Molecular Ecology (2002) **11**, 1–16

INVITED REVIEW

# Strategies for microsatellite isolation: a review

L. ZANE,* L. BARGELLONI* and T. PATARNELLO*†

**Dipartimento di Biologia, Università di Padova Via Ugo Bassi 58/B, 35121, Padova, Italy, †Agripolis/Facoltà di Veterinaria, Università di Padova 35020, Legnaro (Pd), Italy*

**Abstract**

**In the last few years microsatellites have become one of the most popular molecular markers used with applications in many different fields. High polymorphism and the relative ease of scoring represent the two major features that make microsatellites of large interest for many genetic studies. The major drawback of microsatellites is that they need to be isolated *de novo* from species that are being examined for the first time. The aim of the present paper is to review the various methods of microsatellite isolation described in the literature with the purpose of providing useful guidelines in making appropriate choices among the large number of currently available options. In addition, we propose a fast and easy protocol which is a combination of different published methods.**

*Keywords*: enrichment, isolation, microsatellites, review, SSRs

*Received 4 June 2001; revision received 21 September 2001; accepted 21 September 2001*

## Microsatellites: the difficulty of isolating a powerful genetic marker

Microsatellites or simple sequence repeats (SSRs) are tandemly repeated motifs of 1–6 bases found in all prokaryotic and eukaryotic genomes analysed to date. They are present in both coding and noncoding regions and are usually characterized by a high degree of length polymorphism. The origin of such polymorphism is still under debate though it appears most likely to be due to slippage events during DNA replication (Schlötterer & 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 & Luty 1989; Tautz 1989; Weber & 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), but their applications span over different areas ranging from ancient and forensic DNA studies, to population genetics and conservation/management of biological resources (Jarne & Lagoda 1996).

Correspondence: Professor Tomaso Patarnello. Dipartimento di Biologia, Università di Padova Via Ugo Bassi 58/B, 35121, Padova, Italy. Fax: + 39 49 827 6209; E-mail: patarnel@civ.bio.unipd.it

Given their large applicability, there has been an extraordinary increase of interest in SSRs as indicated by the large number of papers published in the last few years that have involved the use of microsatellites (the word 'microsatellite' is found in nearly 8000 records when a search of the 'Current Contents' publication databases for the years 1995–2000 is carried out). The great popularity of SSRs is also demonstrated by the growing number of reports describing the isolation of these markers in many organisms (Box 1), and by the creation of *Molecular Ecology Notes* with its associated database.

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 due to the fact that microsatellites are usually found in noncoding regions where the nucleotide substitution rate is higher than in coding regions. Consequently, the strategy of designing 'universal primers' matching conserved sequences, which was very effective for mitochondrial DNA (Kocher *et al*. 1989), is more problematic for microsatellites. However, the presence of highly conserved flanking regions has been reported for some microsatellite loci in cetaceans (Schlötterer *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 (Ma).

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

**Box 1**

A survey of papers focusing on the isolation of microsatellite loci was carried out considering studies published in the specialized sections of *Molecular Ecology* (1999–March 2001) and *Animal Genetics* (1999–2000). A total of 302 reports were reviewed, recording the species studied, methods used, and the efficiency of isolation protocol(s) whenever possible. The obtained results reveal different trends in the use of microsatellite markers. The low number of new microsatellite reports in *Animal Genetics* appears to indicate that such markers are already commonly used in the genetic studies of domestic animals (Fig. 1). In contrast the increase in 'primer notes' observed in the recent issues of *Molecular Ecology* highlights the fact that microsatellite markers are being used in a rapidly increasing number of species, ranging from fungi to plant and animals (Fig. 2). With regard to isolation protocols, while most authors remain faithful to the traditional methods of library screening, a substantial fraction of papers (more than half in the latest issues of *Molecular Ecology*) describe the use of enhanced protocols ('enrichment methods', see main text). Such use seems to be biased with respect to the investigated species. For some taxonomic groups, microsatellite-enriched libraries are commonly employed, whereas in other taxa they are less frequently used (Fig. 3).

A closer examination of methods used in different taxa suggests that the choice of technique is not always related to the expected microsatellite frequency in the genome of the target species nor to the efficiency of the chosen protocol. For instance, while plant and avian genomes are both known to be poor sources of microsatellite markers (see Box 2), enriched libraries were chosen in the majority of the examined studies for plants only.

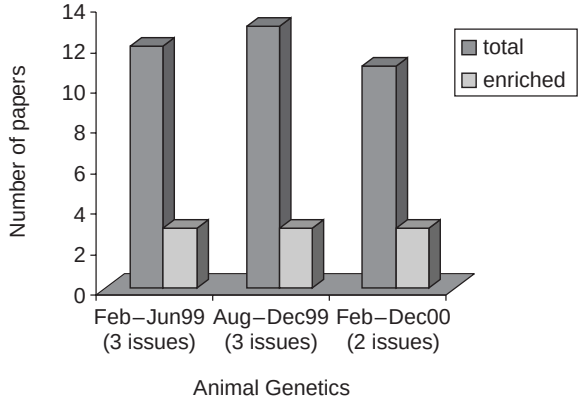


**Fig. 1** Total number of articles published in *Animal Genetics* reporting the isolation of microsatellites (dark grey bars), and number of papers where an 'enriched' protocol was used (light grey bars). Each category on the horizontal axis refers to issues published in the specified time interval.

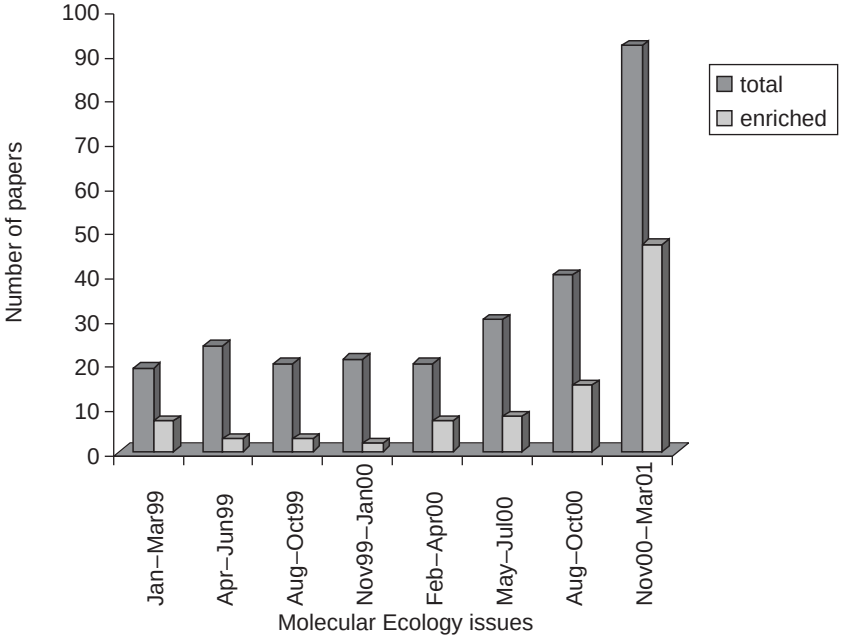


**Fig. 2** Total number of articles published as *Molecular Ecology* primer notes reporting the isolation of microsatellites (dark grey bars), and number of papers where an 'enriched' protocol was used (light grey bars). Each category on the horizontal axis refers to issues published in the specified time interval.

