1458–1465 *Nucleic Acids Research, 1998, Vol. 26, No. 6*

Direct amplification of length polymorphisms (DALP), or how to get and characterize new genetic markers in many species

E. Desmarais*, I. Lanneluc and J. Lagnel[1]

Laboratoire de Génome et Populations CNRS UPR 9060, Université Montpellier 2 CC. 63, Place E.Bataillon, 34095 Montpellier, France and [1]INRA Laboratoire de Zoologie, 2 Place Viala, 34060 Montpellier, France

Received November 11, 1997; Revised and Accepted January 23, 1998 DDBJ/EMBL/GenBank accession nos U40066, AJ000243–AJ000248

ABSTRACT

Direct amplification of length polymorphisms (DALP) uses an arbitrarily primed PCR (AP-PCR) to produce genomic fingerprints and to enable sequencing of DNA polymorphisms in virtually any species. Oligonucleotide pairs were designed to each produce a specific multi-banded pattern and all the fragments thus generated can be directly sequenced with the same two universal M13 sequencing primers. This strategy combines the advantages of a high resolution fingerprint technique and the possibility of characterizing the polymorphisms. The use of family members as templates in the multi-locus detection step allows a direct test of allele transmission, as well as early mapping of the markers or selection of loci associated with some traits or diseases. We used this method to detect micro-deletions/insertions and microsatellite DNA loci useful in population genetics studies, but it could be applied in many other fields of biology, such as genome mapping for definition of polymorphic sequence tagged sites, directly localized on a genetic map.

INTRODUCTION

In population biology the use of genetic markers is today predominantly performed through the PCR technique and is conditioned by *a priori* knowledge of the nucleotide sequences flanking the loci; this is an absolute prerequisite for design of the primers. For a species where no sequence data are available this implies construction of a genomic library, its screening for potentially polymorphic markers and sequencing of numerous clones.

To bypass this preliminary step of sequence recovery several non-targeted amplification methods have been developed, using arbitrarily chosen sequences as primers. These techniques take advantage of the possibility that any given oligonucleotide will pair with templates which are not strictly homologous under relaxed stringency. Then PCR leads to multi-locus amplification instead of targeting a unique locus. Arbitrarily primed PCR (AP-PCR; 1) and randomly amplified polymorphic DNA (RAPD; 2) are the two major methods based on this principle. These protocols are performed at a low annealing temperature (at least in the first cycles) and use either long oligonucleotides for AP-PCR (>18mer) or a single short one for RAPD (10mer). They produce a 'fingerprint' pattern that can be useful in characterizing strains or species and in investigating the level of genetic polymorphism of a new species. However, almost all the detected polymorphisms have a dominant inheritance (presence/absence of bands) and the amplified regions cannot be easily sequenced to design locus-specific primers, even with a cloning step. Moreover, because of the weak stability of the priming reaction the RAPD method, which has been more extensively used, lacks reproducibility and many factors can compromise the results (3).

The amplified fragment length polymorphism technique (AFLP; 4) is another multi-locus approach based on endonuclease digestion of total genomic DNA. Synthetic adapters are ligated to the restriction fragments and selective PCR, anchored on these adapters, is performed to amplify discrete DNA fragments. Although this method may be efficient in revealing polymorphisms, even between closely related individuals, each experiment implies three steps and four different primers for analysis of complex genomes. Furthermore, artifactual polymorphisms can be generated by incomplete cleavage of genomic DNA or imperfect ligation.

In summary, although detection of polymorphisms is possible with these methods, sequencing of the loci is difficult or even unrealisable. Thus definition of mono-locus-specific primers, more reliable for analysis of large samples, is impossible and population studies are performed with the initial arbitrarily chosen primers. A recent method called sequencing with arbitrary primer pairs (SWAPP; 5) was developed to sequence polymorphic bands isolated from a multi-locus pattern obtained by AP-PCR. It uses low resolution agarose gel electrophoresis for a first selection of potentially polymorphic bands and further SSCP analysis of each fragment. While this method is powerful for characterization of polymorphisms, it can be both simplified and improved for detection of more loci.

We propose a new strategy by which a much larger number of polymorphic loci can be detected and isolated for sequencing in only one step. Indeed, this method is not simply another

*To whom correspondence should be addressed. Tel: +33 4 67 14 4684; Fax: +33 4 67 14 4554; Email: desmarais@crit.univ-montp2.fr

supplementary fingerprint technique but was designed from the very beginning to obtain nucleotide sequence information on DNA fragments from any genome with no need for a genomic library. For this purpose a set of oligonucleotides, hereafter called 'selective', was designed, all sharing the same 5′ core sequence: the universal M13 sequencing primer. Each was always used in combination with the same 'reverse' primer: the M13 reverse primer. We show here that such pairs of primers are efficient in producing specific multi-banded patterns where inter-individual length variations, including microsatellite polymorphisms, can be detected. Moreover, because of the special design of the selective primers, each band of these different patterns can be sequenced with the same two oligonucleotides: the universal sequencing primers. We call this new approach DALP, for direct amplification of length polymorphisms.

