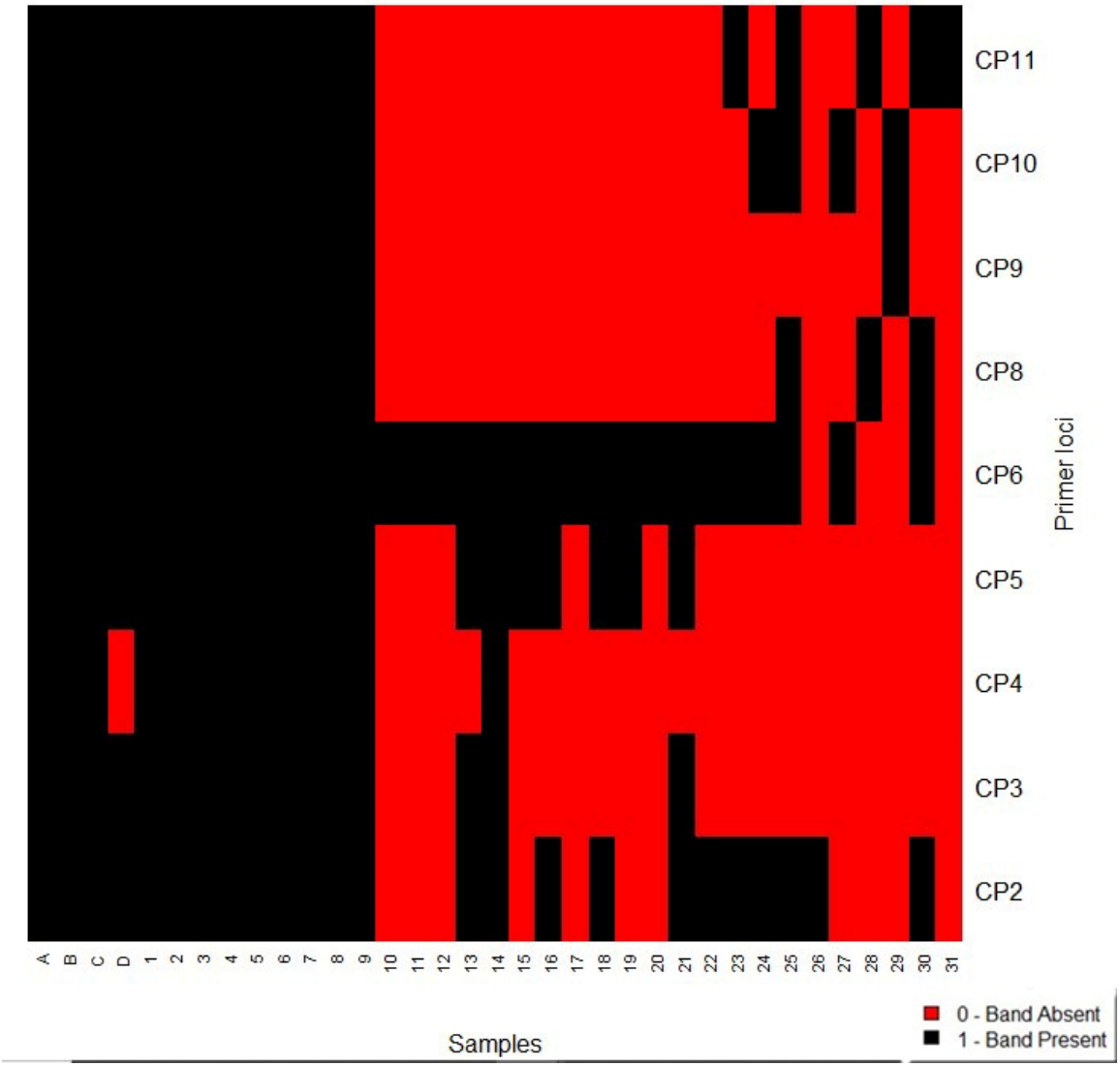


**Figure 1.** A heatmap of band patterns from gel images. (ABC)—CA90 controls, (D)—*C. sativa*, (1–9)—CA90, (10)—*C. mollissima* 60907, (11)—*C. mollissima* E2604, (12)—*C. mollissima* a Y0204, (13)—*C. mollissima* Z1408, (14)—CA90-deifil, (15)—Martainha 2, (16)—Bouche de Betizac 1, (17)—Bouche de Betizac 2, (18)—Martainha 1, (19)—Cota, (20)—Judia, (21)—Marsol, (22)—Precoce migoule, (23–31)—Putative CA90.

### *3.4. Phylogenetic Tree Analysis*

An evolutionary history was inferred with a phylogenetic tree analysis (Figure 2) using the maximum likelihood method and the Tamura-Nei model [36] with a bootstrap consensus tree inferred from 100 and BioNJ algorithms to a matrix of estimated pairwise distances using the Tamura-Nei model. The analysis involved all the 27 nucleotide