DNA RESEARCH 4, 61–66 (1997)

High Efficiency Selection of Full-length cDNA by Improved Biotinylated Cap Trapper

Piero CARNINCI,[1,2] Arthur WESTOVER,[1] Yoko NISHIYAMA,[1] Tomoya OHSUMI,[1] Masayoshi ITOH,[1] Sumiharu NAGAOKA[1] Nobuya SASAKI,[1] Yasushi OKAZAKI,[1] Masami MURAMATSU,[1] Claudio SCHNEIDER,[3] and Yoshihide HAYASHIZAKI[1,2,*]

Genome Science Laboratory, Tsukuba Life Science Centre, The Institute of Physical and Chemical Research (RIKEN), 3-1-1 Koyadai, Tsukuba, Ibaraki 305, Japan,[1] Tsukuba University Medical School, 1-1-1, Tennodai, Tsukuba, Ibaraki 305, Japan,[2] and Laboratorio Nazionale CIB, Area Science Park, 99 Padriciano, 34012 Trieste, Italy[3]

(Received 25 December 1996)

Abstract

We report here an improved protocol for the preparation of full-length cDNA libraries that improves the previously reported method (Carninci, P., Kvam, K., Kitamura, A. et al. 1996, *Genomics*, **137**, 327–336), that allows long cDNAs to be cloned more efficiently. One potential disadvantage of the original biotinylated CAP trapper protocol is the exposure of mRNA to chemical and enzymatic attacks during the biotinylation of the cap structure, before the first-strand cDNA synthesis (and selection of full-length cDNA by biotinylated cap). Here, we show that the biotinylation of the cap structure is very specific and effective even if biotinylation is performed on the mRNA/cDNA hybrid produced by the first-strand cDNA synthesis reaction. Consequently, mRNA remains protected from chemical and enzymatic degradation during the overnight biotinylation step, thus making it possible to select full-length cDNAs of longer average size. We herein report the efficiency and specificity of the new version of the protocol for cap structure biotinylation and capture of full-length cDNA.

Key words: full-length cDNA; biotin technology; cDNA library

1. Introduction

Isolation of full-length cDNAs is a time-consuming task, and is the rate-limiting step in gene cloning and large-scale cDNA sequencing projects. This problem is due to the reduced efficiency with which reverse transcriptase reaches the cap site, especially for mRNAs that present a stable secondary structure, and to the lack of techniques to select the full-length cDNAs. In recent years, a few protocols have been proposed to construct full-length cDNA libraries; a more detailed discussion of these protocols is given in ref.[1] Several methods, however, are not satisfactory in terms of efficiency and bias in the representation of the cDNAs. In fact, the oligo-capping[2] and the Capfinder (CLONTECHniques technical bulletin, January 1996) methods both require the use of PCR, that may lead to the loss of rare or long cDNAs. Additionally, the oligo-capping method involves the use of RNA ligase that can cause an additional biased representation of clones due to the sequence- and secondary structure-dependent activity of this enzyme.[3,4] A subsequent modification of the oligo-capping techniques by omitting the PCR amplification step,[5] still uses RNA ligase, and thus is still subject to the above mentioned drawbacks.

Two alternative approaches to selecting full-length cDNA by the cap structure, including our previous method, have been described.[1,2] By both methods, RNase digestion of the first-strand cDNA reaction product is used to physically remove the ssRNA stretches, together with the cap structure, from the cDNA/mRNA hybrids that do not contain full-length cDNAs. In this way, only the full-length cDNAs are selected through the cap structure of the hybridized mRNA, while partially synthesized cDNAs are lost due to physical removal of the cap structure together with the most proximal 5′ sequence of the mRNA from the partial cDNA/mRNA hybrid. The first method, called CAPture, is based on selection of the cap by a cap-binding protein.[6] This method is, however, not satisfactory due to its low yield and the requirement of reagents that are not commercially available. Additionally, the efficiency of this technique to select the full-length cDNAs has not been completely

Communicated by Mituru Takanami

* To whom correspondence should be addressed. Tel. +81-298-36-9145, Fax. +81-298-36-9098

proven. Consequently, we developed an alternative system, the full-length cDNA selection by the biotinylated cap trapper, and reported the results in our previous paper.[1] This system is based on the chemical addition of a biotin group to the cap structure. This step is followed by RNase I digestion, a ribonuclease that can cleave single-stranded (ss) RNA at any site,[7] followed by selection of full-length cDNA. The library produced by this method contained a very high proportion of full-length cDNAs and was more than satisfactory in terms of the overall yield, starting from a reduced amount of mRNA and without involving PCR amplification.[1] Since the biotinylation step was performed under non-denaturing conditions, in which mRNA may undergo enzymatic or chemical degradation, we report here the improved version of the cap biotinylation protocol. In this new version, the biotinylation takes place at a later stage of the procedure, so that damaging the mRNA does not affect the quality of the library.

