

Review

Simple Sequence Repeat Polymorphisms (SSRPs) for Evaluation of Molecular Diversity and Germplasm Classification of Minor Crops

Yong-Jin Park ¹, Ju Kyong Lee ² and Nam-Soo Kim ^{3,4,*}

¹ Department of Plant Sciences, Kongju National University, Yesan 340-702, Korea

² Division of Bio-Resources Technology, Kangwon National University, Chunchon 200-701, Korea

³ Department of Molecular Biosciences, Kangwon National University, Chunchon 200-701, Korea

⁴ Institute of Biosciences and Biotechnology, Kangwon National University, Chunchon 200-701, Korea

* Author to whom correspondence should be addressed; E-Mail: kimnamsu@kangwon.ac.kr.

Received: 13 September 2009; in revised form: 5 November 2009 / Accepted: 10 November 2009 / Published: 10 November 2009

Abstract: Evaluation of the genetic diversity among populations is an essential prerequisite for the preservation of endangered species. Thousands of new accessions are introduced into germplasm institutes each year, thereby necessitating assessment of their molecular diversity before elimination of the redundant genotypes. Of the protocols that facilitate the assessment of molecular diversity, SSRPs (simple sequence repeat polymorphisms) or microsatellite variation is the preferred system since it detects a large number of DNA polymorphisms with relatively simple technical complexity. The paucity of information on DNA sequences has limited their widespread utilization in the assessment of genetic diversity of minor or neglected crop species. However, recent advancements in DNA sequencing and PCR technologies in conjunction with sophisticated computer software have facilitated the development of SSRP markers in minor crops. This review examines the development and molecular nature of SSR markers, and their utilization in many aspects of plant genetics and ecology.

Keywords: genetic diversity; germplasm; DNA polymorphism; molecular markers; SSR

Introduction

Since the onset of life on earth more than three billion years ago, Mother Nature has generated a plethora of diverse life forms. Although genetic diversity is the most important legacy that one generation can pass on to the next, numerous species have disappeared over the years, and will continue to be lost. For example, it has been suggested that the current climate change may threaten the survival of as much as 12% of the wild relatives of *Solanum* (<http://www.potato2008.org/en/potato/biodiversity.html>). Although modern agricultural practices, urban expansion, deforestation, and other human activities all contribute to the pace with which species are becoming endangered and extinct, significant international efforts have been implemented to impede the erosion of genetic diversity. More than 600,000 plant samples are held by CGIAR (Consultative Group on International Research) to preserve the biodiversity of crop species (http://ftp.fao.org/planttreaty/news/news0003_2n.pdf), and hundreds of thousands of new samples are introduced into the germplasm institutes each year. However, it is necessary to assess the genetic variations in these introductions for redundancy. In an effort to maximize genetic variability and minimize repetitiveness, the concept of ‘core collection’ was introduced to maintain the optimum size of samples in a population [1-3].

Genetic diversity is influenced by selection, mutation, migration, population size, and genetic drift [4,5], and understanding how each of these factors influences the genetic diversity of a population is critical to the conservation of species. Although morphological markers enable the detection of genetic variation, it is often disguised by factors in the environment, and minimized by a paucity of discernible morphological markers. Pleiotropism and the late onset of some morphological markers during plant development also render unequivocal assessment difficult [6]. Significant advancements in molecular biology have shifted the focus of assessment of biodiversity from relying on morphological markers to using isozymes and DNA markers [7-9]. This review will focus on Simple Sequence Repeat Polymorphisms (SSRPs) as a molecular marker system for the assessment of genetic diversity in plants, with particular emphasis on the neglected minor crops or endangered species.

Molecular Genetic Markers in Plants

Variation between individuals in a population or between populations in a species, derived from genes and/or environmental effects, can be easily evaluated through the use of a variety of markers. Genetic markers were described before the discovery of proteins and DNA. In 1913, Alfred Sturtevant mapped six morphological markers that he described as “factors” on the chromosomes of the fruit fly [10]. Karl Sax demonstrated that seed size differences, seed-coat and pigmentation patterns were genetically linked in the common bean, *Phaseolus vulgaris* [11], where the number of easily discernible morphological markers was limited. Morphological traits that exhibit continuous variation between individuals in a population often obscure the evaluation of genetic diversity. Moreover, pleiotropism and a multifactorial basis to morphological traits further obfuscate the characterization of plant populations. The discovery that genes encoded proteins and enzymes led to the utilization of isozymes and other proteins as marker systems for genetic analysis of populations [7,12,13]. Although protein markers circumvent the effects of environment, they have the drawbacks of a limitation in the number of detectable isozymes as well as tissue and development stage specificity.

DNA marker systems, which were introduced to genetic analysis in the 1980s, have many advantages over the traditional morphological and protein markers that are used in genetic and ecological analyses of plant populations: firstly, an unlimited number of DNA markers can be generated; secondly, DNA marker profiles are not affected by the environment, and, thirdly DNA markers, unlike isozyme markers, are not constrained by tissue or developmental stage specificity.

The first generation of DNA marker systems employed Southern blot based markers. RFLPs (restriction fragment length polymorphisms) result from point mutations in restriction enzyme recognition sites. Chromosomal mutations such as insertions, deletions, inversions, and translocations can also cause restriction fragment size polymorphisms. The RFLP technique employs molecular hybridization of cDNA or genomic DNA probes with genomic DNA fragmented by restriction enzymes. Subsequent to the first demonstration of the usefulness of RFLPs in human genetics for linkage analysis [14], the technique was soon adopted by plant research communities [15-17]. Another Southern blot based marker system relied on minisatellite probes for ‘fingerprinting’ individual specific human DNA [18]. Minisatellite DNAs are short stretches of DNA that are present in tandem repeats in eukaryotes. They are highly abundant, and individuals often carry different numbers of tandem repeats which can be detected as VNTRs (variable number of tandem repeat) by PCR amplification [19]. While the RFLP technique utilizes mostly low copy number probes, the fingerprinting technique uses highly repetitive minisatellite DNAs as probes. Plant genomes also harbor many families of minisatellites [20-23]. Human minisatellite probes have been successfully used to distinguish individual Gramineae plants [24]. In modestly equipped laboratories, the technical complexity and high cost of these Southern blot based marker systems often limits their utilization in genetic analyses of large populations.

The second generation DNA markers for genetic analysis were those derived from PCR polymerase chain reaction [25]. PCR revolutionized genetic and ecological analyses of populations in several ways because it had two major advantages over Southern blot based markers. First, it requires only small DNA amounts to allow analysis at very early stages, thus reducing the need for plant nurseries. Second, it is inexpensive, and simple enough that large scale experiments can be carried out rapidly on a large scale. Of the many PCR-marker techniques that have been developed, RAPD, AFLP and SSR are the major systems, with the other systems being modifications of these three. The RAPD (randomly amplified polymorphic DNA) system may be the most efficient and simplest of techniques among the molecular genetic marker systems [26-28]. The RAPD technique utilizes a single arbitrary primer of 10-12 nucleotides (usually 10 nucleotides long) in the PCR reaction and thus, does not require template DNA sequence information. RAPD primers can bind complementary sequences in the genome and amplify the target sequences to produce 1-10 amplicons, depending on the templates and primers, and if the primer binding sites are within amplifying distance (100–1,000 bp). Although widely used in many analyses, the simple RAPD protocol has the significant drawback of a low reproducibility of results [29]. The dominant nature of RAPD markers also limits their application in F2 population and parentage analysis. However, these limitations can be overcome by converting RAPD markers to STS (sequence-tag-site) markers [30-32]. Inter-SSR is a modified version of the RAPD technique with SSR (simple sequence repeat) strategy, and is explained in reviews elsewhere [33,34]. Inter-SSR uses primers consisting of SSR sequences and amplifies template DNA as does

RAPD, but its target sequences are mostly in regions of the genome that harbor SSR sites. Low reproducibility of results with inter-SSR markers also poses as an impediment to its widespread use.

AFLP (amplified fragment length polymorphism), another PCR-based molecular marker system, obviates the need for template DNA sequence information [35]. Genomic DNA digested with two different restriction enzymes is ligated with specific adaptor sequences. The adaptor-ligated restriction fragments are then amplified with adaptor complementary primers that have selective nucleotides at their 3'-ends. The amplified fragments are separated by denaturing polyacrylamide gel electrophoresis and then stained with silver. Reproducibility of AFLP profiles is reasonably high, and each AFLP reaction produces 40-50 anonymous amplicons, ranging from 100–600 bp. The reproducibility of AFLPs was tested throughout a network of European laboratories in which AFLP banding patterns were reproduced in a range of laboratories by rigorous control of all the variables [29]. Partial digestion of the template DNAs by contamination or organ specific methylation produces spurious AFLP bands [36,37]. The anonymous multi-bands render AFLPs a preferred protocol for fingerprinting purposes [35,38]. AFLP markers exhibit dominant inheritance which can be converted to co-dominant STS markers to detect alleles of a given locus [29]. Another popular DNA marker system is the SSR, and since this review focuses on SSRPs, a detailed discussion of SSRs will follow at the end of this section.

The third generation of molecular markers is the system that utilizes SNPs (single nucleotide polymorphisms) [39-41] and microarrays [42]. Compared to the gel-based molecular marker systems, SNP detection and analysis can be carried out with non-gel based assays. The polymorphism of a single base difference can be assessed by through-put analysis, by hybridization with allele-specific oligonucleotides (ASO) [43], primer extension [44], oligonucleotide ligation assay (OLA) [45], and invasive cleavage [46]. The theory behind each of these techniques is reviewed in Sabrino *et al.* [47] and Semagn *et al.* [48]. There are numerous SNPs in plant genomes. In a comparison of genomic sequences from indica type rice and japonica type rice, Yu *et al.* [49] demonstrated one SNP in every 170 bp and one in/del every 540 bp. In a genome wide survey of 877 unigenes, SNPs were estimated to be present in every 200 bp in barley [50]. In maize, which is a cross-fertilizing species, Ching *et al.* [51] showed an even higher frequency of SNPs, an average of one polymorphism per 31 bp in non-coding regions and 1 polymorphism per 124 bp in coding regions. Therefore, we can extrapolate that the frequency of the SNPs can range from approximately one per 30 bp to one per 500 bp in other plant species. Although these new generation marker systems are powerful tools in linkage disequilibrium analysis, germplasm assay by haplotyping, QTL (quantitative trait loci) analysis and a few others [52,53], they are only amenable to use in those species for which extensive nucleotide sequence information is available in major crops. Complete genome sequences for an increasing number of plant species are being enlisted each year with rapid technical advancements in DNA sequencing strategies. As a result, molecular marker systems are being utilized more frequently in plant genomics and plant genetics. Thus, we restrict our discussion on the SSR applications to minor crop species in germplasm activities. Table 1 compares the different marker techniques that are employed in plant genome analysis for a variety of criteria.

