GMR
The Original by FUNPEC-RP

In silico microsatellite transferability from *Psidium guajava* to *Eucalyptus globulus* validated by PCR

F.L.B. Medeiros[1], C.A.F. Santos[2] and A.E.S. Costa[1]

[1] Universidade Federal Rural de Pernambuco, Departamento de Agronomia, Recife, PE, Brasil
[2] Embrapa Semiárido, Petrolina, PE, Brasil

Corresponding author: C.A.F. Santos
E-mail: carlos-fernandes.santos@embrapa.br

Genet. Mol. Res. 20 (4): gmr18985
Received October 26, 2021
Accepted November 28, 2021
Published December 28, 2021
DOI http://dx.doi.org/10.4238/gmr18985

ABSTRACT. BLAST is a genomic local alignment search tool used to identify homology between genotypes and possible orthologous genes. *In vitro* microsatellite transferability is a strategy to enable or increase species molecular fingerprinting, but it is dependent on PCR technique. An initial *in silico* step using BLAST for transferability can be helpful to save resources in pre-selecting markers more likely to amplify. We aligned and transfered SSR sequences from *Psidium guajava* to *Eucalyptus globulus* using BLAST. Twenty-three SSR clone sequences from *P. guajava* (query) were retrieved from the NCBI website and aligned against the whole genome of *E. globulus* (subject) using a cut-off e-value<$1.00e^{-20}$. Another 140 loci retrieved from the GuavaMap project were analyzed using as parameters e-values<1.7 and a maximum distance of 300 nucleotides between forward and reverse sequences. All loci were analyzed using BLASTN with MEGABLAST optimization. DNA extraction of four eucalypt trees was performed with the 2x CTAB protocol containing a sorbitol initial step. Validation of the SSR selected via BLASTN was performed by PCR reactions with 12 loci (seven selected on Blast hits and five without hits) and posterior visualization on polyacrylamide gel. Nine out of 23 microsatellite loci were transferable in the *in silico*, with a mean identity of 87%. With regard to the GuavaMap microsatellite loci, only three showed significant

alignments, among the 140 tested, with the forward and reverse mean identity of 100% and 95%, respectively. All seven SSR with e-values<$1.00e^{-20}$ (mPgCIR001, mPgCIR005, mPgCIR007, mPgCIR009, mPgCIR018, mPgCIR020, and mPgCIR026) showed easy-to-score amplicons on the polyacrylamide gel when using the *in silico* transferability strategy. However, the other five, without significant e-values or hits, showed no amplification. These results highlight the effectiveness of *in silico* transferability for full-length SSR loci, constituting a valid alternative to save time and costs in transferability studies between species.

Key words: BLAST; Guava; Identity; SSR

INTRODUCTION

Simple sequence repeats (SSR) or microsatellites are widely applied in different biology and molecular genetic studies due to their abundance in the genome, codominance, and high level of information (Carrer et al., 2010). Rafalski et al. (1996) were among the first to propose a method for microsatellite identification and primer design that included constructing a genomic library, hybridization screening, determining clone-positive DNA sequences, primer design, PCR analysis, and polymorphism identification. Although it is an expensive and laborious technique (Varshney et al., 2002; 2005), advances in next-generation sequencing allow the faster identification of many microsatellites (Taheri et al., 2018).

In cases where a genomic library for a single species is not available, genetic variability, genetic mapping, fingerprinting, and comparative genomic studies have been made possible by SSR transferability from one species to another (Singleton *et al*., 2020). These markers are conserved DNA regions between closely related species and, therefore, helpful in identifying synteny and collinearity (Shepherd et al., 2006). In this scenario, Fagundes et al. (2016) applied marker transferability to study the biodiversity of different Myrtaceae species from the Atlantic Forest. With regard to the family Myrtaceae, several transferability studies have been conducted with *Eucalyptus* (Santos et al., 2007; Grattapaglia et al., 2011; Hudson et al., 2012; Butler et al., 2017) due to its economic importance and the significant investment in its genomic resources. With the transferability assumption that SSR are conserved among related species, Tuler et al. (2015) proposed a set of such markers for species identification in *Psidium*.

