CHAPTER 5

NORMALIZATION OF cDNA LIBRARIES

ALEX S. SHCHEGLOV*, PAVEL A. ZHULIDOV, EKATERINA A.
BOGDANOVA, DMITRY A. SHAGIN

*Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, 16/10
Miklukho-Maklaya, 117997 Moscow, Russia*
*Corresponding author
Phone: +(7495) 4298020; Fax: +(7495) 3307056; E-mail: jukart@mail.ru

Abstract: In a cellular transcriptome, the number of mRNA copies per gene may differ by several orders. In cDNA libraries, performed from mRNA, these proportions are the same. Normalization methods allow us to equalize numbers of gene's copies in the library. Normalized cDNA libraries are used to discover new genes transcribed at relatively low levels or for functional screenings. Here, we observed different cDNA libraries normalization methods, which were based on hybridization (renaturation) of cDNA or DNA, or RNA. Also we described duplex-specific nuclease (DSN) normalization protocol – simple and effective cDNA libraries normalization method.

Keywords: Genetic regulation, differentially expressed mRNA, mRNA study, poly(A) + RNA, full-length cDNA, incomplete cDNA, cDNA representation, transcript concentration, normalized cDNA libraries, poorly transcribed, clone coverage, saturating hybridization, enzymatic removal, physical separation, biotinylation, biotinylation of initial RNA, frequent cutter, duplex-specific nuclease (DSN), hydroxyapatite column, plasmid vector, phagemid vector, first strand cDNA, second strand cDNA, immobilization.

Abbreviations: cDNA, complementary DNA; dNTP, deoxyribonucleotidetriphosphate; DSN, duplex-specific nuclease; mRNA, messenger RNA; ITR, inverted terminal repeat; PCR, polymerase chain reaction; rRNA, ribosomal RNA; RT, reverse transcription.

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Normalized cDNA Libraries: What are they needed for?

Genetic regulation of various biological processes can be effectively studied by means of examining the involved mRNA pool. Modern methods allow to reveal differentially expressed mRNAs, to examine their spatial or temporal localization in an organism and, finally, to analyze the function of a respective protein. Analysis of a complete set of cellular mRNAs (transcriptome) became very popular in the last decade as an effective alternative for genome sequencing because such a procedure requires less resources and makes it possible to identify

gene structures much more accurately than any kind of bioinformatical predictions based on genome sequence (Guigo et al. 2000).

In eukaryotic cell, mRNA constitutes ~1–5% of total RNA mass and the number of expressing genes varies from several thousands to several tens thousands. In a cellular transcriptome, the number of mRNA copies per gene may differ by several orders. As a rule, 5–10 major housekeeping genes, whose mRNA constitutes totally ~20% of the cellular mRNA mass, 500–2000 genes transcribed at an intermediate level (give 40–60% of the total cellular mRNA mass), and 10,000–20,000 moderately transcribed genes (20–40% of the total cellular mRNA mass) are expressed in an eukaryotic cell (Alberts et al. 1994).

Methods of mRNA study operate mostly not with an mRNA molecule itself, but with its complementary DNA (cDNA). Cloned pool of cDNAs corresponding to mRNA pool of a biological object under study refers to as a cDNA library. Today, several methods providing to generate fitting libraries from limited starting amounts (0.1–1 μg) of total RNA have been developed and are widely in use. The possibility to use total RNA as a template for cDNA synthesis without a stage of labor-intensive and relatively expensive poly(A) + RNA isolation makes these methods advantageous.

cDNA libraries should meet several requirements. First, such libraries must be representative, i.e. they should contain DNA copies for nearly all transcripts that are typical of the biological object under study. In the case of polymerase chain reaction (PCR)-amplified cDNA libraries, this requirement limits the number of the respective PCR cycles. Typically, the maximum of 25 PCR cycles may be used for this purpose; otherwise (is the number of PCR cycles is higher), the library cannot be considered representative (Matz 2002).

Second, the libraries that are enriched in full-length cDNA sequences are preferable for many purposes. For instance, the sequencing of full-length cDNA provides a complete amino acid sequence of the respective protein, whereas in the case of an incomplete cDNA fragment sequencing, special additional procedures may be required to determine the whole structure. Moreover, functional screening of cDNA library becomes impossible if it is not enriched in full-length cDNAs.

Third, for many cDNA library applications (e.g. search for differentially expressed genes), it is essential to minimize the distortion of cDNA representation in a library with respect to initial mRNA. In other words, the content of individual cDNAs in the library must be proportional to the copy number of the initial RNAs. In contrast, for some other tasks the concentrations of different individual cDNAs in a library must be equalized.

Methods to decrease the prevalence of highly abundant transcripts and to equalize transcript concentrations in a cDNA library are called "cDNA normalization" methods. Libraries, in which the disparity in concentrations of cDNAs for various genes is smaller than the initial disparity in concentrations of cDNAs or mRNAs for these genes in the original sample, are referred to as *normalized* ones. In the ideal case (which is never the case in the reality), all cDNA concentrations are equal in the normalized library.

Normalized cDNA libraries are used mostly to discover new genes transcribed at relatively low levels, which are expressed in the biological objects under study. The initial frequency of cDNA abundance for poorly transcribed genes in a general cDNA library may be under 2×10^{-6}, whereas in an ideally normalized library it will be ~25-fold higher (Soares et al. 1994). Therefore, the so-called complete sequencing of normalized libraries (more or less comprehensive library sequencing with at least fivefold clone coverage) requires at least one order less sequenced clones as compared with nonnormalized ones. Also, cDNA normalization significantly reduces the number of analyzed clones in functional screenings.

1.2 Evaluation of cDNA Library Normalization Efficacy

To control the efficacy of cDNA library normalization, it is possible to compare the concentrations of "major" and "minor" transcripts in the respective library before and after normalization step. In such cases one can use quantitative (or semiquantitative) PCR or pseudo Northern blot hybridization. The more efficient is the normalization, the less is the difference between major and minor transcript concentrations after normalization.

Another approach comprises an exhaustive sequencing of a cloned cDNA library and further analysis of the transcripts presented there. Any particular transcript will be presented by *n* copies, where *n* means a whole positive number. Then, it is helpful to plot the relationship between the number of individual transcripts (the number of genes) occurring *n* times, and *n* itself graphically. The closer is this graph to Poisson standard distribution, the more effective is the normalization.

1.3 Basic Approaches to Generate Normalized cDNA Libraries

Today, two basically distinctive approaches to the generation of normalized cDNA libraries have been proposed.

The first approach is a physical isolation of all transcripts in equal quantities from the original cDNA library. Theoretically, denatured genomic DNA of the same organism immobilized on an inert carrier can be used for this task (Weissman 1987). Saturating hybridization of such absorbent with denatured cDNA, followed by the separation and subsequent elution of bound cDNA will result in generation of normalized cDNA library (since almost all genes are presented in the genome by one or few copies, roughly equal quantities of cDNAs for these genes will hybridize with genomic DNA). Nevertheless, rare transcripts escape an analysis in this system, as it is impossible for them to reach highly enough concentrations, which are required for an efficient hybridization with genomic DNA.

The second group of methods is based on denaturation – hybridization of double-stranded cDNA molecules (or heteroduplexes mRNA or first strand

cDNA). For each specific transcript, the hybridization rate will be proportional to the square of its concentration since nucleic acid hybridization kinetics follows a second-order equation. Therefore, the higher the initial concentration of individual cDNA molecules, the more molecules will reassociate after the denaturation (i.e. convert back to double-stranded form). Thus, after denaturation – reassociation step, the cDNA library consists of two pools: (1) the double-stranded fraction, which accumulates the most part of initial cDNA and (2) the single-stranded fraction that includes initial individual cDNAs in substantially equalized rather low concentrations (Young and Anderson 1985; Galau et al. 1977).

For the latter group of methods, the challenging step is the separation of single- stranded pools from double-stranded pools in the cDNA molecules. Until last few years, no simple, efficient, and reliable methods were available to solve this problem, but recently a significant progress in this field was achieved (Zhulidov et al. 2004).

The following approaches are broadly used now for the separation of single- and double-stranded fractions of cDNA libraries: (1) physical separation, (2) selective amplification, and (3) enzymatic removal of the double-stranded pool.

1. Physical separation of single- and double-stranded fractions is mostly realized through column chromatography with hydroxyapatite (Ko 1990; Soares et al. 1994), preliminary immobilization of cDNA first strand on solid carrier beads (Sasaki et al. 1994; Tanaka et al. 1996), and through the biotinylation of initial RNA (Carninci et al. 2000). However, these methods share three substantial disadvantages: they are labor-intensive and complicated, they require large amounts of starting material and, frequently, they are poorly reproducible.

2. Selective amplification of normalized cDNA libraries is based on PCR suppression (Luk'yanov et al. 1996) (see Chapter 2). Being adequate for many research strategies, this method, however, is applicable only for fragmented cDNA libraries (those lacking complete cDNAs) and, moreover, results in a loss of rare transcripts.