Table 1. Comparison of widely using molecular marker systems for plant genome analyses.

	Isozyme	RFLP	RAPD	AFLP	SSR	SNP
Abundance	Low	Medium	Very high	Very high	High	Very high
Types of polymorphism	Amino acid change in polypeptide	Single base change, insertion, deletion, inversion	Single base change, insertion, deletion, inversion	Single base change, insertion, deletion, inversion	Repeat length variation	Single base change
DNA quality	-	High	Medium	High	Medium	Medium
DNA sequence information	-	Not required	Not required	Not required	Required	Required
Level of polymorphism	Low	Medium	High	High	High	High
Inheritance	Co-dominance	Co-dominance	Dominance	Dominance	Co-dominance	Co-dominance
Reproducibility	Medium	High	Low	Medium	High	High
Technical complexity	Medium	High	Low	Medium	Low	Medium
Developmental cost	Medium	High	Low	Low	High in start	High
Species Transferrability	High	Medium	High	High	Medium	Low
Automation	Low	Low	Medium	Medium	High	High

Note: Data on the RFLP, RAPD, AFLP and SSR were modified from Semagn *et al.* [48].

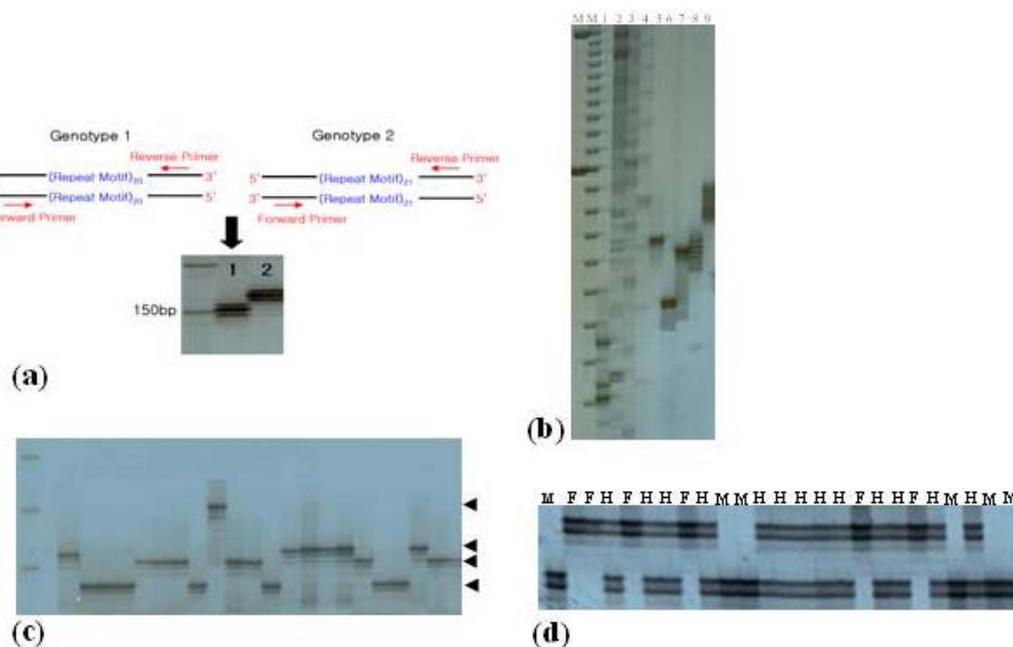
What Are SSRs?

In the 1960s, simple repeats that are scattered throughout eukaryotic genomes, were identified in density gradient centrifugations of randomly sheared genomic DNAs by way of a ‘satellite peak’ [54]. Isolation and sequencing of these satellite DNAs revealed repeat motifs of variable length from just a single base to thousands of bases; a typical satellite DNA is a centromeric sequence with a 100 bp repeat [55]. Subsequently, satellites of 10–30 bp repeat motifs, termed minisatellites, were isolated in mammals [18]. Finally, satellites with even shorter repeat motifs, called microsatellites, were isolated. In 1982, Hamada and his colleagues showed the existence of dinucleotide repeats of poly (C-A) and poly (G-T) in diverse eukaryotic genomes [56]. Since the repeat motifs in microsatellites are as short as 1–6 bases long, microsatellites are referred to as simple sequence repeats or simply, SSRs. Subsequent studies by Tautz and Renz [57] confirmed that SSRs, originally designated as STRs (short tandem repeats), were found to be informative, and abundantly present in human, fruit fly, sea urchin, protozoan, and yeast genomes. Furthermore, Weber and May [58] demonstrated that SSR polymorphisms (SSRPs) could be easily detected by PCR, using two flanking primers, which prompted the development of SSRs in various mammalian species and their subsequent assignment to specific chromosomes [59–61]. In plants, the presence of SSRs was first demonstrated by the hybridization of oligonucleotide probes of poly (G-T) and poly (A-G) on the phage libraries of tropical tree genomes [62]. A search of published DNA sequences reveals that SSRs are also highly abundant

in diverse plant genomes [64,65]. SSRs have therefore become the preferred molecular marker system for analysis in plant genetics and ecology.

The criteria for identifying SSRs differ, depending on the scientists who isolate and characterize them. While some scientists do not recognize single nucleotide repeats as SSRs, some consider the number of repeat units and others consider length of the repeat motifs [63,65-69] as criteria for defining SSRs. Akkaya *et al.* [63] isolated SSRs having more than 4 repeat units of di-, tri-, and tetranucleotide motifs in their analysis with soybean. Temnykh *et al.* [67] isolated 13,989 SSRs with di-, tri-, tetra-nucleotide motifs from 47 Mb of BAC-end sequences of rice, and categorized them into two classes; class 1 - hypervariable markers with motif length ≥ 20 bp, and class 2 - potentially variable markers with a repeat length ≥ 12 bp < 20 bp. However, the issue of repeat motifs is sometimes obscured by slightly imperfect sequences in repeat motif and by composite SSRs having more than one repeat motif. Therefore, describing SSR occurrence in can be specified in each genome sample. In our investigations with minor crop species, we define SSRs as DNA sequences in which a repeat motif length is over 20 bp with lower than 20% sequence imperfection in the repeat motif. Figure 1 describes the protocol for PCR amplification of SSRs with an example of a SSR profile in a population of a *Perrilla* species.

Figure 1. SSR amplification profiles.

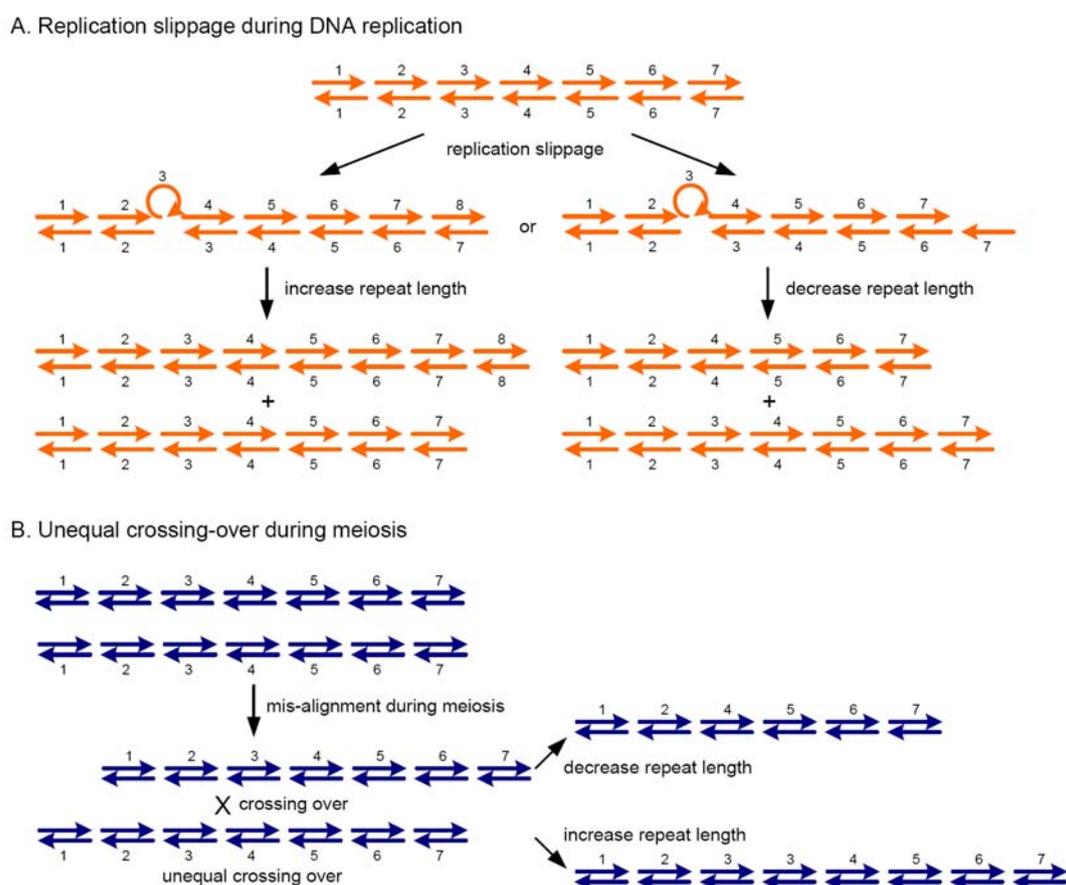


(a) Genotypes 1 and 2 differ by having (AT)₂₀ and (AT)₂₁ at a SSR locus and the PFSSR 25 locus, respectively. After PCR amplification, these alleles revealed polymorphisms upon denaturing polyacrylamide gel electrophoresis; (b) Profiles of different SSR loci in *Perrilla frutescens*. M denotes molecular size markers and subsequent numbers at different SSR loci. The SSRs in the first 4 lanes are duplicated in the genome and produce multiple bands whereas each of the other five SSR loci is a single SSR locus. Note the “stutter” bands in the SSR loci in lanes 8 and 9. Of the 9 SSR loci, the SSRs from lanes 5 to 7 are usable in population analysis while the rest are not good candidates for further utilization; (c) Polymorphic profiles of 21 different accessions of *P. frutescens*. Four alleles among 22 accessions of *P. frutescens* revealed a SSR locus (PFSSR 20). (d) Genetic segregation at a SSR locus in a F₂ segregating mapping population of *Zea mays*. M, F, and H denote maternal, paternal heterozygotes, respectively.