Successful transferability studies have also been developed with this family by Rai et al. (2013) and Sitther et al. (2014). Zou et al. (2020) discussed the obstacles of working with molecular markers, such as null alleles and the need to validate data through PCR reactions. From this perspective, bioinformatics tools, such as local alignment, have become valuable for comparative genomics and optimizing projects.

The basic alignment search tool (BLAST) has been applied in homology searches between species and to identify possible orthologous genes. For example, Varshney et al. (2005) used BLAST to estimate the theoretical transferability of EST-SSR markers mapped in barley against EST-SSR from wheat, rye, rice, maize, sorghum, *Arabidopsis*, and *Medicago*, considering significant the alignments with e-values < $1.00e^{-10}$. For wheat, rye, and rice, the authors found significant homology values of 93.5, 37.3, and 57.3%,

respectively, and estimated the location of putative orthologous regions by comparing physically mapped EST-SRR in barley against other cereal species.

A similar study was developed with the family Myrtaceae by Acuña et al. (2012), in which the authors used BLASTX to attempt to predict the protein function of EST-SSR discovered in *Eucalyptus* and estimate the identities of those regions using a cut-off value of $\leq 1.00e^{-5}$. Sumathi et al. (2018) used BLAST to anchor microsatellites mapped from the *E. tereticornis* × *E. camaldulensis* hybrid into the genome of *E. grandis* and observed high synteny levels, in addition to a 64% collinearity between the compared genomes. From this perspective, this study aimed to apply BLAST to align SSR sequences from *P. guajava* for *in silico* inferences in chromosomes of *E. globulus* to guide the *in vitro* transferability for genomic studies with these two species.

MATERIAL AND METHODS

In silico transferability

Twenty-three SSR clone sequences developed by Risterucci et al. (2005) for *P. guajava* were retrieved from the National Center for Biotechnology Information (NCBI) website and used as queries (Table 1). The whole *E. globulus* genome (subject) (Ferguson et al., 2020), also available on the NCBI website and assembled by chromosomes, was used to search for SSR sequences identical to *P. guajava*. At the NBCI website, the BLASTN algorithm (Zhang et al., 2000) with the MEGABLAST optimization was applied against the *E. globulus* genome in order to locate highly similar sequences. After that, alignments with a cut-off e-value < $1.00e^{-20}$ were selected.

Table 1. SSR locus, EMBL accession number, clone size, and allele size of 23 SSR sequences from Psidium guajava reported by Risterucci et al. (2005).

SSR locus	EMBL accession number	Clone size	Allele size range
mPgCIR01	AJ639775	237	236–250
mPgCIR02	AJ639753	224	202–230
mPgCIR03	AJ639754	158	118–164
mPgCIR04	AJ639755	148	126-150
mPgCIR05	AJ639756	252	224–280
mPgCIR07	AJ639757	149	148–160
mPgCIR08	AJ639758	214	210–224
mPgCIR09	AJ639759	173	156–176
mPgCIR10	AJ639760	261	262–320
mPgCIR11	AJ639761	298	298–314
mPgCIR13	AJ639762	245	240–260
mPgCIR14	AJ639763	185	184–186
mPgCIR15	AJ639764	147	144–172
mPgCIR16	AJ639765	292	268–296
mPgCIR17	AJ639766	231	230–240
mPgCIR18	AJ639767	195	192–204
mPgCIR19	AJ639768	274	258–280
mPgCIR20	AJ639769	266	270–298
mPgCIR21	AJ639770	154	150–164
mPgCIR22	AJ639771	235	236–252
mPgCIR23	AJ639772	185	184–198
mPgCIR25	AJ639773	124	104–130
mPgCIR26	AJ639774	185	180–198

Also, 140 loci made available by the GuavaMap international consortium were used in the BLAST analysis. Short-length sequences were used in the FASTA format, corresponding to the two genomic regions (forward and reverse) and aiming to locate highly similar sequences. Primers were selected based on the score, e-value<1.7, and the distance between forward and reverse primers (maximum of 300 nucleotides considering the subject's start and end).

