1830–1835 *Nucleic Acids Research, 1997, Vol. 25, No. 9*

Primer design for a prokaryotic differential display RT-PCR

R. Fislage*, M. Berceanu+, Y. Humboldt, M. Wendt and H. Oberender

Institut für Medizinische Mikrobiologie, Universität Rostock, Schillingallee 70, D-18057 Rostock, Germany

Received December 3, 1996; Revised and Accepted March 6, 1997

ABSTRACT

We have developed a primer set for a prokaryotic differential display of mRNA in the *Enterobacteriaceae* group. Each combination of ten 10mer and ten 11mer primers generates up to 85 bands from total *Escherichia coli* RNA, thus covering expressed sequences of a complete bacterial genome. Due to the lack of polyadenylation in prokaryotic RNA the type T_{11}VN anchored oligonucleotides for the reverse transcriptase reaction had to be replaced with respect to the original method described by Liang and Pardee [*Science*, 257, 967–971 (1992)]. Therefore, the sequences of both the 10mer and the new 11mer oligonucleotides were determined by a statistical evaluation of species-specific coding regions extracted from the EMBL database. The 11mer primers used for reverse transcription were selected for localization in the 3′-region of the bacterial RNA. The 10mer primers preferentially bind to the 5′-end of the RNA. None of the primers show homology to rRNA or other abundant small RNA species. Randomly sampled cDNA bands were checked for their bacterial origin either by re-amplification, cloning and sequencing or by re-amplification and direct sequencing with 10mer and 11mer primers after asymmetric PCR.

INTRODUCTION

The differential display of mRNA method introduced by Liang and Pardee (1) allows the identification of differentially expressed eukaryotic genes by a non-stringent RT-PCR with different sets of 3′-anchored 13mer RT primers and additional 10mer PCR primers. For research on pathogen–host interactions especially, where only limited amounts of bacterial mRNA are available, the use of a PCR-based technique is desirable. However, the application of differential display PCR to prokaryotes is hampered by the low level of polyadenylation of the mRNA, preventing initiation of cDNA synthesis by the 3′-anchored 13mer primers. This problem may be circumvented by RNA arbitrarily primed (RAP) PCR-based methods (2,3). In these cases either single arbitrary primers (4) or random hexanucleotide mixtures (5) are chosen for the RT reaction. The subsequent PCR reaction contains one or two arbitrary primers. However, the design of primers and the low number of cDNA bands generated pose some difficulties in scanning a complete genome for differentially expressed prokaryotic sequences.

In this study we describe a general method for the design of non-anchored primers with similar statistical properties as compared with those given in the original publication. The primer set was designed and tested to completion for the strain *Escherichia coli* DH5α and partly characterized for seven additional species of *Enterobacteriaceae.* Furthermore, a direct sequencing approach for the identification of differentially expressed genes is outlined.

MATERIALS AND METHODS

Bacterial strains

Reference strains were obtained from the Deutsche Stammsammlung für Mikroorganismen (DSM). Tests for characterization of primers suitable for a differential display of mRNA were done with *E.coli* DH5α. Clinical isolates were characterized using standard microbiological methods. To test the performance of the primer set in related species we isolated RNA from the following strains: *Klebsiella oxytoca* (clinical isolate), *Salmonella enteritidis* (clinical isolate), *Shigella flexneri* type 2 (clinical isolate), *Serratia marcescens* (clinical isolate), *Proteus vulgaris* DSM 30118, *Citrobacter freundii* (clinical isolate), *Enterobacter cloacae* DSM 30054.

Primer design

In general the primers were selected for a high frequency of occurrence within the coding genome of *E.coli.* Furthermore, a high GC content for the 10mers and a lower GC content in the case of the 11mers was preferred. The underlying idea is derived from the fact that *E.coli* genes, though not possessing a poly(A) tract, have an AT-rich 3′-end. If compatible with statistical criteria, the 10mers contain an ATG start codon. The 11mers were selected for harboring a stop codon. Both GC content and the three base motif serve to achieve a preferential binding either at the 5′-end of a gene in the case of the 10mers or at the 3′-end in the case of the 11mers.

