Overcoming limitations of the mRNA differential display technique

Lauren Sompayrac*, Stephen Jane, Timothy C. Burn¹, Daniel G. Tenen² and Kathleen J. Danna

Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, CO 80309, USA,

¹Department of Human Genetics, Integrated Genetics, Inc., Framingham, MA, USA and ²Hematology/Oncology Division, Beth Israel Hospital and Department of Medicine, Harvard Medical School, Boston, MA, USA

Received July 20, 1995; Revised and Accepted October 12, 1995

Differential mRNA display (DD) is potentially a powerful method for identifying genes that are over- or under-expressed in one cell type relative to another. In this technique, as first described by Liang and Pardee (1), total RNA is isolated from two cell types to be compared, and first strand copies of both RNAs are made by reverse transcription, using an oligo-dT primer that has a specific dinucleotide at its 3' end (the 3' amplimer). This 3' amplimer and an arbitrarily-chosen 10mer (the 5' amplimer) are then used in the polymerase chain reaction (PCR) to amplify cDNAs to which the 3' and 5' amplimers both hybridize. A radioactive nucleotide is included in the PCR reactions so that the products can be run side-by-side on a sequencing gel and visualized by autoradiography. Bands that appear on the display of RNA from one cell type but not the other correspond to differentially-expressed mRNAs. These bands can be excised from the gel and amplified with the same primers used for the original display, and the resulting PCR product can be used to probe Northern blots of polyA⁺ RNA to confirm the identification of a differentially-expressed mRNA.

Although this technique has been used to identify differentially-expressed genes, two serious problems have limited its usefulness. First, the pattern of differential expression seen on the 'display' often cannot be reproduced on Northern blots (2,4). These false positives can arise with a frequency >70% (4). A second difficulty is that the cDNAs obtained represent only ~300 bp at the extreme 3' end of the mRNA (usually in the 3' untranslated region). Since this part of the mRNA is frequently not included in GenBank and varies greatly from organism to organism, cDNAs identified by DD cannot always be matched with genes that have already been characterized. To obtain the translated region of the mRNA, experimenters have resorted to the time-consuming process of screening cDNA libraries (2,4), which also can be problematic, because the 3' untranslated regions of different genes can be quite similar.

We have modified the original DD protocol to reduce the frequency of false positives, and have devised a strategy that allows us to easily walk, in steps of ~1 kb, along the mRNA toward its 5' end. Although we have used this walking technique only in connection with DD, our approach should be useful in other situations when access to sequences 5' to known mRNA sequences is desired.

To reduce the number of false positive cDNAs that arise with DD, we use the following strategy. (i) We either a) display PCR products from two 'uninduced' cell lines and two 'induced' lines,

and require that the patterns from the pairs of uninduced or induced lines agree or b) display PCR products from uninduced and induced lines over a time course of induction (5). This redundancy decreases the number of artifactual PCR or RT results which must be tested further. (ii) When possible, we use cytoplasmic RNA rather than total RNA in our RT reactions, thus eliminating false positives due to nuclear RNA that is not transported to the cytoplasm. (iii) To reduce the total number of PCR reactions required for a representative display, we combine the 12 possible 3' amplimers into three sets of four amplimers (3P1: T¹²AA, T¹²AT, T¹¹AC, T¹¹AG; 3P2: T¹¹GA, T¹¹GT, T¹¹CA, T¹¹CT; 3P3: T¹⁰GC, T¹⁰GG, T¹⁰CC, T¹⁰CG). We perform RT reactions with a given set of 3' amplimers (e.g., 3P1), and then use this same set of 3' amplimers plus a 5' amplimer for the PCR reactions. These sets of 3' amplimers have annealing temperatures that are more similar within each set than those recommended in the revised protocol of Liang *et al.* (3).