How Did SSRPs Arise?

SSRPs are mostly length variations by the different number of repeat units. It is not certain how these tandem repeats arose in genomes. Initially, it was thought that the occurrence of unequal crossing-over between repeat units during meiosis accounted for the length variation in minisatellites (SSRs) [70] while DNA replication slippage was responsible for the variation in microsatellites [71]. However, several studies have demonstrated that both types of repeats can be derived by either of the two mechanisms [72]. DNA replication slippage can also occur during *in vitro* amplification of the SSRs to produce ‘stutter bands’, which often obscure non-parental SSR bands in genetically segregating populations [73]. Figure 2 illustrates the mechanisms for producing SSRPs through unequal crossing-over during meiosis and slippage during DNA replication.

Figure 2. The mechanisms of unequal crossing-over and replication slippage for producing SSRPs.

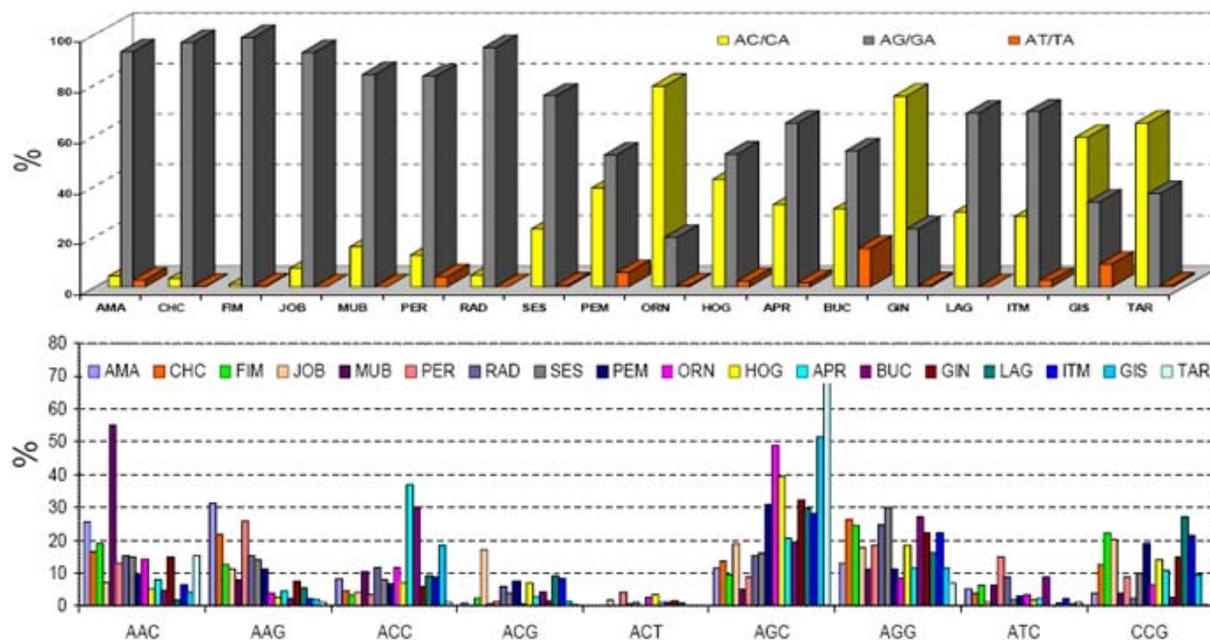


Both unequal crossing over and DNA slippage can either increase or decrease the number of repeat units in genomes. The number of repeat units ranges from 2 to 40, with extreme exceptions being as many as several hundred, as is the case with the Huntington's disease and fragile-X loci in humans. Is there a bias for specific repeat units in the different genomes? Yes, there appears to be a bias for different repeat motifs in different species, as shown in Figure 3. However, it is not clear why different species harbour different SSR motifs preferentially. Are SSRs with a higher number of repeats more polymorphic than those with a lower number of repeats? Generally, SSRs with longer repeats are more polymorphic than those of shorter ones [73].

How Do We Develop SSRs?

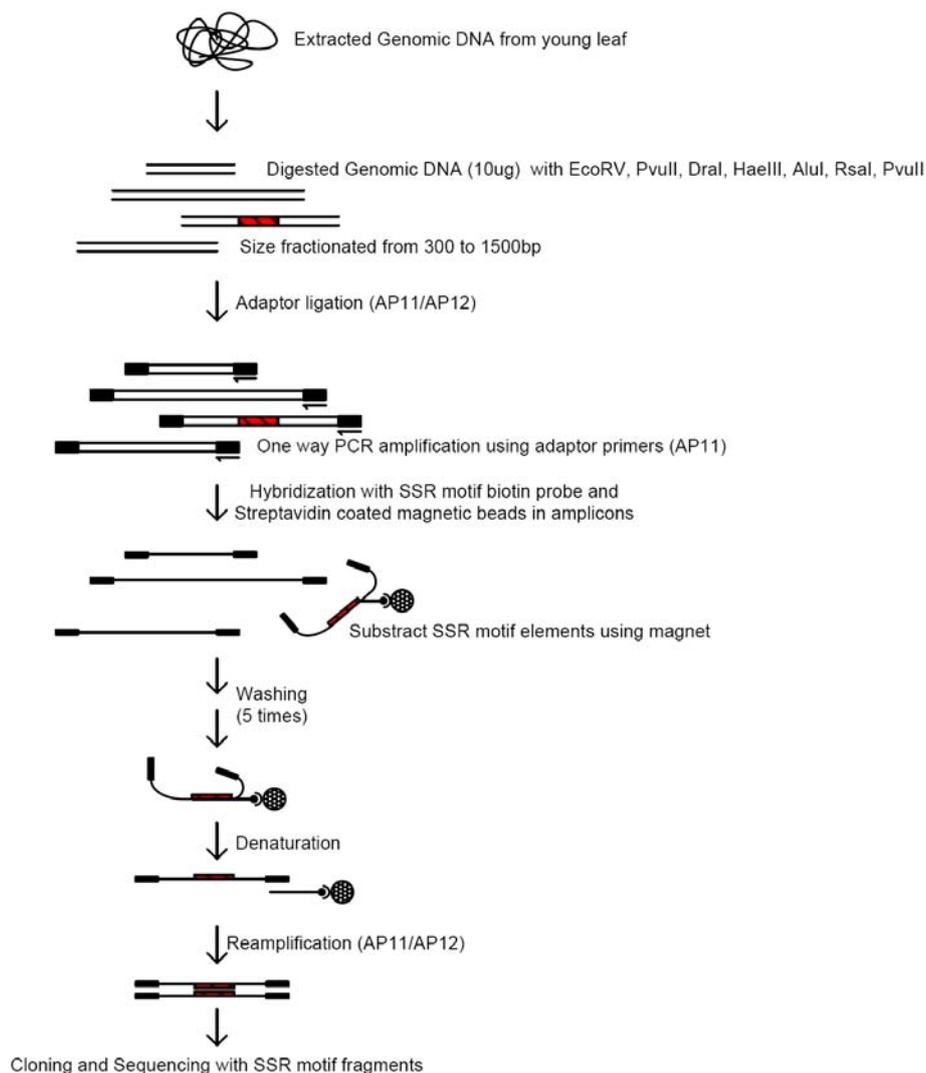
A variety of protocols have been employed to develop SSRs. There are several excellent reviews on the strategies for the isolation of microsatellites [74,75]. Only a fraction of the original clones from a successful isolation of SSR markers will be amenable to PCR amplification, since the clones containing SSR motifs with primer sequences are reduced at each step [75]. Therefore, optimum protocols are necessary for the efficient isolation of SSR markers in a given species.

Figure 3. Frequencies of different classes of di- and tri-nucleotide repeat motifs in 18 underutilized crop species.



AMA, *Amaranthus hypochondriacus*; CHC, *Brassica rapa*; FIM, *Eleusine coracana*; JOB, *Coix lacryma jobi*; MUB, *Vigna radiate*; PER, *Perilla frutescens* Britt.; RAD, *Raphanus sativus*; SES, *Sesamum indicum*; PEM, *Diospyros kaki* Thunb.; ORN, *Citrus raticulata* subsp. *Unshiu*; HOG, *Panicum miliaceum*; APR, *Prunus mume*; BUC, *Fagopyrum esculentum*; GIN, *Zingiber officinale* Rosc.; LAN, *Zoysia japonicum*; ITM, *Setaria italica*; GIS, *Panax ginseng* C.A. Meyer; TAR, *Colocasia esculenta*.

Traditionally, SSRs have been isolated from partial genomic libraries containing small size inserts by colony hybridization with probes that contain SSR sequence motifs [62,76,77]. However, this technique has been inefficient in most cases of species having large genomes, with the frequencies of colonies containing SSR motifs being relatively low. A SSR-enrichment step, prior to the library construction, was incorporated into the protocol to circumvent this problem. The SSR-enriched clones can then be employed by direct sequencing [78-82] or by incorporating one more step of colony hybridization with end-labelled SSR probes [83-85]. Figure 4 describes the general protocol for developing SSR markers with an SSR-enrichment step.

Figure 4. A general protocol for developing SSR markers with a SSR-enrichment step.

Another approach for the isolation of SSRs involves a computational search of the genome databases. Weber and May [58] searched (CA)_n SSRs in a human genome sequence database and demonstrated the abundance of the (CA)_n SSRs in the human genome. In plants, the first report of (CA)_n repeats in soybean sequences was the one from a computer search of the GenBank databases [63]. Around the same time, Morgante and Olivieri [64] showed an abundance of AT repeats in 30 different plant genomes and demonstrated that the repeat number of variations by PCR amplification were highly informative in genome analyses. In 1994, Wang *et al.* [65] surveyed mono-, di-, tri- and tetranucleotide repeats in plant sequences in EMBL and GenBank. They found that mono-, di-, and tetra-nucleotide repeats were all present in non-coding regions, but that 57% of the tri-nucleotide repeats, containing G-C base pairs, resided in the coding region. The high occurrence of tri-nucleotide repeats in coding region was attributed to the other types being eliminated from the coding region because of their ability to cause frame-shift mutations. With the wealth of sequence information of ESTs, fully characterized genes, and full-length cDNAs, computer searches are particularly useful in developing genic SSR markers [68,69,86]. There are a number of web-based SSR search softwares such as: MISA (MicroSatellite, <http://pgrc.ipk-gatersleben.de/misa>), CUGssr (<http://www.genome>

clemson.edu/projects/ssr), Sputnik (<http://abajian.nert/sputnik/index.html>), and SSRSEARCH (<ftp://ftp.gramene.org/pub/gramene/software/scripts/ssr.pl>). However, with species that have scant information on their genomic sequences, as is the case with most small or minor crop species, a computer search would not be useful in developing SSR markers. Library construction with SSR-enriched sequences would be a plausible way to develop SSR markers in these neglected minor crop species.