DNA extraction and quantification

Healthy leaves from four *Eucalyptus sp* plants, available at Embrapa Semiárido campus, were extracted according to the protocol proposed by Inglis et al. (2018), consisting of pre-washing the samples with sorbitol. The 2x CTAB protocol for DNA extraction was applied as described by Costa and Santos (2013). DNA quantification was performed by spectrophotometry using a microdrop plate in the absorbance range of 260 nm. A regression curve was adjusted based on the absorbance values of DNA samples with known concentrations. DNA purity was determined based on the A260/280 ratio. Samples with values close to or higher than 1.8 were used; otherwise, they were re-extracted. After quantification, samples were diluted to the concentration of 60 ng/ µL.

In vitro transferability

Seven loci were selected based on hits in the BLASTN analysis, and another five loci without hits were also selected. Both sets of primers were analyzed in polyacrylamide gel. PCR reactions were carried out for a final 20 µL volume according to the protocol described by Rai et al. (2013), with some modifications: 60ng DNA, 1X Taq DNA polymerase buffer, 2.5 mM MgCl2, 2.5 mM dNTP's, 0.8 µM of each primer, and 1 U enzyme Taq DNA polymerase. Thermocycler programming consisted of an initial denaturation step at 94ºC for 4 min followed by 30 cycles of 94ºC for 45 s, 52ºC for 60 s, 72ºC for 60 s, and a final extension step of 72ºC for 8 min. PCR amplicons after polyacrylamide gel electrophoresis were obtained according to Costa and Santos (2013).

RESULTS

In silico SSR transferability from *P. guajava* to *E. globulus* chromosomes

Among the 23 loci retrieved from Risterucci et al. (2005), nine showed significant alignments (Table 2). High scores ranging from 141 (mPgCIR021) to 435 (mPgCIR007) were observed. All primers showed low e-values, ranging from $3.00e^{-32}$ (mPgCIR0021) to $2.00e^{-120}$ (mPgCIR007). Primer mPgCIR018 showed the highest identity (92.91%) among all loci. The referred loci showed identities greater than 80%, except locus mPgCIR021, with a mean identity value of 87%.

With regard to the 140 SSR loci developed by GuavaMap, only three (mPgCIR261, mPgCIR313, mPgCIR378) showed significant alignments. The mean identity was 100% in the forward sequences and 95.08% in the reverse sequences (Table 3). Markers mPgCIR261, mPgCIR313, and mPgCIR378 showed scores of 44.1, 40.1, and 40.1 (forward

sequences) and 32.2, 34.2, and 32.2 (reverse sequences), respectively. Alignments showed e-values ranging from 0.19 to 0.0002 (Table 3). BLAST-positioned primers mPgCIR261, mPgCIR313, and mPgCIR378 distant 274, 62, and 86 nucleotides, respectively, from the forward and reverse sequences (Table 3), which may result in effective transferability in PCR analysis due to their small distance in the *E. globulus* genome.

Table 2. Score, e-value, and identity of nine SSR sequences (Risterucci et al., 2005) from Psidium guajava to Eucalyptus globulus chromosomes.

Locus from *P. guajava*	Chromosome from *E. globulus*	Score	e-value	Identity (%)
mPgCIR001	6	339	$2.00e^{-91}$	83.64
mPgCIR005	1	259	$2.00e^{-67}$	84.59
mPgCIR007	10	435	$2.00e^{-120}$	91.35
mPgCIR009	4	274	$6.00e^{-72}$	83.08
mPgCIR011	8	372	$3.00e^{-101}$	90.71
mPgCIR018	3	396	$9.00e^{-109}$	92.91
mPgCIR020	11	289	$2.00e^{-76}$	90.83
mPgCIR021	8	141	$3.00e^{-32}$	79.91
mPgCIR026	4	326	$1.00e^{-87}$	83.83

Table 3. Score, e-value, identity (ID), and distance between primers (bp) in three Psidium guajava SSR sequences retrieved from the GuavaMap project in chromosomes (Chr). of Eucalyptus globules.

