

Chapter 3

Construction of cDNA Libraries: Focus on Protists and Fungi

Naiara Rodríguez-Ezpeleta, Shona Teixeira, Lise Forget, Gertraud Burger, and B. Franz Lang

Abstract

Sequencing of cDNA libraries is an efficient and inexpensive approach to analyze the protein-coding portion of a genome. It is frequently used for surveying the genomes of poorly studied eukaryotes, and is particularly useful for species that are not easily amenable to genome sequencing, because they are nonaxenic and/or difficult to cultivate. In this chapter, we describe protocols that have been applied successfully to construct and normalize a variety of cDNA libraries from many different species of free-living protists and fungi, and that require only small quantities of cell material.

Key words: EST, RNA purification, reverse transcriptase, template-switching, normalization, DSN.

1. Introduction

Sequencing of cDNA libraries has been extensively used to determine the expressed portion of protein-coding genes (*Expressed Sequence Tags*; ESTs) in model eukaryotes. It has also gained importance for eukaryotic genome projects, as the precise inference of exon-intron boundaries relies on substantial training sets of EST data from the corresponding organisms. In addition, EST sequencing can now also be applied to poorly growing organisms such as protists (for most part unicellular eukaryotes that are neither animals, fungi, nor plants) and fungi (e.g., (1–8)). In such cases, limited quantities of cell material and small amounts of RNA can be overcome by PCR-based amplification. There are two major advantages of the EST approach compared to genomics. First, bacterial contamination of eukaryotic cultures can be tolerated, as polyA tails of bacterial transcripts are too short to be primed by standard oligo-dT

primers. In fact, it is even feasible to use total RNA instead of purified mRNA for cDNA library construction, a decisive advantage in the case of nonaxenic protists or fungi. The second advantage is that some of these organisms require mechanical methods for cell breakage, because their rigid cell walls resist digestion with commercially available cell wall lysing enzymes. While genomic DNA may become too fragmented through such treatment to be useful for genome sequencing, the significantly smaller mRNAs remain sufficiently intact.

In this chapter, we describe a fast and relatively simple method to construct cDNA libraries from protists and fungi. The protocol presented here describes the RNA extraction and purification, and the cDNA synthesis, amplification, size fractionation, and normalization steps. An overview of the procedure is given in **Fig. 3.1A**. The protocol requires small quantities of cell material, works with both total RNA and purified (polyA)

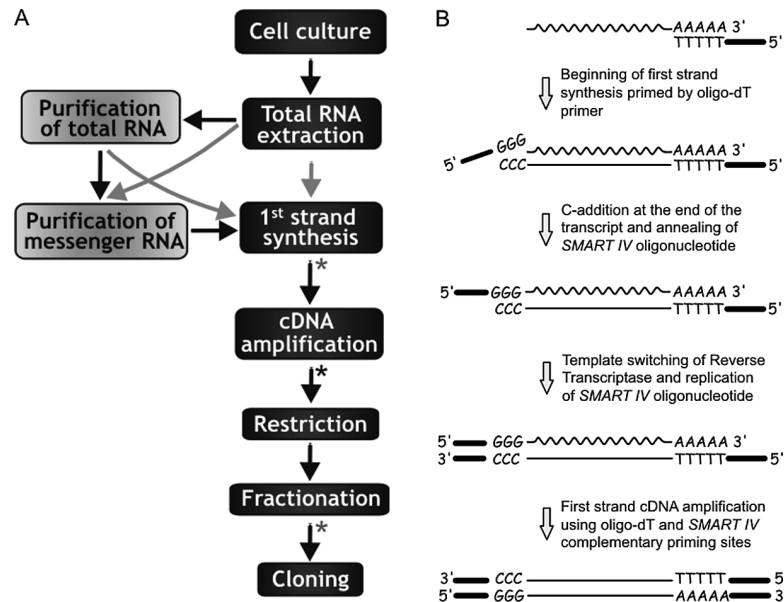


Fig. 3.1. Overview of procedures. (A) cDNA library construction. *Black boxes*, obligatory steps; *grey boxes*, optional steps. *Black arrows* point to protocols described in detail. 'Purification of total RNA' stands for elimination of genomic DNA. A *black asterisk* indicates where a normalization step is usually introduced; a *grey asterisk* indicates where alternative normalization steps can be introduced. **(B)** First strand cDNA synthesis and amplification. First, an oligo-dT containing oligonucleotide is used to prime first strand cDNA synthesis, catalyzed by a RNase H activity-deficient MMLV reverse transcriptase (RT). When the capped 5' end of the mRNA is reached, the RT adds 2 to 5 C residues to the first-strand cDNA, permitting that the SMART IV primer anneals and that DNA synthesis continues until the end of the oligonucleotide. To our experience, noncapped 5' ends undergo the same reaction, albeit at reduced efficiency. Finally, the oligo-dT and SMART IV primers serve for PCR amplification of double-stranded cDNA.

mRNA, and enriches full-length cDNAs. Note, however, that the described protocols involve a PCR amplification step, which is prone to artifacts such as unequal amplification of cDNAs, with a tendency to more efficiently amplify shorter molecules. This potential problem is less relevant in exploratory EST projects. However, when sufficient cell material can be obtained, and when avoiding artifacts is a prime issue, we advise employing cDNA protocols without PCR amplification steps and including prior mRNA purification. In such instances, the readers may follow our procedure from cell culture to mRNA purification, and then continue with one of the protocols described elsewhere in this book.