3. Theoretically, specific enzymatic removal of the double-stranded fraction is considered the most convenient and robust method of cDNA library normalization. The enzyme (or enzyme mix) used should efficiently hydrolyze the double-stranded cDNA while leaving single-stranded cDNA intact. Such enzyme should be active at high temperature conditions (60–70°C), thus allowing avoiding formation of secondary structures by single-stranded cDNAs (to prevent the background digestion by the enzyme of secondary structure-forming single-stranded cDNAs, which would be the case at lower temperatures). The attempt to use a mix of frequent-cutter restriction enzymes turned out to be rather inefficient (Coche and Dewez 1994). Recently, a new enzyme – duplex-specific nuclease (DSN) with above-mentioned features – was isolated and described (Shagin et al. 2002). It was the basis to develop a simple and effective method of full-length-enriched cDNA normalization (so-called duplex-specific nuclease [DSN] normalization, see Section 2) prior to library cloning (Zhulidov et al. 2004).

2. METHODS OF cDNA NORMALIZATION

2.1 cDNA Normalization by Means of Hydroxyapatite Column Chromatography

2.1.1 Normalization of fragmented cDNA

The first method of normalized cDNA library generation was proposed by Minoru Ko and coauthors in 1990 (Ko 1990). The proposed procedure involves the following stages:

1. A double-stranded cDNA sample synthesized on a poly(A) + RNA basis is sonicated to generate shorter fragments
2. 200–400 bp long cDNA fragments are purified from the agarose gel
3. Purified cDNA fragments are ligated to double-stranded oligonucleotide adapter, and are further PCR amplified using the adapter as a site for primer annealing
4. After cDNA denaturation and reassociation, single- and double-stranded pools are separated with standard chromatography on hydroxyapatite columns (Sambrook et al. 1989)
5. The fraction enriched in single-stranded DNA (that which was not bound by the column) is overprecipitated and amplified

It should be noted that his approach requires several cycles of normalization, i.e. the amplified cDNA obtained after the first round of normalization is again redenatured and rehybridized with subsequent column separation. After three rounds of normalization, the amplified cDNA is cloned. In a model experiment, described by Ko and colleagues, the overall normalization was rather efficient: after three rounds, the mean difference between the frequencies of occurrence of major and minor transcripts decreased from ~20,000-fold to 40-fold value.

Independently and almost at the same time, Patanjali and coauthors proposed a very similar protocol (Patanjali et al. 1991). Their protocol provided a sufficiently normalized cDNA library of 400–1600 bp long fragments in just one round of normalization. The authors succeeded to almost completely equalize the concentrations of different individual cDNAs. However, in other hands this method was poorly reproducible.

2.1.2 Normalization of full-length-enriched cDNA libraries

The next step in the evolution of normalization methods based on the physical separation of fractions by means of hydroxyapatite column chromatography was the development of a technique to normalize full-length-enriched cDNA libraries. The method proposed by Soares et al. (1994) involves the following stages (Figure 1):

1. Generation of a directional full-length-enriched cDNA library in a phagemid vector with subsequent isolation of its single-stranded form
2. Synthesis of short (~200 bp) second strand cDNA fragments by means of limited primer extension while using single-stranded cDNA (generated at the previous step) as a template
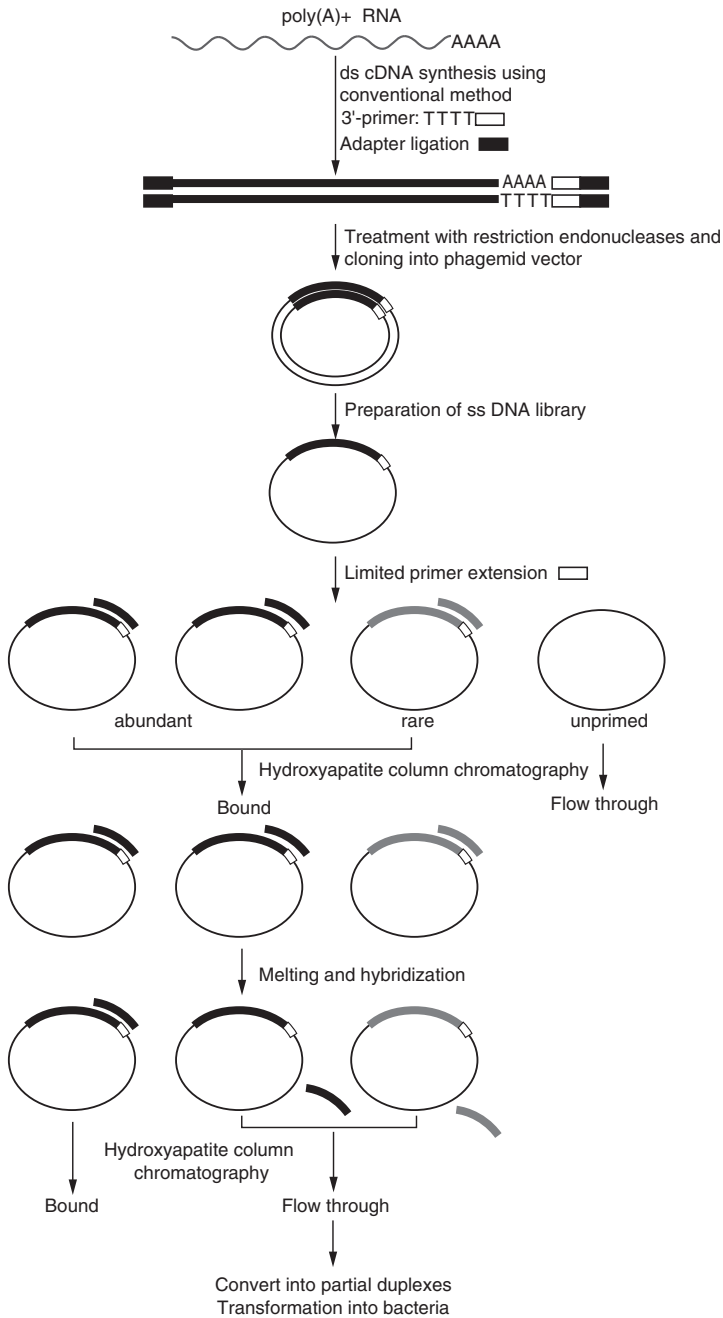
Figure 1. Schematic outline of the normalization of full-length-enriched cDNA libraries using hydroxyapatite column chromatography. (From Soares et al. 1994.) Thick black lines represent abundant transcripts, thick gray line represents rare transcripts. Rectangle represents adapter sequence and its complementary sequence.

3. Elimination of excess circular single-stranded molecules (which were not involved in the synthesis of the second strand cDNA) with hydroxyapatite column
4. Denaturation and rehybridization of double-stranded fragments separated at the previous stage
5. Isolation of a normalized single-stranded DNA pool (from the previous step) by means of hydroxyapatite column chromatography
6. Conversion of single-stranded plasmids isolated at the stage (5) to partial duplexes
7. Electroporation to deliver the plasmids into bacteria to propagate normalized cDNA library

After two rounds of normalization, the authors managed to produce a library with very low (less than tenfold) differences between the frequencies of analyzed transcripts. Marra and colleagues further reported (Marra et al. 1999), that the full-length sequences constitute ~27% of a total cDNA pool in a normalized library generated by means of such protocol.

As compared with other techniques for cDNA normalization, the method proposed by Soares et al. (1994) is the most widely used. For example, it was successfully applied in several expressed sequence tag (EST) large-scale sequencing projects (Urmenyi et al. 1999; Reddy et al. 2002; Caetano et al. 2003).

However, the procedures with hydroxyapatite columns are labor-intensive. To effectively separate fractions on hydroxyapatite columns, 10 µg of initial cDNA are needed. Chromatography-based separation also requires high temperature conditions (60°C) thus complicating manipulations with a sample. Besides, low resolution of hydroxyapatite columns results in contamination of a single-stranded pool with short double-stranded molecules, thus reducing the normalization efficiency. This approach also requires performing several cycles of denaturation – hybridization steps followed by the subsequent chromatography.

Some general shortcomings of this first group of methods should be mentioned. Briefly, the methods proposed by Ko and Pananjali are not applicable to generate normalized full-length cDNA libraries. The method by Soares et al. (1994) includes amplification of the cloned cDNA library before normalization. At this stage, the growth of cDNA-harboring colonies varies depending on the plasmid length; this results in underrepresentation of many long full-length cDNAs after bulk amplification of the library.

2.2 Generation of Normalized Full-Length-Enriched cDNA Libraries Using DNA Immobilization on a Solid Support

Simultaneously with Soares et al. (1994), Sasaki and coauthors proposed an alternative approach to normalize full-length-enriched RNA. The authors hybridized mRNA with short first strand cDNA immobilized on latex beads; this allowed an easy elimination of hybrids by means of sample centrifugation

after hybridization (Sasaki et al. 1994; Tanaka et al. 1996). In this case, normalization is performed in the following manner:

1. First strand cDNA (the so-called driver) is synthesized through the seeding of poly(T) oligonucleotides covalently bound with solid-phase support (latex beads)
2. After cDNA synthesis, the cDNA/RNA heteroduplexes are denatured; cDNAs fixed on latex beads are separated by centrifugation
3. cDNAs are prehybridized with poly(A) oligonucleotides (to reduce the background of the further hybridization); poly(A) + RNA is added to the resulting cDNAs fixed on latex beads, and the whole mix is allowed to hybridize
4. After hybridization, latex beads with DNA/RNA heteroduplexes are eliminated by centrifugation; RNAs remaining in solution are reprecipitated and then utilized in the next round of normalization
5. "Normalized" RNAs are used for cDNA synthesis

Effective normalization required 3–4 consistent cycles of the above procedures. After four cycles, ~96–98% RNA molecules was eliminated. Remaining RNA was considered a normalized one; this fraction was used to synthesize cDNA. The authors of this technique succeeded to generate full-length-enriched cDNA library with less than 30-fold difference of transcript abundances. Using similar protocol to generate normalized cDNA library from human heart muscle, Tanaka et al. (1996) managed to create a library with 1395 nonoverlapping (unique) clones from a total of 3040 sequenced ones (~46% of unique clones in the library).

