4272–4280 Nucleic Acids Research, 1993, Vol. 21, No. 18

Identification of differentially expressed mRNA species by an improved display technique (DDRT-PCR)

David Bauer[1], Heiko Müller[1], Jens Reich[2], Heidemarie Riedel[1], Vibeke Ahrenkiel[3], Peter Warthoe[3] and Michael Strauss[1,3,*]

[1]Max-Planck-Gesellschaft and Humboldt University and [2]Max Delbrück Center for Molecular Medicine, Robert-Rössle-Str. 10, O-1115 Berlin-Buch, Germany and [3]Danish Cancer Society, Division of Cancer Biology, Strandboulevarden 49, DK-2100 Copenhagen, Denmark

Received May 28, 1993; Revised and Accepted August 4, 1993

ABSTRACT

We have significantly improved a method originally developed by Liang and Pardee [Science 257 (1992) 967–971] to display a broad spectrum of expressed genes and to detect differences in expression between different cell types. We have analysed various aspects of the technique and have modified it for both, the application to fast and efficient identification of genes and the use with automatic analysis systems. Based on the mathematical background we have devised the appropriate number of optimal PCR primers. We have also introduced nondenaturating gels for separating double stranded fragments as single bands. By applying the method to regenerating mouse liver, we have identified, out of a total of 38,000 bands, about 70 fragments where the expression of the corresponding genes seems to be differentially regulated at different time points. Application of the method to an automatic DNA sequencer was successfully done. Thus, we have confirmed the usefulness and increased the power of the RNA display technique, which we named differential display reverse transcription PCR (DDRT-PCR), and have extended the range of its application.

INTRODUCTION

The analysis of changes in gene expression in cells which underwent a particular step in differentiation, dedifferentiation or carcinogenesis is of prime interest in molecular biology. Until recently, there were only two alternative approaches to tackle this kind of question. Firstly, subtractive hybridization (1) or differential hybridization (2) are used to identify genes which are expressed in only one cell type of the respective pair of cells. These are mainly qualitative methods which do not allow to detect quantitative changes. In addition, these methods are time consuming and not always satisfying. Secondly, nuclear run-on transcription is applied to analyse changes in the level of expression (3). However, this method can only be applied to the detection of changes in the expression of known genes. Thus, it would be important to have a method which detects all mRNA species expressed in a particular cell. By comparing the patterns of expressed mRNAs from two cell types one should be able to detect both, qualitative and quantitative changes. This kind of method would allow not only to identify new genes but also the diagnosis of any changes in gene expression involved in a particular cellular process.

A method recently published by Liang and Pardee (4) is based on the assumptions that every cell expresses some 15,000 genes and, in principle, every individual mRNA molecule can be reverse transcribed and amplified by the polymerase chain reaction. Their idea was to use a set of arbitrary primers for PCR amplification of cDNA generated by reverse transcription from mRNA.

Every pair of primers has a certain chance of identifying a limited number of target sequences within the pool of cDNAs. The PCR products are resolved on a sequencing gel. In order to display at least 15,000 bands one would need about 100 tracks with an average of 150 bands each. This was the basic idea and the method was designed accordingly. One major trick was to subdivide the total cDNA into 12 fractions with nearly equal numbers of represented mRNA species.

We have extended and improved the RNA display technique in different respects. First of all, we analysed theoretically and experimentally the potential of the method to actually identify every mRNA species expressed. We devised a representative collection of primers and have adopted the method for automatic analysis. Nondenaturing gels were introduced to reduce an artificial complexity of the band pattern. We have named the method, in accordance with the designation used by Liang and Pardee, differential display reverse transcription polymerase chain reaction (DDRT-PCR).

MATERIALS AND METHODS

Preparation of cytoplasmic RNA

For a standard preparation one 10 cm dish of cells ($2-4\times10^6$) was used and washed with PBS. Alternatively, small pieces of mouse liver were taken before and after hepatectomy and

* To whom correspondence should be addressed at: Max-Planck-Gesellschaft, Max-Delbrück-Haus, Robert-Rössle-Str. 10, D-1115 Berlin-Buch, Germany

disintegrated. Since most protocols for RNA preparation are appropriate for this purpose we do not describe our procedure in detail. Most likely, that the use of cytoplasmic RNA is advantagous. Our method is based in NP-40 lysis and the isolation of cytoplasmic RNA according to (5).

Reverse transcription of RNA

Reverse transcription was done in 12 independent reactions using $dT_{11}VN$ primers essentially as described (4). For each reaction 0.1−0.2 μg of RNA, 2.5 μM $dT_{11}VN$ primer, 20 μM dNTP and 300 U MMLV reverse transcriptase were mixed and incubated at 35°C for 60 min. The reaction was stopped by incubation at 95°C for 5 min. Total cDNA was alternatively generated in one reaction using oligodT (12−18) as the primer. For PCR incubations this primer was removed by microdialysis.

Amplification of cDNA

(i) Primers. 10mer deoxyoligonucleotide primers with arbitrary sequences were selected from a list of random sequences considering the following prerequisites: (i) presence of 50% G+C and A+T, and (ii) absence of uninterrupted self-complementarity of more than two nucleotides. Thirteen primers have a GATC sequence at the 5′ end in common (Tab. 1).