Advantages of SSR Analysis

SSR markers have many advantages over the other marker systems. The first advantage is their high reproducibility, which would be the most important in genetic analysis. While reproducibility of the SSR profile is as robust as it is with RFLPs, experimental procedures for SSR analysis are much simpler and require only a small amount of template DNA. Since SSR analysis does not require restriction with enzymes, it can reproduce the same profiles regardless of the state of the template DNA. It also does not require template DNA to be ultra pure, which is a requirement in AFLP analysis since contaminated or impure DNA is often recalcitrant in restriction enzyme digestions to produce nonspecific spurious bands. This is a real benefit when one is dealing with specimens that are dry, contaminated, mummified or even in fossilized form in the wild [87,88].

The second advantage of the SSR marker system is the polymorphic genetic information contents. The hyper-variable nature of SSRs produces very high allelic variations even among very closely related varieties. A literature survey showed that the number of alleles varied from 1 to 37 with diversity indices of 0.29–0.95 in major crop species [89]. The level of genetic variation detected by SSRPs analysis was almost two times higher than that detected by RFLPs, with 61 soybean lines [90]. In a comparative study of the utility of RFLP, RAPD, AFLP, and SSR marker systems for germplasm analysis, SSRs showed the highest expected heterozygosity, while AFLPs had the highest effective multiplex ratio [89].

The third advantage has to do with the co-dominant nature of SSR polymorphisms. Although homoplasious bands can be misleading in scoring SSR profiles, the SSR bands produced from the same set of primers are intuitively orthologous (a more detailed discussion of homoplasmy is provided in the ensuing section). The multiple bands generated from RAPD and AFLP analyses do not permit their designation as allelic or orthologous bands until they are converted into STS markers after sequencing. The co-dominant nature of SSRPs is suitable for genetical analysis in segregating F2 populations or parentage analysis in hybrids [91,92].

The fourth advantage of the SSR marker system is their abundance and distribution in genomes. As more and more genomic sequences are being identified in various eukaryotic species, it is becoming increasingly evident that SSRs are truly abundant in almost all species, and are well distributed throughout their genomes [65,68,93]. Genetic analysis is often frustrated by the fact that large numbers of anonymous RAPD or AFLP markers are clustered in specific locations of chromosomes or linkage maps [94,95]. In search of SSRs longer than 12 bp in a 57.8 Mb, publicly available rice (*Oryza sativa* L.) sequence, Temnykh *et al.* [67] showed that many kinds of SSRs are present every 16 kb. In another survey of SSRs in different eukaryotic genomes, Tóth *et al.* [93] reported that coding and non-coding regions differed significantly in SSR distribution, and characteristic differences also existed between inter-genic regions and introns in eukaryotes from yeasts to mammals to plants. Like the early findings

in plants by Wang *et al.* [65], tri- or hexa-nucleotide SSRs were predominantly present in coding regions, in the study by Tóth *et al.* [93].

A fifth advantage of the SSR marker system is that SSRs are preferentially associated with non-repetitive DNA [68,86,96]. Genomic sites of SSR markers, derived from genomic libraries, fall into either the transcribed region (genic SSRs) or the non-transcribed region (genomic SSRs). The SSRs, derived from ESTs or cDNAs, are mostly genic SSRs, which have the potential for application in such areas as gene function characterization [97], association analysis for gene tagging [98-100], and QTL analysis [101-103]. However, this last advantage can only be applied to those species with large amounts of EST or cDNA sequences that are freely accessible to public.

Problems with SSRPs

Although analysis with SSR markers has many merits compared to other marker systems, there are a few inherent problems associated with this system. Since SSR sequences can cause replication slippage *in vitro*, SSR polymorphisms are sometimes derived from slippage during polymerase chain reactions [73,104]. Slippage during PCR produces ‘stutter bands’ that differ in size from the main product by multiples of the length of repeat unit [104-106]. The stuttering produces many ladder bands in polyacrylamide gel separation, and leads to quasi-scoring, if there are no prominent bands among the ladders. Stuttering also generates ambiguity in SSRs with long stretches of a short repeat unit (1-2 bp) since Taq polymerase slippage increases with the number of repeat units, and is inversely correlated with the length of the repeat unit [107]. By rating the quality of the SSR-PCR products on a scale of 1 to 5 (good to poor), it is suggested that only those SSR primers with a rating of 1 and 2 be used in genomic analyses [106,108].

Homoplasy is an evolutionary term that is used to describe the observation that a character present in two species is not derived from a common ancestry but rather, the similarity is a result of convergence, parallelism or reversion. In the case of genetic markers, homoplasy is a phenomenon wherein different copies of a locus are identical in state despite not being identical by descent [109]. In SSR analysis, homoplasy can occur if two bands are similar in size but not identical in sequence. Table 2 illustrates four electromorphs, each 214 bp in size, from four different *Apis* species [110]. Two sets of identical electromorphs can be identified; one from *A. mellifera* and *A. lingustica* and the other from *A. scutellata* and *A. capensis*, respectively. However, the sequences were slightly different between sets by size homoplasy. The homoplasy can be confirmed by nucleotide sequencing or electrophoresis by single strand conformation polymorphism [111]. Homoplasy can lead to an underestimation of the actual divergence between populations [110,112,113]. However, SSR electromorph homoplasy does not present a significant problem in population genetics analyses since large amounts of variability at SSR loci are often compensated for by their homoplasious evolution, as shown in a computer simulation analysis with a large data set [109]. SSRP detects size variation by the differences in the number of repeat units, which arose by replication slippage or unequal crossing over. If a SSR with 5 repeat units is extended to a SSR with 6 repeat units, the latter can reduce its repeat unit number to 5 by a single ‘back-mutation’. If this happens, SSRP analysis cannot differentiate between the original SSR locus with 5 repeat units and the back mutated SSR locus with 5 repeat units, even though two steps of mutation were involved. This type of homoplasy cannot be checked, and may also lead to an underestimation of the genetic diversity of a population [114,115].

Table 2. SSR size homoplasmy at the locus A113 in different subspecies of *Apis mellifera*.

Sub-species	Electromorph (size in bp)	Core sequence
<i>Apis mellifera</i>	214	(TC) ₂ C(TC) ₂ TT(TC) ₄ TT(TC) ₂ TT(TC) ₉ GTTTCG(TC) ₂
<i>Apis ligustica</i>	214	(TC) ₂ C(TC) ₂ TT(TC) ₄ TT(TC) ₂ TT(TC) ₉ GTTTCG(TC) ₂
<i>Apis scutellata</i>	214	(TC) ₂ C(CT) ₂ TT(TC) ₅ TT(TC) ₁₁ GTTTCG(TC) ₂
<i>Apis capensis</i>	214	(TC) ₂ C(CT) ₂ TT(TC) ₅ TT(TC) ₁₁ GTTTCG(TC) ₂

Note: Data from Estoup *et al.* [110] with modification.

Often, in SSR analyses of a large number of samples from diverse germplasms, a few samples fail to produce PCR products [116-119]. This is a source of frustration to experimenters because they cannot determine whether the absence of PCR products represents true null alleles of the SSR locus or is due to a failure of the PCR reaction. Null SSR alleles represent the phenomena that do not give a PCR product by mutations [119-122]. Any individuals do not show amplification any alleles repeatedly while amplify normally other loci, null alleles may be present in this locus. If re-extracted DNA from the individual fails to show an amplification product, it is highly likely that the individual is homozygous for that allele. SSRs derived from ESTs or cDNA often fail to produce PCR products if one or both primer binding sites happen to be on the splice sites. Presence of large introns or primers designed from chimeric cDNA will not produce successful PCR products. Sometimes, primers designed from poor quality sequence information generate unsatisfactory results. Although the low rate of null alleles may be ignored in most population analyses, they may pose some complexity in parentage analysis as shown by Dakin and Avis [120]. It is recommended that rigorous primer selections be done for consistent amplification across all the samples before large scale analyses are conducted.

Cross-Species Applications

If SSRs are isolated for which primers can be designed, there is no doubt that the SSR marker system has many advantages over other marker systems. If the sequence information is insufficient to develop SSR markers, it may be advantageous to utilize primer sequences identified for one species in the analysis of other closely related species, given the high cost of developing useful SSR markers. SSRs from non-coding regions were not successful in cross-species amplifications due to the sequence variation surrounding SSR motifs, whereas SSRs from coding regions were successful in a wide range of species. In the most extensive cases, 17 SSRs were able to amplify across fish, which diverged about 470 mya (million years ago) [123]; six SSRs isolated from marine turtles amplified freshwater turtles, which diverged about 300 mya [124]. Peakall *et al.* [125] demonstrated that although 31% of soybean SSR loci were transferable to other legume species, the useful transferability was restricted to congeners. In grass species, Chen *et al.* [126] selected 11 SSR markers from *Oryza sativa* with the following criteria: (i) high allelic variation in 13 *O. sativa* cultivars, (ii) a variety of perfect and compound SSR motifs, and (iii) minimal stutter bands. The 11 SSR loci were all amplifiable among the *Oryza* species having the same A genome, whereas 73% (8/11) of primers amplified *Oryza* species having the other genomes B and C, and 27% (3/11) amplified species in other genera. Transferability of EST-SSR markers is high. Gupta *et al.* [119] demonstrated that 43 of 78 EST-SSR markers

exhibited transferability from *Triticum* to *Hordeum*, indicating that the sequences flanking SSR motifs were conserved not only within a single genus but also between related genera in the Poaceae family. However, SSR markers derived from a genomic SSR-enriched library showed poor cross-species amplification between species from a different genus. Only two of eleven SSR markers from an enrichment library of *Swietenia humilis* showed amplification across the Meliaceae family [127]. Regardless, it can be concluded that SSR primers from a SSR-enrichment library are still useful in the analysis of species within a genus. Recently, we isolated 12 SSR loci from *Amaranthus hypochoriacus* and were able to demonstrate cross-amplification of these SSR markers to 18 other wild species in the genus *Amaranthus* [82]. Similar results were obtained with the SSRs isolated from the common buckwheat (*Fagopyrum esculentum*) in cross-species amplifications with other species in the genus *Fagopyrum* [128].

