4520–4523 Nucleic Acids Research, 1995, Vol. 23, No. 21

An analysis of differential display shows a strong bias towards high copy number mRNAs

David J. Bertioli*, Ursula H. A. Schlichter[1], Michael J. Adams[1], Paul R. Burrows, Hans-Henning Steinbiß[2] and John F. Antoniw[1]

Departments of Entomology and Nematology and [1]Plant Pathology, IACR-Rothamsted, Harpenden, Herts AL5 2JQ, UK and [2]Max-Planck-Institut für Züchtungsforschung, Carl-von-Linné-Weg 10, Köln, Germany

Received August 1, 1995; Revised and Accepted September 20, 1995

Differential display of mRNA (DDRT-PCR; refs 1–3), has stimulated much work in the investigation of gene regulation. Although, in principle, the technique is simple and elegant, in practice it is difficult. A problem that has received little attention so far is the sensitivity of DDRT-PCR to rare mRNAs. The distribution of abundances of eukaryotic mRNAs is very skewed towards rare mRNAs (Table 1; refs 4 and 5). Therefore sensitivity to rare mRNAs is vital for DDRT-PCR to work efficiently. In spite of this, no measurements of the sensitivity of DDRT-PCR have yet been published. The following report describes several tests of the sensitivity of DDRT-PCR and the effect of specificity of priming on the sensitivity of DDRT-PCR is discussed.

In previous studies using DDRT-PCR we successfully identified differences in gene expression between tissues which were very different developmentally. For instance, DDRT-PCRs of plant leaves and roots usually showed ~3% difference in banding patterns between these two tissues (unpublished data). It was surprising that there were not more band differences, because other studies have demonstrated that leaves and roots differ in ~25% of their polysomal mRNA species (6). We also used DDRT-PCR to study the roots and leaves of uninfected plants and of plants responding to pathogen attack: barley responding to take-all fungus, *Gaeumannomyces graminis* var. *tritici* and potatoes infected with potato cyst nematode, *Globodera rostochiensis*. Both barley and potato plants show dramatic physiological responses to the pathogen attack (7,8). However, in spite of extensive analysis (65 primer combinations), no differences in gene expression were found between uninfected plants (roots or leaves) and plants responding to pathogen attack (unpublished data). Therefore, it seemed likely that DDRT-PCR may not be efficiently displaying many of the expressed mRNAs in the samples. One cause for this may be a lower sensitivity of DDRT-PCR to less abundant mRNAs. Most (~90%) mRNA species are rare, each species being present at ~0.004% of total mRNA (Table 1 and refs 4 and 5).

The sensitivity of DDRT-PCR was investigated in model biological systems where the expression of defined transcripts are known. A primer (Tob-PR1: GATGTGCTAG) perfectly matching the pathogenesis related PR1a, b and c mRNAs was designed for the detection of mRNAs encoding the PR-1 proteins (accession numbers PR-1a, X05452, X12485, M36691; PR-1b, X05453, X12486, M36692; PR-1c, X05454, X12487). The levels of PR-1 protein mRNAs are ~1% of all mRNAs in tobacco mosaic virus (TMV) infected leaves (9) and <0.01% in uninfected tissue (10) but using this primer DDRT-PCR did not show any bands of the expected sizes (Table 2). Analysis of total leaf protein by non-denaturing polyacrylamide gel electrophoresis confirmed the presence of large amounts of the PR-1 proteins in TMV-infected but not uninfected leaves (data not shown).

Another test was designed for the detection of TobRB7 mRNA which is constitutively expressed in tobacco roots but not in leaves (ref. 11; accession number X54855). The level of root-specific TobRB7 mRNA is ~0.3% in tobacco roots but undetectable in tobacco leaves (11). However, a perfectly matching primer (RB7-759: GATGCCACAC) failed to detect this mRNA in tobacco roots (Table 2).

