*The Agriculturists 13(1): 152-172(2015)* ISSN 2304-7321 (Online), ISSN 1729-5211 (Print)
A Scientific Journal of Krishi Foundation ***Indexed Journal***

***Review Article***

# Development of Microsatellites: A Powerful Genetic Marker

M. Moniruzzaman[1*], R. Khatun[1], Zahira Yaakob[1], M. S. Khan[2] and A. A. Mintoo[3]

*[1]The National University, Malaysia; [2]Bangladesh Council of Scientific and Industrial Research; [3]Lal Teer Seed limited, Bangladesh*

**Corresponding author and Email: monirbge@gmail.com*

Received: 27 April 2014 Accepted: 06 June 2015

**Abstract**

The tandem repeats, conserved short segments of DNA, which are found in all prokaryotic and eukaryotic genomes, are called microsatellites. It is also known as variable number tandem repeats (VNTRs), simple sequence repeats (SSRs) and short tandem repeats (STRs). Microsatellites present in both coding and non-coding regions of a genome. The high polymorphism of microsatellites makes them powerful genetic markers for genome mapping of many organisms. It is also suitable for ancient and forensic DNA studies for population genetics and conservation of biological resources. The major disadvantage of microsatellites is that for the first time they need to be isolated *de novo* from most species being examined. The task of microsatellite isolation is quite cumbersome involving in terms of effort and time, because it traditionally involves screening of genomic libraries. Cross-species amplification, Mining microsatellites from nucleotide sequenced data and Genomic library- based method are general methods of microsatellite isolation. Cross-species method may not effective for all species, Data mining is not applicable if there is no or limited data of DNA sequence. Genomic library based method is the best choice. Traditional protocol, primer extension protocol, selective hybridization, and Fast Isolation by AFLP of Sequences containing repeats (FIASCO) are the protocols of microsatellite development based on genomic library. FIASCO is the best protocol ever developed.

***Keywords:*** microsatellites development, genetic marker, FIASCO, data mining

## 1. Introduction

### *1.1. Microsatellites: the difficulty of isolation*

Microsatellites or simple sequence repeats (SSRs) are tandem repeated motifs of 1-6 bases found in all prokaryotic and eukaryotic genomes analyzed to date. They are present in both coding and no-ncoding regions and are usually characterized by a high degree of length polymorphism (Sawaya *et al*., 2013). The origin of such polymorphism is still under debate though it appears most likely to be due to slippage events during DNA replication (Schltterer and Tautz, 1992). Despite the fact that the mechanism of microsatellite evolution is still unclear, SSRs were being widely employed in many fields soon after their first description (Litt and Luty, 1989; Tautz, 1989; Weber and May, 1989) because of the high variability which makes them very powerful genetic markers. Microsatellites have proven to be an extremely valuable tool for genome mapping in many organisms (Schuler *et al*., 1996; Knapik *et al*., 1998), their applications span over different areas ranging from ancient and forensic DNA studies, to population genetics and

conservation/management of biological resources ( Cavagnaro *et al*., 2011; Cuadrado *et al*., 2008*;* Dalamu *et al*., 2012; Heubl, 2010; Jarne and Lagoda, 1996; Li *et al*., 2012; Park *et al*., 2009*;* Stolle *et al*., 2013*)* as well as commercial application, true hybrid identification (Ashok *et al*., 2011, Manigbas and Villegas, 2004).

The major drawback of microsatellites is that they need to be isolated *de novo* from most species being examined for the first time. This is because microsatellites are usually found in non-coding regions where the nucleotide substitution rate is higher than in coding regions. Consequently, the strategy of designing 'universal primers' matching conserved sequences, which was effective for mitochondrial DNA (Kocher *et al*., 1989), is more problematic for microsatellites. It needs time and intensive labor to identify (Quentin *et al*., 2009). However, the presence of highly conserved flanking regions has been reported for some microsatellite loci in cetaceans (Schltterer *et al*., 1991), turtles (FitzSimmons *et al*., 1995) and fish (Rico *et al*., 1996), allowing cross-amplification from species that diverged as long as 470 million years ago.

It should be noted that during the isolation procedure, loci are selected from the upper end of the repeat length distribution in the genome, the fraction which is known to harbour the most polymorphic markers (Primmer *et al*., 1996). Such bias in loci isolation may likely result in a lower level of polymorphism when orthologous loci are tested in other species (Ellegren *et al*., 1995). Therefore, high polymorphism observed in a species does not guarantee that similar polymorphism will be found in related species especially when increasing the evolutionary distance (Morin *et al*., 1998; Rubinsztein *et al*., 1995).

Reports on birds (Primmer *et al*., 1996) and cattle (Moore *et al*., 1991) suggest a 50% success rate in cross amplification and polymorphism detection in species which diverged from 10 to 20 Ma. This is in agreement with the empirical finding that cross-species amplification works for closely related taxa such as species belonging to the same genus or to recently separated genera (Scribner and Pearce, 2000). The task of microsatellite isolation can be quite involving in terms of effort and time because it traditionally consists of screening genomic libraries with appropriate probes (Rassmann *et al*., 1991). The number of positive clones (containing microsatellites) that can be obtained by means of this traditional method usually ranges from 12% to less than 0.04%. Such an isolation strategy can be effective only in taxa with a high frequency of microsatellites, as in some fish or other vertebrates, and whenever only a relatively low number of microsatellites are needed. This can be the case in population allocation and/or parentage assignment studies, where, given sufficient allelic diversity, a relatively low number of loci (often less than seven) may be sufficient to achieve a high probability of correct assignment as suggested by Bernatchez and Duchesne (2000). However, the statistical power depends not only on the number of scored loci but also on other factors such as the degree of polymorphism of each locus and the sample size, and so the use of a limited number of loci might fail to provide sufficient information.

Traditional strategies are less useful when dealing with taxa with a very low frequency of microsatellites such as birds or plants, or when a large number of microsatellites is required as in the case of studies on genetic distances between populations (Cooper *et al*., 1999, Zhivotovsky and Feldman, 1995) or when constructing a genetic map (Liu, 1997). The recovery rate for useful SSR primers is generally low due to different reasons: a) the primer may not amplify any PCR product; b) the primer may produce very complex, weak or nonspecific amplification patterns; c) the amplification product may not be polymorphic.

Several new protocols, overcoming these limitations, have appeared in the literature during the last few years. This review presents various

methods of microsatellite development so far described with the purpose of providing useful guidelines in choosing the appropriate protocol among the large number of currently available options.

**2. Methods for microsatellite development**

As mentioned above, the major drawback of microsatellites is that they need to be isolated and characterized before being used for the first time. Generally, microsatellites can be developed by the following approaches:

***2.1. Cross-species amplification***

The sequences of flanking region are generally conserved across individuals of the same species and sometimes of different species; a particular microsatellite locus can often be identified by its flanking sequences. The presence of highly conserved flanking regions has been reported for some microsatellite loci in cetaceans, turtles and fish allowing cross-amplification from species that diverged as long as 470 million years ago (Ma).

In cross-species amplification method, the first step is to search published literature and public databases for any existing microsatellite primers for the target species or closely related species. The availability of microsatellite markers for a given species will be a combination of past interest in that species (and related species) and the inherent success rate of microsatellite development for that taxon. There are clear differences in the frequency of microsatellite regions in the genomes of plants, animals, fungi and prokaryotes (Toth *et al.,* 2000) and the success rate of isolating microsatellite markers often scales with their frequency in the genome (Zane *et al.,* 2002).

Currently, many microsatellite markers are reported as primer notes in a specialized journal "Molecular Ecology Notes" (now changed as "Molecular Ecology Resources"). There is a searchable database online for any microsatellite primers published in this journal (http://tomato.bio.trinity.edu/). The sequences themselves are archived in GenBank, and are often submitted long before their use appears in published studies. GenBank can be searched with a web-based engine run by the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/) by typing in the species, genus or family name, the term microsatellite and selecting the Nucleotide database (Benson *et al.,* 2008; Wheeler *et al.,* 2008).

***2. 2. Genomic library- based method***

The development of microsatellite markers involves several distinct steps from obtaining the library to develop a working set of primers that can amplify polymorphic microsatellite loci (Roder *et al.,* 1998). These include:

a) Microsatellite library construction.
b) Identification of unique microsatellite loci.
c) Identifying a suitable area for primer design.
d) Obtaining a PCR product.
e) Evaluation and interpretation of banding patterns.
f) Assessing PCR products for polymorphism.

