3478–3480 *Nucleic Acids Research, 1996, Vol. 24, No. 17*

Preparation of a differentially expressed, full-length cDNA expression library by RecA-mediated triple-strand formation with subtractively enriched cDNA fragments

Theodorus B. M. Hakvoort*, Johannes A. A. Spijkers, Jacqueline L. M. Vermeulen and Wouter H. Lamers

Department of Anatomy and Embryology, Academic Medical Centre, University of Amsterdam, Meibergdreef 15, 1105 AZ Amsterdam, The Netherlands

Received April 12, 1996; Revised and Accepted July 10, 1996

ABSTRACT

We have developed a fast and general method to obtain an enriched, full-length cDNA expression library with subtractively enriched cDNA fragments. The procedure relies on RecA-mediated triple-helix formation of single-stranded cDNA fragments with a double-stranded cDNA plasmid library. The complexes were then captured from the solutions using the digoxigenin–antidigoxigenin paramagnetic beads followed by recovery of the enriched double-stranded cDNA expression library. We have observed a linear relation between the capture of full-length cDNAs in the library and the fold enrichment in the subtracted cDNA population.

One approach to learn about the mechanisms underlying functional or structural changes in cells or tissues is to determine the changes in their pattern or rate of gene expression. Techniques aiming at making an inventory of differentially expressed gene products in two populations of cells being compared often employ (PCR) amplification of all gene products to obtain detectable levels, normally as cDNA fragments. The subsequent identification of these differentially expressed and amplified gene products is generally achieved by one of the following two methods. The differential-display method (1,2) relies on amplification of the entire mRNA populations, followed by a direct comparative display of the resulting populations of cDNA on a sequencing gel to visualize differences in prevalence. The analysis has been simplified by dividing the mRNAs to be amplified into subgroups via the choice of the primers. The other approach is based on the elimination of gene products common for both from the two populations by means of subtractive-hybridization followed by an analysis of the enriched products (3,4). This approach (5), which is a straightforward procedure for obtaining cDNA populations which are enriched for up- or down-regulated gene products, was recently modified in our laboratory by introducing the display method as a way to analyze its results (6). Such an inventory does, however, not allow one to discriminate between the gene products that are responsible for the change in the cellular phenotype and those that result from it. To distinguish between these two groups, a functional test with the full-length cDNA is necessary. The feasibility of successfully executing such a laborious test, like for instance a cell proliferation assay, is largely dependent on the complexity of the population of gene products being studied.

The isolation of specific recombinant plasmids from a cDNA library can be done in various ways, such as colony lifting and phage screening. Both procedures are the options of choice when only a few plasmids have to be obtained but have there limitations when larger numbers of recombinant plasmids need to be isolated. A large number of specific cDNAs can be isolated simultaneously in one single experiment by employing the triple-helix-mediated affinity capture. The approach is based on the formation of triple-helical structures which can then be selectively isolated (8–11). These triple-helical structures are stable complexes between single-stranded and double-stranded DNA molecules which, due to their sequence similarity, are held together by Hoogsteen hydrogen bonding. Triple-helix motifs often used are stretches of homopyrimidines and purines. Although these motifs are not present in our enriched cDNA fragments, we did obtain full-length cDNAs in such a capture experiment but with variable results (not shown). The *Escherichia coli* recombinase protein A (RecA protein) facilitates the parallel positioning of the single-stranded DNA fragment in the major groove of the double-stranded DNA molecules and was introduced to stabilize the complex (7,12,13). Again, in these experiments the typical triple-helix motifs were employed. We extended this approach by using cDNA fragments without any sequence constraint. Here we report that the combination of (i) the subtractive-hybridization method (to enrich a population of cDNA fragments for up- or down-regulated gene products) and (ii) the RecA-mediated formation of triple-stranded complexes between enriched cDNA fragments and double-stranded plasmids containing the full-length cDNA, yields a rapid and powerful procedure to prepare specifically enriched cDNA libraries.

