Marques, M.; Perez Amador, MA. (2012). Construction and analysis of full-length and normalized cDNA libraries from citrus. Functional Genomics: Methods and Protocols. 815:51-65. doi:10.1007/978-1-61779-424-7_5.

Chapter 5

Construction and Analysis of Full-Length and Normalized cDNA Libraries from Citrus

M. Carmen Marques and Miguel A. Perez-Amador

Abstract

We have developed an integrated method to generate a normalized cDNA collection enriched in full-length and rare transcripts from citrus, using different species and multiple tissues and developmental stages. Interpretation of ever-increasing raw sequence information generated by modern genome sequencing technologies faces multiple challenges, such as gene function analysis and genome annotation. In this regard, the availability of full-length cDNA clones facilitates functional analysis of the corresponding genes enabling manipulation of their expression and the generation of a variety of tagged versions of the native protein. The development of full-length cDNA sequences has the power to improve the quality of genome annotation, as well as provide tools for functional characterization of genes.

Key words: Library, cDNA, Citrus, Full-length, Normalized, SMART, DSN nuclease, Gateway technology

1. Introduction

Many methods for the construction of cDNA libraries have been developed in recent years. Conventional cDNA library construction approaches, however, suffer from several major shortcomings. First, the majority of cDNA clones are not full-length, mainly due to premature termination of reverse transcription or blunt-end polishing of cDNA ends prior to subcloning. A number of methods have been developed to overcome this problem and obtain cDNA library preparations enriched in full-length sequences, most of them based on the use of the mRNA cap structure (1–4). However, these methods require high quantities of starting material and complicated multistep manipulations of mRNA and cDNA intermediates, which often results in the degradation of mRNA and the isolation of short clones. The recently described SMART method (switching

Michael Kaufmann and Claudia Klinger (eds.), *Functional Genomics: Methods and Protocols*,
Methods in Molecular Biology, vol. 815, DOI 10.1007/978-1-61779-424-7_5, © Springer Science+Business Media, LLC 2012

mechanism at the 5′ end of the RNA transcript), exploits two intrinsic properties of Moloney murine leukemia virus (MMLV) reverse transcriptase, reverse transcription and template switching of blunt-ended cDNA copies, allowing an easy and efficient production of full-length clones (5). Second, the straightforward random sequencing of clones from standard cDNA libraries is inefficient for discovering rare transcripts, owing to the repeated occurrence of intermediately and highly abundant cDNAs, and a normalization process is often required. This process generally utilizes second-order reaction kinetics of re-association of denatured DNA, so that relative transcript concentration within the remaining single-stranded cDNA fraction is equalized to a considerable extent. A recently described method uses the properties of DSN nuclease to specifically cleave ds-DNA (in both DNA–DNA and DNA–RNA duplexes) allowing the separation of the normalized ss-fraction (6–8). Third, an adaptor-mediated cloning process is still a common approach for cDNA library construction, leading to undesirable ligation by-products and inserts of non-mRNA origin. Directional cloning using *Sfi*I endonuclease minimizes these problems, as it identifies variable target sequences and allows for designing adaptors with noncomplementary ends, thus avoiding their concatenation. As the *Sfi*I recognition sequence is very rare in eukaryotic genomes, the use of *Sfi*I also eliminates the need for methylation during cDNA synthesis (9). In the last place, gene discovery is facilitated by the ability to easily express proteins in both homologous and heterologous biological contexts and thus understanding gene function (10). This entails engineering of multiple expression constructs, which is time-consuming and laborious when using traditional ligase-mediated cloning methods. The recombinational cloning employed in the commercially termed Gateway technology (Invitrogen) exploits the accurate and site-specific recombination system utilized by bacteriophage lambda in order to shuttle sequences between plasmids bearing compatible recombination sites (11–13). This bypasses the need for traditional ligase-mediated cloning while maintaining orientation of the transferred DNA segment and yielding a high proportion of desired clones.

Herein, we describe how we took advantage of the SMART protocol, the DSN nuclease and the Gateway technology to maximize acquisition of full-length and rarely expressed cDNAs from citrus ready to use for functional analysis purposes (14).

2. Materials

2.1. Development of the Gateway-Based Cloning Vector

1. pENTR1A vector (Invitrogen).
2. Restriction enzymes: *Eco*RI, *Xho*I, *Sfi*I.