In our standard procedure, first strand cDNA synthesis and amplification are performed using the CreatorTM SMARTTM cDNA library construction techniques (BD Biosciences, Palo Alto, CA), essentially following the manufacturer's recommendations. The method is based on synthesis of the first cDNA strand with an anchored oligo-dT primer. The terminal C-addition and template switching features of the particular reverse transcriptase allow the second primer to anneal to and extend the RNA template and thus synthesize the first cDNA strand until the end of the primer (**Fig. 3.1B**) (9–11). Subsequently, the product is amplified by PCR using the same primers as before, cut by *Sfi* I (sites are introduced by the primers), and the asymmetrical restriction sites at both cDNA ends are used for directional cloning. A common problem in first-strand cDNA synthesis is premature termination of reverse transcription at mRNA secondary structures. The SMART technique enriches full-length cDNAs because template switching to the 5' primer occurs preferentially when the reverse transcriptase has reached the (usually capped) 5' end of the mRNA. Yet, partial cDNAs will also be produced although less efficiently. In fact, the presence of a certain proportion of 5' end-truncated cDNAs in the library is desirable, as it allows sequences to start at various points of coding regions, thus reducing the requirement for sequencing long cDNA inserts by primer walking.

Since the abundance of the various transcripts in a cell may vary by a factor of a thousand (12), random EST sequencing typically becomes inefficient after a few thousand readings. Therefore, the detection of weakly expressed genes requires normalization of the cDNA libraries, for which we use a simple and efficient procedure involving enzymatic degradation of double stranded DNA or DNA-RNA hybrids (13, 14). We have employed this procedure successfully to construct libraries for protists and fungi. Moreover, except for the culturing and cell lysis steps, these protocols can be applied to animals and plants.

2. Materials

Enzymes, buffers, and reagents such as BSA, DTT, dNTP, and ATP should be stored frozen at -20°C.

2.1. RNA Purification

All solutions must be prepared with RNase-free water, using RNase-free chemicals, glassware and plasticware. Gloves should be worn during manipulation of samples (*see Note 1*).

1. *Trizol*[®] reagent (Invitrogen), or preferentially a home-made substitute (*see Note 2*): 38% stabilized phenol (to stabilize, add 1 mg of hydroxyl quinoline, 2 µL of mercaptoethanol and 0.5 mL of HPLC water per g of phenol, and incubate at 37°C shaking), 0.8 M guanidine thiocyanate, 0.4 M ammonium thiocyanate, 0.1 M sodium acetate, 5% glycerol. Note that *Trizol* contains phenol, which causes heavy skin burns and is toxic on contact or by inhalation of vapors. It should therefore be manipulated under a fume hood, using gloves. According to the manufacturer, commercial *Trizol* is stable at 4°C for at least 9 months (but *see Note 2*).
2. Mixture (1:1) of 150–212 and 425–600 micron-sized glass beads (Sigma); required for species with a tough cell wall (e.g., most fungi, jakobid flagellates, glaucophytes, red and green algae).
3. Chloroform (add a pinch of bicarbonate for stabilization).
4. Isopropanol.
5. ‘Wash ethanol’: Ethanol (75%).
6. ‘Minikit column’, RNeasy[®] Plus MiniKit (QIAGEN).
7. Oligo-dT cellulose (Amersham Biosciences).
8. ‘1X binding buffer’ and ‘2X binding buffer’. Composition of 2X: 20 mM Tris-HCl, pH 7.5, 2 mM EDTA, pH 8, 0.1% SDS, 1 M NaCl. This buffer precipitates at room temperature; heat in a water bath at 65°C before use.
9. ‘Wash buffer’ (1X): 1 mM Tris-HCl, pH 7.5, 1 mM EDTA, pH 8, 0.05% SDS, 0.2 M NaCl. Store at room temperature or heat before use to dissolve precipitated SDS.
10. ‘Elution buffer’ (1X): 1 mM Tris, pH 7.5, 1 mM EDTA, pH 8, 0.05% SDS. Store at room temperature or heat before use to dissolve precipitated SDS.
11. NaCl (4 M).
12. ‘Ethanol-AmAc’: 95% ethanol, 0.5 M ammonium acetate.