Initially, species-specific DNA sequences, except extrachromosomal material, were copied from the EMBL database (6). The coding portions of genomic sequences were extracted and compared with each other to eliminate multiple copies. Then, 5–10mer oligonucleotide frequencies were calculated and sorted in descending order. A total of 875 464 bp of *E.coli* coding

*To whom correspondence should be addressed. Tel: 49 381 494 5901; Fax: 49 381 494 5902; Email: rainer.fislage@medizin.uni-rostock.de

+ Present address: Disciplina Biologie Cellulare si Moleculara, Universitatea de Medicina si Farmacie, Timisoara, Romania

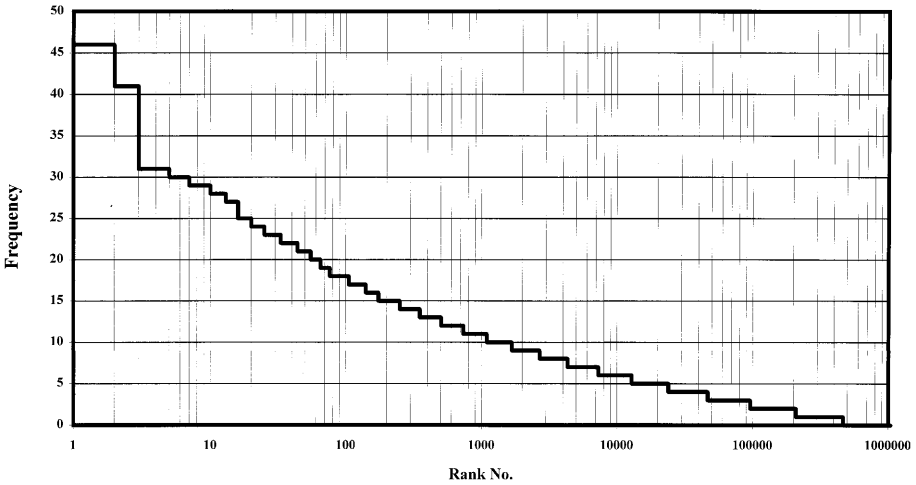

Figure 1. 10mer frequency distribution in 875 464 bp of coding *E.coli* sequences.

sequences were analyzed by computer programs written in the Turbo Pascal (Borland) language. The source code is available on request. Raw frequency distributions were further analyzed using the Access database package (Microsoft). The 10mer primers were derived from the 6mer and 8mer distributions, using the 10 most frequent 6mers of the list to identifiy the 10 most frequent 8mers with the 6mer sequence located at the 5′-end. Six bases from the 3′-end of those 8mers were tested once more against the 8mer distribution to identify 8mers harboring the six bases at their 5′-end. The first 8mer and the two 3′-bases of the second 8mer made up a 10mer primer sequence. For each of the initial 8mers four 10mers with the most frequent two bases at their 3′-end were tested against the EMBL database. Primers with maximum occurrence within the *E.coli* genome and no homology with abundant sequences were chosen for the experiments.

With respect to the 10mer frequency distribution, all of the 10mer primers were located within the first 371 ranks. From the 4^{10} possible 10mer permutations, 258 944 ranks were populated by one member and 580 918 ranks were unpopulated (Fig. 1). The first 500 ranks contained at least 13 members and a maximum of 46 elements.

The ten 5′-bases of the 11mer primers were selected from the first 325 ranks of the 10mer distribution, by searching for the most frequent 10mers harboring a stop codon. The missing base was added as either A, G, C or T to the 3′-end of the 10mers. The oligomer showing most frequent occurrence within the *E.coli* genome and no homology with abundant sequences was chosen for the experiments.

For one initial differential display experiment using RNA from *E.coli* cultures grown to log and stationary phase an additional primer pair was used. The 10mer primer ECs2 (5′-ACGCTG-CTGG-3′) was from rank 1777, populated by nine members, and the first ten 5′-bases of the 11mer ECa1 (5′-TTTTTCGTTTA-3′) were located at rank 7433, populated by six elements.

Since the 11mers were to be used for RT reactions, the reverse complement was synthesized.

No attempt was made to identify primers containing true in-frame start or stop codons in the case of the 10mers and 11mers respectively, for the coding sequences had not been tested to be of full length.