In our laboratory we are studying the altered gene expression in rat liver after partial hepatectomy. cDNAs were prepared from normal and regenerating liver. cDNA fragments of up- and

* To whom correspondence should be addressed

down-regulated gene products were obtained by the subtractive-hybridization method (6). These enriched cDNA fragments were amplified by PCR in the presence of 8 μM digoxigenin-11-dUTP (Boehringer Mannheim). The dATP, dCTP, dGTP and dTTP concentrations were kept at 200 μM. The PCR-amplified, digoxigenin-labelled cDNA fragments were purified on a Qiaquick spin PCR column (Qiagen) according to the manufacturer's instructions. A cDNA expression library was generated from adult rat liver utilizing the 'Great Lengths' cDNA synthesis kit (Clontech) with minor modifications. First strand synthesis was started with a dT-oligomer extended with a *Not*I-restriction site. The double-stranded cDNAs were ligated with *Bst*XI adaptors and, following restriction with *Not*I, cloned unidirectionally into the multiple cloning site of the eukaryotic expression vector pcDNA3 (Invitrogen). The recombinant-plasmid cDNA library was introduced into Epicurian Coli XL1-Blue MRF′ cells (Stratagene) by electrotransformation and cultured on Luria-broth ampicillin (LB/amp) agar plates. Plasmid DNA from 400 000 independent transformants was isolated by the alkaline lysis method. Triple-strand complex formation between the digoxigenin-labelled cDNA-fragments and the double-stranded DNA of the pcDNA3-based expression library was accomplished as described (7,13) with modifications. The digoxigenin-labelled double-stranded cDNA fragments (0.25 μg) were denatured in TE (pH 8.0) in the presence of 10 μg bovine serum albumin (5 min at 100°C) and immediately chilled on ice (5 min). Ten μl of 10× buffer (20 mM $CoCl_2$, 80 mM $MgCl_2$, 300 mM Tris–HCl pH 8.0) and 5 μl (2 μg/μl) RecA protein (Biolabs) were added. The solution was mixed by 'finger tapping' and 16 mM 5′-[γ-thio]ATP (a non-hydrolysable ATP analog) was added after 30–60 s to activate the RecA protein. Following RecA-filament formation and binding of the digoxigenin-labelled single-stranded cDNA fragments (10 min, 37°C), 0.65 μg double-stranded recombinant plasmid was added. After 20 min at 37°C, the formation of triple-stranded complexes was quenched by adding 5 μg *Bst*EII-restricted, phenol-extracted λ DNA (3 min, 37°C). The formation of non-specific triple-stranded complexes was minimized by the addition of NaCl to 50 mM. The RecA filaments were digested by proteinase K (0.2 mg/ml) in the presence of 0.2% $NaDodSO_4$ (7 min, 37°C). Proteinase K was inactivated by the addition of phenylmethylsulfonyl fluoride (PMSF) to 1 mM. The complete incubation mixture was added to prewashed (four times in 10 mM Tris–HCl pH 7.5, 1 mM EDTA and 300 mM NaCl) anti-digoxigenin magnetic particles (Boehringer) (300 μg; 10 mg/ml) for 10 min at room temperature. Utilizing a magnetic-particle separator, the particles were sequentially washed with buffer I (10 mM Tris–HCl pH 7.5, 1 mM PMSF, 1 mM EDTA, 300 mM NaCl and sonicated herring-sperm DNA at 10 μg/ml; 10× 0.5 ml fractions) and buffer II (10 mM Tris–HCl pH 7.5, 1 mM EDTA, 50 mM NaCl; 4× 0.5 ml fractions). The bound recombinant plasmids were eluted in buffer III (30 mM Tris–HCl pH 8.8, 0.1 mM EDTA) at 65°C (2× 0.5 ml fractions). The buffer III eluates were combined and the volume was reduced to 100–200 μl by butanol-2 extraction. Recombinant plasmids were precipitated with 2.5 vol ethanol in the presence of 10 μg glycogen (Boehringer) and 300 mM NaOAc (pH 5.2), overnight at –20°C. The precipitate was washed once with 70% ethanol, air-dried and resuspended in 10 μl TE (pH 8.0). Aliquots of 3 μl of enriched recombinant plasmids were introduced into XL1-Blue MRF′ cells by electrotransformation, plated directly onto 135 mm Hybond-N filters (Amersham) and cultured on LB/amp agar plates (14). A replica of the master filter was made on a second Hybond-N filter and subsequently used for colony screening with selected enriched cDNA fragments as the probe. After each hybridization screening the probe was removed with NaOH and the filter was re-used. The master filter was used for recovery of bacteria containing the selected recombinant plasmid with the full-length cDNA.