***2.2.1. Traditional method for Microsatellites development***

Traditionally, microsatellite loci have been isolated from a partial genomic library of the target species (Figure 1). High quality genomic DNA is fragmented either using restriction enzymes or, less commonly, by sonication. In the former case, the choice of the restriction enzyme depends on the desired average length of DNA fragments, the microsatellite repeat to be found, and the type of ends (cohesive or blunt) of the restriction fragments. Fragmented DNA is then size -selected to preferentially obtain small fragments (300-700 bp).

Depending on the fragmentation method, DNA fragments are ligated into a common plasmid vector either directly or after ligation to specific adaptors. This step is most critical, due to the risk of obtaining low numbers of recombinants and the formation of concatamers between genomic fragments. Transformation of bacterial

cells with ligation product generally yields thousands of recombinant clones that can be subsequently screened for the presence of microsatellite sequences. Screening for positive clones is generally carried out by means of Southern hybridization using repeat-containing probes, after blotting bacterial colonies onto nylon membranes. Colony transfer can be carried out either by classical replica plating or by picking single colonies and ordering them in new arrayed plates. While the later method is more time consuming and limits the total number of screened clones, it avoids the requirement of reprobing positive clones for confirmation.

Repeat-containing probes can be synthesized *de novo*, or alternatively a genomic clone, which contains a microsatellite locus that has already been isolated, can be used. Hybridization probe(s) can be labelled by both radioactive ($^{32}$P, $^{33}$P) and nonradioactive (digoxigenin) methods. Radioactive protocols are generally more sensitive, but the need for dedicated equipment and laboratory space for the manipulation of radionucleotides might pose limitations for those researchers that have no access to such facilities. Moreover, the short-life of radioisotopes makes radio-labelled probes of limited use. Efficiency of nonradioactive labelling techniques has greatly improved in recent years, and these methods allow less stringent and safer working conditions, with the additional bonus of the long-term storage of probes (further information on nonradioactive techniques can be found at http://www.inapg.inra.fr/dsa/microsat/microsat.htm).

Following identification of repeat containing clones, specific primers are designed and PCR conditions are optimized to allow the amplification of each locus from different individuals of a population (Cifarelli *et al.*, 1995).

A different approach (PCR isolation of microsatellite arrays; PIMA), which skips all steps from DNA fragmentation to cloning, has been proposed by Lunt *et al.* (1999). Briefly, several RAPD primers are used to obtain randomly amplified fragments from the target species genome. These amplicons are cloned by using a T vector and arrayed clones are screened using repeat specific and vector primers. This and similar techniques (D'Amato *et al.*, 1999; Ender *et al.*, 1996) take advantage of the fact that RAPD fragments seem to contain microsatellite repeats more frequently than random genomic clones (Cifarelli *et al.*, 1995).

#### *2.2.2. Primer extension method for Microsatellite development*

A different strategy, based on primer extension, has been proposed for the production of libraries enriched in microsatellite loci.

Two protocols have been proposed to produce genomic libraries that are highly enriched for specific microsatellite repeats using a primer extension reaction (Ostrander *et al.*, 1992; Paetkau, 1999). Both methods rely on the construction of a 'primary' genomic library, in which fragmented genomic DNA is inserted into a phagemid or a phage vector (Figure 2a) in order to obtain a single strand DNA (ssDNA) library. ssDNA is then used as a template for a primer extension reaction, primed with repeat-specific oligonucleotides, which generates a double strand product only from vectors containing the desired repeat. The two enrichment procedures diverge in the strategy used to recover primer-extended products (Figure 2b).

In the Ostrander and coworkers approach, 40 000- 60 000 colonies from a phagemid library are eluted from LB-agar plates, grown to saturation in liquid media and super infected with M13 helper phage. Because of the particular genotype of the bacterial host (*dut- ung-*), super infection results in a library of circular ssDNA containing uracil instead of thymine. After the selective conversion of ssDNA to double strand DNA through (CA)*n* or (GT)*n* primer extension, the mixture is used to transform a *dut+ ung+ Escherichia coli* strain. The resulting library is highly enriched for repeat containing inserts

because the native single strand products transform with very low efficiency, and because the uracil containing ssDNA will be degraded by the host uracil-N-glycosilase (*ung+*). In contrast, only double-stranded DNA products can be rescued because the thymidine-containing primer-extended strand allows for the action of host repair mechanisms.

In the Paetkau protocol the primary library is obtained using M13 phage, and circular ssDNA is obtained through elution of 30 000 clear plaques. Primer extension is then performed using 5'biotinylated oligonucleotides and Klenow DNA polymerase. This reaction results, for microsatellite containing phages, in a population of circular DNA molecules whose second strand is a linear primer-extended molecule of DNA with a biotin at one end. These products are selectively recovered from the reaction mix using streptavidin-coated beads and after washing steps, circular phage ssDNA is released by denaturation. Finally, molecules containing the microsatellites are converted to double-stranded molecules with a second round of primer extension and are then used for the final transformation.

#### *2.2.3. Selective hybridization protocols for microsatellites loci isolation*

A very simple strategy for microsatellite isolation using selective hybridization can be outlined based on several published reports (Hamilton *et al*., 1999; Karagyozov *et al*., 1993). The first step is identical to traditional isolation procedures, aimed at producing small genomic fragments that are then ligated to a known sequence, a vector or an adaptor (Figure 3). Because the enrichment strategy is dependent on the ability to recover, after selective hybridization, microsatellite-containing DNA by PCR amplification, this step is very important. Following the fragmentation-ligation step, and depending on the amount of starting DNA, the DNA is hybridized (if necessary after amplification) with the repeat containing probe. The probe can be bound to a nylon membrane (Armour *et al*., 1994; Karagyozov *et al*., 1993) or 5'biotinylated and bound to streptavidin coated beads (Kandpal *et al.*, 1994). After the hybridization step and several washes to remove nonspecific binding, the DNA is eluted and recovered by PCR amplification. Finally, the enriched DNA is cloned into a suitable vector, either by using a restriction site on the known flanking regions or by TA cloning.

Depending on the efficiency of the whole procedure, recombinant clones can be directly sequenced or screened for the presence of repeats by using Southern blotting or PCR strategies.

#### *2.2.4 FIASCO (Fast Isolation by AFLP of Sequences containing repeats) a fast and effective collage protocol, tested in the laboratory*

The method is fast and simple, and many unnecessary steps have been eliminated. The protocol relies on the extremely efficient digestion-ligation reaction of the amplified fragment length polymorphism (AFLP) procedure (Vos *et al*., 1995). This method is being widely used for the development of microsatellite from most species (Aihua *et al.*, 2012; Da-Long *et al*., 2012; Jing-Jing *et al.*, 2013; Jirapong *et al.*, 2012; Shu-Zhen *et al.*, 2010; Takundwa *et al., 2012;* Xiaolian *et al.*, 2011; Zahra *et al.*, 2012; Zalapa *et al*., 2012). FIASCO technique involves the following sections:

***DNA Extraction is the Beginning of Molecular Marker Analysis***

Extraction (isolation) of DNA (nuclear, mitochondrial, and/or chloroplast DNA) from sample to be studied is the first step for all molecular marker types. DNA can be extracted either from fresh, lyophilized, preserved or dried samples but fresh material is ideal for obtaining good quality DNA. There are many alternative protocols for DNA extraction and the choice of a protocol depends on the quality and quantity of DNA needed, nature of samples, and the presence of natural substances that may interfere with the extraction and subsequent analysis. DNA extraction protocols vary from simple and

quick ones (Clancy *et al.,* 1996; Dayteg *et al.,* 1998; Ikeda *et al.*, 2001; von Post *et al.,* 2003) that yield low quality DNA but nevertheless, good enough for routine analyses to the laborious and time-consuming standard methods (Dellaporta *et al.*, 1983; Murray and Thompson, 1980; Saghai-Maroof *et al.*, 1984) that usually produce high quality and quantity of DNA. The most commonly used DNA extraction protocols involve breaking (through grinding) or digesting away cell walls and membranes in order to release the cellular constituents. Removal of membranes lipids is facilitated by using detergent (Semagn *et al.,* 2006) 2541 reagents such as sodium dodecyl sulphate (SDS), Cetyl trimethylammonium bromide (CTAB) or mixed alkyl trimethyl-ammonium bromide (MTAB). The released DNA should be protected from endogenous nucleases and EDTA is often included in the extraction buffer to chelate magnesium ions that is a necessary co-factor for nucleases. DNA extracts often contain a large amount of RNA, proteins, polysaccharides, tannins and pigments, which may interfere with the extracted DNA. Most proteins are removed by adding a protein degrading enzyme (proteinase-K), denaturation at 65$^{o}$C and precipitation using chloroform and isoamyl alcohol. RNAs are normally removed using RNA degrading enzyme called RNase A. Polysaccharide-like contaminants are,however, more difficult to remove. NaCl, together with CTAB is known to remove polysaccharides (Murray and Thompson, 1980; Paterson *et al.,* 1993). Some protocols replace NaCl by KCl (Thompson and Henry, 1995). As DNA will be released along with other compounds (lipids, proteins, carbohydrates, and/or phenols), it needs to be separated from others by centrifugation. The DNA in the aqueous phase will then be transferred into new tubes and precipitated in salt solution (e.g. sodium acetate) or alcohol (100% isopropanol or ethanol), redissolved in sterile water or buffer. Finally, the concentration of the extracted DNA needs to be measured using either 1% agarose gel electrophoresis or spectrophotometer. Agarose gel is useful to check whether the DNA is degraded or not but estimating DNA concentration by visually comparing band intensities of the extracted DNA with a molecular ladder of known concentration is too subjective. Spectrophotometer measures the intensity of absorbance of DNA solution at 260 nm wavelength, and also indicates the presence of protein contaminants but it does not tell whether the DNA is degraded or not. There are three possible outcomes at the end of any DNA extraction:

a) There is no DNA.

b) The DNA appears as sheared (too fragmented), which is an indication of degradation for different reasons.

c) DNA appears as whitish thin threads (good quality DNA) or brownish thread (DNA in the presence of oxidation from contaminants such as phenolic compounds).