2.2. First Strand cDNA Synthesis and PCR Amplification

1. 'SMART IV primer', (10 mM) (Clontech):
5'-AAGCAGTGGTATCAACGCAGACTGGCCATTACG
GCCGGG-3'.
2. 'oligo dT-primer': CDS III/3' (anchored) PCR primer
(10 mM) (Clontech):
5'-ATTCTAGAGGCCGAGGCGGCCGACATG-d(T)₃₀
N₁N-3'
(N₁=A, G, or C; N=A, G, C, or T).
3. 'RT-buffer': First strand buffer (5X) (Clontech): 250 mM
Tris-HCl pH 8.3, 30 mM MgCl₂, 375 mM KCl.
4. 'Reverse transcriptase', *PowerScript*TM (Clontech).
5. DTT (20 mM).
6. dNTP mix (10 mM).
7. RNase I (100 μM).
8. '10X PCR buffer', Advantage[®]-2 (Clontech): 400 mM
Tricine-KOH pH 8.7, 150 mM KOAc, 35 mM Mg(OAc)₂,
37 μg/mL BSA, 0.05% Tween-20, 0.05% Nonidet-P40.
9. '5' PCR primer' (10 mM) (Clontech):
5'-AAGCAGTGGTATCAACGCAGAGT-3'.
10. 'Polymerase mix', Advantage[®]-2 (50X) (Clontech).
11. 'Proteinase K solution', (10 mg/mL) (Boehringer).
12. 'Gel extraction kit', QIAquick (QIAGEN).

2.3. Normalization

1. DSN enzyme (Evrogen JSC) diluted according to the manu-
facturer instructions to 1 Kunitz unit/μL.
2. 'DSN storage buffer', (Evrogen JSC).
3. '5X hybridization buffer': 0.25 M HEPES pH 7.5, 2.5 M NaCl,
1 mM EDTA. This buffer may precipitate; store at room tem-
perature for 20 min or incubate at 37°C for about 10 min before
use.
4. 'DSN buffer' for enzyme reaction; 2X composition: 100 mM
Tris-HCl pH 8.0, 10 mM MgCl₂, 2 mM DTT.
5. EDTA (5 mM).

2.4. Restriction

1. *Sfi* I restriction endonuclease (20 U/μL) (Clontech).
2. '*Sfi* I restriction buffer' (10X) (Clontech).
3. BSA (10 mg/mL).
4. EDTA (0.5 M, pH 8).
5. 'Gel extraction kit', QIAquick (QIAGEN).

2.5. DNA Fractionation

1. 'Low melting agarose', SeaPlaque[®] GTG[®] ultra-pure (Mandel).
2. Formamide (highest quality, Pharmacia; stored under nitrogen or argon).
3. Electroelution chamber (*see* **Note 3**).
4. 'TAE buffer' (1X): 40 mM Tris-acetate, 20 mM sodium acetate, 1 mM EDTA, pH 8.0.

2.6. Cloning

1. 'Ligation buffer' (10X): 200 mM Tris-HCl pH 7.6, 50 mM MgCl₂, 50 mM DTT.
2. 'pDNRLib vector', cut by *Sfi* I and purified (*see* **Note 4**).
3. ATP (10 mM).
4. T4 DNA ligase (5U/μL).
5. Competent cells (DH5α), and LB agar plates containing 10 mg/mL chloramphenicol and 4 μg/mL tetracycline.

3. Methods
3.1. Cell Culture

A large variety of protists require live bacteria as food. To minimize potential RNA degradation by bacterial enzymes, it is important to keep the ratio of eukaryotic versus bacterial cells as high as possible. Bacteria can be partially removed from protist cultures through differential centrifugation, but according to our experience, this is not necessary because cultures in the late logarithmic or stationary phase of growth, which we generally use for RNA extraction, contain only few bacteria. Note that sampling from only one growth condition may restrict the effective number of cDNA sequences in a library. Ideally, cells grown under different conditions should be combined, but this is more difficult if not impossible for bacterivorous protists and for parasites.

3.2. RNA Purification

For total RNA extraction, we use a modified *Trizol* protocol.

3.2.1. Extraction of Total RNA

1. Collect the cells by centrifugation (speed and time vary from one species to another – *see* **Note 5**), or by straining through a fine-mesh nylon coffee filter in case of filamentous fungi.
2. Remove supernatant completely, resuspend cell pellet in *Trizol* and mix; 1 mL of *Trizol* per 5 to 10 × 10⁶ cells is recommended. For filamentous fungi or species with rigid cell walls (e.g., algae), add glass beads (*see* **Note 6**) and shake by hand in a glass bottle (15). Once the cells are broken to >50% (check under light microscope), remove glass beads by repeated rinsing with small volumes of *Trizol*. The cells in *Trizol* may be stored at –80°C for at least one month.