The first application of this method on two subspecies of mouse allowed detection of 2–8 intra-specific polymorphisms per primer used. However, a rapid survey on a panel of very different organisms with a single selective primer revealed that this number can be much higher in some species. The sequencing showed that micro-deletions/insertions are very common, at least in the tested genomes, and that microsatellite loci are only a small fraction of the length polymorphisms that can be detected by this method.

We also investigated the level of specificity of the priming reactions generated with nearly identical primers by an estimation of redundancy in the multi-locus amplifications (i.e. the number of genomic fragments amplified more than once) and analysis of the priming sites of a random selection of bands generated from λ virus DNA. This revealed an interesting capacity for stable non-perfect annealing between the oligonucleotide and the matrix, which offers numerous possibilities for extension of the primer set described in this work.

MATERIALS AND METHODS

DNA samples

DNA extraction was performed by classical proteinase K/phenol/chloroform extraction or, for the mussels, by a rapid protocol using Chelex 100 resin. The DNA concentrations were deliberately not adjusted, ranging from 50 ng/µl to several hundred. These extraction products as well as up to 20-fold dilutions were used in the experiments. DNA from various species were used: mice (*Mus musculus domesticus*, *M.musculus musculus* and *M.spretus*), fish (two tilapia, *Oreochromis niloticus* and *Sarotherodon melanotheron*, and one barb, *Barbus intermedius*), birds (*Larus cachinnans* and *L.audouinii*), mussels (*Mytilus galloprovincialis* and *M.edulis*), a mite (*Amblyseius californicus)* and one virus (λ virus).

Primers

Each reaction was carried out with a combination of one of the 'selective' primers listed in Table 1 and the 'reverse' primer: between two experiments only the selective primer was changed in order to amplify a new selection of genomic fragments.

Amplification

To be directly sequenced a fragment needs to have two different ends. In order to identify such fragments for, each pair of primers two preliminary amplifications were performed in parallel on one sample, with alternate labelling of the two oligonucleotides. The products were loaded onto a gel to sort out bands possessing the same primer at both ends (i.e. revealed by single labelling) from those possessing two different primers at their ends (i.e. present in the two differently labelled products).

The reactions were carried out in a final volume of 20 µl starting from 2 µl DNA, 5 pmol each primer and 0.5 U Taq DNA polymerase in 1.8 mM $MgCl_2$, 100 µM dNTP and 1× reaction buffer. The cycling parameters were: 95°C 1 min; 91°C 30 s; 50 or 55°C 30 s; 72°C 30 s; 29 more cycles to step 2.

Electrophoresis of the multi-locus amplification products

Electrophoresis was performed on denaturing sequencing gels (0.4 mm thick), prepared with high quality acrylamide solution (acrylamide/bisacrylamide 29:1; BioRad), and run on a 50 cm long apparatus (BioRad). This is essential to achieve sufficient resolution between two very close bands in the multi-locus pattern. The samples were mixed with 10 µl 100% formamide loading dye and heated for 5 min at 100°C before loading. The gels were run for 4 or 5 h at 55 W.

Re-amplification of isolated bands

After electrophoresis, drying and autoradiography, the bands were extracted from the gels and re-amplified with the two universal primers (Table 1), no matter which selective primer was used in AP-PCR. An autoradiogram was precisely superimposed on the dried gel following guide marks and the gel slice was cut through the celluloid film. The Whatman paper was removed and the remaining acrylamide piece was directly immersed in 50 µl PCR mixture containing 10 pmol each primer. The reaction was started for 30 cycles at 50 or 58°C. With this extremely simplified protocol we succeeded in re-amplifying nearly 100% of the bands picked up, including some extracted from 2 year old gels (not shown).

After purification the re-amplified products can be sequenced according to standard protocols for PCR products: both radioactive and automatic fluorescent sequencing were used successfully.

RESULTS

Obtaining discrete fragments with two different ends

All the 'selective' oligonucleotides contained the same 5′ core sequence, the M13 –40 universal sequencing primer (Table 1). The 'reverse' primer (DALPR) is also one of the common M13 reverse primers. Then sequencing of all the fragments produced is possible with the universal primers usually supplied in most of the sequencing kits. For example, we frequently used a 24mer primer in our sequencing reactions: the basic 17mer used in all selective primers with seven more bases at its 5′-end.

In the selective primers permutations of *X* (2, 3 or 4) nucleotides have been added at the 3′-end of the common sequence, avoiding thymidine to increase specificity of the priming reaction (6). Thus the difference between two primers (and two experiments) can be only one base.

The set of primers was used in intra- and inter-sub-specific analysis on mouse genomic DNA. In combination with the reverse primer all the selective primers produced specific multi-banded patterns composed of easily analysable discrete bands at 50 or 55°C. Increasing the annealing temperature resulted in a diminution in the signal, with only minor modifications of the profiles (data not shown). Furthermore, differences in DNA concentrations had only a slight influence on the results (Fig. 1).