Marker	Accession number	Chr	Score	e-value	ID (%)	Distance between primers (bp)
mPgCIR261	Unpublished	8	44.1	$2.00e^{-04}$	100	274
mPgCIR313	Unpublished	1	40.1	$2.00e^{-03}$	100	62
mPgCIR378	Unpublished	2	40.1	$2.00e^{-03}$	100	86
mPgCIR261	Unpublished	8	32.2	$4.40e^{-01}$	95	274
mPgCIR313	Unpublished	1	34.2	$1.90e^{-01}$	95	62
mPgCIR378	Unpublished	2	32.2	$6.60e^{-01}$	95	86

In vitro SSR transferability from *P. guajava* to *Eucalyptus sp.* chromosomes

All seven loci with e-values<$3.00xe^{-32}$ showed PCR amplification with good polyacrylamide gel resolution, while the other five loci (mPgCIR003, mPgCIR008, mPgCIR014, mPgCIR017, and mPgCIR023), which had no hits in the *in silico* analysis, did not produce amplicons.

Locus mPgCIR001 (e-value = $2.00e^{-91}$) amplified all four DNA samples with high quality and band definition, which may imply high alignment quality. Three of the four samples tested with locus mPgCIR005 (e-value = $2.00e^{-67}$) were amplified with high resolution. Locus mPgCIR007 ($2.00e^{-120}$) showed the best gel resolution, sample amplification, and monomorphism. Locus mPgCIR009 (e-value = $6.00e^{-72}$) also amplified all samples, although at a lower resolution compared to the other loci. Locus mPgCIR018 (e-value = $9.00e^{-109}$) amplified all samples with good quality and gel resolution. Locus mPgCIR020 (e-value = $2.00e^{-76}$) amplified all samples with optimum quality and gel resolution. Finally, locus mPgCIR026 (e-value = $1.00e^{-87}$) also amplified the four samples.

According to the guava genetic linkage map (Lepitre et al., 2010) and the whole *Eucalyptus* genome sequencing (Ferguson et al., 2020), locus mPgCIR005 was mapped to

linkage group (LG) 1 in *P. guajava* and *E. globulus*. In turn, locus mPgCIR007 was mapped to LG2 in *P. guajava* and LG10 in *E. globulus*. Locus mPgCIR018 was located in LG1 in *P. guajava* and LG3 in *E. globulus*. Locus mPgCIR020 was located in LG7 in *P. guajava* and 11 in *E. globulus*. Finally, locus mPgCIR378 was mapped to LG4 in *P. guajava* and LG2 in *E. globulus*.

DISCUSSION

Studies involving *in silico* approaches should consider statistical analysis and alignment scores as they constitute important information and bring objectivity to sequence comparison and alignment quality assessments (Lobo 2008). According to Pertsemlidis and Fondon (2001), e-values lower than 0.1 or 0.05 are usually biologically significant. Also, according to these authors, high e-value significance implies sequence conservation between two species, which is relevant for comparative genomics and to postulate the protein function.

Markers made available by Risterucci et al. (2005) showed high identity values and a theoretical transferability of nine out of 23. This value was four times higher than that reported by Siqueira (2014) when testing EST-SSR from *Eucalyptus* to *Eugenia klotzschiana*, obtaining a marker transferability of 11 out of 67 tested. Miranda (2014) reported 12 out of 120 transferability in EST-SSR from *Eucalyptus spp.* to *Campomanesia adamantium* and *C. pubescens*. However, the theoretical transferability described here is lower than that reported by Rai et al. (2013) when evaluating the *in vitro* SSR transferability from guava to four other Myrtaceae species (78.2 % in *E. citriodora* and 60.8 % in *E. camaldulensis)*. This difference may be related to the genotype, as described by Varshney et al. (2005), who obtained different transferability levels (64.2 and 72.1%) in two rye genotypes.