(ii) Standard reaction for radioactive detection. In a total volume of 20 μl the PCR mixtures contained 2.5 μM of $dT_{11}VN$ primer, 0.5 μM of the respective upstream primer (see table 1), 2 μM of dNTPs, 2 μCi (60 nM) ^{33}P-dATP, 10 mM Tris−HCl (pH 8.3), 50 mM KCl, 1.2 mM $MgCl_2$, and 1 U Taq polymerase. Amplification was done for 40 cycles with 94°C for 30 sec., 42°C for 1 min., and 72°C for 30 sec. and an additional extension period at 72°C for 5 min.

(iii) Reaction mixture for automatic analysis. For analysis on a sequencing automate the incubation was carried out in two steps. At first, the reaction was set up as under (ii) but run only for 20 cycles. Afterwards, to 1/10 of the reactions upstream primers were added which were modified by addition of either one of the dyes TAMRA, ROX, FAM, or JOE according to the protocols provided by the supplier (Applied Biosystems). The primer concentrations were now changed to be 0.5 μM for the labeled upstream primer and 2.5 μM for the downstream primer. The dNTP concentration was raised to 20 μM.

(iv) Experimental setup for fast and efficient analysis. Incubation mixtures were prepared in 3 microtiter plates (for one cell type) or microtiter-formatted frames with tubes (Perkin-Elmer) by adding 15 μl premixed components corresponding to the 12 cDNA pools from top to the bottom into the wells of each of the 24 rows using a 12-channel micropipette and adding 5 μl of the 24 upstream primers from left to the right to the 12 rows using an 8-channel micropipette. For comparative analysis of two or more cell types an alternative array of the plates and primers turned out to be more convenient.

Gel electrophoresis of DNA fragments

(i) Denaturing gels. PCR mixtures were dried using a Speed Vac and a rotor for microtiter plates. Samples were redissolved in 5 μl of 90% formamide/dye solution, heated for 2 min, and 1.5 μl were run on standard 6% polyacrylamide-urea gels in Tris-borate buffer (6) at 80 W. Standard electrophoreses with radioactively labeled fragments were run in the S2 model from BRL, electrophoresis of unlabeled fragments was done in the blotting-electrophoresis machine from GATC, and electrophoresis of dye-labeled fragments was carried out in the ABI 373A DNA sequencer (Applied Biosystems).

Table 1. List of all 5′ primers used for a complete analysis of expressed RNA by RT-PCR.

No.	Sequence (5′ to 3′)									
1.	T	A	C	A	A	C	G	A	G	G
2.	T	G	G	A	T	T	G	G	T	C
3.	C	T	T	T	C	T	A	C	C	C
4.	T	T	T	T	G	G	C	T	C	C
5.	G	G	A	A	C	C	A	A	T	C
6.	A	A	A	C	T	C	C	G	T	C
7.	T	C	G	A	T	A	C	A	G	G
8.	T	G	G	T	A	A	A	G	G	G
9.	T	C	G	G	T	C	A	T	A	G
10.	G	G	T	A	C	T	A	A	G	G
11.	T	A	C	C	T	A	A	G	C	G
12.	C	T	G	C	T	T	G	A	T	G
13.	G	T	T	T	T	C	G	C	A	G
14.	G	A	T	C	A	A	G	T	C	C
15.	G	A	T	C	C	A	G	T	A	C
16.	G	A	T	C	A	C	G	T	A	C
17.	G	A	T	C	T	G	A	C	A	C
18.	G	A	T	C	T	C	A	G	A	C
19.	G	A	T	C	A	T	A	G	C	C
20.	G	A	T	C	A	A	T	C	G	C
21.	G	A	T	C	T	A	A	C	C	G
22.	G	A	T	C	G	C	A	T	T	G
23.	G	A	T	C	T	G	A	C	T	G
24.	G	A	T	C	A	T	G	G	T	C
25.	G	A	T	C	A	T	A	G	C	G
26.	G	A	T	C	T	A	A	G	G	C

(ii) Native gels. The total PCR incubation mixture was adjusted with glycerol to 5%, applied onto a 6% polyacrylamide gel without urea, and run in Tris-borate buffer at 60 W.

Detection and analysis of band patterns

(i) Radioactive detection. Gels were dried directly on the glass plates or transferred to filter paper before and were exposed to X-ray film overnight. Autoradiographs were scanned using the Howtek Scanmaster.

(ii) Automatic recording of bands. Samples run in the ABI 373A DNA sequencer were automatically recorded using the Genescan software from ABI.

Amplification and sequencing of DNA fragments

For further characterization of individual bands nondenaturing gels were autoradiographed. Bands of interest, e.g. those differing between the patterns from control cells and the test cells, were cut out and were transferred into Eppendorf tubes and the DNA was reamplified in a reaction volume of 50 μl for 20 cycles under the same conditions as before. After adding primers to 10 μM and dNTPs to 100 μM final concentration, amplification was continued for further 20 cycles. DNA sequence analysis (7) was carried out using the fmole-sequencing kit (Promega).