On the relatively infrequent occasions when the initial display reveals a putative differentially-expressed mRNA, we repeat the PCR reactions with the appropriate 5' amplimer together with each of the individual 3' amplimers that were combined for the first DD. If the bands seen in the first DD are not artifactual, this 'breakdown' of the set of 3' amplimers should reproduce the pattern of differential expression for only one of the 3' amplimers.

Repeating the DD with individual 3' amplimers reduces false positives in several ways. In addition to acting as a further screen for PCR artifacts, the breakdown identifies PCR reactions in which the 5' amplimer acts as primer for both ends of the cDNA, since, in such a case, an identical pattern will appear for each of the individual 3' amplimers. This situation can arise, for example, when the expression of rRNA is upregulated in the induced lines, and rRNA acts as a template for reverse transcription, independent of the 3' amplimer. The breakdown also eliminates false positives that arise when an mRNA present, for example, in one of the induced cell lines is amplified by a 3' amplimer, and produces a cDNA that co-migrates with a cDNA that results when a different mRNA, present in the other induced line, is amplified by one of the other three members of the pool of 3' amplimers. Moreover, because the breakdown display will have roughly one quarter as many bands as the original, it is less probable that a co-migrating cDNA will outcompete the desired cDNA when a slice of the differential display gel is excised and the cDNAs it contains are amplified prior to cloning.

* To whom correspondence should be addressed

Figure 1. Strategy to obtain additional 5' sequence.

We have used DD to look for cellular genes whose expression is modulated in SV40-transformed cells. In our early experiments, we used essentially the protocol of Liang and Pardee (3). In our latest experiments, we used the modified protocol described here. Our results are summarized in Table 1, which documents the substantial decrease in the frequency of false positives that can be expected when our modified protocol is used.

Table 1. Comparison of old and new protocols

	Old ^a	New ^b
Amplimer pairs tested	65	96
Putative differentially expressed genes examined by Northern blot	20	10
Pattern of differential expression confirmed by Northern blot	2	8
Percent confirmed by Northern blot	10%	80%

^aLiang and Pardee protocol (3). ^bOur modified protocol.

To obtain additional sequence information 5' of the cDNA fragment that is cloned from the DD gel, we use the strategy depicted in Figure 1. In brief, we use reverse transcriptase and a primer (RT in Fig. 1a) located near the 5' end of the known sequence to make a single-stranded copy of mRNA from a cell in which the gene of interest is expressed. We then make a second strand copy of this cDNA (Fig. 1b), ligate the ends of this double-stranded cDNA to form circles (Fig. 1c), and use PCR primers, PA and PB, to amplify the sequences that are 5' to the known sequences (Fig. 1d). Finally, we clone these PCR products, and determine their sizes by gel electrophoresis. The cDNAs cloned by this method will have a range of sizes that depends on the lengths of the original RT products, but a set of 12 minipreps usually yields several cDNAs that are ~1 kb in length, and which represent sequences 5' to the known mRNA sequence. Once the 5' end of this newly-cloned cDNA has been sequenced, primers can be designed to continue the walk toward the 5' end of the mRNA. We have used this procedure in five experiments to achieve 'walks' of 0.8, 1.3, 0.7, 1.0 and 1.6 kb. A similar strategy has been described as an alternative to the 5' RACE technique (6), but only ~200 bp of new 5' information was obtained.

The detailed protocol for walking is as follows. Mix 2 µg polyA⁺ RNA with 100 pmol of a gene-specific primer (18mer) in