Although *P. guajava* and *E. globulus* are genetically close species (both belong to the family Myrtaceae), they are mainly allogamous and consequently show high heterozygosity, which could explain the different transferability levels described in this study. Marker transfer of SNP were also reported between *P. guajava* and *Eucalyptus* (Costa and Santos 2017) to analyze the genetic divergence and order to orientate guava genetic resources and breeding programs. Also, Cabe and Marshall (2001) stated that the probability of a successful amplification for any DNA sequence is inversely proportional to the evolutionary distance involved, which may justify the low transferability between the genera *Psidium* and *Eucalyptus*. Thornhill et al. (2015) estimate that tribe separation, which originated these two genera, occurred approximately 60-65 million years ago, increasing their evolutionary distance and possibly reducing transferability.

The low transferability obtained with the GuavaMap SSR loci in the present study is attributed to the use of short-length sequences (about 20 nucleotides). This was expected since using these sequences or low-complexity sequences may result in high identity levels, which, however, is not reliable from a biological perspective (Pertsemlidis and Fondon, 2001). The parameters adopted in the present study, such as the choice for alignments from the same chromosome in distances smaller than 300 nucleotides and with low e-values, proved to be helpful to exploit the currently available information in the search for homologies between *P. guajava* and *E. globulus* and can be used as an indication of marker positions in the *P. guajava* genome.

Every computational data must be compared to experimental data for validation (Verli, 2014). Therefore, all loci from *Psidium* predicted to be transferable to *Eucalyptus* were confirmed by *in vitro* analysis, resulting in a high-quality amplification. In contrast, loci with no significant results in the *in silico* analysis did not show *in vitro* amplification. This result was expected since BLAST could not locate those sequences in the eucalypt genome. When using the methodology proposed by Kuleung et al. (2004), Rai et al. (2013) classified PCR products into four classes in order to confirm marker transferability: (1) Strong signal and easy to score; (2) weaker signal but possible to identify; (3) weak signal and difficult to identify; and (4) no signal, with groups 3 and 4 having no amplification pattern. In our study, loci mPgCIR001, mPgCIR005, mPgCIR007 mPgCIR009, mPgCIR018, mPgCIR020, and mPgCIR026 were classified into groups 1 and 2, while loci mPgCIR003, mPgCIR008, mPgCIR014, mPgCIR017, and mPgCIR023 were classified into groups 3 and 4.

The theoretical transferability described here has a significant potential to help genetic mapping, comparative genomics, and transferability studies as it can 'filter' large amounts of molecular information before experimental steps. Jiang et al. (2020) reported that the increase in the number of species sequenced by NGS technologies would increase the information in databases such as GenBank, providing abundant material to be explored using BLAST-like tools. Theoretical transferability may be applied to any marker. However, EST-SSR stands out due to its higher transferability than genomic SSR since it originates from transcription (therefore, more conserved) regions (Wu et al., 2014).

CONCLUSIONS

All seven P. guajava microsatellite loci with Blats low e-values (<1.00e−20) amplified PCR products from Eucalyptus sp. DNA samples, showing easy-to-score amplicons on the polyacrylamide gels. Microsatellite loci with high Blast e-values or hits showed no amplification on the same gels. These results highlight the effectiveness of in silico transferability of SSR loci and constitute a valid alternative to save resources in transferability studies among species.

ACKNOWLEDGMENTS

The authors thank the National Council for Scientific and Technological Development (CNPq) of Brazil for the financial support to this study.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Acuña CV, Fernandez P, Villalba PV, García MN, et al. (2012). Discovery, validation, and in silico functional characterization of EST-SSR markers in *Eucalyptus globulus*. *Tree Genet. Genomes*. 8: 289-301.