THEORETICAL CONSIDERATIONS

The idea of displaying a great portion of the expressed mRNA molecules of a particular cell type was solved by Liang and Pardee (4) using a PCR approach which is based on a suitable number

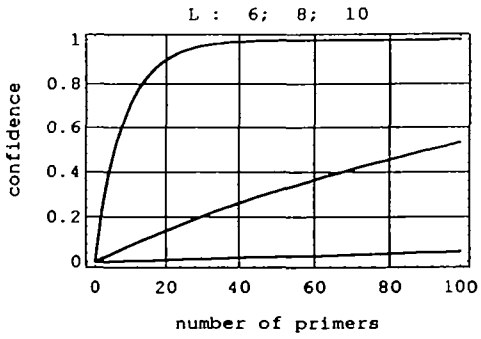

Figure 1. Confidence level that for X primers of the length L any given messenger will be represented by at least one band in a gel. Note that the 95% confidence level for L=6 corresponds to X = 25.

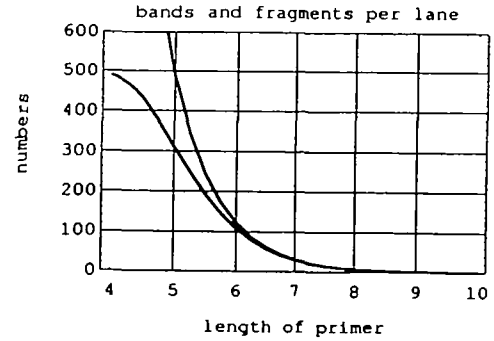

Figure 2. Expected number of bands and of PCR fragments present in one lane of a gel. The number of bands/fragments is shown as a function of the length L of the upstream primer if N = 1000.

of arbitrary primers having a statistical chance of finding a corresponding sequence within the mRNA or cDNA, respectively. If one assumes 15,000 genes being expressed one would like to see at least the same number of bands on a gel. Liang and Pardee (4) have subdivided the total mRNA into 12 fractions by performing the reverse transcription with oligo-dT primers having two additional nucleotides in all possible combinations (excluding T as the penultimate nucleotide) at their 3′ end. In each of the 12 fractions one would expect approximately 1250 individual mRNA species to be present.

On one lane of a sequencing gel one can easily detect some 150 individual bands if not uniformly distributed. This means that one PCR incubation should result in about this number of bands corresponding to individual cDNA species. Provided suitable primers are used 10 to 12 PCR incubations of one mRNA/cDNA fraction would be sufficient to generate this number of bands. However, there are several theoretical problems to be solved. In the following we will give the mathematical background for choosing the length and the number of primers.

Formal treatment

The list of symbols used in this treatment is as follows:

M— number of discernable positions in a lane (approximately 500), each corresponding to a PCR-product of specified length

N— number of cDNA-species produced by one downstream primer (750 to 2000)

X— number of upstream primers applied in independent PCR runs

L— length of an upstream primer

p— probability that a certain upstream primer finds its complementary sequence at a specified position of a given cDNA (if we assume the chance of occurrence at each position of both the primers and the templates for all bases to be 1/4, p would be given by p = 4 ^ (−L))

q— complement to p (q = 1−p)

Consider a specified upstream primer and a given cDNA species. For simplicity, we assume its length to be >M. The probability of priming at least once within the sequence of M nucleotides which could give rise to discernable PCR products is given by

$$1-q\ \hat{}\ M.$$

For a primer with the length of L = 6 this probability would be

$$1-(1-4\ \hat{}\ (-6))\ \hat{}\ 500 = 0.115$$

This example suggests that not one but X independent primers be applied. As long as X << 4 ^ L the probability that one cDNA gives at least one band in X different lanes is

$$1-q\ \hat{}\ (M * X).$$

This relationship is depicted in fig. 1.

One gel band may have been produced from more than one messenger. The probability of k cDNA species beeing present in one position of a lane follows the binomial distribution with the parameters (p, N). The expected mean value of the number of cDNA species per position thus evaluates to

$$N * p.$$

The total number of cDNA's per lane is

$$N * p * M.$$

On the other hand, the probability that a position in a lane is not empty can be calculated from

$$1-q\ \hat{}\ N$$

and the expected number of bands per lane is therefore

$$M * (1-q\ \hat{}\ N).$$

Figure 2 shows the expected number of bands and PCR products, respectively, as a function of L if N = 1000.

The predictions shown in figure 2 are clearly in contradiction to what was observed experimentally. Although a smaller oligonucleotide should display more bands than a longer one, Liang and Pardee (4) detected no band with primers of up to octamers, detected 20 to 30 bands with nonamer primers, and 50 to 100 with the decamer, a paradox. These authors explain this observation by the assumption that the primers hybridize in a degenerate way and, as a decamer has more subsequences capable of stable hybridazation than a nonamer, it displays more bands, accordingly. Sequence comparisons between upstream primers and their putative binding sites in known templates show that terminal mismatches are frequently observed (4,8; and our data). There is little reason to assume, however, that the stability of a hybrid with a nonamer with two terminal mismatches should be far below the stability of a hybrid of a decamer with three terminal mismatches. Computer modeling shows that fragments

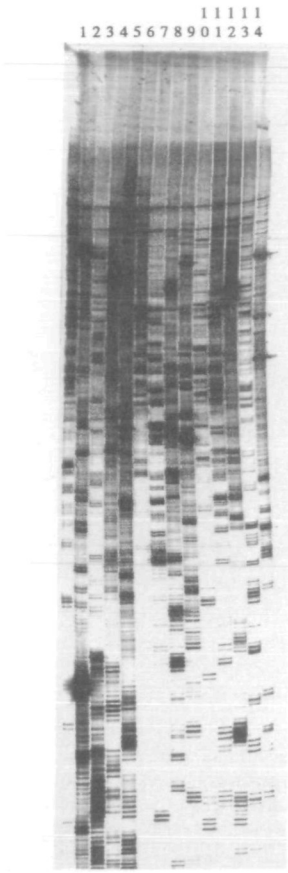

Figure 3. Band pattern derived from RNA of human IMR-90 cells using the RT-PCR technique. cDNA fraction No. 1 (obtained by reverse transcription with primer $T_{11}CA$) and the 5′ primers No. 1 to 14 were used to amplify the cDNA with annealing at 42°C and samples were run on a standard sequencing gel.