Preparation of RNA

Total bacterial RNA was isolated by a method modified from the protocol of Chomczynski and Sacchi (7). Plate grown bacteria were inoculated in 20 ml broth I (17 g/l pancreatic peptone, 3 g/l yeast extract, 5 g/l NaCl) and incubated for 3 h at 37°C under vigorous shaking. The cells of ten 1.5 ml aliquots were sedimented and subsequently lysed in 400 µl GCN solution (4 M guanidinium thiocyanate, 25 mM sodium citrate, pH 7.0, 0.5% *N*-lauroylsarcosine, 100 mM β-mercaptoethanol). After addition of 50 µl 2 M sodium acetate, pH 4.0, the samples were extracted with 500 µl TE-saturated (100 mM Tris–HCl, pH 7.2, 0.5 mM EDTA) phenol and 150 µl chloroform/isoamyl alcohol (49:1) and stored on ice for 15 min. After isopropanol precipitation, the RNA sediment was recovered in 150 µl GCN solution and precipitated again with 150 µl isopropanol. The RNA sediments were joined in a total volume of 80 µl diethylpyrocarbonate (DEPC)-treated water. To remove bacterial DNA contamination the preparation was incubated for 30 min with 5 U RNase-free DNAse (Promega) in RT buffer (AMV; Promega). Following a further phenol/chloroform extraction and isopropanol precipitation, the RNA was recovered in 30–60 µl DEPC-treated water. The RNA concentration of corresponding samples for differential display PCR was adjusted by comparison of ethidium bromide fluorescence intensities.

Reverse transcription of RNA

Freshly isolated bacterial RNA was subjected to reverse transcription within 1 day. The cDNA synthesis was carried out using 8 U AMV reverse transcriptase (Promega) in a total volume of 20 µl with 25 µM dNTP and 6.3 µM 11mer primer. The reaction was run in a thermocycler. A sample of 15 µg total RNA dissolved in DEPC-treated water was denatured for 10 min at 70°C and cooled to 30°C. After addition of the reaction mix the primers were elongated at 30°C for 5 min followed by a 1 h synthesis step at 37°C and denaturation at 95°C for 10 min. Samples were stored for several days at –70°C.

Differential display PCR

The reaction mixture for differential display PCR, in a volume of 20 µl, contained 2 µl cDNA obtained from the RT reaction, 2 µM dNTP, 5 µM 11mer primer, 0.5 µM 10mer primer, 2.5% DMSO, 2.5 U Taq DNA polymerase (Perkin Elmer) and 1.5 µCi [α-^{33}P]dATP in the manufacturer's reaction buffer with additional $MgCl_2$ to a final concentration of 2.5 mM. cDNA was amplified for 40 cycles using 40°C as the annealing temperature. Denatured PCR products were separated on a 6% sequencing gel fixed to a silane A174 (BDH) pre-treated glass plate. If cDNA fragments were to be cut out, the gels were washed in 10% methanol for 15 min prior to drying at 70°C.

Recovery, re-amplification and direct sequencing of cDNA fragments

cDNA fragments were excised from dried gels and stored in 500 µl TE overnight at 4°C to reconstitute the gel material. Next day the samples were incubated for 3 h at 56°C. After centrifugation, the supernatant was extracted twice with 1 ml water by ultrafiltration using Centricon-100 tubes according to the supplier's (Amicon) instructions. The 40 µl concentrate was distributed in two 100 µl PCR reactions containing the primer pair which generated the cDNA fragment. Samples contained 4.4 µM dNTP, 5 µM 11mer primer, 1.3 µM 10mer primer and 4 U Taq DNA polymerase in the manufacturer's reaction buffer with additional $MgCl_2$ to a final concentration of 2.5 mM. In most cases 5 µl amplified DNA was visible on a 3% agarose gel as a clear sharp band below fluorescence saturation after ethidium bromide staining. For further amplification, 6–12 µl of this product was subjected to another 40 cycles of PCR.

For asymmetric PCR (8) 100 µl PCR product was purified by ultrafiltration as described above. The PCR reaction, with 17 cycles and an annealing temperature of 40°C, was performed in a total volume of 100 µl, including 20 µl ultrafiltrate, 200 µM dNTP, 1 µM 11mer primer and 2.5 U Taq DNA polymerase.

The cDNA fragments were directly sequenced after asymmetric PCR using the Sequenase Version 2 kit (Amersham/USB). The 10mers served as sequencing primers using elevated concentrations of 25 µM.

Cloning of PCR products

Re-amplified PCR products were cloned into the *Sma*I site of pBluescript. PCR products were treated with proteinase K prior to ligation in order to improve the cloning efficiency (9). Plasmid inserts were sequenced with universal T3 or T7 primers.