***Restriction enzyme digests and purification***

Isolated DNA has to be digested with the enzymes (commonly used *Msp* 1, *Csp* 6I and *Sau 3A*) as per the supplier's instructions. The digest also has to be cleaned using PCR purification kit.

***Adaptor ligation and amplicon preparation***

An adapter is a short, chemically synthesized, double stranded DNA molecule which is used to link the ends of two other DNA molecules. It may be used to add sticky ends to cDNA allowing it to be ligated into the plasmid much more efficiently. Adapter has to be ligated with small fragment of DNA then ligated DNA has to be amplified with PCR. Takundwa *et al.* (2012) performed adapter ligation and amplification as 500 ng of the restriction-digested DNA, 1 µl of 12 mer adapter, 1µl 24 mer adapter and 3 µl ligation buffer in a final volume of 28 µl. The reaction mixture was heated to 72 °C for 3 min, then cooled by one degree per minute to 4 °C. Two microlitres of (2µl) ligase were added and the reaction incubated at 4 °C for 16 hours. The 12 bp adaptors were removed by heating to 72 °C for five minutes to melt off the 12 mer, followed by purification using PCR purification kit. The ligated DNA was amplified with PCR by combining 5 µl PCR buffer, 5 µl (20 mM) MgCl2, 4 µl (10 mM) dNTPs, 2 µl adapter (100

μM), 34.75 μl water, and 1 μl ligated DNA. The reaction was heated at 72 °C for 5 min; 5 units of *Taq* polymerase enzyme was added and further incubated for 5 min at 72°C. The DNA was amplified for 20 cycles of 95 °C for 30 seconds, and 72 °C for 90 seconds, with a final hold at 72°C for five minutes.

***SSR Enrichment***

This step is for finding SSR enriched DNA fragment. It involves denature and annealing of probes then allow for PCR reaction finally electrophoresis to separate denature DNA. This DNA is SSR enriched. Takundwa *et al*. (2012) developed microsatellite form bean by performing as, the amplified digest was denatured and annealed to the biotinylated primer by combining 20 μl PCR product (200 μg) and 1 μl primer (10 μM) and heating at 95 °C for 5 min. This was followed by incubation at room temperature for 30 min. Before combining the primed DNA with streptavidin beads, 10 μl of unrelated DNA (sheared herring sperm at 1 mg ml-1) was added to minimise non-specific binding. The annealed DNA mixture was then added to 1 mg of magnetic beads and incubated for 30 minutes at room temperature, allowing the streptavidin beads to join with the biotinlyated primers. Five washes with TEN100 (Tris/EDTA/ NaCl) and 5 washes with SSC 0.2X SDS 0.1% were performed to remove non-specific DNA. Then, two denaturation steps were performed to separate DNA containing SSRs from the beads. The first was done by adding 50 μl of TE (Tris-HCl 10 mM, EDTA 1 mM) and heating to 95 °C for 5 minutes. The remaining solution was separated magnetically and stored. The second denaturation used 12 μl 0.5N NaOH, which was neutralised with 12 μl 0.5N HCl and separated magnetically. Each denaturation product (2 per enzyme) was amplified separately with PCR by adding 5 μl PCR buffer, 3 μl MgCl2 (25 mM), 4 μl dNTPs (each 2.5 mM), 2 μl adapter (10 μM, 34.75 μl water, 0.25 μl r*Taq* (2 units) and 1 μl DNA into a PCR tube. The mixture was cycled 20 times from 95 to 72 °C. Electrophoresis was then performed on each of the denaturation products with a 1.5% agarose gel.

***Sequencing of SSR enriched DNA***

All the microsatellite-enriched genomic DNA pools are subjected to sequencing. Form the sequenced data primers are to be designed.

***Microsatellite discovery and primer design***

Primers are then developed around the SSR sites, identifying by using software (described later, 3.2. Primer design for microsatellites).

***SSR primer screening***

Each of the primer sets are screened to DNA of interested individual by PCA amplification.

***SSR data analysis***

The discriminatory power ($D_L$) value represents the probability that two randomly chosen individuals show different allelic patterns at the same microsatellite locus and, thus, are distinguishable from one another. That is, if $p_i$ is the proportion of the population carrying the *i*th banding patterns at the *j*th primer and if $p_i$ are calculated for each pattern generated by the primer (Tessier *et al.*, 1999), then $D_L = 1- \Sigma p_i^2$.

## 3. Mining microsatellites from nucleotide sequences

### *3.1. Sources of the data*

Sequences from both genomic DNA and cDNA can be used for microsatellite mining. Though some researchers produced DNA sequences and kept them in their own laboratories, publicly accessible nucleotide databases are the major source in many studies today including microsatellite mining. The GenBank sequence database is an annotated collection of all publicly available nucleotide sequences and their protein translations. This database is produced at National Center for Biotechnology Information (NCBI) as part of an international collaboration with the European Molecular Biology Laboratory (EMBL) Data Library from the European Bioinformatics Institute (EBI) and the DNA Data Bank of Japan (DDBJ). GenBank and its collaborators receive sequences produced in laboratories throughout the world from more than 100,000 distinct organisms. GenBank

continues to grow at an exponential rate, doubling every 18 months. As of 15 February 2014, GenBank release 200.0 that has 171123749 loci, 157943793171 bases, from 171123749 reported sequences. GenBank is built by direct submissions from individual laboratories, as well as from bulk submissions from large-scale sequencing centers. GenBank nucleotide records are located in separate databases that must be searched independently. These include dbEST and dbGSS, plus multiple databases for the CoreNucleotide division, including nr, htgs and wgs **ESTs** [http://www.ncbi.nlm.nih.gov/dbEST/] are generally short (<1 kb), single-pass cDNA sequences from a particular tissue and/or developmental stage. However, they can also be longer sequences that are obtained by differential display or Rapid Amplification of cDNA Ends (RACE) experiments. ESTs are particularly attractive for marker development since they represent coding regions of the genome and putative function can often be deduced by homology searches although little is known about many of the ESTs. While ESTs provide means for the identification of genes, microsatellites provide high level of polymorphism. Microsatellites identified in ESTs are typically referred to as EST-SSRs or genic SSRs, contrasting to type II SSRs which come from random sequences of the genome. The identification of ESTs has preceded rapidly, with approximately 74 million ESTs sequences now available in public databases (GenBank 1/2013, all species). As a byproduct of EST or BAC sequencing projects in many organisms, microsatellite-mining from SSR-containing ESTs is inexpensive and time-saving, and has proved to be an effective approach to develop microsatellies for genetic map and population genetics studies in animals and plants (Wang *et al.,* 2005; Yue *et al.,* 2004).

**STSs** [http://www.ncbi.nlm.nih.gov/dbSTS/] are short genomic landmark sequences. They are operationally unique in that they are specifically amplified from the genome by PCR amplification. In addition, they define a specific location on the genome and are, therefore, useful for mapping.

**GSSs** [http://www.ncbi.nlm.nih.gov/dbGSS/] are also short sequences but are derived from genomic DNA, about which little is known. They include, but are not limited to, single-pass GSSs, BAC ends, exon-trapped genomic sequences, and AluPCR sequences. EST, STS, and GSS sequences reside in their respective divisions within GenBank, rather than in the taxonomic division of the organism. The sequences are maintained within GenBank in the dbEST, dbSTS, and dbGSS databases. ESTs are particularly attractive for marker development as they represent coding regions of the genome and putative function can often be deduced by homology searches. While ESTs provide means for the identification of genes, microsatellites provide high levels of polymorphism.

