Review Article

Microsatellite markers: what they mean and why they are so useful

Maria Lucia Carneiro Vieira[1], Luciane Santini[1], Augusto Lima Diniz[1] and Carla de Freitas Munhoz[1]

[1]*Departamento de Genética, Escola Superior de Agricultura "Luiz de Queiroz" (ESALQ), Universidade de São Paulo (USP), Piracicaba, SP, Brazil.*

Abstract

Microsatellites or Single Sequence Repeats (SSRs) are extensively employed in plant genetics studies, using both low and high throughput genotyping approaches. Motivated by the importance of these sequences over the last decades this review aims to address some theoretical aspects of SSRs, including definition, characterization and biological function. The methodologies for the development of SSR loci, genotyping and their applications as molecular markers are also reviewed. Finally, two data surveys are presented. The first was conducted using the main database of Web of Science, prospecting for articles published over the period from 2010 to 2015, resulting in approximately 930 records. The second survey was focused on papers that aimed at SSR marker development, published in the American Journal of Botany's Primer Notes and Protocols in Plant Sciences (over 2013 up to 2015), resulting in a total of 87 publications. This scenario confirms the current relevance of SSRs and indicates their continuous utilization in plant science.

Keywords: SSR biological function, genomic distribution, genotyping approaches, molecular marker, practical utility.

Received: February 29, 2016; Accepted: May 13, 2016.

Brief introduction

Ongoing technological advances in all fields of knowledge mean that we cannot be sure which technologies will survive the impact of innovation, and for how long. Over the years, advances in molecular genetics methodology have lead to widespread use of codominant molecular markers, especially Simple Sequence Repeats (SSRs) and, more recently, Single Nucleotide Polymorphisms (SNPs). This paper attempts to present an overview of how the concept of SSRs has evolved and how their biological functions were discovered. We also address the development of methods for identifying polymorphic SSRs, and the application of these markers in genetic analysis. It reveals that much remains to be explored regarding these sequences, particularly in relation to cultivated and wild plants.

Definition and genome occurrence of microsatellites and their use as genetic markers

Microsatellites (1 to 10 nucleotides) and minisatellites (> 10 nucleotides) are subcategories of tandem repeats (TRs) that, together with the predominant interspersed repeats (or remnants of transposable elements), make up genomic repetitive regions. TRs are evolutionarily

relevant due to their instability. They mutate at rates between 10^3 and 10^6 per cell generation i.e., up to 10 orders of magnitude greater than point mutations (Gemayel *et al.*, 2012).

Microsatellites, Simple Sequence Repeats (SSR), Short Tandem Repeats (STR) and Simple Sequence Length Polymorphisms (SSLP) are found in prokaryotes and eukaryotes. They are widely distributed throughout the genome, especially in the euchromatin of eukaryotes, and coding and non-coding nuclear and organellar DNA (Pérez-Jiménez *et al.*, 2013; Phumichai *et al.*, 2015).

There is a lot of evidence to back up the hypothesis that SSRs are not randomly distributed along the genome. In a comparative study, SSR distribution was found to be highly non-random and to vary a great deal in different regions of the genes of *Arabidopsis thaliana* and rice (Lawson and Zhang, 2006). In the major cereals, for instance, authors have tended to categorize microsatellites based on different criteria. In barley and *Avena* species, SSRs were classified in two types: those with unique sequences on either flank and those intimately associated with retrotransposons and other dispersed repetitive elements. The second type was found to be less polymorphic in oat cultivars (Ramsay *et al.*, 1999; Li *et al.*, 2000). Using publicly available DNA sequence information on the rice genome, Temnykh *et al.* (2001) categorized microsatellites based on length and noticed that longer perfect repeats (≥ 20 nucleotides) were highly polymorphic. Microsatellites with SSRs shorter than 12 bp were found to have a mutation potential

Send correspondence to Maria Lucia Carneiro Vieira. USP/ESALQ, P.O. Box 83, 13400-970, Piracicaba, SP, Brazil. E-mail: mlcvieir@usp.br

no different from that of most unique sequences. Moreover, authors reported that ~80% of GC-rich trinucleotides occurred in exons, whereas AT-rich trinucleotides were distributed roughly evenly throughout all genomic components (coding sequences, untranslated regions, introns and intergenic spaces). Tetranucleotide SSRs were predominantly situated in non-coding, mainly intergenic regions of the rice genome. It was later established that the SSR distributions in different regions of the maize genome were non-random, and that density was highest in untranslated regions (UTR), gradually falling off in the promotor, intron, intergenic, and coding sequence regions, in that order (Qu and Liu, 2013).

On the other hand, comparisons of microsatellite distributions in *Rumex acetosa* and *Silene latifolia* chromosomes showed that some motifs (e.g. CAA or TAA) are strongly accumulated in non-recombining regions of the sex chromosome (Y) in both plant species (Kejnovsky *et al.*, 2009). Similarly, a very large accumulation consisting mainly of microsatellites on the heterochromatic W chromosome was reported in a group of fish species (*Leporinus* spp.) that share a ZW sex system, showing an interconnection between heterochromatinization and the accumulation of repetitive sequences, which has been proposed as the basis of sex chromosome evolution (Poltronieri *et al.,* 2014).

Generally speaking, it can be affirmed that the occurrence of SSRs is lower in gene regions, due to the fact that SSRs have a high mutation rate that could compromise gene expression. Studies indicate that in coding regions there is a predominance of SSRs with gene motifs of the tri- and hexanucleotide type, the result of selection pressure against mutations that alter the reading frame (Zhang *et al.*, 2004; Xu *et al.*, 2013b). In humans, the consensus is that SSRs can also originate in coding regions, leading to the appearance of repetitive patterns in protein sequences. In protein sequence database studies, it was reported that tandem repeats are common in many proteins, and the mechanisms involved in their genesis may contribute to the rapid evolution of proteins (Katti *et al.*, 2000; Huntley and Golding, 2000).

Repeat polymorphisms usually result from the addition or deletion of the entire repeat units or motifs. Therefore, different individuals exhibit variations as differences in repeat numbers. In other words, the polymorphisms observed in SSRs are the result of differences in the number of repeats of the motif caused by polymerase strand-slippage in DNA replication or by recombination errors. Strand-slippage replication is a DNA replication error in which the template and nascent strands are mismatched. This means that the template strand can loop out, causing contraction. The nascent strand can also loop out, leading to repeat expansion. Recombination events, such as unequal crossing over and gene conversion, may additionally lead to SSR sequence contractions and expansions. According to several authors, the longer and purer the repeat, the higher the mutation frequency, whereas shorter repeats with lower purity have a lower mutation frequency.

Mutations that have evaded correction by the DNA mismatch repair system form new alleles at SSR loci. For this reason, different alleles may exist at a given SSR locus, which means that SSRs are more informative than other molecular markers, including SNPs.

As for their composition, SSRs can be classified according to motif as: *i)* perfect if composed entirely of repeats of a single motif; *ii)* imperfect if a base pair not belonging to the motif occurs between repeats; *iii)* interrupted if a sequence of a few base pairs is inserted into the motif; or *iv)* composite if formed by multiple, adjacent, repetitive motifs (reviewed in Oliveira *et al.*, 2006; revisited by Mason, 2015).

SSRs have been the most widely used markers for genotyping plants over the past 20 years because they are highly informative, codominant, multi-allele genetic markers that are experimentally reproducible and transferable among related species (Mason, 2015). In particular, SSRs are useful for wild species (*i*) in studies of diversity measured on the basis of genetic distance; (*ii*) to estimate gene flow and crossing over rates; and (*iii*) in evolutionary studies, above all to infer infraspecific genetic relations. On the other hand, for cultivated plants SSRs are commonly used for (*i*) constructing linkage maps; (*ii*) mapping loci involved in quantitative traits (QTL); (*iii*) estimating the degree of kinship between genotypes; (*iv*) using marker-assisted selection; and (*v*) defining cultivar DNA fingerprints (Jonah *et al.*, 2011; Kalia *et al.*, 2011). SSRs have been particularly useful for generating integrated maps for plant species in which full-sib families are used for constructing linkage maps (Garcia *et al.*, 2006; Souza *et al.*, 2013; Pereira *et al.*, 2013), and for combining genetic, physical, and sequence-based maps (Temnykh, 2001), providing breeders and geneticists with a tool to link phenotypic and genotypic variation (see Mammadov *et al.*, 2012; Hayward *et al.*, 2015 for review articles).

These markers are enormously useful in studies of population structure, genetic mapping, and evolutionary processes. SSRs with core repeats 3 to 5 nucleotides long are preferred in forensics and parentage analysis. It is worth noting that a number of SSR search algorithms have been developed, including TRF (Benson, 1999), SSRIT (Temnykh, 2001), MISA (Thiel *et al.*, 2003), SSRFinder (Gao *et al.*, 2003), TROLL (Castelo *et al.*, 2002) and SciRoKo (Kofler *et al.*, 2007).

Detailing the biological functions of SSRs

Despite the wide applicability of SSRs as genetic markers since their discovery in the 1980s, little is known about the biological importance of microsatellites (Tautz and Renz, 1984), especially in plants. Morgante *et al.* (2002) estimated the density of SSRs in *Arabidopsis thaliana*, rice (*Oryza sativa*), soybean (*Glycine max*),

maize (*Zea mays*) and wheat (*Triticum aestivum*) and observed a high frequency of SSRs in transcribed regions, especially in untranslated regions (UTRs). Interestingly, there are substantial data indicating that SSR expansions or contractions in protein-coding regions can lead to a gain or loss of gene function via frameshift mutation or expanded toxic mRNAs. SSR variations in 5'-UTRs could regulate gene expression by affecting transcription and translation, but expansions in the 3'-UTRs cause transcription slippage and produce expanded mRNA, which can disrupt splicing and may disrupt other cellular functions. Intronic SSRs can affect gene transcription, mRNA splicing, or export to cytoplasm. Triplet SSRs located in UTRs or introns can also induce heterochromatin-mediated-like gene silencing. All these effects can eventually lead to phenotypic changes (Li *et al.*, 2004; Nalavade *et al.*, 2013).

In fact, variation in the length of DNA triplet repeats has been linked to phenotypic variability in microbes and to several human disorders, including Huntington's disease which is caused mainly by $(CAG)_n$ expansions. Moreover, the frequencies of different codon repeats vary considerably depending on the type of encoded amino acid. In plants, a triplet repeat-associated genetic defect was identified in a wild variety of *A. thaliana* that carries a dramatically expanded TTC/GAA repeat in the intron of the gene encoding the large subunit 1 of the isopropyl malate isomerase. Expansion of the repeat causes an environment-dependent reduction in the enzyme's activity and severely impairs plant growth, whereas contraction of the expanded repeat can reverse the detrimental effect on the phenotype (Sureshkumar *et al.*, 2009).

Historically, tandem repeats have been designated as nonfunctional DNA, mainly because they are highly unstable. With the exception of tandem repeats involved in human neurodegenerative diseases, repeat variation was often believed to be neutral with no phenotypic consequences (see Gemayel *et al.*, 2012).

The detection of microsatellites in transcripts and regulatory regions of the genome encouraged scientific interest in discovering their possible biological functions. More and more publications have presented evidence that microsatellites play a role in relevant processes, such as the regulation of transcription and translation, organization of chromatin, genome size and the cell cycle (Nevo, 2001; Li *et al.*, 2004; Gao *et al.*, 2013).