sequences. There were 613 positions in the final dataset. The evolutionary analyses were conducted using the MEGA11 software [36]. The phylogenetic tree analysis validated the results obtained by gel electrophoresis and sequencing. These results reaffirmed the close genetic relatedness among samples with bands for primers CP2, CP4, CP6, CP9, and CP10 loci while highlighting considerable genetic diversity among hybrids with bootstrap values

above 95%. The sequences from these loci were considered, suggesting the robustness of the inferred relationship.

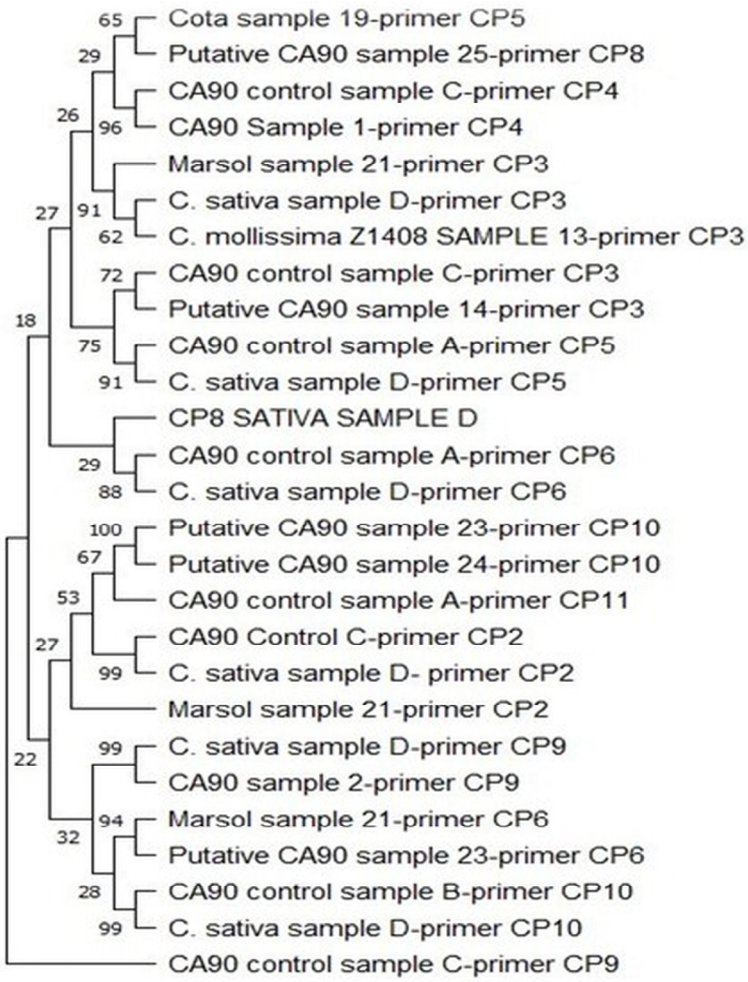


**Figure 2.** Bootstraps consensus phylogenetic tree showing similarities between sequences from primer loci with a Bootstrap value of 100 replicates indicating clustering specificity according the numbers in the tree.

