CHAPTER 7

SUBTRACTIVE HYBRIDIZATION WITH COVALENTLY MODIFIED OLIGONUCLEOTIDES

SHI-LUNG LIN, DONALD CHANG, JOSEPH D. MILLER, SHAO-YAO YING*

Department of Cell and Neurobiology, Keck School of Medicine, University of Southern California, 1333 San Pablo Street, BMT-403, Los Angeles, CA 90033.
**Corresponding author*
Phone: 002-1-323-442-1856; Fax: 002-1-323-442-3466; E-mail: sying@usc.edu or lins@usc.edu

Abstract: The ability to compare two different nucleic acid libraries has permitted inquiries into the role of differentially expressed genes or deleted/inserted genomic sequences involved in the mechanisms of neoplastic transformation, developmental regulation, physiological processing, pathological disorder, and therapeutic efficacy. Subtractive hybridization between two complementary DNA (cDNA) libraries is a powerful tool for identifying differentially expressed genes. In principle, an excess amount of modified subtracter cDNAs derived from cells of a control group are used to bind with tester messenger RNAs (mRNAs) or cDNAs isolated from the cells of interest. Because the subtracter cDNAs are modified to interfere with the amplification processes of reverse transcription (RT) and/or polymerase chain reaction (PCR), all subtracter-bound tester sequences are degraded and only the differentially expressed genes in the tester can be preserved for RT–PCR amplification. To improve the efficiency of subtractive hybridization, we have developed a chemical modification procedure to generate covalently binding cDNAs as the subtracter to capture the homologous tester sequences. We have also proved that the covalently bound duplex hybrids cannot be separated in PCR and thus are removed from the amplified differential gene sequences. Using the novel principle of covalently hybridized subtraction (CHS), we provide an easy, fast, and effective subtractive hybridization method for understanding the alterations of gene expression and/or chromosomal rearrangement in disordered cells in comparison with normal ones, which may reveal targets for gene therapy, eugenic improvement, pharmaceutical drug design, and investigation of etiological mechanisms.

Keywords: Subtracter, covalently modified subtracter, deamination of purines, carboxylation of pyrimidine bases, differential DNA, uracil–DNA glycosylase (UDG), aziridinylbenzoquinone (AZQ), covalent modification, modified pyrimidine, modified purine, covalently hybridized subtraction (CHS), acetic anhydride, alkaline acetic chloride, alkaline potassium permanganate, sodium cyanide/sulfuric acid mixture.

A. Buzdin and S. Lukyanov (eds.), Nucleic Acids Hybridization, 167–186.

Abbreviations: AZQ, aziridinylbenzoquinone; cDNA, complementary DNA; CHS, covalently hybridized subtraction; dNTP, deoxyribonucleotidetriphosphate; dUTP, deoxyuridine triphosphate; mRNA, messenger RNA; PCR, polymerase chain reaction; RB, retinoblastoma; RT, reverse transcription; UDG, uracil–DNA glycosylase.

TABLE OF CONTENTS

1. INTRODUCTION

Cells respond to environmental changes by altering their gene expression patterns to produce the proteins required for cell adaptation to the new environments. Alterations of gene expression patterns in a variety of biological processes as well as in response to external stimuli determine the cell fate and development of all life forms. Therefore, differentially expressed genes often encode the cryptic signals essential for cell adaptation and survival. To find these differentially expressed genes, several methods have been designed to detect and isolate different DNA sequences that are present in one expressed gene library but absent in the other. One of the most commonly used methods to accomplish this purpose is subtractive hybridization, involving the elimination of homologous (common) sequences from the mixture of two mutually compared DNA libraries. This kind of selective isolation can be done either between two complementary DNA (cDNA) libraries [1], or between two genomic DNA libraries [2]. In brief, this method relies upon the generation of cDNA or genomic DNA libraries from both control cells (subtracter DNAs) and cells after experiment, treatment, disorder, or change (tester DNAs). The two DNA libraries are then denatured and hybridized to each other, resulting in subtracter–tester hybrid formation if a sequence is common to both DNA populations. By removing the subtracter-bound common sequences, the remaining DNAs are the desired differential sequences, which are only present in the tester and expressive of the treatment, disorder, or change of interest.

Subtractive hybridization has been successfully used in the discovery of many functional genes and crucial genomic loci, such as T_4 and T_8 lymphocyte-surface glycoproteins [3,4], gamma–interferon-induced cytokines in monocytes [5], choroidermia loci [6], Duchenne muscular dystrophy-related loci [7], and human Y-chromosome-specific DNA [2].