3. Leave the cell/*Trizol* solution for 5 min at room temperature, then add 0.2 mL chloroform per mL of *Trizol*, shake vigorously for 15 sec, and let the mixture settle at room temperature for 2 to 15 min. From here on use 40 mL plastic centrifuge tubes.
4. Centrifuge at 12,000*g* at 4°C for 15 min, and collect the colourless aqueous top phase (about 60% of the total volume). Avoid material from the interface (if this occurs, repeat centrifugation).
5. Add 0.3 mL of isopropanol per mL of the collected aqueous phase to precipitate RNA, mix by inversion and leave 5–10 min at room temperature.
6. Centrifuge at 12,000*g* at 4°C for 15 min and remove the supernatant carefully. The RNA appears as a gel-like or white pellet at the side and bottom of the tube.
7. Wash by adding wash ethanol (1 mL per mL of collected phase) and by inverting the tube a few times. Centrifuge at 12,000*g* at 4°C for 10 min and discard supernatant. Repeat the procedure twice. Remove all traces of ethanol by using a Pasteur pipette, and air-dry briefly. The RNA pellet can be stored at –20°C for at least one year.
8. Dissolve the RNA in RNase-free water. The volume of added water will vary with the quantity of recuperated RNA. It is best to start suspending in a small volume and to continue adding until the RNA is perfectly dissolved (final RNA concentration ~1 mg/mL). From here on use 1.5 mL Eppendorf tubes.
9. Determine the quality and quantity of the RNA by agarose gel electrophoresis, together with an RNA marker of known size and concentration. **Figure 3.2** (lane 1) shows a typical, high-quality RNA extracted by the described method.
10. Remove DNA from RNA preparation by purification on a MiniKit column (*see Note 7*). **Figure 3.2** (lanes 1 and 2) shows the same material before and after this step.

3.2.2. Messenger RNA Purification

This step is performed using oligo-dT cellulose columns. The oligo-dT primers base-pair with the polyA tail of mRNAs, whereas non-polyadenylated RNAs will not bind and flow through the column. The protocol explained here applies to ~1 mg of total RNA; the quantity of reagents should be adapted to the actual RNA quantity. Unless otherwise specified, all centrifugations are performed in a microcentrifuge (for Eppendorf tubes) at 14,000 rpm (maximum speed) for 30 sec at room temperature. A correct pH of the solutions (7.5) is critical for high mRNA yield.

1. Fill 40 mg of oligo-dT cellulose into an Eppendorf tube. Wash the cellulose by mixing it with 600 µL elution buffer, centrifugation, and removal of the supernatant. Repeat this step another two times.

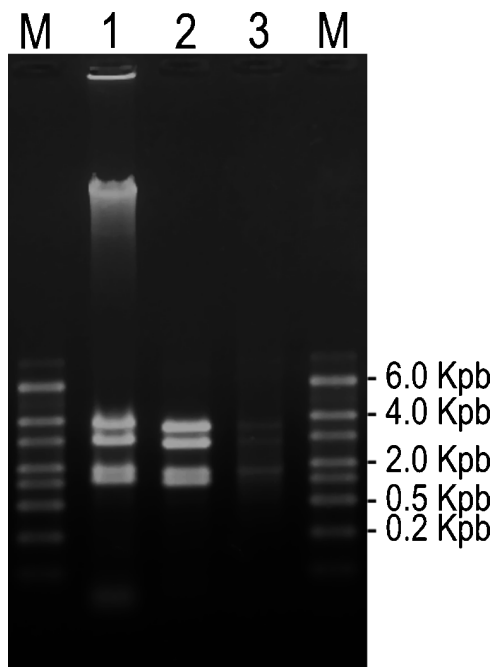


Fig. 3.2. **Agarose gel separation of various stages of a typical RNA purification experiment.** 1.5 μ L of total RNA were loaded before (*lane 1*) and after DNA removal (*lane 2*; **Step 3.2.1.10** in protocol). PolyA mRNA was then purified from 400 μ L of total RNA and recuperated in 15 μ L. 3 μ L of this mRNA fraction were loaded on the gel (*lane 3*; **Step 3.2.2** in protocol). Note a carry-over of rRNAs into the mRNA fraction, which is however negligible as the amount of loaded mRNA corresponds to \sim 50 times more than that of the total RNA. Lane M; RNA ladder, High Range (Fermentas).