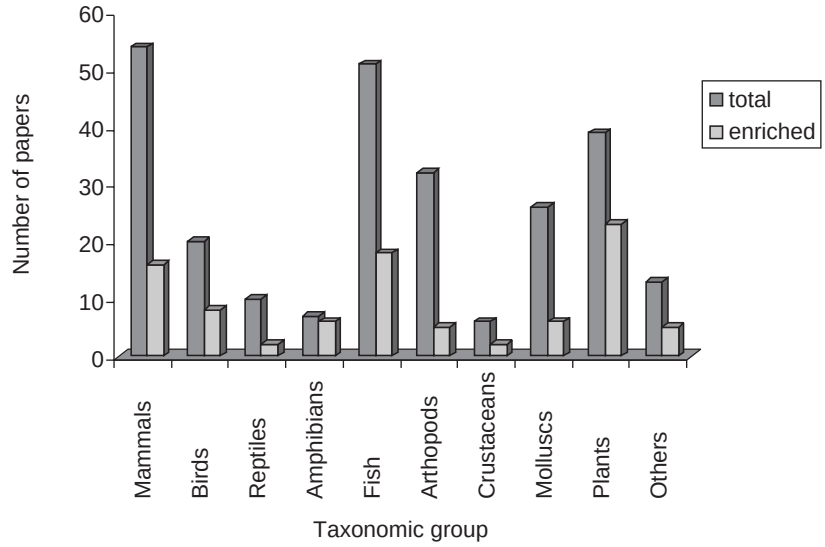


**Fig. 3** Total number of articles published as *Molecular Ecology* primer notes (1999–March 2001) reporting the isolation of microsatellites (dark grey bars), and number of papers where an 'enriched' protocol was used (light grey bars). Each category on the horizontal axis refers to different taxonomic groups.

guarantee that similar polymorphism will be found in related species especially when increasing the evolutionary distance (Rubinsztein *et al*. 1995; Morin *et al*. 1998).

Reports from 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 & 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% (Box 2). 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 is 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 & 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 (Zhivotovsky & Feldman 1995; Cooper *et al*. 1999) or when constructing a genetic map (Liu 1997).

A number of new protocols, overcoming these limitations, have appeared in the literature in the last few years. These methods often present only slight differences from one another and frequently have not been extensively tested. The aim of this review is to present the various methods of microsatellite isolation so far described with the purpose of providing useful guidelines in choosing the appropriate protocol among the large number of currently available options.

## Available methods for microsatellite isolation

Traditionally, microsatellite loci have been isolated from partial genomic libraries (selected for small insert size) of the species of interest (Box 3), screening several thousands of clones through colony hybridization with repeat-containing probes (Rassmann *et al*. 1991). As mentioned above, although relatively simple, especially for microsatellite-rich genomes, this approach can turn out to be extremely tedious and inefficient for species with low microsatellite frequencies. Therefore, several alternative strategies have been devised in order to reduce the time invested in microsatellite isolation and to significantly increase yield (Table 1). The popularity of some of these alternative methods is demonstrated by a survey of papers reporting microsatellite loci isolation, published since 1999 in *Molecular Ecology* and *Animal Genetics* (Box 1).

**Box 2**

Microsatellites are known to be ubiquitous in prokaryotic and eukaryotic genomes and present both in coding and noncoding regions. Distribution of microsatellites, however, is not homogeneous within a single genome, because of different constraints of coding vs. noncoding sequences, historical processes (Arcot *et al*. 1995; Wilder & Hollocher 2001) and the possible different functional roles of different repeats (Valle 1993). Microsatellite genomic frequency also varies across taxa, in terms of both absolute numbers of microsatellite loci and repeat preference (Hancock 1999).

Details on these matters are unfortunately scarce, and very few studies have attempted to address the question of repeat preference and genomic frequency of microsatellites in a broad taxonomic perspective. The only exceptions are reports on model organisms with large sequence coverage, like man, mouse, and fruitfly.

To estimate the expected 'success' rate for microsatellite isolation with traditional cloning strategies we have made use of a recently published report on microsatellite frequency in different eukaryotic genomes (Toth *et al*. 2000). At the same time, we have reviewed experimental data from 267 primer notes published in *Molecular Ecology* to obtain empirical estimates of microsatellite abundance.

*Estimates based on sequence data*

Toth *et al*. (2000) used a computer-based approach to analyse microsatellite frequency in a variety of species ranging from primates to fungi (Table B1). This study

**Table B1** Expected number of microsatellites per megabase (Mbp) of DNA; recalculated from Toth *et al*. (2000)

| Taxonomic group | Number of species* | Sequence cumulated length (Mbp) | Expected microsatellites for megabase of sequence† | Best score motifs (expected number) | | |
|---|---|---|---|---|---|---|
| | | | | dinucleotides | trinucleotides | tetranucleotides |
| Primates | 64 | 160.08 | 223 | AC (40) | AAT (8) | AAAT (17) |
| | | | | AT (15) | AAC(6) | AAAC (11) |
| | | | | AG (14) | AGG (2) | AAAG (9) |
| Rodentia | 81 | 21.26 | 429 | AC (112) | AGG (10) | AAAC (18) |
| | | | | AG (48) | AGC (10) | AGAT (14) |
| | | | | AT (13) | AAC (8) | AAAT (13) |
| Mammalia | 203 | 3.61 | 238 | AC (50) | AGC (10) | AAAT (10) |
| | | | | AG (40) | AGG (6) | AAAG (6) |
| | | | | AT (6) | CCG (6) | ACAG (2) |
| Vertebrata | 353 | 5.47 | 232 | AC (50) | AAT (16) | AGAT (10) |
| | | | | AT (22) | AGG (10) | AAAT (5) |
| | | | | AG (14) | AGC (9) | ACAG (5) |
| Arthropoda | 586 | 28.76 | 245 | AC (46) | AGC (22) | ACAT (5) |
| | | | | AT (18) | AAC (11) | AAAT (4) |
| | | | | AG (13) | AAT (6) | AAAC (2) |
| Embryophyta | 1313 | 48.17 | 154 | AT (29) | AAG (18) | AAAT (3) |
| | | | | AG (16) | AAC (7) | AAAG (2) |
| | | | | AC (5) | ATC (7) | AAAC (1) |
| Fungi | 1164 | 17.78 | 92 | AT (7) | AAT (6) | AAAT (3) |
| | | | | AC (3) | AAC (5) | AAAG (1) |
| | | | | AG (3) | AAG (4) | AAAC (1) |
| *Saccharomyces cerevisiae* | 1 | 15.18 | 99 | AT (24) | AAT (7) | AAAT (2) |
| | | | | AC (2) | AAC (7) | ACAT (1) |
| | | | | AG (1) | AAG (7) | AAAG (1) |
| *Caernorhabditis elegans* | 1 | 81.55 | 88 | AG (10) | AAG (6) | AAAT (3) |
| | | | | AC (8) | AAT (4) | AGGC (1) |
| | | | | AT (7) | ATC (3) | ACCT (1) |

*For some taxonomic groups the largest part of the sequence analysis refers to only a few organisms. The list of 'dominant' species is given below for each taxonomic group, followed (in parenthesis) by the percentage of total sequence information for that species. Primates: *Homo sapiens* (99.43%); Rodentia: *Mus musculus* (73.71%), *Rattus norvegicus* (18.25%); Mammalia: *Bos taurus* (27.26%), *Sus scrofa* (20.72%), *Oryctolagus cuniculus* (19.09%), *Ovis aries* (10.59%), *Canis familiaris* (6.62%); Vertebrata: *Gallus gallus* (32.2%), *Fugu rubripes* (17.76%), *Xenopus laevis* (12.15%); Arthropoda: *Drosophila melanogaster* (84.27%), *Drosophila* sp. (7.93%); Embryophyta: *Arabidopsis thaliana* (79.18%); Fungi: *Schizosaccharomyces pombe* (48.41%).

†Microsatellites are defined as stretch of tandem repeats (2–6 nucleotides) longer than 12 bp; the short cut-off length has been used to avoid loss of loci due to single base substitution resulting in imperfect microsatellites (see Toth *et al*. 2000 for details).

provided several important insights, but data set limitations must be kept in mind. Indeed, the taxonomic categories considered in the study do not represent the existing biological diversity in terms of both the number of taxa and sequence representation of each group (for instance, 99.43% of primate sequences are from the human genome, and 84.27% of arthropod sequences are from the fruitfly genome).