This technique did not become popular due to its multistep protocol, large starting amounts of poly(A) + RNA required (~25 µg), and difficulties in manipulations with RNA. Besides, hybridization of DNA immobilized on a solid-phase support is less effective and slower as compared with hybridization in solution (Young and Anderson 1985).

2.3 Normalization of Full-Length cDNA with the use of Biotinylated RNA

Another approach to generate normalized libraries based on physical separation of fractions has been proposed by Carninci et al. (2000). This procedure involves hybridization of biotinylated RNA and first strand cDNA with subsequent elimination of bio-RNA–DNA hybrids using magnetic beads. The cDNA first strands synthesized via *cap trapper* method (Carninci et al. 1996) are hybridized with biotinylated natural or synthetic RNA. Hybridization in a formamide buffer allows one to reduce the reaction temperature down to 42°C thus preventing the degradation of long single-stranded DNA molecules. To eliminate generated bio-RNA–cDNA hybrids, streptavidine beads are further used. The single-stranded cDNA remaining in solution is used to synthesize cDNA second strand and to further create cDNA library. The basic stages of cDNA synthesis by means of *cap trapper* method and full-length cDNA normalization are illustrated in Figure 2.
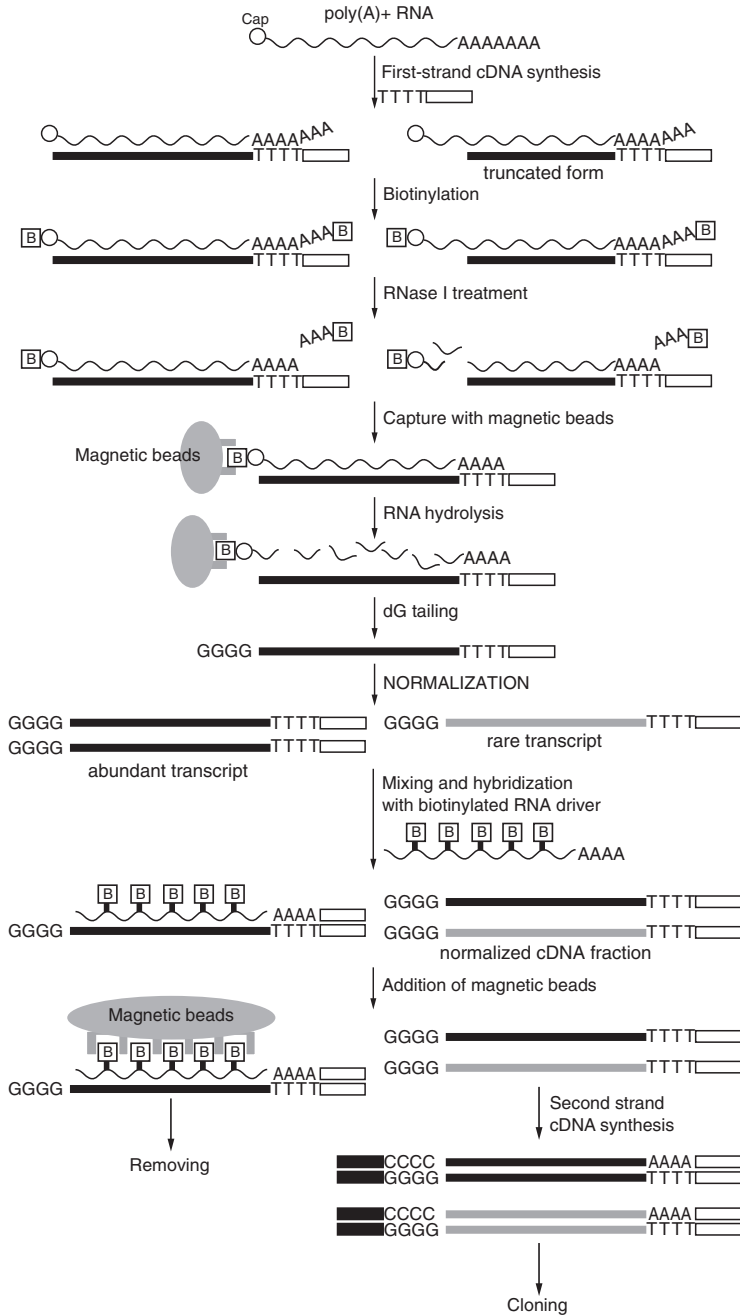
Cap   poly(A)+ RNA

AAAAAAA

First-strand cDNA synthesis

TTTT

AAAAAAA
TTTT

AAAAAAA
TTTT
truncated form

Biotinylation

B   AAAAAAA B
TTTT

B   AAAAAAA B
TTTT

RNase I treatment

AAA B
B   AAAA
TTTT

AAA B
B   AAAA
TTTT

Capture with magnetic beads

Magnetic beads   B   AAAA
TTTT

RNA hydrolysis

B   AAAA
TTTT

dG tailing

GGGG   TTTT

NORMALIZATION

GGGG   TTTT
GGGG   TTTT
abundant transcript

GGGG   TTTT
rare transcript

Mixing and hybridization
with biotinylated RNA driver

B B B B B
AAAA

B B B B B
AAAA
GGGG   TTTT

GGGG   TTTT
GGGG   TTTT
normalized cDNA fraction

Addition of magnetic beads

Magnetic beads
B B B B B
AAAA
GGGG   TTTT

GGGG   TTTT
GGGG   TTTT

Removing

Second strand
cDNA synthesis

CCCC   AAAA
GGGG   TTTT

CCCC   AAAA
GGGG   TTTT

Cloning

Figure 2. Schematic outline of the normalization of full-length-enriched cDNA with the use of biotinylated RNA. (From Carninci et al. 2000.) Black lines represent abundant transcripts, gray lines represent rare transcripts. Rectangle represents adapter sequence and its complementary sequence.

The method described above provides normalized full-length cDNA libraries. This approach was demonstrated to be highly efficient. Representation of major housekeeping gene transcripts in a normalized library decreased 250-fold as compared with the initial library. The difference between representations of moderately abundant and highly abundant transcripts in a normalized library was less than 30-fold. Full-length sequences constituted ~65–70% of all clones. As to disadvantages of this method, large amounts of poly(A) + RNA required and labor-intensiveness because of its multistep protocol and complexities in manipulations with RNA should be mentioned. Besides, several cycles of biotinylation are required due to rather low efficacy of RNA labeling with biotin. Finally, separation of biotin-modified and nonmodified pools usually involves several rounds.

2.4 Normalization of Fragmented cDNA by Means of Selective Amplification

cDNA normalization method free from any kind of physical separation of fractions has been proposed by Luk'yanov et al. (1996). This method is based on a selective amplification of normalized cDNA pool due to selective PCR suppression. As it has been shown, amplification of DNA flanked with inverted terminal repeats (ITR) in PCR with the primer that corresponds to ITR external segment is inhibited (the primer length should be well below ITR length). The principle of PCR suppression is described in detail in Chapter 2.

For normalization, double-stranded cDNA is treated with frequent-cutter restriction enzyme(s) and is separated into two samples. Each sample is ligated with specific suppression adapter oligonucleotide to generate ITRs at the ends of cDNA molecules. After ligation, samples are denatured and allowed to reanneal in different tubes (first hybridization). The double-stranded pool resulting from the first hybridization consists of homoduplexes harboring identical adapter sequences on their both 5′-termini. The two samples are then mixed (without denaturation) and left to hybridize together (second hybridization). Only normalized single-stranded fraction (which did not form homoduplexes at the previous stage) is exposed to the second hybridization. Together with homoduplexes, heteroduplexes flanked with different adapter sequences are generated in the second hybridization as well. Thus, following second hybridization, normalized single-stranded DNA partially converts to heteroduplex fraction. Subsequent to complementary extension of 3′-termini, heteroduplexes may be selectively PCR amplified with primers corresponding to external segments of suppressive adapters. Amplification of hybrids flanked with ITRs is suppressed by stem-loop structures produced in the course of PCR (*selective PCR suppression*; see Chapter 2). Single-stranded molecules are not amplified since they contain only one primer annealing site (Figure 3).