As mentioned above, most of the knowledge acquired on microsatellites occurring in genes was obtained by studying humans and animals, indicating their relationship with the manifestation of disease. In bacteria, maintaining numerous microsatellite variants provides a source of highly mutable sequences that enable prompt generation of novel variations, ensuring the survival of the bacterial population in widely varying environments, and adaptation to pathogenesis and virulence. Nevertheless, few studies have focused on whether the typical instability of microsatellites

is linked to phenotypic effects in plants (Li *et al.*, 2004; Gao *et al.*, 2013). However, thanks to whole genome sequencing the important role repeats might play in genomes is being elucidated.

The consensus is that the biological function of a microsatellite is related to its position in the genome. For instance, SSRs in 5'-UTRs serve as protein binding sites, thereby regulating gene translation and protein component and function, as classically demonstrated for the human genes for thymidylate synthase (Horie *et al.*, 1995) and calmodulin-1 (Toutenhoofd *et al.*, 1998). Ten years later, SSR densities in different regions (5'-UTRs, introns, coding exons, 3'-UTRs, and upstream regions) in housekeeping and tissue-specific genes in human and mouse were compared. Specifically, SSRs in the 5'-UTRs of housekeeping genes are more abundant than in tissue-specific genes. Additionally, it was suggested that SSRs may have an effect on gene expression and may play an important role in contributing to the different expression profiles of housekeeping and tissue-specific genes (Lawson and Zhang, 2008).

In plants, despite the fact that a high density of SSRs has been detected in 5'-UTR regions (Fujimori *et al.*, 2003; Tranbarger *et al.*, 2012; Zhao *et al.*, 2014), there are few studies verifying their effect on the regulation of gene expression. Additionally, tri- and hexanucleotide coding repeats appear to be controlled by stronger mutation pressure in coding regions than in other gene regions. Consequently, in plants there is less allele variability in exonic SSRs than in intronic SSRs. The biased distribution of microsatellites and microsatellite motifs also suggests that microsatellites of different types play different roles in different gene regions, such as within promoters, introns and exons in plants (Li *et al.*, 2004; Gemayel *et al.*, 2012; Gao *et al.*, 2013).

Comparison among SSRs located in CDS, 5' UTR and 3' UTR in the transcriptome of *Sargassum thunbergii*, an economically important brown macroalgae has confirmed that UTR regions harbored more microsatellite compared to the CDS, and the length variation of microsatellite was significantly affected by repeat motif size. Remarkably were the results relative to the function of microsatellite-containing transcripts. After an enrichment analysis, four pathways, i.e. ubiquitin-mediated proteolysis, RNA degradation, spliceosome and terpenoid backbone biosynthesis were obtained, providing new insights into the function and evolution of microsatellite in transcript sequences (Liu *et al.*, 2016).

Microsatellites located in introns can play a role in the transport and alternative splicing of mRNA and in gene silencing, as well as in the regulation of transcription, acting independently or in combination with SSRs present in 5'-UTR regions (Kalia *et al.*, 2011). A number of examples of the effects of intronic SSRs in humans were reviewed by Li *et al.* (2004), including an increase in the expression of the type I collagen alpha2 gene, caused by the presence of

(CA)$_n$ repeats in the 5'-UTR region and (GT)$_n$ repeats in the first intron.

The 3'-UTR region is also subject to alterations due to the presence of SSRs which cause slippage during the transcription or modification of target regions whose translation is controlled by miRNAs (Li *et al.*, 2004; Gao *et al.*, 2013). An example of the effect of polymerase slippage in 3'-UTR regions is the multisystem disorder myotonic dystrophy type 1, caused by expansion of a CTG trinucleotide repeat. Normal alleles have 5 to 34 CTG repeats, but alleles with > 50 CTG repeats are associated with disease manifestations (see Ranum and Day, 2002; Li *et al.*, 2004; Bird, 2015).

Finally, microsatellites are known to affect expression if present in gene promoters and intergenic regions. In the promotor, SSRs render gene expression vulnerable to possible alterations caused by expansion or contraction of repeat sequences. These alterations result in an increase or reduction in the level of gene expression caused by changes in transcription factor linkage sites and can even culminate

in gene silencing. Tandem repeats in intergenic regions can cause changes in the secondary structure of the DNA by forming loops and altering the chromatin, which indirectly results in alterations in the expression of nearby genes (Gao *et al.*, 2013).

In spite of the scarcity of studies on the functional changes brought about by SSRs in plants, their effects are believed to be similar to those found in humans. For instance, the occurrence of trinucleotide repeats in *Arabidopsis* genome was found to be twice as frequent in coding regions, suggesting selection for certain stretches of amino acids (Morgante *et al.*, 2002). Using data generated in our laboratory, we have compared the percentage of SSRs having mono-, di-, tri, tetra-, penta and hexanucleotide motifs in expressed sequences, gene-rich regions, BAC-end sequences and chloroplast genome sequences of *Passiflora edulis*, and identified the prevalent motif in each case. We also noticed the prevalence of tri- and hexanucleotide motifs in expressed sequences (Figure 1).
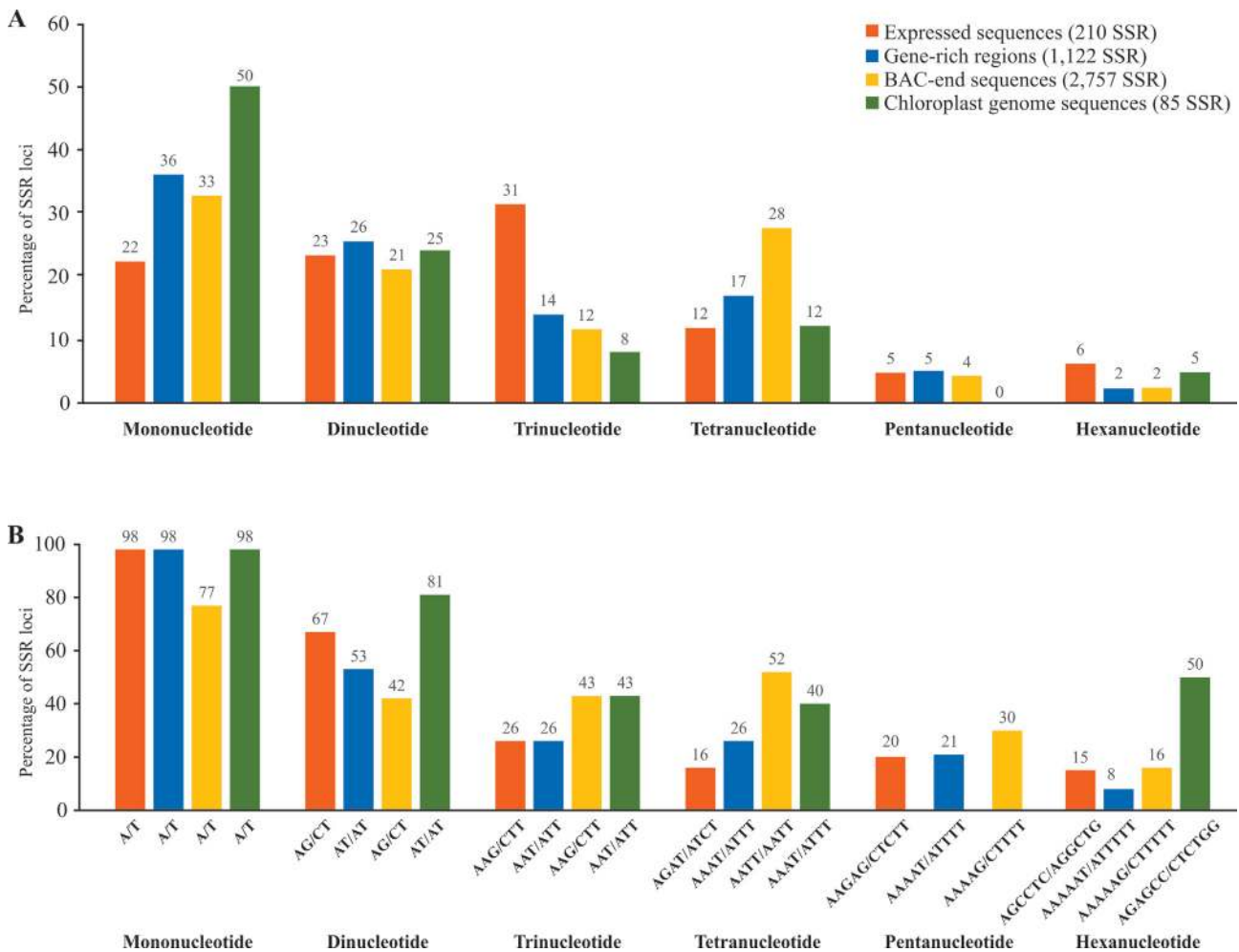

Figure 1 - The percentage of mono-, di-, tri-, tetra-, penta- and hexanucleotides in the microsatellites found in expressed sequences, gene-rich regions, BAC-end sequences and in the chloroplast genome of *Passiflora edulis* (Passifloraceae) (A); the percentage of the most common motif is displayed for each case (B).

Recently, based on the genomes available in the Phytozome database, Zhao *et al.* (2014) analyzed the distribution of tandem repeats in 29 species of terrestrial plants and two species of algae, in which the density of repeat sequences was higher in introns and coding sequences; in plants, 5'-UTR and upstream intergenic 200 nucleotide regions had the first and second highest densities.

In cDNA libraries constructed using plant and reproductive tissues of *Elaeis guineensis*, SSRs were observed in both coding regions and UTRs (Tranbarger *et al.*, 2012). The majority were identified in open reading frames, indicating a possible effect on the gene product and consequently on gene function. On the other hand, mutations in SSRs located in UTRs could affect transcription, translation or transcript splicing (Tranbarger *et al.*, 2012).

An important example of the functioning of SSRs in plants was reported by Liu *et al.* (2014b) using a high-throughput sequencing approach to characterize miRNAs and their targeted transcripts in different tissues of sweet orange. These miRNAs were evenly distributed across the genome in several small clusters, and 69 pre-miRNAs were co-localized with SSRs. Noticeably, the loop size of a particular pre-miRNA was influenced by the repeat number of the CUU codon. Another important aspect is the instability of microsatellites. Studies conducted on transgenic plants of *A. thaliana* showed that this instability increases as the plant ages, mainly due to a drop in the efficiency of DNA repair mechanisms (Golubov *et al.*, 2010). This peculiarity means that SSR markers can be used to assess the impacts of mutagenic contaminants. Mutagenesis induced in *Pisum sativum* by high doses of lead was detected based on the in-stability of microsatellites at a locus involved in metabolizing glutamine (Rodriguez *et al.*, 2013).

Microsatellite alterations associated with diseases in humans are widely known and can give the false impression that the effects of these mutations are predominantly adverse. On the contrary, some examples provide evidence that SSR alleles can offer potential selective advantages (Kashi and King, 2006). It was therefore time to abandon the presumption that SSRs are junk DNA. SSRs are currently qualified as relevant to population adaptation and phenotypic plasticity within and across generations and gene-associated tandem repeats act as evolutionary facilitators, providing abundant, robust variation and thus enabling rapid development of new forms (Nevo, 2001; Kashi and King, 2006).