As reported by Toth *et al*. (2000), the total occurrence of microsatellites expressed as the total length of 1–6 base pairs (bp) repeats per megabase (Mbp) of DNA, varies across taxonomic groups, ranging from 13 889 in rodents to 2139 in the flatworm *Caernorhabditis elegans*. This would mean that, considering the estimated average microsatellite length (available as supplementary material of Toth *et al*. 2000 at http://falco.elte.hu/ssr), and excluding single-base repeats (generally not used as microsatellites loci) about 430 microsatellites are expected to be found for every Mbp of the rodent genomes, whereas 88 are expected in *C. elegans*. Considering that 1 Mbp corresponds to 2500 nonoverlapping clones with an average insert size of 400 bp, traditional methods for microsatellite isolation could yield as much as 17% of positive clones in rodents and 3.5% in *C. elegans*. However, when focusing on specific repeats, the number of expected microsatellites is considerably lower. In fact, the expected frequency of any tri- or tetranucleotide repeat is below 1% of positive clones in all analysed taxa.

Finally it should be noted that repeat preference is very different from species to species, and some of the probes commonly used in screening detect motifs that are largely underrepresented in many organisms. Given this uncertainty, a wise approach could be to use multiple probes in isolating microsatellites (see for example the Estoup and Turgeon protocol at the web site http://www.inapg.inra.fr/dsa/microsat/microsat.htm).

*Empirical success rate of microsatellite isolation*

From 267 primer notes published from 1999 to March 2001, we extracted 170 notes that used traditional isolation methods. We calculated the ratio of positive clones/total number of screened clones for all the entries that provided such information.

**Table B2** Percentage of positive clones from traditional isolation protocols (source: *Molecular Ecology* primer notes 1999– March 2001)

| Taxonomic group | Number of entries | Percentage of positive clones | |
|---|---|---|---|
| | | Average | Min-Max |
| Amphibians | 1 | 0.4 | — |
| Arthropods | 21 | 2 | 0.04–12 |
| Crustaceans | 3 | 2.4 | 0.24–6 |
| Mammals | 27 | 1.67 | 0.13–4.5 |
| Molluscs | 16 | 1.96 | 0.1–6.4 |
| Birds | 9 | 0.46 | 0.025–1.7 |
| Fish | 16 | 3.1 | 0.066–8.92 |
| Reptiles | 5 | 1.4 | 0.2–4 |
| Plants | 16 | 2.3 | 0.059–5.8 |

The general trend seems to be that a percentage of positive clones close to 2–3% should be expected for many taxa (Table B2), with the noticeable exception of birds, which seem to have a lower frequency of microsatellites.

It should be underlined that a large amount of the positive clones are discarded during the isolation–characterization process. This happens because of the lack of suitable sequence for primer design, or absence of the expected repeat, or because of unreliable amplification. For these reasons the percentage of discarded positives can easily be in the order of 50%. A striking difference in the percentage of positives was recorded among taxa, and every group has at least one species with a percentage lower than 0.3% of positive clones. Though some of the lowest percentages are associated with studies searching for tri- and tetranucleotides or with records from birds, many reports found a percentage of positives close to 0.2–0.3% when looking for dinucleotide repeats in a large variety of taxa.

In summary, taking into account the uncertainty associated with the screening efficiency and repeat representation, traditional methods for microsatellite isolation can be prone to a low return for such a significant effort.

To avoid library construction and screening, some authors have proposed modifications of the randomly amplified polymorphic DNA (RAPD, Williams *et al*. 1990) approach for the amplification of unknown microsatellites, by either using repeat-anchored random primers (Wu *et al*. 1994) or using RAPD primers and subsequent Southern hybridization of polymerase chain reaction (PCR) bands with microsatellite probes (Cifarelli *et al*. 1995; Richardson *et al*. 1995). Although not useful for single-locus analyses as no information on microsatellite flanking regions is obtained, these methods inspired alternative strategies for the identification of single microsatellite loci. Based on the observed abundance of repeat regions in RAPD amplicons, isolation of microsatellite regions is achieved simply by

means of Southern hybridization of RAPD profiles with repeat-containing probes, followed by the selective cloning of positive bands (Ender *et al*. 1996), or through the cloning of all the RAPD products and the screening of arrayed clones (Lunt *et al*. 1999). Other 'nonlibrary' PCR-based strategies rely on the use of repeat-anchored primers to isolate and then sequence one (Fisher *et al*. 1996) or both regions (Lench *et al*. 1996; Cooper *et al*. 1997) flanking microsatellite repeats.

**Table 1** Library cost, time investment, and yield compared among the different protocols of microsatellite isolation

| Protocol | Protocol set up* (US dollars) | Library*† (US dollars) | Time† | Yield |
|---|---|---|---|---|
| Traditional | 2000–4000 | < 400 | 1 month | Low |
| RAPD based | 1000 | < 100 | 1 week | Variable |
| Primer extension | 1000–4000 | < 400 | 2 weeks | Medium/High |
| Selective hybridization | 1000–4000 | < 400 | 1–2 weeks | Medium/High |
| Private companies | None | 5000–10 000 | None | High |

*Costs may vary between countries and reagent suppliers.
†Average cost and time to obtain positive clones, excluding sequencing.

**Box 3**

Traditionally, microsatellite loci have been isolated starting from a partial genomic library of the target species (Fig. 4). 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 latter 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.

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 (Ender *et al*. 1996; D'Amato *et al*. 1999) take advantage of the fact that RAPD fragments seem to contain microsatellite repeats more frequently than random genomic clones (Cifarelli *et al*. 1995).