One should be cautious in using transferable SSR markers for assessing species relationships since the maintenance of allele sizes among species is complex. The complexity of the mutation process in SSRs as well as size homoplasmy may complicate the interpretation of SSR variations. Size homoplasmy was frequently detected among cross-species amplified SSR markers between species. In the study of cross-species amplifications in *Oryza* species, Chen *et al.* [126] demonstrated that mutations occurred not only in the repeat units but also in the flanking regions to show allele size homoplasmy as well as cryptic alleles. Similar results have been reported in studies with *Pinus* species [129] as well as legume species [125]. Therefore, it is recommended that inferences vis-a-vis species relationships using SSRPs be accompanied by the information underlying sequences.

Biological Functions of SSRs

SSRs were generally deemed to be evolutionarily neutral. However, numerous lines of evidence have demonstrated that SSRs are not distributed randomly in the genome [86,93]. It is estimated that 14% of the genes in eukaryotic species contain repeated sequences, approximately three times more than in prokaryotes [130]. Incorporation of repeat sequences in eukaryotic genomes may confer an evolutionary advantage of adaptability to new environments [130,131]. Debates on the functional role(s) of the SSRs on species adaptation and survival have been well documented [132,133]. However, the findings of expansion and contraction of the SSR motifs within genes have encouraged the assignment of a biological role to SSRs. Thus far, the best known cases of SSRs with phenotypic effects are the human loci of Huntington's disease, and fragile-X [134]. SSRs on the UTR regions may also be involved in the regulation of expression of nearby genes as shown by a GT repeat in the *Tilapia prolactin 1* gene in fish, in response to a salt-challenged environment [135]. Intronic SSRs can regulate gene expression by influencing mRNA splicing or by translocation of mRNA to cytoplasm, as shown by the CCTG repeat in the first intron of the human zinc finger protein 9 (ZNF9), in which an expansion of the repeat causes one intron splicing to fail, thus leading to myotonic dystrophy [136]. Although biological roles for SSRs in plants have not been reported as yet, similar roles are expected for these molecular markers in plant genes.

If some SSRs are functional and confer an adaptive advantage, are these functional SSRs suitable in the assessment of biodiversity and ecological conservation of endangered species? Most of the molecular markers that have been utilized in population genetics have not undergone selection, and therefore have been essentially neutral. In neutral theory, the frequency of alleles is determined by

purely stochastic processes [137]. In conservation biology, neutral molecular markers may be useful in providing fundamental information about the types of mating in a population, gene flow, and the population history of a species [138]. However, there was a large discrepancy between genetic divergence as measured by neutral RAPD markers and that measured by quantitative genetic traits of the monkey puzzle tree (*Araucaria araucana*), a vulnerable tree endemic to southern South America [139]. Tienderen *et al.* [140] contend that gene-targeted functional markers can contribute to ex-situ management of genetic resources, studies on ecological diversity, and conservation of endangered species. Holderegger *et al.* [141] proffer a theory on the adaptive versus neutral diversity for landscape genetics in which the diversity measured by neutral markers is well suited for the study of processes of gene flow within landscapes, whereas diversity assessed by quantitative genetic experiments using functional markers is best suited for measuring the evolutionary or adaptive potential of a population or species. They concluded that ecologists must recognize these differences between neutral and adaptive genetic variation when interpreting the results of landscape genetic studies. However, it should be remembered that variation in functional genes might reflect the past influence of selection, which can be variable in each gene and can affect the profiles of variation from the history, migration, and drift [141]. While genomic SSR markers are mostly neutral, genic SSRs from EST's or cDNAs may retain some adaptive roles. This duality in selection and adaptation ascribes another advantage to the utilization of SSRs in characterizing the genetic diversity of the resources that are preserved in different germplasm institutes.

Germplasm Collections and Development of SSRs in Minor Crops

Minor crops are also referred to as 'orphan crops' or 'underutilized crops' because of their lack of global cultivation and utilization [142]. Compared to major staples or other economic crops, these minor crops have often been neglected and are therefore on the verge of extinction in some cases. Human activities in the wild have also accelerated the genetic erosion and extinction of these vulnerable plants which are listed in the website of The International Centre of Underutilized Crops (<http://www.icuc-iwmi.org/>). Albeit late, accessions of these neglected minor crops in gene banks and germplasm institutes have been collected worldwide. However, introduction of the accessions of these minor crops without their characterization, limits the maximum preservation of their genetic diversity [143]. The concept of 'core collection' was introduced to counter the problems faced in the management of germplasm (1-3). So, the determination of entries should be done to meet the requirement of least number of accessions with highest diversity of selected subset. At the onset, a core collection was simply defined as a subset of accessions with a good approximation of the genetic diversity of a crop species and its wild relatives from the existing collections. Until the 'core collection' concept was established in early 1980s, germplasm was collected from as many accessions or species as was possible; core collection represents genetic diversity with a minimum of repetitiveness in a crop species and its relatives [1-3]. The assessment of genetic diversity of introductions (accessions) is now a pivotal strategy for their successful and efficient preservation, in situ as well as ex situ [140]. SSRPs are the most suitable markers for the genetic assessment of germplasms because of their hypervariability, attributable to allelic variations [128,144,145]. In a study of 192 accessions of *Medicago trunculata* with five SSR markers, Ellwood *et al.* [144] showed that the core collection was highly diverse; over 90% of the genotypes, with an average of 25 SSR

alleles per locus, defined a subset of accessions ($n = 61$) that will maximize the diversity. Similarly, Zhao *et al.* [145] studied the genetic diversity and population structure of a selected core of garlic accessions with 8 SSR primers; 95 accessions from 613 garlic entries obtained from 36 different countries captured all the alleles in the entire collection of accessions by the model-based heuristic approach [146]. This heuristic method implements the determination of core entries with least number of accessions and 100% coverage of diversity of entire collection in a selected core sub-set, which is completely different from the conventional methods including the stratification and a previous M-strategy. For a similar assessment of Korean Germplasm collections, SSR motifs of various minor crop species have been isolated to aid in efficient germplasm management [147]. Characterization of these collections is currently underway with the SSRs that have been specifically developed for the purpose.

Concluding Remarks

Preservation of genetic diversity is important at both the molecular and species levels, if one generation is to pass on its legacy of intact diverse forms of life to the next generation. Although collections of accessions of neglected or unintended plants or crops are being carried out world-wide in germplasm institutes, a recurrent problem, apart from limited space and other resources, is the redundancy of the introductions, thus necessitating their pre-screening before being placed in the depository. Redundancy screening for molecular diversity through the use of molecular markers is very efficient, and is being widely used. Of the many molecular systems that are available, SSRP is the method of choice because it can also be applied to core collections of species. However, it cannot be directly applied to minor or neglected orphan crops because of the requirement of DNA sequence information for designing primers for PCR. Since the set-up costs for SSR development in new species are exorbitant, it is recommended that DNA sequence information be borrowed from related species for cross-species SSR amplification. In so doing, researchers should be aware that some homoplasious SSR bands can underestimate the molecular diversity. Regardless, SSRPs have more advantages than other molecular marker techniques. With more and more DNA information sequences becoming available through ESTs or whole genome sequencing, the number of available SSR markers is also increasing. Efficient SSR-enrichment library construction protocols are available with various modifications so that researchers can select the protocols to meet their needs. The characterization of the various germplasm introductions through SSRPs, which is well underway, will be of significant benefit in managing the precious genetic diversity of underutilized species.

Acknowledgements

This study was supported by a Bio-Green 21 Project Grant (20080401-034-060-009-04) to YJP and NSK and a Regional Strategic Crop Development from RDA to JKL and NSK. We are indebted to Drs. George Fedak and Wenguang Cao for their suggestions and critical reading of the manuscript. Appreciation is also extended to Dr. Ramesh Bhambhani for reading and critically editing the manuscript.

References

1. Frankel, O.H. Genetic Perspectives of Germplasm Conservation. In *Genetic Manipulation: Impact On Man And Society*. Arber, W.K., Llimensee, K., Peacock, W.J., Starlinger, P., Eds.; Cambridge University Press: Cambridge, UK, 1984; pp. 161–170.
2. Yonezawa, K. A definition of the optimal allocation of effort in conservation of plant genetic resources-with application to sample size determination for field collection. *Euphytica* **1985**, *34*, 345–354.
3. Brown, A.H.D. Core collections: A practical approach to genetic resources management. *Genome* **1989**, *31*, 818–824.
4. Hedrick, P.W. *Genetics of Population*, 3rd Ed.; Jones and Bartlett Pub. Co: Sudbury, MA, USA, 2005.
5. Ouborg, N.J.; Vergeer, P.; Mix, C. The rough edges of the conservation genetics paradigm for plants. *Ecology* **2006**, *94*, 1233–1248.
6. Thomas, R.G. Pleiotropism and the developmental changes correlated with flowering. *Genetica* **1960**, *31*, 329–339.
7. Brown, A.H.D. Isozymes, plant population genetic structure and genetic conservation. *Theor. Appl. Genet.* **1978**, *52*, 145–157.
8. Bretting, P.K.; Widrechner, M.P. Genetic markers and plant genetic resource management. *Plant Breed. Rev.* **1995**, *13*, 11–86.
9. Karp, A.; Edwards, K.J. DNA markers: A Global Overview. In *DNA Markers: Protocols, Applications, And Overviews*; Caetano-Anolles, G., Greshoff, P.M., Eds.; Wiley-VCH Inc.: New York, NY, USA, 1997; pp. 1–13.
10. Sturtevant, A.H. The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of association. *J. Exp. Zool. Part B* **1913**, *14*, 43–59.
11. Sax, K. The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. *Genetics* **1923**, *8*, 552–560.
12. Scandalios, J.G. Genetic control of multiple molecular forms of enzymes in plants: A review. *Biochem. Genet.* **1969**, *3*, 37–79.
13. Hamrick, J.L.; Linhart, Y.B.; Mitton, J.B. Relationships between life history characteristics and electrophoretically detectable genetic variation in plants. *Annu. Rev. Ecol. Syst.* **1979**, *10*, 173–200.
14. Botstein, D.; White, R.L.; Skolnick, M.; Davis, R.W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* **1980**, *32*, 314–331.
15. Helentjaris, T.; King, G.; Slocum, M.; Siedenstrang, C.; Wegman, S. Restriction fragment polymorphisms as probes for plant diversity and their development as tools for applied plant breeding. *Plant Mol. Biol.* **1985**, *5*, 109–118.
16. Chang, C.; Bowman, J.L.; DeJohn, A.W.; Lander, E.S.; Meyerowitz, E.M. Restriction fragment length polymorphism linkage map for *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 6856–6860.
17. Tanksley, S.D.; Young, N.D.; Paterson, A.H.; Bonierbale, M.W. RFLP mapping in plant breeding: New tools for an old science. *Nat. Biotechnol.* **1989**, *7*, 257–264.