Table 1. . Sequences of the primers used

Reverse Primer	DALPR	T T T C A C A C A G G A A A C A G C T A T G A C
Selective Primers : DALP2X_i	DALP221	G T T T T C C C A G T C A C G A C G C
	DALP231	G T T T T C C C A G T C A C G A C A G C
	DALP232	G T T T T C C C A G T C A C G A C G A C
	DALP233	G T T T T C C C A G T C A C G A C A C G
	DALP234	G T T T T C C C A G T C A C G A C C A G
	DALP235	G T T T T C C C A G T C A C G A C C A C
	DALP241	G T T T T C C C A G T C A C G A C T C A G
	DALP242	G T T T T C C C A G T C A C G A C C T A G
Re-amplification and sequencing: U. S. P.	-40USP	G T T T T C C C A G T C A C G A C
	DALPR	T T T C A C A C A G G A A A C A G C T A T G A C

U.S.P., Universal sequencing primers. The grey box indicates the –40 USP sequence.

Whenever the PCR products had the same two ends, unbalanced ratios between the two primers were tested. In all cases it was an increase in the amount of the reverse primer (up to five times the selective one) which proved to be efficient in solving the problem. This suggests that the annealing sites for the two primers are evenly distributed over the genome and that either the concentration of the reverse primer was underestimated or that this primer has a lower affinity for the DNA matrix than the selective primer (maybe due to a larger size). This unbalanced ratio also reduced the smeary background that appeared in some lanes (data not shown).

Amplification of primer-specific patterns

Typically between 30 and 40 workable bands were obtained over 40 cm of gel, ranging from 200 to 1000 bp after a 5 h migration (Fig. 1). For the primer DALP241 the sum of all the fragments represents ~15 kb. Repeated independent experiments as well as tens of simultaneous replicates of amplifications confirmed the stability of the priming reactions (not shown).

Roughly 84% of the bands are shared between *M.m.domesticus* and *M.m.musculus*. However, with *M.spretus* this ratio drops to only 45% (not shown). The majority of these bands are constant (i.e. present in all lanes), but two types of variation can be detected: presence/absence of bands and apparent length polymorphisms (individuals have either a unique fast or slow band whilst some others display two ‘alleles’). The allelism of these bands was verified by performing arbitrary PCR directly on members of a backcross family: all the loci informative in this family had the expected segregation (data not shown). Using such a sample also offered the possibility of directly mapping the loci at this early stage of detection. For constant bands or when the family was not informative extraction and sequencing of the DNA was used to confirm the homology of the sequence between two bands.

A comparison of the products obtained with primer pairs that differ by only one base (DALP234/DALPR versus DALP241/DALPR and DALP234/DALPR versus DALP235/DALPR) showed that the priming reaction can be very specific for the selective primer sequence (Fig. 1). Nevertheless, to test redundancy of the arbitrarily primed amplification at least 12 independent bands, ranging from 150 to 660 bp, from DALP234, DALP235 and DALP241 products were sequenced. All 39 bands produced with these very similar primers were different. This indicates that redundancy between two different primer products as well as within a single PCR product is very low. This means that the portion of a genome analysed with one selective primer is approximately the sum of all the visible bands on a gel, between ~10 and 15 kb. However, sequencing of some polymorphic loci revealed that three bands were produced by two different combinations: DALP232/R and DALP235/R. This could be simply due to contamination of the primer solutions, but it is noticeable that DALP232 was the one which produced the weakest signal. This suggests that the priming reactions initiated by this sequence were less stable and consequently less specific, maybe because of the presence of a repetitive structure at its 3′-end: AC GAC GAC.

Analysis of base pairing between λ DNA and the arbitrary primers

The constitution of the priming sites in the arbitrarily primed amplifications was investigated using λ virus DNA as template. Depending on the selective primer used, 4–10 bands were obtained ranging from 200 bp to 1 kb in size, equivalent to a fraction of the λ genome of 2–4 kb/primer used (data not shown). Three bands produced by three different primers were sequenced and localized on the λ genome. Five out of the six priming sites are composed of at least five consecutive perfect matches (Fig. 2). In four priming sites the homologous stretches are at the 3′-end of the primer, but one has a mismatch at the penultimate position. In the other two sites the matches do not include the extreme 3′ base of the primer: in one of these sites perfect homology is over nine contiguous bases; the other, involving DALP232, contains a maximum of two contiguous perfect matches dispersed along the primer.

Application to other species

In order to test the potentiality of this protocol in working with different species we performed an amplification with DALP231 and DALPR on DNA from a mite, two mollusc species (mussels), two bird species (gulls) and three fish species [one hexaploid African barb species (150 chromosomes) and two tilapia species].