Table 1
Oligonucleotides used in this protocol

Name	Sequence	Step (section)
pENTR-SfiI-F	AATTCGGCCATTATGGCCTGCAGGATCC GGCCGCCTCGGCC	3.1.2
pENTR-SfiI-R	TCGAGGCCGAGGCGGCCGGATCCTGCA GGCCATAATGGCCG	3.1.2
SMART IV	AAGCAGTGGTATCAACGCAGAGT GGCCATTATGGCCGGG	3.3.1
CDSIII/3	ATTCTAGAGGCCGAGGCGGCC GACATG-d$(T)_{30}$NN	3.3.1
M1-5′	AAGCAGTGGTATCAACGCAGAGT	3.4.3
M1-3′	ATTCTAGAGGCCGAGGCGG	3.4.3
M2-5′	AAGCAGTGGTATCAACGCAG	3.4.4
M2-3′	ATTCTAGAGGCCGAGGCG	3.4.4
pENTR-F	GGCTTTAAAGGAACCAATTCAG	3.5.7
pENTR-R	GCAATGCTTTCTTATAATGCCAAC	3.5.7

*Sfi*I recognition sites (GGCCNNNNNGGCC) are underlined

3. Oligonucleotides: pENTR-SfiI-F, pENTR-SfiI-R, pENTR-F, and pENTR-R (Table 1).
4. JM110 *Escherichia coli* competent cells.
5. Shrimp Alkaline Phosphatase (SAP).
6. Qiaquick Gel Extraction kit (Qiagen).
7. TAE 1×.
8. Agarose.

2.2. Preparation of Poly(A+)-RNA

1. Oligotex mRNA kit (Qiagen). It includes the Oligotex resin, Binding Buffer (OBB), Washing Buffer (OW2), and Elution Buffer (OEB).
2. 3 M sodium acetate pH 5.2.
3. Ethanol 96%.
4. GlycoBlue.

2.3. Synthesis of Full-Length cDNAs for the Construction of a Full-Length Enriched Library

1. BD SMART PCR cDNA synthesis kit (BD Biosciences). This kit contains: 7 µl PowerScript Reverse Transcriptase, 200 µl 5× First-strand buffer, 100 µl 5′ PCR Primer IIA (12 µM), 70 µl dNTP mix, 200 µl DTT (20 mM), 5 µl Control Human Placental Total RNA (10 µg/µl), 1 ml deionized water. It also

includes an Advantage long distance PCR kit, containing: 30 µl 50× Advantage 2 Polymerase mix, 200 µl 10× Advantage 2 PCR buffer, 50 µl 50× dNTP mix, 30 µl Control DNA template, 30 µl Control Primer mix, and 2.5 ml PCR Grade water (see Note 1).

2. Oligonucleotides: SMART IV and CDSIII/3′ (Table 1).
3. Qiaquick PCR purification kit (Qiagen).
4. Qiaquick gel extraction kit (Qiagen).
5. T4 DNA ligase (2 U/µl).
6. *Sfi*I restriction enzyme.
7. Proteinase K (10 µg/µl).
8. One Shot MAX Efficiency DH5α-T1 Competent Cells.
9. Kanamycin, stock solution at 50 mg/µl in sterile water.

2.4. Normalization of cDNAs for the Construction of a Normalized Library

1. DSN nuclease (EVROGEN), including DSN enzyme (initially lyophilized and diluted after reception in 50 µl of DSN Storage Buffer to a final concentration of 1 U/µl); 100 µl of 10× Master Buffer; 500 µl of 2× DSN Stop Solution; 20 µl of DSN Control Template (100 ng/µl).
2. Hybridization Buffer: 200 mM HEPES pH 7.5 and 2 M NaCl.
3. Advantage 2 PCR kit (BD Biosciences).
4. Oligonucleotides: M1-5′, M1-3′, M2-5′, and M2-3′ (Table 1).

3. Methods

3.1. Development of the Gateway-Based Cloning Vector

1. Opening the vector.
   Digest 2 µg of purified pENTR1A plasmid (Invitrogen) with 10 U of *Eco*RI and *Xho*I in the appropriate buffer and a final volume of 50 µl by incubating at 37°C for 3 h, then add ten additional units of restriction enzymes and let at 37°C over night (Fig. 1). Dephosphorylate the vector by adding 20 U of SAP to the reaction and incubate at 37°C for 90 min. Then, incubate at 65°C for 45 min to quench the reaction. Run the resultant product in TAE 1× agarose electrophoresis. Purify the band corresponding to the vector with Qiaquick Gel Extraction Kit (see Note 2). Resuspend in water to a concentration of 10 ng/µl.
2. Introduction of the adapters for *Sfi*I recognition sites.
   Heat a 10-µM mix of the synthetic oligonucleotides pENTR-SfiI-F and pENTR-SfiI-R (Table 1) for 10 min at 70°C and let them anneal by slow cooling at room temperature. Digest the