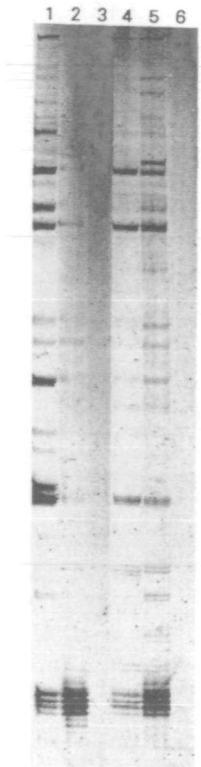

Figure 4. Effect of magnesium and nucleotide concentrations on the complexity of band patterns. Nucleotide concentrations were 2 μM (lanes 1−3) and 4 μM (lanes 4−6). $MgCl_2$ concentrations were 1.25 mM (lanes 1 and 4), 2.5 mM (lanes 2 and 5), and 5 mM (lanes 3 and 6). The PCR incubation was done with cDNA fraction No. 1 as in figure 3 and 5′ primer No. 1 with 40 C as the annealing temperature.

originating from mismatched priming events can be more abundant than fragments from perfect matches, depending on the number of priming events (matched or mismatched) further upstream, even though their priming probability is smaller by two orders of magnitude (data not shown).

Our results described below show that the PCR with decamer primers generates about 120 bands per lane after 40 cycles using an annealing temperature of 40°C. This corresponds to an effective primer length of $L' = 6$ (Fig.2). In order to reach a confidence level of about 0.95 for displaying any messenger as a cDNA fragment on the gel one should use at least 25 upstream primers (Fig. 1).

RESULTS

Chosing suitable upstream PCR primers

On the basis of the theoretical considerations we selected 26 oligonucleotide primers from a list of 50 arbitrary sequences generated by the computer and to test them empirically. These primers were selected to have GC-contents and AT-contents of 50% each. Among these primers we disregarded those exhibiting uninterrupted selfcomplementarity between more than two nucleotides. Thirteen of the primers were chosen to have an identical 5′ end with the sequence GATC in order to find out if this would affect the randomness in the selection of target sequences. After testing 50 potentially suitable primers (not shown) we finally selected 26 primers (Tab. 1).

Using the original procedure of Liang and Pardee (4) we generated 12 cDNA fractions from the RNA of human IMR-90 cells. All cDNA fractions were amplified at 42°C with the first 24 upstream primers using the standard procedure with radioactive labeling. Aliquots of all 288 PCR incubations were run on 6 denaturing gels with 48 samples each. The result is shown in figure 1. First of all, every pair of primers generated a distinct pattern of bands which could be reproduced in a second independent experiment with only slight variations in the intensity of about 5% of the bands (not shown). After scanning the autoradiographs the number of bands in each lane was determined using the Quantity One software from PDI, Inc. Although, the total number of bands in all 288 lanes was 20,160, this is certainly not representing the total number of expressed RNA species. The same experiment was done with RNA isolated from mouse liver. As expected, the band patterns generated were very different from those obtained from the human fibroblasts. However, the total number of bands was very similar.

From these initial results two observations were striking. Firstly, the average number of bands per lane was only 70 with a maximal number of about 140 bands. Thus, conditions should be found which would result in a higher number of bands per lane. Secondly, many fragments appeared to be doublets or even sets of three or four bands with nearly the same intensity suggesting them to represent the two strands of one fragment

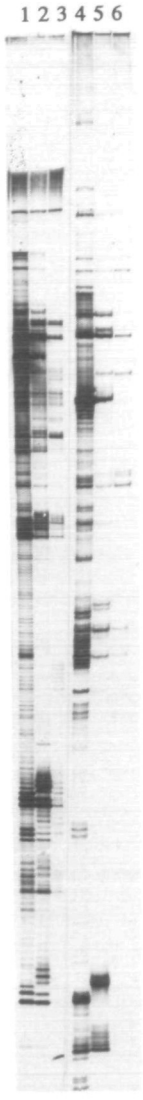

Figure 5. Comparison of denaturating and nondenaturating gel electrophoresis. PCR incubations of the first cDNA fraction with the 5′ primers No. 2, 3 and 7 were run in denaturating gels at 80W (lanes 1–3) and in nondenaturating gels at 60W (lanes 4–6). The numbers of bands in the nondenaturating gel are 186 (lane 4), 72 (lane 5), and 65 (lane 6). Most of the latter are very faint.

and molecules with or without an additional A known to be added by Taq polymerase at the 3′ ends. This assumption was confirmed by eluting some of these bands and sequencing them (not shown). This means that the actual number of genes represented in the pattern was even lower. Therefore, we tried to optimize several parameters.

