Rapid PCR-based characterization of sequences flanking microsatellites in large-insert libraries

J.Rothuizen* and M.van Raak
Utrecht University, Department of Clinical Sciences of Companion Animals, Yalelaan 8, 3508 TD Utrecht, The Netherlands

Received July 18, 1994; Revised and Accepted October 21, 1994

Microsatellite polymorphisms are widely being employed as a major tool in genome analysis. Analysis of length polymorphisms in the PCR depends on the elucidation of the sequence of their flanking regions. Different methods have been developed for the efficient characterization of the flanking sequences. Small insert genomic libraries directly suitable for sequencing of the entire insert and highly enriched for the desired microsatellite sequence may be obtained by a number of strategies (Ostrander et al.; Karagyozev et al., 1993; Kandpal et al., 1994). The limitation of small insert libraries is, however, that the inserts are too small for physical mapping by in situ hybridization. Microsatellites from large insert genomic libraries suitable for the construction of both a genetic and a physical map are routinely analyzed by subcloning into small insert libraries. This elaborate approach may be circumvented by direct outward sequencing with a set of six primers complementary to the repeat motif having one of the six theoretically possible 3' flanking bases (Yuille et al., 1991). Disadvantages of the latter procedure are that for both flanks twelve sequence reactions are required, and that ambiguous results can be obtained if the repeat is imperfect, compound or palindromic. We have developed an alternative method that is more generally applicable to analysis of flanking sequences of microsatellites and have compared the results with those of outward sequencing according to Yuille et al. (1991).

Large insert recombinants were obtained from a previously characterized (Rothuizen and Wolfs, 1994) genomic dog liver library in λ EMBL-3 SP6/T7 (Clontech, Palo Alto, USA). Plaque lifts were detected by chemiluminescence after hybridization with digoxigenin-labeled (CA)15 and ten positive clones were chosen at random. Microsatellite containing clones were grown in liquid culture and DNA was then isolated with Clontech (Princeton, NJ). Hybridization with a microsatellite probe and washing with buffer consisting of 10 mM Tris–HCl, pH 7.5, 1 mM EDTA, and 2 M NaCl. The isolated flanking region was denatured with 0.15 M NaOH for 5 min, and the non-biotinylated strand was removed by washing once with 0.15 M NaOH, and four times with the above wash buffer. Single stranded magnetospherically coupled DNA was then dissolved in 10 μl water and sequenced with the primer complementary to the adaptor, using a standard technique (T7 polymerase sequencing kit, Pharmacia/Biotech). Annealing was allowed for 20 min at 37°C. Samples were electrophoresed for 75 and 150 min. Results were compared with outward sequencing using six primers (CA)15 having a 3' anchoring base for each flank according to Yuille et al. (1991).

Nine of the ten clones gave a clear sequence of the microsatellite flank and part of the (CA)n(GT)n stretch itself following magnetic capture of the biotinylated (CA)n or (GT)n strand of the PCR product. Typically, some 30 bases of the microsatellite could be read; thereafter the sequence reaction lost its specificity due to the fact that annealing-out-of-range (Mariot and Vergnaud, 1992) of the non-anchored biotinylated PCR primer had produced a mixture of products of different length as visualized by agarose gel electrophoresis. The length of the sequence obtained from the end of the fragment to the start of the microsatellite varied from 48 to 220 bases (mean 92 bases),
The present technique permits the rapid isolation of microsatellite-containing regions of large insert recombinants and sequencing of the flanking regions. It is less amenable to ambiguities due to compound or imperfect constitution of the repeats because the sequence reaction is primed at the unique adaptor structure. Only microsatellites composed of palindromic repeats are theoretically not suitable for this approach, since this will result in two different biotinylated products in the PCR and a mixture of two templates in sequencing. In contrast, in methods in which the sequence primer anneals within the microsatellite, imperfections in the repeat may give priming at different sites. The present method is well suited to analysis of large numbers of clones, since restriction enzyme mix, adaptors, and primers for PCR and sequencing of the flanking microsatellites flanks may be routinely used for any clone containing a microsatellite with a given basic unit. The entire procedure can be completed within a day. The number of sequence reactions to obtain a flanking region is three times less than in the method of Yuille et al. We think that the most efficient use is in the sequence determination of one flank for which a flanking PCR primer towards the repeat is then developed. This primer permits a sequence reaction through the repeat to elucidate the structure of the other flank. The entire microsatellite and its two flanks can be analyzed in two sequence reactions instead of twelve as in outward sequencing. The present method gives the flanking sequence directly adjacent to the repeat, permitting the design of PCR primers very close to the microsatellite, and it is relatively insensitive to irregular or complex sequences within the microsatellite.

REFERENCES