

An improved protocol for genomic sequencing and footprinting by ligation-mediated PCR

Jean-Pierre Quivy and Peter B. Becker*

Gene Expression Programme, European Molecular Biology Laboratory, Meyerhofstr. 1, 6900 Heidelberg, Germany

Received March 10, 1993; Revised and Accepted April 23, 1993

In vivo footprinting requires the precise and sensitive mapping of nicks in highly complex genomic DNA. The Ligation-mediated PCR methodology (LM-PCR [1, 2]) has overcome sensitivity problems due to a selective amplification of the DNA fragments under study. The current procedure (3) employs three gene-specific oligonucleotides (P1–P3) to prime three DNA synthesis reactions (Figure 1): First, P1 is extended on denatured genomic DNA to convert all nearby nicks into blunt ends of double-stranded DNA. To these ends a small double-stranded linker is ligated. P2 is then used in conjunction with the long linker oligonucleotide for PCR amplification. Finally, radioactively-labelled P3 is extended to label the set of amplified fragments.

We have introduced various modifications to the current protocol (3) in order to improve the sensitivity, reliability of fragment representation and ease of primer design. First, we exploited the fact that the presence of polyethylene glycol (PEG) during ligations improves the efficiency of blunt end ligation (4). Linker ligation in the presence of 15% PEG resulted in a significant increase in signal and, importantly, improved the equal representation particularly of smaller fragments (data submitted but not shown). Another critical aspect in PCR-based genomic sequencing and footprinting experiments is the reliability of the polymerases involved to complete the various primer extension reactions. In a recent improvement of the procedure Garrity and Wold (5) used the thermostable *Thermococcus litoralis* (Vent) DNA polymerase for the first primer extension and the amplification step and observed an increased sensitivity as well as a more faithful representation of genomic fragments after amplification. We attributed our initial failure to use the Vent polymerase successfully to its known 3'–5' exonuclease activity which can act on double- and single-stranded DNA. Since the Vent polymerase is not recommended by the supplier for sequencing applications, but rather a mutant enzyme lacking the exonuclease activity (Vent Exo⁻) (6), we tested this enzyme for LM-PCR. Figure 2A shows the result of parallel LM-PCR reactions visualising the guanosine-specific cleavage pattern of the genomic *Drosophila hsp27* gene. The three DNA synthesis reactions (primer extension on genomic DNA, PCR amplification and labelling of amplified DNA) were carried out by all possible combinations of Vent and Vent Exo⁻ polymerases. 100 ng of nicked genomic DNA (purified according to ref. 7) and 0.5 pmole P1 (5' AAGCAATTATTAGCTGTC 3') were incubated in 19 μ l Vent buffer (10 mM KCl, 10 mM (NH₄)₂SO₄, 20 mM