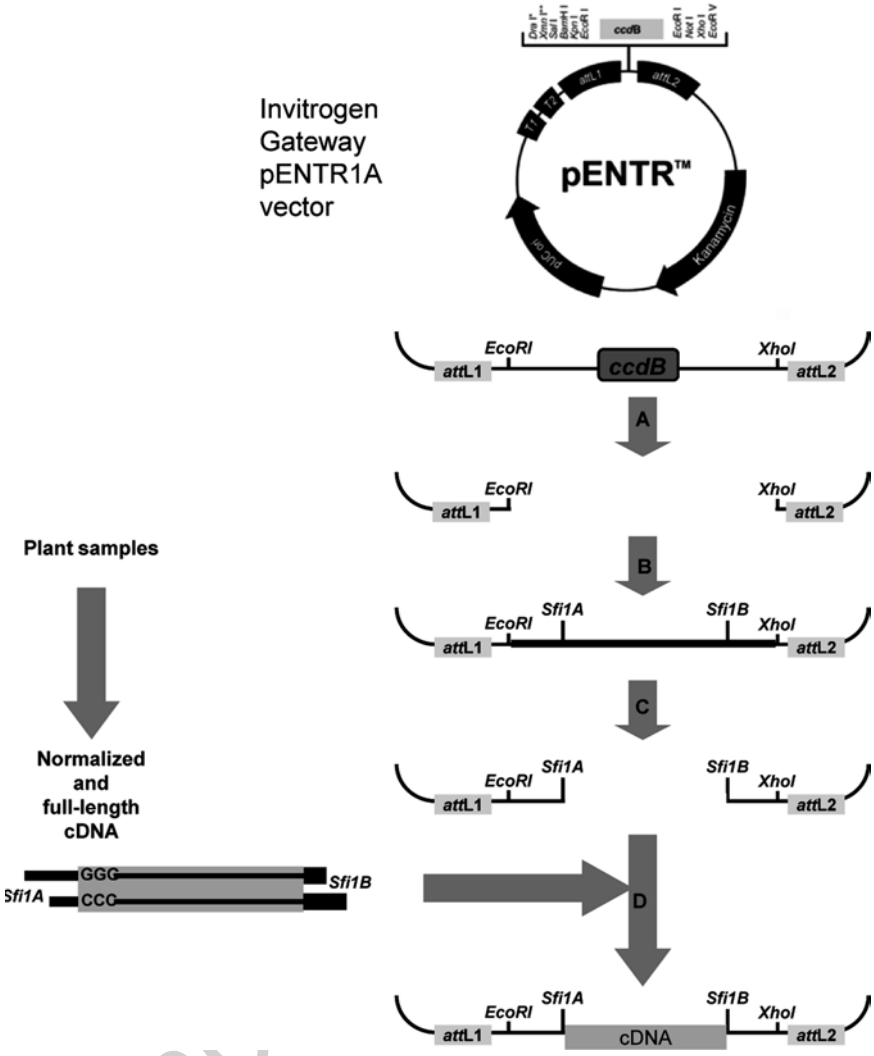

Fig. 1. Generation of the cloning vector pENTR-SfiI. The vector is a modification of the commercial pENTR1A Gateway vector (Invitrogen). Plasmid DNA is digested by *EcoR*I and *Xho*I and dephosphorylated with SAP (**a**). The SfiI adaptor (ds-DNA generated by the annealing of pENTR-SfiI-F and pENTR-SfiI-R oligonucleotides) is then ligated to the vector (**b**). Before ligation with the ds-cDNA, plasmid DNA is digested with *Sfi*I and dephosphorylated (**c**). Ligation with the normalized and full-length enriched cDNA generates the library (**d**).

adapters with 10 U of *EcoR*I and *Xho*I enzymes (see Note 3). Then ligate this double-stranded (ds) oligonucleotide to the opened pENTR1A (10–50 ng) with 10 U of T4 DNA ligase and incubate at 16°C over night. This plasmid constitutes the pENTR-SfiI vector (Fig. 1) (see Note 4).

3. Transformation of the JM110 *E. coli* competent cells. Transform home-made competent JM110 *E. coli* cells with 5 ng of pENTR-SfiI vector, and select transformants by plating onto

50 µg/µl LB-kanamycin plates. Check the construct by digestion/sequencing and make glycerol stocks (see Note 5).

4. Digestion and dephosphorylation of the pENTR1A-SfiI vector. Make a plasmid DNA prep using standard protocol. Add, in a 0.5-ml Eppendorf tube, 5 µg of the pENTR1A-SfiI vector (see step 3.1.3), 4 µl of buffer M 10×, and 2 µl (10 U/µl) of *Sfi*I restriction enzyme, in a final volume of 40 µl. Incubate immediately at 50°C for 4 h, add 2 µl of *Sfi*I and incubate at 50°C for additional 4 h or over night. Dephosphorylate the vector by adding 20 U of SAP to the reaction and incubate at 37°C for 90 min. Then, incubate at 65°C for 45 min to quench the reaction. Run the reaction in TAE 1× agarose electrophoresis, and purify the band corresponding to the vector with Qiaquick Gel Extraction Kit (see Note 2). Quantify and resuspend in water to a final working concentration of 2–5 ng/µl. Check the quality of the preparation by running a ligation of 5 ng of plasmid with T4 DNA ligase and transform (see Note 6).