2. Equilibrate the cellulose by mixing it with 600 μ L of 1X binding buffer, centrifugation, and removal of the supernatant. Repeat another two times.
3. Adjust the RNA solution to 600 μ L with RNase-free water, and heat it at 65°C for 4 min.
4. Add 600 μ L of preheated (65°C) 2X binding buffer and incubate at room temperature for 15 min; constantly invert the tube.
5. Centrifuge briefly and discard supernatant.
6. Wash cellulose twice with 1X binding buffer and twice with wash buffer.
7. Add 250 μ L elution buffer to cellulose, mix gently, and incubate at 37°C for 5 min.
8. Centrifuge for 1 min and transfer the supernatant to a new tube. Add another 250 μ L of elution buffer to cellulose pellet, mix gently, and incubate for 5 min at 37°C.
9. Centrifuge for 1 min and combine the supernatants of **Step 8** and **9**.

10. Using the recovered 500 μL of RNA, repeat the purification cycle at **Step 3**.
11. After two cycles of mRNA purification (**Steps 3–10**), add 40 μL of NaCl (4 M) and 1 mL of Ethanol-AmAc to the RNA and let precipitate at -20°C for 1 h, or over night.
12. Spin down at 14,000 rpm at 4°C for 20 to 30 min, and discard the supernatant.
13. Add 150 μL wash ethanol and centrifuge at 14,000 rpm at 4°C for 10 min.
14. Discard supernatant. Carefully remove ethanol, air dry pellet, and resuspend it in RNase-free water. **Figure 3.2** (lane 3) shows the result of a typical mRNA purification.

3.3. First Strand cDNA Synthesis

1. Mix 1–3 μL of RNA solution (~ 25 –500 ng polyA RNA, or 100–1,000 ng total RNA), 1 μL *SMART IV* primer, 1 μL oligo dT primer, and adjust volume to 5 μL with HPLC water. Mix well by pipetting up and down.
2. Incubate in a heat block (or PCR machine) at 72°C for 2 min; immediately chill on ice for 2 min; spin down briefly to collect droplets.
3. Add 2 μL 5X RT buffer and 1 μL reverse transcriptase; mix and spin briefly.
4. Incubate in a heat block (or a PCR machine) at 42°C for 1 h.
5. After 1 h, heat again at 72°C for 2 min, chill on ice for 2 min, spin briefly, add 1 μL reverse transcriptase, and incubate in a heat block at 42°C for 1 h (*see Note 8*).
6. Chill on ice. At this point, the first strand cDNA synthesis step is completed.

3.4. cDNA Amplification

1. Remove RNA remaining from the first strand cDNA synthesis by adding 0.1 μL of RNase I 100 μM ; leave at room temperature for 10 min (*see Note 9*).
2. Combine 2 μL first strand cDNA, 80 μL HPLC grade water, 10 μL 10X PCR buffer, 2 μL 5' PCR primer, 2 μL oligo-dT primer, and 2 μL Polymerase Mix. Mix well and PCR-amplify under the following conditions: 20 sec at 95°C followed by ~ 20 cycles: 10 sec at 95°C and 6 min at 68°C . The number of cycles depends on the amount of RNA starting material (*see Note 10*).
3. To inactivate the polymerase in the PCR reaction, add 1 μL of Proteinase K solution. Incubate at 45°C for 20 min, followed by 10 min at 65°C .
4. Purify the amplified cDNA with the Gel Extraction kit, and elute the double-stranded cDNA in a final volume of 30 μL (*see Notes 11 and 12*).

3.5. cDNA Normalization

If normalization is not desired, go to **Step 3.6**. Successful normalization requires an optimized concentration of the DSN enzyme for a given cDNA library. It is recommended to use different DSN concentrations and test the efficiency of normalization.

1. Prepare three or more dilutions of the original enzyme (1/4, 1/8, 1/16 ...) using DSN storage buffer.
2. For each sample to be normalized, mix 4 μ L of amplified cDNA (~500 ng) with 1 μ L of 5X hybridization buffer; heat at 98°C for 3 min and at 70°C for 4 h.
3. While keeping the samples at 70°C, add 4 μ L of preheated (70°C) 2X DSN buffer and 1 μ L of DSN enzyme (for the dilutions mentioned above, this makes 0.25, 0.125, and 0.0625 Kunitz units); incubate at 70°C for 20 min.
4. Inactivate DSN enzyme by adding 10 μ L of 5 mM EDTA.
5. Reamplify the DSN-digested cDNA (as in **Section 3.4, Steps 2 to 4**), but elute in 50 μ L final volume.
6. Verify the success of the normalization of each of your samples by gel electrophoresis, and choose that with the desired, even-size distribution for the following step (*see Note 13*).

3.6. Restriction

1. Mix 45 μ L of amplified and purified cDNA, 40 μ L HPLC-grade water, 10 μ L *Sfi* I 10 X restriction buffer, 1 μ L BSA solution, 4 μ L (= 80 U) *Sfi* I restriction enzyme, and incubate for 2 h at 50°C.
2. Stop reaction by addition of 1 μ L 0.5 M EDTA.
3. Purify DNA with the Gel Extraction kit by eluting with a final volume of 30 μ L (*see Note 12*).

