*Article*

# Development of Retrotransposon-Based Molecular Markers for Characterization of *Persea americana* (Avocado) Cultivars and Horticultural Races

**Mario González Carracedo [1,2], Samuel Bello Alonso [1], Rahil Salomé Brito Cabrera [1], David Jiménez-Arias [3] and José Antonio Pérez Pérez [1,2,*]**

1. Department of Biochemistry, Microbiology, Cell Biology and Genetics, University of La Laguna, 38200 Canary Islands, Spain; mgonzalc@ull.edu.es (M.G.C.); samuelbelloalonso@gmail.com (S.B.A.); rahilsalome-10@hotmail.com (R.S.B.C.)
2. Institute of Tropical Diseases and Public Health of the Canary Islands, University of La Laguna, 38206 Canary Islands, Spain
3. ISOPlexis—Center for Sustainable Agriculture and Food Technology, Madeira University, 9020 Madeira, Portugal; david.j.a1983@gmail.com

* Correspondence: joanpere@ull.edu.es; Tel.: +34-92231-6502 (ext. 8678)

**Abstract:** *Persea americana* (avocado) represents one of the most demanded food products worldwide, with an important impact in several agronomy-based economies. The avocado is one of the most salt-sensitive and valuable crops. It is therefore necessary to use salt-tolerant varieties, such as the West Indian, for cultivation in locations with soil salinity problems, such as the Canary Islands. Therefore, characterization of avocado cultivars is in demand, as well as development of molecular tools able to easily identify the main avocado cultivars and horticultural races. In the present work, inter-Primer Binding Site (iPBS) and Inter-Retrotransposon Amplified Polymorphism (IRAP) techniques, which are based on retrotransposon with Long Terminal Repeats (LTR), have been implemented for the first time in *P. americana*, allowing the characterization of genetic variation among cultivars from the three main horticultural races and the identification of potential *P. americana* LTR sequences. The iPBS approach showed clear advantages over its technical implementation, and allowed a better delimitation of horticultural races, especially when focused on West Indian cultivars. However, both techniques generated reproducible genetic fingerprints that not only allowed genetic characterization of each cultivar analyzed, but also revealed potential molecular markers for the identification of avocado cultivars and horticultural races.

**Keywords:** *P. americana*; avocado; LTR-retrotransposon; iPBS; IRAP; molecular markers

**Citation:** Carracedo, M.G.; Alonso, S.B.; Cabrera, R.S.B.; Jiménez-Arias, D.; Pérez Pérez, J.A. Development of Retrotransposon-Based Molecular Markers for Characterization of *Persea americana* (Avocado) Cultivars and Horticultural Races. *Agronomy* **2022**, *12*, 1510. https://doi.org/10.3390/agronomy12071510

Academic Editors: Miguel A. A. Pinheiro De Carvalho, Jan Slaski, Carla Gouveia and Carla Ragonezi

Received: 18 May 2022
Accepted: 22 June 2022
Published: 23 June 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

## 1. Introduction

Avocado (*Persea americana* L.) is one of the most economically important species within the *Lauraceae* family [1]. Its origin has been established in Central America, as avocado seeds found in Mexican excavations have been dated to 7000 B.C. [2]. Local selection of improved genotypes, and their subsequent fixation by vegetative propagation, have allowed the development of hundreds of cultivars, but only a selection is currently being agronomically exploited. Indeed, about 90% of worldwide production relies on the "Hass" cultivar, which originated decades ago at the California University [3]. Classically, avocado cultivars have been classified into three horticultural races, *P. americana* var. drymifolia (Mexican), *P. americana* var. guatemalensis (Guatemalan) and *P. americana* var. americana (Antillean or West Indian), according to morpho-physiological features of trees and fruits [2,4]. However, reproductive biology of *P. americana* is mainly asynchronous (i.e., female flowers bloom first then male ones), which usually favors cross-fertilization, causing a wide genetic diversity in the avocado progeny, as well as the continuous production of new hybrid cultivars [2].

Since 1961, efforts to meet avocado global demand have led to about 23% increase in worldwide production, reaching more than 25 million tons per year in 2018 [5]. The market value of avocado-derived products has also progressively increased [5], not only because of its excellent nutritional properties [6], but also as a consequence of mark-up product development by food, oil, and cosmetic industries [7,8]. Moreover, the anti-inflammatory and analgesic properties of phytochemical compounds found in avocado fruits are being exploited by the pharmaceutical industry [9]. As a consequence, global consumption of this product cannot be actually supplied by the agronomic industry and, therefore, the avocado is currently considered a crop with an excellent profitability. Spain represents the unique European country with significant avocado production, (11,000 ha in 2017) [10], being mainly produced in the south of the Iberian Peninsula, but also in the Canary Islands, in which *P. americana* represents a promising crop to increase the economic impact of the local agronomy-based economy [11,12]. Both the area dedicated to avocado cultivation and its production in the Canary Islands have doubled since 2012, reaching 1965.4 ha and 13,293 tons of fruits produced in 2020 [13]. However, agricultural soils on the islands show a high degree of degradation, especially due to salinization of irrigated soils, which reaches an average of 57% on the islands [14]. Avocado is one of the most salt-sensitive crops [15], and it is noteworthy that the physiological response of 'Hass' avocado to salinity is influenced by the rootstock [16]. In this sense, the West Indian rootstock is able to grow in saline environments [15] and it is resistant to the *Phytophthora cinnamomi* phytopathogen [17], which are the two main reasons why the local administration recommends it for new exploitation [18].

In the same way that the avocado market has been expanded, so interest in development of new molecular markers has grown during recent decades, especially to unequivocally characterize the best cultivars to improve yield, but also to identify cultivars adapted to specific geoclimatic conditions and for development of molecular-assisted breeding programs [2]. In this sense, several publications have addressed the identification of molecular markers in *P. americana*, involving different classic methods such as Restriction Fragment Length Polymorphism (RFLP) [19,20], Amplified Fragment Length Polymorphism (AFLP) [21–24], Random Amplified Polymorphic DNA (RAPD) [25], Single Sequence Repeats (SSRs) [26–34] or Single Nucleotide Polymorphisms (SNPs) [35–40].

Surprisingly, the application of transposable elements for development of new *P. americana* molecular markers has not been exploited so far, as in the case for other agronomically-important plant species [41,42]. Different strategies make use of transposable elements to generate DNA fingerprints, such as Sequence-Specific Amplification Polymorphisms (S-SAP) [43], Inter-Retrotransposon Amplified Polymorphism (IRAP), or Retrotransposon-Microsatellite Amplified Polymorphism (REMAP) [44–47]. However, previous knowledge of nucleotide sequences from target plant species is necessary to apply these strategies. Fortunately, this problem has been successfully solved after development of the so-called inter-Primer Binding Site (iPBS) technique [48,49]. Most of plant transposable elements belong to Class-I retrotransposons, which usually contain Long Terminal Repeats (LTRs) as flanking sequences [50]. These elements show a "copy-paste" transposition mechanism that involves its transcription to an RNA intermediate, which is then reverse transcribed to cDNA and inserted at the target genomic location [51,52]. This transposition mechanism requires the use of host cell tRNAs as primers, which recognize a Primer Binding Site (PBS) sequence placed near to the 5′ LTR of the retrotransposon, to initiate the reverse transcription step [53]. These PBS sequences are usually conserved among species and have been used to design the iPBS nearly-universal primers, which allows a single-primer amplification of DNA fragments placed between two inverted LTR-retrotransposons [48,49].

In the present work, molecular tools based on LTR-retrotransposons (iPBS and IRAP) have been implemented for the first time in *P. americana*. Genetic diversity among 12 avocado cultivars has been evaluated, and phylogenetic relationships were reconstructed in order to compare results obtained by these two techniques.

## 2. Materials and Methods

### 2.1. Plant Samples and DNA Purification

Well-characterized *Persea americana* cultivars were retrieved from *Instituto Canario de Investigaciones Agrarias* (ICIA), as well as from several private collections located in Tenerife (Canary Islands, Spain) (Table 1).

**Table 1.** *Persea americana* cultivars included in the analyses.

| Race [1] | Source | Cultivar |
|---|---|---|
| G | PE | Reed |
| M | ICIA | Thomas |
| W | PE | SS3 |
| GxM | PE | Orotava<br>Fuerte<br>Hass<br>Pinkerton<br>Bacon<br>Zutano |
| | ICIA | Lamb-Hass |
| GxW | ICIA | Choquette<br>Julián |

[1] W (West Indian); G (Guatemalan); M (Mexican); GxW (Guatemalan x West Indian hybrids); GxM (Guatemalan x Mexican hybrids); PE (Private Explotation); ICIA (*Instituto Canario de Investigaciones Agrarias*).

Young leaves were collected from adult trees, without symptoms of disease, chlorosis or wounds. Genomic DNA (gDNA) was purified from 0.1 g of fresh plant material, avoiding petioles and main nerves, with the E.Z.N.A. SP Plant DNA kit (Omega BIO-TEK, Norcross, GA, USA) following manufacturer's instructions. As the first step, leaf samples were homogenized in 2 mL Lysing Matrix-A tubes (M.P. Biomedicals, Irvine, CA, USA) by vigorous shaking twice at 5 m/s for 30 s, in a FastPrep-24 system (M.P. Biomedicals, Irvine, CA, USA), the second time being in the presence of the lysis buffer from the kit. DNA concentration and purity were determined with a DeNovix DS-11 spectrophotometer (Denovix, Wilmington, DE, USA), considering ranges of 1.7–1.9 and 1.8–2.0 for 260/280 and 260/230 absorbance ratios, respectively, as adequate purity references. Each DNA sample was diluted to a final concentration of 10 ng/µL in 10 mM Tris-HCl pH 8.0, and stored at −20 °C. From these stocks, different working dilutions were prepared in the same buffer, as indicated.