Table 1. The abundance distribution of mRNA species and classes in a typical mammalian cell (modified from ref. 4)

	Copies of each species/cell	No. of mRNA species in class	Mean % of each species in class	Mean mass (ng) of each species in 2 µg total RNA
Abundant	12 000	4	3.3	3.3
Intermediate	300	500	0.08	0.08
Rare	15	11 000	0.004	0.004

* To whom correspondence should be addressed

Table 2. The sequences, binding characteristics, target mRNAs, predicted product sizes and detection limits of the primers used in this study

10mer	Target	T-primer	10mer Binding (mismatches in lower case)	Product size (bp)	Minimum level detected with competition (pg)	Minimum level detected without competition (pg)
β-371	β-globin	T_{12}-MC	CCTGGGCAAC	231	400 (1.2%)	0.04
	β-globin	T_{12}-MC	CCTGGgCAAC	315	400 (1.2%)	4
	α-globin	T_{12}-MC	CCTGGgCAAC	201	400 (1.2%)	40
	α-globin	T_{12}-MC	CCTGGgCAAC	141	400 (1.2%)	4
β-δ371	β-globin	T_{12}-MC	gCaGtGCAAC	231	ND	ND
β2-377	β-globin	T_{12}-MC	CtACGTGCTG	225	400 (1.2%)	40
	β-globin	T_{12}-MC	ctAcGTGCTG	351	4000 (11%)	400
	α-globin	T_{12}-MC	CtaCGTGCTG	125	ND	ND
β4-377	β-globin	T_{12}-MC	gtAaGTGCTG	225	ND	4000
	β-globin	T_{12}-MC	GtAaGTGCTG	351	ND	4000
	α-globin	T_{12}-MC	gtaaGTGCTG	126	ND	ND
α-363	α-globin	T_{12}-MC	CCTGGCCAAC	201	ND	4
	α-globin	T_{12}-MC	CCTGGCCAAC	141	400 (1.2%)	4
	β-globin	T_{12}-MC	CCTGGcCAAC	231	ND	ND
	β-globin	T_{12}-MC	CCTGGcCAAC	315	400 (1.2%)	0.4
α-401	α-globin	T_{12}-MC	TGCATGCCTC	163	ND	ND
Tob-PR1	PR1a	T_{12}-MA	GATGTGCTAG	374	ND	NA
	PR1a	T_{12}-MA	GATGTGCTAG	342	ND	NA
	PR1b	T_{12}-MA	GATGTGCTAG	302	ND	NA
	PR1b	T_{12}-MG	GATGTGCTAG	334	ND	NA
	PR1c	T_{12}-MC	GATGTGCTAG	273	ND	NA
	PR1c	T_{12}-MG	GATGTGCTAG	337	ND	NA
RB7-759	TobRB7	T_{12}-MA	GATGCCACAC	179	ND	NA

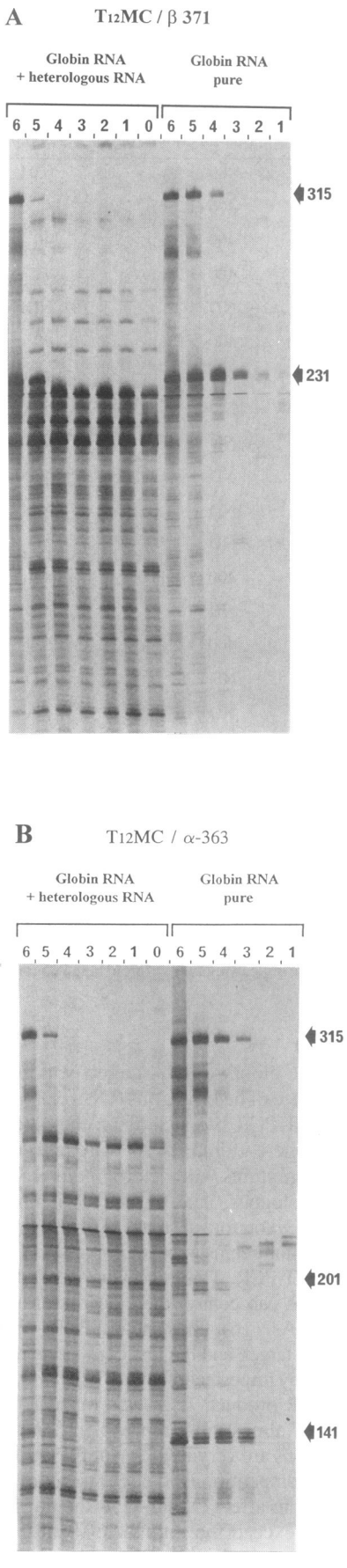
Percentages in parenthesis are the percentage of total mRNA equivalents. (for β-371 where experiments were done using two different heterologous RNAs the lowest detection levels obtained for each target are given). ND is not detected. NA stands for not applicable.