As used here, tester DNAs refer to the DNA library isolated from a treated, mutated, infected, differentiated, or abnormal cell source, while subtracter DNAs refer to the DNA library isolated from a cell source with different status, such as nontreated, undifferentiated, or relatively normal cells (or tissues containing homogeneous cells). The tester library contains desired DNA sequences that are abundant in the tester but very limited in the subtracter. The differential DNA sequences represent the differences between tester and subtracter gene expression patterns (if two cDNA libraries are used as samples for comparison), or those of two compared genomic complexities (if two genomic DNA libraries are used). In this chapter, the isolation of the differential DNA sequences is achieved by using covalently modified subtracter DNAs to remove the tester homologues through covalent hybridization, which refers to a strong heat-stable binding interaction between the modified subtracter and nonmodified tester DNAs. The covalently modified subtracter refers to a DNA library that is chemically modified and thus capable of forming covalent bonds with homologous tester DNAs. The covalent modification comprises two chemical reactions, deamination of purines and carboxylation of pyrimidine bases. An amino-blocking agent is used to block or remove the amino-group of purine bases, such as acetic anhydride and alkaline acetic chloride. Then, a carboxylating agent is used to generate a carboxyl-group on the base structure of the subtracter's pyrimidines, such as a sodium cyanide/sulfuric acid mixture or hot alkaline potassium permanganate. The term "homologues" means DNA sequences that are common to both tester and subtracter DNA libraries.

In some cases, the isolated desired DNAs are so abundant in a cellular source that they can be directly detected and isolated without any enrichment. However, in most cases, the desired differential DNAs are too limited in amount to be detected and a polymerase chain reaction (PCR) is used to enrich the desired DNAs after subtractive hybridization [8,9]. When starting materials are limited, PCR is also used to enrich the subtracter and tester DNA libraries, via amplicon DNAs [10]. In short, such amplification is achieved by ligating a sequence-specific adapter to the ends of an endonuclease-restricted DNA library (amplicon), resulting in the generation of a primer-annealing region in the DNAs for subsequent PCR amplification.

2. SUBTRACTIVE HYBRIDIZATION METHODS

Several methods have been designed to detect and isolate differential DNA sequences which are present in one complementary [11] or genomic DNA [12] library but absent in the other. Representational difference analysis (RDA) was one of the first subtractive hybridization methods, and is particularly effective

in elimination of homologous DNA sequences from two mutually compared DNA libraries. This method relies upon the generation of representative DNA libraries from both control cells (subtracter) and cells of experimental treatments, disorders, or morphological/functional changes (tester). The two DNA libraries are then denatured and mixed, resulting in the formation of subtracter–tester hybrid duplexes when the sequences are common to both tester and subtracter DNA libraries. By removing the common sequences and surplus subtracter, the DNAs remaining in the mixture are the desired DNA sequences only present in the tester library, which is related to the treatment, disorder, or change of interest.

The use of biotinylated subtracter DNAs is an improvement to increase the specificity of subtractive hybridization via streptavidin-based chromatography and to reduce the amount of subtracter needed for repeated hybridization. For example, Straus et al. [13] hybridized biotinylated-deletion-mutant genomic DNAs with restricted-wild-type genomic DNAs, and then subtracted the unwanted common hybrid duplexes with avidin-coated beads. The remaining sequences were ligated to a specific adapter and amplified by PCR, resulting in discovery of genomic deletions present in the mutant but absent in the wild type. Concurrently, Duguid et al. [14] performed a similar experiment but using a biotinylated double-stranded cDNA library isolated from a normal hamster brain to hybridize with a nonmodified cDNA library from a scrapie-infected hamster brain, generating biotinylated hybrid complexes that were removed by biotin-binding avidin resins. The cDNAs remaining in the suspension were amplified and confirmed to be scrapie-infected specific gene sequences. These experiments often require several cycles of subtractive hybridization because of the low efficiency of formation of the biotin–avidin complex and the contamination of subtracter fragments. These drawbacks cause an unfortunate increase in lab work and potential loss of desired DNA sequences during the necessary repeated subtraction steps.