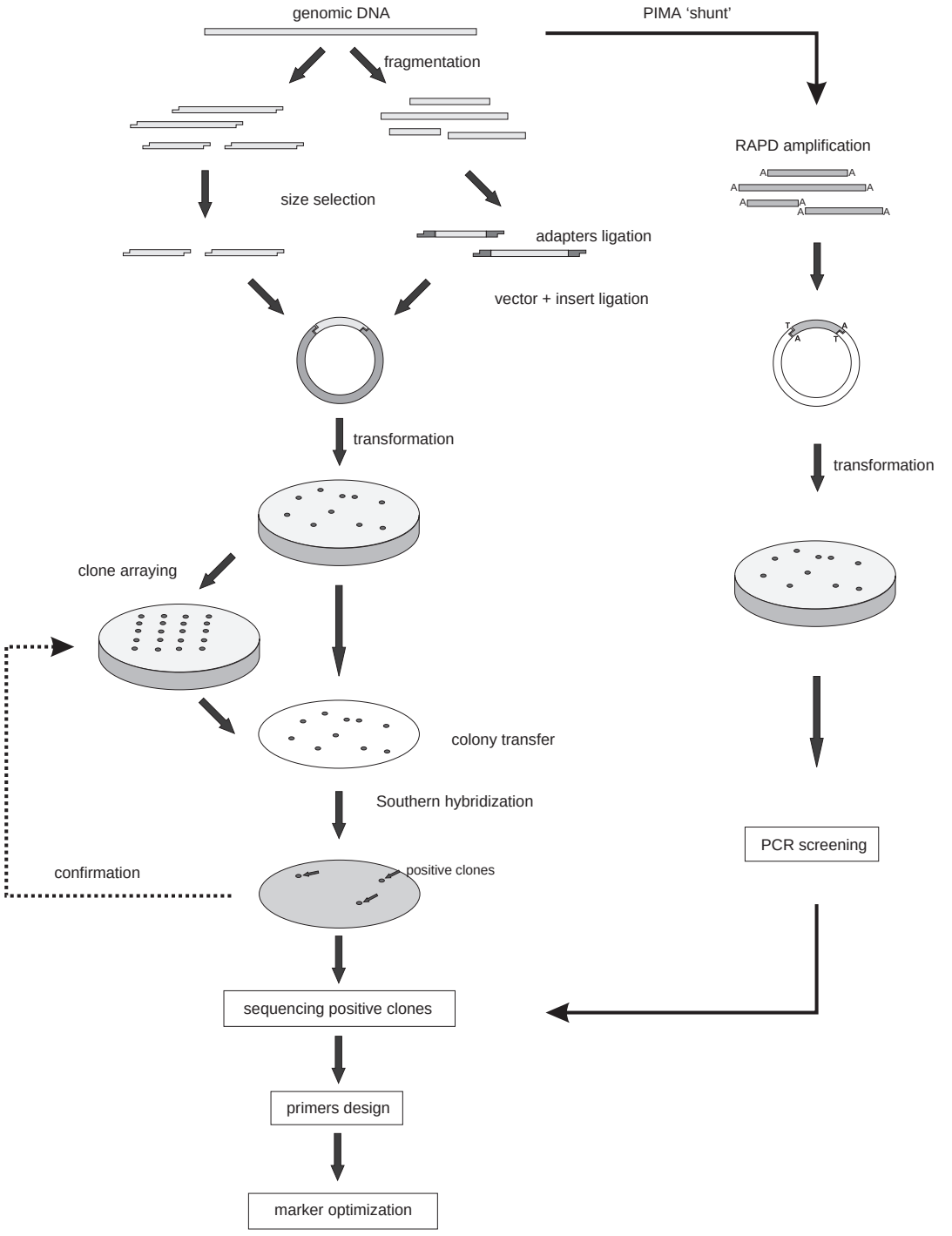


**Fig. 4** Schematic representation of 'traditional' methods for microsatellites isolation, and the alternative PIMA approach (see text for details).

While all these methods provide, if successful, a quick alternative to laborious and time-consuming library screening, their use has not been that frequent. In fact, taken together these protocols account for less than 2% of *Molecular Ecology* primer notes published to date.

A different strategy, based on primer extension, has been proposed for the production of libraries enriched in microsatellite loci (Box 4). Two papers originally described this method which was reported to be very efficient for the enrichment of AC repeats, yielding from 40 to 50%

(Ostrander *et al*. 1992) up to 100% of positive clones (Paetkau 1999). These protocols involve a rather high number of steps (Box 4), which might explain their limited application (at least in *Molecular Ecology* primer notes). To date, three primer notes have reported successful isolation by using the method of Ostrander *et al*. (1992) (one noticeably from a bird), whereas two have successfully used the Paetkau (1999) protocol. Two additional papers have employed a very similar enrichment method (Takahashi *et al*. 1996).

Many of these experiments concerned the isolation of dinucleotide repeat microsatellites, and it is unclear whether the 'primer-extension' approach is effective also for tri- and tetranucleotide. The 'Ostrander' protocol has not been tested for tri- or tetranucleotide repeat enrichment, whereas the 'Paetkau' protocol produced 0–25% positive clones when using a tetranucleotide repeat primer in the extension step. The latter protocol has been reported, in the case of tetranucleotide enrichments, to produce multiple copies of the same clone (Paetkau 1999), which might represent a problem when large numbers of microsatellites are needed. It is worth noting that both

**Box 4**

Two protocols have been proposed that 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 (Fig. 5a) 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 (Fig. 5b).

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 superinfected with M13 helper phage. Because of the particular genotype of the bacterial host (*dut*$^-$ *ung*$^-$), superinfection 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

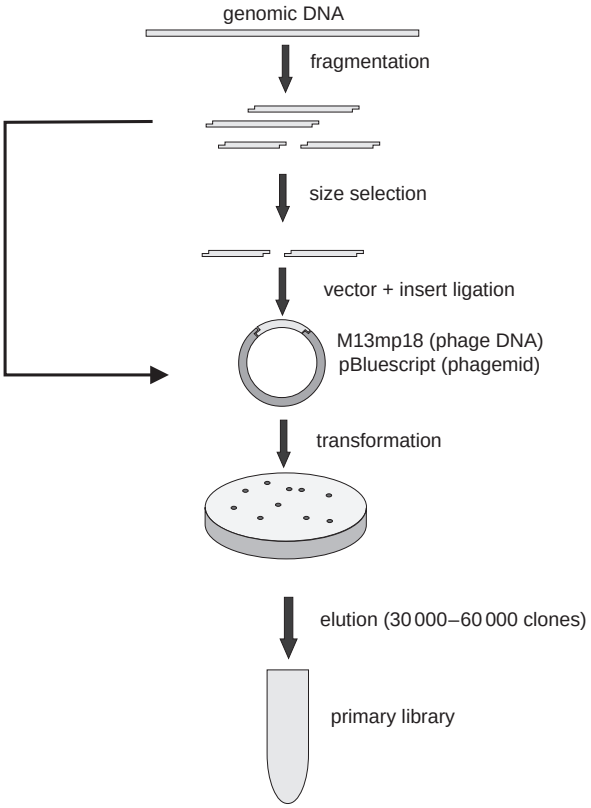


**Fig. 5a** Primer extension enrichment protocols. Schematic representation of the primary library construction.

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.

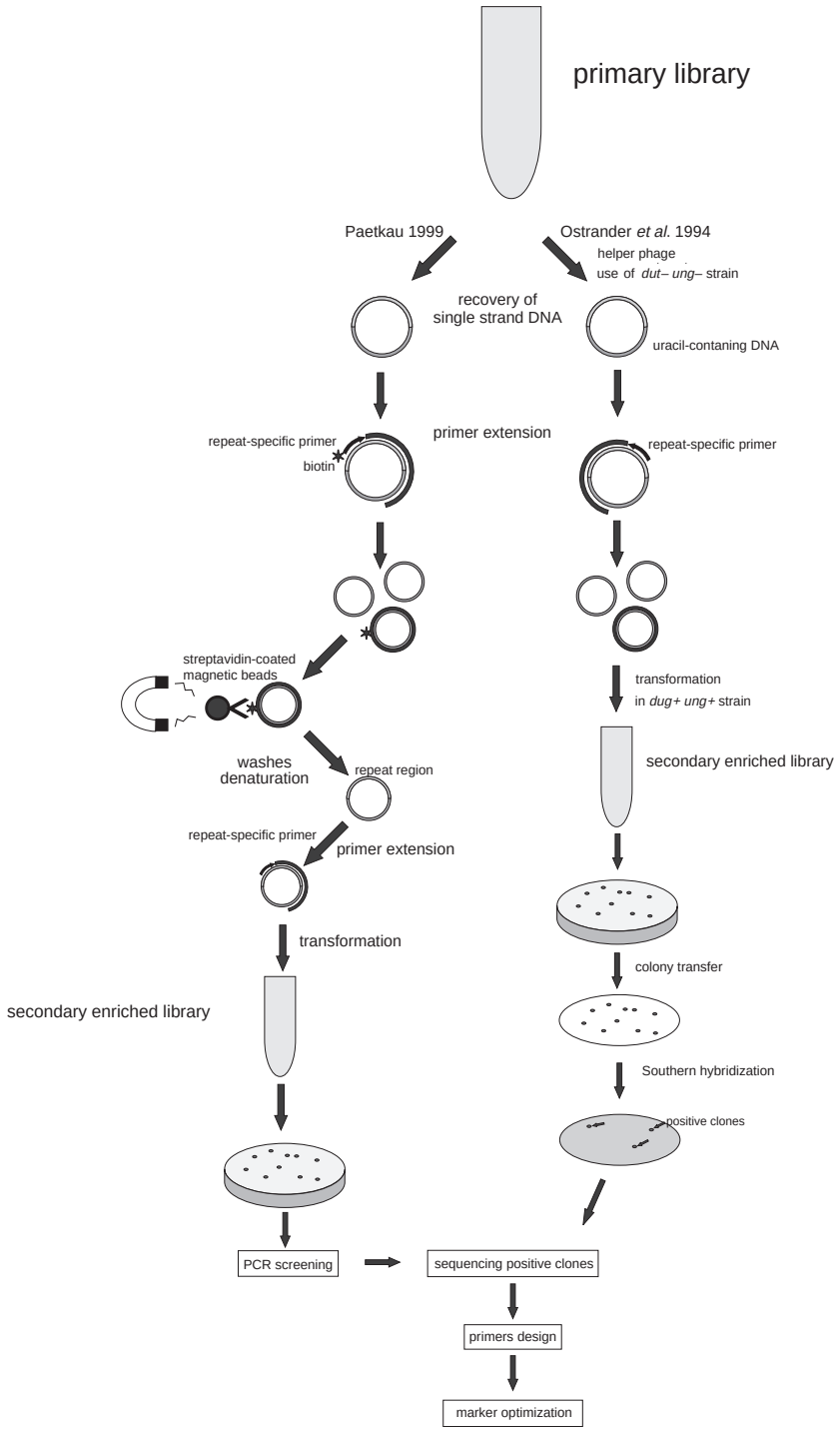


**Fig. 5b** Primer extension enrichment protocols. Schematic representation of protocols from Paetkau *et al*. 1999 (left) and Ostrander *et al*. 1994 (right).