### 2.2. The iPBS Analysis

Implementation of the iPBS strategy was essentially carried out following recommendations of Kalendar et al. [46], making use of a subset of 9 PBS primers (Table 2) [49]. After optimization, PCRs were carried out in a final volume of 20 µL, containing 2 ng of gDNA, 1X Phire HotStar II Reaction Buffer; (ThermoFisher Scientific, Bedford, MA, USA), 0.2 mM each dNTP (VWR, Radnor, PA, USA), 1 µM of one PBS primer, 0.2 µL of Phire HotStar II DNA polymerase, 0.5 µg/µL BSA (VWR, Radnor, PA, USA), and a supplement of $MgCl_2$ (0.5 mM). A ProFlex PCR System (Applied Biosystems, Waltham, MA, USA) was used for incubation of amplification reactions, including an initial denaturation step (98 °C for 30 s), 30 amplification cycles (98 °C for 10 s; annealing temperature described in Table 2, for 30 s; 72 °C for 40 s), and a final extension step (72 °C for 2 min).

**Table 2.** Oligonucleotides tested for iPBS analysis in *P. americana*.

| Primer ID [1] | Sequence (5′-3′) | Ta (°C) [2] |
|---|---|---|
| PBS2228 | CATTGGCTCTTGATACCA | 54.0 |
| PBS2232 | AGAGAGGCTCGGATACCA | 55.4 |
| PBS2237 | CCCCTACCTGGCGTGCCA | 55.0 |
| PBS2239 | ACCTAGGCTCGGATGCCA | 55.0 |
| PBS2242 | GCCCCATGGTGGGCGCCA | 57.0 |
| PBS2251 | GAACAGGCGATGATACCA | 53.2 |
| PBS2373 | GAACTTGCTCCGATGCCA | 51.0 |
| PBS2395 | TCCCCAGCGGAGTCGCCA | 52.8 |
| PBS2415 | CATCGTAGGTGGGCGCCA | 61.0 |

[1] Code assigned to each primer by Kalendar et al. [49]. [2] Annealing temperature recommended by Kalendar et al. [49], which was used in the present work.

PCR products (10 µL) were fractionated by agarose gel electrophoresis in 1X TBE buffer under two different conditions. When target amplicons were in the range of 100–800 bp, they were resolved in 2% agarose gels (10-cm length) at 60 V for 4 h, while larger PCR products (up to 2.5 Kb) were better separated in 1.7% agarose gels (20-cm length) at 120 V for 10 h. The 100 bp DNA Step Ladder (Promega, Madison, WI, USA), was used as the molecular weight marker. Gels were immersed in 1X GelRed (Biotium, Fremont, CA, USA) for 1.5–2 h, and exposed to ultraviolet light in a ChemiDoc XRS+ (BioRad, Hercules, CA, USA) to visualize DNA fragments.

### 2.3. Cloning and Sequencing of iPBS Fragments

Several polymorphic iPBS bands were collected from agarose gels with sterile scalpels. Amplicons with identical lengths from different individuals were pooled, before purification of DNA fragments with the E.Z.N.A. MicroElute Gel Extraction kit (Omega BIO-TEK, Norcross, GA, USA), following manufacturer's instructions. Concentration and purity of DNA preparations were determined spectrophotometrically, as explained before.

Purified iPBS amplicons were cloned into pJET1.2/blunt vector making use of the CloneJet PCR Cloning kit (Thermo Scientific, Bedford, MA, USA), following manufacturer's recommendations. *E. coli* TOP10 cells were transformed with the ligation mixtures following a $CaCl_2$/heat shock transformation protocol. Transformant colonies were screened by PCR to determine insert lengths using the amplification primers supplied by the kit. Recombinant plasmids bearing the targeted iPBS amplicons, were purified with the E.Z.N.A. Plasmid DNA Mini kit (Omega BIO-TEK, Norcross, GA, USA) following manufacturer's indications. Purified plasmids were spectrophotometrically quantified as explained above, and prepared for Sanger sequencing.

### 2.4. Identification of Potential LTRs and IRAP Analysis

To generate a unique consensus sequence for each iPBS fragment, plasmids from at least three positive *E. coli* clones were sequenced. In an attempt to discard locus-specific sequences and to identify all potential LTRs, at this step the first 200 nucleotides at the 5′ end of both forward and reverse DNA chains from all sequenced iPBS amplicons were aligned with ClustalW in MEGAX software [54]. This multiple alignment was used to construct a consensus UPGMA tree from 10,000 bootstrap replicates [55]. Tree branches reproduced in less than 80% of replicates were collapsed, and the remaining clusters were the start point for the next analysis step. The complete iPBS sequences belonging to a certain cluster were retrieved and compared in search of potential LTRs following recommendations of Kalendar et al. [49], which essentially involved the identification of a conserved region at the 5′ end of iPBS sequences, which should start by 5′-TG dinucleotide (as far as 5 residues from the 3′ end of the PBS primer biding site), and finish by CA-3′. This conserved region (potential LTR) was presumed to be followed by a less conserved locus-specific region, as a consequence of LTR-retrotransposon integration at different genome loci (Figure S1).

Potential LTRs were then used to design primers for implementation of IRAP technique. Oligonucleotide sequences were designed with PRIMER3 application included in Gene Runner software [56], maintaining their melting temperature (Tm) at about 65 °C. The IRAP technique was implemented by making use of eight different single primers, which recognized potential avocado LTRs. Preparation of PCR reactions, amplification profiles and electrophoresis conditions were exactly the same as described for the iPBS analysis, with the exception of annealing temperatures (Ta), which were experimentally optimized for each oligonucleotide tested.

### 2.5. Phylogenetic Inferences

Band patterns obtained from IRAP or iPBS analysis were carefully inspected, and converted into binary markers (presence of a band was coded as 1, while its absence as 0). The resulting data matrices were imported into DARwin software [57], which was used to estimate genetic distances between avocado cultivars, applying the Jaccard similarity index. From the distance matrix, phylogenetic trees were obtained using the weighted Neighbor-Joining algorithm contained in the same software, and robustness of each tree node was assessed by 10,000 bootstraps replicates, and a consensus tree was generated, with tree branches reproduced in less than 50% of replicates collapsed.

## 3. Results

### 3.1. Generation of P. americana gDNA Collection

Leaves from 12 different *P. americana* adult trees were collected to generate a gDNA collection that contained the three main avocado horticultural races (West Indian, Guatemalan and Mexican), as well as different hybrid cultivars of agronomic interest (Table 1). DNA concentrations were in the range of 24.1–107.7 ng/µL, with an average of 60.7 ± 32.5 ng/µL. Purity values obtained for A260/A280 ratio varied in the range of 1.38–1.86 (mean ± SD: 1.60 ± 14.0), while the A260/A230 ratio were in the range 0.55–1.55 (mean ± SD: 1.04 ± 0.36). Therefore, DNA purity was suboptimal in some cases, especially for A260/A230, probably caused by exposition of avocado trees to environmental stress conditions that induced production of secondary metabolites [58]. It is known that these compounds are difficult to eliminate during DNA purification, and could act as inhibitors of DNA polymerases [59,60]. Given this, it was necessary to carry out a validation of the iPBS method prior to its use for avocado DNA samples analysis.

### 3.2. iPBS Implementation in P. americana

As a first step in iPBS optimization, three different amounts of gDNA purified from *P. americana* cv. Hass were tested in triplicate for PCR amplification with PBS2251 primer (Table 2). Results showed that 2 ng of gDNA in 20 µL of amplification reaction yielded a higher number of well-defined and repetitive bands (Figure 1a). Moreover, three different gDNA samples, with A260/A280 absorbance ratios between 1.3 and 1.7, were also analyzed in triplicate (Figure 1b). In this case, quality of the observed band patterns did not correlate with possible contaminations revealed by A260/A280 ratios (<1.7), as well-defined and repetitive band patterns were obtained from the three DNA samples.

Interestingly, when the rest of the gDNA samples were tested with PBS2251 primer, several of them failed to amplify, which could have been caused by the presence of contaminants that inhibited DNA polymerase [59,60]. This hypothesis was supported by the existence of low A260/A230 ratios in several samples. To solve this problem, the addition of more $MgCl_2$ (2.0 mM final concentration), BSA (0.5 µg/µL), or both supplements simultaneously to PCR reactions was assayed, using two different DNA templates (SS3 and Bacon) that failed to amplify under initial PCR conditions. Results showed that increasing the $MgCl_2$ concentration did not produce a considerable effect in the amplification pattern, while BSA addition caused a very significant improvement, generating band patterns with good definition. However, increasing $MgCl_2$ concentration in the presence of BSA pro-

duced more intense bands without loss of resolution, and slightly increased repeatability (Figure 1c).