3.2. Starting Material

1. Add RNase-free water to 200 µg of total RNA to a final volume of 250 µl. Poly(A)$^+$ RNA from different citrus tissues is purified using Oligotex mRNA Midi Kit (Qiagen) (see Note 7).
2. Add 250 µl of OBB and mix gently. Add 20 µl of Oligotex and mix gently. Denature this mix by heating at 70°C for 5 min (during this process, shake the tube every 2 min). Hybridize the samples at room temperature for 10 min.
3. Spin 2 min at maximum speed, add 400 µl of OW2 to the pellet, and resuspend with the pipette. Transfer the total volume to the provided column, spin 1 min at maximum speed, and remove the eluate. Wash again with 400 µl of OW2, and give an extra-spin for 1 min and transfer the column to a new tube.
4. Elute Poly(A)$^+$ RNA from the column by adding 75 µl of hot (70°C) OEB to the column, incubating it at 70°C for 2 min and spinning at maximum speed for 1 min. Transfer the eluate to a new tube, repeat the elution again, and pool both aliquots (150 µl final volume). Add 200 µl of RNase-free water to the elute to bring a final volume of 350 µl and quantify the whole volume in a spectrophotometer.
5. Precipitate Poly(A)$^+$ RNA by adding 35 µl of 3 M sodium acetate pH 5.2, 2 µl of glycoBlue and 900 µl EtOH 96%. Incubate at −80°C over night and recover the pellet by centrifuging for 15 min at maximum speed at 4°C. Wash the pellet with EtOH 70% and dry in SpeedVac. Finally, resuspend Poly(A)$^+$ RNA in RNase-free water to obtain a final concentration of 0.17 µg/µl.

3.3. Obtaining Full-Length ds-cDNA

1. First-strand cDNA synthesis, dC tailing, and template switching by reverse transcription (see Note 7b).

1. Combine the following reagents in a sterile 0.5 ml reaction tube: 3 µl Poly(A)$^+$ sample (0.17 ug/µl), 1 µl CDSIII/3′ oligonucleotide (10 µM), and 1 µl SMART IV oligonucleotide (10 µM) (Table 1).
2. Incubate the mix at 72°C for 2 min, cool the tube down on ice for 2 min, and add the following reagents to the reaction tube: 2 µl 5× first-strand buffer, 1 µl DTT (20 mM), 1 µl 50× dNTP (10 mM), and 1 µl PowerScript Reverse Transcriptase.
3. Incubate the tube at 42°C for 1 h to complete first-strand cDNA amplification (see Note 7c).

2. Second-strand synthesis by long-distance PCR (see Note 7d).
   1. Prepare a PCR mix containing the following components in the order shown: 80 µl deionized water, 10 µl 10× Advantage 2 PCR buffer, and 2 µl 50× dNTP Mix, 4 µl 5′ PCR primer IIA, and 2 µl 50× Advantage 2 polymerase mix. Finally, add 2 µl of the first-strand cDNA from the previous step to obtain a final reaction volume of 100 µl.
   2. Place the tube in the preheated (95°C) thermal cycler and commence thermal cycling using the following parameters: an initial preheating at 95°C for 1 min and additional 16 cycles of 5 s at 95°C, 5 s at 65°C, and 6 min at 68°C (see Note 8).

      Make three second-strand synthesis reactions for every full-length cDNA library you want to obtain. Therefore, a total of 300 µl of ds-cDNA is obtained (see Note 9).

3. ds-cDNA polishing. This step contains three procedures: (1) treatment with proteinase K to denature enzymes used in the previous steps, (2) amplification with T4 DNA polymerase to make ds-cDNA blunt-ended, (3) precipitation and concentration of ds-cDNA.
   1. Make 50 µl aliquots of the ds-cDNA obtained in the previous step in 0.5 ml Eppendorf tubes (six tubes). Add 4 µl of proteinase K (10 µg/µl) to each tube and incubate at 45°C for 1 h in order to eliminate the enzymes used in the previous steps that could interfere with the following reactions. Heat the tubes at 90°C for 10 min to inactivate the proteinase K. Then, chill the tubes in ice water for 2 min, add 3.5 µl (15 U) of T4 DNA polymerase and incubate at 16°C for 30 min. Afterward, heat the tubes at 72°C for 10 min to stop the reaction.
   2. Pool together the content of every two tubes and precipitate ds-cDNA by adding 55 µl ammonium acetate 4 M and 420 µl 95% ethanol to each tube. Mix thoroughly by inverting the tubes. Spin immediately at maximum speed

for 20 min at room temperature. Do not chill the tube before centrifuging as it could result in co-precipitation of impurities. Then, wash pellet with 80% ethanol and air dry to evaporate residual ethanol.

3. Collect the polished ds-cDNA contained in the three tubes by resuspension in water to obtain a single aliquot with a final volume of 66 µl. This cDNA is ready for digestion with *Sfi*I and ligation into the appropriate vector to produce the full-length enriched cDNA library.