## 4. Discussion

Genetic variation is crucial for chestnut species to withstand and adapt to adverse environ mental conditions [5]. There are currently no existing profiles/patterns available for chestnut hybrid or varietal identification. Farmers still relay upon morphological features for identification and differentiation. In this study, we conducted a profile for some hybrids and Portuguese chestnut varieties based on band patterns and the existence of motifs using modified SSR marks from Santos et al. [21]. Our studied well-profiled CA90 hybrid (with a 100% genetic identity with our control samples) and *C. sativa* (with a 88.9% genetic resemblance) based on band appearance from all the nine null allelic markers used. *C. sativa* was missing a band from marker CP4 (the most significant locus in this investigation, showing bands exclusively in the CA90 samples). While the gel images in this study showed monomorphic bands, the sequences from all primer loci were characterized by distinct motifs, as detailed in Table 3, proving all nine primer pairs are highly polymorphic. This was a limitation to our study, and we recommend the use of high-resolution capillary electrophoresis in future experimentations to enable a suitable genetic polymorphism type conclusion. This was attributed to the samples' homoplasy, allelic variation, and inherent genetic diversity.

We examined twenty-seven (27) sequences and identified 31 SSRs distributed across 22 sequences containing SSRs. The resulting motifs found revealed that the majority identified microsatellites of (67.74%) trinucleotides, (22.58%) dinucleotides, (6.45%) mononucleotide, and (3.23%) hexanucleotide microsatellites. Our results were compared with those of Santos et al. [21], from which the primers were modified (see Table 3), and showed differences regarding the SSR motifs, allelic size, and number of alleles ranging from 1 to 3 per locus for our study against 1 to 11, except for CP2, which was not modified. The results aligned with the low genetic variability of Portuguese chestnut varieties in the studies by Costa et al. [4]. The range of variability makes it possible to select loci for specific applications. For example, markers with only a few alleles are well-suited for population genetic studies [37]. We also observed some intra-varietal differences within the studied accessions; for example, CA90 has been characterized by the same SSR motifs but different numbers of repeats. This aligns with the results by Dinis et. al. [38]. They used SSR markers to assess the heterogeneity of the genotype of the cultivar "Judia" in Trás-os-Montes.

Sequences from different samples at locus CP4, using forward primer (5′-GCTGCTT CACAACCTTCCTC-3′) and reverse primer (5′-GCAAGAGATTCCCTTTGCTG-3′) consistently exhibited identical SSR motifs [$(GA)_{10}$ and $(GA)_{11}$] and variations in alleles. This validation underscores the success of our objective of developing and validating a genetic identification tool using microsatellite (SSR) markers to distinguish ink disease-resistant chestnut hybrids (particularly CA90) from other varieties. The sequences, when compared with sequences published in the GenBank database using the BLAST algorithm, (https://blast.ncbi.nlm.nih.gov/Blast.cgi, accessed on 13 October 2024) showed nucleotide similarity greater than 95%, belonging to *Quercus lobata*, *Quercus variabilis*, *Quercus suber*, and *Quercus rober*, confirming the lack of DNA sequence data for *Castanea*, which supports the view that most SSRs currently available are developed from repeat-enriched genomic DNA libraries from *Quercus* sequences being helpful in *Castanea* [39,40].

From the similarity phylogenetic tree, we observed differences in the leaf figures, except for CP4, implying that the taxa represented by the leaf nodes were distinct from each other in some way, either genetically, morphologically, or ecologically. A clustering confidence above 70% by bootstrap values indicated robust clustering in several branches, suggesting strong statistical support for these groupings. However, we consider bootstrap values above 85% clustering specificity according to the phylogenetic tree for the selection of the cost-effective marker set in our conclusions for hybrid differentiation. The phylogenetic tree confirmed that *C. mollissima* Z1408 (Chinese hybrid) and Marsol shared genetic relationships with *C. sativa* and CA90 based on their clustering effects.