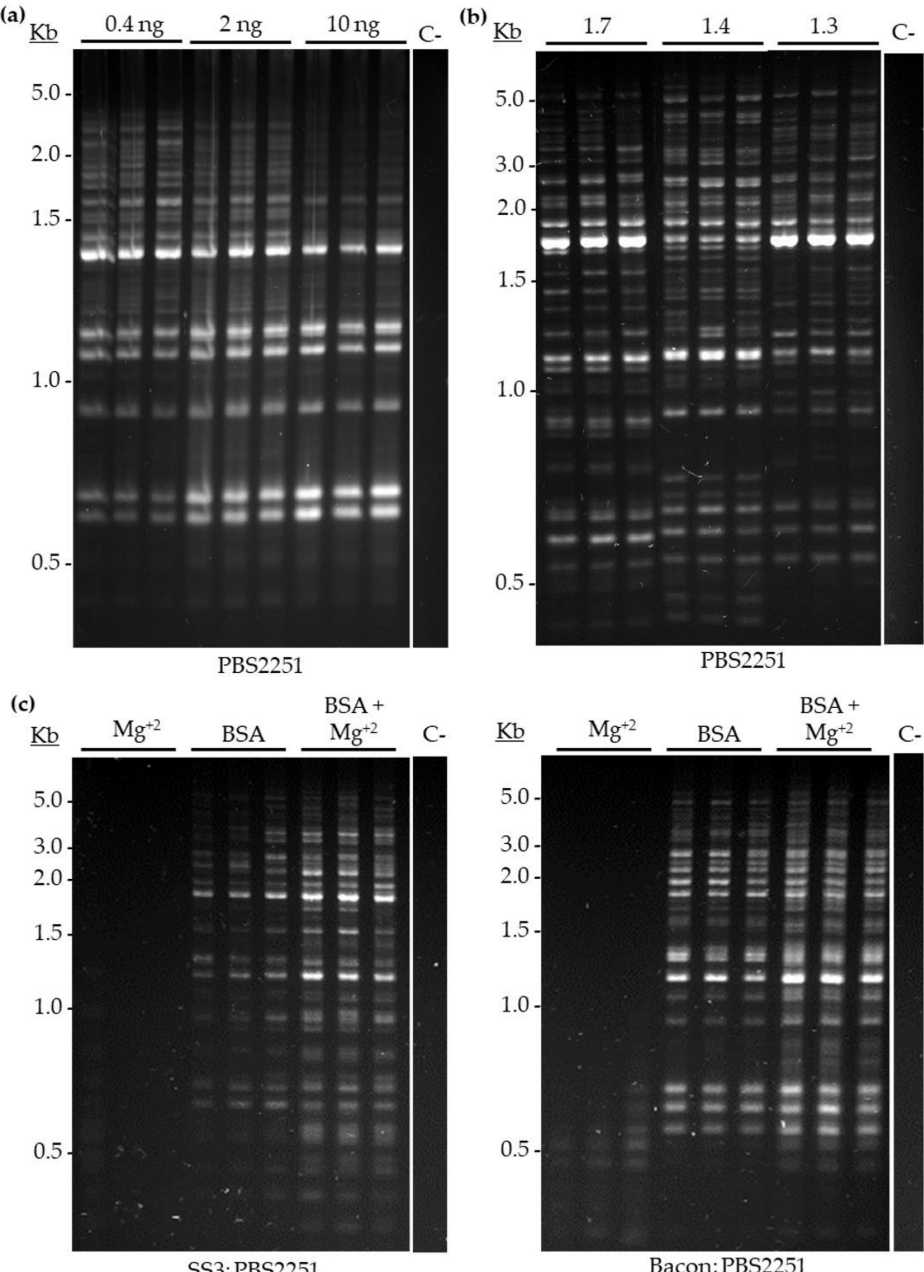


**Figure 1. Optimization of the iPBS method for analysis of avocado germplasm.** (**a**) Indicated amounts of template DNA from *P. americana* cv. Hass were amplified in triplicate with the PBS2251 primer. (**b**) Template DNA (2.0 ng) from three different cultivars, with the indicated A260/A280 absorbance coefficients, were amplified in triplicate with PBS2251 primer. (**c**) Template DNA (2.0 ng) from *P. americana* cv. SS3 (left) and Bacon (right), which failed to amplify under initial PCR conditions, were tested with PBS2251 primer supplementing PCR reactions with 0.5 mM $MgCl_2$, 0.5 µg/µL Bovine Serum Albumin (BSA) or both additives (BSA + $Mg^{+2}$). Kb (Kilobases); C- (negative control).

Consequently, the nine PBS primers (Table 2) were tested for amplification of gDNA purified from Hass and Fuerte cultivars, under optimized PCR conditions (Figure 2). From the nine primers tested, PBS2232, PBS2239 and PBS2251 were selected for subsequent experiments, since they produced clear band patterns with high repeatability. Six primers were excluded, since high amounts of smear were detected (PBS2242 and PBS2395), and

they showed low repeatability, even after PCR optimization (PBS2373), or were unable to amplify under any condition tested (PBS2228, PBS2237 and PBS2415).

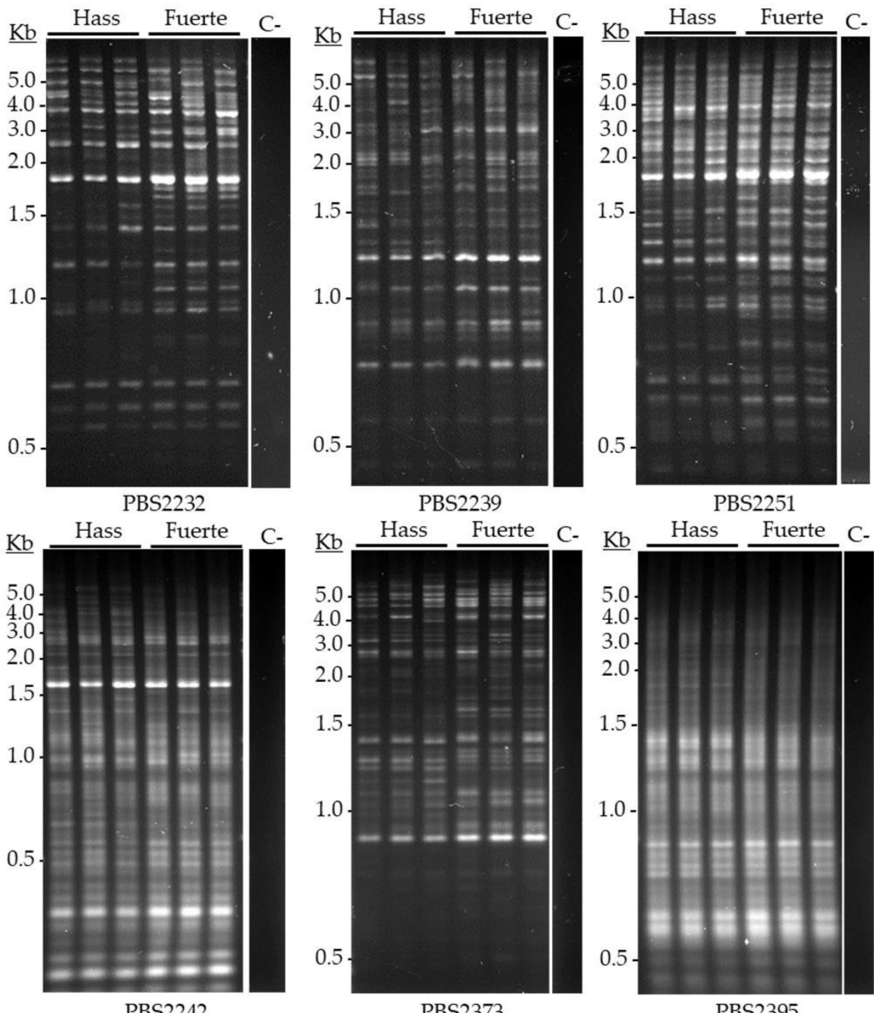


**Figure 2. Evaluation of PBS primers under optimized PCR conditions.** Template DNA (2.0 ng) from *P. americana* cv. Hass and Fuerte were amplified with the indicated PBS primers, supplementing PCR reactions with 0.5 mM $MgCl_2$ and 0.5 µg/µL BSA. Kb (Kilobases); C- (negative control).

The three selected PBS primers were then used for PCR amplification with the set of 12 gDNA samples, which included representatives of the three pure avocado horticultural races, as well as different hybrid cultivars. In all cases, well defined band patterns were obtained (Figure 3a), which were transformed into binary files and analyzed to identify polymorphic and cultivar-specific bands. The number of scored alleles ranged from 23 (PBS2239) and 47 (PBS2232), with an average of 33.7 $\pm$ 12.2 bands per primer. Total number of scored alleles reached 101 with the three primers, being 16.8% monomorphic (present in all cultivars) and 67.3% polymorphic (Table 3). Interestingly, 15.8% of scored alleles were found to be cultivar-exclusive, thus representing potential diagnostic markers.

Phylogenetic analysis of iPBS data allowed differentiation of three clades, supported by bootstrap values higher than 50%. Each clade contained a pure horticultural race, as well as several related hybrids (Figure 3b). The first clade contained the unique Guatemalan representative studied in the present work (*P. americana* cv. Reed), as well as GxM hybrids Orotava, Lamb-Hass and Hass. The two latter cultivars are closely related. The second clade included the Mexican purebred Thomas, closely related with the GxM hybrid Bacon and, to a lesser extent, with another GxM hybrid, the Fuerte cultivar. This clade also

contained another two GxM hybrids (Pinkerton and Zutano), which were found to be more related. Interestingly, GxM hybrids seemed to form two different groups, one of them more related with the Reed cultivar, and the other with Thomas. These could be the consequence of different amounts of their genomes represented by Guatemalan and Mexican breeds. Finally, the third clade contained the SS3 cultivar, as a West Indian purebred representative, and also the GxW hybrids Julian and Choquette. Overall, these results indicated that the iPBS method was able to easily detect genetic variation in *P. americana*, both at cultivar and horticultural race levels.