The procedure described above provides substantially normalized cDNA libraries within just one round of normalization. This technique is rather efficient, easy, and very well-reproducible. However, it is not free from certain imperfections
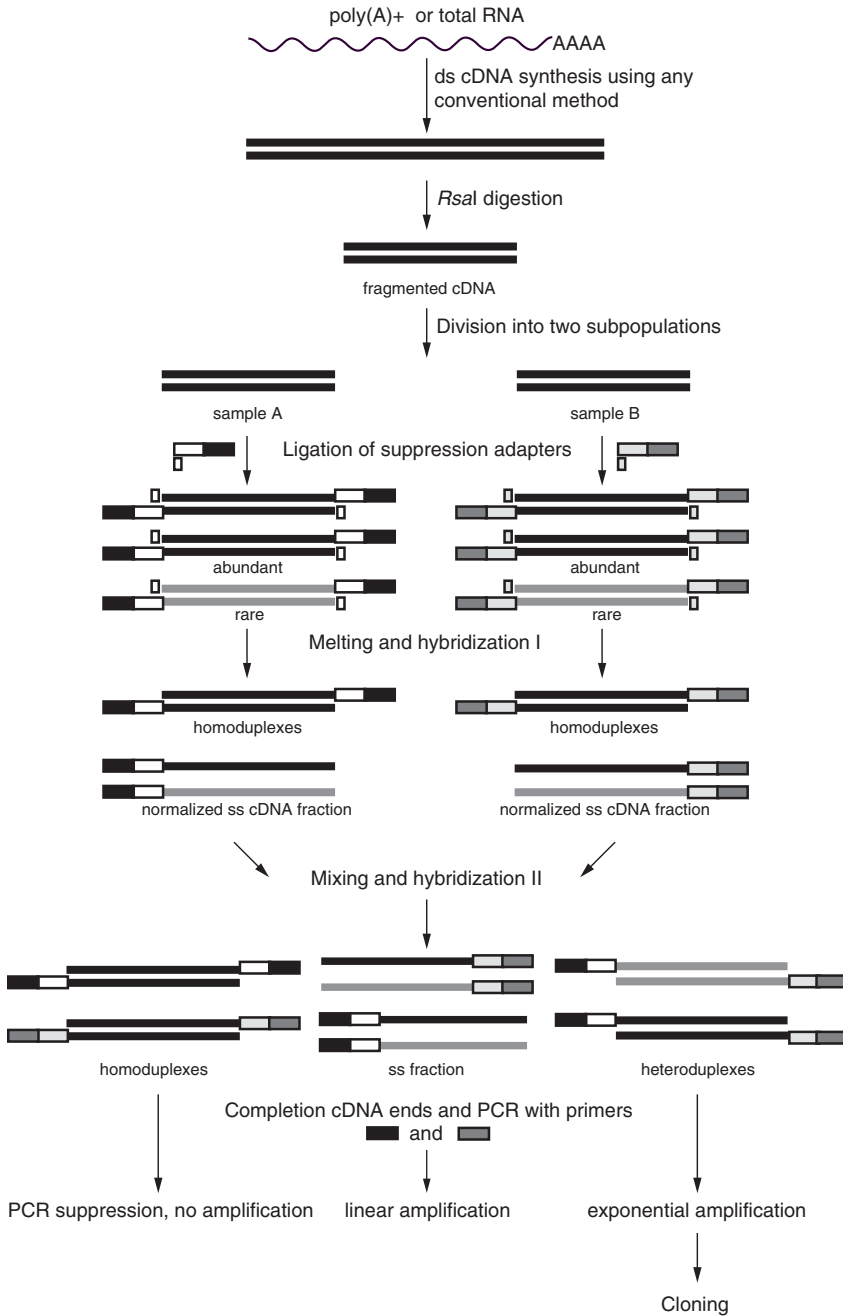
Figure 3. Schematic outline of the normalization of fragmented cDNA libraries using selective amplification. (From Lukyanov et al. 1996.) Black lines represent abundant transcripts, gray lines represent rare transcripts. Rectangle represents adapter sequence and its complementary sequence.

as well: it does not permit to normalize full-length cDNA since the efficiency of PCR suppression is greatly reduced when long DNA molecules are amplified (see Chapter 2). Thus, this method of normalization is applicable only to fragmented cDNA libraries. Moreover, the concentration of low abundant transcripts may be less than it is required for an efficient hybridization, especially during the second hybridization (see above); this frequently results in their underrepresentation in the normalized library.

2.5 cDNA Normalization Using Frequent-Cutter Restriction Enzymes

While seeking to overcome the complexity and low reproducibility of cDNA normalization techniques based on physical separation of single- and double-stranded pools, Coche and Dewez (1994) proposed to use a mixture of frequent-cutter restriction endonucleases to eliminate double-stranded DNA fraction. According to the proposed procedure, first strand cDNA is synthesized through the seeding of oligo(T)-containing oligonucleotides covalently bound with magnetic beads. After second strand synthesis and the subsequent ligation of the cDNA with oligonucleotide adapter, cDNA is further denatured, reannealed, and then exposed to restriction endonucleases. Hydrolyzed fragments of double-stranded cDNA and single-stranded cDNA appear to be immobilized on magnetic beads. Then, second strand cDNAs are synthesized using adapter-specific primer extension and are further PCR amplified. At this stage, restriction endonuclease-digested cDNA fragments will not be amplified as they do not contain two primer annealing sites necessary for exponential PCR. The resulting PCR products were used for the next round of normalization.

Contrary to the expectations, this method was found to be rather inefficient. Even after three normalization steps, concentrations of major transcripts were greatly higher as compared with minor ones. This most probably could be due to a low efficiency of hybridization of cDNA immobilized on a solid-phase carrier. Moreover, treatment with restriction endonucleases is performed at a low temperature. In this case, significant loss of cDNAs tending to secondary structure formation occurs (as the resulting intermolecular duplex may contain a site recognized by the endonuclease).

2.6 Normalization of Full-Length-Enriched cDNA with Duplex-Specific Nuclease (DSN Normalization)

This approach is schematized in the Figure 4. The method involves denaturation of double-stranded cDNA flanked by adapter sequences and its subsequent renaturation, followed by enzymatic degradation of double-stranded DNA fraction (mostly formed by reannealed abundant transcripts) using DSN. Single-stranded fraction containing transcripts in more or less equalized concentrations is further PCR amplified (Zhulidov et al. 2004, 2005).

Similar to previously described techniques, DSN normalization is based on second-order in-solution hybridization kinetics. Standard hybridization conditions
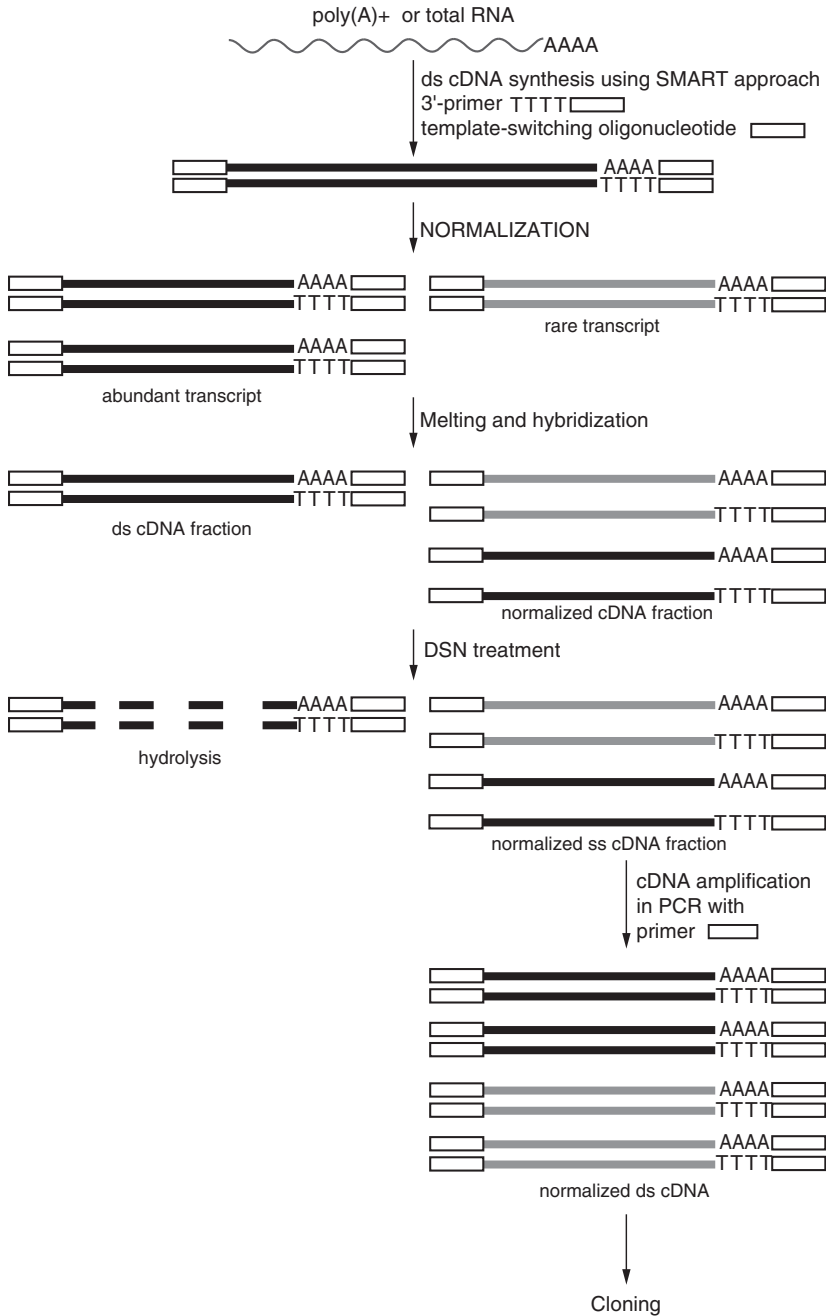
Figure 4. DSN normalization scheme. Black lines represent abundant transcripts, gray lines represent rare transcripts. Rectangle represents adapter sequence and its complementary sequence.

identical to those routinely used for the suppression subtractive hybridization (SSH) (Diatchenko et al. 1996) (Chapter 3) were utilized. Formamide buffer employed by Carninci et al. (2000), which efficiently prevents secondary structures formation could not be used here because it inhibits DSN necessary for double-stranded fraction digestion.