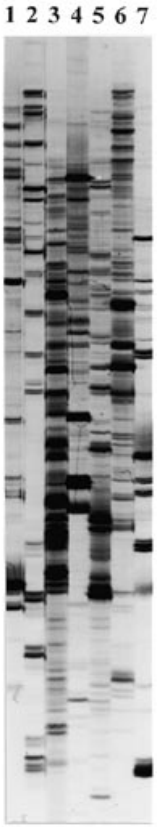

Figure 2. RT-PCR for total *E.coli* RNA. Lanes 1–7, primers Ea3/Es7, Ea3/Es8 and Ea4 with Es2, Es3, Es5, Es7, Es8 respectively.

RESULTS

Primer selection

By statistical analysis 20 primers were chosen for the experiments, resulting in 100 possible combinations of 11mer RT primers with 10mer oligonucleotides. All of the 11mers contained one stop codon and four of the 10mers embodied start codons, as indicated in bold in Table 1. The mean GC content of the 10mers was 61 ± 16% and 43 ± 12% for 11mers. The number of primer combinations may be doubled by using the 10mer primers for the RT reaction.

Table 1. Sequences and GC content of 11mer and 10mer primers for differential display RT-PCR in *Enterobacteriaceae*

RT primer	Sequence (5′→3′)	GC content (%)	PCR primer	Sequence (5′→3′)	GC content (%)
Ea1	TT**TTA**TCCAGC	36	Es1	GCTGGAAAAA	40
Ea2	ACT**TTA**CGCAG	45	Es2	GCTGCTGGCG	80
Ea3	T**TTA**TCCAGCG	45	Es3	GAAGTGCTGG	60
Ea4	**TCA**GCGTTTTA	36	Es5	TGGCGGCGGC	90
Ea5	TT**TCA**GCGCCT	55	Es6	AACTGGCGAA	50
Ea6	TTTTT**TCA**GCA	27	Es7	**ATG**CGCTGGC	70
Ea7	TCTTTT**TTA**CC	27	Es8	TGCCG**ATG**AA	50
Ea8	A**TCA**TCCAGCA	45	Es10	CTGGAAGAAG	50
Ea9	TT**TTA**CCCAGC	45	Es11	**ATG**GCGCTGG	70
Ea10	T**TCA**GCCAGCG	64	Es13	**ATG**GCG**ATG**A	50

Oligonucleotides Ea1–Ea10 denote the 11mer RT primers. Es1–Es13 are the 10mer PCR primers. Stop and start codons are in bold.

Amplification characteristics of *E.coli* RNA

The outcome of differential display RT-PCR for total *E.coli* RNA (Fig. 2) was comparable with results obtained from total eukaryotic RNA using anchored primers. The combinations of 11mer RT primers and 10mer oligonucleotides gave a total number of 4347 bands (Table 2). The number of bands was reproducible from experiment to experiment, as decribed previously for the eukaryotic technique (10).

Table 2. Number of RT-PCR bands for each oligonucleotide pair from total *E.coli* RNA

	Es1	Es2	Es3	Es6	Es5	Es10	Es7	Es8	Es11	Es13
Ea1	40	38	20	23	58	33	39	35	15	22
Ea2	32	50	34	29	60	28	43	38	46	53
Ea3	18	72	58	20	52	18	55	53	8	10
Ea4	46	84	51	29	71	15	70	38	33	24
Ea5	49	66	53	40	48	20	45	37	63	25
Ea6	32	60	33	43	30	43	35	41	59	55
Ea7	80	49	23	85	25	67	46	46	78	21
Ea8	77	36	15	60	41	71	37	46	45	39
Ea9	36	75	33	23	55	29	42	50	38	54
Ea10	32	60	49	41	60	35	27	59	57	67

The primer pairs used in reverse order with respect to the RT reaction resulted in a comparable number of PCR bands with a different fragment pattern.

Water controls showed two types of PCR by-products caused by primer oligomerization and some DNA contamination of the Taq polymerase (data not shown). The latter could be removed by DNase pre-treatment of the Taq polymerase with subsequent heat-inactivation of the DNase. However, both products disappeared when cDNA was present in the reaction mixture.

Amplification characteristics for RNA of related species

RNA of related species was tested with Ea5 and Ea7 RT primers in combination with the Es2, Es3, Es5, Es7 and Es8 primers (Fig. 3). The results shown in Table 3 indicate that the primer set may be suitable for use in the complete group of *Enterobacteriaceae*, although the number of bands generated from *Citrobacter* and *Shigella* RNA was reduced, which may impair scanning of the whole coding genome for these species.