Development of SSR markers, including *de novo* nucleotide sequences for finding SSRs

The development of SSR markers can basically be divided into the following stages: (*i*) prior knowledge of nucleotide sequences in which SSRs occur; (*ii*) design of oligonucleotides (or primers) complementary to the regions flanking the SSR; (*iii*) validation of primers by PCR and electrophoresis of the product of the reaction, and (*iv*) detection of polymorphisms among individuals (Mason, 2015). A schematic workflow showing how an SSR marker can be obtained is given in Figure 2. Interestingly, the efficiency of SSR marker development was found to be associated with the microsatellite class. In rice, for instance, the rate of successful amplification varied from 31.7% (AT repeats) up to 87% (GAA repeats). The following figures
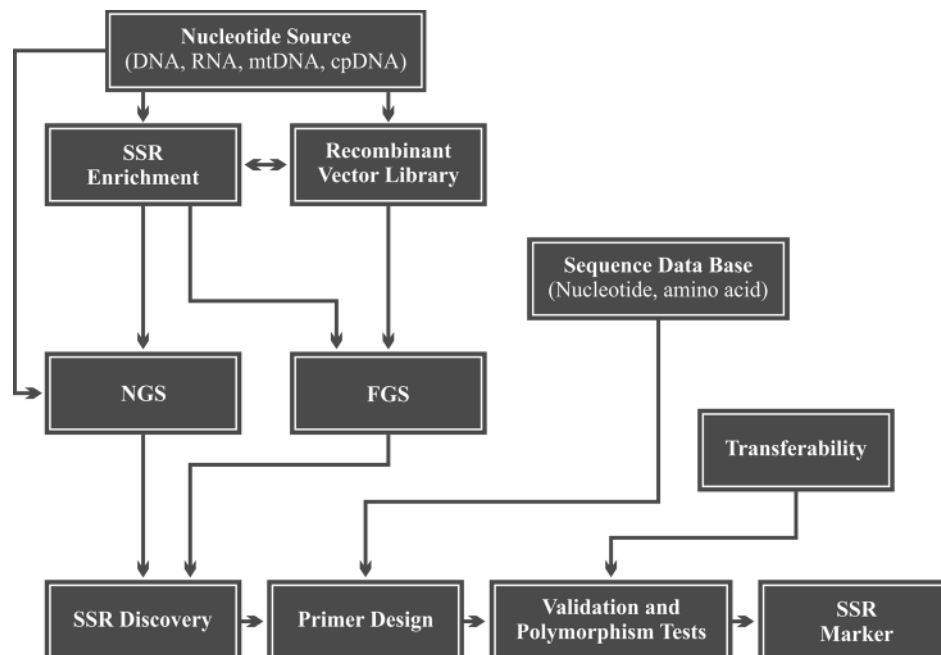

Figure 2 - Workflow steps of SSR marker development.

were observed for other SSR classes: GA, 83.8%; CA, 71.8%; GC-rich trinucleotides, 64.45%; ATT, 78,3%; CAT and CAA, 83,3% and tetranucleotides, 71.4% (Temnykh *et al.*, 2001).

Microsatellites were originally developed from both coding and non-coding regions of plant genomes, and several sources were used to search for SSRs, including a variety of DNA libraries (genomic, genomic-enriched for SSR, bacterial artificial chromosome and cDNA libraries), as well as public databases, including expressed sequence tag (EST) databases (see Hanai *et al.*, 2007).

In prospecting for SSRs, the first step consists of constructing enriched genomic libraries and various enrichment methods have been successfully developed (Billotte *et al.*, 1999; Maio and Castro, 2013). To construct and sequence genomic libraries, the DNA is fragmented, ligated to adaptors and inserted into vectors for transforming *Escherichia coli*. Most protocols involve a stage of enrichment for repetitive sequences that can be achieved using selective hybridization, PCR or both techniques (Senan *et al.*, 2014). In enrichment by hybridization, positive clones are detected using radioactively or chemically labeled SSR probes. Finally, these clones are selected by PCR amplification and sequencing (Semagn *et al.*, 2006; Blair *et al.*, 2009). Another way of enriching a library is to use biotinized SSR probes that are captured by streptavidin-coated beads (Nunome *et al.*, 2006). The captured DNA is eluted, amplified, cloned and sequenced. The enriched libraries are screened to identify clones containing SSRs, producing the subsample of repetitive sequences that is intrinsic to this approach. PCR-based methods can bias the sampling of repetitive sequences in non-enriched libraries, since fragment selection and amplification are dependent on complementarity with specific primers for the SSR and cloning vector. However, non-enriched libraries and alternative methods derived from other molecular markers (e.g. RAPD and AFLP) have also been used to find SSRs (see Senan *et al.*, 2014).

The advances made in Next Generation Sequencing (NGS) have provided a new scenario for detecting microsatellites. Various NGS-based projects have been developed over the last few decades, generating an enormous quantity of sequences made available in public databases and widely used for prospecting for microsatellites. Automation of the original sequencing method proposed by Sanger and Coulson (1975) has made it possible to sequence the complete genome of *A. thaliana* (Arabidopsis Genome Initiative, 2000). However, because of the high cost of the Sanger method when sequencing complete genomes, it has been replaced by NGS platforms or a combination of both methods (Schnable *et al.*, 2009).

NGS has been very useful for various studies, including prospecting for new SSR markers. Successors of the Sanger sequencing method include the 454 FLX (Roche), Solexa (Illumina), SOLiD (Applied Biosystems) and HeliScope True Single Molecule Sequencing (Helicos) platforms. Third generation platforms are also currently available, including a platform developed by Pacific Biosciences (PacBio), based on a new sequencing technology, SMRT sequencing, which has the advantage of producing longer DNA reads.

Each platform has specific characteristics in terms of the number and size of reads generated, run time, as well as the accuracy and cost of each base read, with both advantages and disadvantages compared to other platforms (Egan *et al.*, 2012). In order to advice researchers in sequencing technology choice, Alic *et al.* (2016) published a review about different high-throughput sequencing methods and 50 stand-alone softwares used to control errors. Control error analysis is one of the most important steps in sequencing data analysis, mainly in *de novo* sequencing projects, that lack a reference genome. Furthermore, sequences that contain repetitive regions are challenges to be overcome by error correction methods, due to their vulnerability to errors. Initiatives for sequencing the complete genomes of various species use combinations of different platforms with the aim of incorporating the best features of each and extracting the maximum amount of information.

Currently, 454 and Illumina are the NGS platforms most widely used for developing SSR markers. However, the PacBio SMRT sequencing technology is being considered an economically viable alternative for discovering microsatellites (Grohme *et al.*, 2013).

In-silico prospecting and transferability of SSR markers

With the advent of NGS, it was necessary to create databases for storing the information generated. In addition to genomic sequences, a large quantity of expressed sequence tags (EST) derived from cDNA libraries (i.e. originating from mRNA) were also generated. The online database platforms for nucleotide, protein and transcript data available for the majority of plant species are relatively small when compared to model species, such as *A. thaliana*, *Glycine max, O. sativa* and *Z. mays*. Since the protocols for obtaining and isolating *de novo* SSR loci can be expensive and not viable in some cases, the investigation of these elements *in silico* (i.e. in the actual databases) is a promising strategy. This approach is possible only because SSR loci primers are transferable among different, phylogenetically matching species (Kuleung *et al.*, 2004).

The possibility of interchanging this genetic information is ascribed to the synteny between matching species. Although there are some exceptions resulting from structural rearrangements, synteny is an import attribute of plant genomes and is inversely proportional to the phylogenetic distance between species (Kaló *et al.*, 2004). The conservation of this information could indicate that these loci confer

evolutionary advantages, and are therefore subject to low selection pressure (Zhu, 2005).

Microsatellites found in the chloroplast genome of higher plants (cpSSRs) consist basically of mononucleotide repeats (A and T) (Bryan *et al.*, 1999). Contrarily, we have found 50, 25, 8, 12 and 5% of mono-, di-, tri-, tetra- and hexanucleotides respectively in the microsatellites of the chloroplast genome of *Passiflora edulis* (Figure 1A), but we have confirmed that mononucleotide repeats consisted predominantly of A/T repeats (98%, Figure 1B). In terms of transferability, cpSSRs are particularly promising for the study of phylogenetically distant species, since the regions flanking them are strongly conserved, so that universal primers can be developed (Weising and Gardner, 1999; Ebert and Peakall, 2009).

Genotyping

After identifying the sequences containing SSRs, specific primers must be synthesized (18 and 25 bp in length), complementary to the flanking regions, followed by amplification and polymorphism testing. According to Guichoux *et al.* (2011), a number of experimental problems can arise during SSR amplification, which can compromise allele calling and binning, resulting in increased error rates or the need for extensive manual corrections. These authors itemized possible solutions for aiding researchers to solve these problems, such as stuttering or shadow bands, non-template addition of a nucleotide by the *Taq* polymerase, primer mispriming, etc.

Once the SSR markers have been produced, genotyping can begin. It is a relatively easy and low-cost procedure. The allele variants of a given SSR locus can be identified by agarose gel electrophoresis (AGE) or polyacrylamide gel electrophoresis (PAGE), low-complexity methods used routinely in molecular genetics laboratories. PAGE genotyping is more labor intensive but provides better resolution, allowing identification of given polymorphisms for a single base pair (Penha *et al.*, 2013; Mason, 2015). Alternatively, marked SSR primers can be synthesized with fluorescent markers for genotyping by capillary electrophoresis using conventional sequencers (Araújo *et al.*, 2007; Csencsics *et al.*, 2010; Agarwal *et al.*, 2015). In this case, each DNA sample is loaded into a capillary containing a polyacrylamide matrix in which the electrophoresis is performed. The fluorescence emitted by the marked primer is captured and the molecular mass of the amplified fragment is determined. The result is an electropherogram showing luminescence peaks corresponding to each amplified allele. Lastly, the genotyping stage consists of comparing the electropherograms of different individuals (see Culley *et al.*, 2013; Mason, 2015), a technique that is particularly widely used when working with complex genome species, such as sugarcane and other polyploids (Morais TBR de, 2012, Doctoral Thesis. Escola Superior de Agricultura

"Luiz de Queiroz, University of São Paulo, Piracicaba, SP, Brazil).

The most appropriate genotyping method for each project is defined according to the species under investigation, the sensitivity required in determining allele variations, the availability of the equipment and cost effectiveness. The amplification and genotyping stages can be perfected to multiplex different SSR loci, cutting costs and saving time, and allowing large scale analysis (Brown *et al.*, 1996; Guichoux *et al.*, 2011; Lepais and Bacles, 2011). There are two ways of performing multiplexed analysis of microsatellite loci. The first is by multiplexed PCR, in which different SSR primers are placed in the same reaction tube. The following stages are essential: *i)* determining the length (in bp) of the alleles at each SSR locus; *ii)* selecting loci whose allele lengths are not superimposed; *iii) in silico* testing at melting temperature (T_m) and the possible formation of secondary structures between the primers of the SSR loci selected. The second multiplexed SSR loci analysis method entails multiplexed genotyping. In this case, amplifications are performed separately, but the amplified products of a biological sample are mixed and loaded into the same electrophoresis gel channel or sequencing capillary.

Guichoux *et al.* (2011) have published an outstanding analysis of current trends in microsatellite genotyping. Several aspects are reviewed, including the overall cost of SSR genotyping as a function of the degree of multiplexing and the number of genotyped samples. For instance, the most widely cited commercial kit has a cost per sample of 1.88. The authors then suggest solutions to cut the final cost per sample. According to these authors, most of the work done to develop and optimize SSR multiplexing actually consists of phases common to all SSR development projects.