The rationale of this method is degradation of double-stranded fraction using DSN. A number of specific DSN features make it ideal for removing double-stranded DNA from complex mixtures of nucleic acids. DSN displays a strong preference for cleaving double-stranded DNA in both DNA–DNA and DNA–RNA hybrids, as compared to single-stranded DNA and RNA, irrespective of the sequence length. Moreover, the enzyme remains stable over a wide range of temperatures and displays optimal activity at 55–65°C (Shagin et al. 2002). Consequently, degradation of double-stranded DNA-containing fraction by this enzyme may occur at elevated temperatures, thereby avoiding loss of transcripts due to formation of secondary structures and nonspecific hybridization involving adapter sequences.

DSN normalization includes a step of PCR amplification of normalized single stranded cDNA fraction. To overcome PCR tendency to amplify shorter DNA fragments more efficiently than longer ones two basic approaches were applied. The first is the use of well-known long-distance PCR system (Barnes 1994) and the second is regulation of average length of complex PCR product by partial PCR suppression (Shagin et al. 1999). The last approach requires that cDNAs to be normalized have identical inverted adapter sequences at both termini (see Chapter 2 for the detailed description of the PCR suppression effect).

Adapter sequences can be introduced at the cDNA termini using various approaches, for instance, by adapter ligation or during cDNA synthesis utilizing the *SMART* approach (Zhu et al. 2001) or its modifications. SMART-prepared cDNA is enriched with full-length sequences and could be obtained both from poly(A) + and total RNA (even if only small amount of starting material is available). Synthetic adapter sequences are introduced to both 5′- and 3′-ends of cDNA during cDNA synthesis.

3. cDNA PREPARATION FOR DSN NORMALIZATION

3.1 RNA Requirements

The sequence complexity and the average length of the normalized cDNA library noticeably depend on the quality of experimental RNA starting material. There are several methods to isolate RNA providing stable RNA preparation from a majority of biological objects, for example, Trizol method (GIBCO/Life Technologies), Chomczynski and Sacchi method (Chomczynski and Sacchi 1987), and RNeasy kits (QIAGEN). After RNA isolation, RNA quality should be evaluated by means of denaturing formaldehyde or agarose gel electrophoresis (Sambrook et al. 1989).

The following characteristics indicate successful RNA isolation:

– For mammalian total RNA, two intensive bands at ~4.5 and 1.9 kb should be observed against a light smear. These bands represent 28S and 18S ribosomal RNAs (rRNAs). The ratio of intensities of these bands should be ~1.5–2.5:1. Intact mammalian poly A + RNA appears as a smear sized from 0.1 to 4–7 kb (or more) with faint 28S and 18S rRNA bands.

– With RNA from other sources (plants, insects, yeast, amphibians), the normal mRNA smear may not exceed 2–3 kb. Moreover, the overwhelming majority of invertebrates have 28s rRNA with a "hidden break" (Ishikawa 1977). In some organisms, interaction between the parts of 28s rRNA is rather weak, so the total RNA preparation exhibits a single 18s-like rRNA band. In other species, the 28s rRNA is more robust, so it is still visible as the second band.

If experimental RNA is shorter than expected and/or degraded according to electrophoresis data, preparation of fresh RNA is recommended after checking the quality of RNA purification reagents. In some cases, partially degraded RNA is only available (e.g. tumor samples or hard treated tissues). This RNA can be used to synthesize cDNA; however, the cDNA sample will contain reduced number of full-length molecules.

As representation of the resulting amplified cDNA depends on the initial amount of RNA used for the first strand cDNA synthesis, we recommend using higher starting amounts of RNA indicated in the selected cDNA synthesis protocol. In our experiments, the best results were achieved when cDNA synthesis was started with 0.5–1.5 μg of poly(A) + or total RNA.

3.2 cDNA Synthesis

The DSN normalization protocol provided below is optimized for cDNA prepared by the SMART approach (Zhu et al. 2001) or its modifications. However, some other full-length cDNA synthesis systems might be used for this purpose as well, e.g. RLM FirstChoice cDNA synthesis system produced by Ambion, which is based on the specific enzymatic modification of RNA 5′-cap by the tobacco acid pyrophosphatase and further ligation of an RNA adapter to it; this ensures full-length double-stranded cDNA production, if oligo d(T)-containing primer is used for the first strand cDNA synthesis. Also, an appropriate double-stranded cDNA can be prepared using the following commercially available cDNA synthesis kits:

– Mint cDNA synthesis kit (Evrogen Catalog #: SK001)
– SMART PCR cDNA synthesis kit (Clontech Catalog #: 634902)
– Super SMART PCR cDNA synthesis kit (Clontech Catalog #: 635000)

cDNA obtained is flanked by the identical adapter sequences (e.g. 5′-AAG CAG TGG TAT CAA CGC AGA GTA CT-3′) at both termini and allows nondirectional cloning of cDNA library only. Do not perform cDNA ends polishing before normalization procedure.

   If a directionally cloned cDNA library is required, we recommend using commercially available kits SMART cDNA library construction kit (Clontech Catalog # 634901) or Creator SMART cDNA library construction kit (Clontech Catalog # 634903) except for CDS III/3′ PCR primer. During first strand cDNA synthesis, the CDS III/3′ PCR primer must be replaced by CDS-3M adapter (5′-AAG CAG TGG TAT CAA CGC AGA GTG GCC GAG GCG GCC $(T)_{20}$VN-3′, where N = A, C, G, or T; V = A, G, or C). This adapter has the same exterior sequence as SMART IV oligonucleotide (provided in the kits listed above) and allows following amplification of the first strand cDNA in PCR with only one PCR primer (5′ PCR primer: 5′-AAG CAG TGG TAT CAA CGC AGA GT-3′).

   First strand cDNA synthesis should be performed according to the manufacturer protocol (SMART cDNA synthesis by LD PCR, section A, first-strand cDNA synthesis) with foregoing modification. For PCR amplification of double-stranded cDNA the kit manufacturer protocol should be slightly modified: the PCR mixture should contain doubled amount of 5′ PCR primer and should not contain CDS III/3′ PCR primer.

   Synthetized cDNAs are flanked by adapter sequences with a common exterior part and different interior sequences with *SfiIA* or *SfiIB* restriction sites for cloning.

   The use of other cDNA synthesis kits may require additional optimization of the DSN normalization protocol provided.

   It is important to realize that adapters used to synthesize starting cDNA for directional cloning are longer than those used for nondirectional cloning. Using longer adapters leads to a reasonable decrease in the cDNA average length during PCR and often causes the appearance of a low-molecular weight fraction in the resulting normalized cDNA (which, in turn, makes it necessary to include a size-separation procedure to remove short cDNA fragments before cloning). Therefore, if directional cloning of cDNA library is not crucial for research purposes, we recommend preparing cDNA with shorter adapters (that are suitable for nondirectional cloning only).

3.3 cDNA Purification

Before normalization, double-stranded cDNA must be purified from primer excess, deoxyribonucleotidetriphosphate (dNTPs), and salts. Several commercial kits are suitable for cDNA purification like QIAquick PCR purification kit (Catalog # 28104, 28106, QIAGEN Inc.) or equivalent. It is necessary that the kit used effectively removes the primer excess.

4. DSN NORMALIZATION PROTOCOL

4.1 Materials

– Double-stranded cDNA flanked by the sequence 5′-AAG CAG TGG TAT CAA CGC AGA GT-3′ at both ends (see Section 5.3 for details)
– 4X Hybridization buffer (200 mM HEPES, pH 7.5; 2 M NaCl)*

- DSN enzyme (Evrogen)* in 50 mM Tris-HCl, pH 8.0 with 50% glycerol (1 U/μl)
- DSN storage buffer (50 mM Tris-HCl, pH 8.0)*
- 2X DSN master buffer (100 mM Tris-HCl, pH 8.0; 10 mM $MgCl_2$; 2 mM DTT)*
- DSN stop solution (5 mM EDTA)*
- Thermostable single-stranded DNA binding (SSB) protein (1.5–5 μg/μl; optional, see Section 5.4.4, step 3)
- PCR primer M1 (10 μM, 5′-AAG CAG TGG TAT CAA CGC AGA GT-3′)*
- PCR primer M2 (10 μM, 5′-AAG CAG TGG TAT CAA CGC AG-3′)*
- Long-distance (LD) PCR enzyme mix with 10X PCR buffer (hot start must be used to reduce nonspecific DNA synthesis during the PCR setup. PCR kits that are adapted for long-distance PCR and also include automatic hot start are recommended, for example, Advantage 2 PCR kit, Clontech Catalog #: 639206, 639207)
- 50X dNTP Mix (10 mM each nucleotide)
- Sterile nuclease free (millipore-filtered) water
- Mineral oil
- Blue ice
- Sterile 0.5 ml PCR tubes (thin-wall PCR tubes recommended. These tubes are optimized to ensure more efficient heat transfer and to maximize thermal-cycling performance)
- Sterile microcentrifuge 1.5 ml tubes
- Pipettors (P20, P200)
- Pipette tips
- Base and tray or retainer for holding tubes
- Vortex mixer
- Microcentrifuge
- Agarose gel electrophoresis equipment and reagents (1X TAE buffer, 1.2–1.5% agarose with ethidium bromide, 1 kb DNA size markers)
- PCR thermal cycler
- 98% ethanol
- 80% ethanol
- 3M NaAc (sodium acetate), pH 4.8

*These reagents are available from Evrogen Trimmer kits (Evrogen Catalog # NK001, NK002)

4.2 cDNA Precipitation

Note: Do not use any coprecipitants in the precipitation procedure.
1. Aliquot cDNA solution containing ~700–1300 ng of purified cDNA (see Section 5.3 for details) into a separate sterile tube. Store the remaining cDNA solution at −20°C.
2. Add 0.1 volumes of 3 M NaAc, pH 4.8, to the reaction tube.

