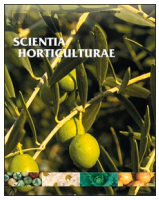
Scientia Horticulturae 338 (2024) 113775

Contents lists available at ScienceDirect

Scientia Horticulturae

journal homepage: www.elsevier.com/locate/scihorti



Research Paper

Towards improved markers for molecular characterization in kiwifruit

Maria Teresa Marrazzo, Andrea Amilcare Passerelli, Guido Cipriani, Gloria De Mori *

Department of Agriculture, Food, Environmental and Animal Sciences, University of Udine, Via delle Scienze 206, 33100 Udine, Italy

ARTICLE INFO

Keywords:
Microsatellite
SSR
Fingerprinting
Actinidia

ABSTRACT

Actinidia, a genus of kiwifruit, has 54 species, with the commercially dominant species being one with two subspecies: *Actinidia chinensis* var. *chinensis* and *A. chinensis* var. *deliciosa*. The global kiwifruit industry has used a single variety for decades, but in the last 20 years, new varieties have been introduced, with breeders and propagation rights concerned about protecting them from fraudulent use. The genome of *A. chinensis* var. *chinensis* 'Red5' was searched for perfect microsatellites with repeat motifs of no less than two to ten bases. Out of 216,456 possible perfect microsatellite loci, 82 were chosen to be spread across all 29 chromosomes. Twenty microsatellite loci with repeat motifs over two have been developed for the varietal characterization of kiwifruit. The markers consist of repeating motifs of at least three bases that suffer less from problems in interpreting electrophoretic profiles due to stuttering. A single amplification protocol valid for all loci was developed. The markers were tested on a sample of 100 genotypes that included diploid, tetraploid, hexaploid, and octoploid individuals. The selected markers were able to clearly discriminate all genotypes except for two clonal mutations. This indicates that the microsatellite loci are highly effective in distinguishing between different kiwifruit varieties, regardless of their ploidy level. The ability to accurately characterize kiwifruit genotypes using these markers can greatly benefit breeding programs and conservation efforts for this important fruit crop.

1. Introduction

Since 1904, when some *Actinidia chinensis* var. *deliciosa* berries from a wild population on the banks of the Yangtze River in China were brought to New Zealand, the crop has become more popular in fresh fruit markets worldwide. This occurred subsequent to the initial domestication, followed by the development of commercial cultivars (Ferguson, 2004).

The *Actinidia* genus belongs to the *Actinidiaceae* family (order *Ericales*) and the distribution of wild individuals ranges from southwestern China, as mentioned above, to Siberia and eastwards to the Korean peninsula and Japan (Huang, 2016). Kiwifruit has been collected by populations in these regions since ancient times for consumption. Explorer Robert Fortune introduced the fruit to Europe in 1847, but botanist E.H. Wilson's second attempt was more successful. The first seedlings were presented to the public in 1903, attracting interest more as an ornamental plant than a crop. The species also spread to America, Australia, and New Zealand, where favorable environmental conditions led to the establishment of most commercial cultivars (Ferguson, 2004).

The different selective environments, the numerous inter- and intraspecific hybridizations, and the recurrent naturally occurring polyploidization events have resulted in an extremely complex evolutionary network within the *Actinidia* genus, making taxonomic studies complicated. One of the most recent reviews, carried out by Li et al. (2007) and revised by Huang (2016), allowed the identification of 54 species (52 endemic to China, 1 from Japan, and 1 from Nepal) and 21 botanical varieties.

The ploidy of individuals belonging to the *Actinidia* genus is extremely variable, and the basic number of chromosomes is $x = 29$. It is thought that this number is so high as a result of open-pollinated crosses and subsequent re-diploidization events that occurred in wild ancestral progenitors (Ferguson and Huang, 2007). Flow cytometry studies identified diploid, tetraploid, hexaploid, heptaploid, and octoploid individuals (Ferguson et al., 2009). Geographical separation of these individuals can hinder genetic material exchange, but also provide breeders opportunities to combine genetic pools separated by geographical and biological barriers.

Over the last 20 years, the kiwifruit has been the crop with the greatest per capita growth among the main fruit species ("Statistics | FAO | Food and Agriculture Organization of the United Nations"). Although kiwifruit production represents only 0.25 % of all fruit produced worldwide (Wu, 2019), it is thought that this growth trend will

* Corresponding author.
E-mail address: gloria.demori@uniud.it (G. De Mori).

https://doi.org/10.1016/j.scienta.2024.113775
Received 9 August 2024; Received in revised form 24 October 2024; Accepted 24 October 2024
Available online 1 November 2024

also be observable in the next decade, given the challenges that breeders are facing to meet the emerging needs of consumers and to resolve the issues related to the spread of bacterial canker (*Pseudomonas syringae* pv. *actinidiae*; PSA) epidemics of kiwifruit in Europe and New Zealand (Yao et al., 2022). Despite the abundant availability of germplasm within the *Actinidia* genus, kiwifruit breeding programs for commercial and production purposes have focused almost exclusively on *A. chinensis*, which includes the varieties *A. chinensis* var. *chinensis* kiwifruit with yellow, green, or bicolour red/yellow flesh) and *A. chinensis* var. *deliciosa* (green flesh kiwifruit). Together, these varieties represent approximately 68 % of global kiwifruit production (Pinto and Vilela, 2018). Other significant species include *A. arguta* (kiwi berry), and *A. eriantha*. To address the limited genetic variability in cultivated kiwifruit and the problems associated with PSA, kiwifruit breeding activities have been intensified in recent years in all the major producing countries. The goal is to obtain new cultivars that meet consumer preferences and are resistant/tolerant to biotic stress. Consequently, the need emerged to identify methods that allow, in an objective and standardized way, to discriminate between the different cultivars, both to facilitate breeders in the selection process and to protect intellectual property rights, as provided for by the Trade agreement -Related Intellectual Property Rights- (TRIPs), signed between member countries of the World Trade Organization (WTO) (Archak, 2000). Over time, among the varietal characterization techniques, the use of microsatellite molecular genetic markers is the one that has become most established.

Microsatellites (Litt and Luty, 1989), also known as Short Tandem Repeats (STRs) (Edwards et al., 1991), Simple Sequence Repeats (SSRs) or Simple Sequence Length Polymorphism (SSLP) (Tautz, 1989), are short nucleotide sequences of 1–6 base pairs (bp) repeated in tandem and uniformly distributed within the prokaryotic and eukaryotic genomes (Zane et al., 2002). SSRs are characterized by a low number of repeats per locus (5–100), a random distribution of approximately 10^4–10^5 per genome (Tautz, 1993) and high-length polymorphism (Zane et al., 2002) . The latter is due to differences in the number of repeats of the basic motif in the microsatellite loci, making them identifiable markers with PCR and highly reproducible. The genesis of microsatellites is a dynamic and extremely complex evolutionary process (Ellegren, 2004) and it is widely accepted that any mechanism involving new DNA synthesis (replication, recombination, and DNA repair mechanisms) can generate a variation in the length of the tandem repeats (Richard et al., 2008). Currently, the most accepted hypothesis is that SSR polymorphism occurs as a result of Slipped Strand Mispairing (SSM) phenomena. The repetitive nature of the nucleotide sequence favors secondary DNA structures that could be responsible for the mutational process (Richard et al., 2008) . The formation of hairpins, triple helices, and the inhibition of binding to Single Strand Binding Proteins (SSB) can cause an arrest of the replication protein complex. In this situation, the lagging and leading strands can detach and repair, not always occurring correctly, and it can cause one of the two filaments to slip over the other. The slippage of DNA polymerase III on the lagging strand determines a change in expansion, while the slippage on the leading strand causes a contraction if the mismatches are not repaired by the mismatch repair system (Wang et al., 2003). The study by Bhargava and Fuentes (2009) highlighted how even unequal crossing-over can lead to variations in the number of repeats of the core repeat. This phenomenon occurs during homologous recombination when two sequences are not correctly aligned. The greater the similarity between the sequences, the greater the probability that an unequal crossing-over will occur. This type of mutation involves the loss of a DNA fragment on one strand and duplication on the other. Unlike replication slippage, the mechanism just described would better explain large-scale mutations involving many repeats. The theories previously reported do not exclude that SSRs can also originate from retrotransposition events. Retrotransposons are DNA fragments rich in repeated sequences inserted into chromosomes following a retrotranscription phenomenon starting from an RNA molecule (Kalia et al., 2010). In support of this mechanism of SSR formation, Parida et al. (2010) observed that, in sugarcane (*Saccharum officinarum*), 23.1 % of genomic sequences containing microsatellites present significant sequence homology with retrotransposons.

A work by Guichoux et al. (2011) has highlighted that, generally in population genetic studies relating to both the animal and plant worlds, di-nucleotide microsatellites are predominantly searched for and used. However, this type of SSR tends to produce one or more stutter bands during the analysis phase, i.e., multiple PCR products, usually shorter than one or a few repetitions, of the same amplified fragment (Chambers and MacAvoy, 2000). This phenomenon is due to slippage of the polymerase during amplification and makes allelic recognition more difficult, especially in heterozygous individuals with adjacent alleles. In contrast, polymerase slippage occurs less likely in regions with SSRs having core repeats of 3–5 nucleotides (Edwards et al., 1991).

On average, SSRs that present a high number of repeats of the basic motif are characterized by a higher mutation rate: selecting loci with a sufficient number of repeats is necessary to ensure a certain degree of polymorphism in different individuals. To confirm what has been said, van Asch et al. (2010) recommend selecting tetra-nucleotide microsatellites with several repeats between 11 and 16. The lower limit was identified because the mutation rate of alleles with more than 11 repeats is higher, allowing the identification of more polymorphic loci. The upper limit was set because it was thought that alleles with more than 16 repeats are more likely to have core motifs that are broken, which could make it hard to figure out what the results mean.