2. Material and Methods

2.1. RNA preparation

Total RNA was prepared from brain mouse B6/J by modification of a standard procedure. Briefly, the tissues were homogenized in 10 ml Solution D[8] followed by the addition of 1 ml of 2 M sodium acetate (pH 4.0) and organic extraction with one volume of acid phenol/chloroform (5:1). After extraction, RNA was precipitated from the aqueous phase by adding 1 volume of isopropanol. The samples were then incubated for 1 hr on ice, after which the RNA was pelleted by centrifugation at 4,000 rpm for 15 min in a refrigerated bench centrifuge. The pellet was washed with 70% ethanol and the RNA was resuspended in 4 ml of water. To remove polysaccharides, selective CTAB (cetyltrimethylammonium bromide)[9] the mRNA was precipitated. After RNA resuspension in 4 ml of H_2O, 1.3 ml of 5 M NaCl was added and the RNA was precipitated again by adding 16 ml of a solution containing 1% CTAB, 4 M UREA, 50 mM TRIS, pH 7.0. After 15 min of centrifugation at 4,000 rpm at room temperature, RNA was resuspended in 4 ml of 7 M guanidine-Cl. RNA was finally precipitated by adding two volumes of ethanol. After 1 hr of incubation on ice, the samples were centrifuged at 4,000 rpm for 15 min. The pellet was washed with 70% ethanol and resuspended in water. RNA purity was monitored by reading the OD ratio 260/280 (> 1.8) and 230/260 (< 0.45).[10] Messenger RNA was prepared using a commercial kit, PolyA-Quick (Stratagene, USA), starting from CTAB-purified total RNA.

2.2. First-strand cDNA preparation

First, 1–2 μg of mRNA was denatured at 65°C for 10 min, together with 1 μg of the first-strand *Xho* I primer-adapter 5′$(GA)_8$AACTAGTCTCGAG $(T)_{16}$MN3′ (where N is any nucleotide and M is G, A, or C). Subsequently, the reverse transcription reaction was carried out in a final volume of 20 μl of buffer containing 50 mM Tris-HCl, pH 8.3, 75 mM KCl, 4.5 mM $MgCl_2$, 10 mM DTT, dATP, dTTP, dGTP, and dCTP at a final concentration of 3 mM (total). The reaction also contained 1 μl of [α-^{32}P]dGTP (3000 Ci/mMol, 10 mCi/μl, Amersham, UK) for radio labelling. The annealing of the primer was done on a thermal cycler: samples were preheated at 45°C in a MJ Research thermal cycler and then approximately 400 units of Superscript II (Gibco BRL, USA) was added to the reaction mixture. The annealing profile included an initial negative ramp of 1°C per min until 35°C was reached. After further incubation at 35°C for 5 min, the temperature was raised to 45°C for 1 hr, followed by 10 min of incubation at 50°C.

To follow the yield of the first-strand synthesis, 0.5 μl of the radiolabelled reaction aliquot was spotted on DE-81 paper and the incorporation (cpm) was calculated by measuring the radioactivity before and after three washings with 0.5 M phosphate buffer, pH 7.0.[10]