Because DDRT-PCR failed in these model biological systems, we then looked at a simpler system, the detection of the very abundant rabbit α- and β-globin mRNAs (accession numbers V00875 and V00879 respectively). In a serial dilution of globin RNA, several primers with different degrees of match (Table 2), were able to detect 4 pg of target (equivalent to the amount of a rare mRNA in 2 μg total RNA, Fig. 1). However, in a more realistic system with globin RNA mixed with a heterologous total RNA no primer/target combination tested could detect α- or β-globin mRNA at <400 pg (1.2% of mRNA, equivalent to an intermediate/ abundant mRNA; Table 1 and 2; Fig. 1).

Therefore in our tests, using a total of 23 primer/target combinations DDRT-PCR could not detect target at <1.2% of mRNA (Table 2). Although it is possible that some other primer combinations or conditions could show greater sensitivity, our results show that DDRT-PCR has a strong bias towards higher abundance mRNAs. It is interesting to note that RAPD-PCR (12,13) which also uses 10mers of arbitrary sequence as primers for PCR with a complex genomic DNA as template, gives bands which mostly represent repetitive DNA (14,15).

Sensitivity to the globin mRNA targets was 10–10 000 times greater in the absence of heterologous RNA/cDNA (Table 2), also some of the DDRT-PCR products from the heterologous RNA were not visible where globin RNAs were present at the highest concentrations in mixtures with heterologous RNA (Fig. 1A and B, globin + heterologous, lane 6). These findings can both be explained by competition for substrates by the many PCR products. Therefore, competition seems to be an important determinant of the sensitivity of DDRT-PCR. This is similar to standard PCR where competing reactions can control the outcome of amplifications of less abundant DNAs (16). In standard PCR, primers normally perfectly match the target, and the few competing reactions are less efficiently primed by imperfect matches. However, in DDRT-PCR, the number of PCR products is greater, mismatches are essential for the priming of almost all templates (ref. 3 and below), and dNTP concentrations are very low (1% of the amounts in standard PCR); therefore, competition is likely to be more severe.

For competition to affect DDRT-PCR the number of potential PCR bands must exceed the number of bands visualised on a DDRT-PCR gel. The number of potential bands is determined by



the specificity of the primers. This was investigated further using a computer simulation which generated 15 000 random sequences, each 1200 nucleotides long, to represent the different mRNA species present in a cell. Sequences were searched for matches to an arbitrarily chosen 10mer primer. In our tests (Table 2) DDRT-PCR bands were produced with up to three mismatches, but others have reported that as many as four mismatches can occur (3). Therefore, in the simulation 0–4 mismatches were allowed between primer and template. Mismatching increases the number of potential binding sites much more rapidly than using an equivalent, shorter primer (Table 3). Thus a 10mer primer with two mismatches binds nearly half of the different mRNA/cDNA molecules in a cell, with three mismatches there are on average 3–4 binding sites/cDNA, and with four mismatches an average of 19–20 binding sites/cDNA!

Table 3. Priming characteristics of primers of different length, and (by computer simulations) of arbitrary 10mers with different degrees of mismatch

Primer length (nts)	Expected no. of binding sites	Mismatches in 10mer	Observed no. of binding sites of 10mer with mismatches
6	4394	4	290 025
7	1098	3	55 102
8	274	2	6 917
9	68	1	509
10	17	0	18

However, not all of the binding sites will be efficient priming sites. The efficiency of priming DNA synthesis depends on the number and positions of the mismatches. Binding sites with fewer mismatches are probably better priming sites than those with more. Also binding sites where mismatches are closer to the 5′ end would probably be more efficient priming sites than those with mismatches closer to the 3′ end. However, in our experiments a mismatch at position 6 from the 5′ end does not prevent the priming of cDNA synthesis (Table 2). Non-degenerate anchored T-primers are tolerant of a mismatch at their 3′ penultimate base (3), and even mismatches at the 3′ end do not necessarily prevent DNA synthesis (17). Therefore there are probably many more priming sites than previously predicted (1).