primer-extension protocols involve the production of a primary library in order to obtain a pool of single-strand circular DNA molecules for subsequent enrichment (Box 4). In this step, for practical reasons, only a limited portion of the investigated genome is cloned, and so the population of inserts undergoes a severe bottleneck that results in loss of rare repeat motifs. With 60 000 clones in the primary library (Ostrander *et al*. 1992), in the case of a specific repeat motif with genomic frequency lower than 1% (Box 2), only 600 loci (containing the desired repeat motif) will

be represented in the enriched library. Thus, even if the enriched library is composed of thousands of clones, only 600 different loci can be (at best) isolated. A significant amount of redundancy might therefore affect the above protocols.

A further class of isolation methods is based on selective hybridization (Box 5). Selective hybridization protocols appear to be extremely popular being used in over 25% of all reviewed primer notes and 70% of those employing enrichment procedures. The basic protocol as proposed by Karagyozov *et al*. (1993), Armour *et al*. (1994), Kijas *et al*. (1994), is relatively straightforward, although several modifications have been independently suggested by various authors in an attempt to further optimize crucial steps or to remove unnecessary procedures (see for instance the Travis Glenn Web page at http://129.252.89.20/Msats/Microsatellites.html). The most frequently quoted protocols in *Molecular Ecology* primer notes are understandably those presented in the earliest papers on the topic (14 papers were based on the Armour *et al*. 1994 protocol, with seven more citations on its modification by Edwards *et al*. 1996; 10 papers were based on Kijas *et al*. 1994; four papers were based on Karagyozov *et al*. 1993, and four on Kandpal *et al*. 1994). In these studies, enrichment efficiency ranged from 20% to 90%, in a large variety of taxa, from plants to vertebrates, using di-, tri-, and tetranucleotide probes. These protocols appear therefore to be efficient and widely applicable, and if working with microsatellite-rich organisms and dinucleotide probes, enrichment may even be so efficient as to allow microsatellite identification by directly sequencing random recombinant clones alone. Although relatively simple all the selective hybridization methods can require some time in order to have the entire procedure up and running. In our own experience, protocol set up may take a few months as confirmed by recent discussions on the 'evoldir' and 'microsat' mailing lists (addresses and copies of the messages are available from the authors). The start-up investment in time and money (Table 1) may well be worthwhile for a group planning to work on several different species or on a large number of loci. Otherwise, it might be less expensive to seek help from other laboratories or commercial companies. Enriched microsatellite libraries for the species of interest can be purchased from an increasing number of suppliers, both commercial and academic (current prices are within the ranges of 5000–10 000 US dollars).

## Selective hybridization protocols for microsatellites loci isolation: an overview

As in traditional methods, the first step is DNA fragmentation followed by vector or adaptor ligation (Box 5). This step is important because low yield of ligated inserts and formation of concatamers can limit the steps that follow.

The alternative options available are basically similar to those already described in Box 3 for nonenriched protocols. DNA is fragmented either by sonication or by digestion with restriction enzymes. The length of DNA fragments produced by sonication (Karagyozov *et al*. 1993; Kandpal *et al*. 1994) is less dependent on genomic nucleotide composition, but requires an additional step to obtain blunt-end fragments. This may be achieved by either filling overhangs with T4 DNA polymerase or removing them with mung bean nuclease. On the other hand, when using restriction enzymes, the average fragment length depends on genome base composition and endonuclease recognition sequence. Moreover, differences in nucleotide composition within genomes might determine unequal sampling of genomic regions. This problem seems to be negligible as several protocols ignore it (Armour *et al*. 1994; Kandpal

**Box 5**

A very simple strategy for microsatellite isolation using selective hybridization can be outlined based on several reports that have been published in the last 10 years (Karagyozov *et al*. 1993; Hamilton *et al*. 1999).

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 (Fig. 6). 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 (Karagyozov *et al*. 1993; Armour *et al*. 1994) or 5′ biotinylated and bound to streptavidin coated beads (Kandpal *et al*. 1994; Kijas *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.