Sequencing of cDNA fragments

To assess specificity and origin of amplified cDNA fragments 10 randomly chosen samples were sequenced either after cloning or directly using the 10mer oligonucleotide as the sequencing primer. Gels with directly sequenced PCR products showed elevated background signals, probably caused by minor contaminants of the main PCR product.

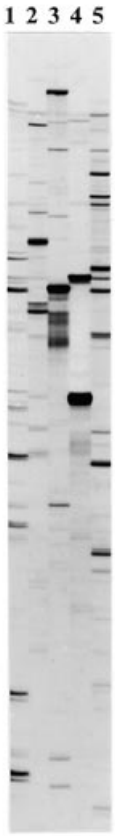

Figure 3. RT-PCR for total *P.vulgaris* RNA. Lanes 1–5, primer Ea5 with Es2, Es3, Es5, Es7, Es8 respectively.

One attempt was made to compare RNA prepared from *E.coli* cultures grown to log phase and stationary phase respectively (Fig. 4). Here the *aldB* gene was identified, which is known to be transcriptionally induced at the transition from the exponential to the stationary growth phase (11).

In total, seven out of 10 fragments were directly sequenced. Four of the 10 fragments were revealed to be known genes [*xylA* (12,13), *trkA* (14), *aldB* (11) and *rimI* (15)]. The positions and matching of the primers to known sequences is shown in Table 4. The length of the six unknown re-amplified PCR products was between 350 and 200 bp. Although the 11mer primers bound preferentially in the 3′-region of the known genes, none of the oligonucleotides bound to a start or stop codon. The PCR product from the *rimI* locus was synthesized by the 10mer primer only. A 4-fold mismatched putative 11mer binding site was identified at the 3′-end of the gene.

Table 3. Number of bands for different members of the *Enterobacteriaceae* group

	Citrobacter freundii	*Proteus vulgaris*	*Enterobacter cloacae*	*Klebsiella oxytoca*	*Serratia marcescens*	*Shigella flexneri*	*Salmonella enteritidis*	*Escherichia coli*
Es2	16/10	46/ND	63/86	44/35	42/30	24/ND	61/44	66/49
Es3	12/15	41/43	42/69	49/24	38/40	15/44	81/ND	53/23
Es5	15/11	36/31	63/66	79/51	82/46	28/36	71/40	48/25
Es7	28/19	35/24	45/55	54/23	54/50	15/15	41/ND	45/46
Es8	21/15	61/50	34/57	51/24	52/56	28/51	61/ND	37/46

The results denote the combinations of oligonucleotides Es2, Es3, Es5, Es7 and Es8 with the RT primers Ea5 or Ea7 respectively. ND, not determined.

Table 4. Matching of primer sequences to known templates

Locus	RT primer	Position	PCR primer	Position	Coding sequence (Start, Stop)	EMBL accession no.	Length of PCR product (bp)
aldB	Eca1		Ecs2				
	5′-TTTTTCGTTTA		5′-ACGCTGCTGG				
	I I IIIIII		I IIIIIIII				
	5′-TCTACCGTTTA	7397	5′-AAGCTGCTGG	6700	5565–7101	AE000463/rev	698
rimI	Eca1		Ecs2				
	5′-TTTTTCGTTTA		5′-ACGCTGCTGG				
	II IIIII		IIIIIIIII				
	5′-CGTTGAGTTTA	11833	5′-GCGCTGCTGG	11380	11134–11617	AE000507	288
			Ecs2/rev				
			5′-ACGCTGCTGG				
			IIIIII II				
			5′-CCGCTGCAGG	11659			
trkA	Ea5		Es7				
	5′-TTTCAGCGCCT		5′-ATGCGCTGGC				
	II I IIII I		I III II I				
	5′-TTCCCGCGCAT	1713	5′-ACGCGTTGTC	1337	162–1552	X52114	374
xylA	Ea4		Es11				
	5′-TCAGCGTTTTA		5′-ATGGCGCTGG				
	I II I IIII		IIIIIIII				
	5′-TAAGAGCTTTA	3576	5′-GGGGCGCTGG	3271	3217–3505	K01996	315

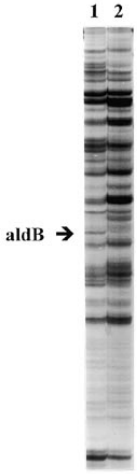

Figure 4. Differential display RT-PCR of *E.coli* RNA from log phase and stationary growth phase with primers Eca1/Ecs2. Lane 1, stationary-phase; lane 2, log phase.