Lin et al. [15] and Bjourson et al. [16] devised a further improvement in subtractive hybridization that employed a biotinylated primer and a uracil-containing deoxynucleotide mixture (e.g. mixture of dATP, dCTP, dGTP, and dUTP) to generate biotinylated uracil-containing subtracter DNAs (U-DNAs) for enzymatic subtraction. In these cases, control and experimental DNA libraries were isolated from cells under different conditions, restricted by an endonuclease, and ligated to different primer-specific adapters. Then, a special PCR, using the uracil-containing deoxynucleotide mixture, was performed to produce the biotinylated subtracter U-DNAs, which were then hybridized with nonmodified tester DNAs, resulting in the formation of biotinylated and uracil-containing heterohybrid duplexes that were common to both libraries. Because the biotinylated heterohybrids were removed by streptavidin–phenol–chloroform extraction and surplus subtracter U-DNAs were further digested by uracil–DNA glycosylase (UDG), the remaining tester DNAs were the desired differential DNA sequences. However, this method still required tedious work in biotinylation and at least two rounds of extraction and chromatography.

Finally, subtraction with covalent affinity was invented to simplify the process of subtractive hybridization, in which an aziridinylbenzoquinone (AZQ) interstrand cross-linking agent was used to covalently subtract common sequences in both tester and subtracter libraries [17,18]. Single-stranded tester was firstly hybridized with single-stranded subtracter to form hybrid duplexes, and then the AZQ was added to generate covalent bonds between the hybridized duplexes, caused by the cross-linking interaction of guanine and cytosine. Because the AZQ cross-links all double-stranded nucleic acid sequences, this kind of external covalent binding greatly facilitates homolog subtraction after hybridization. However, during subtractive hybridization only the single-stranded tester and subtracter can be used as starting materials due to the interstrand cross-linking action of the AZQ-like agents, which prevents analysis of genomic DNA samples, limits the experiment to the starting materials, and prevents adapter-specific amplification of the final results. These disadvantages impose more restrictions in sample selection, less stability of sample storage and less sensitivity in the final detection step in comparison with traditional subtraction hybridization. Further, detection of the final desired DNA sequences is accomplished by a nonspecific random-primer extension reaction, which lowers the specificity of the final results.

To reduce the drawbacks of AZQ-like cross-linking agents, we have developed a chemical modification procedure to generate covalent binding DNAs as the subtracter to capture complementary tester mRNAs or DNAs, as shown in Figures 1 and 2 [19]. Modified subtracter DNAs were generated by carboxylating the base structures of certain subtracter DNA nucleotides to introduce strong covalent affinity between the modified subtracter and the homologous tester DNAs. After that, the desired differential (heterologous) DNA sequences remained in the hydrogen-binding form, whereas the hybridized common (homologous) DNA sequences were covalently bound. Since the covalent binding cannot be broken in a PCR, there is no amplification of the homologous sequences but great amplification of the desired differential sequences. The desired DNA sequences found after such a covalently hybridized subtraction (CHS) and subsequent selective amplification are the DNA sequences that only exist in the tester but not in the subtracter DNA library. This technique is designed for the subtractive hybridization of differential sequences between two DNA libraries from distinct cell sources and will allow more efficient isolations in experiments on cancer formation, development of gene therapy, and understanding of pathological status and developmental regulation.

3. COVALENT MODIFICATION

As used here, covalent modification refers to a chemical reaction in which direct covalent bonding with nonmodified tester sequences and modified subtracter sequences is generated by amino-blocking and carboxylating reagents. The amino-blocking reagent is a chemical, which can block or remove the amino-group of a