2.3. Biotinylation of diol groups of RNA

The biotinylation reaction was performed either before or after the cDNA synthesis. When performed after the first-strand synthesis, to the sample 0.8 μl 0.5 M EDTA and 1.3 μl 5 M NaCl were added. The sample was extracted once with phenol chloroform and chloroform, followed by ethanol precipitation. The pellet was washed with 70% and 80% ethanol and resuspended in H_2O. The common biotinylation protocol is as follows: the samples were subsequently oxidized on the diol groups of mRNA in 25 μl of 66 mM sodium acetate buffer, pH 4.5, containing 5 mM $NaIO_4$ (freshly prepared) as the oxidizing agent. The oxidation reaction was carried out on ice in the dark for 45 min. The oxidized mRNA/cDNA hybrid was subsequently precipitated by the addition of 2.5 μl of 5 M LiCl, 0.5 μl of 10% SDS and 1 volume of isopropanol. After 30 min incubation at −20°C, mRNA/cDNA hybrid was precipitated by centrifugation at 15,000 rpm for 15 min at +4°C. The pellet was washed with 70% ethanol and then resuspended in 25 μl of water. Next, 2.5 μl of 1 M sodium acetate, pH 6.1, 2.5 μl of 10% SDS and 75 μl of 10 mM biocytin hydrazide (freshly dissolved in water) were added to the sample. The oxidized diol group on RNA was then biotinylated by overnight incubation at room temperature (22 to 26°C). Finally, nucleic acids was precipitated again by addition of 2.5 μl of 5 M NaCl, 37.5 μl of 1 M sodium acetate, pH 6.1, and 2.5 volumes of ethanol. After a 1-hr incubation on ice, biotinylated sam-

ple was precipitated by 15 min centrifugation at +4°C. The pellet was washed once with 70% ethanol and once with 80% ethanol and finally resuspeded in RNase-free water.

2.4. *RNase protection of full-length cDNA and TAP treatment*

RNase digestion of the 0.5 μg aliquots of the first-strand cDNA reaction was performed with RNase I (RNase ONE,™ Promega, USA) in a volume of 20 μl water and 20 units of RNase I in the buffer recommended by the maker at 37°C for 30 min. To remove cap, 150 units of TAP were used (Nippon Gene, Japan) in 20 μl volume in the buffer recommended by the manufacture.

2.5. *Blocking of magnetic beads and capturing of nucleic acids*

First-strand full-length cDNA/mRNA hybrid was captured on magnetic porous glass (MPG) particles coated with streptavidin (CPG, New Jersey, USA). Before binding of the nucleic acids, 100 μl of beads (1% suspension; 1 mg of beads can bind 800 pmol of a biotinylated 25-mer oligonucleotide) were blocked by adding 10 μl of 40 μg/μl DNA-free tRNA and incubating on ice for 1 hr with occasional gentle vortexing. Just before nucleic acid capture, the beads were separated using a magnetic stand and the supernatant was removed by pipetting. All subsequent capture, washings and release procedures were performed with the help of a magnetic stand. After the blocking step, the beads were washed three times with 500 μl 2 M NaCl and 50 mM EDTA, pH 8.0, and finally resuspended in 100 μl of 2 M NaCl and 50 mM EDTA, pH 8.0. Finally, cDNA (approximately 0.5 μg) was captured at room temperature for 30 min with continuous gentle mixing to prevent bead sedimentation, and in the presence of 100 μg of tRNA as a carrier. After removal of unbound cDNA, the beads were washed four times with 500 μl of 2 M NaCl and 50 mM EDTA, pH 8.0, containing 100 μg/ml tRNA. Finally, cDNA was eluted from beads by treatment with 100 mM NaOH, 5 mM EDTA and was then loaded on an alkaline gel.[10]

3. Results and Discussion

In the original full-length cDNA cloning by biotinylated cap trapper cloning techniques,[1] mRNA was used as the substrate for the chemical biotinylation of the cap structure. Biotin groups were added to the diol groups of mRNA, at the cap structure and the 3′ end of mRNA (Fig. 1A) before synthesis of the first-strand cDNA. The two-stage chemical biotinylation (Fig. 1B) was the initial step in this protocol. Our new protocol is shown in Fig. 1C. From the RNase I protection of hybridized mRNA to cloning