In this work, we propose a set of new microsatellite markers with bases characterizing the repeat of no less than three. Several versions of high-quality genomes have been published for *Actinidia chinensis* in recent years (Han et al., 2023; Wu et al., 2023; Yue et al., 2023). An exhaustive analysis was conducted on the sequenced genome of the diploid genotype of *A. chinensis* var. *chinensis* 'Red5' (Pilkington et al., 2018) to identify all the microsatellite sequences on which to choose for molecular analyses to identify the markers that produced fewer interpretative difficulties, having a high power discriminatory between the varieties analysed.

2. Material and methods

2.1. Plant material

Leaves from young shoots were harvested during Spring 2023 from the *Actinidia* germplasm repository of the University of Udine, Italy (Latitude: 46.038261 | Longitude: 13.221413). One hundred samples were collected, 87 belonged to the species *Actinidia chinensis* (54 *Actinidia chinensis* var. *chinensis* and 33 *Actinidia chinensis* var. *deliciosa*), and 13 belonged to different species of the genus *Actinidia*. Fifteen samples were diploid, 51 were tetraploid, 33 were hexaploid, and one was octoploid (Supplementary Table 1). Leaves were lyophilized for one week and immediately stored at 4 °C until processing.

2.2. DNA extraction, quantification, and quality checking

DNA was extracted from lyophilized leaves using the commercial kit NucleoSpin® Plant II (Macherey-Nagel) following the manufacturer's instructions. DNA samples were quantified using both the spectrophotometer NanoDrop ND-1000 (Thermo Fisher Scientific) and the fluorimeter Qbit® 3.0 (Invitrogen™). DNA integrity was checked on 1 %, 0.5x TBE agarose gel stained with SYBR™ Safe DNA Gel Stain (ThermoFisher Scientific).

2.3. SSR research and primer design

The search for SSRs was carried out with the open-source software GMATA (Genome-wide Microsatellite Analyzing Toward Application) developed by Xuewen Wang (Wang and Wang, 2016) . Thanks to this tool, it was possible to identify the SSRs that met certain criteria present

in the reference genome of *A. chinensis* var. *chinensis* genotype 'Red5′ (2n=58) (Pilkington et al., 2018).

The sequences of each of the 29 non-homologous chromosomes were downloaded in FASTA format from the NCBI database (accession: GCA_003024255.1) and individually entered into the software. Parameters were then set to identify all the SSRs and their related forward and reverse primers required for the subsequent amplification steps.

The main setting parameters were as follows:

- a basic motif between 2 and 10 nucleotides in length, repeated at least 5 times;
- a length of the amplicon between 120 and 400 bp;
- primer with an optimal annealing temperature of 60 °C.

For each chromosome (Chr), the information reported relating to all the SSRs present and their primers are:

- A nucleotide sequence that constitutes the basic motif of the SSR;
- start and end positions of the SSR within the Chr sequence;
- length and number of repetitions of the basic SSR sequence;
- SSR forward and reverse primer sequences;
- annealing temperatures of the forward and reverse primers of the SSR;
- expected length of the SSR amplification product.

The physical distribution of SSR markers along the chromosomes was visualized using the 'karyoploteR' (Gel and Serra, 2017) and 'GenomicRanges' (Lawrence et al., 2013) packages in R. A distinct colour was assigned to each marker class to enhance their identification. To visualize the frequency of SSR markers per megabase (Mb), the percentage of microsatellite classes within the 'Red5' genome, and the proportions of di-nucleotide marker classes across the 'Red5' genome, the 'ggplot2' (Wickham, 2016) package in R was used.

2.4. In silico preliminary selection of SSR loci

All the sequences containing perfect microsatellites with a repetition motif of 3–6 bp were ranked for each Chr. Loci were randomly selected among those that in the genotype 'Red5' had the larger number of repetitions avoiding polyN (E.g. (AAA)n for a trinucleotide motif) and the shortest amplified length. At least one locus was selected for each Chr. Three cycles of loci selection were subsequently performed, in which the success of the amplification by PCR was tested and the robustness of the analysis carried out after electrophoretic separation on a capillary sequencer was determined.

2.5. Loci validation

Four samples of different ploidy levels (2n, 4n, and 6n) were selected for the preliminary screening of the primers (Merck – Sigma). The forward primers of each putative SSR marker were added with M13 sequence for pig-tailing PCR amplification (Schuelke, 2000). PCR reactions were carried out in 10 μl total volume using 200 μM each dNTP, 20 ng of genomic DNA, 0.008 μM forward primer, 0.2 μM reverse primer, 0.2 μM M13 e 0.5 U of HotMaster Taq DNA polymerase (QuantaBio). The PCR reactions were carried out in a 2720 thermal cycler (Applied Biosystem) with the following thermal profile: 94 °C for 2 min followed by 10 touch-down cycles at 94 °C for 20 s, 55 °C for 10 s, with −0.5° reduction for cycle, and 65 °C for 30 s; followed by 30 cycles at 94 °C for 20 s, 50 °C for 20 s,65 °C for 30 s and a final elongation step of 1 hour at 65 °C. Amplified products were separated into 1 %, 0.5 % TBE agarose gel. Poor, erratic or no amplificated loci were discarded, and no further analyses were tested to change PCR conditions.

2.6. Fingerprinting analysis

Validated loci were separated on the SeqStudio Genetic Analyzer (ThermoFisher Scientific) capillary sequencer using the manufacturer's standard run parameters. The forward primers of each putative SSR marker were added with M13 sequence for pig-tailing PCR amplification. M13 primer was labeled with four different fluorophores (6-FAM, NED, PET, VIC, ThermoFisher Scientific). PCR products labeled with Fam, Ned and Vic of the four samples were diluted 1:100 with MilliQ® water, while those labeled with Pet were diluted 1:50. 1 μL of the diluted solution was added to 0,05 μL of GeneScan™ 600 LIZ™ *dye Size Standard* (ThermoFisher Scientific) and 7,95 μL formamide. Low-quality loci, mainly because of stuttering peaks, were discarded. After the double primer selection process, the remaining SSR loci were amplified on all 100 kiwifruit samples using the same PCR and analysis protocol.

2.7. Allele calling and genotyping

Alleles identification and calling were performed with the Microsatellite Analysis Software (MAS) available on the website of Thermofisher Scientific. Parameters were set for the automatic identification of the alleles of 2n, 4n, 6n, and 8n ploidy levels of the kiwifruit genotypes. Misidentification correction and binning were performed manually after visualization of all peaks on each of the 100 samples. The allele size, height, and area of each peak were recorded for further analysis. Allele frequencies, polymorphism information content (PIC), and genetic distances were calculated using the Polygene software (Huang et al., 2020) on three data sets obtained by the analysis of diploid, tetraploid, and hexaploid genotypes. Euclidean distance matrices were calculated on the three datasets of different ploidy levels using Polygene. Phenetic classification analysis and UPGMA (Unweighted Pair Group Method with Arithmetic mean) hierarchical clustering trees were carried out with MEGA11: Molecular Evolutionary Genetics Analysis version 11 (Tamura et al., 2021) .

3. Results

3.1. Microsatellite in 'Red5' genome

Out of the 29 chromosomes of the cultivar 'Red5′, 216,456 potential perfect microsatellite loci with motifs ranging from two to ten bp were identified (Supplementary Table 2). SSRs were evenly spread across the *A. chinensis* var. *chinensis* 'Red5' genome (Supplementary Fig. 1). Chi-square tests ($p > 0.05$) revealed no significant differences between the observed and expected frequencies of microsatellite loci distribution on each chromosome, calculated based on the total SSRs identified and the genome length in millions of base pairs (Mbp). As the number of bases in the motif increased, the distribution uniformity in the chromosomes decreased (Fig. 1A). In the genome of the 'Red5′ genotype, on average, the frequency of microsatellites is 394.89 per million base pairs (Mbp). The highest number of microsatellites per Mbp was detected in Chr9 (432.91 SSRs/Mbp) and the lowest in Chr19 (333.72 SSRs/Mbp) (Fig. 1B).