### *3.2. Finding and characterizing repeat motifs*

Traditionally, SSR isolation has relied on the screening of genomic libraries using repetitive probes and sequencing of positive clones in order to develop locus-specific primers. These processes are necessary for many organisms but normally time-consuming and labor intensive. Mining SSR from public databases has been streamlined with technological advancement and protocol optimization to make the process cheaper, more efficient and more successful, and has proved to be an effective approach to develop microsatellites for genetic map and population genetics studies in animals (Chen *et al.*, 2005; Maneeruttanarungroj *et al.*, 2006; Pérez *et al.*, 2005; Serapion *et al.*, 2004; Wang *et al.*, 2005; Yue *et al.*, 2004) and plants (Chen *et al.*, 2006; Cordeiro *et al.*, 2001; Kantety *et al.*, 2002).

The steps of SSRs mining from common carp EST data are as follows:

***a. Download EST sequences from public databases***

The target ESTs from the NCBI dbEST database has to be downloaded into VectorNTI software (InforMax Inc.). ‘‘common carp EST’’ can be

used as a keyword to search nucleotide sequences at the NCBI databases (http://www.ncbi.nlm.nih.gov). All matched sequences can be downloaded by changing the ''display'' window to FASTA, and the ''send to'' window to FILE. Finally the sequences have to be saved as a text file.

**a. 1. Tools for microsatellite mining**

In general, microsatellite-finding tools can be classified broadly into three subcategories based on their architecture: first, such as MISA and TROLL *etc*; second, Tandem-Repeats Finder (TRF) etc; third, ATR and ETR, *etc* (Table 1). (Prakash *et al.*, 2007).

***b. Related bioinformatics work***

**b. 1. Clustering analysis**

EST sequences are subjected to analyze by cluster analysis using the ContigExpress module in VectorNTI package (available at http://download.invitrogen.com) etc. and linear assembly algorithm was applied. The criteria for clustering may be set at a minimum overlap of 30 bases (default is 20 bases). Each cluster has to be inspected visually to ensure the fidelity of alignment to avoid pseudo-clusters caused by repetitive elements or long strings of microsatellite repeats. After clustering and assembly unique microsatellite-containing ESTs is identified.

**b. 2. Identification of the known genes**

The unique ESTs were then subjected to BLASTx search against the GenBank (protein database) for putative identification of gene function. When accumulated probability of sequence similarity was less than 1×104, the tentative identities were established.

**b. 3. Primer design for microsatellites**

Primer can be designed by using several software (Table 2) for the amplification of repeat regions of interest across the flanking regions. During the primer designing, the range of annealing temperature is set between 45 and 55 °C, and the expected size of PCR products 150-250 bp. A single pair of "best" primers is designed and synthesized for each unique EST or gene that contains SSR, and no repeated designs and syntheses of primers are carried out.

***c. Laboratory verification of predicted microsatellites***

**c. 1. PCR amplification and polymorphism test for microsatellites**

After primer designing, the primer(s) can be amplified using thermo cycler. The optimum PCR condition has to be adjusted. Tong *et al*. (2005) perform the experiment the condition as 94 °C for 5 min, followed by 34 cycles of 94 °C for 35s, appropriate annealing temperature for 35s, and 72°C for 50s, and a final extension of 72°C for 10 min. The PCR reactions were performed in a 25 μl-reaction mixture, which contained 2.5 μl 10×reaction buffer, 2 μl Mg2＋(1.5mmol.L-1), 1μldNTP (10mmol.L-1), 0.5U *Taq* polymerase(2U/μl), 2 μl template DNA , 0.25μl each of the primer (5μmol.L-1), and 17 μl sterile water. PCR products have to be separated in 6% denaturing polyacrylamide gel and visualized by silver staining. Allele sizes can be determined by comparison with pBR322 DNA/*Msp* I markers.

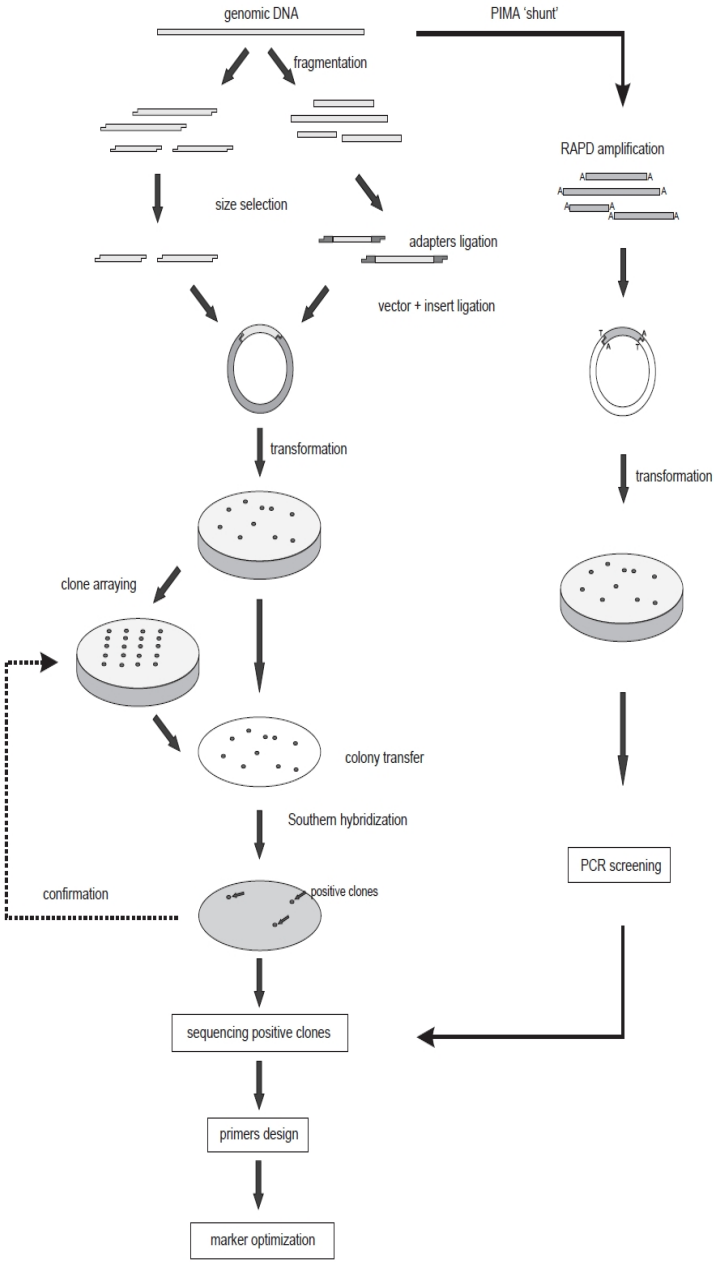


**Figure 1.** Schematic representation of 'traditional' methods for microsatellites isolation, and the alternative PIMA approach (Lunt *et al*. 1999)

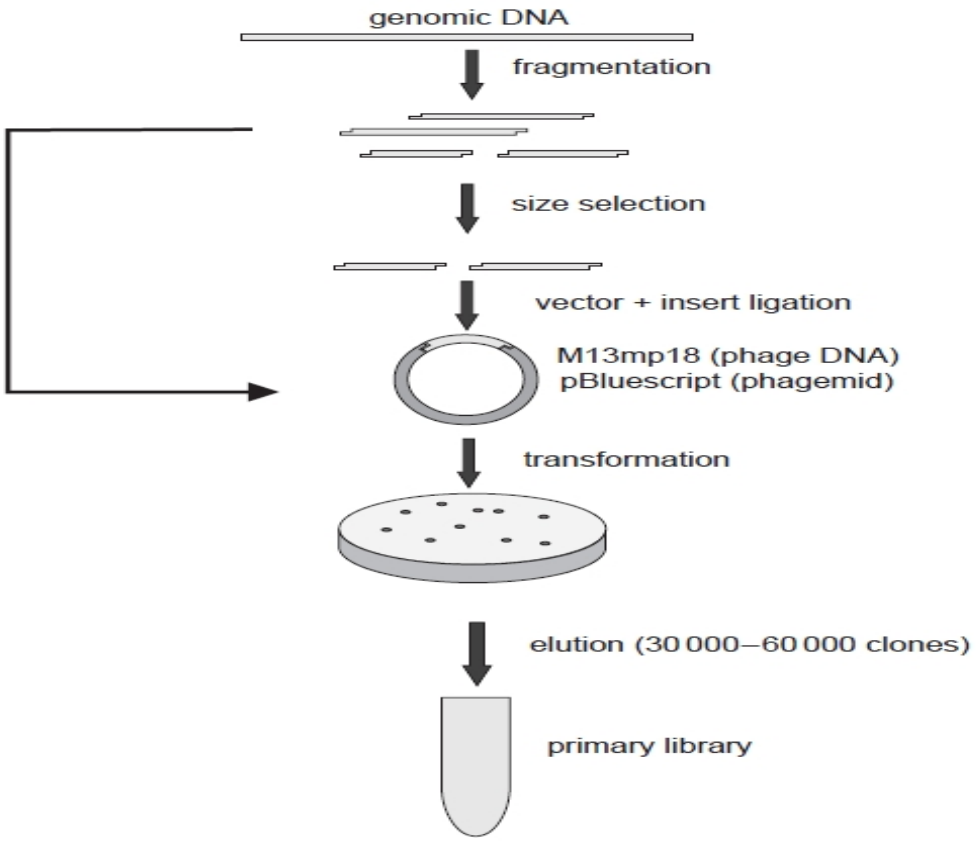


**Figure 2a.** Primer extension enrichment protocols. Schematic representation of the primary library construction (Ostrander *et al.*, 1992; Paetkau, 1999)