18. Jeffreys, A.J.; Wilson, V.; Thein, S.L. Individual-specific 'fingerprints' of human DNA. *Nature* **1985**, *316*, 76–79.
19. Jeffreys, A.J.; Wilson, V.; Neumann, R.; Keyte, J. Amplification of human minisatellites by the polymerase chain reaction: Towards DNA fingerprinting of single cells. *Nucleic Acids Res.* **1988**, *16*, 10953–10971.
20. Weising, K.; Nybom, H.; Wolff, K. DNA Fingerprinting in Plants: Principles, Methods, and Applications. CRC Press: London, UK, 2005.
21. Nybom, H.; Hall, H.K. Minisatellite DNA 'fingerprints' can distinguish *Rubus* cultivars and estimate their degree of relatedness. *Euphytica* **1991**, *53*, 107–114.
22. Ouborg, N.J.; Piquot, Y.; Van Groenendael, J.M. Population genetics, molecular markers and the study of dispersal in plants. *J. Ecol.* **1999**, *551*–568.
23. Kim, N.S.; Park, N.I.; Kim, S.H.; Kim, S.T.; Han, S.S.; Kang, K.Y. Isolation of TC/AG repeat microsatellite sequences for fingerprinting rice blast fungus and their possible horizontal transfer to plant species. *Mol. Cells* **2000**, *10*, 127–134.
24. Dallas, J.F. Detection of DNA "Fingerprints" of cultivated rice by hybridization with a human minisatellite DNA Probe. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 6831–6835.
25. Mullis, K.; Faloona, F.; Scharf, S.; Saiki, R.; Horn, G.; Erlich, H. Specific enzymatic amplification Of Dna *in vitro*: The polymerase chain reaction. *Cold Spring Harb. Symp. Quant. Biol.* **1986**, *51*, 263–273.
26. Williams, J.G.; Kubelik, A.R.; Livak, K.J.; Rafalski, J.A.; Tingey, S.V. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.* **1990**, *18*, 6531–6535.
27. Welsh, J.; McClelland, M. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res.* **1990**, *18*, 7213–7218.
28. Lynch, M.; Milligan, B.G. Analysis of population genetic structure with RAPD markers. *Mol. Ecol.* **1994**, *3*, 91–99.
29. Jones, C.J.; Edwards, K.J.; Castaglione, S.; Winfield, M.O.; Sala, F.; Van de Wiel, C.; Bredemeijer, G.; Vosman, B.; Matthes, M.; Daly, A. Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Mol. Breed.* **1997**, *3*, 381–390.
30. Monna, L.; Miyao, A.; Inoue, T.; Fukuoka, S.; Yamazaki, M.; Zhong, H. S.; Sasaki, T.; Minobe, Y. Determination of RAPD markers in rice and their conversion into sequence tagged sites (STSs) and STS-specific primers. *DNA Res.* **1994**, *1*, 139–148.
31. Brady, J.L.; Scott, N.S.; Thomas, M.R. DNA typing of hops (*Humulus lupulus*) through application of RAPD and microsatellite marker sequences converted to sequence tagged sites (STS). *Euphytica* **1996**, *91*, 277–284.
32. Komatsuda, T.; Nakamura, I.; Takaiwa, F.; Oka, S. Development of STS markers closely linked to the *vrs1* locus in barley, *Hordeum vulgare*. *Genome* **1998**, *41*, 680–685.
33. Godwin, I.D.; Aitken, E.A.B.; Smith, L.W., Application of inter simple sequence repeat (ISSR) markers to plant genetics. *Electrophoresis* **1997**, *18*, 1524–1528.
34. Nagaraju, J.; Kathirvel, M.; Kumar, R.R.; Siddiq, E.A.; Hasnain, S.E. Genetic analysis of traditional and evolved Basmati and non-Basmati rice varieties by using fluorescence-based ISSR-PCR and SSR markers. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5836–5841.

35. Vos, P.; Hogers, R.; Bleeker, M.; Reijans, M.; Van De Lee, T.; Hornes, M.; Frijters, A.; Pot, J.; Peleman, J.; Kuiper, M. AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Res.* **1995**, *23*, 4407–4414.
36. Donini, P.; Koebner, R.M.D.; Elias, M.L.; Bougourd, S.M. AFLP fingerprinting reveals pattern differences between template DNA extracted from different plant organs. *Genome* **1997**, *40*, 521–526.
37. Dowling, T.E.; Moritz, C.; Palmer, J.D.; Rieseberg, L.H. Nucleic acids III: Analysis of fragments and restriction sites. In *Molecular Systematics*. Hillis D.M., Moritz C., Richardson L., Eds.; Sinauer, Sunderland, MA, USA, 1996; pp. 249–321.
38. Mueller, U.G.; Wolfenbarger, L.L.R. AFLP genotyping and fingerprinting. *Trends Ecol. Evol.* **1999**, *14*, 389–394.
39. Kruglyak, L.; Nickerson, D.A. Variation is the spice of life. *Nat. Genet.* **2001**, *27*, 234–236.
40. Nasu, S.; Suzuki, J.; Ohta, R.; Hasegawa, K.; Yui, R.; Kitazawa, N.; Monna, L.; Minobe, Y. Search for and analysis of single nucleotide polymorphisms (SNPs) in rice (*Oryza sativa*, *Oryza rufipogon*) and establishment of SNP markers. *DNA Res.* **2002**, *9*, 163–171.
41. Choi, I.Y.; Hyten, D.L.; Matukumalli, L.K.; Song, Q.; Chaky, J.M.; Quigley, C.V.; Chase, K.; Lark, K.G.; Reiter, R.S.; Yoon, M.S. A soybean transcript map: Gene distribution, haplotype and single-nucleotide polymorphism analysis. *Genetics* **2007**, *176*, 685–696.
42. Lindroos, K.; Sigurdsson, S.; Johansson, K.; Ronnblom, L.; Syvanen, A.C. Multiplex SNP genotyping in pooled DNA samples by a four-colour microarray system. *Nucleic Acids Res.* **2002**, *30*, e70.
43. Hashimoto, K.; Hashimoto, M.; Mishiro, S.; Oota, Y. Method of detecting nucleic acid relating to disease. *EU Patent App. 1375672*, 2004.
44. Sybänen, A.C. From gels to chips: “Minisequencing” primer extension for analysis of point mutations and single nucleotide polymorphisms. *Hum. Mutat.* **1999**, *13*, 1–10.
45. Iannone, M.A.; Taylor, J.D.; Chen, J.; Li, M.S.; Rivers, P.; Slentz-Kesler, K.A.; Weiner, M.P. Multiplexed single nucleotide polymorphism genotyping by oligonucleotide ligation and flow cytometry. *Cytometry* **2000**, *39*, 131–140.
46. Lyamichev, V.; Mast, A.L.; Hall, J.G.; Prudent, J.R.; Kaiser, M.W.; Takova, T.; Kwiatkowski, R.W.; Sander, T.J.; de Arruda, M.; Arco, D.A. Polymorphism identification and quantitative detection of genomic DNA by invasive cleavage of oligonucleotide probes. *Nat. Biotechnol.* **1999**, *17*, 292–296.
47. Sobrino, B.; Bri n, M.; Carracedo, A. SNPs in forensic genetics: A review on SNP typing methodologies. *Forensic Sci.Int.* **2005**, *154*, 181–194.
48. Semagn, K.; Bjørnstad; Ndjiondjop, M.N. An overview of molecular marker methods for plants. *Afr. J. Biotechnol.* **2006**, *5*, 2540–2568.
49. Yu, J.; Hu, S.; Wang, J.; Wong, G.K.S.; Li, S.; Liu, B.; Deng, Y.; Dai, L.; Zhou, Y.; Zhang, X. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* **2002**, *296*, 79–92.
50. Rostoks, N.; Mudie, S.; Cardle, L.; Russell, J.; Ramsay, L.; Booth, A.; Svensson, J.T.; Wanamaker, S.I.; Walia, H.; Rodriguez, E. M. Genome-wide SNP discovery and linkage analysis in barley based on genes responsive to abiotic stress. *Mol. Genet. Genomics* **2005**, *274*, 515–527.

51. Ching, A.; Caldwell, K.S.; Jung, M.; Dolan, M.; Smith, O.S.H.; Tingey, S.; Morgante, M.; Rafalski, A.J. SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. *BMC Genet.* **2002**, *3*, 19.
52. Gupta, P.K.; Roy, J.K.; Prasad, M. Single nucleotide polymorphisms: A new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. *Curr. Sci.* **2001**, *80*, 524–535.
53. Rafalski, A. Applications of single nucleotide polymorphisms in crop genetics. *Curr. Opin. Plant Biol.* **2002**, *5*, 94–100.
54. Moustacchi, E.; Williamson, D.H. Physiological variations in satellite components of yeast DNA detected by density gradient centrifugation. *Biochem. Biophys. Res. Commun.* **1966**, *23*, 56–61.
55. Pardue, M.L.; Gall, J.G. Chromosomal localization of mouse satellite DNA. *Science* **1970**, *168*, 1356–1358.
56. Hamada, H.; Petrino, M.G.; Kakunaga, T. A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 6465–6469.
57. Tautz, D.; Renz, M. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Res.* **1984**, *12*, 4127–4138.
58. Weber, J.L.; May, P.E. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am. J. Hum. Genet.* **1989**, *44*, 388–396.
59. Weissenbach, J.; Gyapay, G.; Dib, C.; Vignal, A.; Morissette, J.; Millasseau, P.; Vaysseix, G.; Lathrop, M. A second-generation linkage map of the human genome. *Nature* **1992**, *359*, 794–801.
60. McAleer, M.A.; Aitman, T.J.; Cornall, R.J.; Ghosh, S.; Hall, J.R.S.; Hearne, C.M.; Love, J.M.; Prins, J.B.; Ramachandran, S.; Rodrigues, N. Linkage analysis of 84 microsatellite markers in intra- and interspecific backcrosses. *Mamm. Genome* **1992**, *3*, 457–460.
61. Serikawa, T.; Kuramoto, T.; Hilbert, P.; Mori, M.; Yamada, J.; Dubay, C.J.; Lindpainter, K.; Ganten, D.; Guenet, J.L.; Lathrop, G.M. Rat gene mapping using PCR-analyzed microsatellites. *Genetics* **1992**, *131*, 701–721.
62. Condit, R.; Hubbell, S.P. Abundance and DNA sequence of two-base repeat regions in tropical tree genomes. *Genome* **1991**, *34*, 66–71.
63. Akkaya, M.S.; Bhagwat, A.A.; Cregan, P.B. Length polymorphisms of simple sequence repeat DNA in soybean. *Genetics* **1992**, *132*, 1131–1139.
64. Morgante, M.; Olivieri, A.M. PCR-amplified microsatellites as markers in plant genetics. *Plant J.* **1993**, *3*, 175–182.
65. Wang, Z.; Weber, J.L.; Zhong, G.; Tanksley, S.D. Survey of plant short tandem DNA repeats. *Theor. Appl. Genet.* **1994**, *88*, 1–6.
66. Kresovich, S.; Szewc-McFadden, A. K.; Bliet, S.M.; McFerson, J.R. Abundance and characterization of simple-sequence repeats (SSRs) isolated from a size-fractionated genomic library of *Brassica napus* L. (rapeseed). *Theor. Appl. Genet.* **1995**, *91*, 206–211.
67. Temnykh, S.; DeClerck, G.; Lukashova, A.; Lipovich, L.; Cartinhour, S.; McCouch, S. Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): Frequency, length variation, transposon associations, and genetic marker potential. *Genome Res.* **2001**, *11*, 1441–1452.