In the past, alternative methods have been developed to facilitate genotyped PCR multiplexing by capillary electrophoresis, such as the M13 tailed primer method (Oetting *et al.*, 1995). In this method, the sequencing reaction is performed as a multiplexed PCR using the M13 (reverse) primer, conjugated with a fluorescent colorant and various modified SSR (forward) primers. The SSR primers are modified by a 19-bp extension at the 5' end, identical to the M13 nucleotide sequence. In the first PCR cycle, amplification is based on the SSR primers, forming an M13 annealing site at the 3' end, used in the second amplification cycle. A variant of this technique (Multiplex-Ready PCR) was subsequently published with the aim of cutting the cost of primer marking, which is usually 5 to 10 times that of conventional primer synthesis (Hayden *et al.*, 2008).

Current overview

Microsatellite genomic distribution, biological function and practical utility have been reviewed in a number of articles over the past two decades, some of which are highlighted here: Jarne and Lagoda (1996); Schlötterer (1998);

Li *et al.* (2002); Buschiazzo and Gemmell (2006); Oliveira *et al.* (2006); Sun *et al.* (2009); Guichoux *et al.* (2011); Gemayel *et al.* (2012); Senan *et al.* (2014); Mason (2015). With the aim of investigating the use of microsatellite markers over the period from 2010 to 2015 in the genetic analysis of cultivated plants, we conducted a search in the main database of Web of Science (Web of Science™ Core Collection). We entered "microsatellite" or "SSR marker" in the title field and "crop*" in the topic field. To avoid selecting records related to plant pathogens and insect pests, the following terms were excluded from the topic field: bacteria (bacter*), fungi (fung*), insect (insect*) and pathogen (pathogen*). Finally, the search was refined by selecting the field of Plant Science, and all resulting hits were manually checked. We found 933 unique records (Figure 3, Supplementary Material Table S1) showing that microsatellites continue to be used as high-relevance molecular markers in the genetic analysis of cultivated plants. The number of publications rose steadily until 2012, and then fell back, possibly due to the ease with which genetic studies could be carried out using SNPs.

Recent studies have shown that the easiest way of identifying SSR loci is by using NGS to sequence the genome or transcriptome. Zalapa *et al.* (2012) reviewed papers published in the American Journal of Botany's Primer Notes and Protocols in Plant Sciences, an important monthly journal that centralizes a significant number of publications related to the discovery and use of SSRs in plants. Note that the use of the Sanger method was predominant up to this time, as well as the use of genomic libraries enriched with sequences containing microsatellites. Since then, there is a tendency to replace this method by NGS genome or transcriptome sequencing.

With the aim of comparing this scenario with the current situation, we conducted a similar survey based on papers published in the AJB from January 2013 to December 2015, selecting only those papers in which sequences were generated by developing SSR markers. A total of 87 papers were published during this period, the majority of which involved using the Sanger method to sequence genomic libraries enriched with SSRs. It is worth noting that the use of NGS for prospecting for and generating SSR markers has been on the increase, surpassing the Sanger method in 2015 (Table 1). We also realized that the enrichment stage might no longer be advantageous, due to the number of sequences generated by NGS. On the contrary, since the composition of the nucleotide base and the frequency of SSR motifs can actually vary among plant genomes (Li *et al.*, 2002), the enrichment stage with a small number of motifs should allow curtailment or subsampling of the population of SSRs to be discovered.

Another interesting trend is that the Illumina platform is being routinely used for transcriptome sequencing. The advantage of developing SSR markers from transcribed sequences includes the possibility of finding associations with genes and phenotypes (Li *et al.*, 2002). As observed by Zalapa *et al.* (2012), a common factor of all the papers, irrespective of the sequencing method, is that only a small fraction of the SSR loci discovered have been assessed. As mentioned earlier, obtaining a sequence is only the first stage in the marker development process. Primer design and PCR optimization still represent a bottleneck. Furthermore, there is always the possibility that the locus is monomorphic, i.e. non-informative.

One strategy for working around this limitation is to track loci polymorphisms *in silico*, during the stage at which regions that contain SSRs are identified. This can be done using two or more genetically contrasting individuals or their progeny (F_1) for performing NGS, increasing the possibility of sampling alleles based on the alignment of the sequences obtained, and thereby avoiding the synthesis and testing of primers for monomorphic loci (Iorizzo *et al.*, 2011).
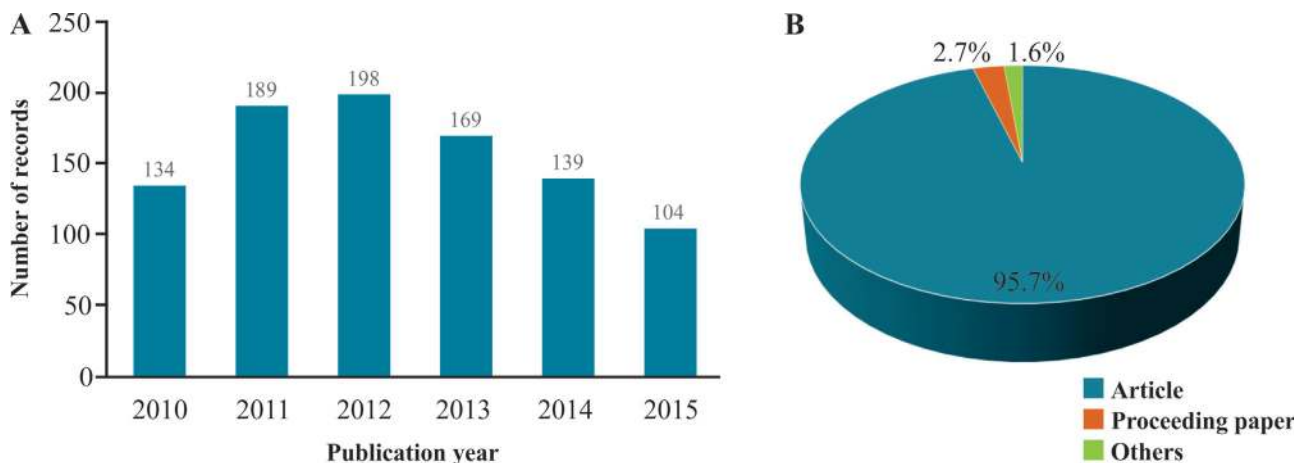

Figure 3 - Number of publications relating to the use of microsatellites in crop genetic studies from 2010 to 2015 according to the Web of Science database (A). Distribution of records according to the type of publication (B).

Table 1 - Recent studies involved in the detection and development of SSR markers in plants, using different sequencing technologies.

Technology	Source[1]	Library[2]	Enrichment	Species	Reference
Sanger	G	Y	CA repeats and $(GA)_{10}$	*Aniba rosaeodora*	Angrizani *et al.*, 2013
	T	**N**	Expressed sequence tags from roots	*Callerya speciosa*	Li *et al.*, 2013
	G	Y	$(AG)_{10}$, $(GT)_{15}$, $(CAG)_8$ and $(AC)_6(AG)_5$	*Canavalia cathartica* and *C. lineata*	Yamashiro *et al.*, 2013
	G	Y	$(CT)_8$ and $(GT)_8$	*Cariniana legalis*	Tambarussi *et al.*, 2013
	G	Y	CA, AAC, ATG, and TAGA	*Ceanothus megacarpus*	Ishibashi *et al.*, 2013
	G	Y	$(GT)_{12}$	*Cornus sanguinea*	Wadl *et al.*, 2013
	G	Y	$(AC)_{15}$, $(AG)_{15}$, and $(AAG)_{10}$	*Diplarche multiflora*	Zhang *et al.*, 2013
	G	Y	$(CT)_8$ and $(GT)_8$	*Encholirium horridum*	Hmeljevski *et al.*, 2013
	G	Y	Not informed	*Lagerstroemia indica*	Liu *et al.*, 2013c
	G	Y	$(AC)_6(AG)_7$ or $(AC)_6(TC)_7$	*Leptospermum recurvum*	Ando *et al.*, 2013
	G	Y	$(AC)_6(AG)_5$ or $(TC)_6(AC)_5$	*Lilium longiflorum*	Sakazono *et al.*, 2013
	G	Y	$(AG)_{15}$ and $(AC)_{15}$	*Melastoma dodecandrum*	Liu *et al.*, 2013b
	G	Y	TC_n, TG_n, and $GATA_n$	*Miscanthus* ssp. and *Saccharum* ssp.	Hodkinson *et al.*, 2013
	G	Y	$(AC)_{18}/(AG)_{18}/(ATG)_{12}$	*Myriophyllum spicatum*	Wu *et al.*, 2013
	G	Y	$(GT)_{15}$ and $(AG)_{15}$	*Phellodendron amurense*	Yu *et al.*, 2013
	G	Y	$(CAT)_{11}$, $(GCA)_6$, $(GATA)_{11}$, $(AAC)_{12}$, $(ATTT)_{10}(GC)_8$, $(GCGA)_5$, $(TTC)_{15}$ and $(GGT)_7$	*Pinus edulis* and *P. monophylla*	Krohn *et al.*, 2013
	G	Y	$(AC)_n$	*Pinus wangii*	Dou *et al.*, 2013
	T	**N**	Expressed sequence tags	*Pisum sativum*	Jain and McPhee 2013
	G	Y	$(GA)_{12}$ and $(CA)_{12}$	*Prunus sibirica*	Liu *et al.*, 2013a
	G	Y	$(AG)_{13}$ and $(TC)_{13}$	*Rhizophora mangle*	Ribeiro *et al.*, 2013
	G	Y	$(CCG)_6$, $(AAG)_8$, $(AGG)_6$, $(CT)_{13}$, $(AGC)_6$, $(AC)_{10,}$ and $(ATC)_6$	*Rhodiola ssp*	You *et al.*, 2013
	G	Y	$(CT)_8$ and $(GT)_8$	*Smilax brasiliensis*	Martins *et al.*, 2013
	G	Y	$(AC)_6(AG)_5$ or $(TC)_6(AC)_5$	*Tricyrtis macrantha*	Ohki and Setoguchi 2013
	G	Y	$(AC)_{13}$ and $(AG)_{13}$	*Vernicia fordii*	Pan *et al.*, 2013
Illumina	T	**N**	Expressed sequence tags from leaves	*Firmiana danxiaensis*	Fan *et al.*, 2013
	G	**N**	*	*Gleditsia triacanthos*	Owusu *et al.*, 2013
	T	**N**	Expressed sequence tags	*Spartina alterniflora*	Guo *et al.*, 2013
454	G	Y	CT and GT	*Anthyllis vulneraria*	Kesselring *et al.*, 2013
	G	Y	CT and GT	*Berberis microphylla*	Varas *et al.*, 2013
	G	Y	$(GT)_8$ $(TC)_{9.5}$, $(GTT)_{6.33}$, $(TTC)_7$, $(GTA)_{8.33}$, $(GTG)_{4.67}$, $(TCC)_5$, $(TTTG)_{2.5}$, $(TTTC)_6$, $(TTAC)_{6.75}$, and $(GATG)_{4.25}$	*Elaeagnus angustifolia*	Gaskin *et al.*, 2013
	G	**N**	*	*Melaleuca argentea*	Nevill *et al.*, 2013
	G	Y	Not informed	*Pachyrhizus* Rich. ex DC.	Delêtre *et al.*, 2013
	G	Y	$(GT)_{15}$, $(CT)_{15}$, $(GATA)_{10}$, $(GACA)_8$, and $(GATGT)_5$	*Phoradendron californicum*	Arroyo *et al.*, 2013
	T	**N**	Expressed sequence tags from stem	*Pisum sativum*	Zhuang *et al.*, 2013
	G	**N**	*	*Prosopis alba* and *P. chilensis*	Bessega *et al.*, 2013
	G	Y	$(CT)_{13}$ and $(GT)_{13}$	*Sebaea aurea*	Kissling *et al.*, 2013
	G	Y	TG,TC, AAC, AAG, AGG, ACG, ACAT and ACTC	*Thuja occidentalis*	Xu *et al.*, 2013a
	G	**N**	*	*Virola sebifera*	Wei *et al.*, 2013
	G	Y	$(GA)_{15}$, $(GTA)_8$, and $(TTC)_8$	*Argania spinosa*	Bahloul *et al.*, 2014

Table 1 (cont.)