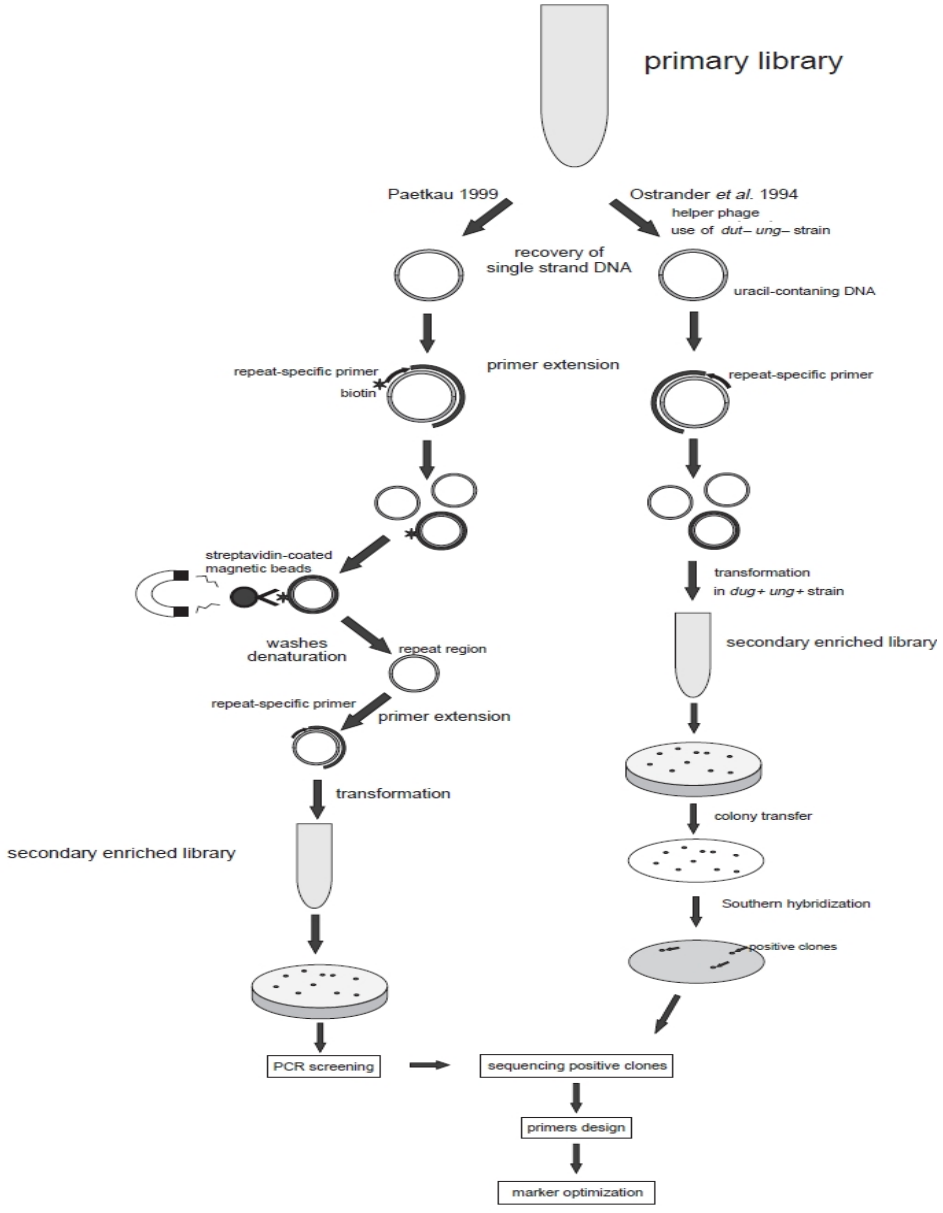


**Figure 2b.** Primer extension enrichment protocols. Schematic representation of protocols from Paetkau *et al.* 1999 (left) and Ostrander *et al.* 1992 (right).