## 5. Conclusions

The genetic profiles and easy mode of molecular identification for selection of chestnut hybrids and varieties has a significant impact on their continuous existence. While rootstock vendors may sometimes offer *C. sativa* plants while asserting that they are hybrid cultivars, unfortunately, our study only identified one marker (CP4) that can be used to detect such cases of fraud in routine analyses on agarose gels. From the gel image analysis on the heatmap, Locus CP4 emerged as the most significant locus in this investigation, showing bands exclusively in the CA90 samples. The absence of bands in *C. sativa* at locus CP4 and the presence in all loci of CA90 ultimately confirm this loci's significance in differentiating this hybrid. Therefore, for time and cost management, the primer loci CP2, CP4, CP6, CP9, and CP10 without sequencing or capillary electrophoresis could serve as effective primer combinations for differentiating CA90 hybrids and *C. Sativa* since they gave good amplifications and clustering specificity according to the phylogenetic tree.

We intend to collaborate with researchers and institutions in the chestnut field to expand the sample size across Portugal. Additionally, we aim to utilize higher-resolution electrophoresis techniques to obtain well-characterized genetic profiles of Portuguese chestnut hybrids and varieties through our set of markers.

Molecular marker-based characterization is faster, cheaper, and more reliable than traditional morphological classification, although both methods have complementary values. To enhance the effectiveness of SSR markers for chestnut identification, it would be beneficial to evaluate markers designed for species closely related to *Castanea* and those within the *Fagaceae* family. Implementing these technologies in Portugal's chestnut industry could bring several advantages: (1) certification of cultivars and germplasm; (2) precise identification of species, cultivars, and rootstocks; and (3) improved screening of selection in breeding programs. This information will aid in identifying pollinators that influence fruit set, yield, size, and quality. Further research is recommended to expand the SSR marker set for enhanced accuracy across various chestnut species because of the lack of certification and validation programs for cultivars.

**Supplementary Materials:** The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/agronomy15030543/s1. Figure S1: Gel image of primer locus CP4; Figure S2: Gel image of primer loci. A = CP2, B = CP3, C = CP5, D = CP6; Figure S3: Gel image of primer locus E = CP8, F = CP9, G = CP10, H = CP11.

**Author Contributions:** Conceptualization, A.B.C. and T.S.Y.; Funding Acquisition, E.G.; Investigation, T.S.Y. and A.B.C.; Methodology, T.S.Y., A.B.C. and N.E.d.C.; Visualization, T.S.Y.; Writing—Original Draft, T.S.Y.; Writing—Review and Editing, A.B.C., E.G. and V.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by IPB Project—Experimental Program for Biological Control of Chestnut Blight no. 167 and PDR2020—Prevenção da Floresta Contra Agentes Bióticos e Abióticos, CP/02/2022, Portugal.