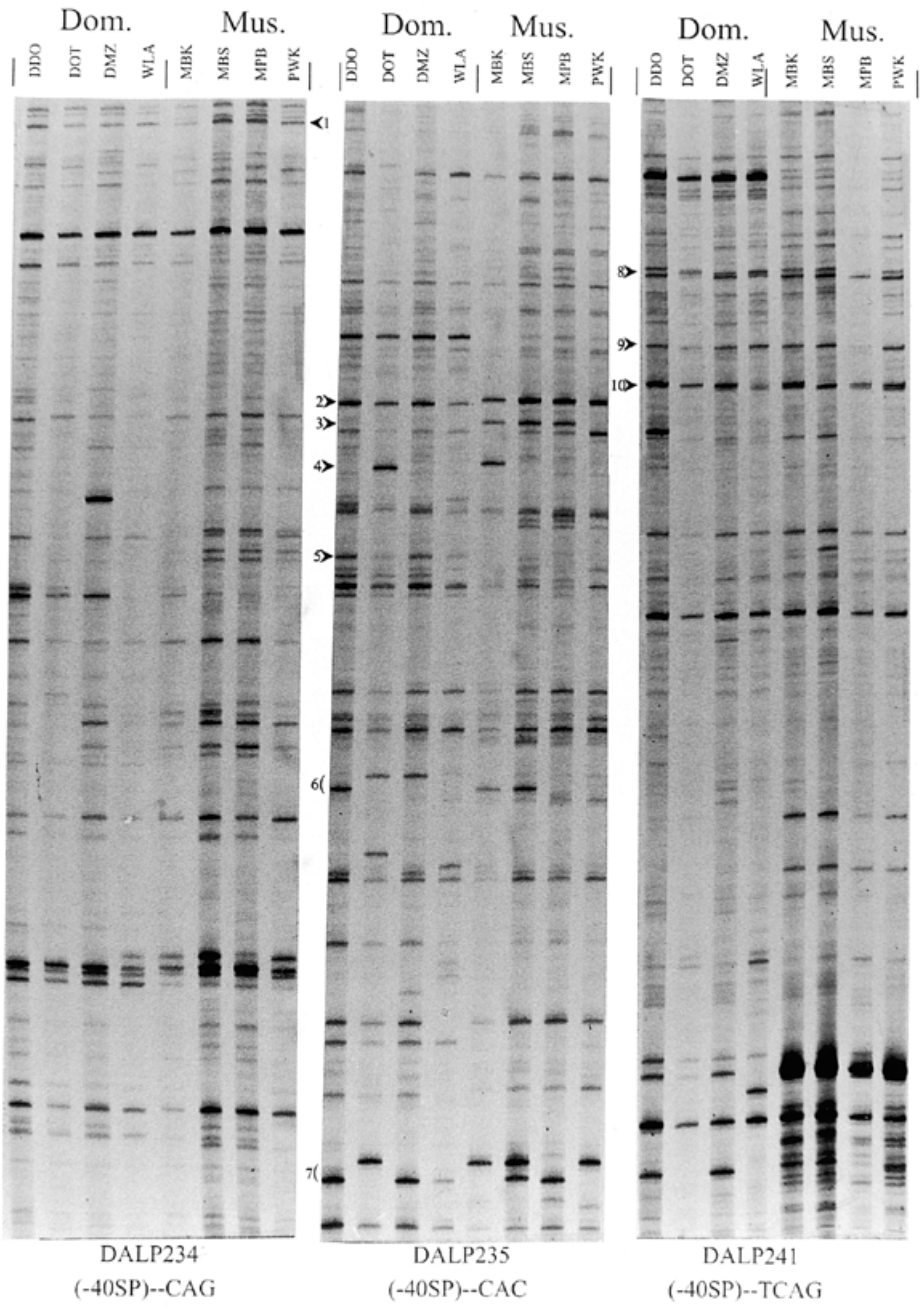

Figure 1. Acrylamide gel electrophoresis of mouse genomic DNA amplified with three different selective primers in combination with DALPR. The three primers differ from each other by only one nucleotide at their 3′-end. From left to right: DALP234, M13 –40 sequencing primer + CAG; DALP235, –40 SP + CAC; DALP241, –40 SP + TCAG. The products were run on a 5% acrylamide gel for 5 h in 0.5× TBE, 7 M urea. Aliquots of 6 µl from a 20 µl PCR were loaded after addition of 4 µl formamide dye and heat denaturation. Dom., four strains of *M.m.domesticus* of various origins (DDO, Denmark; DOT, Tahiti; DMZ, Morocco; WLA, Poland); Mus., four strains of *M.m.musculus* (MBK and MBS, Bulgaria; MPB and PWK, Poland). Quantities of DNA used in the PCR ranged from 5 (PWK and WLA) to 60 ng (MBK). Lane 1, band Mus234-R38 (~800 bp); lane 2, locus MUS232-1 (here amplified with DALP235); lane 3, locus MUS235-1; lane 4, locus MUS235-4 (a presence/absence polymorphism combined with a length polymorphism); lane 5, band Mus235-R14 (460 bp); lane 6, locus MUS235-3; lane 7, locus MUS235-2; lane 8, band Mus241-R9 (660 bp); lane 9, band Mus241-R4 (580 bp); lane 10, locus MUS241-1.

All the reactions were carried out in the same experiment and strictly under the conditions used previously for mice. In all cases the DALP method was able to produce a ladder of discrete bands (Fig. 3), even with DNA from the barbs, a species which have an especially highly complex genome and for which a smear could have been generated. For the mite, since the individual DNA extraction yield is very low, the resulting signal is lower than for the other species. This experiment also showed that DNA prepared with Chelex 100 can be used as a template for DALP, with no decrease in size of the amplified fragments (see mussel lanes in Fig. 3).