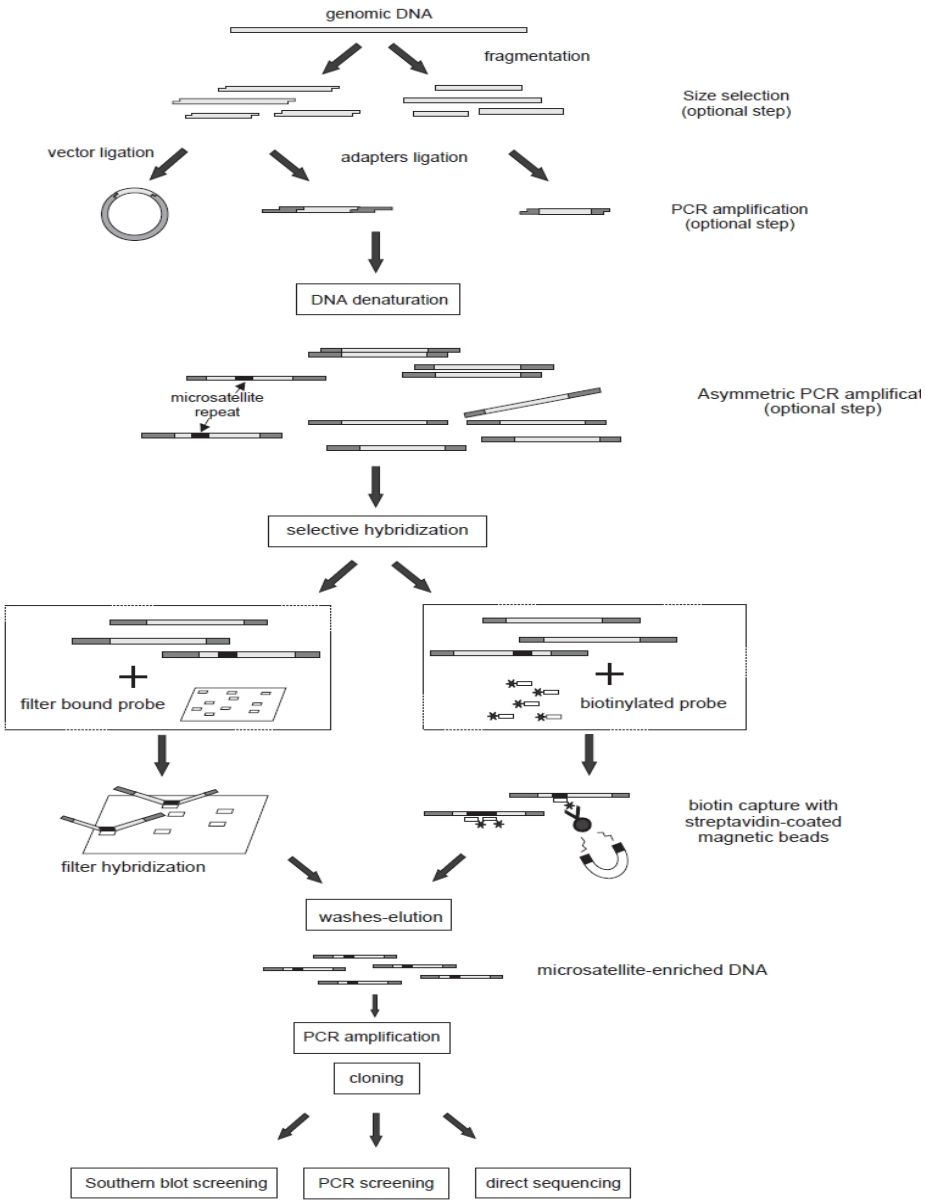


**Figure 3.** Schematic representation of selective hybridization protocol (Karagyozov *et al*., 1993).

**Table 1.** Characteristics of some important microsatellite search tools

| Name, acronym and weblink of the tool | Salient features | Limitations |
|---|---|---|
| Repeatmasker www.repeatmasker.org | Available online and standalone; mines perfect, imperfect and compound repeats; accepts data in multiple formats; presents statistical analysis; returns flanking sequences; MaskerAid, a performance enhancement is available | Runs only on Unix/Linux systems; not specific for microsatellites |
| Sputnik (http://espressosoftware.com/pages/sputnik.jsp and http://cbi.labri.fr/outils/Pise/sputnik.html) | C-language program available online and stand-alone; mines perfect, imperfect and compound repeats; accepts data in multiple formats; improved versions include Modified Sputnik-I and Modified Sputnik-II | Automated statistical analysis files not generated; runs only on Unix/Linux systems; Hexanucleotide repeats are not screened |
| Tandem Repeats Finder (TRF) (http://tandem.bu.edu/trf/trf.html) | Both online and stand-alone versions are GUI; mines perfect, imperfect and compound repeats; platform independent | Accepts input as fasta files only; automated statistical analysis file not generated (TRAP; www.coccidia.icb.usp.br /trap/ [54] can be used); process limited-size files only; output files are numerous and difficult to manage |
| Repeatfinder (www.cbcb.umd.edu/software/RepeatFinder/) | Available online and standalone; mines perfect, imperfect and compound repeats; accepts multiple formats as input | Runs on Unix/Linux systems; not specific for microsatellites |
| eTandem and eQuicktandem (http://bioweb.pasteur.fr/seqanal/interfaces/etandem.html) | Perl script available online and stand-alone; parts of EMBOSS suite; mines perfect, imperfect and compound repeats; accepts input in multiple formats; generates statistics | Runs only on SGI Irix, Linux, Sun solaris and Tru64 Unix |
| REPuter (http://bibiserv.techfak.uni-ielefeld.de/reputer/) | Available online and standalone; stand-alone version can handle large genomic sequences; output cataloged in a format similar to BLAST; statistical and graphical analysis provided; excellent connectivity to BLAST, FASTA. | Limited capacity of online version; accepts data in fasta/plain format only; runs only on sUnix; not specific for microsatellites |
| Simple-Sequence Repeat Identification Tool (SSRIT) and Clemson University Genomics Institute | Perl scripts available online and stand-alone; platform independent (CUGIssr is a modified version of SSRIT) | Finds only perfect repeats; accepts only fasta-formatted files; automated statistical analysis not generated |

| | | |
|---|---|---|
| Simple- Sequence Repeat Tool (CUGIssr)(www.gramene.org/db/searches/ssrtool) | | |
| Tandem Repeats Occurrence Locator (TROLL) (http://wsmartins.net/cgilocal/webtroll/troll.cgi) andWebTROLL (http://wsmartins.net/webtroll/troll.html) | C++ program available online and stand-alone (TROLL downloadable, WebTROLL web interface); identifies perfect, imperfect and compound repeats; alsodesigns primers | Accepts fasta-formatted files only as input; executes only on Linux systems; statistical analysis not provided |
| Microsatellite Analysis Server (MICAS) (http://210.212.212.7/MIC /index.html) | An exclusively web-based Utility | Scans only one file at a time; compound and imperfect repeats are not identified; statistical analysis is not performed |
| MISA (http://pgrc.ipkgatersleben .de/misa/) | Perl script executing only offline; large sequences are handled easily; statistical analysis is generated; platform independent; can design primers using Primer3 by running supplementary scripts | Inappropriate clustering of microsatellite motifs in statistical analysis file; only fasta-formatted files are taken as input; identifies only perfect repeats and compound repeats |
| Mreps (http://bioinfo.lifl.fr/mreps/mreps.php and http://bioweb.pasteur.fr/seqanal/ interfaces/mreps.html) | Available online and standalone; identifies compound and imperfect repeats; accepts data in multiple formats; platform independent; can design primers | Statistical analysis is not performed |
| Search for Tandem Repeats in Genomes (STRING) (http://www.caspur.it/_castri/STRING/) | C-language program availableonline and stand-alone; finds perfect, imperfect and compound repeats; runs well with large genomic sequences; platform independent | Only fasta files taken as input; no automated statistical analysis |
| Search for Tandem Approximate Repeats (STAR) (http://atgc.lirmm.fr/star) | Available online and standalone; searches for 'approximate' tandem repeats of a given motif; platform independent | Does not generate statistical analysis |
| MicrosatDesign (http://daphnia.cgb.indiana.edu/wfleabase/software) | Perl scripts executing as a stand-alone tool; builds database and designs primers from the nascent DNA-sequencer outputs; DNA-sequence trace files are taken as an input; combination of phredPhrap, Primer 3 and GCG software/eTandem software; identifies compound repeats and imperfect | Specific in its use; does not generate statistical analysis |

| | | |
|---|---|---|
| | repeats as well | |
| Poly (http://bioinformatics.org/poly/) | Downloadable Python script; statistical analysis is provided; platform independent | Slow |
| Exact Tandem Repeats Analyzer (E-TRA) and Tandem Repeats Analyzer (TRA) (ftp.akdeniz.edu.tr/Araclar/) | C++ program available online and stand-alone; search microsatellites in ESTs combining with key-word match searches; multiple sequences and multiple files can be handled simultaneously; provide flanking sequences and capable of designing primers; fast; GUI; find perfect, imperfect and compound repeats; accept input in multiple formats; provides statistical Analysis | Redundancy in output |
| msatcommander (http://code.google.com/p/msatcommander/) | Python script available for download; GUI; capable of searching perfect, imperfect and compound repeats with flexibility; output in CSV format; platform independent; primer designing utility available | No online interface; only fasta formatted files accepted as input; statistical analysis is not generated automatically |
| SciRoko (www.kofler.or.at/bioinformatics/SciRoKo/index.html) | C-language program available for stand-alone execution; identifies perfect, imperfect and compound repeats; highly flexible; extremely fast; GUI; provides statistical analysis; platform independent | Depends on .NET Framework |
| Imperfect Microsatellite Extraction (IMEx) (http://203.197.254.154/IMEX/) | C-language program executing stand-alone; finds perfect and imperfect repeats; efficient, fast and user- friendly; returns the coding/ noncoding information of microsatellites; highly flexible; can design primers as well; statistics are generated | Executes on Linux |

**Table 2.** Characterize important software for microsatellites design

| Name of the tools | Features | Limitations |
|---|---|---|
| Primer 3 | Work on line. (http://www.genome.wi.mit.edu/cgi-bin/primer/) | web;C-language |
| Primer 5 | Designing primers for long PCR of sequences up to 50 kb is possible. | Windows |
| Oligo 6 | The graphic features allow screens to be displayed in either a bar or a dot graph. | Windows; Macintosh |
| DNAstar | Sequence assembly and SNP discovery; gene finding; utility for importing unusual file types. Primer design function included. | Windows |
| FASTPCR | Automatically SSR loci detection; direct PCR primers design | Windows |

**4. Conclusions**

Microsatellite is co-dominant, highly distributed throughout the genome and polymorphic marker. Among different DNA markers, Microsatellites have been getting attention to scientist because its features those make it a powerful marker for genetic studies. The main drawback of the use of this marker is, it must be developed *de novo* from most species being examined for the first time. Thus, to develop microsatellite, it requires work efficiency and adequate time. Cross-species amplification is the simple method of marker development and validation. But it may not workable for all species. Data mining is another way, but it will not be fruitful if the sequence data of desired species or related species are unavailable or limited. It is cost effective because many tools for data mining are freely available. Microsatellite development based on genomic library is the right path. It is expensive but reliable. Different scientists develop many protocol based genomic library method. FIASCO (Fast Isolation by AFLP of Sequences Containing repeats) is fast, efficient, simple method.

**References**

Aihua ,W., Ding, Y., Zhenhua, H., Lin, C., Wang, S., Wang B., Zhang, H. and Zhou, G. 2012. Isolation and Characterization of 13 New Polymorphic Microsatellite Markers in the *Phaseolus vulgaris* L. (Common Bean) Genome. *Internal Journal of Molecular Sciences*,13(9):11188–11193.

Armour J. A., Neumann R., Gobert S., Jeffreys A.J. 1994. Isolation of human simple repeat loci by hybridization selection. *Human Molecular Genetics*, 3: 599-565.

Ashok B. D., Mamta P. R., Manoj R B., Kshanada J. 2011. Identification and genetic purity testing of Cotton F1 hybrid using molecular markers. *Indian Journal of Biotechnology*, 10: 301-306.