Optimizing the incubation and electrophoresis conditions

Several parameters were varied in parallel incubations using three upstream primers (No. 1, 2, 3) and one cDNA fraction (No. 1): (i) annealing temperatures 36°, 38°, 40°, 42°, and 44°C; (ii) dNTP concentrations between 0.5 μM and 5 μM; (iii) magnesium concentrations between 1.0 and 5.0 mM. The effect of the annealing temperature was considerable. Above 42°C the number of bands decreases dramatically and below 40°C the lanes are nearly filled with bands in every position (data not shown). The optimal magnesium concentration varied slightly with the primers

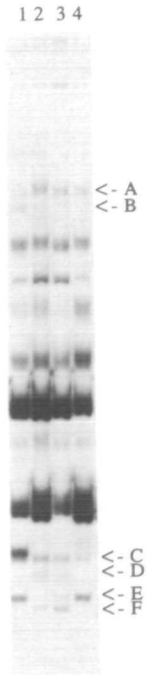

Figure 6. Comparative analysis of band patterns generated with RNA from different stages of liver regeneration. The figure shows part of a gel in which several differences were identified. Lane 1, normal liver; lane 2, liver at 24 h after hepatectomy; lane 3, 48 h after hepatectomy; lane 4, 72 h after hepatectomy. Bands corresponding to differently expressed genes are marked A to F.

and was also dependent on the concentration of nucleotides (Fig. 4). However, 1.25 mM $MgCl_2$ seemed to be a good compromise for a dNTP concentration of 2 μM.

In order to avoid the complexity of the patterns on denaturing polyacrylamide gels caused by the several bands derived from one DNA species we have run the fragments on nondenaturing gels. The comparison of identical PCR samples in both gel systems (Fig. 5) clearly shows the reduced complexity in nondenaturating gels and favors their use.

Using these optimized conditions, the analysis was done with 4 different preparations of RNA from mouse liver corresponding to different time points of regeneration (see below). We have run 4×312 PCR reactions with all 26 upstream primers and all 12 cDNA fractions and have counted the number of bands detectable on the film after an overnight exposure. The total number of bands is 38,004 (Tab. 2). Each figure represents the mean value of the four different RNA samples. These figures were highly reproducible with less than 5% deviation which was most likely due to real differences in the pattern of expressed genes. The variation in band numbers is remarkably low.

Are all expressed genes detectable?

It seems obvious from the formal considerations that theoretically all expressed genes should be detectable using this method. The question arises, if the collection of primers and the corresponding set of PCR reactions is really sufficient to generate a representative pattern of bands. To this end, we have done a computer search in different genes of interest to find out if our primers would detect sufficient homology to amplify respective fragments. The results of this search are summarized in table

3. Interestingly, none of the primers matches completely with one of the genes looked at. We have to allow for at least one or even more mispairings to occur to have the chance of detecting one or more bands within the length limit of 500 nucleotides. In order to find out which of the potential fragments would actually be generated under the experimental conditions we have carried out nested PCR incubations for the identification of the potential fragments related to alpha1-antitrypsin, albumin and to β-actin. These results confirm our theoretical considerations as well as the data of Liang and Pardee (4). A recent paper of Pardee and coworkers (8) also shows that at least three mismatches are tolerated if clustered at the 5′ end of the primer. On the other hand, a primer with two internal mismatches failed to detect alpha1-antitrypsin. Thus, it seems to be possible to predict the position(s) of (a) band(s) corresponding to a known gene of interest within the display pattern on the basis of the six-nucleotide homology at the 3′ end of the primer. However, it cannot be ruled out that one or two internal mismatches are tolerated in some instances.

Differences in gene expression during liver regeneration

The most interesting application of this technique is the comparison of cells in different physiological situations or developmental stages. One of the most complex changes in gene expression probably occurs during regeneration of a tissue like the liver. We have analysed RNA from normal and regenerating liver using the technique in all its variations. We looked at three time points of regeneration after hepatectomy and compared it with the resting liver. To this end we have performed 4×312 PCR incubations and run 28 gels with the corresponding incubations from the four different stages of the liver side by side. Part of the analysis is given in figure 6 and shows clearly that some bands disappear in the regenerating liver and others appear at different time points. Other regions of the gels did not show any difference between the four situations. In total we detected some 70 fragments which are different between the four stages. We have cut many of the bands which are different and have tried to sequence them. In many instances the sequence ladder clearly showed the simultaneous presence of more than one DNA species suggesting that most places in the gels are occupied by more than one fragment. This problem could be avoided by either direct cloning of the fragments (4) or by using the amplified band as a probe to isolate the corresponding cDNA from a library. We are currently working on the latter aspect using the magnetic bead technique (9). From the 20 fragments analysed up to now, only one had a known sequence (10) which corresponds to a cDNA from a human liver cell (HepG2) library (Fig. 7). This sequence again shows that two mispairings at the 5′ end of the 10mer upstream primer do not impair the primer function. The differential expression of most of the fragments has been confirmed by slot blot and partly also by nuclear run-on analyses (not shown).

Automatic recording of the RNA display pattern

As the method creates reproducible band patterns it might be useful for general diagnostic purposes. However, the routine use

Table 2: Number of PCR bands obtained from mouse liver RNA (cDNA) on nondenaturating gels.