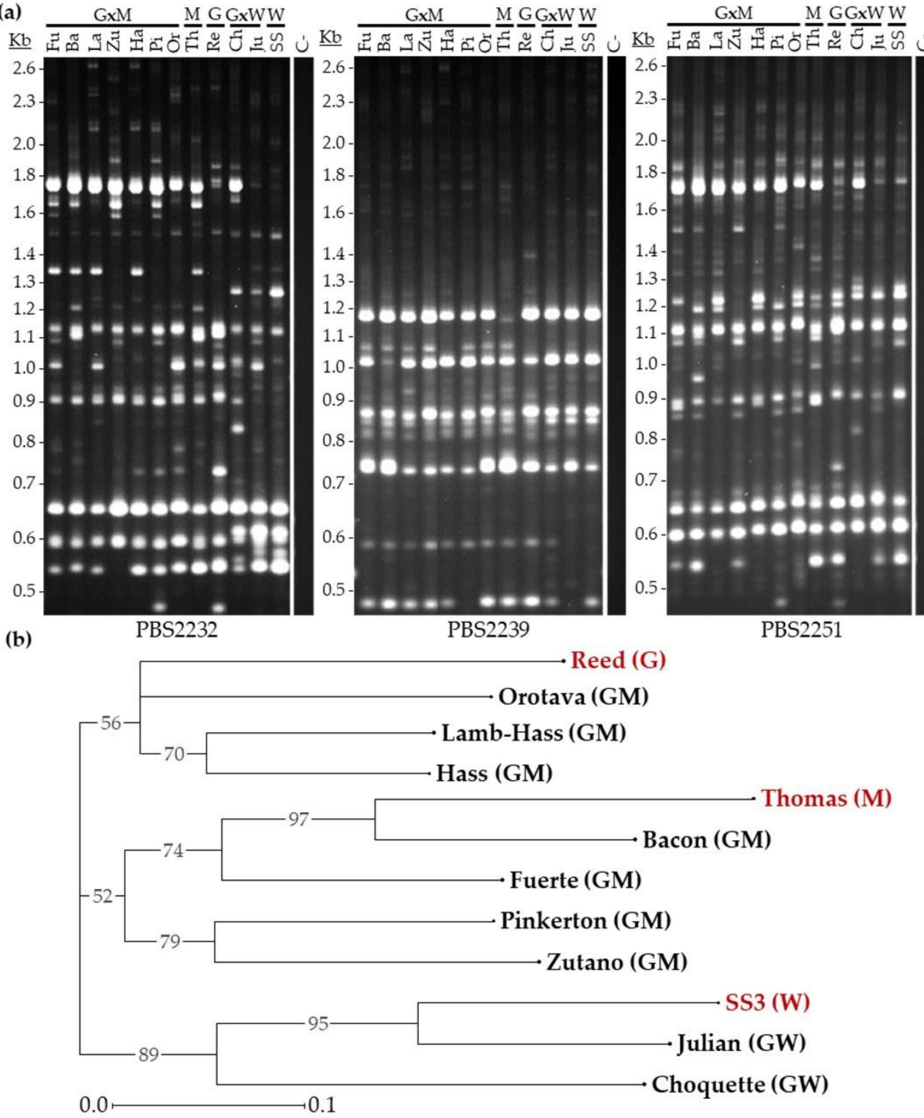


**Figure 3. iPBS analysis of *P. americana* cultivars.** (**a**) Template DNA from *P. americana* cv. Fuerte (Fu), Bacon (Ba), Lamb-Hass (La), Zutano (Zu), Hass (Ha), Pinkerton (Pi), Thomas (Th), Reed (Re), Orotava (Or), Choquette (Ch), Julian (Ju) or SS3 (SS) were amplified with the three selected PBS primers (as indicated below each gel image). Each experiment was repeated three times with identical results. (**b**) Phylogenetic tree generated from concatenated binary data obtained from iPBS gels showed in (**a**). Nodes supported by less than 50% bootstrap values were collapsed. W (West Indian); G (Guatemalan); M (Mexican); GxW (Guatemalan x West Indian hybrid); GxM (Guatemalan x Mexican hybrid); Kb (Kilobases); C- (negative control).

**Table 3.** Number of scored iPBS alleles (bands) from 12 avocado cultivars.

| Race | Cultivar | PBS2232 | PBS2239 | PBS2251 |
|---|---|---|---|---|
| M | Thomas | 21 | 14 | 16 |
| G | Reed | 21 | 13 | 17 |
| W | SS3 | 22 | 10 | 10 |
| GxM | Fuerte | 22 | 12 | 18 |
| | Bacon | 19 | 14 | 16 |
| | Lamb-Hass | 20 | 13 | 13 |
| | Zutano | 18 | 12 | 14 |
| | Hass | 18 | 11 | 15 |
| | Pinkerton | 20 | 11 | 16 |
| | Orotava | 20 | 12 | 15 |
| GxW | Choquette | 19 | 12 | 13 |
| | Julián | 24 | 10 | 13 |
| **Alleles** | | **47** | **23** | **31** |
| **Monomorphic alleles** | | **7 (14.9)** | **5 (21.7)** | **5 (16.1)** |
| **Polymorphic alleles** | | **30 (63.8)** | **18 (78.3)** | **20 (64.5)** |
| **Cultivar-specific alleles [1]** | | **10 (21.3)** | **0 (0.0)** | **6 (19.4)** |

Percentages are shown in brackets; [1] No. of alleles (percentage) absent/present in only one cultivar.

### 3.3. Potential LTR Identification

For identification of potential LTRs, 33 polymorphic bands were extracted from PBS2232 and PBS2251 gels, with amplicon sizes between 430 and 1380 bp. Fragments were cloned into *E. coli* plasmids and sequenced. The initial dataset included 123 sequences, from which a curated set of 26 non-redundant iPBS consensus sequences was generated. To search for potential LTRs, both direct and reverse iPBS sequences were trimmed to 200 bp from their 5′ end, and a multiple alignment of 52 sequences was obtained. The UPGMA cladogram (Figure S2) revealed the presence of 12 different LTR clusters, from which it was possible to define eight potential LTRs, after filtering, following recommendations from Kalendar et al. [49] (Table 4). However, potential LTRs were not clearly identified from LTR clusters 5, 8, 9 and 10, as the expected conserved region flanked by 5′-TG and CA-3′ dinucleotides, was not found to be as expected.

Lengths of potential LTRs were variable, in the range of 146–313 bp, while identity percentages between LTRs of the same cluster were in the range 66.0–97.1%. Analysis of locus-specific sequences from each cluster revealed that potential LTRs were always inserted in at least two different genome positions, revealing four different retrotransposon insertion events (cluster 11). Finally, potential LTR sequences (Dataset S1) were used to design primers for implementation of the Inter-Retrotransposon Amplification Polymorphism (IRAP) technique in *P. americana* (Table 5).

**Table 4.** Features of Potential LTRs.

| LTR Cluster | iPBS Sequences | Potential LTRs | | |
|---|---|---|---|---|
| | | Length (bp) | Identity (%) | Gap (%) |
| 1 | PBS2232.01<br>PBS2251.08 | 278<br>280 | 97.1 | 0.7 |
| 2 | PBS2232.13<br>PBS2251.01 | 297<br>303 | 79.4 | 10.1 |
| 3 | PBS2232.12(-)<br>PBS2251.11(-) | 313<br>307 | 82.7 | 5.0 |
| 4 | PBS2251.06<br>PBS2251.15 | 234 | 66.0 | 17.9 |
| 6 | PBS2232.08<br>PBS2232.08(-) | 225<br>224 | 77.5 | 5.6 |
| 7 | PBS2232.12<br>PBS2232.10(-)<br>PBS2251.06(-) | 146 | 87.2 | 0.0 |
| 11 | PBS2251.14<br>PBS2232.11(-)<br>PBS2232.14(-)<br>PBS2251.08(-) | 181 | 96.4 | 0.0 |
| 12 | PBS2232.05(-)<br>PBS2251.17(-) | 179<br>176 | 83.8 | 1.7 |

**Table 5.** Primers tested for implementation of IRAP in *P. americana*.

| Primer Code | Potential LTR [1] | Primer Sequence (5′-3′) | Tm (°C) |
|---|---|---|---|
| PaIRAP-1 | PBS2251.01 | AGAAAGGAAAACCATCTAATTGTATC | 65.3 |
| PaIRAP-3 | PBS2232.12(-) | CTAGCTGGACTGGATTGATGG | 63.8 |
| PaIRAP-4 | PBS2251.06 | ATTAAATTGGATTGGGGTGTAAC | 64.7 |
| PaIRAP-5 | PBS2251.15 | TTTGGGGCTGGGGTGTAAC | 66.9 |
| PaIRAP-6 | PBS2251.14,<br>PBS2232.11(-)<br>PBS2232.14(-)<br>PBS2251.08(-) | GTAAGGGTGTAAGCTCTACATATAAAC | 63.7 |
| PaIRAP-8 | PBS2232.10(-) | GGGCTCGACCACAATTATGAC | 66.1 |
| PaIRAP-9 | PBS2232.13 | GGGCTTTGGGCCTATTTAAC | 65.6 |
| PaIRAP-10 | PBS2232.01<br>PBS2251.08 | GGGCTTTTGGCCTGTTTAAC | 66.2 |

[1] Table 4 and Dataset S1.

### 3.4. IRAP Implementation in P. americana

In order to implement the IRAP technique for the first time in *P. americana*, the eight IRAP primers designed in the present work (Table 5) were tested in duplicate PCR reactions with gDNAs from *P. americana* cv. Reed, SS3 and Zutano. Starting PCR conditions were set as previously optimized for iPBS analysis, while three different annealing temperatures (57, 61 and 65 °C) were experimentally tested. Three primers (PaIRAP-5, PaIRAP-6 and PaIRAP-8) showed adequate results, considering the number of discrete bands, repeatability of PCR duplicates and the amount of genetic variation detected in trial experiments (Figure 4). The other five primers listed in Table 5 were discarded, since a reduced number of bands, high background noise and/or low repeatability was/were observed. Between the three selected IRAP primers, the most robust results were obtained when the annealing temperature was

set to 57 °C. Taken together, these results validated the use of the iPBS method to identify potential LTRs in *P. americana*, since they confirmed the presence of multiple copies of the potential LTRs.