3.4. Obtaining Normalized Full-Length ds-cDNA

1. Purification of ds-cDNA.
   Aliquot 100 µl of the ds-cDNA obtained in the step 3.3.2 (see Note 9) into two Eppendorf tubes, containing 50 µl each, and make two reactions of purification using the Qiaquick PCR Purification Kit (Qiagen), according to manufacturer's instructions. Elute cDNA from each column with 50 µl of 1 mM TE. Mix both elutes (100 µl) and concentrate in a SpeedVac to obtain a final concentration of 100 ng/µl approximately (see Note 9b).
2. Normalization step contains three procedures: (1) denaturing, (2) hybridization, and (3) degradation of re-natured ds-cDNA.
   1. Combine in a sterile tube 12 µl of ds-cDNA from the previous step and 4 µl of 4× Hybridization Buffer (see Note 10). Aliquot 4 µl of the reaction mixture into four 0.5 ml Eppendorf tubes.
   2. Denature ds-cDNA by incubating all the tubes at 98°C for 2 min and let rehybridize at 68°C for 5 h. Immediately, add 5 µl of hot Master Buffer 2× (68°C) and incubate the tubes at the same temperature for an additional 10 min (see Note 11).
   3. Dilute 1 µl of the DSN enzyme to 1/2 and 1/4 in DSN Storage Buffer. Add 1 µl (1 U) of DSN enzyme to the first tube, 1 µl of enzyme diluted to 1/2 (0.5 U) to the second, and 1 µl of the enzyme diluted 1/4 (0.25 U) to the third. Add 1 µl of Storage Buffer to the fourth tube to have a control of the normalization. Label each tube appropriately to avoid mistakes (see Note 12).
   4. Incubate at 68°C for 25 min. To quench the reaction add 10 µl of hot Stop Solution 2× to each tube and incubate at 68°C for additional 5 min. Finally, cool the tubes down on ice for several min and add 20 µl of water to each tube to obtain a final volume of 40 µl (see Note 13).
3. First round of amplification of the normalized cDNA and elucidation of the optimal number of cycles.
   1. Each DSN-treated cDNA from the previous step will be amplified separately.

Combine the following reagents in a tube to prepare a PCR Master Mix: 156 µl of water, 20 µl of 10× Advantage PCR buffer, 4 µl of 50× dNTP mix, 6 µl of primer M1-5′ (10 µM), 6 µl of primer M1-3′ (10 µM), and 4 µl of 50× Advantage Polymerase mix.

2. Aliquot 49 µl of this PCR master mix into four sterile 0.5 ml tubes and label them as in the previous step. Add 1 µl of the DSN-treated cDNA from the previous step to their corresponding reaction tube. Place the tubes in a preheated (95°C) thermal cycler and start thermal cycling using the following parameters: 7 s at 95°C, 10 s at 66°C, and 6 min at 72°C (see Note 14).
3. To establish the optimal number of cycles, transfer an aliquot (10 µl) of each PCR reaction to a clean tube after 7, 9, 11, 13 and 15 PCR cycles, obtaining a series of five tubes from every initial PCR reaction (20 tubes) (see Note 15).
4. Electrophorese 5 µl aliquots from each tube in a TAE 1.5× agarose gel to determine the efficiency of normalization (Fig. 2a) (see Note 16).

4. Second round of amplification of the normalized cDNA. Make a tenfold dilution of the reaction that best fit in the normalization parameters. If you plan to estimate the normalization efficiency you should also amplify control (non-normalized) cDNA in parallel. Prepare a PCR master mix by combining the following reagents in the order shown: 76 µl of sterile water, 10 µl of 10× Advantage PCR buffer, 2 µl of 50× dNTP mix, 4 µl of primer M2-5′ (10 µM), 4 µl of primer M2-3′ (10 µM), and 2 µl of 50× Advantage Polymerase Mix. Finally, add 2 µl of the tenfold dilution of the normalized cDNA (see Note 17).
5. cDNA polishing. Proceed as described in step 3.3.3.

3.5. Cloning cDNA into the Plasmid Vector

1. Digestion of ds-cDNA. Combine, in two independent reactions, the following reagents: 33 µl of ds-polished-cDNA (see Note 18), 4 µl of buffer M 10×, and 2 µl (10 U/µl) of *Sfi*I restriction enzyme. Incubate immediately at 50°C for 1 h, add 1 µl of *Sfi*I, and incubate at 50°C for additional 3 h.
2. Purification of the digested cDNA. Purify each digestion using the Qiaquick PCR purification kit (Qiagen), by eluting with 50 µl of Tris–HCl 1 mM. Combine both eluates and concentrate to a final volume of 30 µl.
3. Electrophoresis of the digested cDNA. Electrophorese the 30 µl of digested cDNA obtained in the previous step in TAE 1× agarose gel. Purify cDNA of the appropriate size by cutting the gel in blocks containing cDNAs longer than 1,000 bp and shorter than 5,000 bp (see Note 19).