3. Add 2.5 volumes of 98% ethanol to the reaction tube.
4. Vortex the mixture thoroughly.
5. Centrifuge the tube for 15 min at 12,000–14,000 rpm at room temperature.
6. Remove the supernatant carefully.
7. Gently overlay the pellets with 100 μl of 80% ethanol.
8. Centrifuge the tubes for 5 min at 12,000–14,000 rpm at room temperature.
9. Carefully remove the supernatant.
10. Repeat steps 4.2.6–4.2.8.
11. Air dry the pellet for 10–15 min at room temperature. Be sure the pellet has dried completely.
12. Dissolve the pellet in sterile water to the final cDNA concentration of ~100–150 ng/μl.
13. To check the cDNA quality and concentration, analyze 1 μl of cDNA solution using gel electrophoresis alongside 0.1 μg of 1 kb DNA size markers on a 1.5% agarose/EtBr gel in 1X TAE buffer.

4.3 Hybridization

Note: Before starting hybridization, make sure 4X hybridization buffer has been allowed to stay at room temperature for at least 15–20 min. The buffer used should have no visible pellet or precipitate. If necessary, warm the buffer at 37°C for about 10 min to dissolve any precipitate.

1. For each sample to be normalized combine the following reagents in a sterile 1.5 ml tube: 4–12 μl double-stranded cDNA solution from step 4.2.12 (~600–1200 ng of cDNA); 4 μl 4X hybridization buffer; and sterile water to the total volume of 16 μl.
2. Mix the contents and spin the tube briefly in a microcentrifuge.
3. Aliquot 4 μl of the reaction mixture into each of the four appropriately labeled (e.g. see Table 1) sterile PCR tubes.

Table 1. Setting up DSN treatment

Component\tube	Experimental			Control
	Tube 1	Tube 2	Tube 3	Tube 4
	(S1 DSN1) (μl)	(S1 DSN1/2) (μl)	(S1 DSN1/4) (μl)	(S1 control) (μl)
DSN enzyme in storage buffer	1	–	–	–
1/2 DSN dilution	–	1	–	–
1/4 DSN dilution	–	–	1	–
DSN storage buffer	–	–	–	1

4. Overlay the reaction mixture in each tube with a drop of mineral oil and centrifuge the tubes at 14,000 rpm for 2 min.
5. Incubate the tubes in a thermal cycler at 98°C for 2 min.
6. Incubate the tubes at 68°C for 5 h, and then proceed immediately to DSN treatment. Samples may be hybridized for as little as 4 h, or as long as 7 h. Do not allow the incubation to proceed for more than 7 h. Do not remove the samples from the thermal cycler before DSN treatment.

4.4 DSN Treatment

1. Shortly before the end of the hybridization procedure, prepare the following dilutions of the DSN enzyme in two sterile tubes:
   (a) Add 1 μl of DSN storage buffer and 1 μl of DSN solution to the first tube. Mix by gently pipetting the reaction mixture up and down. Label the tube as 1/2 DSN.
   (b) Add 3 μl of DSN storage buffer and 1 μl of DSN solution to the second tube. Mix by gently pipetting the reaction mixture up and down. Label the tube as 1/4 DSN.
   (c) Place the tubes on ice.
2. Preheat the DSN master buffer at 68°C.
3. Add 5 μl of the preheated DSN master buffer to each tube containing hybridized cDNA (see Section 5.4.3, step 6), spin the tube briefly in a microcentrifuge and return it to the thermal cycler. Do not remove the tubes from the thermal cycler except for the time necessary to add preheated DSN master buffer.

Note: When integrity of very long transcripts (more than 5 kb) is crucial, 1 μl of the thermostable SSB protein (with concentration 1–5 μM) can also be added to the reaction mixture. In this case, less PCR cycles are required (for 3–4 cycles) to amplify normalized cDNA (see Section 5.4.5, step 11).

4. Incubate the tubes at 68°C for 10 min.
5. Add DSN enzyme as specified in Table 1. After DSN adding return the tubes immediately to the thermal cycler. Do not remove the tubes from the thermal cycler except for the time necessary to add DSN enzyme. When the tube is left at room temperature after DSN adding, nonspecific digestion of secondary structures formed by ss DNA may occur to decrease the efficacy of normalization.
6. Incubate the tubes in the thermal cycler at 68°C for 25 min.
7. Add 10 μl of DSN stop solution, mix the contents, and spin the tubes briefly in a microcentrifuge.
8. Incubate the tubes in the thermal cycler at 68°C for 5 min.
9. Take the tubes off the thermal cycler and place them on ice.
10. Add 20 μl of sterile water to each tube. Mix the contents and spin the tubes briefly in a microcentrifuge. Place the tubes on ice. The samples obtained can be stored at −20°C for up to 2 weeks and used afterwards to prepare more normalized cDNA.

4.5 First Amplification of Normalized cDNA

1. Prepare a PCR master mix for all reaction tubes. In a sterile 1.5 ml tube, combine the following reagents in the order shown (per rxn): 40.5 μl sterile water, 5 μl 10X PCR buffer, 1 μl 50X dNTP mix, 1.5 μl PCR primer M1, 1 μl 50X polymerase mix (total volume is 49 μl).

2. Mix well by vortexing and spin the tube briefly in a microcentrifuge.

3. Aliquot 1 μl of each diluted cDNA (from Step 4.4.10) into an appropriately labeled sterile PCR tube.

4. Aliquot 49 μl of the PCR master mix into each of these reaction tubes.

5. Mix the contents by gently flicking the tubes. Spin the tubes briefly in a microcentrifuge.

6. If the thermal cycler used is not equipped with a heated cover, overlay each reaction with a drop of mineral oil. Close the tubes, and place them into the preheated thermal cycler.

7. Subject the tubes to 7 PCR cycles using the following program: 95°C for 7 s, 66°C for 30 s, 72°C for 4 min. Please note than this program is optimized for MJ Research PTC-200 thermal cycler and polymerase mixtures noted in Section 5.4.1. Optimal PCR parameters may vary with different thermal cyclers, polymerase mixes, and templates.

8. Use the control tube (see Table 1) to determine the optimal number of PCR cycles as described in following steps 9–10 (below). Store other tubes on ice.

9. For each control tube, determine the optimal number of PCR cycles as shown in Figure 5):
   (a) Transfer 12 μl from the 7-cycle PCR tube to a clean microcentrifuge tube (for agarose/EtBr gel analysis).
   (b) Run two additional cycles (for a total of 9) with the remaining 38 μl of the PCR mixture.
   (c) Transfer 12 μl from the 9-cycle PCR tube to a clean microcentrifuge tube (for agarose/EtBr gel analysis).
   (d) Run two additional cycles (for a total of 11) with the remaining 26 μl of the PCR mixture.
   (e) Transfer 12 μl from the 11-cycle PCR tube to a clean microcentrifuge tube (for agarose/EtBr gel analysis).
   (f) Run two additional cycles (for a total of 13) with the remaining 14 μl of the PCR mixture.

10. Use 5 μl of each aliquot of each PCR reaction (from Step 9) for gel electrophoresis alongside 0.1 μg of 1 kb DNA size markers on a 1.5% agarose/EtBr gel in 1X TAE buffer. Store the remaining material on ice. Use Figure 7 to determine optimal number of PCR cycles required for amplification of each of the control tubes ("X"). Choosing the optimal number of PCR cycles ensures that the double-stranded cDNA will remain in the exponential phase of amplification. When the yield of PCR products stops increasing with every additional cycle, the reaction has
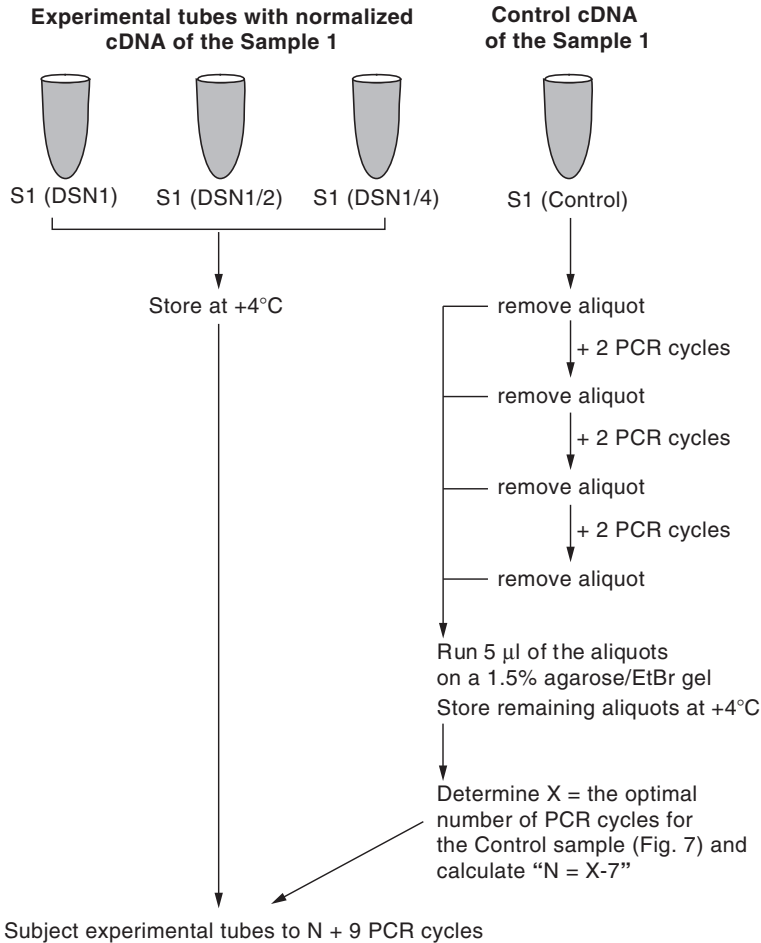
Experimental tubes with normalized
cDNA of the Sample 1