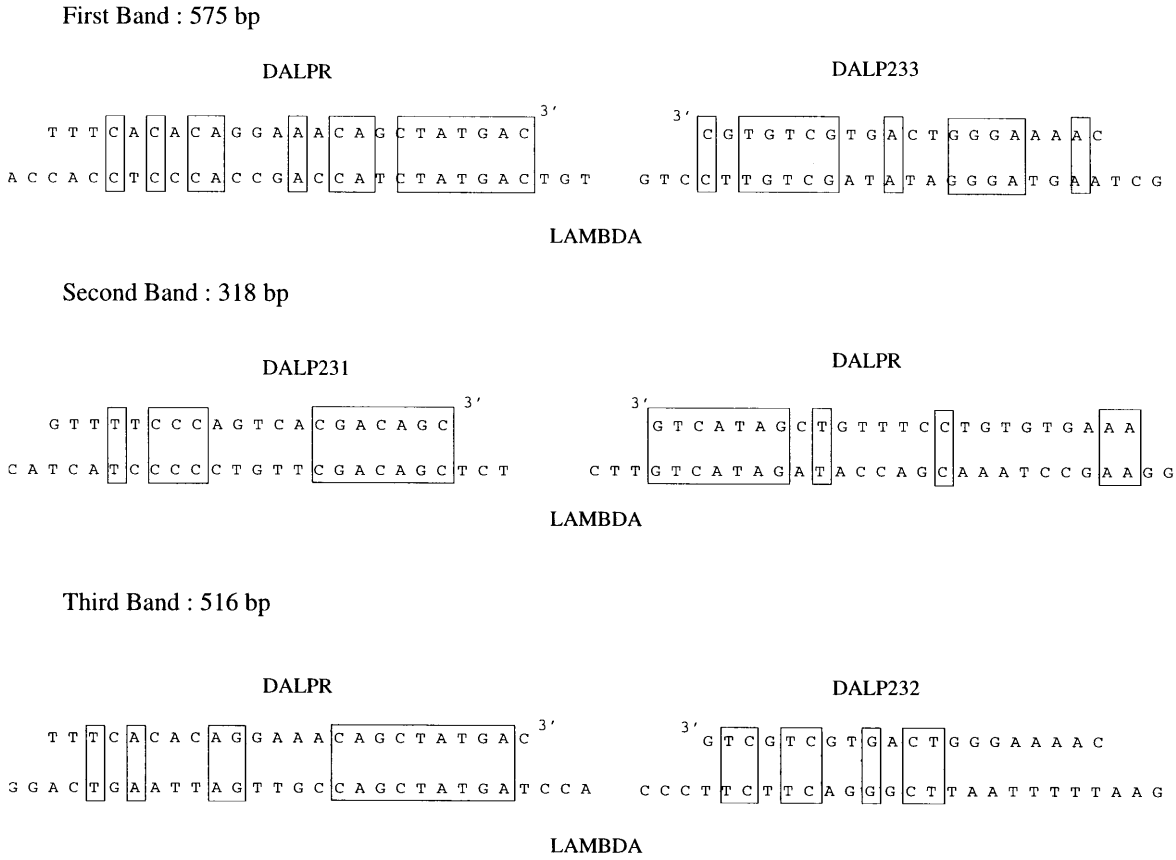

Figure 2. Annealing sites on λ genomic DNA. Three bands were randomly chosen among the λ amplification products obtained with DALP231, DALP232 and DALP233. They were totally sequenced and their priming sites inferred from the λ sequence.

Nature of the polymorphic loci

In all the species the two kinds of intra-specific polymorphisms described earlier were observed (presence/absence and length polymorphisms). The electrophoresis conditions used allowed detection of differences of as little as 2 bp between fragments up to 700 bp in length. In the mouse, and depending on the selective primer used, from two to eight polymorphic loci were detected using only two to four individuals per species or sub-species (Table 2). However, it of course depends on the species: for the mussels the intra-specific variability level was so high that it became difficult to identify alleles from a single locus except by sequencing, whereas in the gulls only inter-specific variations can be visualized. It is important to note that with this very limited experiment we were able to rapidly assess the global polymorphism of the species, thus confirming the preliminary results of previous mono-locus studies.

Among the mouse loci, some were polymorphic in both *M.m.domesticus* and *M.m.musculus*, while others seemed to be polymorphic in only one of the two sub-species. Several invariant bands were also found in only one of the two sub-species and might be diagnostic. Some other clear and intense bands occurred repeatedly in only a few lanes, suggesting a presence/absence polymorphism and not a priming artefact. Sequencing two of these bands did not reveal any homology with a known genomic mobile element, which could have been a simple explanation for this polymorphism.

We sequenced seven loci to determine the origin of the length variation. In six cases this was due to a short insertion/deletion of 2–4 bp, whereas the other locus contained a microsatellite motif (Table 3). Primers were designed and genomic amplifications were performed for three loci in mice and one in tilapia. In the two mouse sub-species and in some other full species the three loci were amplified successfully, showing a relative conservation. For the fish, while amplification confirmed segregation of this polymorphism in natural populations of *O.niloticus*, in which it was first detected, it gave some unclear results in *S.melanotheron*.