68. Varshney, R.K.; Graner, A.; Sorrells, M.E. Genic microsatellite markers in plants: Features and applications. *Trends Biotechnol.* **2005**, *23*, 48–55.
69. Simko, I. Development of EST-SSR Markers for the study of population structure in Lettuce (*Lactuca sativa* L.). *J. Hered.* **2009**, *100*, 256–262.
70. Jeffreys, A.J.; Murray, J.; Neumann, R. High-resolution mapping of crossovers in human sperm defines a minisatellite-associated recombination hotspot. *Mol. Cell* **1998**, *2*, 267–273.
71. Strand, M.; Prolla, T.A.; Liskay, R.M.; Petes, T.D. Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. *Nature* **1993**, *365*, 274–276.
72. Richard, G.F.; Kerrest, A.; Dujon, B. Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. *Microbiol. Mol. Biol. Rev.* **2008**, *72*, 686–727.
73. Ellegren, H. Microsatellites: Simple sequences with complex evolution. *Nat. Rev. Genet.* **2004**, *5*, 435–445.
74. Zane, L.; Bargelloni, L.; Patarnello, T. Strategies for microsatellite isolation: A review. *Mol. Ecol.* **2002**, *11*, 1–16.
75. Squirrell, J.; Hollingsworth, P.M.; Woodhead, M.; Russell, J.; Lowe, A.J.; Gibby, M.; Powell, W. How much effort is required to isolate nuclear microsatellites from plants? *Mol. Ecol.* **2003**, *12*, 1339–1348.
76. Kelly, A.J.; Willis, J.H. Polymorphic microsatellite loci in *Mimulus guttatus* and related species. *Mol. Ecol.* **1998**, *7*, 769–774.
77. Dick, C.W.; Hamilton, M.B. Microsatellites from the Amazonian tree *Dinizia excelsa* (Fabaceae). *Mol. Ecol.* **1999**, *8*, 1765–1766.
78. Green, J.M.; Edwards, K.J.; Usher, S.L.; Barker, J.H.A.; Marshall, E.J.P.; Froud-Williams, R.J.; Karp, A. Microsatellites for Barren Brome (*Anisantha sterilis*). *Mol. Ecol.* **2000**, *9*, 2195–2197.
79. Hughes, M.; Hollingsworth, P.M.; Squirrell, J. Isolation of polymorphic microsatellite markers for *Begonia sutherlandii* Hook. f. *Mol. Ecol. Notes* **2002**, *2*, 185–186.
80. De La Rosa, R.; James, C.M.; Tobutt, K.R. Isolation and characterization of polymorphic microsatellites in olive (*Olea europaea* L.) and their transferability to other genera in the *Oleaceae*. *Mol. Ecol. Notes* **2002**, *2*, 265–267.
81. Kwon, S.J.; Lee, J.K.; Kim, N.S.; Yu, J.W.; Dixit, A.; Cho, E.G.; Park, Y.J. Isolation and characterization of microsatellite markers in *Perilla frutescens* Brit. *Mol. Ecol. Notes* **2005**, *5*, 455–457.
82. Lee, J.R.; Hong, G.Y.; Dixit, A.; Chung, J.W.; Ma, K.H.; Lee, J.H.; Kang, H.K.; Cho, Y.H.; Gwag, J.G.; Park, Y.J. Characterization of microsatellite loci developed for *Amaranthus hypochondriacus* and their cross-amplifications in wild species. *Conserv. Genet.* **2008**, *9*, 243–246.
83. Iwata, H.; Konuma, A.; Tsumura, Y. Development of microsatellite markers in the tropical tree *Neobalanocarpus heimii* (Dipterocarpaceae). *Mol. Ecol.* **2000**, *9*, 1684–1685.
84. Lian, C.; Miwa, M.; Hogetsu, T. Isolation and characterization of microsatellite loci from the Japanese red pine, *Pinus densiflora*. *Mol. Ecol.* **2000**, *9*, 1186–1187.
85. Gaiotto, F.A.; Brondani, R.P.V.; Grattapaglia, D. Microsatellite markers for heart of palm-*Euterpe edulis* and *E. oleracea* Mart.(Arecaceae). *Mol. Ecol. Notes* **2001**, *1*, 86–88.

86. Morgante, M.; Hanafey, M.; Powell, W. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nature Genet.* **2002**, *30*, 194–200.
87. Manen, J.F.; Bouby, L.; Dalnoki, O.; Marival, P.; Turgay, M.; Schlumbaum, A. Microsatellites from archaeological *Vitis vinifera* seeds allow a tentative assignment of the geographical origin of ancient cultivars. *J. Archaeol. Sci.* **2003**, *30*, 721–729.
88. Boder, P.; Deak, T.; Bacso, R.; Velich, I.; Bisztray, G.D.; Fascar, G.; Gyulai, P. Morphological and genetic investigation of medieval grape seeds. *Acta Horticulture (ISHS)* **2006**, 713–718.
89. Powell, W.; Morgante, M.; Andre, C.; Hanafey, M.; Vogel, J.; Tingey, S.; Rafalski, A. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Mol. Breed.* **1996**, *2*, 225–238.
90. Morgante, M.; Rafalski, A.; Biddle, P.; Tingey, S.; Olivieri, A.M. Genetic mapping and variability of seven soybean simple sequence repeat loci. *Genome* **1994**, *37*, 763–769.
91. Scott, K.D.; Egger, P.; Seaton, G.; Rossetto, M.; Ablett, E. M.; Lee, L.S.; Henry, R.J. Analysis of SSRs derived from grape ESTs. *Theor. Appl. Genet.* **2000**, *100*, 723–726.
92. Slavov, G.T.; Howe, G.T.; Gyaourova, A.V.; Birkes, D.S.; Adams, W.T. Estimating pollen flow using SSR markers and paternity exclusion: Accounting for mistyping. *Mol. Ecol.* **2005**, *14*, 3109–3121.
93. Tóth, G.; Gáspári, Z.; Jurka, J. Microsatellites in different eukaryotic genomes: Survey and analysis. *Genome Res.* **2000**, *10*, 967–981.
94. Vuylsteke, M.; Mank, R.; Antonise, R.; Bastiaans, E.; Senior, M.L.; Stuber, C.W.; Melchinger, A.E.; Lbberstedt, T.; Xia, X. C.; Stam, P. Two high-density AFLP® linkage maps of *Zea mays* L.: Analysis of distribution of AFLP markers. *Theor. Appl. Genet.* **1999**, *99*, 921–935.
95. Kwon, S.; Hong, S.; Son, J.; Lee, J.K.; Cha, Y.; Eun, M.; Kim, N. CACTA and MITE transposon distributions on a genetic map of rice using F15 RILs derived from Milyang 23 and Gihobyoe hybrids. *Mol. Cells* **2006**, *21*, 360–366.
96. Andersen, J.R.; Liberstedt, T. Functional markers in plants. *Trends Plant Sci.* **2003**, *8*, 554–560.
97. Ronning, C.M.; Stegalkina, S.S.; Ascenzi, R.A.; Bougri, O.; Hart, A.L.; Utterbach, T.R.; Vanaken, S.E.; Riedmuller, S.B.; White, J.A.; Cho, J. Comparative analyses of potato expressed sequence tag libraries. *Plant Physiol.* **2003**, *131*, 419–429.
98. Szalma, S.J.; Buckler, E.S.; Snook, M.E.; McMullen, M.D. Association analysis of candidate genes for maysin and chlorogenic acid accumulation in maize silks. *Theor. Appl. Genet.* **2005**, *110*, 1324–1333.
99. Shin, J.H.; Kwon, S.J.; Lee, J.K.; Min, H.K.; Kim, N.S. Genetic diversity of maize kernel starch-synthesis genes with SNAPs. *Genome* **2006**, *49*, 1287–1296.
100. Crossa, J.; Burgueno, J.; Dreisigacker, S.; Vargas, M.; Herrera-Foessel, S.A.; Lillemo, M.; Singh, R.P.; Trethowan, R.; Warburton, M.; Franco, J. Association analysis of historical bread wheat germplasm using additive genetic covariance of relatives and population structure. *Genetics* **2007**, *177*, 1889–1913.
101. Buerstmayr, H.; Lemmens, M.; Hartl, L.; Doldi, L.; Steiner, B.; Stierschneider, M.; Ruckenbauer, P. Molecular mapping of QTLs for Fusarium head blight resistance in spring wheat. I. Resistance to fungal spread (Type II resistance). *Theor. Appl. Genet.* **2002**, *104*, 84–91.