Technology	Source[1]	Library[2]	Enrichment	Species	Reference
Sanger	G	Y	$(CT)_8$, $(GT)_8$	*Byrsonima cydoniifolia*	Bernardes *et al.*, 2014
	Cp	**N**	*	*Lemna minor*	Wani *et al.*, 2014
	G	Y	CT	*Lobelia inflata*	Hughes *et al.*, 2014
	G	Y	$(GT)_8$ and $(CT)_8$	*Passiflora* ssp.	Cerqueira-Silva *et al.*, 2014
	G	Y	$(CT)_8$ and $(GT)_8$	*Piper solmsianum*	Yoshida *et al.*, 2014
	G	Y	$(AC)_6(AG)_5$ or $(TC)_6(AC)_5$	*Scrophularia incisa*	Wang *et al.*, 2014
	G	Y	$(AC)_{15}$ and $(AG)_{15}$	*Spiraea* ssp	Khan *et al.*, 2014
	G	Y	$(AC)_6(AG)_5$, $(TC)_6(AC)_5$	*Vitex rotundifolia*	Ohtsuki *et al.*, 2014
	G	Y	$(GA)_n$ and $(GT)_n$	*Xanthosoma sagittifolium*	Cathebras *et al.*, 2014
Illumina	T	**N**	Expressed sequence tags from roots	*Buxus* spp.	Thammina *et al.*, 2014
	G	**N**	*	*Macadamia* ssp.	Nock *et al.*, 2014
	T	**N**	Expressed sequence tags from leaves	*Ostryopsis* ssp.	Liu *et al.*, 2014a
	G	**N**	*In silico* mining	*Phoenix dactylifera*	Aberlenc-Bertossi *et al.*, 2014
	G	**N**	*	*Solidago* L.	Beck *et al.*, 2014
	G	**N**	*	*Saxifraga granulata*	Meer *et al.*, 2014
454	G	Y	$(GA)_{15}$, $(GTA)_8$, and $(TTC)_8$	*Argania spinosa*	Bahloul *et al.*, 2014
	G	**N**	*	*Agave utahensis*	Byers *et al.*, 2014
	G	**N**	*	*Bidens alba*	Lu *et al.*, 2014
	G	Y	CT and GT	*Nephroma* ssp.	Belinchón *et al.*, 2014
	G	Y	TG, TC, AAC, AAG, AGG, ACG, ACAT, and ACTC	*Parietaria judaica*	Bossu *et al.*, 2014
Sanger	G	Y	$(CT)_8$ and $(GT)_8$	*Cabomba aquatica*	Barbosa *et al.*, 2015
	G	Y	$(AT)_8$, $(GA)_8$, and $(GAA)_8$	*Calibrachoa heterophylla*	Silva-Arias *et al.*, 2015
	G	Y	GA, GT, AGA, ACT, and ATC	*Campanula pyramidalis*	Radosavljeviæ *et al.*, 2015
	G	Y	$(AC)_{15}$ and $(AG)_{15}$	*Commelina communis*	Li *et al.*, 2015
	G	Y	$(AG)_{10}$	*Ilex chinensis*	Chen *et al.*, 2015
	G	Y	Not informed	*Fothergilla intermedia*	Hatmaker *et al.*, 2015
	G	Y	$(AC)_6(AG)_5$ or $(GA)_5(CA)_5$	*Hepatica nobilis* var. *japonica*	Kameoka *et al.*, 2015
	G	Y	$(CT)_8$ and $(GT)_8$	*Philcoxia minensis*	Scatigna *et al.*, 2015
	G	Y	$(AG)_{17}$, $(AC)_{17}$, $(AAC)_{10}$, $(CCG)_{10}$, $(CTG)_{10}$, and $(AAT)_{10}$	*Psittacanthus schiedeanus*	González *et al.*, 2015
	G	Y	$(AG)_{17}$, $(AC)_{17}$, $(AAC)_{10}$, $(CCG)_{10}$, $(CTG)_{10}$, and $(AAT)_{10}$	*Quillaja saponaria*	Letelier *et al.*, 2015
	G	Y	$(AC)_{15}$ and $(AG)_{15}$	*Saxifraga egregia*	Zhang *et al.*, 2015
	G	Y	$(TTC)_{10}$, $(CG)_{10}$, and $(GT)_{10}$	*Vellozia squamata*	Duarte-Barbosa *et al.*, 2015
Illumina	T/Cp	**N**	Expressed sequence tags from leaves	*Artocarpus moraceae*	Gardner *et al.*, 2015
	T	**N**	Expressed sequence tags from leaves	*Bombax ceiba*	Ju *et al.*, 2015
	T	**N**	Expressed sequence tags from leaves	*Carallia brachiata*	Qiang *et al.*, 2015
	G	**N**	*	*Dendrobium calamiforme*	Trapnell *et al.*, 2015
	T	**N**	Expressed sequence tags from leaves	*Lablab purpureus, Lathyrus sativus*	Chapman 2015
				Psophocarpus tetragonolobus and *Vigna subterranea*	

Table 1 (cont.)

Technology	Source[1]	Library[2]	Enrichment	Species	Reference
	T	**N**	Expressed sequence tags from leaves and cambium	*Thujopsis dolabrata var. hondae*	Sato *et al.*, 2015
454	G	**N**	*	*Cyperus fuscus*	Böckelmann *et al.*, 2015
	G	**N**	*In silico* mining	*Metasequoia glyptostroboides*	Jin *et al.*, 2015
	G	**N**	*	*Pilosella alpicola*	Vít *et al.*, 2015
	G	**N**	*	*Pulsatilla vulgaris*	DiLeo *et al.*, 2015
	G	Y	$(AG)_{10}$, $(AC)_{10}$, $(AAC)_8$, $(ACG)_8$, $(AAG)_8$, $(AGG)_8$, $(ACAT)_6$, and $(ATCT)_6$	*Quercus variabilis*	Wang *et al.*, 2015
	G	Y	CT and GT	*Salix humboldtiana*	Bozzi *et al.*, 2015
	G	**N**	*	*Silene acaulis*	Müller *et al.*, 2015
	G	Y	TG, TC, AAC, AAG, AGG, ACG, ACAT, and ACTC	*Veronica* subsect. *Pentasepalae*	López-González *et al.*, 2015
	G	**N**	*	*Vinca minor*	Moeller *et al.*, 2015

[1] G = genome, T = transcriptome, Cp = chloroplast DNA
[2] SSR enrichment library: Y, yes; **N**, no
* Total genomic DNA sequencing

Finally, visiting the website of the 24th edition of the Plant and Animal Genome (PAG) Conference (San Diego, CA) held in January 2016, we were able to find 63 workshops, abstracts and posters in which the term SSR was employed. We categorized these studies according to the groups of species analyzed and found the great majority of them (~90%) related to plants. We also checked references to SNPs and found about 150 studies, two-thirds related to plants and a third to domesticated animals (cattle, chicken, horse, pig, sheep and fish). A few (1.5% for SSRs and 5% for SNPs) proposed advances in experimental approaches or novel bioinformatics tools (https://pag.confex.com/pag/xxiv/meetingapp.cgi). Are SNPs destined to replace SSRs as the preferred marker? It seems clear that this will occur, but we do believe that SSRs will still be applicable in future plant genetic and genomic studies.

References

Aberlenc-Bertossi F, Castillo K, Tranchant-Dubreuil C, Chérif E, Ballardini M, Abdoulkader S, Gros-Balthazard M, Chabrillange N, Santoni S, Mercuri A *et al.* (2014) *In silico* mining of microsatellites in coding sequences of the date palm (Arecaceae) genome, characterization, and transferability. Appl Plant Sci 2:1300058.

Agarwal G, Sabbavarapu MM, Singh VK, Thudi M, Sheelamary S, Gaur PM and Varshney RK (2015) Identification of a non-redundant set of 202 *in silico* SSR markers and applicability of a select set in chickpea (*Cicer arietinum* L.). Euphytica 205:381-394.

Alic AS, Ruzafa D, Dopazo J and Blanquer I (2016) Objective review of *de novo* stand-alone error correction methods for NGS data: Objective review of de novo stand-alone NGS correctors. Wiley Interdiscip Rev Comput Mol Sci 6:111-1146.

Ando S, Kaneko S, Isagi Y, Repin R and Kitayama K (2013) Development of SSR markers for the tropical alpine tree species *Leptospermum recurvum* (Myrtaceae) on Mount Kinabalu in Borneo. Appl Plant Sci 1:1200010.

Angrizani RC, Contim LAS and Lemes MR (2013) Development and characterization of microsatellite markers for the endangered Amazonian tree *Aniba rosaeodora* (Lauraceae). Appl Plant Sci 1:1200516.

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796-815.

Araújo IS, Intorne AC, Pereira MG, Lopes UV and Filho GA de S (2007) Development and characterization of novel tetra-, tri- and di-nucleotide microsatellite markers in cacao (*Theobroma cacao* L.). Mol Breed 20:73-81.

Arroyo JM, Munguia-Vega A, Rodríguez-Estrella R and Bascompte J (2013) Isolation of 18 microsatellite loci in the desert mistletoe *Phoradendron californicum* (Santalaceae) via 454 pyrosequencing. Appl Plant Sci 1:1300048.

Bahloul YE, Dauchot N, Machtoun I, Gaboun F and Cutsem PV (2014) Development and characterization of microsatellite loci for the Moroccan endemic endangered species *Argania spinosa* (Sapotaceae). Appl Plant Sci 2:1300071.

Barbosa TDM, Trad RJ, Bajay MM and Amaral MCE (2015) Microsatellite markers isolated from *Cabomba aquatica* s.l. (Cabombaceae) from an enriched genomic library. Appl Plant Sci 3:1500076.

Beck JB, Semple JC, Brull JM, Lance SL, Phillips MM, Hoot SB and Meyer GA (2014) Genus-wide microsatellite primers for the goldenrods (Solidago; Asteraceae). Appl Plant Sci 2:1300093.

Belinchón R, Ellis CJ and Yahr R (2014) Microsatellite loci in two epiphytic lichens with contrasting dispersal modes:

Nephroma laevigatum and *N. parile* (Nephromataceae). Appl Plant Sci 2:1400080.

Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res 27:573-580.

Bernardes V, Anjos DE dos, Gondim SG de CA, Murakami DM, Bizão N and Telles MP de C (2014) Isolation and characterization of microsatellite loci in *Byrsonima cydoniifolia* (Malpighiaceae) and cross-amplification in *B. crassifolia*. Appl Plant Sci 2:1400016. d

Bessega CF, Pometti CL, Miller JT, Watts R, Saidman BO and Vilardi JC (2013) New microsatellite loci for *Prosopis alba* and *P. chilensis* (Fabaceae). Appl Plant Sci 1:1200324.

Billotte N, Lagoda PJL, Risterucci A-M and Baurens F-C (1999) Microsatellite-enriched libraries: applied methodology for the development of SSR markers in tropical crops. Fruits 54:277-288.