DISCUSSION

Genome coverage of the primer set

Under the simple assumption that one cDNA band represents one gene, the 4347 observed fragments cover 90% of the coding sequences, provided that the *E.coli* genome of 4.7 Mbp contains a total number of ~4700 coding sequences. If one takes the operon structure of prokaryotic genes into account, the coverage may rise by a factor of two as the number of independent transcripts decreases. A more thorough view of the 11mer frequency distribution allows calculation of a value for the minimum genome coverage achieved by the primer set. If each gene in the genome is to be reverse transcribed the spacing of the 11mer primers within the coding sequences becomes crucial, therefore, we determined all the binding sites for the 11mers showing not more than one mismatch at each position except the 3′-base (Table 5). Overall there are 528 binding sites per 875 464 bp of coding sequence in our dataset, giving a mean binding site spacing of 1.7 kb. Assuming a mean length of 1 kb for a coding sequence, at least 59% of all genes would be reverse transcribed. Taking into account the more highly degenerate binding sites and the operon structure of prokaryotic genes, coverage of transcribed sequences will further increase.

Table 5. Sum of all mono-degenerate and non-degenerate binding sites for 11mer primers in 875 464 bp of coding *E.coli* sequences

Primer	Number of binding sites	Spacing (kb)
Ea1	59	15
Ea2	27	32
Ea3	62	14
Ea4	32	27
Ea5	63	14
Ea6	87	10
Ea7	50	18
Ea8	55	16
Ea9	42	21
Ea10	51	17

Primer design and primer performance in prokaryotes

Statistical analysis of coding sequences allows selection of short oligonucleotide sequences with above average occurrence in the prokaryotic genome. The resulting primers described here are suitable for differential display PCR with a genome coverage in the range 59–90% in the *Enterobacteriaceae* group. They may also be useful for DNA fingerprinting methods. This method of primer construction should be applicable to all prokaryotic and eukaryotic species with cDNA sequences available. For completely sequenced prokaryotes (16–19) especially, transcriptional regulation of all genes with unknown functions may be assessed under varying environmental conditions.

The concept of excluding primers showing homology to expressed repetitive sequences may also be helpful in the development of 10mers for eukaryotic systems, since 10% of polyadenylated human transcripts are of repetitive origin (20).

The number of bands generated from total *E.coli* RNA did not correlate with the absolute ranking position of the partial or total primer sequence within the 10mer distribution. The mean values of rows and columns in Table 2 revealed no statistically significant differences, indicating that 10mers and 11mers contributed equally to the number of bands. This fact may be explained by the degenerate binding mode of short oligonucleotides, leading to non-statistical behavior (1). In the case of the 11mers especially, where non-stringent conditions for the RT reaction were applied, a non-statistical behavior is to be expected. Another explanation is the small difference of a factor of four between maximum and minimum oligonucleotide frequencies within the first 500 ranks of the 10mer distribution, in which all 10mer and the partial 11mer primers are located. These findings suggest that the complicated construction of 10mer primers by setting up chained sequences may not be necessary and may be replaced by a direct selection from the first 500 ranks of the 10mer distribution. The initial method was only chosen to ensure that the primers contained frequent 6–8mer oligonucleotides as a potential region of perfect homology when the 10mer oligonucleotides bind in a degenerate fashion to the cDNA. The incorporation of start and stop codons had no effect on the location of primers within the coding sequences, although the 11mer primers were directed to the 3′-ends of the genes.

Re-amplification and sequencing of PCR products

Three of the four cDNA fragments of known genes from *E.coli* were located within transcription units, as jugded from the database entries, indicating that RNA was amplified by the RT-PCR. All of the fragments were located within the second halves of the coding sequences. To avoid false positive results when performing true differential display experiments, care has to be taken in removing the contaminating genomic DNA, as is the case in the eukaryotic system. One criterion to differentiate between bands generated by DNA contamination and true cDNA fragments may be the search for prokaryotic transcription terminator sequences if known genes are identified. For a fragment resulting from mRNA the 11mer sequence should be located in the 5′-direction of the stem–loop sequence. However, not all terminator sequences for the transcripts are defined in the databases. An independent re-testing of differentially expressed genes by Northern hybridization or quantitative RT-PCR is strongly recommended to exclude false positive results from unreproducible bands.