102. Breseghello, F.; Sorrells, M.E. Association analysis as a strategy for improvement of quantitative traits in plants. *Crop Sci.* **2006**, *46*, 1323–1330.
103. Zeng, Y.; Yang, S.; Cui, H.; Yang, X.; Xu, L.; Du, J.; Pu, X.; Li, Z.; Cheng, Z.; Huang, X. QTLs of cold-related traits at the booting stage for NIL-RILs in rice revealed by SSR. *Genes Genom.* **2009**, *31*, 143–145.
104. Hauge, X. Y.; Litt, M. A study of the origin of 'shadow bands' seen when typing dinucleotide repeat polymorphisms by the PCR. *Hum. Mol. Genet.* **1993**, *2*, 411–415.
105. Murray, V.; Monchawin, C.; England, P.R. The determination of the sequences present in the shadow bands of a dinucleotide repeat PCR. *Nucleic Acids Res.* **1993**, *21*, 2395–2398.
106. Smulders, M.J.M.; Bredemeijer, G.; Rus-Kortekaas, W.; Arens, P.; Vosman, B. Use of short microsatellites from database sequences to generate polymorphisms among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species. *Theor. Appl. Genet.* **1997**, *94*, 264–272.
107. Shinde, D.; Lai, Y.; Sun, F.; Arnheim, N. Taq DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: (CA/GT) n and (A/T) n microsatellites. *Nucleic Acids Res.* **2003**, *31*, 974–980.
108. Ramsay, L.; Macaulay, M.; Degli Ivanissevich, S.; MacLean, K.; Cardle, L.; Fuller, J.; Edwards, K.J.; Tuveesson, S.; Morgante, M.; Massari, A. A simple sequence repeat-based linkage map of barley. *Genetics* **2000**, *156*, 1997–2005.
109. Estoup, A.; Jarne, P.; Cornuet, J.M. Homoplasmy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Mol. Ecol.* **2002**, *11*, 1591–1604.
110. Estoup, A.; Tailliez, C.; Cornuet, J. M.; Solignac, M. Size homoplasmy and mutational processes of interrupted microsatellites in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae). *Mol. Biol. Evol.* **1995**, *12*, 1074–1084.
111. Angers, B.; Estoup, A.; Jarne, P. Microsatellite size homoplasmy, SSCP, and population structure: A case study in the freshwater snail *Bulinus truncatus*. *Mol. Biol. Evol.* **2000**, *17*, 1926–1932.
112. Jarne, P.; Lagoda, P. J. L. Microsatellites, from molecules to populations and back. *Trends Ecol. Evol.* **1996**, *11*, 424–429.
113. Curtu, A. L.; Finkeldey, R.; Gailing, O. Comparative sequencing of a microsatellite locus reveals size homoplasmy within and between European oak species (*Quercus* spp.). *Plant Mol. Biol. Rep.* **2004**, *22*, 339–346.
114. Goldstein, D.B.; Linares, A.R.; Cavalli-Sforza, L.L.; Feldman, M.W. An evaluation of genetic distances for use with microsatellite loci. *Genetics* **1995**, *139*, 463–471.
115. Selkoe, K.A.; Toonen, R.J. Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. *Ecol. Lett.* **2006**, *9*, 615–629.
116. Thiel, T.; Michalek, W.; Varshney, R.; Graner, A. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* **2003**, *106*, 411–422.
117. Kota, R.; Varshney, R.K.; Thiel, T.; Dehmer, K.J.; Graner, A. Generation and comparison of EST-derived SSRs and SNPs in barley (*Hordeum vulgare* L.). *Hereditas* **2001**, *135*, 145–151.

118. Saha, M.C.; Mian, M.A.R.; Eujayl, I.; Zwonitzer, J.C.; Wang, L.; May, G.D. Tall fescue EST-SSR markers with transferability across several grass species. *Theor. Appl. Genet.* **2004**, *109*, 783–791.
119. Gupta, P.K.; Rustgi, S.; Sharma, S.; Singh, R.; Kumar, N.; Balyan, H.S. Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. *Mol. Genet. Genomics* **2003**, *270*, 315–323.
120. Dakin, E.E.; Avise, J.C. Microsatellite null alleles in parentage analysis. *Heredity* **2004**, *93*, 504–509.
121. Fraser, L.G.; Harvey, C.F.; Crowhurst, R.N.; Silva, H.N.D. EST-derived microsatellites from *Actinidia* species and their potential for mapping. *Theor. Appl. Genet.* **2004**, *108*, 1010–1016.
122. Rungis, D.; Berube, Y.; Zhang, J.; Ralph, S.; Ritland, C.E.; Ellis, B.E.; Douglas, C.; Bohlmann, J.; Ritland, K. Robust simple sequence repeat markers for spruce (*Picea* spp.) from expressed sequence tags. *Theor. Appl. Genet.* **2004**, *109*, 1283–1294.
123. Rico, C.; Rico, I.; Hewitt, G. 470 million years of conservation of microsatellite loci among fish species. *Proc. R. Soc. Lond., B, Biol. Sci.* **1996**, 549–557.
124. FitzSimmons, N.N.; Moritz, C.; Moore, S.S. Conservation and dynamics of microsatellite loci over 300 million years of marine turtle evolution. *Mol. Biol. Evol.* **1995**, *12*, 432–440.
125. Peakall, R.; Gilmore, S.; Keys, W.; Morgante, M.; Rafalski, A. Cross-species amplification of soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: Implications for the transferability of SSRs in plants. *Mol. Biol. Evol.* **1998**, *15*, 1275–1287.
126. Chen, X.; Cho, Y.; McCouch, S. Sequence divergence of rice microsatellites in *Oryza* and other plant species. *Mol. Genet. Genomics* **2002**, *268*, 331–343.
127. White, G.; Powell, W. Cross-species amplification of SSR loci in the Meliaceae family. *Mol. Ecol.* **1997**, *6*, 1195–1197.
128. Ma, K.H.; Kim, N.S.; Lee, G.A.; Lee, S.Y.; Lee, J.K.; Yi, J.Y.; Park, Y.J.; Kim, T.S.; Gwag, J.G.; Kwon, S.J. Development of SSR markers for studies of diversity in common buckwheat. *Theor. Appl. Genet.* **2009**, *119*, 1247–1254.
129. Kostia, S.; Varvio, S.L.; Vakkari, P.; Pulkkinen, P. Microsatellite sequences in a conifer, *Pinus sylvestris*. *Genome* **1995**, *38*, 1244–1248.
130. Marcotte, E.M.; Pellegrini, M.; Yeates, T.O.; Eisenberg, D. A census of protein repeats. *J. Mol. Biol.* **1999**, *293*, 151–160.
131. Wren, J.D.; Forgacs, E.; Fondon, J.W.; Pertsemliadis, A.; Cheng, S.Y.; Gallardo, T.; Williams, R.S.; Shohet, R.V.; Minna, J.D.; Garner, H.R. Repeat polymorphisms within gene regions: Phenotypic and evolutionary implications. *Am. J. Hum. Genet.* **2000**, *67*, 345–356.
132. Li, Y.C.; Korol, A.B.; Fahima, T.; Beiles, A.; Nevo, E. Microsatellites: Genomic distribution, putative functions and mutational mechanisms: A review. *Mol. Ecol.* **2002**, *11*, 2453–2465.
133. Li, Y.C.; Korol, A.B.; Fahima, T.; Nevo, E. Microsatellites within genes: Structure, function, and evolution. *Mol. Biol. Evol.* **2004**, *21*, 991–1007.
134. Cummings, C.J.; Zoghbi, H.Y. Fourteen and counting: Unraveling trinucleotide repeat diseases. *Hum. Mol. Genet.* **2000**, *9*, 909–916.

135. Streelman, J.T.; Kocher, T.D. Microsatellite variation associated with prolactin expression and growth of salt-challenged tilapia. *Physiol. Genomics* **2002**, *9*, 1–4.
136. Liquori, C.L.; Ricker, K.; Moseley, M.L.; Jacobsen, J.F.; Kress, W.; Naylor, S.L.; Day, J.W.; Ranum, L.P.W. Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science* **2001**, *293*, 864–867.
137. Li, W.H.; Graur, D. *Fundamentals of Molecular Evolution*; Sinauer Pub.: Sunderland, MA, USA, 1991.
138. Ennos, R.A. Utilising Genetic Information in Plant Conservation Programmes. In *Aspects of the Genesis and Maintenance of Biological Diversity*; Hochberg, M.E., Clibert, J., Barbault, R., Eds.; Oxford University Press: Oxford, UK, 1996; pp. 278–291.
139. Bekessy, S.A.; Ennos, R.A.; Burgman, M.A.; Newton, A.C.; Ades, P.K. Neutral DNA markers fail to detect genetic divergence in an ecologically important trait. *Biol. Conserv.* **2003**, *110*, 267–275.
140. Van Tienderen, P.H.; De Haan, A.A.; Van der Linden, C.G.; Vosman, B. Biodiversity assessment using markers for ecologically important traits. *Trends Ecol. Evol.* **2002**, *17*, 577–582.
141. Holderegger, R.; Kamm, U.; Gugerli, F. Adaptive vs. neutral genetic diversity: Implications for landscape genetics. *Landsc. Ecol.* **2006**, *21*, 797–807.
142. Hammer, K.; Heller, J. Promoting the conservation and use of underutilized and neglected crops. *Schriften Genet. Resour.* **1997**, *8*, 223–227.
143. Padulosi, S.; Eyzaquirre, P.; Hodgkin, T. Challenges and Strategies in Promoting Conservation and Use of Neglected and Underutilized Crop Species. In *Perspectives on New Crops and New Uses*; Janick, J., Ed.; ASHS Press: Alexandria, VA, USA, 1999; pp. 140–145.
144. Ellwood, S.R.; D’Souza, N.K.; Kamphuis, L.G.; Burgess, T.I.; Nair, R.M.; Oliver, R.P. SSR analysis of the *Medicago truncatula* SARDI core collection reveals substantial diversity and unusual genotype dispersal throughout the Mediterranean basin. *Theor. Appl. Genet.* **2006**, *112*, 977–983.
145. Zhao, W.G.; Chung, J.W.; Lee, G.A.; Ma, K.H.; Kim, H.H.; Chung, I.M.; Kang, H.K.; Kim, T.S.; Lee, J.K.; Kim, N.S.; Park, Y.J. Molecular genetic diversity and population structure of a selected core set in garlic (*Allium sativum* L.) using novel SSR markers. *BMC Plant Biol.* **2009**, Submitted.
146. Kim, K.W.; Chung, H.K.; Cho, G.T.; Ma, K.H.; Chandrabalan, D.; Gwag, J.G.; Kim, T.S.; Cho, E.G.; Park, Y.J. PowerCore: A program applying the advanced M strategy with a heuristic search for establishing core sets. *Bioinformatics* **2007**, *23*, 2155–2162.
147. Yu, J.W.; Dixit, A.; Ma, K.H.; Chung, J.W.; Park, Y.J. A study on relative abundance, composition and length variation of microsatellites in 18 underutilized crop species. *Genet. Resour. Crop Evol.* **2009**, *56*, 237–246.

Sample availability: Not available.