Considering the total number of 216,456 microsatellites, 191,133 are characterized by di-, 16,848 tri-, 5326 tetra-, 1418 penta- and 1728 hexanucleotide repeats. Only on Chr 2, two 7-nucleotide motifs and one 8-nucleotide motif were detected. Nine-nucleotide and 10-nucleotide motifs were not identified. In the whole genome of 'Red5', the percentages of different motifs were 88.3 % for di-nucleotides, 7.78 % for tri-nucleotides, 2.46 % for tetra-nucleotides, 0.66 % for penta-nucleotides, and 0.80 % for hexa-nucleotides, respectively (Fig. 1C). Di-nucleotide SSRs were the most abundant class, highlighting their prevalence compared to tri-, tetra-, penta-, and hexa-nucleotide core repeats (Fig. 1C). The percentage of SSRs for the seven different motifs in each chromosome remained relatively constant, with a standard deviation ranging from 0.08 to 0.57 for penta- and di-nucleotides, which

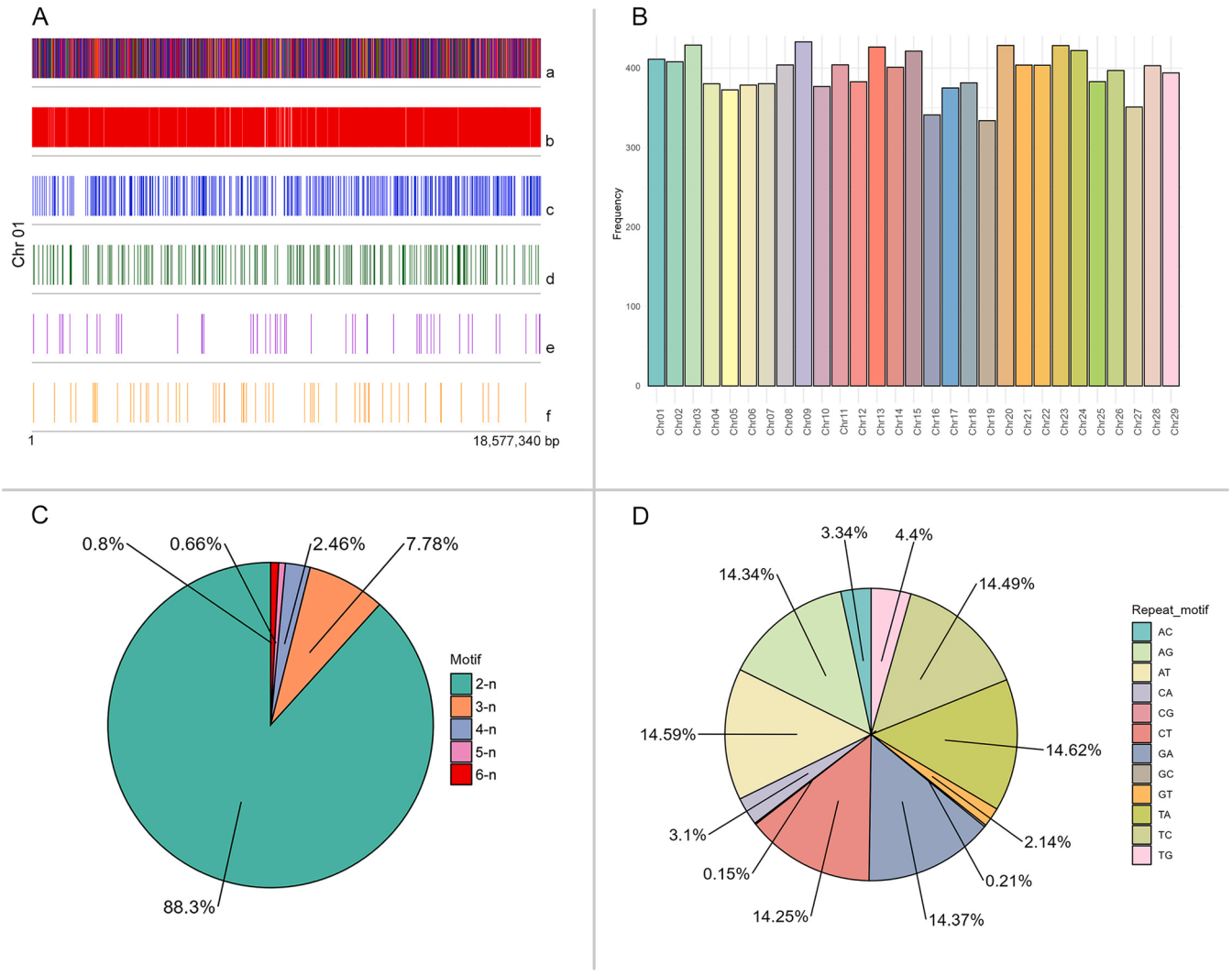


Fig. 1. A: Physical distribution of the different classes of SSRs loci in *Actinidia chinensis* var. *chinensis* 'Red5' chromosome 1. (**a**) All classes distribution; (**b**) di-nucleotide distribution; (**c**) tri-nucleotide distribution; (**d**) tetra-nucleotide distribution; (**e**) penta-nucleotide distribution; (**f**) hexa-nucleotide distribution. **B:** frequencies of the SSRs/Mb in the 29 'Red5' chromosomes. **C:** percentage of the different classes of SSRs in the whole 'Red5' genome. **D:** percentages of the different classes of the di-nucleotide SSRs in the whole 'Red5' genome.

were in the most extreme classes (Supplementary Table 2).

The percentage of observed classes of repeat motifs on the theoretical number of classes decreased from 100 % for di- and trinucleotide to 18 % in hexanucleotide motifs. (Table 1). There were 12 different classes of di-nucleotide repeat motifs in all chromosomes, aligning with the theoretical number of classes and excluding motifs composed of the same two bases (AA, CC, GG, TT) (Fig. 1D). The mean percentage of different class motifs within all chromosomes ranged from 89.4 % (Std Dev 4.78) to 1.29 (Std Dev 0.24) in the trinucleotide and hexanucleotide SSR motifs, respectively (Supplementary Table 3). Among the 12 classes of di-nucleotide microsatellites observed, the core repeat TA was the most frequent (mean 14.67 %, Standard Deviation 0.78), whereas the core repeat GC was the least frequent (mean 0.14 %, Standard Deviation 0.04).

The motifs AAT, AAAT, AAAAT, AAAAAT were the most frequent motifs of tri-, tetra-, penta- and hexanucleotide in the 'Red5' genome, respectively (data not shown).

3.2. Selection and validation of SSR loci

The main objective of this work was the identification of potential microsatellite loci useful in the varietal characterization of kiwifruit, which would guarantee unambiguous identification of the alleles, avoiding, as far as possible, the appearance of stuttering peaks. For this purpose, 82 loci were selected. These loci are distributed across all 29 chromosomes. The distribution of loci ranged from 1 to 4 per

Table 1
Percentage of observed classes of microsatellite motifs on the total theoretical number of classes in the 29 *Actinidia chinensis* var. *chinensis* 'Red5' chromosomes.

Motif length	2-n	3-n	4-n	5-n	6-n	7-n	8-n
Theoretical combinations	16	64	256	1024	4096	16,384	65,536
Theoretical combinations (without motifs made of the same nucleotide)	12	60	252	1020	4092	16,380	65,532
Observed combinations considering all Chromosomes	12	60	174	350	745	2	1
Percentage observed combination/Theoretical combination	1	1	0.6905	0.3431	0.1821	0.0001	2E-05

Row 1 identifies the kind of motif from di-nucleotide to octo-nucleotide.

chromosome (Table 2).

A total of 82 loci were tested: 37 loci with trinucleotide repeats (45 %), 16 with tetranucleotide repeats (20 %), 15 with pentanucleotide repeats (18 %), and 14 with hexanucleotide repeats (17 %). During the selection process to identify the most suitable loci for varietal characterization, 63 loci (77 %) were discarded. Of these, 19 loci (23 %) were eliminated after the preliminary analysis due to issues such as lack of amplification, unexpected amplicon lengths, or weak amplification. Another 32 loci (39 %) were rejected after the second phase of selection, which involved testing on a capillary sequencer with a small sample size (4). These loci were excluded mainly because of stuttering or difficulties in determining allele lengths. Finally, 12 loci (15 %) were discarded after testing the full set of 100 samples, with the main reason being ambiguity in assigning alleles. In the end, 19 primer pairs were selected for the fingerprinting analysis of kiwifruit varieties (Table 3).

These primer pairs amplified 20 microsatellite loci (the primer pair identified on Chr23 amplifies 2 loci, UDA_Chr23A and UDA_Chr23B) and are sufficiently robust for the unambiguous assignment of amplified alleles with common PCR conditions for all primers proposed in this work. Trinucleotide repeats were present in 13 selected loci (65 %), tetranucleotide repeats were present in 2 selected loci (10 %) and hexanucleotide repeats were present in 5 selected loci (25 %). None of the loci characterized by pentanucleotide repeats passed the three selection steps (Table 2).

3.3. Fingerprinting analysis

Diversity indices were calculated using Polygene software (Huang et al., 2020) individually for each ploidy level of the analysed genotypes, except for the unique octoploid sample belonging to the species *A. arguta* var. *purpurea*. In this case, no more than six alleles were detected in any of the loci, and the analysis was carried out together with that of the hexaploid individuals.

Overall, 160 alleles were detected, averaging 8.05 alleles per locus, considering all the 100 genotypes (Supplementary Table 4). Among the 15 diploid individuals in our analysed samples, the number of diploid individuals analysed with the 20 SSR loci varied. It ranged from 10 individuals at the locus UDA_Chr10, where five samples had issues with amplification or interpretation, to 15 individuals at 12 loci (60 %) where all samples were successfully amplified and interpreted clearly. On average, 14.15 diploid individuals were analysed using all 20 SSR loci (Table 4). The number of alleles per locus ranged from three (10 % of the loci) to 11 for the locus UDA_Chr2, averaging 7 on the 20 loci (Table 4). The average values for H0, He, and PIC were 0.48, 0.75, and 0.72, respectively. There was no correlation between the SSR motif and the PIC value. For example, the highest and the lowest PIC values, 0,87 and 0.39, were both found in the two tetranucleotide motif SSRs.

Within the 51 tetraploid individuals in our analysed genotypes pool, the number of individuals successfully analysed with the 20 SSR loci varied. For the locus UDA_Chr10, 46 individuals were analysed, five samples either failed to amplify, had poor amplification, or could not be clearly interpreted. In contrast, all 51 individuals were successfully amplified and clearly interpreted for 14 loci (70 %). The average of tetraploid individuals analysed with all the 20 SSR loci was 50.35 (Table 5). The number of alleles per locus ranged from three (10 % of the loci) to 18 for the loci UDA_Chr17 and UDA_Chr29, averaging 9.60 on the 20 loci (Table 5). The average H0, He, and PIC were 0.56, 0.67, and 0.64, respectively. There was no correlation between the SSR motif and the PIC value. The highest and the lowest PIC values, 0,90 and 0.32, were found in the loci UDA_Chr29 and UDA_Chr25, respectively.