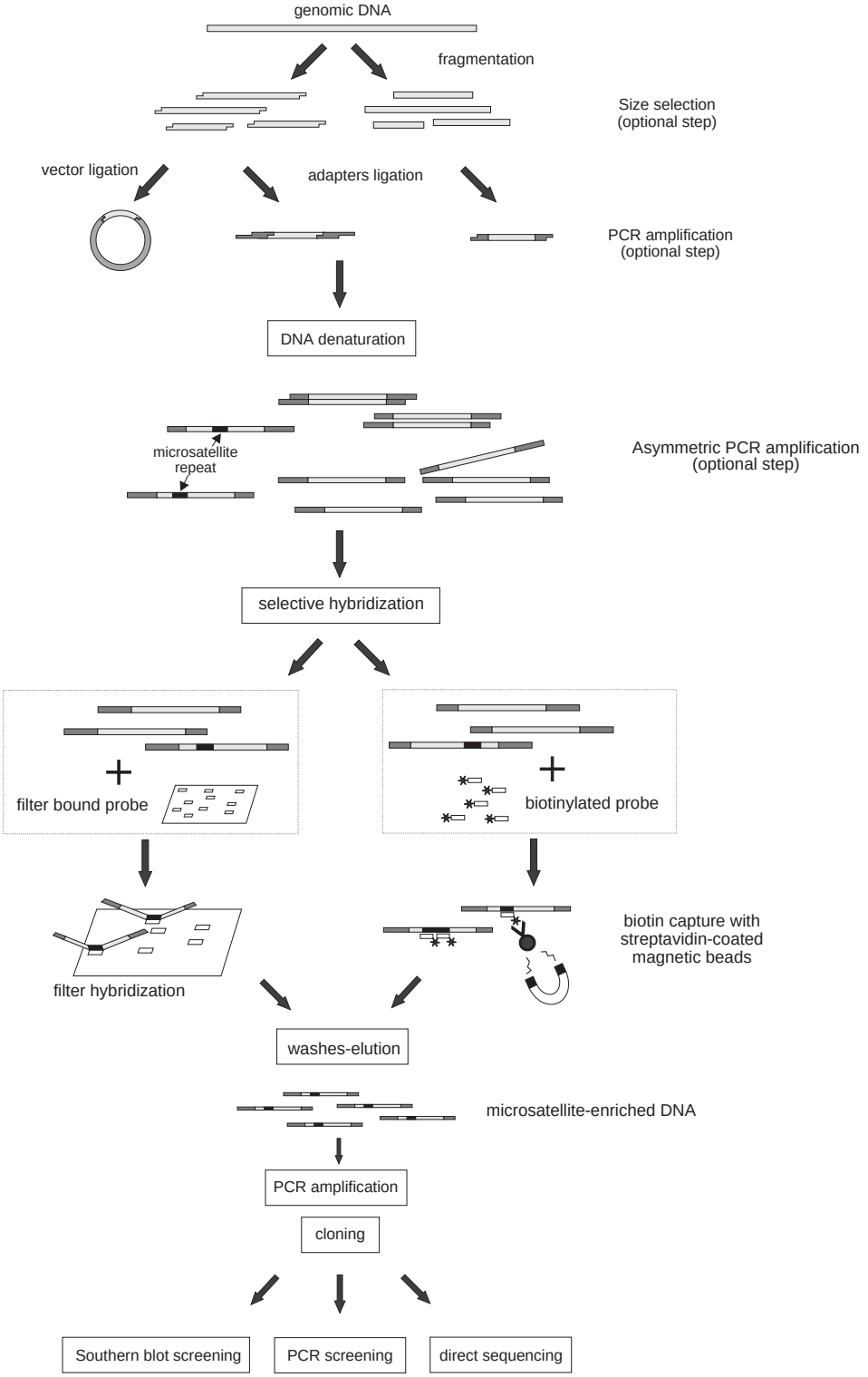


**Fig. 6** Schematic representation of selective hybridization protocols.

**Table 2** List of published adaptors that allow PCR amplification of fragmented DNA

| Adaptors sequence (5′–3′)* | Ligated termini | Restriction site for cloning | Reference |
|---|---|---|---|
| CTCTTGCTTGAATTCGGACTA<br>pTAGTCCGAATTCAAGCAAGAGCACA | blunt | *Eco*RI | Karagyozov *et al*. (1993) |
| CTCTTGCTTACGCGTGGACTA<br>pTAGTCCACGCGTAAGCAAGAGCACA | blunt | *Mlu*I | Edwards *et al*. (1996) |
| CGTAGTACTCGTGCGAATTCTGC<br>pGATCGCAGAATTCGCACGAGTACTAC | *Mbo*I | *Eco*RI | Kandpal *et al*. (1994) |
| CTAAGGCCTTGCTAGCAGAAGC<br>pGCTTCTGCTAGCAAGGCCTTAGAAAA | blunt | *Stu*I, *Nhe*I† | Hamilton *et al*. (1999) |
| GGCCAGAGACCCCAAGCTTCG<br>pGATCCGAAGCTTGGGGTCTCTGGCC | *Sau*3AI | (TA cloning) | Refseth *et al*. (1997) |
| GCGGTACCCGGGAAGCTTGG<br>pGATCCCAAGCTTCCCGGGTACCGC | *Mbo*I | (*Mbo*I) | Armour *et al*. (1994) |
| CGGAATTCTGGACTCAGTGCC<br>AATTGGCACTGAGTCCAGAATTCCG | *Tsp*509I | *Eco*RI | Tenzer *et al*. (1999) |

*p indicates a phosphorylated end.
†Adaptor concatamerization creates a *Xmn*I site.

*et al*. 1994; Kijas *et al*. 1994; Refseth *et al*. 1997). Other authors, however, have proposed to overcome this limitation by using multiple restriction enzymes to digest genomic DNA (Hamilton *et al*. 1999). In this case, multiple digestions can be carried out either simultaneously, which results in a smaller average size of fragments, or performed separately and then pooled together thereby producing longer fragments.

The latter option is useful in obtaining fragments in the range of 200–1000 bp, ideal for successful cloning and in recovering enough flanking regions to design primers for the amplification of individual microsatellites. However, if multiple digestions are not performed with blunt-end enzymes, the products need to be treated with T4 DNA ligase or mung bean nuclease to produce blunt termini for subsequent blunt-end ligation. At the same time, the use of multiple restriction enzymes can increase the chance of recovering multicopy sequences (e.g. satellite DNA), which can be detected as faint bands on a smear when digested DNA is subjected to agarose gel electrophoresis. Cloning these sequences can result in a library containing a large fraction of clones with the same insert, thus reducing the yield of useful loci. This problem can be by-passed by selecting a different set of restriction enzymes.

The size range of digested fragments can be controlled by adding a size selection step. DNA is separated according to its dimension by agarose gel electrophoresis, and fragments of the desired size are extracted from the gel and purified. Size selection can be performed after the digestion step (Kijas *et al*. 1994), or after the ligation step (Kandpal *et al*. 1994). In this latter case size selection is useful in removing free linkers, if needed.

After a reliable DNA fragmentation is obtained, the successive ligation step depends on the nature of the termini created on fragmented genomic DNA. As in any ligation, optimal experimental conditions should be found in order to maximize efficiency and to minimize unwanted concatamerization (Sambrook *et al*. 1989). To this end, many possible variants have been described, ranging from cohesive-end ligation into a dephosphorylated plasmid (Kijas *et al*. 1994) to linker-mediated ligation of blunt-ended sonicated DNA into a dephosphorylated λgt10 vector (Kandpal *et al*. 1994).

In other protocols ligation involves adaptors that are long enough to allow PCR amplification using primers designed for the adaptor sequence (Table 2). It should be recalled here that if adaptors are not phosphorylated (as in the case of unmodified synthetic oligonucleotides), or if the inserts have been dephosphorylated to prevent concatamer formation, insert-adaptor ligation results in DNA molecules carrying a nick in one of the two strands. This nick has to be filled before proceeding, because the following step involves DNA denaturation, which would result in loss of adaptor sequence. Nick repair can be easily obtained by extending the nicked strand with *Taq* DNA polymerase, in a PCR step in which the primers are omitted (Travis Glenn http://129.252.89.20/Msats/Microsatellites.html). PCR amplification is also a convenient way to obtain a sufficient amount of DNA for selective hybridization, when the starting amount of material is limited. In this case however, care must be taken to avoid overamplification of digested genomic DNA, which may lead to unequal representation of genomic fragments.

As stated above, before proceeding with selective hybridization, it is possible to size select genomic DNA, after ligation or after ligation-amplification. However, given that small fragments are expected to be preferentially lost

during selective hybridization, as they are likely to contain no (or only few) sequences that hybridize with selective probe(s), this step is often omitted.

Selective hybridization is performed by using an oligonucleotide containing several tandem repeats of the motif to be enriched as a probe (Box 5). The probe can be cross-linked to a nylon membrane or can be biotinylated at the 5′ end, so that DNA hybridized with the probe can be selectively removed by using streptavidin-coated paramagnetic beads. The use of a biotinylated probe is generally preferable because in the liquid medium the probe is fully available for hybridization. In contrast, the nylon bound probe is partially cross-linked to the membrane, and therefore hybridizes less efficiently with the target DNA.

Although different probe length, different hybridization and washing conditions are reported in the literature, the effect of these differences on microsatellite enrichment efficiency has not been extensively investigated. To our knowledge only one study has standardized temperatures for stringency washes (Kandpal *et al*. 1994).

Both for nylon bound and for biotinylated probes a further optimization involves the use of multiple probes in the hybridization step (Edwards *et al*. 1996; Travis Glenn http://129.252.89.20/Msats/Microsatellites.html), which seems to increase the overall enrichment efficacy.

After selective hybridization, captured fragments are recovered by PCR and cloned using standard methods. Finally recombinant clones are directly sequenced or Southern blotted and probed. Alternatively, PCR screening of recombinants seems to be a good approach for mildly enriched libraries (Waldbieser 1995). This approach involves two PCR reactions for every clone, using one primer for the vector and a second repeat-containing oligonucleotide.

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

Here we present our own contribution to previously published protocols.

We have tested this procedure in different organisms such as birds (*Passera lagia*), fish (*Sparus aurata* and *Lophius americanus*), crustacean (*Meganyctiphanes norvegica*) and red coral (*Corallium rubrum*). The percentage of clones containing dinucleotide repeats varied from a minimum of 50% (*Passera lagia*) to a maximum of 95% (*Sparus aurata*). These results are indicative of a high microsatellite isolation efficiency though the information concerning the frequency of polymorphic loci among the positive ones is currently under study in the different organisms cited above.

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 procedure (AFLP, Vos *et al*. 1995). DNA is simultaneously digested with *Mse*I and ligated to *Mse*I AFLP adaptor (5′-TACTCAGGACTCAT-3′/5′-GACGATGAGTCCTGAG-3′) using the following conditions: 25–250 ng of genomic DNA, buffer OnePhorAll 1X (Pharmacia), DTT 5 mM, BSA 50 μg/mL, adaptor 1 μM, ATP 200 μM, 2.5 units of *Mse*I (New England Biolabs), and 1 unit of T4 DNA ligase (Amersham-Pharmacia), in a total volume of 25 μL. The reaction is then incubated for 3 h at 37 °C.