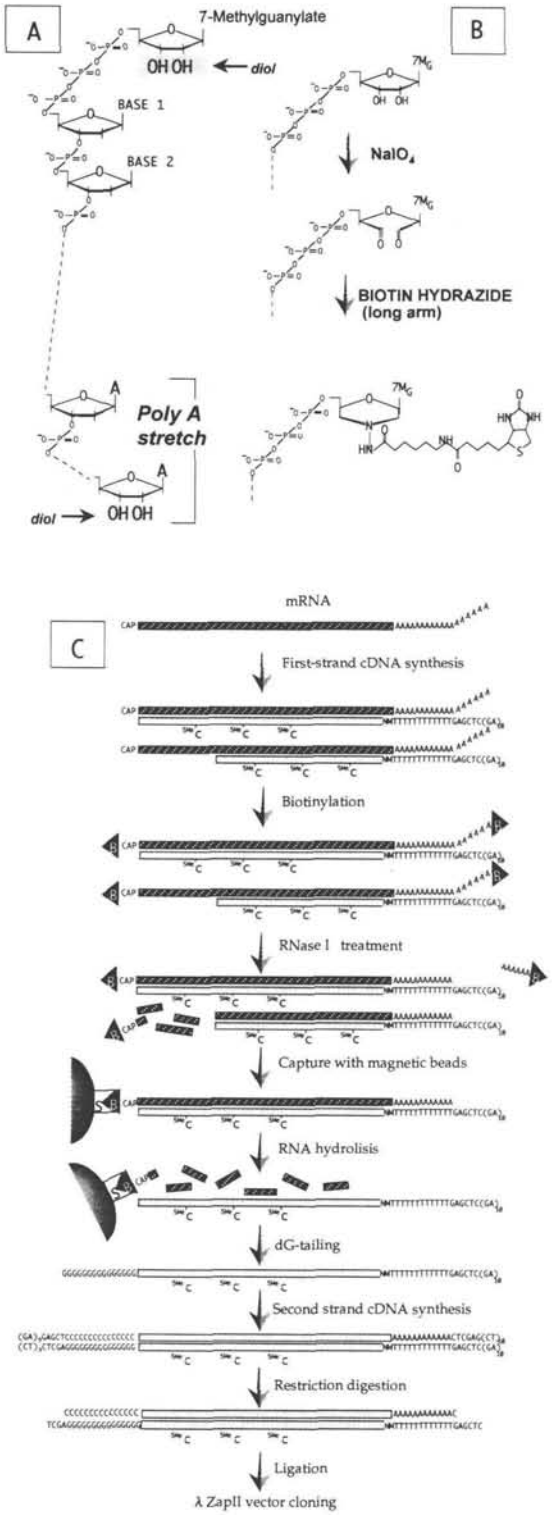

Figure. 1 Flowchart of the proposed protocol for the preparation of the cDNA library by using the modified biotinylated cap trapper. A, two diol groups are present on the cap structure and 3′ end of mRNA. B, biotinylation reaction of diol groups (shown only for the cap structure). C, flowchart of the preparation of a cDNA library by improved biotinylated cap trapper.

of cDNA, the procedures were the same as the previous protocol, which included the following steps: capture of full-length cDNA by streptavidin-coated magnetic beads, release of full-length cDNA from beads, oligo dG tailing by terminal transferase, and second-strand cDNA synthesis and cloning in lambda phage vector.[1] However, during the initial biotinylation step, mRNA was incubated overnight at room temperature in a slightly acidic buffer. This incubation could cause a partial degradation of mRNA and by the selection of full-length cDNA, it may result in the selective loss of cDNAs corresponding to longer full-length cDNAs. To ensure the higher reproducibility of the protocol and to clone even long cDNAs at high efficiency, we modified the previous protocol by biotinylating the cap structure of mRNA after the synthesis of the first-strand cDNA (Fig. 1). With this improved protocol, the first step of the library preparation is synthesis of the first-strand cDNA, followed by the two-step biotinylation of the cap structure and the 3′ end of the hybridized mRNA. In the new protocol, full-length cDNAs thus protects mRNAs from the degradation due to any damaging effect of chemical reagents during the biotinylation step. Under the protection of mRNA by forming a cDNA/RNA hybrid, even if biotinylation reagents cause a nick on the RNA strand, it is not likely to affect the subsequent full-length cDNA capture. The efficiency of the subsequent cDNA cloning steps was previously described and discussed.[1]

To demonstrate the advantage of the improved protocol, we first show that the first-strand cDNAs are longer when the biotinylation reaction was performed after, instead than before (as in the previous protocol), the synthesis of cDNA. In fact, when mRNA hybridized to cDNA was biotinylated after the first-strand cDNA reaction, the average size of cDNA was longer, up to 9.4 Kb (Fig. 2, lane 1) than the sample in which the cDNA was synthesized by using biotinylated mRNA (Fig. 2, lane 2). These cDNAs are on the average shorter, and extends up to the 6.5 Kb marker. This increase in the average size of the cDNA is likely to correspond to increased abundance of long full-length cDNAs in the obtained library.