Benson D. A., Karsch-Mizrachi I., Lipman D.J., Ostell J., Wheeler D.L. 2008. GenBank. *Nucleic Acids Research*, 36: 25-30.

Bernatchez L., Duchesne P. 2000. Individual-based genotype analysis in studies of parentage and population assignment: how many loci, how many alleles? *Canadian Journal of Fisheries and Aquatic Sciences*, 57: 1-12.

Cavagnaro, F., Chung, S. M., Manin, S., Yildiz, M., Ali A., Maria, S. A., Massimo I., Douglas A. S., and Philipp, W. S., 2011. Microsatellite isolation and marker development in carrot - genomic distribution, linkage mapping, genetic diversity analysis and marker transferability across Apiaceae. *BMC Genomics*, 12: 386.

Chen, C. X., Zhou, P., Choi, Y. A., Huang, S., Gmitter, F.G. 2006. Mining and characterizing microsatellites from citrus ESTs. *Theoritical and Applied Genet*ics, 112: 1248-1257.

Chen, S.L., Liu, Y. G., Xu, M. Y., Li, J. 2005. Isolation and characterization of polymorphic microsatellite loci from an EST-library of red sea bream (*Chrysophrys major*) and cross-species amplification. *Molecular Ecology Notes 5*: 215-217.

Cifarelli, R. A., Gallitelli, M., Cellini F. 1995. Random amplified hybridization microsatellites (RAHM): isolation of a new class of microsatellite-containing DNA clones. *Nucleic Acids Research*, 23: 3802-3803.

Clancy, J. A., Jitkov V. A., Han, F., Ullrich S. E. 1996. Barley tissue as direct template for PCR: a practical breeding tool. *Molecular Breeding*, 2: 181-183.

Cooper, G., Amos, W., Bellamy R. 1999. An empirical exploration of the (dm)2 genetic distance for 213 human microsatellite markers. *American Journal of Human Genetics*, 65:1125-1133.

Cordeiro, G. M., Casu, R., McIntyre, C. L., Manners, J. M., Henry, R. J. 2001. Microsatellite markers from sugarcane (*Saccharum* spp.) ESTs cross transferable to erianthus and sorghum. *Plant Science*, 160: 1115-1123.

Cuadrado, A., Cardoso, M., Jouve, N. 2008. Physical organisation of simple sequence repeats (SSRs) in Triticeae: structural, functional and evolutionary implications. *Cytogenet and Genome Resarch,*120: 3-4,

D-Amato, M. E., Lunt, D. H., Carvalho, G. R. 1999. Microsatellite markers for the hake *Macruronus magellanicus* amplify other gadoid fish. *Molecular Ecology*, 8: 1086-1087.

Dalamu, T. K., Behera, A. B., Gaikwad, Swati Saxena, C., Bharadwaj, A. D. 2012. Morphological and molecular analysis define the genetic diversity of Asian bitter gourd (*Momordica charantia* L.). *Australian Journal of Crop Science*, 6(2): 261-267.

Da-Long, G., Zhang, J. P., Xue, Y. M., Hou, X. G. 2012. Isolation and characterization of 10 SSR markers of M*omordica charantia* (Cucurbitaceae). *American Journal of Botany*, 182–183.

Dayteg, C., Vos Post L., Lund R., Tuvesson S. 1998. Quick DNA extraction method for practical plant breeding programmes. Plant and Animal Genome VI. San Diego, Ca., USA, Jan 18–22, p. 39.

Dellaporta, S. L., Wood, J., Hicks, J. B. 1983. A plant DNA minipreparation: version II. *Plant Molecular Biology Reporter*, 1: 19-21.

Ellegren, E. H., Primmer, C. R., Sheldon, B. 1995. Microsatellite evolution: directionality or bias in locus selection? *Nature Genetics*, 11: 60-62.

Ender, A., Schwenk, K., Stdler, T., Streit, B., Schierwater, B. 1996. RAPD identification of microsatellites in *Daphnia*. *Molecular Ecology*, 5: 437-441.

FitzSimmons, N. N., Moritz, C., Moore, S. S. 1995. Conservation and dynamics of microsatellite loci over 300 million years of marine turtle evolution. *Molecular Biology and Evolution*, 12: 432-440.

Hamilton, M. B., Pincus, E. L., Di-Fiore, A., Fleischer, R. C. 1999. Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. *Biotechniques*, 27: 500-507.

Heubl, G. 2010. New aspects of DNA-based authentication of Chinese medicinal plants by molecular biological techniques. *Planta Medica*. 76(17): 1963-1974.

Ikeda, N., Bautista N. S., Yamada, T., Kamijima, O., Ishii, T. 2001. Ultrasimple DNA extraction method for marker-assisted selection using microsatellite markers in rice. *Plant Molecular Biology Reporter*,19: 27–32.

Jarne, P., Lagoda, P. J. L. 1996. Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution*, 11: 424-429.

Jing-Jing D., Ren-Chao Z., An-Jun T., Xue-Jun G., Wei W. 2013. Development and characterization of nine microsatellites for an endangered tree, *Pinus wangii* (Pinaceae). Applications in Plant Sciences, 1(2) 2013.

Jirapong, J., Tetsuya, K., Yoshiyuki, Y., Sachiyo, S. M., Kazuki, M., Kosuke, T., Satoru, K., Seigo, K., Masahiro, U., Yoshitaka, S., Kimiko, Y., Masaya, M., Hideshi, Y. 2012. A Simple Sequence Repeat- and Single-Nucleotide Polymorphism-Based Genetic Linkage Map of the Brown Planthopper, *Nilaparvata lugens*. *DNA Research*, 1–14.

Kandpal, R. P., Kandpal, G., Weissman, S. M. 1994. Construction of libraries enriched for sequence repeats and jumping clones, and hybridization selection for region-specific markers. *Proceedings of the National Academy of Sciences of the USA*, 91: 88-92.

Kantety, R. V., Rota, M. L., Matthews, D. E., Sorrells, M. E. 2002. Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Molecular Biology*, 48: 501-510.

Karagyozov, L., Kalcheva, I. D., Chapman, V. M. 1993. Construction of random small-insert genomic libraries highly enriched

for simple sequence repeats. *Nucleic Acids Research*, 21: 3911-3912.

Knapik, E. W., Goodman, A., Ekker, M. 1998. A microsatellite genetic linkage map for zebrafish (*Danio rerio*). *Nature Genetics*, 18: 338-343.

Kocher, T. D., Thomas, W. K., Meyer, A. 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences of the USA*, 86: 6196-6200.

Li, J. W., Zhou, S. S., Huang, F. 2012. Microsatellite markers and their application in genetic diversity research of Plasmodium spp. *Chinese Journal of Schistosomoasis Control*, 24(2): 232-4, 238.

Litt, M., Luty, J. A. 1989. A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American Journal of Human Genetics*, 44: 397-401.

Liu, B. H. 1997. *Statistical Genomics: Linkage, Mapping and QTL Analysis*. CRC Press, 359-362.

Lunt, D. H., Hutchinson, W. F., Carvalho, G. R., 1999. An efficient method for PCR-based identification of microsatellite arrays (PIMA). *Molecular Ecology*, 8: 893-894.

Maneeruttanarungroj, C., Pongsomboon, S., Wuthisuthimethavee, S., Klinbunga, S., Wilson, K.J., Swan, J., Li, Y., Whan, V., Chu, K.H., Li, C.P., Tong, J., Glenn, K., Rothschild, M., Jerry, D., Tassanakajon, A. 2006. Development of polymorphic expressed sequence tag-derived microsatellites for the extension of the genetic linkage map of the black tiger shrimp (*Penaeus monodon*). *Animal Genetics*, 37: 363- 368.

Manigbas, L., and Lucille, C. V. 2004. Microsatellite markers in hybridity tests to identify true hybrid of Sugarcane, *Philippine Journal of Crop Science*, 29(2): 23-32

Moore, S. S., Sargeant, L. L., King, T. J., Mattick, J. S., Georges, M., Hetzel, J. S. 1991. The conservation of dinucleotide microsatellites among mammalian genomes allows the use of heterologous PCR primer pairs in closely related species. *Genomics*, 10: 654-660.

Morin, P. A., Mahboubi, P., Wedel, S., Rogers, J. 1998. Rapid screening and comparison of human microsatellite markers in baboons: allele size is conserved, but allele number is not. *Genomics*, 53: 12-20.

Murray, M. G. and Thompson, W.F. 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research*, 8: 4321-4325.

Ostrander, E. A., Jong, P. M., Rine, J., Duyk, G. 1992. Construction of small-insert genomic DNA libraries highly enriched for microsatellite repeat sequences. *Proceedings of the National Academy of Sciences of the USA*, 89: 3419-3423.