The enrichment obtained with the subtractive-hybridization and the RecA-mediated triple strand formation were compared. The fold enrichment resulting from the successive subtractive-hybridization steps, was assessed by Southern blot analysis. Equal amounts of amplified cDNA fragments (2 μg) from the initial and the final round of subtraction were, after electrophoresis on 1.5% agarose and transfer to Hybond-N, probed with a selected enriched cDNA fragment. Probes were generated from up-regulated cDNA fragments, extracted from the display gel, PCR amplified and subcloned into a TA-cloning vector (Invitrogen). The insert was isolated and 100 ng was random-prime labelled with [^{32}P]dCTP. Hybridization signals, after a high stringency wash (0.1× SSC/0.1% $NaDodSO_4$ at 68°C), were quantitated employing a phosphorimager. The enrichment was determined for each selected probe separately by calculating the ratio of the hybridization signal found after the final round of subtractive hybridization over that present in the initial material.

The fold enrichment in the full-length cDNA library was determined by colony-hybridization screening with the same set of selected probes used in the Southern blot analysis described above. Aliquots of recombinant plasmids from the original and of the enriched population were introduced into XL1-Blue MRF′ cells and propagated on Hybond-N filters. Colonies were screened with the respective up-regulated cDNA fragments for the presence of the corresponding full-length cDNA. The number of positive colonies was counted and divided by the total number of colonies present on the filter. The fold enrichment in the cDNA library was calculated as the ratio of the frequency of positive colonies in the enriched over that in the initial recombinant-plasmid population. Sequence analysis of the 5′-end and restriction mapping of the selected known cDNAs was used to verify that the positive clones in the enriched cDNA library indeed represented full-length cDNAs (data not shown).

Figure 1 shows the fold enrichment in the full-length cDNA library that resulted from selection with the triple-strand complexation method, as a function of the fold enrichment of cDNA fragments that was achieved by applying the subtractive-hybridization procedure. The data were obtained using cDNA fragments from transferrin, α_1-antitrypsin and two novel liver regeneration-associated cDNAs (R-18-4 and R-5-25). All four gene products were up-regulated after partial hepatectomy as was demonstrated by Northern blot analysis (not shown). Product R-5-25 is enriched 45-fold in the subtracted cDNA fragments and in the cDNA library. To demonstrate the dynamic range of the RecA-mediated triple-strand selection, we also performed this method with a cDNA fragment of a moderately abundant (frequency 0.15%) hepatic gene product, viz. glutamine synthetase (GS). The GS-fragment (nt 2186–2697) was labelled with digoxigenin–dUTP and used in the triple-strand complexation method, under conditions as described for the population of enriched cDNA fragments. The initial liver cDNA library and the enriched GS-cDNA plasmids were introduced into XL1-Blue MRF′ cells and colonies were screened with the 700 bp 5′-*Eco*RI fragment of the rat GS cDNA. The frequency of positive clones for GS in

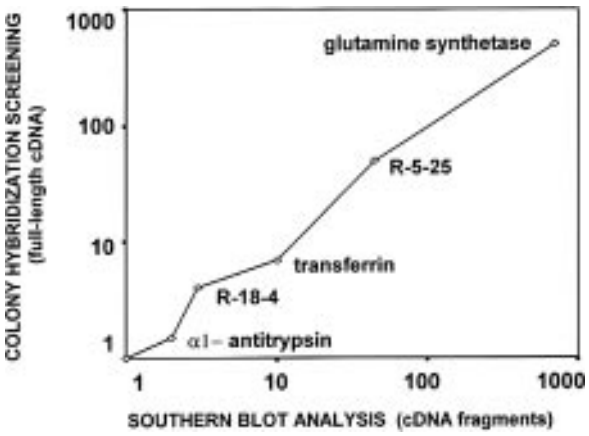

Figure 1. Enrichment of a full-length cDNA library as a result of selection with a population of subtractively enriched cDNA fragments via the RecA-mediated affinity capture method. The fold enrichment of transferrin, α_1-antitrypsin and two novel liver regeneration-associated cDNAs (R-5-25 and R-18-4) by subtractive hybridization was determined by Southern blot analysis of equal amounts of cDNA fragments before and after subtraction. The prevalence of full-length cDNAs in the original and the enriched libraries was determined by colony hybridization screening. The cDNAs shown were selected from a display gel, except for the glutamine–synthetase cDNA fragment which was included to demonstrate the dynamic range of the triple-strand complexation method.