Finally, 39 independent bands (13.5 kb in total) were sequenced and compared with the sequence database using the BLAST algorithm. Only two fragments were found to be homologous to murine repetitive elements (L1 and B1) and five others were homologous (from 64 to 100%) to known coding or non-coding regions (Table 3).

Sequences of the polymorphic loci have been submitted to GenBank for the tilapia locus, which we called DON-1 (accession no. U40066), and to EMBL for the mouse loci, called MUS232-1 (accession no. AJ000243), MUS241-1 (accession no. AJ000244), MUS235-1 (accession no. AJ000245), MUS235-2 (accession no. AJ000246), MUS235-3 (accession no. AJ000247) and MUS235-4 (accession no. AJ000248).

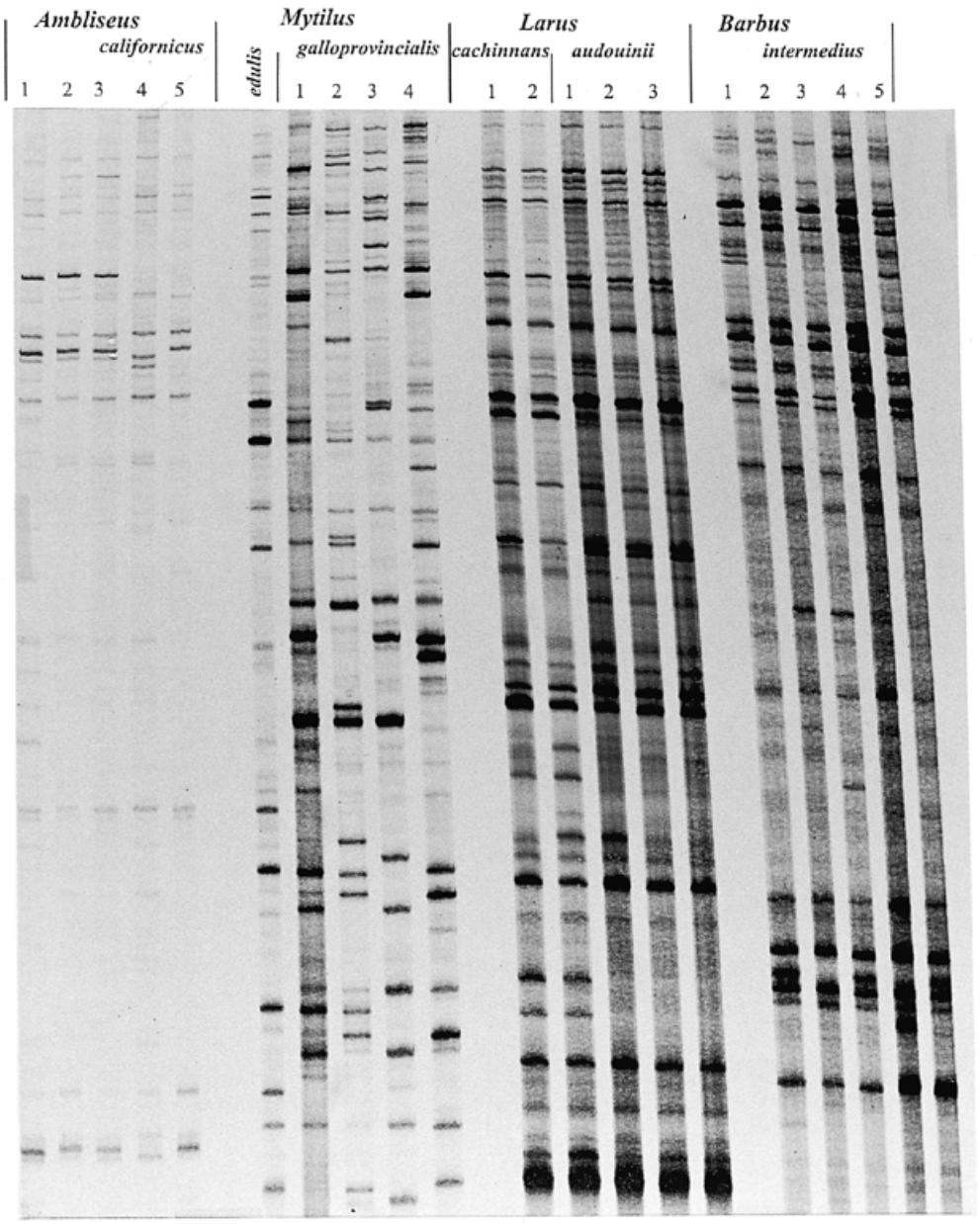

Figure 3. Multi-locus amplification on various species. DNA from very different species was used to test the ability of the protocol to produce discrete analysable bands in many organisms. The test was performed with DALP231 in combination with DALPR, under the same conditions for all samples (see text). From left to right: mite (*Amblyseius californicus)*; mussels (*Mytilus galloprovincialis* and *M.edulis*); birds (*Larus cachinnans* and *L.audouinii*); fish (*Barbus intermedius*). In the mite, fish and especially in the mussels many polymorphic loci can be seen in spite of the low number of samples used for each species. In birds only inter-specific differences can be observed.