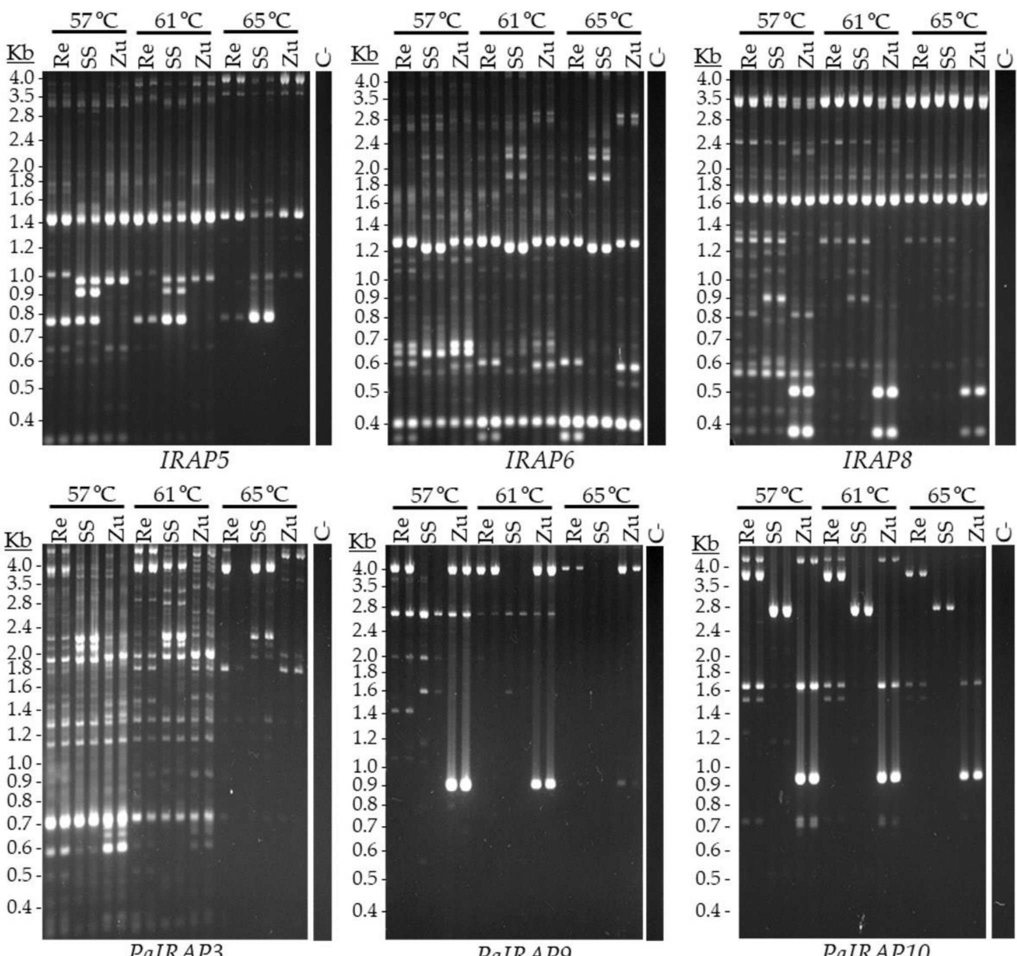


**Figure 4. Evaluation of IRAP primers and PCR conditions.** Template DNA from *P. americana* cv. Reed (Re), SS3 (SS) or Zutano (Zu) were amplified with the indicated IRAP primers by duplicate. Three different annealing temperatures were tested, as indicated. Kb (Kilobases); C- (negative control).

Once PCR conditions were validated, the 12 gDNA samples previously analyzed by the iPBS approach were used for IRAP experiments, in such a way that both techniques required a similar effort in terms of PCR reactions, electrophoresis, and data analysis. As expected, well defined band patterns were obtained for the 12 cultivars analyzed with the three primers (Figure 5a). A summary of the genetic data obtained is shown in Table 6. In this case, the number of scored alleles ranged from 26 (PaIRAP-5) to 43 (PaIRAP-8), with an average of 29.6 ± 4.0 alleles per primer. The total number of scored bands reached 89 with the three primers, being 15.7% monomorphic and 64.0% polymorphic (Table 3). In addition, the IRAP experiments allowed the identification of cultivar-exclusive alleles (20.2%), which represented potential diagnostic markers for the identification of *P. americana* cultivars.

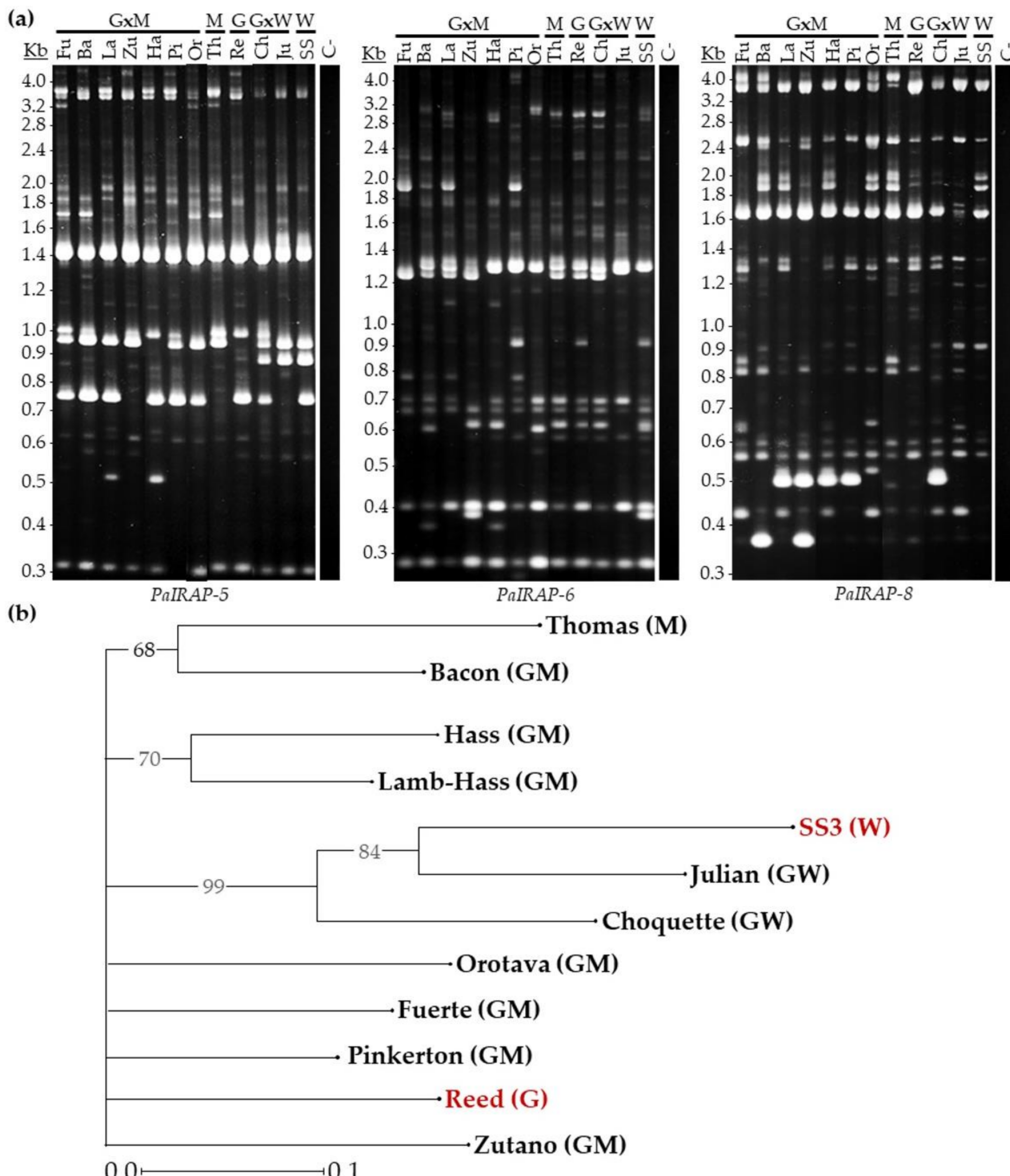


**Figure 5. IRAP analysis of *P. americana* cultivars.** (**a**) Template DNA from *P. americana* cv. Fuerte (Fu), Bacon (Ba), Lamb-Hass (La), Zutano (Zu), Hass (Ha), Pinkerton (Pi), Thomas (Th), Reed (Re), Orotava (Or), Choquette (Ch), Julian (Ju) or SS3 (SS) were amplified with the three selected IRAP primers (as indicated below each gel image), and setting annealing temperature at 57 °C. Each experiment was repeated three times with identical results. (**b**) Phylogenetic tree generated from concatenated binary data obtained from IRAP gels showed in (**a**) Nodes supported by less than 50% bootstrap values were collapsed. W (West Indian); G (Guatemalan); M (Mexican); GxW (Guatemalan x West Indian hybrids); GxM (Guatemalan x Mexican hybrids); Kb (Kilobases); C- (negative control).

Unfortunately, phylogenetic analysis of IRAP results (Figure 5b) did not allow a clear differentiation of the three clades previously observed with iPBS analysis (Figure 3b). In general, bootstrap values were slightly lower, and condensed tree topology was unable to separate the Guatemalan purebred (Reed) and four of the GM hybrids (Fuerte, Pinkerton, Zutano and Orotava) in a monophyletic group. Interestingly, the West Indian clade was extremely well supported (99% bootstrap), thus confirming iPBS results, as SS3 cultivar (W) was closely related with Julian (GxW) and, to a lesser extent, with Choquette (GxW) cultivars. Moreover, the relationship between Hass and Lamb-Hass (GxM hybrids) cultivars was confirmed, and well supported (70% bootstrap). Finally, the Mexican purebred repre-

sentative (Thomas) was confirmed to be closely related with Bacon (GxM), as previously observed in iPBS analysis.

**Table 6.** Number of scored IRAP alleles (bands) from 12 avocado cultivars.

| Race | Cultivar | PaIRAP-5 | PaIRAP-6 | PaIRAP-8 |
|---|---|---|---|---|
| M | Thomas | 8 | 12 | 15 |
| G | Reed | 15 | 13 | 13 |
| W | SS3 | 11 | 10 | 14 |
| GxM | Fuerte | 14 | 15 | 15 |
| | Bacon | 14 | 11 | 16 |
| | Lamb-Hass | 13 | 11 | 14 |
| | Zutano | 10 | 12 | 13 |
| | Hass | 14 | 11 | 16 |
| | Pinkerton | 13 | 11 | 13 |
| | Orotava | 11 | 11 | 14 |
| GxW | Choquette | 13 | 10 | 12 |
| | Julián | 13 | 13 | 16 |
| **Alleles** | | **26** | **29** | **34** |
| **Monomorphic alleles** | | **1 (3.8)** | **4 (13.8)** | **9 (26.5)** |
| **Polymorphic alleles** | | **20 (76.9)** | **15 (51.7)** | **22 (64.7)** |
| **Cultivar-specific alleles** [1] | | **5 (19.2)** | **10 (34.5)** | **3 (8.8)** |

Percentages are shown in brackets. [1] No. of alleles (percentage) absent/present in only one cultivar.