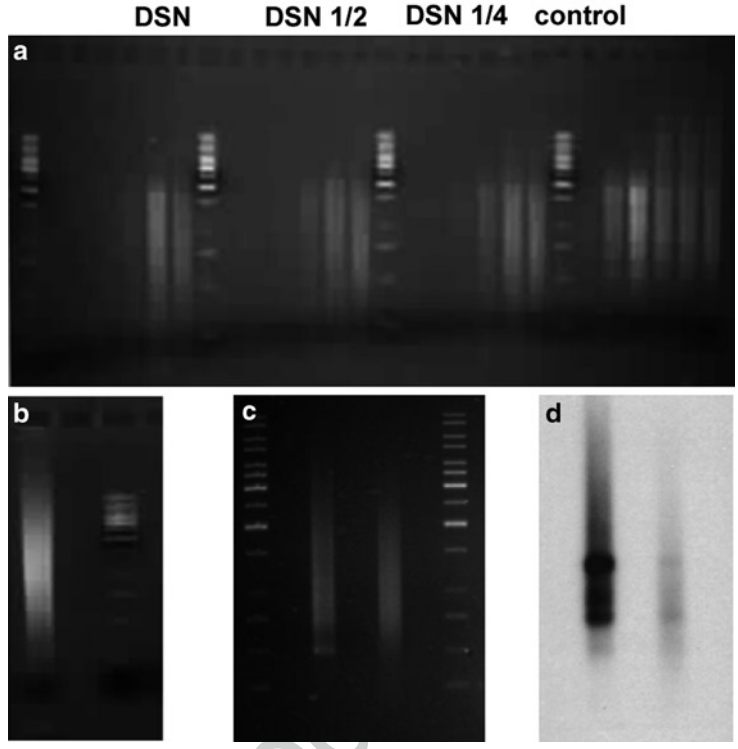

Fig. 2. Evaluation of the efficiency of normalization of cDNA libraries. (**a**) Gel electrophoretic analysis of 5 µl aliquots from the first amplification of the normalized cDNA taken at 7, 9, 11, 13, 15, and 17 PCR cycles (step 3.4.3). (**b**) Gel electrophoretic analysis of the normalized cDNA (as in step 3.5.3) utilized in the construction of the normalized cDNA library RVDevelopN (14). (**c**) Gel electrophoretic analysis of a non-normalized cDNA population (*left*) and a normalized cDNA population (*right*) from the same RNA sample. (**d**) Virtual northern of the cDNA smear blotted and hybridized with the highly abundant citrus clone *C32009H03*.

4. Purification of gel blocks. Purify cDNA from the gel blocks using the Qiaquick gel extraction kit (Qiagen) following manufacturer's instructions. Concentrate the eluted cDNA to a final volume of 20 µl (see Note 20).
5. Ligation into pENTR1A-SfiI vector. Combine, in a 0.5-ml Eppendorf tube, 5 ng of the digested pENTR1A-SfiI vector (see step 3.1.4), 6 µl deionized water, 2 µl cDNA from the previous step, 1 µl ligase buffer 10×, and 1 µl T4 DNA ligase. Incubate at 16°C over night. A control ligation without cDNA must be carried out in parallel (see Note 6).
6. Transformation of DH5α *E. coli* competent cells. We used One Shot MAXEfficiency DH5α-T1 competent cells (Invitrogen) and performed the transformation according to manufacturer's instructions, using 5 µl of the ligation. Plate the transformation onto two 50 µg/µl LB-kanamycin plates and grow at 37°C (see Note 21).

7. Check the quality of the library. Carry out colony PCR from 2 to 300 kanamycin-resistant colonies, using plasmid oligonucleotides pENTR-F and pENTR-R (Table 1). Check the size of the PCR fragments by electrophoresis (see Note 22).
8. Selection of recombinant clones. Select colonies, grow in LB-kanamycin media, and purify plasmid DNA using a 96-well plate format method (Eppendorf or Millipore). Sequence the corresponding cDNA inserts using plasmid oligonucleotide pENTR-F (Table 1) to generate an EST collection (see Note 23).

3.6. Virtual Northern

To get a better assessment of the normalization efficiency, carry out a virtual northern to estimate the relative concentration of a highly abundant clone in both the non-normalized and the normalized cDNA populations obtained from the second run of amplification (Fig. 2). Electrophorese, in a TAE 1.5× gel, equivalent quantities of cDNA corresponding to the non-normalized and normalized samples subjected to the second run of amplification. Transfer DNA to a nitrocellulose membrane, and run a standard Southern blot analysis. Obtain a probe of a highly abundant clone by carrying out a PCR of the corresponding cDNA (see Note 23).

4. Notes

1. This kit also provides seven CHROMA-SPIN-1000 columns and seven microfiltration columns (0.45 μm), but they are not used in this procedure. Please, note that the kit supplies oligonucleotides SMART IIA and 3′ SMART CDS primer II A, however, these oligonucleotides contain *Rsa*I cloning site. Since we are using the properties of *Sfi*I site for cloning purposes, we employ SMART IV and CDSIII/3′ oligonucleotides instead.
2. This process removes the ccdB gene which allows for negative selection of expression clones and lets two binding sites for *Sfi*I adapters.
3. The adapters are ready to be inserted in the opened pENTR1A vector. They provide two recognition sites (*SfiI*A and *SfiI*B, underlined in Table 1) which, once cut with *Sfi*I restriction enzyme, generate two nonsymmetrical ends ready for directional subcloning.
4. The developed pENTR1A-SfiI vector allows both, effective directional cloning by taking advantage of the nonsymmetrical cleavage of the *Sfi*I restriction enzyme and the ease of subcloning by the Gateway System.
5. Although the polylinker does not contain dam or dcm methylation-susceptible sequences, we observed that the digestion of

the vector with *Sfi*I was more efficient in plasmid obtained from JM110 *E. coli* cells.