Tris–Cl pH 8.8, 4 mM MgSO₄, 0.01% TritonC-100) for 5 min at 95°C and, after a quick spin, for 30 min at 50°C. 1 μ l (2 u) of Vent or Vent Exo⁻ DNA polymerase (Biolabs) was added, the reaction was incubated at 76°C for 10 min and then chilled. To this, 30 μ l of a premix containing 85.3 mM Tris–HCl pH 7.5, 16.6 mM MgCl₂, 16.6 mM DTT, 1.66 mM ATP, 42 μ g/ml BSA, 25.3% PEG 8000 and 100 pmoles of annealed linker (short oligo: 5' GAATTCAGATC 3', long oligo: 5' CACCCGGGAGATCTGAATTC 3') was added (PEG was added from a 40% stock solution in water, filtered through a 0.22 μ m filter). After addition of 1 μ l of T4 DNA ligase (400 u, Biolabs) the reaction was incubated overnight at 17°C. The reaction was diluted with 150 μ l TE, extracted with 150 μ l phenol/chloroform/isoamylalcohol (25:24:1) and precipitated. The pellet after centrifugation, 80% ethanol wash and drying was resuspended in 20 μ l of water. PCR amplifications were done on a Perkin-Elmer thermal cycler. 20 μ l of ligated DNA was added to 28.5 μ l of a premix containing 1.75 \times Vent buffer, 350 μ M each dNTP, 10 pmoles each of P2 (5' TAAGCGACATTTAAAAGCCCC-TTTG 3') and the long linker primer, followed by 1.5 μ l (3 u) of either Vent or Vent Exo⁻ polymerase. The reaction was incubated for 2.5 min at 95°C and then subjected to 20 thermocycles: 1 min at 95°C, 2 min at 60°C, 3 min at 76°C with a 5 sec increase in incubation time at 76°C per cycle. The amplification ended with a 10 min incubation at 76°C. The labelling was done with either enzyme according to the 'Linker Tag Selection' procedure described below using labelled P3 (5' TAAAAGCCCCCTTTGAATACGCCCGTAAT 3'). When all three reactions were performed with the Vent polymerase, a high background obscured the G-specific pattern, presumably due to partial degradation of the template or the newly synthesised DNA (Figure 2A, lane 1). The clearest G-specific pattern (lane 7) and the best yield of larger fragments was obtained when the Vent Exo⁻ polymerase was used in the initial primer extension on genomic DNA and in the final labelling reaction, but the Vent enzyme was employed during the exponential amplification step. In the latter reaction heterogeneous 3' ends after a DNA synthesis reaction will be repaired by the following primer annealing and extension. The Exo⁻ enzyme performed better for all single primer extension steps, where heterogeneity of fragment ends will spoil the original pattern. This is surprising in light of the recent observation that this enzyme adds an extra nucleotide to a fraction of the newly synthesised strands (6). Clearly, under

* To whom correspondence should be addressed

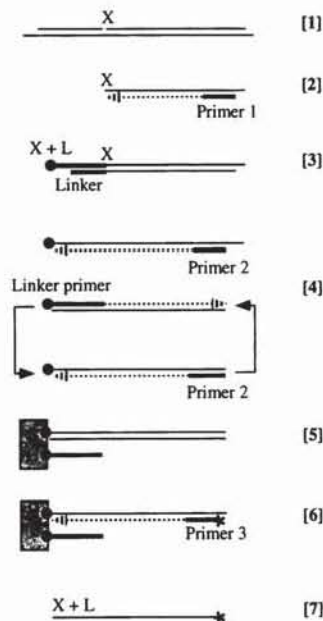


Figure 1. A schematic outline of the Linker Tag Selection LM-PCR procedure. X: Location of a nick in genomic DNA. Broken arrows: DNA synthesis reactions. Filled circle on the long linker primer: Biotin tag. L: long linker oligonucleotide. Hatched block: Streptavidin-coated paramagnetic bead.

our conditions, this does not seem to be the case. The optimal combination of both polymerases employs the fidelity and superior performance of Vent during PCR, the faithfulness of the Exo⁻ enzyme in the primer extension steps and the high reaction temperature common to both enzymes enabling stringent primer annealing conditions and minimising inhibitory secondary structures in the template. The performance of Vent polymerase in single primer extensions was not improved when the last (3') phosphodiester bond of the primer was substituted for a phosphorothioate (8, not shown).

Efficient labelling of the amplified strands by extension of P3 is only achieved when P3 competes efficiently with a large excess of P2 (from the PCR reaction) for annealing. Therefore, P3 is usually designed such that it overlaps P2 partially, but has a significantly higher melting point (T_m) than P2 (3). This requirement severely limits the options for the design of primer combinations that work, in particular when studying the AT-rich intergenic regions of *Drosophila*. To avoid competition between P3 and P2 during labelling we designed the 'Linker Tag Selection' (LTS) which employs a biotinylated linker oligonucleotide during ligation and PCR allowing the rapid immobilisation of the amplified DNA on paramagnetic streptavidin-coated beads via a magnet (Figure 1, [5]), effectively removing P2. Subsequent manipulations of the immobilised DNA are easy and quantitative since organic extractions and precipitations are eliminated and optimal primer extension conditions in a small volume are established reproducibly. The long linker oligonucleotide was biotinylated at its 5' end during the synthesis on a ABI 394 DNA synthesiser by incorporation of a biotin-2-o-propyl phosphoramidite. Concentrating and washes of the paramagnetic beads was performed with an MPC-E magnetic particle concentrator (Dynal) according to the manufacturer's