a total volume of 10 µl, incubate at 70°C for 10 min, and transfer to ice. Add 4 µl of 5× 1st strand buffer (supplied with Superscript II enzyme), 2 µl of 0.1 mM DTT, and 1 µl of a mix of all four dNTPs (each at 10 mM); incubate for 2 min at 45°C; add 3 µl of Superscript II reverse transcriptase (BRL, 200 U/µl); and continue the incubation for 60 min. To synthesize the second strand, transfer the sample to ice; add 90 µl dH₂O, 30 µl of 5× 2nd strand buffer [100 mM Tris (pH 6.9), 450 mM KCl, 23 mM MgCl₂, 0.75 mM β-NAD, 50 mM (NH₄)₂SO₄], 3 µl of 10 mM dNTP mix, 5.6 µl *Escherichia coli* DNA polymerase I (Promega, 7 U/µl), and 1 µl *E. coli* RNase H (BRL, 2 U/µl); and incubate at 16°C for 2 h. To blunt the ends of the DNA, add 2 µl T4 DNA polymerase (5 U/µl), and continue the incubation for an additional 5 min. Add 20 µl of 0.25 M EDTA, extract with 180 µl of phenol:chloroform:isoamyl alcohol (25:24:1), and precipitate the DNA with ethanol after adding 20 µg glycogen. Spin down the precipitate, wash it with 70% ethanol, resuspend it in 70 µl TE, and subject the cDNA to exclusion chromatography on a CHROMASPIN+TE-1000 column (Clontech) to eliminate short RT products.

Phosphorylate 54 µl cDNA (out of a total of 70 µl) by adding 6.75 µl of 10 mM ATP, 6.75 µl of 10× KL buffer [0.66 M Tris (pH 7.2), 10 mM spermidine, 0.1 M MgCl₂, 0.15 M DTT, 1.35 mg/ml gelatin], and 3.4 µl of T4 kinase (10 U/µl). Incubate for 30 min at 37°C and then for 5 min at 65°C. Ligate the kinased cDNA to form circles by adding 82 µl dH₂O, 40 µl of 5× ligase buffer (BRL) and 8 µl T4 DNA ligase (BRL, 1 U/µl). After an overnight incubation at room temperature, extract the sample with phenol:chloroform:isoamyl alcohol, ethanol precipitate the DNA with 20 µg glycogen, and resuspend the precipitate in 10 µl TE.

Perform a 40 µl PCR reaction that includes 2 µl of the circularized DNA, 2 µl each of primers PA and PB (10 µM), 4 µl of 10× PCR buffer (supplied with Amplitaq polymerase), 3.2 µl MgCl₂ (25 mM), 0.8 µl of a mixture of dNTPs (10 mM each), and 1.3 µl Amplitaq polymerase (Perkin Elmer, 5 U/µl). Amplify as follows: 30 s at 94°C, 30 s at an annealing temperature 3°C below the primer melting temperature, and 3 min at 72°C. After 30 cycles, extend for an additional 5 min at 72°C.

To clone the PCR product, add TE to bring the volume to 69 µl, add 1 µl of glycogen (10 µg/µl), and subject the DNA to exclusion chromatography on a CHROMASPIN+TE-1000 column. To 4 µl of the column eluate (out of 70 µl), add 1 µl pGem-T vector (Promega, 50 ng/µl), 3 µl dH₂O, 2 µl BRL 5× ligase buffer, and 1 µl T4 DNA ligase (BRL, 1 U/µl). Incubate the sample overnight at room temperature, add 40 µl dH₂O, and use 1 µl of the diluted ligation reaction for bacterial electroporation.

ACKNOWLEDGEMENTS

This work was supported by a grant to L.S. from the Council for Tobacco Research (#3307) and NIH grants to L.S. (CA-53322) and D.T. (DK-48660).

REFERENCES

- Liang, P. and Pardee, A.B. (1992) *Science*, **257**, 967-971.
- Liang, P. et al. (1992) *Cancer Res.*, **52**, 6966-6968.
- Liang, P. et al. (1993) *Nucleic Acids Res.*, **21**, 3269-3274.
- Sun, Y. et al. (1994) *Cancer Res.*, **54**, 1139-1144.
- Burn, T.C. et al. (1994) *Blood*, **84**, 2776-2783.
- Zeiner, M. and Gehring, U. (1994) *Biotechniques*, **17**, 1051-1053.