AFLP adaptors have two important features ensuring high efficiency: (i) adaptors are not phosphorylated after being synthesized, and this prevents self ligation; and (ii) their sequences are designed so that ligation of the adaptor with digested DNA does not restore a *Mse*I site, thus allowing digestion and ligation to be performed simultaneously.

The digestion-ligation mixture is diluted (1:10), and directly amplified in a total volume of 20 μL with AFLP adaptor-specific primers (5′-GATGAGTCCTGAGTAAN-3′: hereafter referred to as *Mse*I-N). PCR conditions are (Vos *et al*. 1995): *Taq* DNA polymerase buffer 1× (Promega), $MgCl_2$ 1.5 mM, primer *Mse*I-N 120 ng, dNTPs 200 μM each, 0.4 units of *Taq* DNA polymerase (Promega) and 5 μL of a 1/10 dilution of digested-ligated DNA. The reaction is incubated in a GeneAmp PCR system 9700 (Perkin Elmer) set to 94 °C 30 s, 53 °C 1 min, 72 °C 1 min, for 14–26 cycles (see below).

Hot start PCR of DNA is avoided in order to allow the *Taq* DNA polymerase to fill the nicks present on the ligated DNA (see previous section), during the first ramping step.

In the AFLP protocol a similar amplification (the so called ‘preselective’ amplification) is performed by using a primer with a ‘selective’ nucleotide at the 3′ end that matches the first nucleotide beyond the original restriction site. In our protocol the amplification is performed by mixing primers carrying all four possible ‘selective’ bases (*Mse*I-N), thus allowing amplification of all fragments flanked by *Mse*I sites, providing only that they have an appropriate size for PCR. This procedure offers an important advantage: in the case of undesired bands appearing in the PCR amplification it is possible to go back just one step and perform a new PCR using an optimized combination of the four primers, thereby getting rid of the unwanted bands. This is extremely important because in some organisms amplification of digested–ligated DNA often generates one or more discrete bands that probably represent multicopy sequences in the original genome. These bands can be over-represented in the final PCR product, and they tend to be carried over during enrichment, especially if they cross hybridize with the biotinylated probe, accounting for a significant fraction of the obtained recombinant clones.

The number of cycles in the PCR amplification needs to

be optimized because over-amplification was found to change the average size of amplified fragments. To reduce, at least partly, the problem of biased amplification it is recommended that parallel PCR amplifications are progressively performed increasing the number of cycles (14–17–20–23–26 cycles). PCR conditions producing a visible product on agarose gel (in the form of a smear) are considered optimal and are selected for further use. PCR products, from the organisms tested to date, are always larger than 200 bp, thus eliminating the need for size selection.

PCR amplification under optimal conditions is replicated to obtain several hundred nanograms of amplified DNA. Unless the yield of the DNA amplification is very poor, sample concentration is not recommended because it can result in loss or low recovery of DNA.

DNA is then hybridized according to the Travis Glenn protocol (http://129.252.89.20/Msats/Microsatellites.html), but with a biotinylated $(AC)_{17}$ probe. In detail, 250–500 ng of amplified DNA are mixed with 50–80 pmol of biotinylated oligonucleotide in a total volume of 100 μL of SSC 4.2X, SDS 0.07%. DNA is denatured at 95 °C (3 min), and annealing is performed at room temperature for 15 min.

DNA molecules hybridized to biotinylated probes are selectively captured by streptavidin coated beads (Streptavidin Magnetic Particles, Boehringer Mannheim), prepared as follows: 1 mg of beads is extensively washed in $TEN_{100}$ (10 mM Tris-HCl, 1 mM EDTA, 100 mM NaCl, pH 7.5) and resuspended in 40 μL of the same buffer. To minimize nonspecific binding of genomic DNA, 10 μL (corresponding to approximately 1 μg) of an unrelated PCR product is mixed to the beads before adding the hybridization mixture. The prepared beads are mixed to the DNA-probe hybrid molecules (diluted with 300 μL of $TEN_{100}$) and incubated for 30 min at room temperature with constant gentle agitation.

The beads-probe-DNA complex is separated by a magnetic field from the hybridization buffer, which is then discarded. Nonspecific DNA is removed by three nonstringency washes and three stringency washes. Each wash is carried out for 5 min at room temperature with gentle mixing. DNA is recovered by magnetic field separation each time. Nonstringency washes are performed by adding 400 μL of $TEN_{1000}$ (10 mM Tris-HCl, 1 mM EDTA, 1 M NaCl, pH 7.5), while stringency washes are performed by adding 400 μL of SSC 0.2X, 0.1% SDS to the DNA. The last nonstringency wash and the last stringency wash are stored for further use (see below).

DNA is separated from the beads-probe complex by two denaturation steps. In the first step 50 μL of TE (Tris-HCl 10 mM, EDTA 1 mM, pH 8) is added to the beads, which are then incubated at 95 °C for 5 min. The supernatant, containing target DNA, is quickly removed and stored. The second denaturation step is performed by treating beads with 12 μL of 0.15 M NaOH; in this case the recovered supernatant must be neutralized, before storage, by the addition of an appropriate amount of acetic acid. This is determined in advance by titrating the NaOH stock solution with 0.1667 M acetic acid. TE is then added to reach a final volume of 50 μL.

The last nonstringency wash, the last stringency wash and the two elutions obtained from the denaturation steps should harbour an increasing proportion of DNA fragments containing the selected repeat and should carry the *Mse*I-N primer target site at each end.

DNA recovered from washing and denaturation steps is precipitated with one volume of isopropanol and sodium acetate (0.15 M final concentration), and resuspended in 50 μL of water. Two microliters of each recovered fraction are amplified by 30 cycles of PCR using the *Mse*I-N primer under the conditions described above. Agarose gel visualization of the amplified fragments should display in each of the four PCRs a smear above 200 bp (ideally the PCR of the last stringency wash should not yield any product, indicating complete removal of nonspecifically bound DNA).

The PCR products of the two elution steps are the best candidates for producing a highly enriched microsatellite library, because they are likely to contain the largest proportion of repeat-containing fragments. Cloning PCR amplicons can be conveniently carried out by using the TOPO-TA cloning kit (Invitrogen), with an expected yield of 1000–4000 recombinant colonies. It is recommended to screen by PCR at least 20 clones, using vector primers, to ensure that inserts of different size have been cloned. PCR products can be purified by Exonuclease-Phosphatase (PCR products presequencing kit, Amersham-Pharmacia) and directly sequenced. It is advisable to add DMSO (5% final concentration) to the sequencing reaction mix in order to prevent poor sequencing results due to the presence of repeats.

## Concluding remarks

Having considered in detail the problems of microsatellite isolation, the question remains as to which is the best strategy to adopt.

The possible options are to isolate microsatellites in one's own laboratory (either by using traditional methods or enrichment procedures) or to choose a commercial supplier.

Traditional methods have a low efficiency and they can be time-consuming. For organisms with high numbers of microsatellites (e.g. fish) together with optimizations of procedures (e.g. screening with multiple probes) the efforts required can be affordable, although some experience in library screening is needed. However, it is possible that microsatellites have a low frequency in a specific genome (Box 2).

Enrichment protocols appear preferable because, besides the advantage of being fast and efficient, they require

only basic skills in molecular biology (cloning is the most difficult step) and limited laboratory equipment, in addition to what is required for subsequent microsatellite screening.

In addition the initial cost of reagents is limited (with less than 2000 US dollars the FIASCO protocol allows one to perform 10 enrichments). By starting from DNA that has already been extracted, cloned products ready to be sequenced can be obtained in about three days, given that the method has been already set up in one's laboratory. However, for research groups that use microsatellites only occasionally, or that focus their research on one or few organisms, a commercial supplier can be a good option. The cost can be quite variable depending on the request. An enriched library with at least 50% of positive clones containing repeats (and 5–20 clones sequenced) can cost between 5000 and 10 000 US dollars. The price is much higher if the supplier is required to carry out the whole procedure and provide optimized microsatellites loci. In fact, the task of sequencing and optimizing microsatellites loci is, by itself, time- and labour-consuming.