DISCUSSION

The multi-locus amplifications described in this work were performed at a relatively higher annealing temperature, lower concentration of magnesium chloride and with fewer cycles than previously described (7). Despite the higher stringency of annealing, the reaction succeeded in producing multi-locus patterns with arbitrarily chosen primers. However, probably because of these conditions, the annealing sites for a given arbitrary primer were constant across experiments and highly dependent on the sequence of the oligonucleotide. In addition, all fragments (instead of a maximum of 68%; 7) possessed two different oligonucleotides at their ends, allowing direct sequencing.

Production of such patterns is the result of imperfect annealing between templates and primers and of the 'high tolerance of the PCR to mismatches'. This property of the PCR was established from an analysis of databases and simulations on production of non-targeted products in 'mono-locus' PCR (8). Here we deliberately wanted to produce non-targeted products and we fully used this capacity, as the analysis of some priming sites

Table 2. Number of polymorphisms detected per primer in the mouse sub-species

	DALP231	DALP232	DALP233	DALP234	DALP235	DALP241	DALP242
Sample size and species	Mm X Md	4 Mm 4 Md	Mm X Md	4 Mm 4 Md	4 Mm 4 Md	4 Mm 4 Md	Mm X Md
Length polymorphisms	4	4	1	3	6	5	2
Presence/ absence	1	0	1	2	2	1	0

Mm × Md, offspring of a back-cross between *M.m.musculus* (Mm) and *M.m.domesticus* (Md).

Table 3. Summary of the polymorphic loci and/or the database homologies

Locus (*)	Homology *	Motif	Primers for the mono-locus amplifications	Size*
MUS241-1 (10)		i-TA	F : GCCAACATGAGGTACAAAGC R : TACTAAGCCAACCAAAAGCC	151
MUS235-1 (3)		$(GAAT)_n$	F : CATTAATAGAGTCTCCGAGG R : GCTCTGTAAAATGGGAATGG	122
MUS235-2 (7)	Mouse cysteine protease	i-ACA	F : CGTCTATGGGAACCATTCCT R : TAATGACCCTGAATGCATGG	95
MUS235-3 (6)		i-CAGC		
MUS235-4 (4)		i-CT		
MUS232-1 (2)		i-AT		
DON1		i-GT	F : TTTCTCCCAGGTGTGTTCCC R : AGTAATTAGGCTGGACGAGC	336
Mus241-R4 (9)	α A-crystallin			
Mus241-R17	L1 Mouse repeated element			
Mus235-R14 (5)	Mouse Immunity Associated Protein 38			
Mus232-R3	Mouse AcylCoA deshydrogenase			
Mus241-R9 (8)	B1 Mouse repeated element			
Mus234-R38 (1)	Rat Olfactory receptor			

Size, size of the amplified product; i, insertion or deletion of the motif.
Homology, >60% on at least 50 bp.
*Band number in Figure 1.

revealed up to 13 mismatches for a 20mer oligonucleotide. But we also showed that reproducibility and specificity can be reached in such priming reactions; by only changing a single 3′ base in a single primer we produced new patterns in which more polymorphisms could be detected. Then the design of a wide range of supplementary primers on the same model will allow screening of many other genomic fragments; omitting thymidine for the last two 3′ bases, there remain 144 possible permutations of four bases.

Analysis of the priming sites on λ DNA reveals that initiation of polymerization is possible even if the extreme 3′ base is not perfectly paired. However, it is noticeable that in two such cases (Fig. 2) the mismatches involve a thymidine on the λ DNA. Such a mismatch was demonstrated to have very little effect on amplification of a precise locus (6). This finding is confirmed here with non-targeted priming, but may also be extended to the overall stability of the duplex, since internal mismatched thymidines again seem to have little effect, as suggested by the priming site of DALP232 on λ, where nine out of 13 mispaired bases are a T. Moreover, in hybridization reactions with 19 bp oligonucleotides G-T and G-A mismatches were shown to have only a slight effect on stability of the duplex (9) and it was assumed that hydrogen bonds can be formed in such base pairs: our results fully confirm this hypothesis, since the DALP232 annealing site on λ contains eight of these mismatches and, if they are scored together with the perfect matches, then the duplex is formed over 15 nt starting from the 3′-end of the primer. This is sufficiently stable to successfully initiate PCR. In such a case avoiding false priming in mono-locus PCR becomes very difficult and production of the expected band is only the result of a good ‘signal-to-noise’ ratio. This means that in many cases, if the resolution of electrophoresis were increased, numerous bands should then appear. This is also an explanation for the difficulty in sequencing RAPD bands. But this also means that the real difficulty in approaches like AP-PCR is not amplifying enough

targets, but, on the contrary, amplifying too many targets, resulting in a smear rather than discrete bands.

Detection of new polymorphisms

Among the seven polymorphic loci we sequenced, six contained a micro-deletion/insertion and only one a variable number of a repeated motif. How we identified the polymorphisms on gels obviously facilitated detection of loci with small differences between the alleles. However, detection of loci with alleles of very different sizes should also be possible. In this case some individual patterns would lack some bands, whereas some isolated bands would appear somewhere else in the same lane. We found such bands, but not associated with any absence at other locations and the sequencing did not reveal any structure that suggested the existence of other alleles of very different size. These bands were assumed to be a presence/absence polymorphism, but this could be confirmed only by definition of the primers and a detailed analysis of the results of genomic amplification on a larger sample.