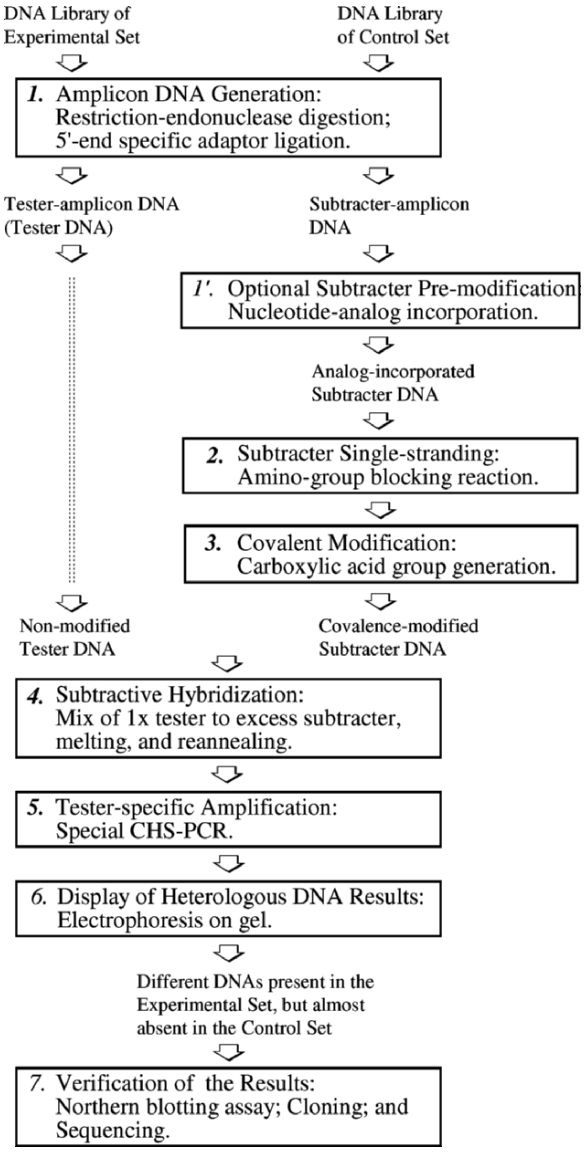


Figure 1. A flowchart protocol for the covalently hybridized subtraction (CHS) assay, illustrating the covalently bonded hybrid formation between tester and subtracter DNAs, and differential amplification steps after subtractive hybridization. The process is shown up to the final products of the first round CHS. To iterate another round of subtraction, the first difference products are used as tester following the same scheme to generate the second difference products and so on.

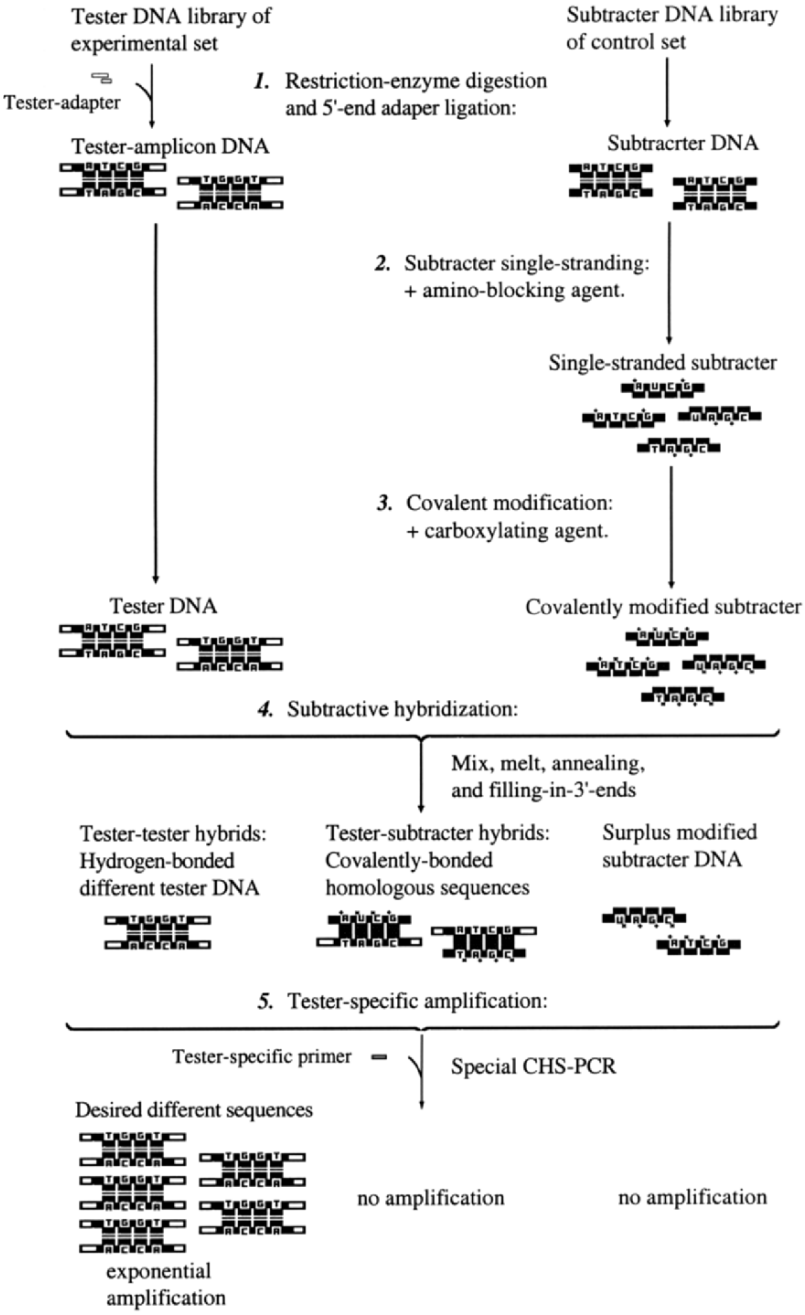


Figure 2. A schematic illustration of CHS subtractive hybridization in Figure 1.

nucleotide base, such as acetic anhydride or alkaline acetic chloride, while the carboxylating reagent is a chemical such as sodium cyanide/sulfuric acid mixture or hot alkaline potassium permanganate that can generate a carboxyl-group on the base structure of a modified subtracter sequence to allow covalent bonding with a nonmodified tester sequence.