All the 33 hexaploid individuals analysed with the 20 SSR loci were amplified giving a non-ambiguous interpretation (Table 6). The number of alleles per locus ranged from three (10 % of the loci) to 14, for the locus UDA_Chr29, averaging 7.6 on the 20 loci (Table 6). The average H0, He, and PIC were 0.56, 0.61, and 0.57, respectively. There was no correlation between the SSR motif and the PIC value. The highest and the lowest PIC values, 0,82 and 0.22, were found in the loci UDA_Chr15 and UDA_Chr3, respectively.

Considering all 100 cultivars and selection, irrespective of the ploidy level, the lowest allele frequency was 0.01, observed at the locus UDA_Chr6, and the highest allele frequency was 0.87, observed at the locus UDA_Chr3.

Table 2
Distribution of the 82 microsatellite loci tested among the 29 *Actinidia chinensis* var. *chinensis* 'Red5' chromosomes.

Chr	Primer tested	3n-motif	4n- motif	5n-motif	6n-motif	Number of SSR Selected	Selected motif	Notes
Chr1	3	1		1	1	1	3n	
Chr2	1	1				1	3n	
Chr3	4	3		1		1	3n	
Chr4	4	2	1	1		1	3n	
Chr5	3	1		1	1	0		
Chr6	2	1		1		1	3n	
Chr7	4	1	2	1		0		
Chr8	4	2			2	0		
Chr9	2	1		1		1	3n	
Chr10	2	1	1			1	4n	
Chr11	2		1		1	1	6n	
Chr12	3	2	1			1	3n	
Chr13	2	1		1		1	3n	
Chr14	4	2	2			0		
Chr15	2	1	1			1	3n	
Chr16	4	3	1			0		
Chr17	2	1			1	1	3n	
Chr18	2				2	1	6n	
Chr19	3		1	2		0		
Chr20	4		2	2		0		
Chr21	4	2		1	1	0		
Chr22	3	1	2			1	4n	
Chr23	2	1			1	1	6n	two loci
Chr24	3			2	1	0		
Chr25	4	4				1	3n	
Chr26	4	3			1	0		
Chr27	3	1	1		1	1	3n	
Chr28	1				1	1	6n	
Chr29	1	1				1	3n	
Tot	82	37	16	15	14	19		

Table 3
Identification of the microsatellites on *Actinidia chinensis* var. *chinensis* 'Red5' genome and primer pairs of the 20 SSR loci selected for *Actinidia* fingerprinting.

Locus name	Primer pair	Core repeat	Product Size	Position within 'Red5' genome
UDA_Chr1	For: AACGCGTCAAATTTCCTGAC Rev:TGATCTTCTGCTCATTGGTGA	(AGC)10	215	737,564 - 737,993
UDA_Chr2	For: CGACATCACGACCACATTCT Rev: TTCATCCCTGCATATGACCA	(AAT)30	229	6,108,231 - 6,108,720
UDA_Chr3	For: GGAGAAAATCCGGAGAAAGG Rev: CTTTTACCCGTAGGCCACAA	(GCT)8	137	3,837,293 - 3,837,716
UDA_Chr4	For: CGCACATCATCATCATCCTC Rev: TCGCTGTCTGGAACCTTCTT	(TCG)6	227	8,472,049 - 8,472,466
UDA_Chr6	For: GCATCGTCTTCGTCGAGTTT Rev: CACACCATTTTCGTGCTCTC	(GAT)10	245	14,351,823 - 14,352,252
UDA_Chr9	For: AAAGCAAGCTCCAACCTCAA Rev: CTTGAGAATCCAGTGCACCA	(GCA)8	269	13,556,854 - 13,557,277
UDA_Chr10	For: TTGAGCAGCAAGGAAAGGTT Rev: CATCACTTTGCTACCGCAGA	(ACAT)7	262	18,208,025 - 18,208,452
UDA_Chr11	For: ACGGTGTCGGTGTAGGTAGG Rev: TGCAGATCCCTCTCTTTGCT	(GGAGTC)7	164	7,298,549 – 7,298,990
UDA_Chr12	For: ACGGAGTCCTGTTCAAATGC Rev: TCACCTTGCTGATGCAGAAC	(GCA)6	183	6,536,029 - 6,536,446
UDA_Chr13	For: GCTAATGGTGGCGGTAAGAA Rev: AAATTGCTGATAGGGCTGGA	(GAC)8	274	5,969,194 - 5,969,617
UDA_Chr15	For: AAAAGACAGGGTGTCGGATG Rev: GTCCTTCCAATCCCTTCACA	(ATG)7	333	14,939,561 - 14,939,981
UDA_Chr17	For: TCGTCATCGTCATCGTCATT Rev: CACTCATGCCATTCATACCG	(TCA)8	151	15,696,268 - 15,696,691
UDA_Chr18	For: AGTGTAGGCAGGCCGTAATG Rev: GGCCATATCACCGATACCAA	(ATCGGT)5	361	19,635,197 - 19,635,626
UDA_Chr22	For: GCAGGTTCGGTTGTTTGATT Rev: GCTTCCACTGAGGTTCAAGG	(CATA)7	223	17,078,051 - 17,078,478
UDA_Chr23	For: CTTTGAGCGAGATCCTACCG Rev: ATTCGGGGAAGGTTACAAGC	(CTGGCA)6	203	23,607,751 - 23,608,186
UDA_Chr25	For: AAGAGGGGTTTTGGTGGTCT Rev: GCAGGCTGATCACAAGTTCA	(GAT)7	307	12,075,178 - 12,075,598
UDA_Chr27	For: GGAGAGCGTGAGAGATCTGG Rev: CGCCTCTTTCTCACACACAA	(TGA)6	195	6,742,949 - 6,743,366
UDA_Chr28	For: TCCTTTGCATGAAACATACCC Rev: CTTGCAAAATTCGACCAACA	(GACTCT)7	270	13,374,635 - 13,375,076
UDA_Chr29	For: GAAGGTTGGGACACAGGAAA Rev: TTCCTTGTTTTCGTCGCTCT	(GAA)81	370	16,784,830 - 16,785,472

Product Size is calculated in *A. chinensis* var. *chinensis* 'Red5' genome.

Table 4
Polymorphism indexes of the 20 SSRs loci in the diploid *Actinidia* genotypes.

Ploidy	Locus_name	k	n	Ho	He	PIC	Motif
2	UDA_Chr1	8	15	0.533333	0.802222	0.781993	3n
2	UDA_Chr2	11	13	0.461538	0.866864	0.853668	3n
2	UDA_Chr3	6	15	0.266667	0.797778	0.768205	3n
2	UDA_Chr4	10	14	0.571429	0.84949	0.833198	3n
2	UDA_Chr6	7	14	0.571429	0.806122	0.779812	3n
2	UDA_Chr9	8	15	0.533333	0.795556	0.769526	3n
2	UDA_Chr10	4	10	0.3	0.415	0.389388	4n
2	UDA_Chr11	8	15	0.533333	0.837778	0.818694	6n
2	UDA_Chr12	5	15	0.866667	0.704444	0.653598	3n
2	UDA_Chr13	7	15	0.533333	0.791111	0.761368	3n
2	UDA_Chr15	5	13	0.461538	0.757396	0.717622	3n
2	UDA_Chr17	8	15	0.333333	0.751111	0.725575	3n
2	UDA_Chr18	8	15	0.4	0.808889	0.784281	6n
2	UDA_Chr22	10	15	0.8	0.877778	0.865568	4n
2	UDA_Chr23A	5	14	0.357143	0.693878	0.636772	6n
2	UDA_Chr23B	3	13	0.461538	0.636095	0.558957	6n
2	UDA_Chr25	4	15	0.2	0.473333	0.436867	3n
2	UDA_Chr27	3	15	0.133333	0.56	0.4992	3n
2	UDA_Chr28	10	15	0.666667	0.866667	0.852252	6n
2	UDA_Chr29	10	12	0.666667	0.854167	0.837746	3n
	Average	7	14.15	0.482564	0.747284	0.716214	

k, the number of distinct alleles at a locus; **n**, the total number of individuals genotyped at a locus; **Ho**, the observed heterozygosity; **He**, the expected heterozygosity; **PIC**, the polymorphic information content.

Three phenetic similarity trees were generated using the UPGMA algorithm to assess the differentiation of genotypes across the three ploidy levels, diploid, tetraploid and hexaploidy, with the 20 microsatellite loci (Supplementary Fig.2 (diploid tree), Fig. 2 (tetraploid tree), Supplementary Fig.3 (hexaploidy tree). The goal was to assess the performance of the new SSR markers proposed for kiwifruit

Table 5
Polymorphism indexes of the 20 SSRs loci in the tetraploid *Actinidia* genotypes.