6. It is very important to test the quality of the *Sfi*I-digested pENTR-SfiI vector prior to the ligation with the ds-cDNA, as well as every time you run a ligation with ds-cDNA. For a control ligation (without insert), mix 5 ng (1–2 µl) of *Sfi*I-digested vector in a 0.5-ml Eppendorf tube with 6 µl deionized water, 1 µl ligase buffer 10×, and 1 µl T4 DNA ligase. Incubate at 16°C over night. Transform One Shot MAXEfficiency DH5α-T1 competent cells using 5 µl of the ligation, plate onto two LB-kanamycin (50 µg/µl) plates, and grow at 37°C. Number of colonies is expected to be very low (less that 50 per plate). Larger numbers mean that probability to get nonrecombinant vectors during sequencing is high, which diminished the quality and efficiency of the library. A new vector preparation has to be obtained. According to our results only 1–5% of the colonies lacked an insert when sing cDNA as insert.

7. It is important to warm OEB at 70°C and Oligotex suspension at 37°C before starting the protocol.

7b. In this step the PowerScript Reverse Transcriptase (RT) provided in the BD SMART PCR Synthesis Kit synthesizes first-strand cDNA primed by CDSIII/3′, which contains a 30-mer oligo dT. This RT also promotes dC tailing (addition of three cytosines at the 3′ end of the cDNA when the first-strand reaches mRNA 5′ end). Furthermore, the addition of SMART IV oligonucleotide, which contains three guanines at its 3′ end, allows template switching needed for next steps.

7c. This first-strand cDNA (10 µl) can be stored at –20°C for up to 3 months.

7d. Double-stranded cDNA is generated with PCR catalyzed by a long-distance polymerase mixture which ensures processive second-strand synthesis and amplification while maintaining accurate size representation. This reaction uses the 5′ anchor primer, which is complementary to the *Sfi*A sequence and the CDS primer that contains the *Sfi*B sequence.

8. PCR parameters for long-distance PCR require short denaturing and annealing steps. Times are adjusted to a Perkin Elmer 9400 thermal cycler, if other apparatus are to be used parameters should be adjusted depending on the ramp rate needed to acquire the running temperature.

9. Three reactions are needed to prepare a full-length cDNA library. The normalization procedure requires ds-cDNA from this step as starting material. So, if you are planning to normalize the cDNAs two additional reactions should be performed, with one of them used in the normalization and the other kept intact in order to make future comparisons.

9b. Usually a final volume of 25 μl renders the appropriate concentration.

10. Heat Hybridization Buffer at 37°C for 10 min to dissolve any precipitate.

11. Note that during the process of normalization the temperature must be kept constant and tubes cannot be removed from the incubator for more than 30 s. Prepare and heat the dilutions and/or buffers to be used shortly before they are to be used.

12. For the degradation of the ds-fraction formed during re-association of cDNA using the DSN nuclease assay different enzyme concentrations in each tube, as the appropriate quantity cannot be known a priori.

13. This DSN-treated cDNA can be stored at −20°C for up to 2 weeks.

14. There is a well-known tendency of PCR to amplify shorter fragments more efficiently than longer ones. Thus, the cDNA sample should be somewhat biased toward longer cDNAs to obtain a natural length distribution upon cloning. This can be done by using a process of regulation of average length which combines the use of an enzyme mixture for long and accurate PCR, the design of primers with complementary sequences at their ends that tend to anneal to each other and compete with primer annealing, being this competition more pronounced in short molecules, and lowering the primer concentration to shift the equilibrium toward intramolecular annealing and therefore increase the suppression.

15. Choosing an optimal number of cycles ensures that the ds-cDNA will remain in the exponential phase of amplification. When the yield of PCR products stops increasing with every additional cycle, the reaction has reached its plateau. The optimal number of cycles for the experiment should be one or two cycles less than that needed to reach the plateau. It is better to use fewer cycles than too many.

16. The profile of an efficiently normalized and amplified cDNA is the one that (1) its overall signal intensity of the smear is similar to that shown for the control (not treated with DSN) but does not contain distinguishable bands, (2) the signal intensity of smear has reached its plateau, (3) the upper boundary of the cDNA smear do not exceed 4.5 kb. This amplified normalized cDNA can be stored at −20°C for up to a month (see Fig. 2).