**Data Availability Statement:** The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest.

## References

1. Clark, S.L.; Marcolin, E.; Patrício, M.S.; Loewe-Muñoz, V. A Silvicultural Synthesis of Sweet (*Castanea sativa*) and American (*C. dentata*) Chestnuts. *For. Ecol. Manag.* **2023**, *539*, 121041. [CrossRef]
2. Ribeiro, M.; Freitas, M.; Domínguez-Perles, R.; Barros, A.I.R.N.A.; Ferreira-Cardoso, J.; Igrejas, G. Nutriproteomics Survey of Sweet Chestnut (*Castanea sativa* Miller) Genetic Resources in Portugal. *Food Biosci.* **2020**, *36*, 100622. [CrossRef]
3. Braga, M.; Araújo, S.d.S.; Sales, H.; Pontes, R.; Nunes, J. Portuguese *Castanea sativa* Genetic Resources: Characterization, Productive Challenges and Breeding Efforts. *Agriculture* **2023**, *13*, 1629. [CrossRef]
4. Costa, R.; Valdiviesso, T.; Marum, L.; Borges, O.; Soeiro, J.; Assunção, A.; Soares, F.M.; Sequeira, J. Characterisation of Traditional Portuguese Chestnut Cultivars by Nuclear SSRs. *Acta Hortic.* **2005**, *693*, 437–440. [CrossRef]
5. Freitas, T.R.; Santos, J.A.; Silva, A.P.; Fraga, H. Influence of Climate Change on Chestnut Trees: A Review. *Plants* **2021**, *10*, 1463. [CrossRef] [PubMed]
6. Rigling, D.; Prospero, S. *Cryphonectria parasitica*, the Causal Agent of Chestnut Blight: Invasion History, Population Biology and Disease Control. *Mol. Plant Pathol.* **2018**, *19*, 7–20. [CrossRef] [PubMed]
7. Anagnostakis, S.L. Chestnut Breeding in the United States for Disease Insect Resistance. *Plant Dis.* **2012**, *96*, 1392–1403. [CrossRef]
8. Vannini, A.; Vettraino, A.M. Ink Disease in Chestnuts: Impact on the European Chestnut. *For. Snow Landsc.* **2001**, *76*, 345–350.
9. Burgess, T.I.; Scott, J.K.; Mcdougall, K.L.; Stukely, M.J.C.; Crane, C.; Dunstan, W.A.; Brigg, F.; Andjic, V.; White, D.; Rudman, T.; et al. Current and Projected Global Distribution of *Phytophthora cinnamomi*, One of the World's Worst Plant Pathogens. *Glob. Change Biol.* **2017**, *23*, 1661–1674. [CrossRef] [PubMed]
10. Anagnostakis, S.L. Chestnut Blight: The Classical Problem of An Introduced Pathogen. *Mycologia* **1987**, *79*, 23–37. [CrossRef]