T_{11}	CA	CG	CT	CC	GA	GG	GT	GC	AA	AG	AT	AC	Σ	X
Primer	1	2	3	4	5	6	7	8	9	10	11	12		
1	92	188	73	137	186	182	67	189	121	142	129	147	1653	138
2	148	66	131	82	156	150	176	126	103	179	154	125	1596	133
3	81	172	172	121	89	82	135	131	114	97	145	124	1463	122
4	105	126	99	71	118	107	96	130	67	170	124	174	1387	116
5	94	182	104	157	101	91	70	91	115	89	148	103	1345	112
6	125	137	111	95	161	112	98	121	65	64	124	152	1365	114
7	62	93	129	60	169	146	176	64	104	175	123	156	1457	121
8	120	133	93	94	110	158	188	102	151	108	173	60	1490	124
9	183	157	124	148	173	167	180	167	76	146	90	87	1698	142
10	112	118	183	174	101	73	75	104	165	110	136	111	1462	122
11	98	180	174	159	140	139	123	101	147	73	176	138	1648	137
12	111	129	164	81	126	135	133	89	180	96	152	141	1537	128
13	82	139	149	135	124	146	71	73	136	153	86	90	1384	115
14	73	146	180	93	90	134	167	80	160	70	72	126	1391	116
15	174	179	72	105	138	153	83	88	130	130	104	87	1443	120
16	77	109	145	162	67	127	164	135	132	108	104	120	1450	121
17	65	105	135	136	129	170	151	85	130	106	159	178	1549	129
18	121	185	74	134	117	122	100	66	123	72	74	110	1298	108
19	165	73	80	179	114	169	91	134	65	122	75	121	1388	116
20	165	141	84	69	71	170	62	73	87	68	101	135	1226	102
21	80	137	156	120	94	179	66	126	161	131	188	89	1527	127
22	68	127	121	182	115	85	137	166	107	127	108	181	1524	127
23	156	141	131	78	164	183	92	157	123	95	138	84	1542	129
24	165	62	89	163	107	99	144	93	145	123	138	63	1391	116
25	82	185	187	71	114	184	140	103	142	101	116	170	1595	133
26	81	87	96	153	70	106	177	83	81	119	160	77	1290	108
Σ	2885	3497	3152	3159	3144	3569	3167	2877	3130	2974	3297	3149		
X	111	135	121	122	121	137	122	111	120	114	127	121	X=	122

The 5′ primers 1 to 26 are those listed in Table 1. The 3′ primers (T_{11}VN) 1 to 12 are specified by the respective dinucleotide sequence. Each figure is a mean value of four electrophoretic runs representing four stages of liver regeneration. Sums and mean numbers of bands were estimated for both, vertical and horizontal lines. The total number of bands is 38,004 and the mean number of bands per lane is 122.

for the comparison of test cells with a standard cell would require also a highly reproducible system for band recording. For this purpose autoradiographs are not well suited. Scanning of band patterns is very time consuming and error-prone. Therefore, we have attempted to use an automatic recording system together with dye-labeling of the fragments. The PCR reactions were set up with dye-labeled primers using different dyes for different cell types and the reactions were run in the ABI 373A automatic sequencer which allows to run four differently labeled samples in one lane of the gel. We have applied this technique to the analysis of liver regeneration and again compared the three time points of regeneration with the resting liver. One example of a comparison between two samples is given in figure 8. This clearly shows that differences in the band patterns of two differently labeled samples can be recorded simultaneously in one lane. Actually, it is also possible to record all four colors simultaneously either as a four-color curve or as a hard copy of the gel image (not shown). The major problem to be solved is the effect exerted by the different fluorescent dyes on the priming abilites of the oligonucleotides and on the simultaneous recording of differently colored fragments in one gel position. Generally, all dyes cause changes in the priming abilites of the respective primer if used in the PCR incubation from the beginning. This problem was solved by adding the dye-labeled primers only after 20 initial rounds of amplification (see Materials and Methods) thus allowing the labeled primer to act as a decamer. This solution works well for TAMRA (yellow) and FAM (blue) but less perfect for the other two colors. The simultaneous recording of different colors has its inherent problems. Quantitative differences between two different dyes in one gel position cannot be correlated with real differences in the amount of DNA. Thus, in its present stage the automatic analysis of RNA display patterns is a qualitative method allowing to identify genes which are completely switched on or off in the course of the particular process. However, the typical example shown in figure 8 clearly shows that the analysis on the automatic sequencer allows to identify real differences in patterns of gene expression.

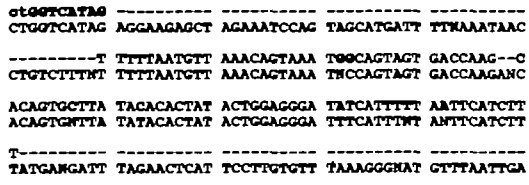

Figure 7. Partial sequence of the DNA obtained from one fragment of mouse liver cDNA. The fragment was cut from the gel, amplified, and sequenced as described in Materials and Methods. The sequence was found to be homologous to that of a cDNA clone HSOS14F08 derived from the human liver cell line HepG2 (10). This sequence is given in the lower line for comparison. The sequence of the upstream primer (No. 9) used to generate this fragment is given in bold letters and the two mismatched nucleotides at the 5′ end are indicated by lower case letters.