Ploidy	Locus_name	k	n	Ho	He	PIC	Motif
4	UDA_Chr1	12	51	0.5256	0.758495	0.726304	3n
4	UDA_Chr2	14	47	0.44708	0.807097	0.788685	3n
4	UDA_Chr3	8	51	0.598099	0.713664	0.67108	3n
4	UDA_Chr4	11	51	0.649531	0.690172	0.658301	3n
4	UDA_Chr6	12	51	0.471043	0.526068	0.510002	3n
4	UDA_Chr9	7	51	0.602411	0.597116	0.527348	3n
4	UDA_Chr10	8	46	0.476742	0.788579	0.762543	4n
4	UDA_Chr11	11	51	0.708713	0.766653	0.744748	6n
4	UDA_Chr12	7	51	0.674397	0.65267	0.589204	3n
4	UDA_Chr13	7	50	0.508307	0.55777	0.533221	3n
4	UDA_Chr15	10	51	0.499386	0.602545	0.575027	3n
4	UDA_Chr17	18	51	0.495822	0.843176	0.824401	3n
4	UDA_Chr18	11	50	0.741081	0.820569	0.798148	6n
4	UDA_Chr22	6	51	0.516008	0.590216	0.554572	4n
4	UDA_Chr23A	3	51	0.446943	0.431346	0.34656	6n
4	UDA_Chr23B	4	50	0.427508	0.653559	0.60169	6n
4	UDA_Chr25	5	51	0.28734	0.34192	0.320138	3n
4	UDA_Chr27	7	51	0.406889	0.497656	0.471391	3n
4	UDA_Chr28	13	51	0.837124	0.877854	0.865695	6n
4	UDA_Chr29	18	50	0.845175	0.909161	0.902213	3n
	Average	9.6	50.35	0.55826	0.671314	0.638564	

k, the number of distinct alleles at a locus; **n**, the total number of individuals genotyped at a locus; **Ho**, the observed heterozygosity; **He**, the expected heterozygosity; **PIC**, the polymorphic information content.

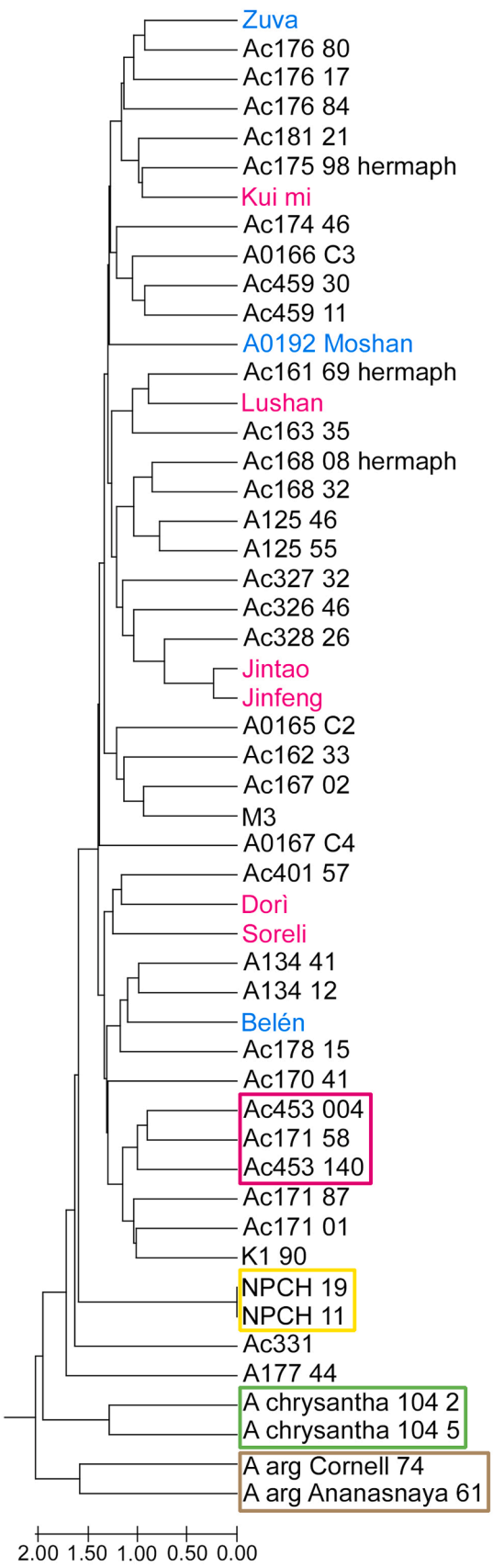
Table 6
Polymorphism indexes of the 20 SSRs loci in the hexaploid *Actinidia* genotypes.

Ploidy	Locus_name	k	n	Ho	He	PIC	Motif
6	UDA_Chr1	12	34	0.657715	0.776277	0.748986	3n
6	UDA_Chr2	11	34	0.702418	0.739172	0.697956	3n
6	UDA_Chr3	5	34	0.184296	0.22897	0.216036	3n
6	UDA_Chr4	5	34	0.403728	0.556271	0.50084	3n
6	UDA_Chr6	5	34	0.46705	0.454838	0.417026	3n
6	UDA_Chr9	6	34	0.611957	0.598038	0.520441	3n
6	UDA_Chr10	5	33	0.099706	0.260941	0.2404	4n
6	UDA_Chr11	8	34	0.868575	0.807282	0.779572	6n
6	UDA_Chr12	11	34	0.793915	0.751624	0.710308	3n
6	UDA_Chr13	6	34	0.625416	0.672484	0.619187	3n
6	UDA_Chr15	11	34	0.855097	0.841429	0.821952	3n
6	UDA_Chr17	11	34	0.371986	0.675547	0.636406	3n
6	UDA_Chr18	7	34	0.743336	0.744398	0.699273	6n
6	UDA_Chr22	6	34	0.468691	0.503399	0.466214	4n
6	UDA_Chr23A	3	34	0.446621	0.4411	0.3999	6n
6	UDA_Chr23B	5	34	0.416334	0.534067	0.495561	6n
6	UDA_Chr25	7	34	0.37293	0.39387	0.369148	3n
6	UDA_Chr27	5	34	0.557204	0.589851	0.534727	3n
6	UDA_Chr28	8	34	0.865156	0.815734	0.790789	6n
6	UDA_Chr29	14	34	0.70663	0.78093	0.758501	3n
	Average	7.55	33.95	0.560938	0.608311	0.571161	

k, the number of distinct alleles at a locus; **n**, the total number of individuals genotyped at a locus; **Ho**, the observed heterozygosity; **He**, the expected heterozygosity; **PIC**, the polymorphic information content.

fingerprinting, rather than to investigate relationships among genotypes. The 20 SSR loci were able to uniquely identify all the diploid genotypes. Moreover, the analysis included 15 samples of five different species (*A. chinensis, A. lanceolata, A. polygama, A. eriantha* and *A. latifolia*) to assess the transferability of the loci among *Actinidia* species. Consistent clustering is observed within specific groups; for instance, all the *A. chinensis* var. *rufopulpa* genotypes (A202_32, A201_32, A201_52, A201_66) were grouped in the same sub-cluster, and all genotypes of the same species clustered together. The diploid *A. chinensis* is known to be strictly related to the tetraploid *A. chinensis* var. *chinensis* and the hexaploidy *A. chinensis* var. *deliciosa* from which all the cultivated varieties were selected. The 51 tetraploid samples represent three different species: *A. arguta, A. chrysantha* and *A. chinensis* var *chinensis*. All the genotypes could be uniquely identified (Fig. 2) except for samples 'NPCH 11′ and 'NPCH 19′ which both originated from a colchicine-induced chromosome duplication of the same diploid *A. chinensis* var. *rufopulpa* genotype (De Mori et al., in press). It is well-established that SSR markers are not ideal for distinguishing closely related genotypes or clones. All commercial yellow flesh or bicolour varieties were selected from tetraploid *A. chinensis* genotypes. In these tetraploid genotypes, the grouping is consistent; for example, the two sister genotypes 'Ac453_004′ and 'Ac453_140′ are clustered with their mother plant 'Ac171_58′ (Fig. 2), from which they originated through crossing (Crosses made by Cipriani). The analysis included six commercial female varieties ('Jintao', 'Jinfeng', 'Dorì', 'Soreli', 'Lushan', 'Kui-mi') and three commercial male varieties ('Belen', 'Zuva', 'A0192 Moshan'), all of which were distinctly separated by the SSR fingerprinting analysis (Fig. 2). *A. arguta* and *A. chrysantha* genotypes were grouped and clustered separately from the main group of *A. chinensis* var. *chinensis* genotypes, indicating distinct clustering patterns among different species (Fig. 2).

All the hexaploidy genotypes are classified as *A. chinensis* var.



(caption on next column)

Fig. 2. UPGMA dendrogram of 51 tetraploid *Actinidia* genotypes. The six commercial female varieties are reported in red. The three commercial male varieties are reported in blue. The two genotypes, resulting from colchicine-induced chromosome duplication of the same diploid genotype of *A. chinensis* var. *rufopulpa*, are highlighted in the yellow box. The two sister genotypes 'Ac453_004′ and 'Ac453_140′, clustered with their mother plant 'Ac171_58′, are highlighted in the magenta box. The genotypes of *A. arguta* and *A. chrysantha*, grouped separately from the main cluster of *A. chinensis* var. *chinensis* genotypes, are highlighted in the brown and green boxes, respectively.

deliciosa, including the well-known green flesh 'Hayward' variety. All genotypes were uniquely identified, except for the samples 'Ecor m' and 'Ecor f'. 'Ecor f' is a female genotype that arose from a natural mutation of 'Ecor m', the male genotype (Testolin et al., 2004).

A. arguta var. *purpurea* was the only octoploid genotype included in the analysis. Though octoploid, no more than six alleles could be identified in this sample and, for the purpose of this work, it was analysed together with the hexaploidy samples.