Butler JB, Vaillancourt RE, Potts BM, Lee DJ, et al. (2017). Comparative genomics of *Eucalyptus* and *Corymbia* reveals low rates of genome structural rearrangement. *BMC Genomics*. 18: 397.

Cabe PR and Marshall KE (2001). Microsatellite loci from the house wren (*Troglodytes aedon*). *Mol. Ecol. Notes*. 1: 155-156.

Carrer H, Barbosa AL and Ramiro DA (2010). Biotecnologia na agricultura. *Estud. Avancados*. 24: 149-164.
Costa SR and Santos CAF (2013). Allelic database and divergence among *Psidium* accessions by using microsatellite markers. *Genet. Mol. Res*. 12: 6802-6812.
Costa SR and Santos CAF (2017). Genetic divergence among Psidium accessions based on single nucleotide polymorphisms developed for *Eucalyptus*. *Genet. Mol. Res*. 16: 10.4238/gmr16029566.
Santos KL, Welter LJ, Dantas AC de M, Guerra MP, et al. (2007). Transference of microsatellite markers from *Eucalyptus spp.* to *Acca sellowiana* and the successful use of this technique in genetic characterization. *Genet. Mol. Biol*. 30: 73-79.
Fagundes BS, da Silva LF, Giacomin RM, Secco D, et al. (2016). Transferability of microsatellite markers among Myrtaceae species and their use to obtain population genetics data to help the conservation of the Brazilian Atlantic Forest. *Trop. Conser. Sci*. 9: 408-422.
Ferguson S, Jones A and Borevitz J (2020). Eucalyptus pangenome project. In: National Center for Biotechnology Information [Internet]. Available at [https://www.ncbi.nlm.nih.gov/nuccore/JABKBK000000000.1/]. Accessed: April 16, 2021.
Grattapaglia D, Silva-Junior OB, Kirst M, de Lima BM, et al. (2011). High-throughput SNP genotyping in the highly heterozygous genome of *Eucalyptus*: Assay success, polymorphism and transferability across species. *BMC Plant. Biol*. 11: 65.
Hudson CJ, Kullan ARK, Freeman JS, Faria DA, et al. (2012). High synteny and colinearity among Eucalyptus genomes revealed by high-density comparative genetic mapping. *Tree Genet. Genomes*. 8: 339-352.
Inglis PW, Pappas M de CR, Resende LV and Grattapaglia D (2018). Fast and inexpensive protocols for consistent extraction of high-quality DNA and RNA from challenging plant and fungal samples for high-throughput SNP genotyping and sequencing applications. *PLoS One*. 13: e0206085.
Jiang Y, Xu S, Wang R, Zhou J, et al. (2020). Characterization, validation, and cross-species transferability of EST-SSR markers developed from *Lycoris aurea* and their application in genetic evaluation of *Lycoris* species. *BMC Plant Biol*. 20: 522-536.
Kuleung C, Baenziger PS and Dweikat I (2004). Transferability of SSR markers among wheat, rye, and triticale. *Theor. Appl. Genet*. 108: 1147-1150.
Lepitre V, Nansot G, Grangeon R, Pomies V, et al. (2010). The microsatellite (SSR)/AFLP reference linkage map of guava. *Acta Hortic*. 849: 183-192.
Lobo I (2008). Basic local alignment search tool (BLAST). *Biol. Fertil. Soils* 1: 215-223.
Miranda EAGC (2014). Transferibilidade e validação de marcadores microssatélites derivados de EST para duas espécies de *Campomanesia* (Myrtaceae) do Cerrado. Master's thesis. Universidade Federal de Goiás, Goiânia. Available at: [https://repositorio.bc.ufg.br/tede/handle/tede/5595].
Pertsemlidis A and Fondon JW (2001). Having a BLAST with bioinformatics (and avoiding BLASTphemy). *Genome Biol*. 2: REVIEWS2002.