Bird TD (2015) Myotonic dystrophy type 1. 1999 Sep 17 [Updated 2015 Oct 22]. In: Pagon RA, Adam MP, Ardinger HH, Wallace SE, Amemiya A, Bean LJ, Bird TD, Fong C-T, Mefford HC, Smith RJ, *et al.* (eds) GeneReviews, University of Washington, Seattle (WA), http://www.ncbi.nlm.nih.gov/books/NBK1165/.

Blair MW, Torres MM, Pedraza F, Giraldo MC, Buendía HF and Hurtado N (2009) Development of microsatellite markers for common bean (*Phaseolus vulgaris* L.) based on screening of non-enriched, small-insert genomic libraries. Genome 52:772-782.

Böckelmann J, Wieser D, Tremetsberger K, Šumberová K and Bernhardt K-G (2015) Isolation of nuclear microsatellite markers for *Cyperus fuscus* (Cyperaceae). Appl Plant Sci 3:1500071.

Bossu A, Bertaudière-Montès V, Dubut V and Manel S (2014) Microsatellite primers in *Parietaria judaica* (Urticaceae) to assess genetic diversity and structure in urban landscapes. Appl Plant Sci 2:1400036.

Bozzi JA, Liepelt S, Ohneiser S, Gallo LA, Marchelli P, Leyer I, Ziegenhagen B and Mengel C (2015) Characterization of 23 polymorphic SSR markers in *Salix humboldtiana* (Salicaceae) using next-generation sequencing and cross-amplification from related species. Appl Plant Sci 3:1400120.

Brown SM, Hopkins MS, Mitchell SE, Senior ML, Wang TY, Duncan RR, Gonzalez-Candelas F and Kresovich S (1996) Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum [*Sorghum bicolor* (L.) Moench]. Theor Appl Genet 93:190-198.

Bryan GJ, McNicoll J, Ramsay G, Meyer RC and De Jong WS (1999) Polymorphic simple sequence repeat markers in chloroplast genomes of Solanaceous plants. Theor Appl Genet 99:859-867.

Buschiazzo E and Gemmell NJ (2006) The rise, fall and renaissance of microsatellites in eukaryotic genomes. Bioessays 28:1040-1050.

Byers C, Maughan PJ, Clouse J and Stewart JR (2014) Microsatellite primers in *Agave utahensis* (Asparagaceae), a keystone species in the Mojave Desert and Colorado Plateau. Appl Plant Sci 2:1400047.

Castelo AT, Martins W and Gao GR (2002) TROLL—Tandem Repeat Occurrence Locator. Bioinformatics 18:634-636.

Cathebras C, Traore R, Malapa R, Risterucci A-M and Chaïr H (2014) Characterization of microsatellites in *Xanthosoma sagittifolium* (Araceae) and cross-amplification in related species. Appl Plant Sci 2:1400027.

Cerqueira-Silva CBM, Santos ESL, Vieira JGP, Mori GM, Jesus ON, Corrêa RX and Souza AP (2014) New microsatellite markers for wild and commercial species of *Passiflora* (Passifloraceae) and cross-amplification. Appl Plant Sci 2:1300061.

Chapman MA (2015) Transcriptome sequencing and marker development for four underutilized legumes. Appl Plant Sci 3:1400111.

Chen W-W, Xiao Z-Z, Tong X, Liu Y-P and Li Y-Y (2015) Development and characterization of 25 microsatellite primers for *Ilex chinensis* (Aquifoliaceae). Appl Plant Sci 3:1500057.

Csencsics D, Brodbeck S and Holderegger R (2010) Cost-effective, species-specific microsatellite development for the endangered dwarf bulrush (*Typha minima*) using next-generation sequencing technology. J Hered 101:789-793.

Culley TM, Stamper TI, Stokes RL, Brzyski JR, Hardiman NA, Klooster MR and Merritt BJ (2013) An efficient technique for primer development and application that integrates fluorescent labeling and multiplex PCR. Appl. Plant Sci. 1:1300027.

Delêtre M, Soengas B, Utge J, Lambourdière J and Sørensen M (2013) Microsatellite markers for the yam bean *Pachyrhizus* (Fabaceae). Appl Plant Sci 1:1200551.

DiLeo MF, Graf R, Holderegger R, Rico Y and Wagner HH (2015) Highly polymorphic microsatellite markers in *Pulsatilla vulgaris* (Ranunculaceae) using next-generation sequencing. Appl Plant Sci 3:1500031.

Dou J-J, Zhou R-C, Tang A-J, Ge X-J and Wu W (2013) Development and characterization of nine microsatellites for an endangered tree, *Pinus wangii* (Pinaceae). Appl Plant Sci 1:1200134.

Duarte-Barbosa M, Bajay MM, Zucchi MI and Pivello VR (2015) Development and characterization of 47 novel microsatellite markers for *Vellozia squamata* (Velloziaceae). Appl Plant Sci 3:1400087.

Ebert D and Peakall R (2009) A new set of universal de novo sequencing primers for extensive coverage of noncoding chloroplast DNA: new opportunities for phylogenetic studies and cpSSR discovery. Mol Ecol Resour 9:777-783.

Egan AN, Schlueter J and Spooner DM (2012) Applications of next-generation sequencing in plant biology. Am J Bot 99:175-185.

Fan Q, Chen S, Li M, He S, Zhou R and Liao W (2013) Development and characterization of microsatellite markers from the transcriptome of *Firmiana danxiaensis* (Malvaceae s.l.). Appl Plant Sci 1:1300047.

Fujimori S, Washio T, Higo K, Ohtomo Y, Murakami K, Matsubara K, Kawai J, Carninci P, Hayashizaki Y, Kikuchi S *et al.* (2003) A novel feature of microsatellites in plants: a distribution gradient along the direction of transcription. FEBS Lett 554:17-22.

Gao C, Ren X, Mason AS, Li J, Wang W, Xiao M and Fu D (2013) Revisiting an important component of plant genomes: microsatellites. Funct Plant Biol 40:645.

Gao L, Tang J, Li H and Jia J (2003) Analysis of microsatellites in major crops assessed by computational and experimental approaches. Mol Breed 12:245-261.

Garcia FAA, Kido AE, Meza NA, Souza BHM, Pinto RL, Pastina MM, Leite SC, Silva GJA da, Ulian CE, Figueira A *et al.* (2006) Development of an integrated genetic map of a sugarcane (*Saccharum* spp.) commercial cross, based on a maximum-likelihood approach for estimation of linkage and linkage phases. Theor Appl Genet 112:298-314.

Gardner EM, Laricchia KM, Murphy M, Ragone D, Scheffler BE, Simpson S, Williams EW and Zerega NJC (2015) Chloroplast microsatellite markers for *Artocarpus* (Moraceae) developed from transcriptome sequences. Appl Plant Sci 3:1500049.

Gaskin JF, Hufbauer RA and Bogdanowicz SM (2013) Microsatellite markers for Russian olive (*Elaeagnus angustifolia*; Elaeagnaceae). Appl Plant Sci 1:1300013.

Gemayel R, Cho J, Boeynaems S and Verstrepen KJ (2012) Beyond Junk-Variable Tandem repeats as facilitators of rapid evolution of regulatory and coding sequences. Genes 3:461-480.

Golubov A, Yao Y, Maheshwari P, Bilichak A, Boyko A, Belzile F and Kovalchuk I (2010) Microsatellite instability in *Arabidopsis* increases with plant Development. Plant Physiol 154:1415-1427.

González C, Harvey N and Ornelas JF (2015) Development and characterization of microsatellite loci in the mistletoe *Psittacanthus schiedeanus* (Loranthaceae). Appl Plant Sci 3:1400099.

Grohme MA, Soler RF, Wink M and Frohme M (2013) Microsatellite marker discovery using single molecule real-time circular consensus sequencing on the Pacific Biosciences RS. BioTechniques 55:253-256.

Guichoux E, Lagache L, Wagner S, Chaumeil P, LéGer P, Lepais O, Lepoittevin C, Malausa T, Revardel E, Salin F *et al.* (2011) Current trends in microsatellite genotyping. Mol Ecol Resour 11:591-611.

Guo W, Huang Y, He Z, Yan Y, Zhou R and Shi S (2013) Development and characterization of microsatellite loci for smooth cordgrass, *Spartina alterniflora* (Poaceae). Appl Plant Sci 1:1200211.

Hanai LR, de Campos T, Camargo LEA, Benchimol LL, de Souza AP, Melotto M, Carbonell SAM, Chioratto AF, Consoli L, Formighieri EF *et al.* (2007) Development, characterization, and comparative analysis of polymorphism at common bean SSR loci isolated from genic and genomic sources. Genome 50:266-277.

Hatmaker EA, Wadl PA, Mantooth K, Scheffler BE, Ownley BH and Trigiano RN (2015) Development of microsatellites from *Fothergilla intermedia* (Hamamelidaceae) and cross transfer to four other genera within Hamamelidaceae. Appl Plant Sci 3:1400123.

Hayden MJ, Nguyen TM, Waterman A and Chalmers KJ (2008) Multiplex-ready PCR: A new method for multiplexed SSR and SNP genotyping. BMC Genomics 9:80.

Hayward AC, Tollenaere R, Dalton-Morgan J and Batley J (2015) Molecular marker applications in plants. In: Batley J (ed) Plant Genotyping. Springer, New York, NY, pp 13-27

Hmeljevski KV, Ciampi MB, Baldauf C, Reis MSD and Forzza RC (2013) Development of SSR markers for *Encholirium horridum* (Bromeliaceae) and transferability to other Pitcairnioideae. Appl Plant Sci 1:1200445.

Hodkinson TR, Cesare M de and Barth S (2013) Nuclear SSR markers for *Miscanthus*, *Saccharum*, and related grasses (Saccharinae, Poaceae). Appl Plant Sci 1:1300042.

Horie N, Aiba H, Oguro K, Hojo H and Takeishi K (1995) Functional analysis and DNA polymorphism of the tandemly repeated sequences in the 5'-terminal regulatory region of the human gene for thymidylate synthase. Cell Struct Funct 20:191-197.

Hughes PW, Jaworski AF, Davis CS, Aitken SM and Simons AM (2014) Development of polymorphic microsatellite markers for Indian tobacco, *Lobelia inflata* (Campanulaceae). Appl Plant Sci 2:1300096.

Huntley M and Golding GB (2000) Evolution of simple sequence in proteins. J Mol Evol 51:131-140.

Iorizzo M, Senalik DA, Grzebelus D, Bowman M, Cavagnaro PF, Matvienko M, Ashrafi H, Van Deynze A and Simon PW (2011) De novo assembly and characterization of the carrot transcriptome reveals novel genes, new markers, and genetic diversity. BMC Genomics 12:389.

Ishibashi CDA, Shaver AR, Perrault DP, Davis SD and Honeycutt RL (2013) Isolation of microsatellite markers in a chaparral species endemic to Southern California, *Ceanothus megacarpus* (Rhamnaceae). Appl Plant Sci 1:1200393.

Jain S and McPhee KE (2013) Isolation and characterization of novel EST-derived genic markers in *Pisum sativum* (Fabaceae). Appl Plant Sci 1:1300026.

Jarne P and Lagoda PJ (1996) Microsatellites, from molecules to populations and back. Trends Ecol Evol 11:424-429.

Jin Y, Bi Q, Guan W and Mao J-F (2015) Development of 23 novel polymorphic EST-SSR markers for the endangered relict conifer *Metasequoia glyptostroboides*. Appl Plant Sci 3:1500038.