4. Discussion

Although SNP markers are making significant advances in the molecular characterization of many important crop plants, microsatellite markers remain widely used. There are several reasons for this preference: the high costs of developing SNP markers in species with limited or no genomic resources for variant detection, the high level of polymorphism provided by SSRs, their widespread use in many laboratories, and the availability of databases for comparing electrophoretic profiles across different laboratories. In the human field, the use of SSRs is still prevalent due to the availability of data accumulated over years of analysis in criminal cases in which DNA analysis has constituted evidence to accuse or exonerate guilty suspects (Ruitberg et al., 2001). Indeed, the 20 short tandem repeat (STR) markers of the Combined DNA Index System (CODIS) are the basis of the vast majority of forensic genetics in the United States and other countries.

Microsatellites have been widely used in the molecular characterization of numerous fruit species with different purposes, such as varietal identification, population studies, and the determination of somatic stability during the vegetative propagation process (Cipriani et al., 2008; Dettori et al., 2015; Mohsenipoor et al., 2010; Nickravesh et al., 2023; Nybom and Schaal, 1990; Sadat-Hosseini et al., 2019).

For many years, kiwifruit cultivation was based on a green-fleshed variety, and only a little more than 20 years ago the first yellow-fleshed variety appeared on Western markets. The number of varieties available to fruit growers and consumers is increasing significantly and their molecular characterization to protect plant rights and to discover fraudulent uses of varieties is increasingly highly requested. There are examples of genetic characterization of the new kiwifruit variety determined using molecular DNA analysis, based on SSR markers (Mavromatis et al., 2010). The availability of large amounts of data relating to genomic DNA sequences obtained with NGS (Next Generation Sequencing) techniques allows for a rigorous selection of the best SSRs by considering various aspects related to the nature of these molecular markers.

In population genetics research, di-nucleotide microsatellites are commonly used, but they often generate stutter bands during analysis due to polymerase slippage (Guichoux et al., 2011). This complicates the identification of different alleles, especially in heterozygous individuals because the separation between neighbouring alleles is difficult to analysed and results in less reliable electropherogram interpretation and allele identification (Amos et al., 2007; Meldgaard and Morling, 1997). Various PCR amplification conditions, including the use of next-generation polymerases and modifications to temperature regimes, have been proposed to reduce stuttering (Guichoux et al., 2011; Olejniczak and Krzyzosiak, 2006; Seo et al., 2014). However, Ding et al. (2017) found in their preliminary studies that neither high-fidelity Pfu

DNA polymerase nor purified template DNA significantly reduced stuttering interference. Consequently, they concluded that the most effective approach to addressing the stuttering issue is the selection of suitable SSR motifs (Ding et al., 2017). Polymerase slippage is less likely with SSRs containing 3–5 nucleotides (Edwards et al. in 1991; Chambers and MacAvoy, 2000). Moreover, di-nucleotide repeats, when regular, have a 2-bp interval between adjacent alleles, and this short distance makes it challenging to accurately define the bins, that is the range of variation in the size of each allele (Testolin et al., 2023). In contrast, neighbouring alleles in tetra and penta-nucleotide SSRs are easier to differentiate than those in di-nucleotide SSRs. Selecting loci that have a significant number of repetitions is a fundamental condition in order to guarantee a particular degree of polymorphism in various individuals. SSRs that contain a large number of core motif repeats are, on average, characterized by a greater mutation rate. The selection of tetra-nucleotide microsatellites with a number of repeats ranging from 11 to 16 is something that van Asch and his colleagues advocate (Van Asch et al., 2010).

In this work we propose the use of a series of SSR markers that have been developed with the intent of reducing the problems of interpretation of electrophoretic profiles, using repeat sequences not less than three bases of the motif. In our experience, genotyping 100 *Actinidia* cultivars and selections using perfect microsatellite loci, the best compromise between two constraints, easy interpretable electrophoretic profiles, and sufficient polymorphism was obtained with 3n motif loci. Because the ploidy level of cultivars and selections was previously known, this provides a crosscheck for the accuracy of scoring alleles. The 19 primer pairs selected amplified 20 microsatellite loci of which 65 % had a trinucleotide motif, 10 % had a tetranucleotide motif and 25 % had a hexanucleotide motif. Although 15 different primer pairs were tested that amplified pentanucleotide-type loci no locus with a pentanucleotide motif was selected due to the ambiguous electrophoretic profiles, mainly for the presence of unexpected stuttering bands. During the microsatellite loci selection process, 75 % of loci were discarded due to issues such as failed amplification, poor amplification, or challenges in interpreting the electrophoretic profile using allele recognition software or with researcher assistance. The interpretation of the electrophoretic profiles was easier in the 15 diploid genotypes, where the appearance of secondary peaks was practically absent using any microsatellite with a repeat pattern greater than three (Supplementary Fig. 4). On the other hand, stuttering made it harder to interpret many profiles in the tetraploid and hexaploid genotypes. As a result, only 25 % of the loci were applicable across all ploidy levels. Failures in amplification and weak amplification products led to the exclusion of a high number of potential microsatellite loci useful for molecular fingerprinting. By adjusting the amplification conditions, it would have been possible to determine an optimal profile for each single locus, maximizing the number of markers usable in kiwifruit fingerprinting. However, it was decided to prioritize the uniqueness of the amplification profile, in order to analyze all the loci using the same method.

Numerous studies on plant genomes have highlighted that SSRs are uniformly distributed within chromosomes (Li et al., 2004; Ramsay et al., 2000; Tang et al., 2002). In plants, it is generally expected to find at least one SSR locus in every 10 kb of DNA sequence (Tautz, 1989). Cavagnaro et al. (2010) summarized the content of perfect microsatellites in the genomic sequences of cucumber and seven other plant species. The density of microsatellites was found to be as high as 428.5 per Mb, on average, for eight selected plant species (Cavagnaro et al., 2010). This trend was also confirmed in the genotype of *A. chinensis* var *chinensis* 'Red5' (Supplementary Fig. 1). The average density of microsatellites is 394.89 SSRs/Mbp, which aligns well with observations in other species. There were no large differences between the chromosomes in the relative frequencies of the microsatellite loci that were found using chi-square tests. Di-nucleotide SSRs were the most common, indicating that di-, tri-, tetra-, penta-, and hexa-nucleotide core repeats are generally less abundant, as reported in previous studies on *Vitis* (Cipriani et al., 2008) and peach (Ding et a.., 2017). The results obtained are consistent with those of several studies which have highlighted the greater abundance of TA and GA repeats in plant genomes (Morgante et al., 2002; Tóth et al., 2000). The classes of microsatellites increase with the repeating base motif and, among the most numerous classes of the tri-, tetra-, penta- and hexanucleotide microsatellites, the core repeats AAT, AAAT, AAAAT, and AAAAAT were the most frequent. Cipriani et al. (2008) obtained similar results in a study that aimed to identify a set of SRRs to minimize genotyping errors in grapevine (*Vitis spp*.). Although data on these classes of SSRs are scarce in the literature, analyses on the 'Red5' genotype revealed that as the repeated motif length increases, the nucleotide composition shows a pattern of repeated adenine bases followed by a single thymine base.

A variety of parameters were used to evaluate marker informativeness. The most fundamental criterion is the number of alleles; markers with more alleles are more likely to be polymorphic for a particular set of genotypes. A more precise indicator of polymorphism is expected heterozygosity, which assesses how those alleles are distributed throughout the germplasm under study (Jones et al., 2007). Analysis of 48 varieties and selections of *A. chinensis* var. *chinensis* and *A. chinensis* var. *deliciosa* with di- and tri- microsatellite loci identified a mean number of alleles between 2.6 and 3.5 and observed heterozygosity between 0.546 and 0.671, respectively (Zhen et al., 2004). We looked at 100 cultivars and selections with the new 20 SSR loci set and found an average of 8.05 alleles per locus. The number of alleles ranged from 7 to 9.65 in diploid and tetraploid *A. chinensis* var. *chinensis*. The observed heterozygosity ranged from 0.48 to 0.56 (average 0.51) in diploid *A. chinensis* var. *chinensis* and in hexaploid *A. chinensis* var. *deliciosa*, respectively. The new set of 20 microsatellite loci was therefore comparable in terms of defining the genetic differences present in the *Actinidia* genotypes to the one used previously. Molecular markers, including microsatellite markers, have been valuable in various population genetics studies, identifying varieties, and establishing core germplasm collections to enhance genetic diversity in on-site repositories, leading to space and cost savings in maintaining large plants like kiwifruit (Hu et al., 2022).

5. Conclusions

Molecular markers are widely used to determine the DNA molecular profile of plants and animals. This study presents a novel set of 20 microsatellite markers developed from the *Actinidia chinensis* var. *chinensis*'Red5′ genome, which are distributed across the 29 linkage groups of the diploid kiwifruit genome. These markers, primarily composed of trinucleotide, tetranucleotide, and hexanucleotide motifs, have been optimized to minimize the interpretative challenges often associated with dinucleotide markers, such as stuttering, and generate optimal genotypic fingerprinting profiles for unique characterization of Actinidia germplasm.

The newly developed markers were tested on 100 genotypes of kiwifruit including varieties and selections, successfully distinguishing all genotypes except for two clonal mutations, demonstrating their robustness and applicability for varietal characterization. These SSR markers are easily interpretable and amplify consistently using a single protocol across different loci, making them highly suitable for kiwifruit fingerprinting.

Furthermore, these markers will be valuable not only for breeders seeking to protect intellectual property but also for conservation efforts aiming to maintain kiwifruit biodiversity. Their high level of polymorphism and cross-species applicability also make them suitable for population studies and the creation of core collections, contributing to the efficient management and conservation of kiwifruit germplasm.

Funding

De Mori G. Grant PON 2014–2020 AZIONE IV.6 GREEN.