11. Pereira-Lorenzo, S.; Lourenço Costa, R.; Serdar Ondokuz, U.; Üniversitesi, M.; Ramos-Cabrer, A. Interspecific Hybridization of Chestnut. In *Polyploidy and Hybridization for Crop Improvement*, 1st ed.; Annaliese, S.M., Ed.; CRC Press: Boca Raton, FL, USA, 2016; pp. 379–408.
12. Santos, C.; Nelson, C.D.; Zhebentyayeva, T.; Machado, H.; Gomes-Laranjo, J.; Costa, R.L. First Interspecific Genetic Linkage Map for *Castanea sativa x Castanea crenata* Revealed QTLs for Resistance to *Phytophthora cinnamomi*. *PLoS ONE* **2017**, *12*, e0184381. [CrossRef] [PubMed]
13. Serrazina, S.; Santos, C.; Machado, H.; Pesquita, C.; Vicentini, R.; Pais, M.S.; Sebastiana, M.; Costa, R. *Castanea* Root Transcriptome in Response to *Phytophthora cinnamomi* Challenge. *Tree Genet. Genomes* **2015**, *11*, 6. [CrossRef]
14. Hardham, A.R. *Phytophthora cinnamomi* . *Mol. Plant Pathol.* **2005**, *6*, 589–604. [CrossRef] [PubMed]
15. Boccacci, P.; Akkak, A.; Torello, D.; Bounous, G.; Botta, R. Typing European Chestnut (*Castanea sativa* Mill.) Cultivars Using Oak Simple Sequence Repeat Markers. *HortScience* **2004**, *39*, 1212–1216. [CrossRef]
16. Ramos-Cabrer, A.M.; Pereira-Lorenzo, S. Genetic Relationship between *Castanea sativa* Mill. Trees from North-Western to South Spain Based on Morphological Traits and Isoenzymes. *Genet. Resour. Crop Evol.* **2005**, *52*, 879–890. [CrossRef]
17. Goulao, L.; Valdiviesso, T.; Santana, C.; Oliveira, C.M. Comparison between Phenetic Characterisation Using RAPD and ISSR Markers and Phenotypic Data of Cultivated Chestnut *(Castanea Sativa* Mill). *Genet. Resour. Crop Evol.* **2001**, *48*, 329–338. [CrossRef]
18. Santos, C.; Machado, H.; Serrazina, S.; Gomes, F.; Gomes-Laranjo, J.; Correia, I.; Zhebentyayeva, T.; Duarte, S.; Bragança, H.; Fevereiro, P.; et al. Comprehension of Resistance to Diseases in Chestnut. *Revista de Ciências Agrárias* **2016**, *39*, 189–193. [CrossRef]
19. Costa, R.; Santos, C.; Tavares, F.; Machado, H.; Gomes-Laranjo, J.; Kubisiak, T.; Nelson, C.D. Mapping and Transcriptomic Approches Implemented for Understanding Disease Resistance to *Phytophthora cinammomi* in *Castanea* sp. *BMC Proc.* **2011**, *5*, O18. [CrossRef]
20. Fernandes, P.; Tedesco, S.; da Silva, I.V.; Santos, C.; Machado, H.; Costa, R.L. A New Clonal Propagation Protocol Develops Quality Root Systems in Chestnut. *Forests* **2020**, *11*, 826. [CrossRef]
21. Santos, C.; Zhebentyayeva, T.; Serrazina, S.; Nelson, C.D.; Costa, R. Development and Characterization of EST-SSR Markers for Mapping Reaction to *Phytophthora cinnamomi* in *Castanea* spp. *Sci. Hortic.* **2015**, *194*, 181–187. [CrossRef]
22. Martin, M.A.; Alvarez, J.B.; Mattioni, C.; Cherubini, M.; Villani, F.; Martin, L.M. Identification and Characterisation of Traditional Chestnut Varieties of Southern Spain Using Morphological and Simple Sequence Repeat (SSRs) Markers. *Ann. Appl. Biol.* **2009**, *154*, 389–398. [CrossRef]
23. Rasoarahona, R.; Wattanadilokchatkun, P.; Panthum, T.; Thong, T.; Singchat, W.; Ahmad, S.F.; Chaiyes, A.; Han, K.; Kraichak, E.; Muangmai, N.; et al. Optimizing Microsatellite Marker Panels for Genetic Diversity and Population Genetic Studies: An Ant Colony Algorithm Approach with Polymorphic Information Content. *Biology* **2023**, *12*, 1280. [CrossRef]
24. Bai, X.; Zhang, S.; Wang, W.; Chen, Y.; Zhao, Y.; Shi, F.; Zhu, C. Genetic Relationships of 118 *Castanea* Specific Germplasms and Construction of Their Molecular ID Based on Morphological Characteristics and SSR Markers. *Plants* **2023**, *12*, 1438. [CrossRef] [PubMed]
25. Alessandri, S.; Ramos Cabrer, A.M.; Martìn, M.A.; Mattioni, C.; Pereira-Lorenzo, S.; Dondini, L. Genetic Characterization of Italian and Spanish Wild and Domesticated Chestnut. *Sci. Hortic.* **2022**, *295*, 110882. [CrossRef]
26. Tumpa, K.; Šatović, Z.; Liber, Z.; Vidaković, A.; Idžojtić, M.; Ježić, M.; Ćurković-Perica, M.; Poljak, I. Gene Flow between Wild Trees and Cultivated Varieties Shapes the Genetic Structure of Sweet Chestnut (*Castanea sativa* Mill.) Populations. *Sci. Rep.* **2022**, *12*, 15007. [CrossRef]
27. Nishio, S.; Yamamoto, T.; Terakami, S.; Sawamura, Y.; Takada, N.; Saito, T. Genetic Diversity of Japanese Chestnut Cultivars Assessed by SSR Markers. *Breed. Sci.* **2011**, *61*, 109–120. [CrossRef]
28. Zhao, M.; Shu, G.; Hu, Y.; Cao, G.; Wang, Y. Pattern and Variation in Simple Sequence Repeat (SSR) at Different Genomic Regions and Its Implications to Maize Evolution and Breeding. *BMC Genom.* **2023**, *24*, 136. [CrossRef] [PubMed]
29. Divakar, S.; Jha, R.K.; Kamat, D.N.; Singh, A. Validation of Candidate Gene-Based EST-SSR Markers for Sugar Yield in Sugarcane. *Front. Plant Sci.* **2023**, *14*, 1273740. [CrossRef] [PubMed]
30. Yadav, S.; Carvalho, J.; Trujillo, I.; Prado, M. Microsatellite Markers in Olives (*Olea europaea* L.): Utility in the Cataloging of Germplasm, Food Authenticity and Traceability Studies. *Foods* **2021**, *10*, 1907. [CrossRef]
31. Arbeiter, A.B.; Hladnik, M.; Jakše, J.; Bandelj, D. Identification and Validation of Novel EST-SSR Markers in Olives. *Sci. Agric.* **2017**, *74*, 215–225. [CrossRef]
32. Thunder, R.; Conrad, A.O.; Burdine, C.; Yang, J.; Lhotka, J.M.; Abbott, A.G.; Nelson, D. *White Oak (Quercus alba L.) Microsatellite Markers for Genetic Diversity Studies*; U.S. Department of Forest Service: Washington, DC, USA, 2022; Volume 12.
33. Marinoni, D.; Akkak, A.; Bounous, G.; Edwards, K.J.; Botta, R. Development and Characterization of Microsatellite Markers in *Castanea sativa* (Mill.). *Mol. Breed.* **2003**, *11*, 127–136. [CrossRef]
34. Botta, R.; Marinoni, D.; Beccaro, G.; Akkak, A.; Bounous, G. Development of a DNA Typing Technique for the Genetic Certification of Chestnut Cultivars. *For. Snow Landsc.* **2001**, *76*, 425–428.