## 4. Discussion

Transposable elements have been identified as one of the main sources of genetic variation in plants [50,61,62]. In land plants, retrotransposons are present as high copy number elements, conforming a substantial part of their genomes [63], and are currently considered as important drivers in evolution [64]. Accordingly, molecular techniques able to detect mobile element-based genetic variation have been validated as useful tools for generation of molecular markers [48]. Unlike the IRAP technique, which has been clearly contrasted in the literature [44–47], the iPBS strategy has been described more recently [49]. As far as we know, only one previous study, based on retrotransposons, has been successfully implemented in *P. americana*. In that work, polymorphism detection and discrimination ability was compared between the retrotransposon-based technique Inverse Sequence Tagged Repeat (ISTR), with SSR and AFLP analysis, concluding that the number of average polymorphic bands obtained by AFLP and ISTR analysis was higher than with SSR [65]. Nevertheless, iPBS has also been used to characterize genetic variation present in several genera of agronomic interest, such as *Prunus* [66], *Vitis* [67,68], *Psidium* [69], *Phoenix* [70], *Nicotiana* [71], *Solanum* [72], *Allium* [73,74], *Gnetum* [75] or *Musa* [76]. The clear advantage of the iPBS strategy is that prior knowledge of nucleotide sequences from target species is not required, since a relatively small set of "universal" primers can be used for analysis of any eukaryotic organism [46,49]. Moreover, co-dominant markers, as microsatellites or SNPs, require allele dosage determination, which is especially difficult for partially heterozygous genotypes in polyploid species. This limitation is usually overcome after transformation of co-dominant genotypes into dominant ones, thus providing essentially the same binary information as iPBS or IRAP. Therefore, another important advantage arises when iPBS and IRAP are implemented for analysis of polyploid species, as dominant multi-locus genotypes are directly obtained, which can be immediately analyzed with most available bioinformatic tools [46,49]. Finally, their set-up simplicity, low cost, and the lower need for scarce laboratory resources, make iPBS and IRAP very attractive tools for population genetic analysis in the field of agronomy.

In the present work, iPBS strategy was applied, for the first time, to analyze genetic variability in *Persea americana* cultivars of agronomic interest, showing adequate repeatabil-

ity for three PBS primers, and revealing a high proportion of polymorphic alleles (63.8 to 78.3%). The haploid genome size for the Hass cultivar ranges between 1.33–1.63 Gb [40] approximately, and draft versions of *P. americana* cv. Drymifolia (Mexican horticultural race) and Hass (GxM hybrid) genomes have been published, confirming the presence of 12 chromosomes [77], but these nucleotide sequences were not considered in the present work because they did not include a reference West Indian genome. Selection and sequencing of iPBS fragments allowed us to characterize twelve different *P. americana* potential LTR clusters, which were then used for design and validation of IRAP primers. As far as we know, the present work also represents the first time in which the IRAP technique has been applied to investigate genetic variability between *P. americana* purebreds and hybrid cultivars. To find out if genetic information obtained through IRAP analysis compensates the effort required for LTR characterization from iPBS fragments, the same set of avocado gDNA samples were analyzed with three iPBS or IRAP primers, in such a way that application of both techniques required similar effort. With respect to the amount of genetic variation detected, the total number of alleles was about 12% higher for iPBS (101 alleles) than for IRAP (89 alleles), while the percentage of polymorphic alleles was similar (67.3% and 64.0% for iPBS and IRAP, respectively).

Moreover, the usefulness of the data generated by both strategies was compared by respective phylogenetic analysis. Briefly, data provided by the iPBS and IRAP techniques allowed clear distinguishing of cultivars with a West Indian component (SS3, Julián and Choquette) from the rest (89% and 99% bootstrap replicates for iPBS and IRAP, respectively). In addition, a clade formed by GxM hybrids Hass and Lamb-Hass, as well as another cluster formed by the GxM cultivar Bacon and the Mexican purebred Thomas, were concordant with both techniques. However, the iPBS technique was better able to define genetic relationships among the rest of the cultivars. Overall, we could conclude that the results provided by the two techniques showed concordant results. However, the IRAP strategy showed less ability to discriminate between the Guatemalan purebred and several GxM hybrids.

It must be considered that the bibliography available about phylogenetic relationships between avocado cultivars is certainly confusing and sometimes contradictory. Therefore, it was difficult to compare our results with other studies but, in general, we found that the iPBS and IRAP phylogenies reached plausible conclusions. For example, SS3 and Julian cultivars, which were found to be mother and descendant (ICIA, personal communication), were grouped into a well-differentiated clade. The same coherence was observed with Choquette and Julian cultivars, both GxW hybrids, which are known to be closely related with the West Indian purebred [35]. However, an important discordance, related to the clade composed by Hass and Lamb-Hass cultivars, was found. This clade showed relatively high bootstrap values by both iPBS and IRAP-based phylogenies, while microsatellite-based phylogeny suggested a higher association between Hass and Pinkerton cultivars [78]. Another surprising result related to Zutano and Bacon cultivars, which are closely related, according to bibliography [36,78]. Nevertheless, Bacon was found to be more related with Thomas by both iPBS and IRAP phylogenies. These discordances could be explained because molecular markers, with different change rates, were analyzed in these studies. In addition, botanical characterization of avocado horticultural races has usually been based on a limited set of phenotypic characters, which were not necessarily related to the large number of molecular markers analyzed. Another explanation could be related to the difficulty in certifying the authenticity of certain hybrid cultivars, due to frequent seed propagation without control of the male parent involved in the fertilization event, in such a way that two cultivars with the same breed denomination could actually be genetically different. This genetic drift would occur even if fertilization processes are controlled, since parent cultivars are usually not genetically pure.

In conclusion, the IRAP results were quite similar to those obtained by iPBS, regarding the amount of genetic variation detected. However, the ability of our IRAP primers to determine phylogenetic differences was found to be lower. In addition, the time-consuming

work that requires LTR identification makes this technique less attractive than iPBS. The clear advantage of iPBS relies on the generation of potential LTR sequences, which could be used in combination with locus-specific primers to design RBIP molecular markers that allow a simple PCR characterization of *P. americana* cultivars and horticultural races, especially in those cultivars with a genetic component from the West Indian race.

**Supplementary Materials:** The following are available online at https://www.mdpi.com/article/10.3390/agronomy12071510/s1. Figure S1: Description of iPBS technique and pipeline for potential LTR identification. Figure S2: Identification of LTR clusters and potential LTRs. Dataset S1: Sequences of potential LTRs.

**Author Contributions:** Conceptualization, J.A.P.P.; methodology, J.A.P.P., M.G.C., S.B.A., R.S.B.C. and D.J.-A.; formal analysis, J.A.P.P. and M.G.C.; data curation, M.G.C.; writing—original draft preparation, M.G.C.; writing—review and editing, J.A.P.P., M.G.C. and D.J.-A.; visualization, J.A.P.P., M.G.C., S.B.A., R.S.B.C. and D.J.-A.; project administration, J.A.P.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. APC was carried out by Universidad de La Laguna (ULL) and Consejeria de Economía, Conocimiento y Empleo (CEI-Canarias). Grant number: SD19/02-2021/0001530.

**Institutional Review Board Statement:** Not applicable. This study does not involve humans or animals.

**Informed Consent Statement:** Not applicable. This study does not involve humans.

**Data Availability Statement:** Data is contained within the article or supplementary material.

**Acknowledgments:** Authors want to specially acknowledge *Pedro Hernández Delgado* from ICIA, and to Dª. Esperanza Hernández de Agroviveros el Rincón S.L., for providing samples of the different avocado cultivars.

**Conflicts of Interest:** The authors declare 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. The Angiosperm Phylogeny Group; Chase, M.W.; Christenhusz, M.J.M.; Fay, M.F.; Byng, J.W.; Judd, W.S.; Soltis, D.E.; Mabberley, D.J.; Sennikov, A.N.; Soltis, P.S.; et al. An Update of the Angiosperm Phylogeny Group Classification for the Orders and Families of Flowering Plants: APG IV. *Bot. J. Linn. Soc.* **2016**, *181*, 1–20. [CrossRef]
2. Lahav, E.; Lavi, U. Avocado Genetics and Breeding. In *Breeding Plantation Tree Crops: Tropical Species*; Jain, S.M., Priyadarshan, P.M., Eds.; Springer: New York, NY, USA, 2009; pp. 247–285. ISBN 978-0-387-71199-7.
3. Schaffer, B.A.; Wolstenholme, B.N.; Whiley, A.W. *The Avocado: Botany, Production and Uses*; CAB International: London, UK, 2013.
4. Bergh, B.; Ellstrand, N. Taxonomy of the Avocado. *Calif. Avocado Soc.* **1986**, *70*, 135–146.
5. Sommaruga, R.; Eldridge, H.M. Avocado Production: Water Footprint and Socio-Economic Implications. *EuroChoices* **2021**, *20*, 48–53. [CrossRef]
6. Fulgoni, V.L.; Dreher, M.; Davenport, A.J. Avocado Consumption Is Associated with Better Diet Quality and Nutrient Intake, and Lower Metabolic Syndrome Risk in US Adults: Results from the National Health and Nutrition Examination Survey (NHANES) 2001–2008. *Nutr. J.* **2013**, *12*, 1. [CrossRef]
7. Araújo, R.G.; Rodriguez-Jasso, R.M.; Ruiz, H.A.; Pintado, M.M.E.; Aguilar, C.N. Avocado By-Products: Nutritional and Functional Properties. *Trends Food Sci. Technol.* **2018**, *80*, 51–60. [CrossRef]
8. Akan, S. Phytochemicals in Avocado Peel and Their Potential Uses. *Food Health* **2021**, *7*, 138–149. [CrossRef]
9. Adeyemi, O.O.; Okpo, S.O.; Ogunti, O.O. Analgesic and Anti-Inflammatory Effects of the Aqueous Extract of Leaves of *Persea americana* Mill (Lauraceae). *Fitoterapia* **2002**, *73*, 375–380. [CrossRef]
10. Moreno-Ortega, G.; Pliego, C.; Sarmiento, D.; Barceló, A.; Martínez-Ferri, E. Yield and Fruit Quality of Avocado Trees under Different Regimes of Water Supply in the Subtropical Coast of Spain. *Agric. Water Manag.* **2019**, *221*, 192–201. [CrossRef]
11. Sauco, G. Avocado Culture in the Canary Islands. *Calif. Avocado Soc.* **1985**, *69*, 73–80.
12. Alonso Gonzalez, J. *Los Cultivos Agrícolas en Canarias: Los Cambios y sus Causas*; Repositorio Institutional Universidad de La Laguna (RIULL): Canary Islands, Spain, 2020.