17. To increase the cDNA concentration, perform three reactions of amplification for the normalized cDNA.

18. This cDNA proceeds from step 3.3.3 for full-length cDNA libraries or from step 3.4.4 for normalized libraries.

19. Elimination of fragments shorter than 1 kb allows enrichment of full-length clones and excludes those obtained in conventional libraries. On the other hand, we excluded fragments longer than 5,000 bp as full-length cDNAs do not seem to be larger than 4 kb.
20. The final concentration is approximately 10 ng/µl.
21. The library is completed. Approximately 40,000 kanamycin-resistant colonies (2,000 colonies per transformation, 2 transformations per ligation, 10 ligations per cDNA synthesis) can be obtained per assay.
22. Expected size of the PCR products (corresponding to the cloned cDNAs) range between 500 bp to 2 kb. Although cDNAs between 1 and 5 kb are purified from the agarose gel (see step 3.5.3), the average size of the cDNAs is smaller due to low cloning efficiency of large cDNAs.
23. Plasmid DNA preps, sequencing of ESTs, and Virtual northern are carried out by standard protocols.

Acknowledgments

The authors would like to thank to all participants in the Spanish Citrus Functional Genomic Project, specially to Drs. Javier Forment, Jose Gadea, and Vicente Conejero. This work was funded by grants from the Spanish Government GEN2001-4885-CO5-01 and GEN2001-4885-CO5-02.

References

1. Carninci P, Kvam C, Kitamura A, Ohsumi T, Okazaki Y, Itoh M, Kamiya M, Shibata K, Sasaki N, Izawa M, Muramatsu M, Hayashizaki Y, and Schneider C (1996) High-efficiency full-length cDNA cloning by biotinylated CAP trapper. *Genomics* **37**, 327–336.
2. Carninci P, Shibata Y, Hayatsu N, Sugahara Y, Shibata K, Itoh M, Konno H, Okazaki Y, Muramatsu M, and Hayashizaki Y (2000) Normalization and subtraction of CAP-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. **10**, 1617–1630.
3. Suzuki Y, and Sugano S (2003) Construction of a full-length enriched and a 5'-end enriched cDNA library using the Oligo-capping method. *Methods Mol. Biol.* **221**, 73–91.
4. Clepet C, Le Clainche I, and Caboche M (2004) Improved full-length cDNA production based on RNA tagging by T4 DNA ligase. *Nucleic Acids Res.* **32,** e6.
5. Zhu YY, Machleder EM, Chenchik A, Li R, and Siebert PD (2001) Reverse transcriptase template switching: a SMART approach for full-length cDNA library construction. *Biotechniques* **30,** 892–897.
6. Zhulidov PA, Bogdanova EA, Shcheglov AS, Vagner LL, Khaspekov GL, Kozhemyako VB, Matz MV, Meleshkevitch E, Moroz LL, Lukyanov SA, and Shagin DA (2004) Simple cDNA normalization using Kamchatka crab duplex-specific nuclease. *Nucleic Acids Res.* **32,** e37.
7. Zhulidov PA, Bogdanova EA, Shcheglov AS, Shagina IA, Wagner LL, Khazpekov GL, Kozhemyako VV, Lukyanov SA, and Shagin DA (2005) A method for the preparation of normalized cDNA libraries enriched with full-length sequences. *Russian J. Bioorg. Chem.* **31,** 170–177.

8. Anisimova VE, Rebrikov DV, Zhulidov PA, Staroverov DB, Lukyanov SA, and Shcheglov AS (2006) Renaturation, activation, and practical use of recombinant duplex-specific nuclease from Kamchatka crab. *Biochem.-Moscow* **71**, 513–519.
9. Toru M, Matsui T, Heidaran MA, and Aaronson SA (1989) An efficient directional cloning system to construct cDNA libraries containing full-length inserts at high frequency. *Gene* **83**, 137–146.
10. Castelli V, Aury JM, Jaillon O, Wincker P, Clepet C, Menard M, Cruaud C, Quétier F, Scarpelli C, Schächter V, Temple G, Caboche M, Weissenbach J, and Salanoubat M (2004) Whole genome sequence comparisons and "Full-length" cDNA sequences: a combined approach to evaluate and improve Arabidopsis genome annotation. *Genome Res.* **14**, 406–413.
11. Earley KW, Haag JR, Pontes O, Opper K, Juehne T, Song KM, and Pikaard CS (2006) Gateway-compatible vectors for plant functional genomics and proteomics. *Plant J.* **45**, 616–629.
12. Karimi M, Inzé D, and Depicker A (2002) GATEWAY™ vectors for *Agrobacterium*- mediated plant transformation. *TRENDS Plant Sci.* **7**, 193–195.
13. Hartley JL, Temple GF, and Brasch MA (2000) DNA cloning using *in vitro* site-specific recombination. *Genome Res.* **10**, 1788–1795.
14. Marques M C, Alonso-Cantabrana H, Forment J, Arribas R, Alamar S, Conejero V, and Perez-Amador MA (2009) A new set of ESTs and cDNA clones from full-length and normalized libraries for gene discovery and functional characterization in citrus. *BMC Genomics* **10**, 428.