Statement

During the preparation of this work, the authors used Quillbot Service in order to improve the readability and language of the manuscript. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

CRediT authorship contribution statement

Maria Teresa Marrazzo: Investigation, Formal analysis. **Andrea Amilcare Passerelli:** Investigation. **Guido Cipriani:** Writing – original draft, Funding acquisition, Conceptualization. **Gloria De Mori:** Writing – review & editing, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Gloria De Mori reports financial support was provided by University of Udine Department of Agricultural Food Environmental and Animal Sciences. Maria Teresa Marrazzo reports financial support was provided by University of Udine Department of Agricultural Food Environmental and Animal Sciences. Guido Cipriani reports financial support was provided by University of Udine Department of Agricultural Food Environmental and Animal Sciences. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors thank Giorgio Comuzzo for maintaining the kiwifruit germplasm repository at the Azienda Agraria Sperimentale A. Servadei of the University of Udine and for assistance in preparing the figures. The authors would like to thank the anonymous reviewers who helped improve the manuscript.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.scienta.2024.113775.

Data availability

Data will be made available on request.

References

Amos, W., Hoffman, J.I., Frodsham, A., Zhang, L., Best, S., Hill, A.V.S., 2007. Automated binning of microsatellite alleles: problems and solutions. Mol. Ecol. Notes. 7, 10–14. https://doi.org/10.1111/J.1471-8286.2006.01560.X.

Archak, S., 2000. Plant DNA fingerprinting: an overview. AgBiotechNet 2.

Bhargava, A., Fuentes, F.F., 2009. Mutational dynamics of microsatellites. Mol. Biotechnol. 2009 44 (3), 250–266. https://doi.org/10.1007/S12033-009-9230-4, 44.

Cavagnaro, P.F., Senalik, D.A., Yang, L., Simon, P.W., Harkins, T.T., Kodira, C.D., Huang, S., Weng, Y., 2010. Genome-wide characterization of simple sequence repeats in cucumber (Cucumis sativus L.). BMC. Genomic. 11, 1–18. https://doi.org/10.1186/1471-2164-11-569/FIGURES/5.

Chambers, G.K., MacAvoy, E.S., 2000. Microsatellites: consensus and controversy. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 126, 455–476. https://doi.org/10.1016/S0305-0491(00)00233-9.

Cipriani, G., Marrazzo, M.T., Di Gaspero, G., Pfeiffer, A., Morgante, M., Testolin, R., 2008. A set of microsatellite markers with long core repeat optimized for grape (*Vitis* spp.) genotyping. BMC. Plant Biol. 8, 1–13. https://doi.org/10.1186/1471-2229-8-127/FIGURES/3.

De Mori, G., Montefiori, M., Santi, G., Navacchi, O., Cipriani, G., In press. Chromosomal duplication in *Actinidia chinensis* var. *Rufopulpa*. Acta Hortic.

Dettori, M.T., Micali, S., Giovinazzi, J., Scalabrin, S., Verde, I., Cipriani, G., 2015. Mining microsatellites in the peach genome: development of new long-core SSR markers for genetic analyses in five Prunus species. Springerplus. 4, 1–18. https://doi.org/10.1186/S40064-015-1098-0/FIGURES/3.

Ding, M., Wang, K., Wang, W., Chen, M., Wu, D., Xu, C., Chen, K., 2017. Development of high quality EST-SSR markers without stutter bands in peach and their application in cultivar discrimination and hybrid authentication. HortScience 52, 24–30. https://doi.org/10.21273/HORTSCI11314-16.

Edwards, A., Civitello, A., Hammond, H.A., Caskey, T., 1991. DNA typing and genetic mapping with Trimeric and Tetrameric tandem repeats. Am. J. Hum. Genet 49, 746–756.

Ellegren, H., 2004. Microsatellites: simple sequences with complex evolution. Nature Rev. Genetic. 2004 5 (6), 435–445. https://doi.org/10.1038/nrg1348, 5.

Ferguson, A.R., 2004. 1904—The year that kiwifruit (*Actinidia deliciosa*) came to New Zealand. N Z J Crop Hortic Sci 32, 3–27. https://doi.org/10.1080/01140671.2004.9514276.

Ferguson, A.R., Zhang, J.-L., Duffy, A.M., Beatson, R.A., Cheng, C.-H., Datson, P.M., Lowe, R.G., McNeilage, M.A., Harris-Virgin, P.M., Seal, A.G., Wu, J.-H., 2009. Ploidy and the use of flow cytometry in kiwifruit breeding. Italus Hortus 16 (5), 78–83.

Ferguson, R.A., Huang, H., 2007. Genetic Resources of Kiwifruit: domestication and Breeding. Hortic. Rev. (Am. Soc. Hortic. Sci) 33, 1–121. https://doi.org/10.1002/9780470168011.CH1.

Gel, B., Serra, E., 2017. KaryoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. Bioinformatics. 33, 3088–3090. https://doi.org/10.1093/BIOINFORMATICS/BTX346.

Guichoux, E., Lagache, L., Wagner, S., Chaumeil, P., Léger, P., Lepais, O., Lepoittevin, C., Malausa, T., Revardel, E., Salin, F., Petit, R.J., 2011. Current trends in microsatellite genotyping. Mol. Ecol. Resour. 11, 591–611. https://doi.org/10.1111/J.1755-0998.2011.03014.X.

Han, X., Zhang, Y., Zhang, Q., Ma, N., Liu, X., Tao, W., Lou, Z., Zhong, C., Deng, X.W., Li, D., He, H., 2023. Two haplotype-resolved, gap-free genome assemblies for *Actinidia latifolia* and *Actinidia chinensis* shed light on the regulatory mechanisms of vitamin C and sucrose metabolism in kiwifruit. Mol. Plant 16, 452–470. https://doi.org/10.1016/j.molp.2022.12.022.

Huang, H., 2016. Systematics and Genetic Variation of *Actinidia*. Kiwifruit. Academic Press. https://doi.org/10.1016/B978-0-12-803066-0.00001-0.

Huang, K., Dunn, D.W., Ritland, K., Li, B., 2020. polygene: population genetics analyses for autopolyploids based on allelic phenotypes. Methods Ecol. Evol. 11, 448–456. https://doi.org/10.1111/2041-210X.13338.

Hu, G., Jiang, Q., Wang, Z., Li, Z., Liao, W., Shen, D., Zhong, C., 2022. Genetic diversity analysis and core collection construction of the *Actinidia chinensis* complex (Kiwifruit) based on SSR markers. Agronomy 2022 12 (12), 3078. https://doi.org/10.3390/AGRONOMY12123078. Page3078.

Jones, E.S., Sullivan, H., Bhattramakki, D., Smith, J.S.C., 2007. A comparison of simple sequence repeat and single nucleotide polymorphism marker technologies for the genotypic analysis of maize (*Zea mays* L.). Theor. Appl. Genetic. 115, 361–371. https://doi.org/10.1007/S00122-007-0570-9/FIGURES/4.

Kalia, Rajwant K, Rai, Manoj K, Kalia, Sanjay, Singh, R., Dhawan, •.A.K., Kalia, R.K., Rai, M.K., Singh, Á.R., Kalia, S., Dhawan, A.K., 2010. Microsatellite markers: an overview of the recent progress in plants. Euphytica 2010 177 (3), 309–334. https://doi.org/10.1007/S10681-010-0286-9, 177.

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T., Carey, V.J., 2013. Software for computing and annotating genomic ranges. PLoS. Comput. Biol. 9. https://doi.org/10.1371/JOURNAL.PCBI.1003118.

Li, J.Q., Li, X.W., Soejarto, D.D., 2007. A Revision of the genus *actinidia* from China. Acta Hortic. 753, 41–44. https://doi.org/10.17660/ACTAHORTIC.2007.753.2.

Litt, M., Luty, J.A., 1989. A Hypervariable microsatellite revealed by In Vitro amplification of a dinucleotide repeat within the cardiac muscle actin Gene. Am. J. Hum. Genet 44, 397–401.

Li, Y.C., Korol, A.B., Fahima, T., Nevo, E., 2004. Microsatellites within genes: structure, function, and evolution. Mol. Biol. Evol. 21, 991–1007. https://doi.org/10.1093/MOLBEV/MSH073.

Mavromatis, A.G., Arvanitoyannis, I., Nanos, G., Sakellariou, M., Ilanidis, C., Korkovelos, A., 2010. Molecular fingerprinting of a new Kiwifruit cultivar (cv. Tsehelidis) and comparative analysis with cv. Hayward according to physicochemical properties. Sci. Hortic. 125, 277–282. https://doi.org/10.1016/J.SCIENTA.2010.03.010.

Meldgaard, M., Morling, N., 1997. Detection and quantitative characterization of artificial extra peaks following polymerase chain reaction amplification of 14 short tandem repeat systems used in forensic investigations. Electrophoresis 18, 1928–1935. https://doi.org/10.1002/ELPS.1150181107.

Mohsenipoor, S., Vahdati, K., Amiri, R., Mozaffari, M.R., 2010. Study of the genetic structure and gene flow in Persian walnut (Juglans regia L.) using SSR markers. Acta Hortic. 861, 133–142. https://doi.org/10.17660/ACTAHORTIC.2010.861.17.

Morgante, M., Hanafey, M., Powell, W., 2002. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. Nature Genetic. 2002 30 (2), 194–200. https://doi.org/10.1038/ng822, 30.

Nickravesh, M.H., Vahdati, K., Amini, F., Di Pierro, E.A., Amiri, R., Woeste, K., Arab, M. M., 2023. Reliable propagation of persian walnut varieties using SSR Marker-based True-to-type validation. HortScience 58, 64–66. https://doi.org/10.21273/HORTSCI16851-22.