Paetkau, D. 1999. Microsatellites obtained using strand extension: An enrichment protocol. *Biotechniques*, 26: 690-697.

Park, Y. J., Lee, J. K., Kim, N. S. 2009. Simple sequence repeat polymorphisms (SSRPs) for evaluation of molecular diversity and germplasm classification of minor crops. *Molecules*. 10; 14(11): 4546-4569.

Paterson, A. H., Brubaker, C. L., Wendel, J. F. 1993. A rapid method for extraction of Cotton (Gossypium spp) genomic DNA suitable for RFLP or PCR analysis. *Plant Mololecilar Biology Reporter*, 11:122-127.

Pérez, F., Ortiz, J., Zhinaula, M., Gonzabay, C., Calderón, J., Volckaert, F. A .M .J. 2005. Development of EST-SSR markers by data mining in three species of shrimp: *Litopenaeus vannamei*, *Litopenaeus stylirostris*, and *Trachypenaeus birdy*. *Marine Biotechnology*, 7: 554-569.

Prakash, C. 2007. Mining microsatellites in eukaryotic genomes. *Trends in biotechnology*, 490-498.

Primmer, C. R., Mller, A. P., Ellegren, H. 1996. A wide-ranging survey of cross-species

amplification in birds. *Molecular Ecology*, 5: 365-378.

Quentin, C., Martin P. A. C., Emma, T. S., Osmond X. M., Gifty N. A. H., Michael J. W., and Brenda D. W. 2009. Microsatellite discovery by deep sequencing of enriched genomic libraries. *BioTechniques*, 46:217-223.

Ramasamy, Y., Modhumita, G., Ramasamy, S., Krish, G. 2005. Cross-species amplification of eucalyptus SSR markers in Casuarinaceae. *Acta Botanica Croatica*, 64 (1): 115–120.

Rassmann, K., Schltterer, C., Tautz, D. 1991. Isolation of simplesequence loci for use in polymerase chain reaction-based DNA fingerprinting. *Electrophoresis*, 12: 113-118.

Rico, C., Rico, I., Hewitt, G. 1996. 470 million years of conservation of microsatellite loci among fish species. *Proceedings of Royal Society of London B: Biological Sciences*, 263, 549-557.

Roder, M. S., Korzun, V., Wendehake, K., Plaschke, J., Tixier, M. H., Leroy, P., Ganal, M. W. 1998. A microsatellite map of wheat. *Genetics*, 149: 2007–2023.

Rubinsztein, D. C., Amos, W., Leggo, J. 1995. Microsatellite evolution - Evidence for directionality and variation in rate between species. *Nature Genetics*, 10: 337-343.

Saghai-Maroof, M. A., Soliman, K. M., Jorgensen, R. A, Allard, R. W. 1984. Ribosomal DNA sepacer-length polymorphism in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proceedings of National Academy of Science, USA* 81, 8014-8019.

Sawaya, S., Bagshaw, A., Buschiazzo, E., Kumar, P., Chowdhury, S., 2013. Microsatellite Tandem Repeats Are Abundant in Human Promoters and Are Associated with Regulatory Elements. PLoS ONE 8(2): 54710.

Schltterer, C., Amos, B., Tautz, D. 1991. Conservation of polymorphic simple sequence loci in cetacean species. *Nature*, 354: 63-65.

Schltterer, C. and Tautz, D. 1992. Slippage synthesis of simple sequence DNA. *Nucleic Acids Research*, 20: 211-215.

Schuler, G. D., Boguski, M. S., Stewart, E. A. 1996. A gene map of the human genome. *Science*, 274: 540-546.

Scribner, K. T. and Pearce, J. M. 2000. Microsatellites: evolutionary and methodological background and empirical applications at individual, population, and phylogenetic levels. In: Molecular Methods in Ecology (ed. Baker A), pp. 235-271. Blackwell Science Limited, London.

Semagn, K., Bjørnstad, Å., Skinnes, H., Marøy, A. G., Tarkegne, T., William, M. 2006. Distribution of DArT, AFLP, and SSR markers in a genetic linkage map of a doubled haploid hexaploid wheat population. *Genome*, 49:545–555.

Serapion,, J., Kucuktas, H., Feng, J. N., Liu, Z. J., 2004. Bioinformatic mining of type I microsatellites from expressed sequence tags of channel catfish (*Ictalurus punctatus*). *Marine Biotechnology*, 6: 364-377.

Shu-Zhen Wang, Lei Pan, Kan Hu, Chan-You Chen, and Yi Ding, 2010. Development and characterization of polymorphic microsatellite markers in *Momordica charantia* (Cucurbitaceae), *American Journal of Botany*, 75–78.

Stolle, E., Kidner, J. H., Moritz, R. F. 2013. Patterns of evolutionary conservation of microsatellites (SSRs) suggest a faster rate of genome evolution in Hymenoptera than in Diptera. *Genome Biology Evolution*, 5(1): 151-162.

Takundwa, M., Nepolo, E., Khola ,M., Kandawa-Schulz, M.A., Cullis, A.C., Kunert, K., Jackson-Malete, J. J., Chiwona Karltun L., Chimwamurombe, P. M. 2012. Development and use of microsatellite markers in Marama bean. *African Crop Science Journal*, 20: 95-105.

Tautz, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research*, 17: 6463-6471.

Tessier, C., David, J., This, P., Boursiquat, J.M., Charrier, A., 1999. Optimization of the choice of molecular markers for varietal identification in *Vitis vinifera* L. *Theoretical Application Genetics*, 98: 171-177.

Thompson, D., Henry, R. J. 1995. Single step protocol for preparation of plant tissue for analysis by PCR. *BioTechniques*, 19:394-397.

Tong, J., Yu, X., Liao, X. 2005. Characterization of a highly conserved microsatellite marker with utility potentials in cyprinid fishes. *Journal of Applied Ichthyology*, 21: 232-235.

Toth, G., G. Spari, Z., Jurka, J. 2000. Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Research* ,967-981.

Von, P. R., Von, P. L., Dayteg, C., Nilsson, M., Forster, B. P. and Tuvesson, S. 2003. A high-throughput DNA extraction method for barley seed. *Euphytica*, 130: 255–260.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van, D. L., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., Zabeau, M., 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, 23: 4407-4414.

Wang, H. X., Li, F. H., Xiang, J. H. 2005. Polymorphic EST-SSR markers and their mode of inheritance in *Fenneropenaeus chinensis*. *Aquaculture*, 249: 107-114.

Weber, J. L., May, P. E. 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics*, 44: 388-396.

Wheeler, D. L., Barrett, T., Benson, D. A., Bryant, S. H., Canese, K., Chetvernin, V., Church, D. M., DiCuccio, M., Edgar, R., Federhen,, S., Feolo M., Geer, L.Y., Helmberg, W., Kapustin, Y., Khovayko, O., Landsman, D., Lipman, D. J., Madden, T. L., Maglott, D. R., Miller, V., Ostell, J., Pruitt, K. D., Schuler, G. D., Shumway, M., Sequeira, E., Sherry, S. T., Sirotkin, K., Souvorov, A., Starchenko, G., Tatusov, R. L., Tatusova, T.A., Wagner, L. & Yaschenko, E. 2008. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, 36 (Database issue):, 13-21.

Xiaolian, L.., Wei, L., Cong Z., Weimin W. and Zexia G. 2011. Isolation of New 40 Microsatellite Markers in Mandarin Fish (*Siniperca chuatsi*) . *International Journal of Molicular Science*, *12:* 4180-4189

Yue, G. H., Ho, M. Y., Orban, L., Komen, J. 2004. Microsatellites within genes and ESTs of common carp and their applicability in silver crucian carp. *Aquaculture*, 234: 85-98.

Zahra, N., Mehrshad, Z., Mohsen, M., Seyed, M. P., Seyed, H. M., Seyed M. K. 2012. Isolation and characterization of a first set of polymorphic microsatellite markers in saffron, C*rocus sativus* (Iridaceae). *American Journal of Botany*. 340–343.

Zalapa, J. E., Cuevas, H., Zhu, H., Steffan, S., Senalik, D., Zeldin, E., McCown, B., Harbut, R., Simon, P. 2012. Using next-generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences. *American Journal of Botany*, 99(2): 193-208.

Zane, L., Bargelloni, L. and Patarnello, T. 2002. Strategies for microsatellite isolation: a review. *Molecular Ecology*, 11: 1–16.

Zhivotovsky, L. A. and Feldman, M. W. 1995. Microsatellite variability and genetic distances. *Proceedings of the National Academy of Sciences of the USA*, 92, 11549-11552.