In conclusion microsatellites have become a 'must' for many genetic studies, but some issues about their isolation are still open. Will the future provide us with a well established universal protocol, or will completely new approaches become available due to a better knowledge of microsatellite evolution combined with new technical advances? The situation at present is such that a careful evaluation of the experimental strategy has to be carried out case by case.

## Acknowledgements

We thank Dr A. Grapputo, Dr L. Maccatrozzo, C. Papetti, and Dr F. Tinti for the help in testing the FIASCO protocol in various organisms. This work was supported by the Italian MURST and by the University of Padova funding.

## References

Arcot SS, Wang Z, Weber JL, Deininger PL, Batzer MA (1995) Alu repeats: a source for the genesis of primate microsatellites. *Genomics*, **29**, 136–144.

Armour JA, Neumann R, Gobert S, Jeffreys AJ (1994) Isolation of human simple repeat loci by hybridization selection. *Human Molecular Genetics*, **3**, 599–565.

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.

Cifarelli RA, 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.

Cooper G, Amos W, Bellamy R *et al.* (1999) An empirical exploration of the $(\delta\mu)^2$ genetic distance for 213 human microsatellite markers. *American Journal of Human Genetics*, **65**, 1125–1133.

Cooper S, Bull CM, Gardner MG (1997) Characterization of microsatellite loci from the socially monogamous lizard *Tiliqua rugosa* using a PCR-based isolation technique. *Molecular Ecology*, **6**, 793–795.

D'Amato ME, Lunt DH, Carvalho GR (1999) Microsatellite markers for the hake *Macruronus magellanicus* amplify other gadoid fish. *Molecular Ecology*, **8**, 1086–1087.

Edwards KJ, Barker JHA, Daly A, Jones C, Karp A (1996) Microsatellite libraries enriched for several microsatellite sequences in plants. *Biotechniques*, **20**, 758.

Ellegren H, Primmer CR, Sheldon B (1995) Microsatellite evolution: directionality or bias in locus selection? *Nature Genetics*, **11**, 60–62.

Ender A, Schwenk K, Städler T, Streit B, Schierwater B (1996) RAPD identification of microsatellites in *Daphnia*. *Molecular Ecology*, **5**, 437–441.

Fisher PJ, Gardner RC, Richardson TE (1996) Single locus microsatellites isolated using 5'-anchored PCR. *Nucleic Acids Research*, **24**, 4369–4371.

FitzSimmons NN, Moritz C, Moore SS (1995) Conservation and dynamics of microsatellite loci over 300 million years of marine turtle evolution. *Molecular Biology and Evolution*, **12**, 432–440.

Hamilton MB, Pincus EL, Di-Fiore A, Fleischer RC (1999) Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. *Biotechniques*, **27**, 500–507.

Hancock JM (1999) Microsatellites and other simple sequences: genomic context and mutational mechanisms. In: *Microsatellites. Evolution and Applications* (eds Goldstein DB, Schlötterer C), pp. 1–9. Oxford University Press, New York.

Jarne P, Lagoda PJL (1996) Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution*, **11**, 424–429.

Kandpal RP, Kandpal G, Weissman SM (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.

Karagyozov L, Kalcheva ID, Chapman VM (1993) Construction of random small-insert genomic libraries highly enriched for simple sequence repeats. *Nucleic Acids Research*, **21**, 3911–3912.

Kijas JM, Fowler JC, Garbett CA, Thomas MR (1994) Enrichment of microsatellites from the citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. *Biotechniques*, **16**, 656–662.

Knapik EW, Goodman A, Ekker M *et al.* (1998) A microsatellite genetic linkage map for zebrafish (*Danio rerio*). *Nature Genetics*, **18**, 338–343.

Kocher TD, Thomas WK, Meyer A *et al.* (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.

Lench NJ, Norris A, Bailey A, Booth A, Markham AF (1996) Vectorette PCR isolation of microsatellites repeat sequences using anchored dinucleotide repeat primers. *Nucleic Acids Research*, **24**, 2190–2191.

Litt M, Luty JA (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 BH (1997) *Statistical Genomics: Linkage, Mapping and QTL Analysis*. CRC Press, Boca Raton, FA.

Lunt DH, Hutchinson WF, Carvalho GR (1999) An efficient method for PCR-based identification of microsatellite arrays (PIMA). *Molecular Ecology*, **8**, 893–894.

Moore SS, Sargeant LL, King TJ, Mattick JS, Georges M, Hetzel JS (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 PA, 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.

Ostrander EA, Jong PM, 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.

Primmer CR, Møller AP, Ellegren H (1996) A wide-ranging survey of cross-species amplification in birds. *Molecular Ecology*, **5**, 365–378.

Rassmann K, Schlötterer C, Tautz D (1991) Isolation of simple-sequence loci for use in polymerase chain reaction-based DNA fingerprinting. *Electrophoresis*, **12**, 113–118.

Refseth UH, Fangan BM, Jakobsen KS (1997) Hybridization capture of microsatellites directly from genomic DNA. *Electrophoresis*, **18**, 1519–1523.

Richardson T, Cato S, Ramser J, Kahl G, Weising K (1995) Hybridization of microsatellites to RAPD: a new source of polymorphic markers. *Nucleic Acids Research*, **23**, 3798–3799.

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.

Rubinsztein DC, Amos W, Leggo J *et al.* (1995) Microsatellite evolution — Evidence for directionality and variation in rate between species. *Nature Genetics*, **10**, 337–343.

Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, New York.

Schlötterer C, Amos B, Tautz D (1991) Conservation of polymorphic simple sequence loci in cetacean species. *Nature*, **354**, 63–65.

Schlötterer C, Tautz D (1992) Slippage synthesis of simple sequence DNA. *Nucleic Acids Research*, **20**, 211–215.

Schuler GD, Boguski MS, Stewart EA *et al.* (1996) A gene map of the human genome. *Science*, **274**, 540–546.

Scribner KT, Pearce JM (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.

Takahashi H, Nirasawa K, Furukawa T (1996) An efficient method to clone chicken microsatellite repeat sequences. *Japanese Poultry Science*, **33**, 292–299.

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

Tenzer I, degli Ivanissevich S, Morgante M, Gessler C (1999) Identification of microsatellites markers and their application to population genetics of *Venturia inequalis*. *Phytopathology*, **89**, 748–753.

Toth G, Gáspari Z, Jurka J (2000) Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Research* 967–981.

Valle G (1993) TA-repeat microsatellites are closely associated with ARS consensus sequences in yeast chromosome III. *Yeast*, **9**, 753–759.

Vos P, Hogers R, Bleeker M *et al.* (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, **23**, 4407–4414.

Waldbieser GC (1995) PCR-based identification of AT-rich tri and tetranucleotide repeat loci in an enriched plasmid library. *Biotechniques*, **19**, 742–744.

Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics*, **44**, 388–396.

Wilder J, Hollocher H (2001) Mobile elements and the genesis of microsatellites in dipterans. *Molecular Biology and Evolution*, **18**, 384–392.

Williams JG, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research*, **18**, 6531–6535.

Wu K, Jones R, Danneberger L, Scolnik PA (1994) Detection of microsatellite polymorphisms without cloning. *Nucleic Acids Research*, **22**, 3257–3258.

Zhivotovsky LA, Feldman MW (1995) Microsatellite variability and genetic distances. *Proceedings of the National Academy of Sciences of the USA*, **92**, 11549–11552.

---

Lorenzo Zane, Luca Bargelloni, and Tomaso Patarnello are primarily interested in evolutionary processes and part of their research activity is currently focusing on population genetics of marine organisms and conservation genetics of endangered species.

---