Jonah P, Bello L, Lucky O, Midau A and Moruppa S (2011) Review: the importance of molecular markers in plant breeding programmes. Glob J Sci Front Res 11:eV-vers1.

Ju M-M, Ma H-C, Xin P-Y, Zhou Z-L and Tian B (2015) Development and characterization of EST-SSR markers in *Bombax ceiba* (Malvaceae). Appl Plant Sci 3:1500001.

Kalia RK, Rai MK, Kalia S, Singh R and Dhawan AK (2011) Microsatellite markers: an overview of the recent progress in plants. Euphytica 177:309-334.

Kaló P, Seres A, Taylor SA, Jakab J, Kevei Z, Kereszt A, Endre G, Ellis THN and Kiss GB (2004) Comparative mapping between *Medicago sativa* and *Pisum sativum*. Mol Genet Genomics 272:235-246. doi: 10.1007/s00438-004-1055-z

Kameoka S, Higashi H and Setoguchi H (2015) Development of polymorphic microsatellite loci in the perennial herb *Hepatica nobilis* var. *japonica* (Ranunculaceae). Appl Plant Sci 3:1400114.

Kashi Y and King D (2006) Simple sequence repeats as advantageous mutators in evolution. Trends Genet 22:253-259.

Katti MV, Sami-Subbu R, Ranjekar PK and Gupta VS (2000) Amino acid repeat patterns in protein sequences: Their diversity and structural-functional implications. Protein Sci 9:1203-1209.

Kejnovsky E, Hobza R, Cermak T, Kubat Z and Vyskot B (2009) The role of repetitive DNA in structure and evolution of sex chromosomes in plants. Heredity 102:533-541.

Kesselring H, Hamann E, Stöcklin J and Armbruster GFJ (2013) New microsatellite markers for *Anthyllis vulneraria* (Fabaceae), analyzed with spreadex gel electrophoresis. Appl Plant Sci 1:1300054.

Khan G, Zhang F, Gao Q, Jiao X, Fu P, Xing R, Zhang J and Chen S (2014) Isolation of 16 microsatellite markers for *Spiraea alpina* and *S. mongolica* (Rosaceae) of the Qinghai-Tibet Plateau. Appl Plant Sci 2:1300059.

Kissling J, Bachmann O, Thali MR and Segarra-Moragues JG (2013) Novel microsatellite loci for *Sebaea aurea* (Gentianaceae) and cross-amplification in related species. Appl Plant Sci 1:1300056.

Kofler R, Schlotterer C and Lelley T (2007) SciRoKo: a new tool for whole genome microsatellite search and investigation. Bioinformatics 23:1683-1685.

Krohn AL, Flores-Rentería L and Gehring CA (2013) Microsatellite primers in the foundation tree species *Pinus edulis* and *P. monophylla* (Pinaceae). Appl Plant Sci 1:1200552. doi: 10.3732/apps.1200552

Kuleung C, Baenziger PS and Dweikat I (2004) Transferability of SSR markers among wheat, rye, and triticale. Theor Appl Genet 108:1147-1150.

Lawson MJ and Zhang L (2006) Distinct patterns of SSR distribution in the *Arabidopsis thaliana* and rice genomes. Genome Biol 7:R14.

Lawson MJ and Zhang L (2008) Housekeeping and tissue-specific genes differ in simple sequence repeats in the 5'-UTR region. Gene 407:54-62.

Lepais O and Bacles CFE (2011) *De novo* discovery and multiplexed amplification of microsatellite markers for black alder (*Alnus glutinosa*) and related species using SSR-enriched shotgun pyrosequencing. J Hered 102:627-632.

Letelier L, Harvey N, Valderrama A, Stoll A and González-Rodríguez A (2015) Isolation and characterization of 12 microsatellite loci in soapbark, *Quillaja saponaria* (Quillajaceae). Appl Plant Sci 3:1500024.

Li CD, Rossnagel BG and Scoles GJ (2000) The development of oat microsatellite markers and their use in identifying relationships among *Avena* species and oat cultivars. Theor Appl Genet 101:1259-1268.

Li J-K, Song Y-P, Xu H, Zhu J-Y and Tang L-L (2015) Development and characterization of microsatellite loci for the Pseudometallophyte *Commelina communis* (Commelinaceae). Appl Plant Sci 3:1400098.

Li L, Li Z, Li K, Huang B and Xu L (2013) Development and characterization of EST-SSR markers in the Chinese medicinal plant *Callerya speciosa* (Fabaceae). Appl Plant Sci 1:1200345.

Li Y-C, Korol AB, Fahima T, Beiles A and Nevo E (2002) Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Mol Ecol 11:2453-2465.

Li Y-C, Korol AB, Fahima T and Nevo E (2004) Microsatellites within genes: structure, function, and evolution. Mol Biol Evol 21:991-1007.

Liu B-B, Tian B, Ma H, Lu Z-Q, Qiu Q, Mao K-S and Liu J-Q (2014a) Development and characterization of EST-SSR

markers in *Ostryopsis* (Betulaceae). Appl Plant Sci 2:1300062.

Liu F, Hu Z, Liu W, Li J, Wang W, Liang Z, Wang F and Sun X (2016) Distribution, function and evolution characterization of microsatellite in *Sargassum thunbergii* (Fucales, Phaeophyta) transcriptome and their application in marker development. Sci Rep 6:18947.

Liu H-B, Liu J, Wang Z, Ma L-Y, Wang S-Q, Lin X-G, Wu R-L and Pang X-M (2013a) Development and characterization of microsatellite markers in *Prunus sibirica* (Rosaceae). Appl Plant Sci 1:1200074.

Liu T, Dai S, Wu W, Zhang R, Fan Q, Shi S and Zhou R (2013b) Development and characterization of microsatellite markers for *Melastoma dodecandrum* (Melastomataceae). Appl Plant Sci 1:1200294.

Liu Y, He D, Cai M, Tang W, Li X-Y, Pan H-T and Zhang Q-X (2013c) Development of microsatellite markers for *Lagerstroemia indica* (Lythraceae) and related species. Appl Plant Sci 1:1200203.

Liu Y, Wang L, Chen D, Wu X, Huang D, Chen L, Li L, Deng X and Xu Q (2014b) Genome-wide comparison of microRNAs and their targeted transcripts among leaf, flower and fruit of sweet orange. BMC Genomics 15:695. doi: 10.1186/1471-2164-15-695.

López-González N, Mayland-Quellhorst E, Pinto-Carrasco D and Martínez-Ortega MM (2015) Characterization of 12 polymorphic SSR markers in *Veronica* Subsect. *Pentasepalae* (Plantaginaceae) and cross-amplification in 10 other subgenera. Appl Plant Sci 3:1500059.

Lu Y-B, Huang D-L, Wang X, Wu Z-J and Tang S-Q (2014) Microsatellite markers for the invasive species *Bidens alba* (Asteraceae). Appl Plant Sci 2:1400008.

Maio AD and Castro OD (2013) SSR-Patchwork: An optimized protocol to obtain a rapid and inexpensive SSR library using first-generation sequencing technology. Appl Plant Sci 1:1200158.

Mammadov J, Aggarwal R, Buyyarapu R and Kumpatla S (2012) SNP markers and their impact on plant breeding. Int J Plant Genomics 2012:1-11.

Martins AR, Abreu AG, Bajay MM, Villela PMS, Batista CEA, Monteiro M, Alves-Pereira A, Figueira GM, Pinheiro JB, Appezzato-Da-Glória B *et al.* (2013) Development and characterization of microsatellite markers for the medicinal plant *Smilax brasiliensis* (Smilacaceae) and Related Species. Appl Plant Sci 1:1200507.

Mason AS (2015) SSR Genotyping. In: Batley J (ed) Plant Genotyping. Springer, New York, NY, pp 77-89.

Meer S van der, Houdt JKJV, Maes GE, Hellemans B and Jacquemyn H (2014) Microsatellite primers for the gynodioecious grassland perennial *Saxifraga granulata* (Saxifragaceae). Appl Plant Sci 2:1400040.

Morgante M, Hanafey M and Powell W (2002) Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. Nat Genet 30:194-200.

Müller E, Hlaváčková I, Svoen ME, Alsos IG and Eidesen PB (2015) Characterization of 14 microsatellite markers for *Silene acaulis* (Caryophyllaceae). Appl Plant Sci 3:1500036.

Nalavade R, Griesche N, Ryan DP, Hildebrand S and Krau S (2013) Mechanisms of RNA-induced toxicity in CAG repeat disorders. Cell Death Dis 4:e752.

Nevill PG, Williams A, Krauss S, Bradbury D, Samaraweera S and Gardner MG (2013) Development of microsatellite loci for the riparian tree species *Melaleuca argentea* (Myrtaceae) using 454 sequencing. Appl Plant Sci 1:1200401.

Nevo E (2001) Evolution of genome-phenome diversity under environmental stress. Proc Natl Acad Sci U S A 98:6233-6240.

Nock CJ, Elphinstone MS, Ablett G, Kawamata A, Hancock W, Hardner CM and King GJ (2014) Whole genome shotgun sequences for microsatellite discovery and application in cultivated and wild *Macadamia* (Proteaceae). Appl Plant Sci 2:1300089.

Nunome T, Negoro S, Miyatake K, Yamaguchi H and Fukuoka H (2006) A protocol for the construction of microsatellite enriched genomic library. Plant Mol Biol Report 24:305-312.

Oetting WS, Lee HK, Flanders DJ, Wiesner GL, Sellers TA and King RA (1995) Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers. Genomics 30:450-458.

Ohki N and Setoguchi H (2013) New microsatellite markers for *Tricyrtis macrantha* (Convallariaceae) and cross-amplification in closely related species. Appl Plant Sci 1:1200247.

Ohtsuki T, Shoda T, Kaneko Y and Setoguchi H (2014) Development of microsatellite markers for *Vitex rotundifolia* (Verbenaceae), an endangered coastal plant in Lake Biwa, Japan. Appl Plant Sci 2:1300100.

Oliveira EJ, Pádua JG, Zucchi MI, Vencovsky R and Vieira MLC (2006) Origin, evolution and genome distribution of microsatellites. Genet Mol Biol 29:294-307.

Owusu SA, Staton M, Jennings TN, Schlarbaum S, Coggeshall MV, Romero-Severson J, Carlson JE and Gailing O (2013) Development of genomic microsatellites in *Gleditsia triacanthos* (Fabaceae) using Illumina sequencing. Appl Plant Sci 1:1300050.

Pan Y, Pan L, Chen L, Zhang L, Nevo E and Peng J (2013) Development of microsatellite markers in the oil-producing species *Vernicia fordii* (Euphorbiaceae), a potential biodiesel feedstock. Appl Plant Sci 1:1200004.

Penha HP, Pereira GS, Zucchi MI, Diniz AL, Diniz AL and Vieira MLC (2013) Microsatellite markers in sweet passion fruit, and identification of length and conformation polymorphisms within repeat sequences. Plant Breed 132:732-735.

Pereira GS, Nunes ES, Laperuta LDC, Braga MF, Penha HA, Diniz AL, Munhoz CF, Gazaffi R, Garcia AAF and Vieira MLC (2013) Molecular polymorphism and linkage analysis in sweet passion fruit, an outcrossing species: Molecular map in sweet passion fruit. Ann Appl Biol 162:347–361. doi: 10.1111/aab.12028.

Pérez-Jiménez M, Besnard G, Dorado G and Hernandez P (2013) varietal tracing of virgin olive oils based on plastid DNA variation profiling. PLoS One 8:e70507.