Control cDNA
of the Sample 1

S1 (DSN1)   S1 (DSN1/2)   S1 (DSN1/4)   S1 (Control)

Store at +4°C

remove aliquot

+ 2 PCR cycles

remove aliquot

+ 2 PCR cycles

remove aliquot

+ 2 PCR cycles

remove aliquot

Run 5 μl of the aliquots
on a 1.5% agarose/EtBr gel
Store remaining aliquots at +4°C

Determine X = the optimal
number of PCR cycles for
the Control sample (Fig. 7) and
calculate "N = X-7"

Subject experimental tubes to N + 9 PCR cycles

Figure 5. Optimizing PCR parameters for normalized cDNA amplification. Scheme of the experi-
mental procedures.

reached its plateau. The optimal number of cycles for your experiment
should be one or two cycles less than that needed to reach the plateau.
Figure 6 provides an example of how the analysis should proceed. Be con-
servative: when in doubt, it is better to use fewer cycles than too many.

11. Retrieve the 7-PCR tubes from ice, return them to the thermal cycler, and if
necessary, subject them to additional N cycles (where $N = X - 7$). Then, imme-
diately, subject the tubes to additional 9 cycles. Altogether, control tube should
be subjected to X PCR cycles, whereas experimental tubes should be subjected
to $X + 9$ PCR cycles, where X is the optimal number of PCR cycles determined
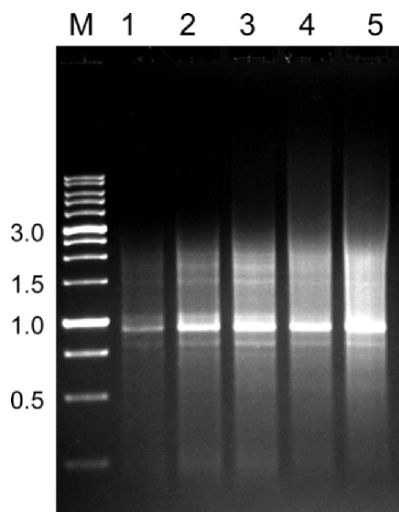for the control tube. In the example shown in Figure 6, the optimal number of

Figure 6. Analysis for optimizing PCR parameters. 5 μl of the aliquots from the control tube were analyzed using gel electrophoresis on a 1.5% agarose/EtBr gel in 1X TAE buffer: 1 – after 7 PCR cycles; 2 – after 9 cycles; 3 – after 11 cycles; 4 – after 13 cycles; 5 – after 15 PCR cycles. Lane M: 1 kb DNA ladder size markers, 0.1 μg loaded. In this experiment, after 11 cycles, a smear appears in the high-molecular weight region of the gel, indicating that the reaction is overcycled. Because the plateau is reached after 11 cycles, the optimal number of cycles for this experiment is 9.

PCR cycles determined for control cDNA in the control tube was 9. Thus, in this example $X = 9$, and $N = 9–7 = 2$. Hence, in this example, 7-PCR experimental tubes should be subjected to $2 + 9$ additional PCR cycles.

Note: If SSB protein is used during DSN treatment (see Section 5.4.4, step 3), less PCR cycles are required to amplify normalized cDNA, e.g. in this case, experimental tubes should be subjected to $N + 6$ additional PCR cycles.

12. When the cycling is completed, analyze 5 μl from each tube using gel electrophoresis alongside 5 μl aliquot from control PCR tube (with the optimal PCR cycle number) and 0.1 μg of 1 kb DNA size markers on a 1.5% agarose/EtBr gel in 1X TAE buffer.

13. Select the tube(s) with efficient normalization as described in Section 5.4.6. For comparison, Figure 7 shows a characteristic gel profile of normalized human placenta cDNA. If cDNA from two or more tubes seems well normalized, combine the contents of these tubes in one sterile 1.5 ml tube, mix well by vortexing and spin the tube briefly in a microcentrifuge. Amplified normalized cDNA obtained can be stored at −20°C for up to 1 month and used afterwards to prepare more normalized cDNA.

4.6 Preliminary Analysis of the Normalization Results

1. Compare the intensity of the banding pattern of your PCR products from experimental tubes with that from the control tube and with the 1 kb DNA ladder size markers (0.1 μg run on the same gel). Use Figure 7 as an example.
   – If the smear from the experimental tubes is much fainter than that shown for the control, PCR undercycling could be the problem. Subject experimental
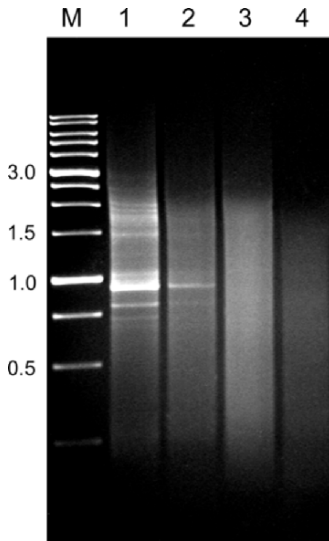
Figure 7. Analysis of cDNA normalization results. 5 µl aliquots of the PCR products were analyzed using gel electrophoresis on a 1.5% agarose/EtBr gel. Lane M: 1 kb DNA ladder size markers, 0.1 µg loaded. Lane 1: cDNA from control tube Lane 2: cDNA from S1_DSN1/4 Tube. Lane 3: cDNA from S1_DSN1/2 tube. Lane 4: cDNA from S1_DSN1 tube. In this experiment efficient normalization was achieved in the S1_DSN1/2 tube (lane 3). In the S1_DSN1/4 tube (lane 1) normalization was not completed, in the S1_DSN1 tube (lane 4) DSN treatment was excessive, resulting in partial cDNA degradation.

tubes to two or three additional PCR cycles and repeat electrophoresis. If there is still a strong difference between the overall signal intensity of PCR products from all experimental tubes and from the control tube, it may indicate that normalization process was superfluous.

– If the overall signal intensity of PCR products from the experimental tubes is much stronger than that shown for the Control, especially if the bright bands are distinguishable, it may indicate that normalization process was not successful. Test DSN activity and repeat normalization with doublet amounts of DSN.

– If the overall signal intensity of PCR products from the experimental tubes is similar to that in the control tube, select the tube(s) with efficient normalization using the instruction below.

2. A typical result, indicative of efficient normalization, should have the following characteristics:

– The pattern of PCR products from the experimental tube(s) containing efficiently normalized cDNA looks like smears without clear bands, whereas a number of distinct bands are usually present in the pattern of PCR products from the nonnormalized control tube.

– The average length of PCR products from the experimental tube(s) containing efficiently normalized cDNA is congruous with the average length of the PCR products from the nonnormalized control tube.

4.7 Second Amplification of Normalized cDNA

Reamplification of normalized cDNA allows to avoid cDNA degradation due to residual DSN activity and to prepare more cDNA for library cloning. If you plan to estimate normalization efficiency before cloning, it is essential to amplify control nonnormalized cDNA simultaneously.