35. Beier, S.; Thiel, T.; Münch, T.; Scholz, U.; Mascher, M. MISA-Web: A Web Server for Microsatellite Prediction. *Bioinformatics* **2017**, *33*, 2583–2585. [CrossRef]
36. Tamura, K.; Nei, M. Estimation of the Number of Nucleotide Substitutions in the Control Region of Mitochondrial DNA in Humans and Chimpanzees. *Mol. Biol. Evol.* **1993**, *10*, 512–526. [PubMed]
37. Yang, H.C.; Lin, H.C.; Huang, M.C.; Li, L.H.; Pan, W.H.; Wu, J.Y.; Chen, Y.T. A New Analysis Tool for Individual-Level Allele Frequency for Genomic Studies. *BMC Genom.* **2010**, *11*, 415. [CrossRef] [PubMed]
38. Dinis, L.-T.J.; Peixoto, F.; Costa, R.; Gomes-Laranjo, J. Molecular Characterization of "Judia" (*Castanea sativa* Mill.) from Several Trás-Os-Montes Regions by Nuclear Microsatellite Markers. *Acta Hort.* **2010**, *866*, 225–232. [CrossRef]
39. Aldrich, P.; Michler, C.H.; Romero-Severson, J. Pedigree Analysis of Maize Inbreds. *New For.* **2003**, *26*, 43–49.
40. Barreneche, T.; Casasoli, M.; Russell, K.; Akkak, A.; Meddour, H.; Plomion, C.; Villani, F.; Kremer, A. Comparative Mapping between *Quercus* and *Castanea* Using Simple-Sequence Repeats (SSRs). *Theor. Appl. Genet.* **2004**, *108*, 558–566. [CrossRef] [PubMed]