We next demonstrated that the biotinylation reaction of the cap structure and the 3′ end of cDNA is effective when mRNA is hybridized and protected by the first-strand cDNA. To examine the specificity of biotinylation, and to exclude any reactivity from other groups of mRNA or cDNA under new conditions, we treated aliquots of the hybrid biotinylated mRNA/radiolabelled cDNA with RNase I, that cleaves ssRNA at any base, and/or TAP (tobacco acid pyrophosphatase), that removes the cap structure. The results are shown in Fig. 3. After the four combination of treatment with and/or without RNase I and TAP, the cDNA/mRNA hybrid was captured by streptavidin-coated magnetic beads. Subsequently, the captured cDNAs were released by alkali treatment and

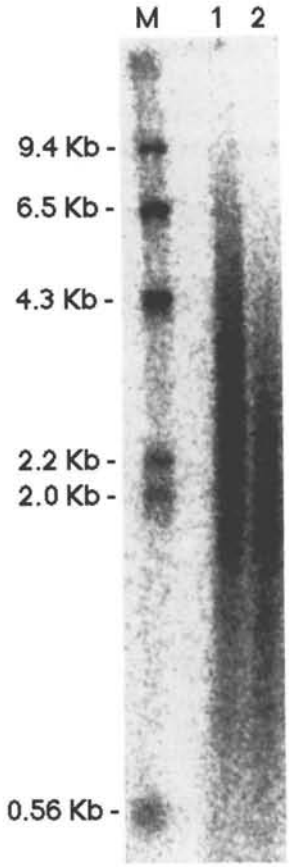

Figure 2. Increase of the length of the cDNA when the biotinylation follows synthesis of first-strand cDNA. Lane 1, cDNA synthesis before biotinylation. Lane 2, cDNA synthesis after biotinylation. M, Lambda-*Hind*III molecular weight marker, the length in Kb is indicated on the side.

loaded onto alkaline gel electrophoresis to visualize the size and the final yield of the captured cDNA (Fig. 3). In lane 1, the biotinylated mRNA/cDNA was captured without any treatment by RNase I and TAP, suggesting visualization of total cDNA including cDNA with partial and complete length. cDNA was recovered at very high yield, as confirmed by measuring the radioactivity (cpm) captured by streptavidin-coated magnetic beads, that reached approximately 90% of the input. In lane 2, the mRNA/cDNA hybrid was treated with RNase I, that removes the biotin group at the 3′ end and the cap structure in the absence full-length cDNA synthesis. This cDNA corresponds to the full-length cDNA, and constituted approximately 10% of the starting aliquot of cDNA as evaluated by the radioactivity captured by the beads. In lane 3, cDNA was treated with TAP, to remove only the cap structure. Here, cDNA was captured through the biotin group at the 3′ end of mRNA. The efficiency of capture was high, since the entire population of cDNA

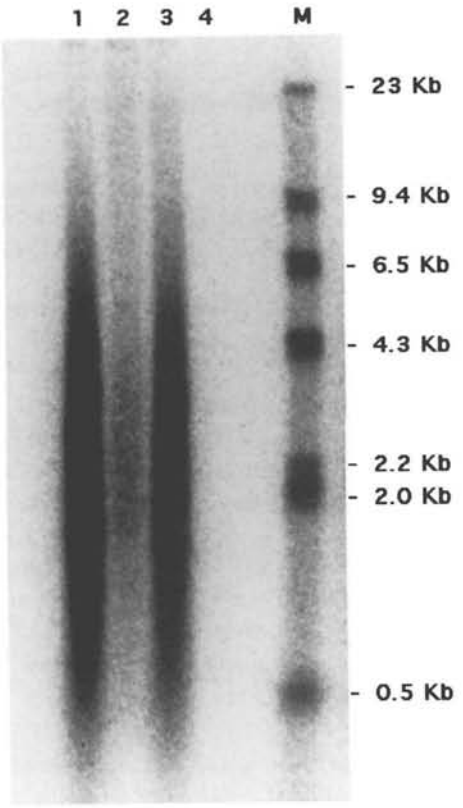

Figure 3. Specificity of the biotinylation for cap and the 3′ end of mRNA. Lane 1, cDNA is captured through both the biotinylated structures present on the cap and the 3′ end of the hybridized mRNA. Lane 2, full-length cDNA is captured through the biotinylated cap structure of the hybridized mRNA after RNase I digestion. Lane 3, cDNA is captured by the 3′ end of the hybridized mRNA after removal of the cap structure by TAP. Lane 4, cDNA is failed to be captured after both RNase I treatment and cap removal by TAP. M, molecular weight markers.