1. Aliquot 2 μl of normalized cDNA (see Section 5.4.5, step 13) into a sterile 1.5 μl tube; add 20 μl of sterile water to the tube, mix well by vortexing and spin the tubes briefly in a microcentrifuge.
2. Aliquot 2 μl of control cDNA (from aliquot with optimal PCR cycling, see Section 5.4.5, steps 9 and 10) into another sterile 1.5 μl tube; add 20 μl of sterile water to the tube, mix well by vortexing and spin the tubes briefly in a microcentrifuge.
3. Aliquot 2 μl of diluted normalized cDNA from step 1 into an appropriately labeled sterile PCR tube.
4. Aliquot 2 μl of diluted control cDNA from step 2 into another appropriately labeled sterile PCR tube.
5. Preheat a thermal cycler to 95°C.
6. Prepare a PCR master mix for all reaction tubes. In a sterile 1.5 ml tube, combine the following reagents in the order shown (per rxn): 80 μl sterile water; 10 μl 10X PCR buffer; 2 μl 50X dNTP mix; 4 μl PCR primer M2; 2 μl 50X polymerase mix (total reaction volume 98 μl).
7. Aliquot 98 μl of the PCR master mix into each of the reaction tubes (from steps 3 and 4).
8. Mix contents by gently flicking the tubes. Spin the tubes briefly in a microcentrifuge.
9. If your thermal cycler is not equipped with a heated cover, overlay each reaction with two drops of mineral oil. Close the tubes, and place them into the preheated thermal cycler.
10. Commence thermal cycling using the following program: 95°C for 7 s, 66°C for 30 s, 72°C for 4 min. Subject the tubes to 12 cycles. Please note than this program is optimized for MJ Research PTC-200 thermal cycler and polymerase mixtures noted in Section 5.4.1. Optimal PCR parameters may vary with different thermal cyclers, polymerase mixes, and templates.
11. When the cycling is completed, analyze 5 μl of the PCR product using gel electrophoresis alongside 0.1 μg of 1 kb DNA size markers on a 1.5% agarose/EtBr gel in 1X TAE buffer to check the PCR quality and concentration. If necessary, subject the PCR tubes to 1–2 additional PCR cycles.
12. Now, you have normalized double-stranded cDNA that can be used for cloning into a vector of your choice. This cDNA can be stored at −20°C for up to 3 months. To estimate normalization efficiency after cDNA library cloning, sequence 100 randomly picked clones from the library. In a well-normalized library, redundancy of the first 100 sequences should not exceed 5%. You can also estimate normalization efficiency before cloning using
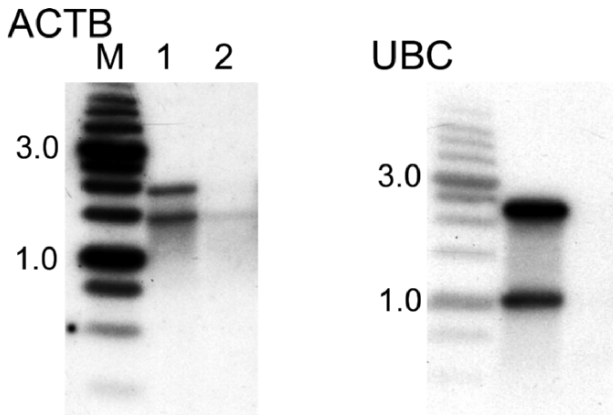
Figure 8. Virtual Northern blot analysis of abundant (ACTB, UBC) transcripts in nonnormalized (lane 1) and normalized (lane 2) cDNAs.

quantitative PCR or Virtual Northern blotting (Franz et al. 1999) with marker genes of known abundance. In both cases, it is done by comparing the abundance of known cDNAs before and after normalization. A typical result of Virtual Northern blot of nonnormalized and normalized cDNA with *ACTB*- and *UBC*-derived probes is shown in Figure 8.

Note: ACTB and *UBC* genes are expressed at high levels in most human tissues and cell lines; however, there could be some exceptions. If in the samples of your particular interest *ACTB* and *UBC* transcripts belong to intermediate or low abundance groups, unchanged or slightly increased concentration of these transcripts in normalized cDNA could be observed. In this case, select other marker genes that are highly abundant your samples to test normalization efficiency.

REFERENCES

Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD (1994) Molecular Biology of the Cell. Garland Publishing, New York

Barnes WM (1994) PCR amplification of up to 35-kb DNA with high fidelity and high yield from lambda bacteriophage templates. Proc Natl Acad Sci USA 91:2216–2220

Caetano AR, Johnson RK, Pomp D (2003) Generation and sequence characterization of a normalized cDNA library from swine ovarian follicles. Mamm Genome 14:65–70

Carninci P, Kvam C, Kitamura A, Ohsumi T, Okazaki Y, Itoh M, Kamiya M, Shibata K, Sasaki N, Izawa M, Muramatsu M, Hayashizaki Y, Schneider C (1996) High-efficiency full-length cDNA cloning by biotinylated CAP trapper. Genomics 37:327–336

Carninci P, Shibata Y, Hayatsu N, Sugahara Y, Shibata K, Itoh M, Konno H, Okazaki Y, Muramatsu M, 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

Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem 162:156–159

Coche Th, Dewez M (1994) Reducing bias in cDNA sequence representation by molecular selection. Nucleic Acids Res 22:4545–4546

Diatchenko L, Lau Y-FC, Campbell AP, Chenchik A, Mogadam F, Huang B, Lukyanov S, Lukyanov K, Gurskaya N, Sverdlov ED, Siebert PD (1996) Suppression subtractive hybridization, a method for generating differentially regulated or tissue-specific cDNA probes and libraries. Proc Natl Acad Sci USA 93:6025–6030

Franz O, Bruchhaus II, Roeder T (1999) Verification of differential gene transcription using virtual northern blotting. Nucleic Acids Res 27:e3

Galau GA, Klein WH, Britten RJ, Davidson EH (1977) Significance of rare mRNA sequences in liver. Arch Biochem Biophys 179:584–599

Guigo R, Agarwal P, Abril JF, Burset M, Fickett JW (2000) An assessment of gene prediction accuracy in large DNA sequences. Genome Res 10:1631–1642

Ishikawa H (1977) Evolution of ribosomal RNA. Comp Biochem Physiol B 58:1–7

Ko MS (1990) An "equalized cDNA library" by the reassociation of short double-stranded cDNAs. Nucleic Acids Res 18:5705–5711

Luk'yanov KA, Gurskaya NG, Matts MV, Khaspekov GL, D'vachenko LB, Chenchik AA, D'evich-Stuchkov SG, Luk'yanov SA (1996) A method for obtaining equalized cDNA libraries based on polymerase chain reaction suppression. Rus J Bioorg Chem 22:587–591

Marra M, Hillier L, Kucaba T, Allen M, Barstead R, Beck C, Blistain A, Bonaldo M, Bowers Y, Bowles L, Cardenas M, Chamberlain A, Chappell J, Clifton S, Favello A, Geisel S, Gibbons M, Harvey N, Hill F, Jackson Y, Kohn S, Lennon G, Mardis E, Martin J, Mila L, McCann R, Morales R, Pape D, Person B, Prange C, Ritter E, Soares M, Schurk R, Shin T, Steptoe M, Swaller T, Theising B, Underwood K, Wylie T, Yount T, Wilson R, Waterston R (1999) An encyclopedia of mouse genes. Nat Genet 21:191–194

Matz MV (2002) Amplification of representative cDNA samples from microscopic amounts of invertebrate tissue to search for new genes. Methods Mol Biol 183:3–18

Patanjali SR, Parimoo S, Weissman SM (1991) Construction of a uniform-abundance (normalized) cDNA library. Proc Natl Acad Sci USA 88:1943–1947

Reddy AR, Ramakrishna W, Sekhar AC, Ithal N, Babu PR, Bonaldo MF, Soares MB, Bennetzen JL (2002) Novel genes are enriched in normalized cDNA libraries from drought-stressed seedlings of rice (*Oryza sativa* L. subsp. *indica* cv. Nagina 22). Genome 45:204–211

Sambrook J, Fritsch EF, Maniatis T (1989). Molecular Cloning: A Laboratory Manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Sasaki YF, Ayusawa D, Oishi M (1994) Construction of a normalized cDNA library by introduction of a semi-solid mRNA-cDNA hybridization system. Nucleic Acids Res 22:987–992

Shagin DA, Lukyanov KA, Vagner LL, Matz MV (1999) Regulation of average length of complex PCR product. Nucleic Acids Res 27(18):e23

Shagin DA, Rebrikov DV, Kozhemyako VB, Altshuler IM, Shcheglov AS, Zhulidov PA, Bogdanova EA, Staroverov DB, Rasskazov VA, Lukyanov S (2002) A novel method for SNP detection using a new duplex-specific nuclease from crab hepatopancreas. Genome Res 12:1935–1942

Soares M, Bonaldo M, Jelene P, Su L, Lawton L, Efstratiadis A (1994) Construction and characterization of a normalized cDNA library. Proc Natl Acad Sci USA 91:9228–9232

Tanaka T, Ogiwara A, Uchiyama I, Takagi T, Yazaki Y, Nakamura Y (1996) Construction of a normalized directionally cloned cDNA library from adult heart and analysis of 3040 clones by partial sequencing. Genomics 35:231–235

Urmenyi TP, Bonaldo MF, Soares MB, Rondinelli E (1999) Construction of a normalized cDNA library for the *Trypanosoma cruzi* genome project. J Eukaryot Microbiol 46:542–544

Weissman SM (1987) Molecular genetic techniques for mapping the human genome. Mol Biol Med 4:133–143

Young BD, Anderson MLM (1985) Quantitative analysis of solution hybridisation. In: Hames BD, Higgins SJ (eds) Nucleic Acid Hybridisation. IRL Press, Oxford-Washington DC, pp 47–71

Zhu YY, Machleder EM, Chenchik A, Li R, Siebert PD (2001) Reverse transcriptase template switching: a SMART approach for full-length cDNA library construction. Biotechniques 30:892–897

Zhulidov PA, Bogdanova EA, Shcheglov AS, Vagner LL, Khaspekov GL, Kozhemyako VB, Matz MV, Meleshkevitch E, Moroz LL, Lukyanov SA, Shagin DA (2004) Simple cDNA normalization using kamchatka crab duplex-specific nuclease. Nucleic Acid Res 32:e37

Zhulidov PA, Bogdanova EA, Shcheglov AS, Shagina IA, Wagner LL, Khazpekov GL, Kozhemyako VV, Lukyanov SA, Shagin DA (2005) A method for the preparation of normalized cDNA libraries enriched with full-length sequences. Rus J Bioorg Chem 31:170–177