Rafalski JA, Vogel JM, Morgante M, Powell IW, et al. (1996). Analysis of Nonmammalian Genomes: A Practical Guide In: Generating and Using DNA markers in Plants (Birren B and Lai E, eds.). Academic Press, Boca Raton, 75-134.
Rai MK, Phulwaria M and Shekhawat NS (2013) Transferability of simple sequence repeat (SSR) markers developed in guava (*Psidium guajava* L.) to four Myrtaceae species. *Mol. Biol. Rep*. 40: 5067-5071.
Risterucci AM, Duval MF, Rohde W and Billotte N (2005) Isolation and characterization of microsatellite loci from *Psidium guajava* L. *Mol. Ecol. Notes* 5: 745-748.
Shepherd M, Kasem S, Lee D and Henry R (2006). Construction of micro- satellite linkage maps for *Corymbia*. *Silvae Genet*. 55: 228-238.
Singleton JJ, Mangat PK, Shim J, Vavra C, et al. (2020). Cross-species transferability of Solanum spp. DNA markers and their application in assessing genetic variation in silverleaf nightshade (*Solanum elaeagnifolium*) populations from Texas, USA. *Weed Sci*. 68: 396-404.
Siqueira MN de (2014). Transferibilidade e variabilidade genética nos microssatélites gênicos em Eugenia klotzschiana BERG (Myrtaceae). Master's thesis. Universidade Federal de Goiás, Goiânia. Available at: [https://repositorio.bc.ufg.br/tede/handle/tede/4172].
Sitther V, Zhang D, Harris DL, Yadav AK, et al. (2014). Genetic characterization of guava (*Psidium guajava* L.) germplasm in the United States using microsatellite markers. Genetic Resour. *Crop Evol*. 61: 829-839.
Sumathi M, Bachpai VKW, Mayavel A, Dasgupta MG, et al. (2018). Genetic linkage map and QTL identification for adventitious rooting traits in red gum eucalypts. *3 Biotech*. 8: 242.
Taheri S, Abdullah TL, Yusop MR, Hanafi MM, et al. (2018). Mining and development of novel SSR markers using Next Generation Sequencing (NGS) data in plants. *Molecules*. 23: 399-417.
Thornhill AH, Ho SYW, Külheim C and Crisp MD (2015). Interpreting the modern distribution of Myrtaceae using a dated molecular phylogeny. *Mol. Phylogenet. Evol*. 93: 29-43.
Tuler AC, Carrijo TT, Nóia LR, Ferreira A, et al. (2015). SSR markers: a tool for species identification in *Psidium* (Myrtaceae). *Mol. Biol. Res*. 42: 1501-13.
Varshney RK, Sigmund R, Börner A, Korzun V, et al. (2005). Interspecific transferability and comparative mapping of barley EST-SSR markers in wheat, rye and rice. *Plant Sci*. 168: 195-202.

Varshney RK, Thiel T, Stein N, Langridge P, et al. (2002). In Silico Analysis on Frequency and Distribution of microsatellites in ESTs of some cereal species. *Cell. Mol. Biol. Lett.* 7: 537-546.

Verli H (2014). Bioinformática da Biologia à Flexibilidade Molecular. 1st edn. Sociedade Brasileira de Bioquímica e Biologia Molecular Press, São Paulo.

Wei L, Liu Y, Dubchak I, Shon J, et al. (2002). Comparative genomics approaches to study organism similarities and differences. *J. Biomed. Infor.* 35: 142-150.

Wu J, Cai C, Cheng F, Cui H, et al. (2014). Characterisation and development of EST-SSR markers in tree peony using transcriptome sequences. *Mol. Breed.* 34: 1853-1866.

Zhang Z, Schwartz S, Wagner L and Miller W (2000). A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* 7: 203-214.

Zou C, Karn A, Reisch B, Nguyen A, et al. (2020). Haplotyping the Vitis collinear core genome with rhAmpSeq improves marker transferability in a diverse genus. *Nat. Commun.* 11: 413.