13. Instituto Canario de Estadística (ISTAC). Gobierno de Canarias Estadísticas de La Comunidad Autónoma de Canarias. Series Anuales de Agricultura. Municipios, Islas y Provincias de Canarias. 1999–2020. Available online: http://www.gobiernodecanarias.org/istac/jaxi-istac/tabla.do (accessed on 18 November 2021).
14. Hernandez-Moreno, J.; Tejedor, M.; Jiménez, C. Effects of Land Use on Soil Degradation and Restoration in the Canary Islands. In *Soils of Volcanic Regions in Europe*; Springer Nature: Berlin/Heidelberg, Germany, 2007; pp. 565–579. [CrossRef]
15. Celis, N.; Suarez, D.L.; Wu, L.; Li, R.; Arpaia, M.L.; Mauk, P. Salt Tolerance and Growth of 13 Avocado Rootstocks Related Best to Chloride Uptake. *HortScience* **2018**, *53*, 1737–1745. [CrossRef]
16. Acosta-Rangel, A.M.; Li, R.; Celis, N.; Suarez, D.L.; Santiago, L.S.; Arpaia, M.L.; Mauk, P.A. The Physiological Response of 'Hass' Avocado to Salinity as Influenced by Rootstock. *Sci. Hortic.* **2019**, *256*, 108629. [CrossRef]
17. Llobet, G.; Zárate, P. Evaluación en campo de patrones clonales de aguacate de raza mexicana y antillana tolerante-resistentes a *Phytophthora cinnamomi* Rands. In Proceedings of the V World Avocado Congress, Málaga, Spain, 19–24 October 2003; pp. 573–578.
18. *Boletín Oficial de Canarias Num. 125*; Orden de 18 de Junio de 2012, Por la que se aprueban las normas técnicas específicas de producción integrada del Aguacate, Mango, Papaya y Piña Tropical en Canarias; Gobierno de Canarias: Canarias, Spain, 2012; pp. 12044–12088.
19. Davis, J.; Henderson, D.; Kobayashi, M.; Clegg, M.T.; Clegg, M.T. Genealogical Relationships among Cultivated Avocado as Revealed through RFLP Analyses. *J. Hered.* **1998**, *89*, 319–323. [CrossRef]
20. Furnier, G.R.; Cummings, M.P.; Clegg, M.T. Evolution of the Avocados as Revealed by DNA Restriction Fragment Variation. *J. Hered.* **1990**, *81*, 183–188. [CrossRef]
21. Cañas-Gutiérrez, G.P.; Galindo-López, L.F.; Arango-Isaza, R.; Saldamando-Benjumea, C.I. Diversidad Genética de Cultivares de Aguacate (*Persea americana*) En Antioquia, Colombia. *Agron. Mesoam.* **2015**, *26*, 129–143. [CrossRef]
22. Farrow, B.J.; Fourie, G.; Van den Berg, N. DNA Profiling of *Persea americana* Using AFLP Markers. *S. Afr. J. Bot.* **2010**, *76*, 410. [CrossRef]
23. Taah, K.J.; Alderson, P.G.; Power, J.B. Molecular Approaches for the Characterization of Ghanaian Avocado Pear (*Persea americana* Mill.) Germplasm. In Proceedings of the V World Avocado Congress (Actas V Congreso Mundial del Aguacate), Málaga, Spain, 19–24 October 2003; pp. 19–24.
24. Douhan, G.W.; Fuller, E.; McKee, B.; Pond, E. Genetic Diversity Analysis of Avocado (*Persea americana* Miller) Rootstocks Selected under Greenhouse Conditions for Tolerance to *Phytophthora* Root Rot Caused by *Phytophthora cinnamomi*. *Euphytica* **2011**, *182*, 209. [CrossRef]
25. Fiedler, J.; Bufler, G.; Bangerth, F. Genetic Relationships of Avocado (*Persea americana* Mill.) Using RAPD Markers. *Euphytica* **1998**, *101*, 249–255. [CrossRef]
26. Cuiris-Pérez, H.; Guillén-Andrade, H.; Pedraza-Santos, M.E.; López-Medina, J.; Vidales-Fernández, I. Genetic Variability within Mexican Race Avocado (*Persea americana* Mill.) Germplasm Collections Determined by ISSRs. *Rev. Chapingo Ser. Hortic.* **2009**, *15*, 169–175. [CrossRef]
27. Borrone, J.W.; Olano, C.T.; Kuhn, D.N.; Brown, J.S.; Schnell, R.J.; Violi, H.A. Outcrossing in Florida Avocados as Measured Using Microsatellite Markers. *J. Am. Soc. Hortic. Sci.* **2008**, *133*, 255–261. [CrossRef]
28. Acheampong, A.K.; Akromah, R.; Ofori, F.A.; Takrama, J.F.; Saada, D.; Bitton, I.; Lavi, U. Genetic Characterization of Ghanaian Avocados Using Microsatellite Markers. *J. Am. Soc. Hortic. Sci.* **2008**, *133*, 801–809.
29. Canas-Gutierrez, G.P.; Arango-Isaza, R.E.; Saldamando-Benjumea, C.I. Microsatellites Revealed Genetic Diversity and Population Structure in Colombian Avocado (*Persea americana* Mill.) Germplasm Collection and Its Natural Populations. *JPBCS* **2019**, *11*, 106–119. [CrossRef]
30. Schnell, R.J.; Brown, J.S.; Olano, C.T.; Power, E.J.; Krol, C.A.; Kuhn, D.N.; Motamayor, J.C. Evaluation of Avocado Germplasm Using Microsatellite Markers. *J. Am. Soc. Hortic. Sci.* **2003**, *128*, 881–889. [CrossRef]
31. Borrone, J.W.; Schnell, R.J.; Violi, H.A.; Ploetz, R.C. Primer note: Seventy Microsatellite Markers from *Persea americana* Miller (Avocado) Expressed Sequence Tags: PRIMER NOTE. *Mol. Ecol. Notes* **2006**, *7*, 439–444. [CrossRef]
32. Ashworth, V.E.T.M.; Kobayashi, M.C.; De La Cruz, M.; Clegg, M.T. Microsatellite Markers in Avocado (*Persea americana* Mill.): Development of Dinucleotide and Trinucleotide Markers. *Sci. Hortic.* **2004**, *101*, 255–267. [CrossRef]
33. Gross-German, E.; Viruel, M.A. Molecular Characterization of Avocado Germplasm with a New Set of SSR and EST-SSR Markers: Genetic Diversity, Population Structure, and Identification of Race-Specific Markers in a Group of Cultivated Genotypes. *Tree Genet. Genomes* **2013**, *9*, 539–555. [CrossRef]
34. Boza, E.J.; Tondo, C.L.; Ledesma, N.; Campbell, R.J.; Bost, J.; Schnell, R.J.; Gutiérrez, O.A. Genetic Differentiation, Races and Interracial Admixture in Avocado (*Persea americana* Mill.), and *Persea* spp. Evaluated Using SSR Markers. *Genet. Resour. Crop Evol.* **2018**, *65*, 1195–1215. [CrossRef]
35. Chen, H.; Morrell, P.L.; Ashworth, V.E.T.M.; de la Cruz, M.; Clegg, M.T. Tracing the Geographic Origins of Major Avocado Cultivars. *J. Hered.* **2008**, *100*, 56–65. [CrossRef]
36. Ge, Y.; Zhang, T.; Wu, B.; Tan, L.; Ma, F.; Zou, M.; Chen, H.; Pei, J.; Liu, Y.; Chen, Z.; et al. Genome-Wide Assessment of Avocado Germplasm Determined from Specific Length Amplified Fragment Sequencing and Transcriptomes: Population Structure, Genetic Diversity, Identification, and Application of Race-Specific Markers. *Genes* **2019**, *10*, 215. [CrossRef]
37. Rubinstein, M.; Eshed, R.; Rozen, A.; Zviran, T.; Kuhn, D.N.; Irihimovitch, V.; Sherman, A.; Ophir, R. Genetic Diversity of Avocado (*Persea americana* Mill.) Germplasm Using Pooled Sequencing. *BMC Genom.* **2019**, *20*, 379. [CrossRef]