Because of the abundance of bands in some parts of a gel, it happened that some bands that we thought to be allelic appeared, after sequencing, to be totally different from each other. One way to avoid such confusing patterns would be to reduce the number of amplified bands per primer: this could increase the distance between two loci. This could be better achieved by reducing the availability of the reagents needed for the polymerization reaction than by increasing the temperature. Indeed, our experiments showed that this only leads to a reduction in the signal, with no change in the patterns. Furthermore, Kwok *et al*. (6) showed that some mismatches are sensitive to a reduction in the deoxynucleotide concentration. But a reduction in the number of bands per primer could also mean that more primer is required to detect the same number of polymorphisms.

Finally, the fact that six out of the seven loci we sequenced contained a deletion/insertion suggests that this kind of polymorphism could be very frequent in the genomes we examined. It would be interesting to verify such an hypothesis in the future and to investigate the mechanisms leading to these mutations. Moreover, these kinds of genetic markers could be good alternatives when a large number of alleles at microsatellite loci is a handicap.

CONCLUSION

Using the DALP approach described here we were able to selectively amplify discrete fragments of DNA from a genome and to obtain genomic sequence information on several polymorphic loci without cloning. Presumably, by increasing the sample size or by selecting individuals from very distant populations, the number of polymorphic loci could have been increased. Moreover, the oligonucleotide set can be easily extended by choosing other permutations according to the pre-determined rule for the use of thymidine and in the light of our analysis of priming sites.

This protocol is efficient in very different species (virus, vertebrate, invertebrate, etc.) and is totally independent of any *a priori* knowledge of the flanking region of the loci or of their nature (microsatellite, deletion/insertion, presence/absence). Polymorphism of the loci is directly assessed. All these properties make this new approach a tool well-suited for population biology studies on new species. Moreover, in mapping projects the polymorphic loci can be rapidly localized, since multi-locus amplification can be performed on members of a family. Applied to a large sample of individuals which can be run on the same gel (60–80) the technique has a certain statistical power, in the localization of quantitative trait loci for example. This can also be very useful for definition of sequence tagged sites (STS), since in a single step physical genomic landmarks can be defined and integrated into a genetic map. This can be achieved with only the minimum material needed in a laboratory dedicated to population biology, a PCR machine and sequencing gel apparatus.

The problem of non-reproducibility, even if encountered for some primers, has only minor consequences, as everything in this protocol was designed for sequencing the polymorphic loci to define mono-locus primers and not for direct use in population surveys. Moreover, a by-product of this method could be conservation of the genomic material of precious and rare individuals, as we were able to re-amplify bands from dried gels after 2 years. DNA can probably be preserved for an even longer period as long as re-hydration is avoided. A replicate gel can also be transferred onto a filter membrane and hybridized with specific probes to pick up a gene of interest in a new species, even if no length polymorphism is present. Furthermore, some substitutions could also be detected if the DNA fragments are run on denaturing gradient electrophoresis gels or under the conditions for detecting single-strand conformation polymorphisms, as in the SWAPP method.

ACKNOWLEDGEMENTS

The authors wish to thank F.Bonhomme for his support and his comments on this work, B.Dod and D.Hedgecock for their helpful discussions on the manuscript, A.Orth, A.Chenuil, C.Daguin, C.Delsert and M.Genovart for providing the DNA samples. This work was partially supported by the Groupement de Recherche et d'Etude sur les Génomes.

REFERENCES

1 Welsh,J. and McClelland,M. (1990) *Nucleic Acids Res*., **18**, 7213–7218.
2 Williams,J.G.K., Kubelik,A.R., Livak,K.J., Rafalski,J.A. and Tingey,S.V. (1990) *Nucleic Acids Res*., **18**, 6531–6535.
3 Schweder,M.E.,Jr, Shatters,R.,G.,Jr, West,S.H. and Smith,R.L. (1995) *BioTechniques*, **19**, 38–42.
4 Vos,P., Hogers,R., Bleeker,M., Reijans,M., Lee,T.V.D., Hornes,M., Frijters,A., Pot,J., Peleman,J., Kuiper,M. and Zabeau,M. (1995) *Nucleic Acids Res*., **23**, 4407–4414.
5 Burt,A., Carter,D.A., White,T.J. and Taylor,J.W. (1994) *Mol. Ecol*., **3**, 523–525.
6 Kwok,S., Kellogg,D.E., McKinney,N., Spasic,D., Goda,L., Levenson,C. and Sninsky,J.J. (1990) *Nucleic Acids Res*., **18**, 999–1005.
7 Welsh,J. and McClelland,M. (1991) *Nucleic Acids Res*., **19**, 5275–5279.
8 Rubin,E. and Levy,A.A. (1996) *Nucleic Acids Res*., **24**, 3538–3545.
9 Ikuta,S., Takagi,K., Wallace,R.B. and Ikatura,K. (1987) *Nucleic Acids Res*., **15**, 797–811.