Thus, in the PCR step, a single 10mer primer with mismatches probably primes many more PCR products than there are bands visible on a DDRT-PCR gel. These products would be amplified

Figure 1. DDRT-PCR of different amounts of globin RNA in the presence and absence of 2 μg heterologous total RNA [total RNA from *G.rostochiensis* (**A**) or potato leaf (**B**)]. Amounts of globin RNA are: 6, 10 ng; 5, 1 ng ; 4, 100 pg; 3: 100 pg; 2, 10 pg; 1, 1 pg; 0, no globin RNA. When mixed with 2 μg heterologous RNA these amounts of globin RNA contributed 11, 1.2, 0.12, 0.012, 0.0012 and 0.00012% α- or β-mRNA to total mRNA respectively (assuming 5% of heterologous total RNA is mRNA, and that the average size of a mRNA is ~3× that of α- or β-globin). Predicted globin products and their sizes are indicated by arrow heads. DDRT-PCR was done as previously described (3). Total RNA was extracted as previously described (21). Globin mRNA purified from rabbit reticulocyte polyribosomes was purchased from Life Technologies Inc. and consists of ~40% α-globin mRNA, ~40% β-globin mRNA and ~20% of a mixture of other mRNAs.

by the PCR until one of the substrates is depleted. There are only enough dNTPs to synthesise ~0.17 pmol DNA, but enough 10mer and T-primer to make 10 and 50 pmol DNA respectively (assuming a mean product size of 240 bp and one of each primer in each DNA fragment). Because of this the dNTPs will probably be depleted before the 40 cycles are complete. The PCR products that are displayed are those that have been amplified to a level necessary to expose the X-ray film. The PCR products representing abundant mRNAs will reach this level earlier because the cDNAs from which they were amplified were present in greater amounts at the start of PCR. Thus the DDRT-PCR procedure is strongly biased towards detecting abundant mRNAs.

However, competition is not the only factor limiting the sensitivity of DDRT-PCR as shown by tests using pure globin mRNA. Under these conditions there must be little competition. DDRT-PCR with many primer combinations was unable to detect globin targets in amounts equivalent to a rare mRNA present in a standard DDRT-PCR template (~4 pg, Tables 1 and 2). Therefore, DDRT-PCR is limited by an inability to reliably amplify low copy number templates, perhaps because 10mers are inefficient primers for PCR as suggested by Liang and Pardee (1).

It has been suggested that a selection of 312 primer combinations would allow an almost complete analysis of all mRNAs in a cell (2). Our results suggest that the standard DDRT-PCR technique is not able to display efficiently the vast majority of mRNAs in a cell i.e. the rare mRNAs, regardless of how many 10mer primers are used. On the other hand our results also suggest that only a few primer combinations will be needed to display all abundant mRNAs.

Although any mRNA display technique that displays more than one band from each reaction is likely to show bias towards more abundant mRNAs, this bias in DDRT-PCR may be reduced and the sensitivity to less abundant mRNAs/cDNAs improved in three ways. Firstly, by reducing the competition in PCR, secondly, by increasing the efficiency of PCR and thereby its sensitivity to low copy number templates and thirdly, by the addition of more of the limiting substrates (but this approach alone will increase gel background). Competition could be reduced by using simpler templates for PCR and/or more specific priming strategies for PCR. The sensitivity to low copy number templates may be improved by the use of more efficient primers for PCR. All of these could be achieved by the use of longer primers which have higher binding energies and are more selective in their priming (when used at higher annealing temperatures).

It is interesting to note that a different RNA fingerprinting strategy (18) does use longer primers for both reverse transcription and PCR, but is known not to reliably detect rare mRNAs (19). A modification of this method based on nested PCR has been designed to improve sensitivity to lower copy number mRNAs, but has not yet been extensively tested (20).