Nybom, H., Schaal, B.A., 1990. DNA "fingerprints" applied to paternity analysis in apples (Malus x domestica). Theoretic. Appl. Genetic. 79, 763–768. https://doi.org/10.1007/BF00224242/METRICS.

Olejniczak, M., Krzyzosiak, W.J., 2006. Genotyping of simple sequence repeat factors implicated in shadow band generation revisited. Electrophoresis 27, 3724–3734. https://doi.org/10.1002/ELPS.200600136.

Parida, S.K., Pandit, A., Gaikwad, K., Sharma, T.R., Srivastava, P.S., Singh, N.K., Mohapatra, T., 2010. Functionally relevant microsatellites in sugarcane unigenes. BMC. Plant Biol. 10, 1–19. https://doi.org/10.1186/1471-2229-10-251/TABLES/3.

Pilkington, S.M., Crowhurst, R., Hilario, E., Nardozza, S., Fraser, L., Peng, Y., Gunaseelan, K., Simpson, R., Tahir, J., Deroles, S.C., Templeton, K., Luo, Z., Davy, M., Cheng, C., McNeilage, M., Scaglione, D., Liu, Y., Zhang, Q., Datson, P., De Silva, N., Gardiner, S.E., Bassett, H., Chagné, D., McCallum, J., Dzierzon, H., Deng, C., Wang, Y.Y., Barron, L., Manako, K., Bowen, J., Foster, T.M., Erridge, Z.A., Tiffin, H., Waite, C.N., Davies, K.M., Grierson, E.P., Laing, W.A., Kirk, R., Chen, X., Wood, M., Montefiori, M., Brummell, D.A., Schwinn, K.E., Catanach, A., Fullerton, C., Li, D., Meiyalaghan, S., Nieuwenhuizen, N., Read, N., Prakash, R., Hunter, D., Zhang, H., McKenzie, M., Knäbel, M., Harris, A., Allan, A.C., Gleave, A., Chen, A., Janssen, B.J., Plunkett, B., Ampomah-Dwamena, C., Voogd, C., Leif, D., Lafferty, D., Souleyre, E.J.F., Varkonyi-Gasic, E., Gambi, F., Hanley, J., Yao, J.L., Cheung, J., David, K.M., Warren, B., Marsh, K., Snowden, K.C., Lin-Wang, K., Brian, L., Martinez-Sanchez, M., Wang, M., Ileperuma, N., Macnee, N., Campin, R., McAtee, P., Drummond, R.S.M., Espley, R.V., Ireland, H.S., Wu, R., Atkinson, R.G., Karunairetnam, S., Bulley, S., Chunkath, S., Hanley, Z., Storey, R., Thrimawithana, A.H., Thomson, S., David, C., Testolin, R., Huang, H., Hellens, R.P., Schaffer, R.J., 2018. A manually annotated *Actinidia chinensis* var. *chinensis* (kiwifruit) genome highlights the challenges associated with draft genomes and gene prediction in plants. BMC. Genomics. 19, 1–19. https://doi.org/10.1186/s12864-018-4656-3.

Pinto, T., Vilela, A., 2018. Kiwifruit, a botany, chemical and sensory approach a review. Adv. Plants Agric. Res. Volume 8. https://doi.org/10.15406/APAR.2018.08.00355.

Ramsay, L., Macaulay, M., Degli Ivanissevich, S., MacLean, K., Cardle, L., Fuller, J., Edwards, K.J., Tuvesson, S., Morgante, M., Massari, A., Maestri, E., Marmiroli, N., Sjakste, T., Ganal, M., Powell, W., Waugh, R., 2000. A simple sequence repeat-based linkage map of barley. Genetics 156, 1997–2005. https://doi.org/10.1093/GENETICS/156.4.1997.

Richard, G.-F., Kerrest, A., Dujon, B., 2008. Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. Microbiol. Mol. Biol. Rev. 72, 686–727. https://doi.org/10.1128/MMBR.00011-08.

Ruitberg, C.M., Reeder, D.J., Butler, J.M., 2001. STRBase: a short tandem repeat DNA database for the human identity testing community. Nucleic. Acids. Res. 29, 320–322. https://doi.org/10.1093/NAR/29.1.320.

Sadat-Hosseini, M., Vahdati, K., Leslie, C.A., 2019. Germination of persian walnut somatic embryos and evaluation of their genetic stability by ISSR fingerprinting and flow cytometry. HortScience 54, 1576–1580. https://doi.org/10.21273/HORTSCI13787-18.

Schuelke, M., 2000. An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology 2000 18 (2), 233–234. https://doi.org/10.1038/72708, 18.

Seo, S.B., Ge, J., King, J.L., Budowle, B., 2014. Reduction of stutter ratios in short tandem repeat loci typing of low copy number DNA samples. Forensic Sci. Int. Genet. 8, 213–218. https://doi.org/10.1016/J.FSIGEN.2013.10.004.

Statistics | FAO | Food and Agriculture Organization of the United Nations, https://www.fao.org/statistics/en/(accessed 8.7.24).

Tamura, K., Stecher, G., Kumar, S., 2021. MEGA11: molecular evolutionary genetics analysis version 11. Mol. Biol. Evol. 38, 3022–3027. https://doi.org/10.1093/MOLBEV/MSAB120.

Tang, S., Yu, J.K., Slabaugh, M.B., Shintani, D.K., Knapp, S.J., 2002. Simple sequence repeat map of the sunflower genome. Theor. Appl. Genetic. 105, 1124–1136. https://doi.org/10.1007/S00122-002-0989-Y/METRICS.

Tautz, D., 1993. Notes on the definition and nomenclature of tandemly repetitive DNA sequences. EXS. Birkhäuser 21–28. https://doi.org/10.1007/978-3-0348-8583-6_2.

Tautz, D., 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic. Acids. Res. 17, 6463–6471. https://doi.org/10.1093/NAR/17.16.6463.

Testolin, R., Messina, R., Cipriani, G., De Mori, G., 2023. SSR-based DNA fingerprinting of fruit crops. Crop. Sci. 63, 390–459. https://doi.org/10.1002/CSC2.20896.

Testolin, R., Messina, R., Lain, O., Cipriani, G., 2004. A natural sex mutant in kiwifruit (*Actinidia deliciosa*). N Z J. Crop. Hortic. Sci. 32, 179–183. https://doi.org/10.1080/01140671.2004.9514294.

Tóth, G., Gáspári, Z., Jurka, J., 2000. Microsatellites in different eukaryotic genomes: survey and analysis. Genome Res. 10, 967–981. https://doi.org/10.1101/GR.10.7.967.

Van Asch, B., Pinheiro, R., Pereira, R., Alves, C., Pereira, V., Pereira, F., Gusmão, L., Amorim, A., 2010. A framework for the development of STR genotyping in domestic animal species: characterization and population study of 12 canine X-chromosome loci. Electrophoresis 31, 303–308. https://doi.org/10.1002/ELPS.200900389.

Wang, H., Yang, Y., Schofield, M.J., Du, C., Fridman, Y., Lee, S.D., Larson, E.D., Drummond, J.T., Alani, E., Hsieh, P., Erie, D.A., 2003. DNA bending and unbending by MutS govern mismatch recognition and specificity. Proc. Natl. Acad. Sci. U.S.A 100, 14822–14827. https://doi.org/10.1073/PNAS.2433654100/SUPPL_FILE/3654FIG8.PDF.

Wang, X., Wang, L., 2016. GMATA: an integrated software package for genome-scale SSR mining, marker development and viewing. Front. Plant Sci. 7, 215951. https://doi.org/10.3389/FPLS.2016.01350/BIBTEX.

Wickham, H., 2016. ggplot2, Use R! Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-24277-4.

Wu, B., Yu, Q., Deng, Z., Duan, Y., Luo, F., Gmitter, F., 2023. A chromosome-level phased genome enabling allele-level studies in sweet orange: a case study on citrus Huanglongbing tolerance. Hortic. Res. 10. https://doi.org/10.1093/HR/UHAC247.

Wu, J., 2019. Cultivar, environment and integration of cultural practices will determine the future of the kiwifruit industry. Scripta Hortic. 20, 171–178.

Yao, X., Wang, S., Wang, Z., Li, D., Jiang, Q., Zhang, Q., Gao, L., Zhong, C., Huang, H., Liu, Y., 2022. The genome sequencing and comparative analysis of a wild kiwifruit *Actinidia eriantha*. Molecular Horticulture 2, 1–14. https://doi.org/10.1186/S43897-022-00034-Z/FIGURES/4.

Yue, J., Chen, Q., Wang, Yingzhen, Zhang, L., Ye, C., Wang, X., Cao, S., Lin, Y., Huang, W., Xian, H., Qin, H., Wang, Yanli, Zhang, S., Wu, Y., Wang, S., Yue, Y., Liu, Y., 2023. Telomere-to-telomere and gap-free reference genome assembly of the kiwifruit *Actinidia chinensis*. Hortic. Res. 10. https://doi.org/10.1093/HR/UHAC264.

Zane, L., Bargelloni, L., Patarnello, T., 2002. Strategies for microsatellite isolation: a review. Mol. Ecol. 11, 1–16. https://doi.org/10.1046/J.0962-1083.2001.01418.X.

Zhen, Y., Li, Z., Huang, H., Wang, Y., 2004. Molecular characterization of kiwifruit (*Actinidia*) cultivars and selections using SSR markers. J. Am. Soc. Hortic. Sci. 129, 374–382. https://doi.org/10.21273/JASHS.129.3.0374.