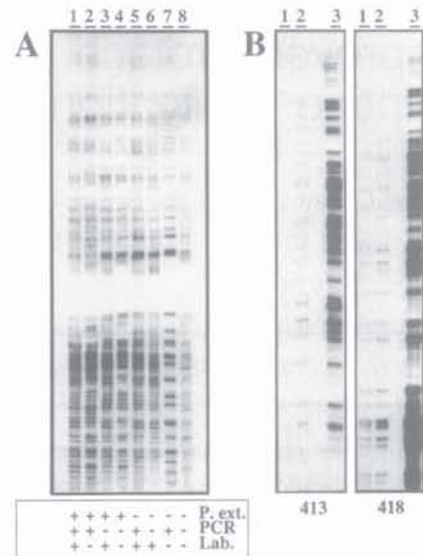


Figure 2. A. Optimal use of Vent and Vent Exo⁻ polymerases. Genomic *Drosophila* DNA that had been subjected to partial G-specific cleavage *in vitro* was analysed in duplicate LM-PCR reactions visualising the specific pattern of the hsp27 gene promoter. All possible combinations of Vent or Vent Exo⁻ polymerases were used for first strand extension on genomic DNA (P. ext.), PCR amplification (PCR) and the labelling step (Lab.). The use of Vent polymerase in a reaction is symbolised by (+) in the matrix, the use of Vent Exo⁻ by (-). B. A comparison between different LM-PCR strategies. Duplicate genomic DNA samples as in A were processed according to the modified protocol. Before the labelling step each reaction was split and aliquots were processed according to the three different LM-PCR strategies: (1) Standard protocol; (2) Standard protocol + gel filtration to remove P2 after the PCR amplification and (3) the Linker Tag Selection procedure.

specifications. Bead sedimentation during incubations was avoided by occasional gentle agitation. Care was taken to resuspend the beads well after every wash. 400 µg streptavidin-coated paramagnetic beads (Dynabeads M-280 streptavidin, Dynal) per sample were washed in bulk twice with 100 µl of PBS pH 7.4, once with 100 µl of PBS pH 7.4/0.1% BSA and twice with BW solution (2.5 M NaCl, 5 mM Tris-Cl pH 8.5, 0.5 mM EDTA) and finally suspended in 50 µl of BW solution. 20 µl of the PCR reaction were mixed with the bead suspension avoiding mineral oil contamination (a chloroform extraction was omitted as traces of this solvent have adverse effects on subsequent steps), and incubated for 30 min at room temperature on a rotator. The beads were washed with 100 µl of BW solution and resuspended in 100 µl freshly diluted 150 mM NaOH. Denaturation occurred upon incubation for 5 min at room temperature and a further 2 min at 50°. The complementary strands were removed by washing with 100 µl of 150 mM NaOH, and after further two washes with 100 µl TE (pH 8.5) and one with Vent buffer the beads were finally resuspended in 19 µl of labelling mix (Vent buffer plus 200 µM dNTPs and 0.2 pmoles labelled P3). The reaction was heated at 95° for 30 sec followed by a 30 min incubation at 52° (P413) or 69° (P418). 1 µl of Vent Exo⁻ was added and the reaction was incubated at 76°C for 10 min. After chilling on ice the beads were concentrated, washed in 100 µl of water and resuspended in 4 µl of a fresh 2:1 mix of formamide loading buffer (7) and 150 mM NaOH. Samples

and then chilled on ice. Beads were concentrated and 0.5–1 μ l of the supernatant was analysed on a 6% sequencing gel, which was dried and exposed to X-ray film.