the initial cDNA library was taken for the Southern blot analysis value. The fold enrichment in the GS-cDNA population was calculated as the ratio of the frequency of positive colonies in the enriched over that in the initial cDNA library. In order to obtain full-length cDNAs a high quality cDNA library is required. For the RecA-mediated affinity capture of GS-cDNA, a 3′-end fragment was used to select for GS. We noticed that using a 5′-end hybridization probe two populations of positive stained colonies (strong and weak) were present. The strong staining intensities represented the complete GS-cDNA and accounted for 11% of the total positive clones.

The near linear relation in Figure 1 shows that the step from enriched cDNA fragments to an enriched plasmid-cDNA library was made without loss of the enrichment that was achieved during the subtractive-hybridization procedure. This demonstrates that the conversion step approaches a near 100% efficiency. The linearity of the relation between the fold enrichment in the colony-hybridization procedure and the fold enrichment in the Southern blot signal suggest that the method should be applicable over another 1–2 decades, which is necessary to identify the very rare messages. Furthermore, the entire procedure, as described above, does not require the preparation of single-stranded plasmid DNA and subsequent conversion to double-stranded DNA as often required in other methods. The presented RecA-mediated affinity capture therefore provides a fast-single step conversion of enriched cDNA fragments from a subtractive-hybridization procedure, or the selected fragments from the differential display PCR method, to a corresponding enriched full-length cDNA library. The availability of such a subset of enriched full-length cDNAs makes functional screening assays for e.g. proliferative activity manageable. Such assays are currently in progress. The procedure as described above is, however, applicable more generally.

ACKNOWLEDGEMENT

The authors like to thank Dr R. A. F. M. Chamuleau for his advice and careful reading of this manuscript.

REFERENCES

1 Liang, P. and Pardee, A.B. (1995) *Curr. Opin. Immunol.*, **7**, 274–280.
2 McMlelland, M., Mathieu-Daude, F. and Welsh, J. (1995) *Trends Genet.*, **11**, 242–246.
3 Fargnoli, J., Holbrook, N.J. and Fornace, A.J. (1990) *Anal. Biochem.*, **187**, 364–373.
4 Lisitsyn, N.A. (1995) *Trends Genet.*, **11**, 303–307.
5 Wang, Z. and Brown, D.D. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 11505–11509.
6 Hakvoort, T.B.M., Leegwater, A.C.J., Michiels, F.A.M., Chamuleau, R.A.F.M. and Lamers, W.H. (1994) *Nucleic Acid Res.*, **22**, 878–879.
7 Honigberg, S.M., Rao, B.J. and Radding, C.M. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 9586–9590.
8 Tagle, D.A., Swaroop, M., Lovett, M. and Collins, F.S. (1993) *Nature*, **361**, 751–7453.
9 Rounds, D., Brueckner, M. and Ward, D.C. (1995) *Genomics*, **29**, 616–622.
10 Sonti, S.V., Griffor, M.C., Sano, T., Narayanswami, S., Bose, A., Cantor, C.R. and Kausch, A.P. (1995) *Nucleic Acid Res.*, **23**, 3995–3996.
11 Johnson, A.F., Wang, R., Ji, H., Chen, D., Guilfoyle, R.A. and Smith, L.M. (1996) *Anal. Biochem.*, **234**, 83–95.
12 Radding, C.M. (1991) *J. Biol. Chem.*, **266**, 5355–5358.
13 Rigas, B., Welcher, A.A., Ward, D.C. and Weissman, S.M. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 9591–9595.
14 Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1987) *Current Protocols in Molecular Biology*, Wiley Interscience, New York.