3.7. cDNA Sizing

1. Depending on the capacity of the wells of the electrophoresis system used, the volume of the sample may have to be reduced from 30 μ L to a smaller volume by evaporation in a speed-vac (*see Note 12*).
2. Add formamide to a final concentration of 10% and loading buffer to the cDNA, incubate for 10 min at 50°C, then chill on ice. This step will reduce aggregation of DNA.
3. Load the sample on a low-melting agarose gel (1.2%; TAE buffer) together with a size marker; start by migrating slowly (1.5 V/cm) for a few minutes, then increase to ~3 V/cm. Migrating at higher voltages may overheat and deform the gel matrix. Excise agarose blocks containing DNA fragments of desired size (e.g., 0.5 to 1 kbp; 1 to 5 kbp).
4. Electroelute each cDNA fraction (*see Note 14*), and check their yield and size distribution by agarose gel electrophoresis loading about 1/10 of the recuperated material (*see Note 15*).

3.8. Cloning

Small inserts clone more efficiently. Therefore, to avoid a bias toward small inserts in cDNA libraries, cDNAs should be size-fractionated and fractions cloned separately.

3.8.1. Ligation

The molecule-ratio of insert to vector should be $\sim 3:1$ (the vector size is 4 kbp). When calculating the size of the inserts, one needs to consider that the fragment sizes are not necessarily uniformly distributed. For example, in the 1 to 5 kbp fraction, there can be overrepresentation of fragments from 1 to 1.5 kbp. The size distribution should be assessed based on agarose gel migration and the mean size of the fragments estimated accordingly. The final DNA concentration of insert and vector DNA combined should be 5 ng/ μL , and the total volume of the reaction should be 3 to 4 μL . See **Table 3.1** for two examples of ligation recipes.

1. For each fraction to be cloned, mix well adequate quantities of insert and vector DNA.
2. Heat mix at 50°C for 10 min, place on ice for a few minutes, and let stand at room temperature for several minutes.
3. Add ligation buffer, ATP (final concentration 0.5–1 mM), HPLC grade water, and 0.5 U of T4 DNA ligase.
4. Place the ligation mix in a 14°C incubator overnight.

3.8.2. Transformation

The above-described ligation mix can be used for 10 transformations (200 μL of competent cells per transformation). Competent cells are prepared and transformation is conducted by standard procedures. Transformed cells should be plated onto chloramphenicol-containing agar plates (about 10 plates per transformation).

Table 3.1
Examples for ligation reactions

	0.5 to 1 kbp (mean: 750 bp)	1 to 5 kbp (mean: 1.5 kbp)
Insert ¹	0.5 μL (10 ng/ μL)	0.5 μL (20 ng/ μL)
Vector (10 ng/ μL)	1 μL	1 μL
Ligation Buffer (10X)	0.3 μL	0.4 μL
ATP (10 mM)	0.27 μL	0.36 μL
Water	0.83 μL	1.64 μL
T4 DNA ligase (5 U/ μL)	0.1 μL	0.1 μL
FINAL VOLUME	3 μL	4 μL

¹Numbers between parentheses correspond to the initial concentration of insert.

On average, 1,000 to 2,000 colonies are expected for each transformation with ~ 20 ng ligation mix (when using high-quality competent cells).

4. Notes



1. A major difficulty in handling RNA is the prevention of degradation by contaminant RNases. Autoclaving glassware, tips, tubes, and solutions is often insufficient to inactivate RNases. For additional measures, glassware may be baked at 180°C overnight, and plasticware, tubes, and solutions be treated with diethylpyrocarbonate (DEPC). DEPC reacts with histidine residues of proteins and thus inactivates RNases. Add DEPC to solutions (water, buffers) at a final concentration of 0.05–0.1%, incubate for several hours, and autoclave at least 45 min (the characteristic DEPC scent should disappear). Note that DEPC also reacts with RNA; therefore, it has to be completely removed from all materials before use. Moreover, DEPC can react with chemicals containing primary amine groups, such as Tris. Therefore, these chemicals should be added to the solution only once DEPC is removed. DEPC is a suspected carcinogen; take appropriate precautions when handling it (e.g., always wear gloves and handle it under a fume hood). Water purified by a well-maintained MilliQ system is virtually RNase-free, without further treatment. To verify if MilliQ water is indeed RNase-free, dissolve high-quality RNA in this water, incubate it at 37°C for several hours, and compare the RNA before and after incubation by gel electrophoresis.
2. Highest quality RNA (and high molecular weight DNA) is regularly obtained with home-made *Trizol*, but not with commercial *Trizol* sources. An apparent reason is that the recipe for home-made *Trizol* calls for phenol of highest purity (supplied in light-protected glass bottles under a protective gas), and that it is protected from oxidation by additives. Advanced phenol oxidation, which is recognizable by a reddish-pink color, causes slight RNA but severe DNA degradation. Stabilized phenol or *Trizol* prepared by our recipe may be stored frozen at -20°C for many years. Once in use, we recommend to keep it for <1 month at 4°C , in light-protected bottles (brown glass or unstained glass wrapped with aluminum foil).
3. There are numerous techniques for electroeluting DNA from agarose, and diverse devices are commercially available that will not be described here. Relevant information specific to each