DISCUSSION

We have improved the RNA display method introduced by Liang and Pardee (4). Based on mathematical theory, we have devised a set of PCR primers which produced an average of 70 bands per lane at 42°C annealing temperature. This number of bands is increased by more than twofold using 40°C as the annealing temperature and denaturating gels. Using similar conditions as Liang and Pardee (4) we could confirm the high reproducibility of individual band patterns. However, we find that conventional denaturating sequencing gels produce artificially crowded band patterns due to the simultaneous presence of alternative forms of some or even most fragments including the two separate strands and molecules with and without an added A at the 3′ end. This leads definitely to the occupation of the same position in the gel by DNA molecules derived from different fragments. Using nondenaturating gels we could reduce the complexity of band

Table 3. Predicted fragments within the RNA display pattern for known genes

Gene	matched primer[1]	position[2]	fragment size[3]	detec.[4]
albumin	(19) g a t c ATAGCC	2008	90	yes
	(23) g a T c TGACTG	1867	230	no
	(24) c ATCATGGTC	1633	470	yes
AAT	(26) g a t c TAAGGC	1067	270	yes
	(17) GATC t g ACAC	1233	100	no
β-actin	(2) t gga TTGGTC	1794	100	no
	(7) T c Ga TACAGG	1558	330	no
	(12) C t g CTTGATG	177	8110	yes
	(14) g A t c AAGTCC	1564	330	no
	(22) GA t c GCATTG	1652	240	yes
	(24) GA t c ATGGTC	1609	280	yes

All gene sequences analysed are derived from mouse. AAT is alpha1-antitrypsin, Rb is the retinoblastoma susceptibility gene, and PGK is phospho glycerate kinase.

[1]The numbers of the primers given in brackets correspond to those used in table1. Mismatches between primer and target sequences are indicated by lower case. For AAT and PGK we have listed one example of primers which have two internal mismatches with the target sequence. AAT is not detected by this type of mispairing. Otherwise, all potential fragments generated by hexanucleotide homology are given. Note that PGK would only be detectable as a 800 base-pair fragment. Larger potential fragments also exist for the other genes, but are not listed.

[2]Positions refer to the nucleotide number within the respective gene where the first 5′ nucleotide of the primer would annael.

[3]Fragment sizes correspond to the distances between the 5′ nucleotides of both primers.

[4]Analysis of the actual existence of the fragments was done by nested PCR using gene specific internal primers. Bands were identified by agarose gel electrophoresis.