38. Kämper, W.; Trueman, S.J.; Cooke, J.; Kasinadhuni, N.; Brunton, A.J.; Ogbourne, S.M. Single-Nucleotide Polymorphisms That Uniquely Identify Cultivars of Avocado (*Persea americana*). *Appl. Plant Sci.* **2021**, *9*, e11440. [CrossRef]
39. Kuhn, D.N.; Livingstone, D.S.; Richards, J.H.; Manosalva, P.; Van den Berg, N.; Chambers, A.H. Application of Genomic Tools to Avocado (*Persea americana*) Breeding: SNP Discovery for Genotyping and Germplasm Characterization. *Sci. Hortic.* **2019**, *246*, 1–11. [CrossRef]
40. Talavera, A.; Soorni, A.; Bombarely, A.; Matas, A.J.; Hormaza, J.I. Genome-Wide SNP Discovery and Genomic Characterization in Avocado (*Persea americana* Mill.). *Sci. Rep.* **2019**, *9*, 20137. [CrossRef] [PubMed]
41. Schulman, A.H.; Flavell, A.J.; Ellis, T.H.N. The Application of LTR Retrotransposons as Molecular Markers in Plants. In *Mobile Genetic Elements: Protocols and Genomic Applications*; Methods in Molecular Biology; Miller, W.J., Capy, P., Eds.; Humana Press: Totowa, NJ, USA, 2004; pp. 145–173. ISBN 978-1-59259-755-0.
42. Kalendar, R.; Flavell, A.J.; Ellis, T.H.N.; Sjakste, T.; Moisy, C.; Schulman, A.H. Analysis of Plant Diversity with Retrotransposon-Based Molecular Markers. *Heredity* **2011**, *106*, 520–530. [CrossRef] [PubMed]
43. Waugh, R.; McLean, K.; Flavell, A.J.; Pearce, S.R.; Kumar, A.; Thomas, B.B.T.; Powell, W. Genetic Distribution of Bare–1-like Retrotransposable Elements in the Barley Genome Revealed by Sequence-Specific Amplification Polymorphisms (S-SAP). *Mol. Gen. Genet.* **1997**, *253*, 687–694. [CrossRef]
44. Kalendar, R.; Grob, T.; Regina, M.; Suoniemi, A.; Schulman, A. IRAP and REMAP: Two New Retrotransposon-Based DNA Fingerprinting Techniques. *Theor. Appl. Genet.* **1999**, *98*, 704–711. [CrossRef]
45. Kalendar, R.; Muterko, A.; Boronnikova, S. Retrotransposable Elements: DNA Fingerprinting and the Assessment of Genetic Diversity. In *Molecular Plant Taxonomy: Methods and Protocols*; Methods in Molecular Biology; Besse, P., Ed.; Springer: New York, NY, USA, 2021; pp. 263–286. ISBN 978-1-07-160997-2.
46. Kalendar, R.; Schulman, A.H. Transposon-Based Tagging: IRAP, REMAP, and IPBS. In *Molecular Plant Taxonomy*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 233–255.
47. Kalendar, R.; Schulman, A.H. IRAP and REMAP for Retrotransposon-Based Genotyping and Fingerprinting. *Nat. Protoc.* **2006**, *1*, 2478–2484. [CrossRef]
48. Kalendar, R.; Amenov, A.; Daniyarov, A. Use of Retrotransposon-Derived Genetic Markers to Analyse Genomic Variability in Plants. *Funct. Plant Biol.* **2019**, *46*, 15–29. [CrossRef]
49. Kalendar, R.; Antonius, K.; Smýkal, P.; Schulman, A.H. IPBS: A Universal Method for DNA Fingerprinting and Retrotransposon Isolation. *Theor. Appl. Genet.* **2010**, *121*, 1419–1430. [CrossRef]
50. Galindo-González, L.; Mhiri, C.; Deyholos, M.K.; Grandbastien, M.-A. LTR-Retrotransposons in Plants: Engines of Evolution. *Gene* **2017**, *626*, 14–25. [CrossRef]
51. Seberg, O.; Petersen, G. A Unified Classification System for Eukaryotic Transposable Elements Should Reflect Their Phylogeny. *Nat. Rev. Genet.* **2009**, *10*, 276. [CrossRef]
52. Wicker, T.; Sabot, F.; Hua-Van, A.; Bennetzen, J.L.; Capy, P.; Chalhoub, B.; Flavell, A.; Leroy, P.; Morgante, M.; Panaud, O. A Unified Classification System for Eukaryotic Transposable Elements. *Nat. Rev. Genet.* **2007**, *8*, 973–982. [CrossRef]
53. Griffiths, J.; Catoni, M.; Iwasaki, M.; Paszkowski, J. Sequence-Independent Identification of Active LTR Retrotransposons in Arabidopsis. *Mol. Plant* **2018**, *11*, 508–511. [CrossRef]
54. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [CrossRef]
55. Sneath, P.H.; Sokal, R.R. Unweighted Pair Group Method with Arithmetic Mean. In *Numerical Taxonomy*; CAB International: Wallingford, UK, 1973; pp. 230–234.
56. Buquicchio, F.; Spruyt, M. Gene Runner. Available online: www.generunner.net (accessed on 17 May 2019).
57. Jacquemoud-Collet, J.-P.; Perrier, X. DARWin: Dissimilarity Analysis and Representation for Windows. Available online: https://darwin.cirad.f (accessed on 17 May 2006).
58. Castro-López, C.; Bautista-Hernández, I.; González-Hernández, M.D.; Martínez-Ávila, G.C.G.; Rojas, R.; Gutiérrez-Díez, A.; Medina-Herrera, N.; Aguirre-Arzola, V.E. Polyphenolic Profile and Antioxidant Activity of Leaf Purified Hydroalcoholic Extracts from Seven Mexican *Persea americana* Cultivars. *Molecules* **2019**, *24*, 173. [CrossRef]
59. Pandey, R.N.; Adams, R.P.; Flournoy, L.E. Inhibition of Random Amplified Polymorphic DNAs (RAPDs) by Plant Polysaccharides. *Plant Mol. Biol. Rep.* **1996**, *14*, 17–22. [CrossRef]
60. Ramírez, I.M.; Rodríguez, N.N.; Valdés-Infante, J.; Capote, M.; Becker, D.; Rohde, W. Isolation of Genomic DNAs from the Tropical Fruit Trees Avocado, Coconut, Guava and Mango for PCR-Based DNA Marker Application. *Cultiv. Trop.* **2004**, *25*, 33–38.
61. Bennetzen, J.L. Transposable Element Contributions to Plant Gene and Genome Evolution. *Plant Mol. Biol.* **2000**, *42*, 251–269. [CrossRef]
62. Lisch, D. How Important Are Transposons for Plant Evolution? *Nat. Rev. Genet.* **2012**, *14*, 49–61. [CrossRef]
63. Schulman, A.H. Retrotransposon Replication in Plants. *Curr. Opin. Virol.* **2013**, *3*, 604–614. [CrossRef]
64. Lanciano, S.; Mirouze, M. Transposable Elements: All Mobile, All Different, Some Stress Responsive, Some Adaptive? *Curr. Opin. Genet. Dev.* **2018**, *49*, 106–114. [CrossRef]
65. Medina, N.N.R.; Lorenzo, J.L.F.; Arbelo, O.C.; Fiallo, V.R.F.; Pérez, I.M.R.; Becker, D.; García, I.R.; Arencibia, C.G.; Martín, X.X.; Gutiérrez, M.I.R. Agro-Morphologic Traits, Isoenzyme and DNA Markers for Estimating the Polymorphism Levels, Discriminating Capacity and Informativeness in Avocado. *Rev. CENIC Cienc. Biol.* **2009**, *40*, 63–74.

66. Baránek, M.; Meszáros, M.; Sochorová, J.; Čechová, J.; Raddová, J. Utility of Retrotransposon-Derived Marker Systems for Differentiation of Presumed Clones of the Apricot Cultivar Velkopavlovická. *Sci. Hortic.* **2012**, *143*, 1–6. [CrossRef]
67. Guo, D.-L.; Guo, M.-X.; Hou, X.-G.; Zhang, G.-H. Molecular Diversity Analysis of Grape Varieties Based on IPBS Markers. *Biochem. Syst. Ecol.* **2014**, *52*, 27–32. [CrossRef]
68. Milovanov, A.; Zvyagin, A.; Daniyarov, A.; Kalendar, R.; Troshin, L. Genetic Analysis of the Grapevine Genotypes of the Russian Vitis Ampelographic Collection Using IPBS Markers. *Genetica* **2019**, *147*, 91–101. [CrossRef]
69. Mehmood, A.; Luo, S.; Ahmad, N.M.; Dong, C.; Mahmood, T.; Sajjad, Y.; Jaskani, M.J.; Sharp, P. Molecular Variability and Phylogenetic Relationships of Guava (*Psidium guajava* L.) Cultivars Using Inter-Primer Binding Site (IPBS) and Microsatellite (SSR) Markers. *Genet. Resour. Crop Evol.* **2016**, *63*, 1345–1361. [CrossRef]
70. Al-Najm, A.; Luo, S.; Ahmad, N.M.; Trethowan, R. Molecular Variability and Genetic Relationships of Date Palm (*Phoenix dactylifera* L.) Cultivars Based on Inter-Primer Binding Site (IPBS) Markers. *Aust. J. Crop Sci.* **2016**, *10*, 732. [CrossRef]
71. Yaldiz, G.; Camlica, M.; Nadeem, M.A.; Nawaz, M.A.; Baloch, F.S. Genetic Diversity Assessment in *Nicotiana tabacum* L. with IPBS-Retrotransposons. *Turk. J. Agric. For.* **2018**, *42*, 154–164. [CrossRef]
72. Demirel, U.; Tındaş, İ.; Yavuz, C.; Baloch, F.S.; Çalışkan, M.E. Assessing Genetic Diversity of Potato Genotypes Using Inter-PBS Retrotransposon Marker System. *Plant Genet. Resour.* **2018**, *16*, 137–145. [CrossRef]
73. Khapilina, O.; Raiser, O.; Danilova, A.; Shevtsov, V.; Turzhanova, A.; Kalendar, R. DNA Profiling and Assessment of Genetic Diversity of Relict Species *Allium Altaicum* Pall. on the Territory of Altai. *PeerJ* **2021**, *9*, e10674. [CrossRef]
74. Khapilina, O.; Turzhanova, A.; Danilova, A.; Tumenbayeva, A.; Shevtsov, V.; Kotukhov, Y.; Kalendar, R. Primer Binding Site (PBS) Profiling of Genetic Diversity of Natural Populations of Endemic Species *Allium ledebourianum* Schult. *BioTech* **2021**, *10*, 23. [CrossRef]
75. Doungous, O.; Kalendar, R.; Filippova, N.; Ngane, B.K. Utility of IPBS Retrotransposons Markers for Molecular Characterization of African Gnetum Species. *Plant Biosyst.-Int. J. Deal. All Asp. Plant Biol.* **2020**, *154*, 587–592. [CrossRef]
76. González Carracedo, M.; Tejera-Pérez, H.; Hernández Ferrer, M.; Jiménez Arias, D.; Pérez Pérez, J.A. Comparative Assessment of Microsatellite and Retrotransposon-Based Markers for Genetic Characterization of Commercial Banana Cultivars (*Musa* spp.). *Plant Breed.* **2021**, *140*, 968–980. [CrossRef]
77. Rendón-Anaya, M.; Ibarra-Laclette, E.; Méndez-Bravo, A.; Lan, T.; Zheng, C.; Carretero-Paulet, L.; Perez-Torres, C.A.; Chacón-López, A.; Hernandez-Guzmán, G.; Chang, T.-H.; et al. The Avocado Genome Informs Deep Angiosperm Phylogeny, Highlights Introgressive Hybridization, and Reveals Pathogen-Influenced Gene Space Adaptation. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 17081–17089. [CrossRef] [PubMed]
78. Ashworth, V.E.T.M. Microsatellite Markers in Avocado (*Persea americana* Mill.): Genealogical Relationships Among Cultivated Avocado Genotypes. *J. Hered.* **2003**, *94*, 407–415. [CrossRef]