could be captured by the 3′ end-biotin residue. In lane 4, the biotinylated mRNA/cDNA was treated with both RNase I and TAP. This treatment is supposed to remove the biotin groups at both the 3′ end and at the cap site. Accordingly, no cDNA could be captured and subsequently observed on alkaline gel. These data lead us to conclude as follows. The first conclusion is that the biotinylation reaction shows very high specificity for the cap structure and for the 3′ end of mRNA. In fact, any nonspecific biotinylation of cDNA or mRNA is excluded by this result, since there is no captured cDNA in lane 4 of Fig. 3. The second conclusion is that the average size of the cDNA selected by cap trapping is higher than non-selected cDNA. In fact, the average size of cDNA treated with RNase I seems higher than the untreated or TAP-treated sample. In fact, by evaluating the intensity of the cDNA smear, in lane 2 the quantity of cDNA of size around 500 bp is much reduced if compared to lanes 1 or 3, but around 10 Kb or higher, cDNA concentration in lane 2 seems similar to the concentration of the non-treated cDNA of lanes 1 and 3. Thus, this protocol is likely to select full-length cDNAs and increase the average size of the captured cDNAs.

The full-length cDNA selection by improved biotinylated cap trapper thus appears to be a very promising technique to construct cDNA libraries containing a large proportion of long full-length cDNAs. This protocol will contribute to the construction of a second generation of cDNA libraries, that we define as the unbiased full-length cDNA libraries. These cDNA libraries will have the following two characteristics: (i) they will contain a very high proportion of full-length cDNAs and (ii) the frequency of individual will reflect the original frequency of the corresponding mRNAs into the cell without any bias or loss of specific sequences caused by the cloning system.

Acknowledgments: We thank Kimi Akama and Mari Itoh for their technical and secretarial assistance. This study was supported by Special Coordination Funds and a Research Grant for the Genome Exploration Research Project from the Science and Technology Agency of Japan, and a Grant-in-Aid for Scientific Research on Priority Areas and Human Genome Program, from the Ministry of Education and Culture of Japan to Y. H. This work was also supported by a Grant for Research on Ageing and Health, a Grant-in-Aid for a Second Term Comprehensive 10-Year Strategy for Cancer Control and a grant for cardiomyopathy from the Ministry of Health and Welfare of Japan.

References

1. Carninci, P., Kvam, K., Kitamura, A. et al. 1996, High-efficiency full-length cDNA cloning by biotinylated cap trapper, *Genomics*, **137**, 327–336.
2. Maruyama, K. and Sugano, S. 1994, Oligo-Capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides, *Gene*, **138**, 171–174.
3. England, T. E., Bruce, A. G., and Uhlenbeck, O. C. 1980, Specific labeling of 3′ termini of RNA with T4 RNA ligase, *Methods Enzymol.*, **65**, 65–74.
4. England, T. E. and Uhlenbeck, O. 1978, Enzymatic oligoribonucleotide synthesis with T4 RNA ligase, *Biochemistry*, **17**, 2069–2076.
5. Kato, S., Sekine, S., and Oh, S.-W. et al. 1994, Construction of a human full-length cDNA bank, *Gene*, **150**, 243–250.
6. Ederly, I., Chu, L. L., Sonenberg, N., and Pelletier, J. 1995, An efficient strategy to isolate full-length cDNAs based on an mRNA cap retention procedure (CAPture), *Mol. Cell. Biol.*, **15**, 3363–3371.
7. Meador, J. III, Cannon, B., Cannistraro, V. J., and Kennel, D. 1990, Purification and characterization of *Escherichia coli* RNase I. Comparisons with RNase M, *Eur. J. Biochem.*, **187**, 549–553.
8. Chomczynski, P. and Sacchi, N. 1987, Single-step method

of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction, *Anal. Biochem.*, **162**, 156–159.

9. Del Sal, G., Manfioletti, G., and Schneider, C. 1989, The CTAB-DNA precipitation method: a common mini-scale preparation of template DNA from phagemids, phages or plasmids suitable for sequencing, *BioTechniques*, **7**, 514–519.
10. Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989, *Molecular Cloning*, CSH Laboratory Press.