technique is easily available. The least complicated procedure is electroelution in a closed dialysis tube. Electroelution chambers are available from Schleicher and Schüll (Elutrap), Millipore (Centrilutor), EMD BioSciences (D-tube electroelution), and RPI Research Products (GeneCapsule), to mention only some of the more popular devices.

4. We have noted on several occasions that the 220-bp stuffer fragment is not removed from the commercially distributed, 'ready-to-use' pDNRLib cloning vector. For highest cloning efficiency, the *Sfi*I-digested vector should be purified by electrophoresis on low-melting agarose, followed by electroelution of the 4.2 kbp fragment.
5. The centrifugation conditions for pelleting cells depend on multiple factors that are specific to each culture (density of the medium, type of cells, etc). Small cell pellets, in particular those of small and/or flagellated eukaryotes, tend to dissolve quickly and have to be decanted immediately after centrifugation, under close visual control.
6. A number of protists and fungi contain a rigid cell wall that is not (or only for a small fraction of cells) dissolved by *Trizol*. In such cases, cells have to be broken mechanically. We recommend cell disintegration in the presence of *Trizol*, as RNA will otherwise be degraded by intra-cellular (and if present, bacterial) RNases. Cells of filamentous fungi are broken by grinding together with sand or glass beads in a mortar; cells of unicellular organisms may be disintegrated by manual shaking together with glass beads in a glass bottle (e.g., (15)), or by other suitable disruption methods. Because the volume increases by the addition of glass beads, more *Trizol* has to be used in this case (we use ~ 10 mL of *Trizol* and 10 mL of glass beads for 1 g of cells). Once >50% of cells are broken (check by microscopy), decant the glass beads and collect the supernatant. Repeatedly (2 to 4 times) rinse the glass beads with small volumes of *Trizol* to collect a maximum of the cell lysate.
7. Total RNA extractions contain variable amounts of genomic DNA (depending on the organism and the extraction conditions), which should be eliminated to avoid undesirable PCR products.
8. The reverse transcriptase reaction is repeated once to increase cDNA length. By heating to 72°C after the first reaction cycle, secondary structures in mRNA are destabilized and elongation of the first strand may proceed in the subsequent cycle.
9. RNA should be digested after first strand synthesis to permit optimal synthesis of a second DNA strand, and to avoid interference in the following PCR amplification step.

10. The amount of RNA starting material versus the number of PCR cycles recommended by manufacturer is as follows (total RNA/mRNA/number of cycles): 1.0–2.0 µg/0.5–1.0 µg/18–20; 0.5–1.0 µg/0.25–0.5 µg/20–22; 0.25–0.5 µg/0.125–0.25 µg/22–24; 0.05–0.25 µg/0.025–0.125 µg/24–26. We recommend minimizing the number of amplification cycles to avoid PCR artifacts.
11. A proteinase K digestion prior to the gel extraction is recommended. Other PCR purification methods that efficiently remove dNTPs, salts, and long primers may be used as well (note that the longest primer used here is 59 nt long).
12. In order to maximize DNA recuperation, the elution volume may be increased to 100 µL, and subsequently reduced in a speed-vac.
13. The nonnormalized and the normalized samples generated with different DSN concentrations should be compared by agarose gel electrophoresis. For best results, the discrete bands of highly expressed mRNAs should have disappeared, and fragment sizes should be evenly distributed and not be smaller than in the nonnormalized sample.
14. For the extraction of the sized DNA fragments from the agarose gel, we discourage the use of gel extraction kits, because cloning efficiencies may be reduced by 1 to 3 orders of magnitude, compared to electroelution.
15. Size fractionation of cDNA is sometimes difficult because remaining contaminants (e.g., polysaccharides) are carried over despite purification. In such cases, DNA fragments tend to aggregate, causing contamination of the larger-size cDNA with small fragments in electrophoresis. Separation of smaller cDNA quantities (to avoid overloading of the gel) will often help to reduce, although not eliminate, the problem.