A difference from eukaryotic differential display is the increased amount of RNA needed to generate signals from total prokaryotic RNA. Especially when grown to the end of log phase, the apparent quality of the RNA is impaired, although the yield, as measured by fluorescence intensity of the ribosomal bands, is only slightly reduced. A reason may be the high turnover rate of prokaryotic mRNA causing a decline in steady-state concentrations when metabolic activity is reduced.

The quality of the re-amplification products was improved by washing the sequencing gels in 10% methanol instead of acetic acid prior to drying. This procedure may prevent depurination caused by heat and the low pH of acetic acid.

Ultrafiltration of eluted cDNA fragments increased recovery rates and completely removed urea, which is known to inhibit Taq DNA polymerase.

Direct sequencing of re-amplified PCR products is faster than cloning and sequencing, but the signal-to-noise ratio is diminished by background contaminating PCR products. On the other hand, the sequence of the major product alone is determined, thus eliminating the need to sequence several plasmid inserts obtained from cloning of cDNA bands. A strategy for fast determination of cDNA sequences from re-amplified differential display products may include an initial attempt at direct sequencing and thereafter a cloning and sequencing step if the sequence was unknown or direct sequencing failed if the PCR product was generated by only one of the primers.

ACKNOWLEDGEMENT

This work was supported by the internal FORUN program of the Medical Faculty of the University of Rostock.

REFERENCES

1 Liang,P. and Pardee,A.B. (1992) *Science*, **257**, 967–971.
2 Welsh,J. and McClelland,M. (1990) *Nucleic Acids Res.*, **18**, 7213–7218.
3 Williams,J.G., Kubelik,A.R., Livak,K.J., Rafalski,J.A. and Tingey,S.V. (1990) *Nucleic Acids Res.*, **18**, 6531–6535.
4 Wong,K.K. and McClelland,M. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 639–43.
5 Kwaik,Y.A. and Pederson,L.L. (1996) *Mol. Microbiol.*, **21**, 543–556.
6 Etzold,T. and Argos,P. (1993) *Comput. Appl. Biosci.*, **9**, 49–57
7 Chomczynski,P. and Sacchi,N. (1987) *Anal. Biochem.*, **162**, 156–159.
8 Gyllensten,U.B. and Erlich,H.A. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 7652–7656
9 Crowe,J.S., Cooper,H.J., Smith,M.A., Sims,M.J., Parker,D. and Gewert,D. (1991) *Nucleic Acids Res.*, **19**, 184.
10 Bauer,D., Müller,H., Reich,J., Riedel,H., Ahrenkiel,V., Warthoe,P. and Strauss,M. (1993) *Nucleic Acids Res.*, **21**, 4272–4280.
11 Xu,J. and Johnson,R.C. (1995) *J. Bacteriol.*, **177**, 3166–3175.
12 Lawlis,V.B., Dennis,M.S., Chen,E.Y., Smith,D.H. and Henner,D.J. (1984) *Appl. Environ. Microbiol.*, **47**, 15–21.
13 Schellenberg,G.D., Sarthy,A., Larson,A.E., Backer,M.P., Crabb,J.W., Lidstrom,M., Hall,B.D. and Furlong,C.E. (1984) *J. Biol. Chem.*, **259**, 6826–6832.
14 Schloesser,A., Hamann,A., Bossemeyer,D. and Bakker,E.P. (1993) *Mol. Microbiol.*, **9**, 533–543.
15 Yoshikawa,A., Isono,S., Sheback,A. and Isono,K. (1987) *Mol. Gen. Genet.*, **209**, 481–488.
16 Fleischmann,R.D. *et al.* (1995) *Science*, **269**, 496–512.
17 Fraser,C.M. *et al.* (1995) *Science*, **270**, 397–403.
18 Bult,C.J. *et al.* (1996) *Science*, **273**, 1058–1073.
19 Kaneko,T., Sato,S., Kotani,H., Tanaka,A., Asamizu,E., Nakamura,Y., Miyajima,N., Hirosawa,M., Sugiura,M. and Sasamoto,S. (1996) *DNA Res.*, **3**, 109–136.
20 Crampton,J.M., Davies,K.E. and Knapp,T.F. (1981) *Nucleic Acids Res.*, **9**, 3821–3834.