Figure 2B illustrates the power of LTS. Parallel LM-PCR reactions were performed as above with either labelling primer 418 (27 nt long, 59% GC, $T_m = 68^\circ$), which overlaps and competes efficiently with P2 (25 nt long, 56% GC, $T_m 65^\circ$), or P413, which cannot be used for labelling in standard reactions due to its poor features (19 nt long, 31% GC, $T_m = 50^\circ$, no overlap with P2). When the standard protocol (3) was used (Figure 2B, lanes 1) a G-specific ladder could be detected with P418 after a 12 hour exposure, but not at all with P413. Removal of the competing P2 by gel filtration through a spin column prior to labelling (lanes 2) improved the signal from P413 about 3 fold but not significantly with P418. The Linker Tag Selection procedure (lanes 3) resulted in a further 10 fold increased signal regardless of the specific features of the labelling primer. This was due to both removal of competing P2 and alkali denaturation of immobilised DNA, allowing for the removal of the complementary strands and efficient primer extension on single-stranded DNA. It should be emphasised that the 'Linker Tag Selection' does not only quantitatively improve the labelling reaction, but introduces a new quality into the protocol since under conventional conditions a primer like P413 could not be used as labelling primer at all. Our procedure differs from the previously published 'Extension Product Capture' LM-PCR (9), which uses a biotinylated P1 for isolation of the products of the first primer extension by streptavidin-beads in order to reduce the complexity of the DNA in the PCR amplification. In principle, both strategies are complementary and could be combined for optimal results.

In summary, we have refined the LM-PCR protocol by improving the linker ligation conditions, by optimising the use of heat-stable polymerases for each DNA synthesis step and by introducing the 'Linker Tag Selection'. The new protocol (a step-by-step version is available upon request) increases the sensitivity of the LM-PCR significantly, improves the faithfulness of band representation and creates freedom in primer design. We have used it successfully with a variety of primers on both the *Drosophila* hsp26 and hsp27 gene promoters.

ACKNOWLEDGEMENTS

We thank Drs G.Pfeifer, G.Rigaud and T.Gränge, J.R.Jackson and C.Benyajati for LM-PCR protocols, M.Truss and G.Dahan for helpful discussions, G.Wall and Dr F.Stewart for comments on the manuscript and the EMBL oligonucleotide service for perfect performance. J.-P.Q. was supported by a fellowship from the European Molecular Biology Organisation.

REFERENCES

1. Mueller,P.R. and Wold,B. (1989) *Science* **246**, 780–810.
2. Pfeifer,G.P., Steigerwald,S.D., Mueller,P.R., Wold,B.J. and Riggs,A.D. (1989) *Science* **246**, 810–813.
3. Mueller,P.R., Garrity,P.A. and Wold,B.J. (1992) In, Ausubel *et al.* (eds), *Current Protocols*, **Supplement 20**, Wiley Interscience, New York.
4. Pfeiffer,B.H. and Zimmermann,S.B. (1983) *Nucleic Acids Res.* **11**, 7853–7871.
5. Garrity,P.A. and Wold,B.J. (1992) *Proc. Natl. Acad. Sci USA* **89**, 1021–1025.

1965–1975.

7. Becker,P.B. and Schütz,G. (1988) In, Setlow,J.K. (ed.) *Genetic Engineering*. **10**, 1–19, Plenum Press, New York.
8. Skerra,A. (1992) *Nucleic Acids Res.* **20**, 3551–3554.
9. Törnänen,V.T., Swiderski,P.M., Kaplan,B.E., Pfeifer,G.P. and Riggs,A.D. (1992) *Nucleic Acids Res.* **20**, 5487–5488.