Acknowledgments

We wish to thank Jean-François Bouffard, Jung Hwa Seo, Zhang Wang, and Yun Zhu for excellent technical assistance. This research has been funded in part by grants from Génome Québec/Canada (GB; BFL) and CIHR Canada (BFL). NRE has been supported by the ‘Programa de Formación de Investigadores del Departamento de Educación, Universidades e Investigación’ (Government of Basque Country). Salary and interaction support from the Canadian Institute for Advanced Research, Program in Evolutionary Biology (BFL; GB), is gratefully acknowledged.

References

1. Crepineau, F., Roscoe, T., Kaas, R., Kloareg, B., and Boyen, C. (2000) Characterisation of complementary DNAs from the expressed sequence tag analysis of life cycle stages of *Laminaria digitata* (Phaeophyceae). *Plant Mol Biol.* **43**, 503–13.
2. Howe, D. K. (2001) Initiation of a *Sarcocystis neurona* expressed sequence tag (EST) sequencing project: a preliminary report. *Vet Parasitol.* **95**, 233–9.
3. Nikaido, I., Asamizu, E., Nakajima, M., Nakamura, Y., Saga, N., and Tabata, S. (2000) Generation of 10,154 expressed sequence tags from a leafy gametophyte of a marine red alga, *Porphyra yezoensis*. *DNA Res.* **7**, 223–7.
4. Qutob, D., Hraber, P. T., Sobral, B. W., and Gijzen, M. (2000) Comparative analysis of expressed sequences in *Phytophthora sojae*. *Plant Physiol.* **123**, 243–54.
5. Broecker, K., Bernard, F., Moerschbacher, B. M., Brown, D. W., Cheung, F., Proctor, R. H., et al. (2006) An EST library from *Puccinia graminis* f. sp. tritici reveals genes potentially involved in fungal differentiation. *FEMS Microbiol Lett.* **256**, 273–81.
6. Brown, D. W., Cheung, F., Proctor, R. H., Butchko, R. A., Zheng, L., Lee, Y., et al. (2005) Comparative analysis of 87,000 expressed sequence tags from the fumonisin-producing fungus *Fusarium verticillioides*. *Fungal Genet Biol.* **42**, 848–61.
7. Felipe, M. S., Andrade, R. V., Petrofeza, S. S., Maranhao, A. Q., Torres, F. A., Albuquerque, P., et al. (2003) Transcriptome characterization of the dimorphic and pathogenic fungus *Paracoccidioides brasiliensis* by EST analysis. *Yeast* **20**, 263–71.
8. Keon, J., Antoniw, J., Rudd, J., Skinner, W., Hargreaves, J., Hammond-Kosack, K., et al. (2005) Analysis of expressed sequence tags from the wheat leaf blotch pathogen *Mycosphaerella graminicola* (anamorph *Septoria tritici*). *Fungal Genet Biol.* **42**, 376–89.
9. Schramm, G., Bruchhaus, I., and Roeder, T. (2000) A simple and reliable 5'-RACE approach. *Nucleic Acids Res.* **28**, E96.
10. Schmidt, W. M., and Mueller, M. W. (1999) CapSelect: a highly sensitive method for 5' CAP-dependent enrichment of full-length cDNA in PCR-mediated analysis of mRNAs. *Nucleic Acids Res.* **27**, e31.
11. Zhu, Y. Y., Machleder, E. M., Chenchik, A., Li, R., and Siebert, P. D. (2001) Reverse transcriptase template switching: a SMART approach for full-length cDNA library construction. *Biotechniques* **30**, 892–7.
12. Galau, G. A., Klein, W. H., Britten, R. J., and Davidson, E. H. (1977) Significance of rare m RNA sequences in liver. *Arch Biochem Biophys.* **179**, 584–99.
13. Shagin, D. A., Rebrikov, D. V., Kozhemyako, V. B., Altshuler, I. M., Shcheglov, A. S., Zhulidov, P. A., et al. (2002) A novel method for SNP detection using a new duplex-specific nuclease from crab hepatopancreas. *Genome Res.* **12**, 1935–42.
14. Zhulidov, P. A., Bogdanova, E. A., Shcheglov, A. S., Vagner, L. L., Khaspekov, G. L., Kozhemyako, V. B., et al. (2004) Simple cDNA normalization using kamchatka crab duplex-specific nuclease. *Nucleic Acids Res.* **32**, e37.
15. Lang, B., Burger, G., Doxiadis, I., Thomas, D. Y., Bandlow, W., and Kaudewitz, F. (1977) A simple method for the large-scale preparation of mitochondria from microorganisms. *Anal Biochem.* **77**, 110–21.

Expressed Sequence Tags (ESTs)

Generation and Analysis

Parkinson, J. (Ed.)

2009, X, 332 p. 29 illus., Hardcover

ISBN: 978-1-58829-759-4

A product of Humana Press