This study has highlighted the bias of DDRT-PCR towards abundant mRNAs and has provided experimental data showing that this bias is extreme. We hope that this study will help in modifications of DDRT-PCR such that it can become a more efficient tool in the analysis of gene expression.

ACKNOWLEDGEMENTS

We thank Ray White, David Hornby, Mike Hahn and Arie Rosner (Rothamsted) for advice and support. We are grateful to the Gottlieb Daimler- and Karl Benz-Stiftung for a research fellowship and the Commission of the European Communities for a fellowship in the framework of the EEC biotechnology programme (grant no. 040870) for U. Schlichter. IACR receives grant-aided support from the Biotechnology and Biological Sciences Research Council of the United Kingdom.

REFERENCES

1 Liang,P., and Pardee,A. (1992) *Science*, **257**, 967–971.
2 Bauer,D., Muller,H., Reich,J., Reidel,H., Ahrenkiel,V., Warthoe,P. and Strauss,M. (1993) *Nucleic Acids Res.*, **21**, 4272–4280. .
3 Liang,P., Averboukh,L., and Pardee,A. (1993) *Nucleic Acids Res.*, **21**, 3269–3275.
4 Alberts,B., Bray,D., Lewis,J., Raff,M., Roberts,K. and Watson,J.D. (1989) *Molecular biology of the cell.* 2nd edition. Garland Publishing Inc.
5 Buffard,D., Vaillant,V. and Esnault,R. (1982) *Eur. J. Biochem.*, **126**, 129–134.
6 Kamalay,J.C. and Goldberg,R.B. (1980) *Cell*, **19**, 935–946.
7 Asher,M.J.C. and Shipton,P.J. (eds) (1981) Introduction in: *Biology and Control of Take-all: Take-all Symptoms.* Academic Press, London.
8 Goddijn,O.J.M., Lindsey,K., van der Lee,F.M., Klap,J.C. and Sijmons, P. (1993) *Plant J.*, **4**, 863–873.
9 Pfitzner,U.M. and Goodman,H.M. (1987) *Nucleic Acids Res.*, **15**, 4449–4465.
10 Hooft van Huijsduijnen,R.A.M., Cornelissen,B.J.C., van Loon,L.C., van Boom,J.H., Tromp,M. and Bol,J.F. (1985) *EMBO J.*, **4**, 2167–2171.
11 Conkling,M.A., Cheng,C., Yamamoto,Y.T. and Goodman,H.M. (1990) *Plant Physiol.*, **93**, 1203–1211.
12 Williams,J.G.K., Kubelik,A.R., Livak,K.J., Rafalski,J.A. and Tingey,S.V. (1990) *Nucleic Acids Res.*, **18**, 6531–6535.
13 Welsh,J., Chada,K., Dalal,S.S., Cheng,R., Ralph,D. and McClelland,M. (1992) *Nucleic Acids Res.*, **20**, 4965–4970.
14 Anderson,P.A., Tyler,B. and Pryor,A. (1992) *Exp. Mycology*, **16**, 302–307.
15 Arnau,J., Housego,A.P. and Oliver,R.P. (1994) *Curr. Genet.*, **25**, 438–444
16 Chou,Q., Russell,M., Birch,D.E., Raymond,J. and Bloch,W. (1992) *Nucleic Acids Res.*, **20**, 1717–1723.
17 Kwok,S., Kellogg,D.E., McKinney,N., Spasic, D., Goda, L., Levenson,C. and Sninsky,J.J. (1990) *Nucleic Acids Res.*, **18**, 999–1005.
18 Welsh,J. and McClelland,M. (1990) *Nucleic Acids Res.*, **18**, 7213–7218.
19 McClelland,M., Mathieu-Daude,F. and Welsh,J. (1995) *Trends Genet.*, **11**, 242–246.
20 Ralph,D., McClelland,M. and Welsh,J. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 10710–10714.
21 Logeman,J., Schell,J. and Willmitzer,L. (1987) *Anal. Biochem.*, **163**, 16–20.