Phumichai C, Phumichai T and Wongkaew A (2015) Novel chloroplast microsatellite (cpSSR) markers for genetic diversity assessment of cultivated and wild *Hevea* rubber. Plant Mol Biol Report 33:1486-1498.

Poltronieri J, Marquioni V, Bertollo LAC, Kejnovsky E, Molina WF, Liehr T and Cioffi MB (2014) Comparative chromosomal mapping of microsatellites in *Leporinus* species (Characiformes, Anostomidae): Unequal accumulation on the W chromosomes. Cytogenet Genome Res 142:40-45.

Qiang Y, Xie H, Qiao S, Yuan Y, Liu Y, Shi X, Shu M, Jin J, Shi S, Tan F *et al.* (2015) Development of microsatellite markers for *Carallia brachiata* (Rhizophoraceae). Appl Plant Sci 3:1400125.

Qu J and Liu J (2013) A genome-wide analysis of simple sequence repeats in maize and the development of polymorphism markers from next-generation sequence data. BMC Res Notes 6:403.

Radosavljeviæ I, Jakse J, Satovic Z, Javornik B, Jankoviæ I and Liber Z (2015) New microsatellite markers for *Campanula pyramidalis* (Campanulaceae) and cross-amplification in closely related species. Appl Plant Sci 3:1400117.

Ramsay L, Macaulay M, Cardle L, Morgante M, Ivanissevich S degli, Maestri E, Powell W and Waugh R (1999) Intimate association of microsatellite repeats with retrotransposons and other dispersed repetitive elements in barley. Plant J 17:415-425.

Ranum LP and Day JW (2002) Dominantly inherited, non-coding microsatellite expansion disorders. Curr Opin Genet Dev 12:266-271.

Ribeiro DO, Vinson CC, Nascimento DSS, Mehlig U, Menezes MPM, Sampaio I and Silva MB (2013) Isolation of microsatellite markers for the red mangrove, *Rhizophora mangle* (Rhizophoraceae). Appl Plant Sci 1:1300003.

Rodriguez E, Azevedo R, Moreira H, Souto L and Santos C (2013) Pb^{2+} exposure induced microsatellite instability in *Pisum sativum* in a locus related with glutamine metabolism. Plant Physiol Biochem 62:19-22.

Sakazono S, Hiramatsu M, Watanabe M and Okubo H (2013) Development and characterization of microsatellite markers for *Lilium longiflorum* (Liliaceae). Appl Plant Sci 1:1300014.

Sanger F and Coulson AR (1975) A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J Mol Biol 94:441-448.

Sato M, Hasegawa Y, Mishima K and Takata K (2015) Isolation and characterization of 22 EST-SSR markers for the genus *Thujopsis* (Cupressaceae). Appl Plant Sci 3:1400101.

Scatigna AV, Oliveira FA, Mantello CC, Francisco PM, Souza AP and Simões AO (2015) Microsatellite markers for studies with the carnivorous plant *Philcoxia minensis* (Plantaginaceae). Appl Plant Sci 3:1500035.

Schlötterer C (1998) Genome evolution: Are microsatellites really simple sequences? Curr Biol 8:R132-R134.

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L and Graves TA (2009) The B73 maize genome: complexity, diversity, and dynamics. Science 326:1112-1115.

Semagn K, Bjørnstad Å and Ndjiondjop MN (2006) An overview of molecular marker methods for plants. Afr J Biotechnol 5:2540-2568.

Senan S, Kizhakayil D, Sasikumar B and Sheeja TE (2014) Methods for development of microsatellite markers: an overview. Not Sci Biol 6:1-13.

Silva-Arias GA, Mäder G, Bonatto SL and Freitas LB (2015) Novel microsatellites for *Calibrachoa heterophylla*

(Solanaceae) endemic to the south Atlantic coastal plain of South America. Appl Plant Sci 3:1500021.

Souza LM, Gazaffi R, Mantello CC, Silva CC, Garcia D, Le Guen V, Cardoso SEA, Garcia AAF and Souza AP (2013) QTL Mapping of growth-related traits in a full-sib family of rubber tree (*Hevea brasiliensis*) evaluated in a sub-tropical climate. PLoS One 8:e61238.

Sun JX, Mullikin JC, Patterson N and Reich DE (2009) Microsatellites are molecular clocks that support accurate inferences about history. Mol Biol Evol 26:1017-1027.

Sureshkumar S, Todesco M, Schneeberger K, Harilal R, Balasubramanian S and Weigel D (2009) A genetic defect caused by a triplet repeat expansion in *Arabidopsis thaliana*. Science 323:1060-1063.

Tambarussi EV, Sebbenn AM, Moreno MA, Ferraz EM, Kageyama PY and Vencovsky R (2013) Microsatellite markers for *Cariniana legalis* (Lecythidaceae) and their transferability to C. estrellensis. Appl Plant Sci 1:1200493.

Tautz D and Renz M (1984) Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucleic Acids Res 12:4127-4138.

Temnykh S (2001) Computational and experimental analysis of microsatellites in rice (Oryza sativa L.): Frequency, length variation, transposon associations, and genetic marker potential. Genome Res 11:1441-1452.

Temnykh S, DeClerck G, Lukashova A, Lipovich L, Cartinhour S and McCouch S (2001) Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. Genome Res 11:1441-1452.

Thammina CS, Olsen RT, Malapi-Wight M, Crouch JA and Pooler MR (2014) Development of polymorphic genic-SSR markers by cDNA library sequencing in boxwood, *Buxus* spp. (Buxaceae). Appl Plant Sci 2:1400095. d

Thiel T, Michalek W, Varshney R and Graner A (2003) Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). Theor Appl Genet 106:411-422.

Toutenhoofd SL, Garcia F, Zacharias DA, Wilson RA and Strehler EE (1998) Minimum CAG repeat in the human calmodulin-1 gene 5' untranslated region is required for full expression. Biochim Biophys Acta BBA - Gene Struct Expr 1398:315-320.

Tranbarger TJ, Kluabmongkol W, Sangsrakru D, Morcillo F, Tregear JW, Tragoonrung S and Billotte N (2012) SSR markers in transcripts of genes linked to post-transcriptional and transcriptional regulatory functions during vegetative and reproductive development of *Elaeis guineensis*. BMC Plant Biol 12:e1.

Trapnell DW, Beasley RR, Lance SL, Field AR and Jones KL (2015) Characterization of microsatellite loci for an *Australian* epiphytic orchid, *Dendrobium calamiforme*, using Illumina sequencing. Appl Plant Sci 3:1500016.

Varas B, Castro MH, Rodriguez R, Baer DV, Mardones C and Hinrichsen P (2013) Identification and characterization of microsatellites from calafate (*Berberis microphylla*, Berberidaceae). Appl Plant Sci 1:1200003.

Vít P, ingliarová B, Zozomová-Lihová J, Marhold K and Krak K (2015) Microsatellite markers for the *Pilosella alpicola*

group (*Hieraciinae*, Asteraceae) and their cross-amplification in other Hieraciinae genera. Appl Plant Sci 3:1500048.

Wadl PA, Hatmaker EA, Fussi B, Scheffler BE and Trigiano RN (2013) Isolation and characterization of microsatellite loci for *Cornus sanguinea* (Cornaceae). Appl Plant Sci 1:1300012.

Wang R-H, Chen C, Ma Q, Li P and Fu C-X (2014) Development of microsatellite loci in *Scrophularia incisa* (Scrophulariaceae) and cross-amplification in congeneric species. Appl Plant Sci 2:1300077.

Wang X, Li J and Li Y (2015) Isolation and characterization of microsatellite markers for an endemic tree in East Asia, *Quercus variabilis* (Fagaceae). Appl Plant Sci 3:1500032.

Wani GA, Shah MA, Reshi ZA, Atangana AR and Khasa DP (2014) cpDNA Microsatellite markers for *Lemna minor* (Araceae): Phylogeographic implications. Appl Plant Sci 2:1300099.

Wei N, Dick CW, Lowe AJ and Gardner MG (2013) Polymorphic microsatellite loci for *Virola sebifera* (Myristicaceae) derived from shotgun 454 pyrosequencing. Appl Plant Sci 1:1200295.

Weising K and Gardner RC (1999) A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. Genome 42:9-19.

Wu Z-G, Yu D and Xu X-W (2013) Development of microsatellite markers in the hexaploid aquatic macrophyte, *Myriophyllum spicatum* (Haloragaceae). Appl Plant Sci 1:1200230.

Xu H, Tremblay F and Bergeron Y (2013a) Development and multiplexed amplification of SSR markers for *Thuja occidentalis* (Cupressaceae) using shotgun pyrosequencing. Appl Plant Sci 1:1200427. 7

Xu J, Liu L, Xu Y, Chen C, Rong T, Ali F, Zhou S, Wu F, Liu Y, Wang J *et al.* (2013b) Development and characterization of simple sequence repeat markers providing genome-wide coverage and high resolution in maize. DNA Res 20:497-509.

Yamashiro A, Yamashiro T and Tateishi Y (2013) Isolation and characterization of microsatellite markers for *Canavalia cathartica* and *C. lineat*a (Fabaceae). Appl Plant Sci 1:1200111.

Yoshida NC, Lima PF, Priolli RHG, Kato MJ and Colombo CA (2014) Isolation and characterization of nine polymorphic microsatellite loci in *Piper solmsianum* (Piperaceae). Appl Plant Sci 2:1300092.

You J, Liu W, Zhao Y, Zhu Y, Zhang W, Wang Y, Lu F and Song Z (2013) Microsatellite markers in *Rhodiola* (Crassulaceae), a medicinal herb genus widely used in traditional Chinese Medicine. Appl Plant Sci 1:1200219.

Yu J-H, Chen C-M, Han S-J, Guo X-R, Yuan S-S, Wang C-J and Zu Y-G (2013) Development and characterization of polymorphic microsatellite loci in *Phellodendron amurense* (Rutaceae). Appl Plant Sci 1:1200321.

Zalapa JE, Cuevas H, Zhu H, Steffan S, Senalik D, Zeldin E, McCown B, Harbut R and Simon P (2012) Using next-generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences. Am J Bot 99:193-208.

Zhang F, Li Y, Gao Q, Lei S, Khan G, Yang H and Chen S (2015) Development and characterization of polymorphic microsatellite loci for *Saxifraga egregia* (Saxifragaceae). Appl Plant Sci 3:1500037.

Zhang L, Yuan D, Yu S, Li Z, Cao Y, Miao Z, Qian H and Tang K (2004) Preference of simple sequence repeats in coding and non-coding regions of *Arabidopsis thaliana*. Bioinformatics 20:1081-1086.

Zhang L, Zhang Z-R and Gao L-M (2013) Isolation and characterization of 27 microsatellite markers for the endemic species *Diplarche multiflora* (Ericaceae). Appl Plant Sci 1:1200235.

Zhao Z, Guo C, Sutharzan S, Li P, Echt CS, Zhang J and Liang C (2014) Genome-wide analysis of tandem repeats in plants and green algae. G3 Genes Genomes Genet 4:67-78.

Zhuang X, McPhee KE, Coram TE, Peever TL and Chilvers MI (2013) Development and characterization of 37 novel EST-SSR markers in *Pisum sativum* (Fabaceae). Appl Plant Sci 1:1200249.

Zhu H (2005) Bridging model and crop legumes through comparative genomics. Plant Physiol 137:1189-1196.

Supplementary material

The following online material is available for this article:
Table S1 - List of 933 publications on microsatellite markers (WOS 2010-2015).

Associate Editor: Klaus Hartfelder