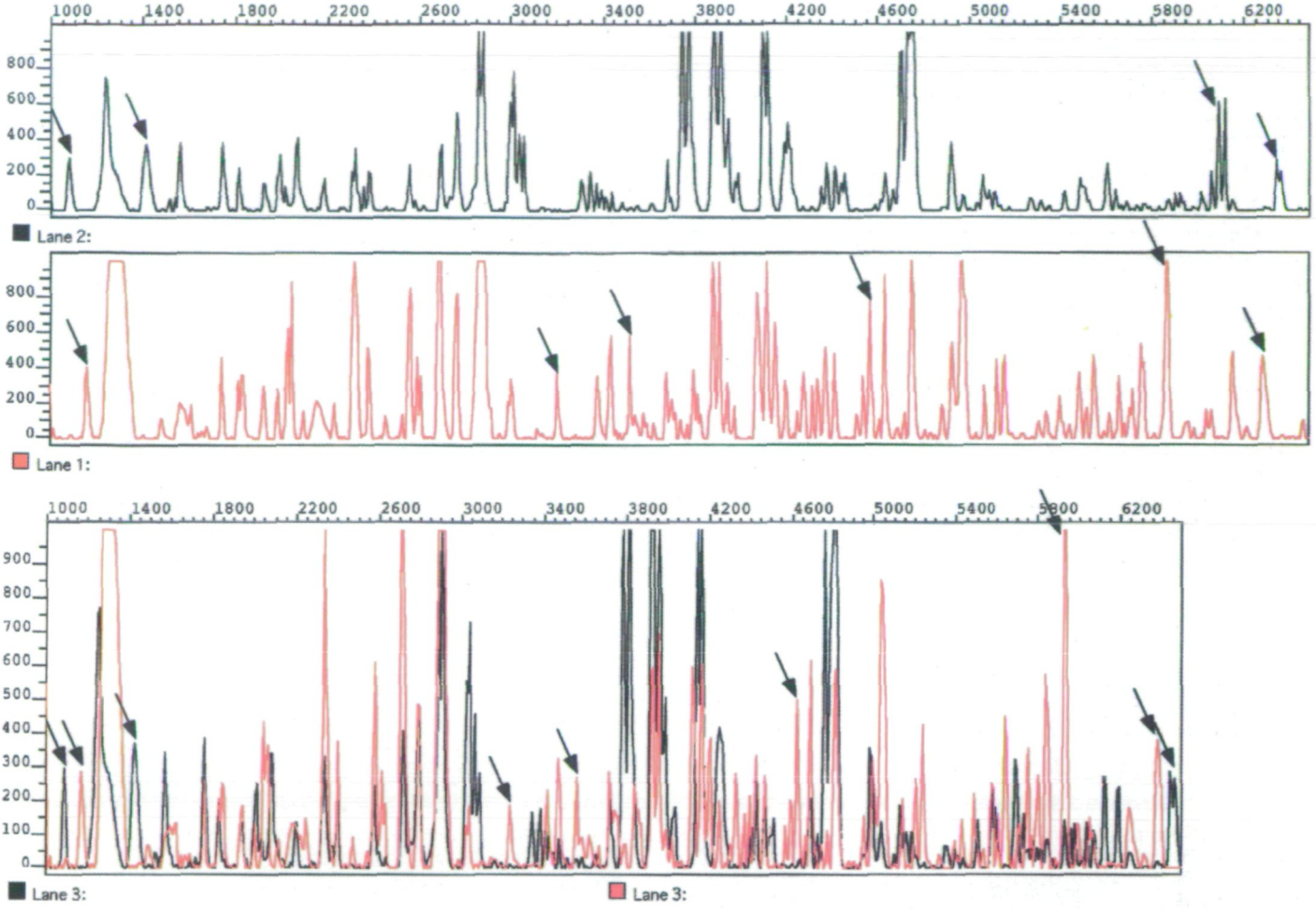

Figure 8. Band profiles obtained from RNA of resting and regenerating mouse liver using the ABI 373A DNA sequencer. cDNA fractions No. 1 from both, normal and regenerating liver 24h after hepatectomy, were amplified in the presence of upstream primer No. 3 which was labeled with different fluorescent dyes in both cases. Samples were run separately (normal liver, lane 1; regenerating liver, lane 2) and together (lane 3). Note that the differences in the patterns of the two samples are clearly detectable in the mixed sample.

patterns. These gels were not uniformly covered with bands. The mean number of bands obtained per lane was 122. A complete analysis with all 26 5′ primers in 312 PCR incubations generated a total of about 38,000 bands. This number is more than twofold higher than the predicted number of expressed genes. According to the theoretical considerations, this should be sufficient to have almost all RNA species represented by at least one band. Probably, the actual number of DNA fragments is even higher due to occupation of the same band positions by more than one fragment.

Published data (4,8) as well as our results strongly suggest that at least the six nucleotides at the 3′ end of the upstream primer must be perfectly matched. This six-nucleotide homology probably accounts for the majority of amplified fragments as predicted by the theory (Fig. 2). One to four mismatches in the 5′ end of the primer are obviously tolerated. This allows to predict the band positions for most known genes. Thus, we can conclude that the RNA display technique is able to identify almost all expressed genes using the protocol and the primers described here. This is the prerequisit for both promising aspects of application of this method, the generation of a complete diagnostic band pattern and for the identification of new genes. However, in its present state the method is only a qualitative one. Semiquantitative analyses will be possible by running different numbers of PCR cycles.

For diagnostic purposes we have adopted the RNA display technique to the ABI 373A DNA sequencer. The primers were labeled with four different fluorescent dyes. This allows to run up to three test RNA (cDNA) samples in parallel to the reference RNA. Reproducible results were obtained if the colors were scanned separately. However, problems with the quantitative recording occured when different colors were to be detected in one lane. Thus, the software has to be modified to allow for a quantitative comparison of the differently colored bands within one lane of the gel. The reproducible intensity of all unchanged bands can serve as an internal standard. Thus, the RNA display technique can actually be applied to the routine analysis of gene expression on a qualitative and semiquantitative level, at least with the two dyes TAMRA and FAM.

The most exciting aspect of the DDRT-PCR technique is the possibility to isolate new genes on a systematic basis. In contrast to the rather random sequencing of all members of a cDNA library, one could envisage the systematic analysis of the band patterns and assignement of cDNA clones and genes to individual

bands. The first step in this direction will be the isolation of genes differently expressed in various cell types, in different stages of development, different phases of the cell cycle, or in malignant tissues. In combination with transfection of regulatory genes or inactivation of regulatory gene function by antisense mechanisms, the technique will allow for identification of genes regulated by master genes. The gene isolation strategy requires a method which can deal with many samples at the same time. Magnetic beads covered with streptavidin (9,11) are ideally suited for this purpose and are currently used in our laboratories to isolate the genes differentially regulated during liver regeneration and those regulated by tumour suppressor genes.

ACKNOWLEDGEMENT

This work was supported by the Fonds der Chemischen Industrie.

REFERENCES

1. Zimmerman,C.R., Orr,W.C., Leclerc,R.F., Barnard,E.C. and Timberlake,W.E. (1980) Cell, **21**, 709–715
2. St. John,T.P. and Davis,R.W. (1979) Cell, **16**, 443–452
3. Strauss,M., Hering,S., Lubbe,L. and Griffin,B.E. (1990) Oncogene, **5**, 1223–1229
4. Liang,P. and Pardee,A.B. (1992) Science, **257**, 967–971
5. Chomczynski,P. and Sacchi,N. (1987) Anal. Biochem., **167**, 157–159
6. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) Molecular Cloning:A Laboratory Manual. Second Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
7. Sanger,F., Nicklen,S. and Coulson,A.R. (1977) Proc.Natl.Acad.Sci., **74**, 5463–5467
8. Liang,P., Averboukh,L., Keyomarsi,K., Sager,R. and Pardee,A.B. (1992) Cancer Res., **52**, 6966–6968
9. Tagle,D.A., Swaroop,M., Lovett,M. and Collins,F.S. (1993) Nature, **361**, 751–753
10. Okubo,K., Hori,N., Matoba,R., Niiyama,T., Fukushima,A., Kojima,Y. and Matsubara,K. (1992) Nature Genet., **2**, 173–179
11. Korn,B., Sedlacek,Z., Manca,A., Kioschis,P., Konecki,D., Lehrach,H. and Poustka,A. (1992) Hum. Molec. Genet., **1**, 235–242