The advantages of covalently modified subtracter sequences are as follows: First, during hybridization, the affinity between homologues can be greatly enhanced by covalent modification, such as the carboxyl-group on the C-5/C-6 of modified pyrimidines, resulting in peptide-like binding with the activating amino-group on the C-6/C-2 of nonmodified purines, respectively (Figure 3). Such covalent bonding between homologues fully inhibits any further reaction of the homologues and therefore reduces contamination with common homologues and surplus subtracter sequences. Second, the covalently modified subtracter sequences are single-stranded and inert to each other, resulting in a high binding efficiency in heterohybrid formation between the modified subtracter and non-modified tester DNA strands rather than two modified subtracter strands. Third, because the covalent binding is an interstrand interaction occurring either between adenine and modified uracil or between guanine and modified cytosine, covalently pairing significantly increases the specificity of CHS, which occurs only between tester and subtracter sequences with highly matched base pairs.

In experiments (Figures 1 and 2), a subtracter DNA library is first prepared from the control samples, in the following steps: (a) restricting the initial DNA library with a restriction enzyme to generate 5′-cohesive termini on both ends; (b) ligating a specific adapter to the ends of the restricted DNAs to form a short template for binding with a specific PCR primer; and (c) incubating the adapter-ligated DNAs in PCR to permit the primer-dependent enrichment of the

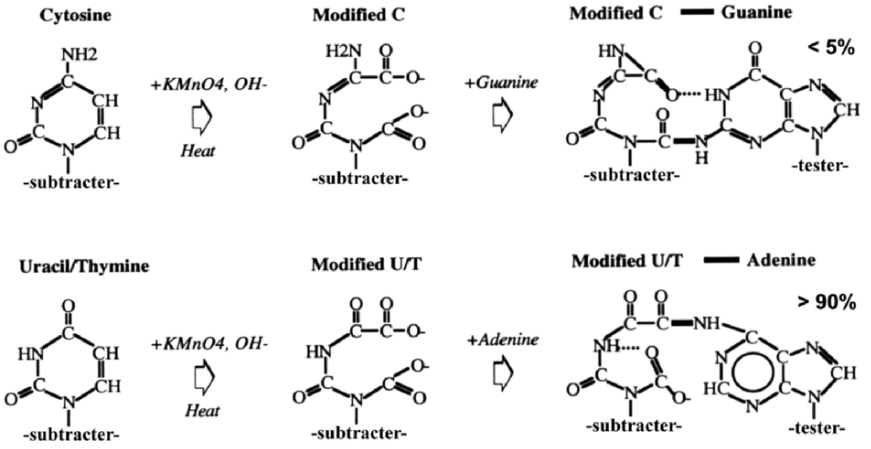


Figure 3. A detailed illustration of interstrand covalent bond formation in step 3 of Figure 2.

subtracter amplicon library. A subtracter amplicon library can be made from either a cDNA library or a genomic DNA library. The specific adapters and primers for PCR amplification are shown in Table 1.

Because covalent modification can be greatly facilitated by using some nucleotide analogs in the subtracter, we preferably incorporate deoxyuridine triphosphates (dUTP) into the subtracter amplicon sequences during PCR. For example, when 2′-deoxy-dUTPs instead of deoxythymidine triphosphates is used to generate the subtracter amplicon, the carboxylation reaction will occur only on the C-4 of uracil rather than the C-2, which is sometimes carboxylated if deoxythymidine triphosphates are used. Some alternative analog formulae are shown below, in which A, B, D, E, and F are selected from either a N or a CH group, G is a 2′-deoxy-D-ribose triphosphates, and X is a methyl group while Y is a H group and vice versa.

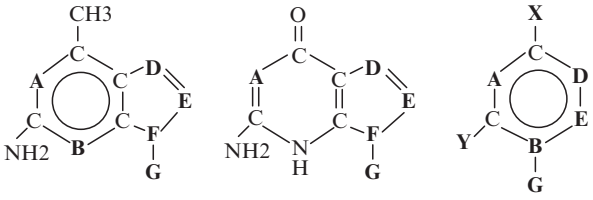


Table 1. The adapters and primers used in the CHS assay

Name	Application	Sequence
T-dpn2–24 mer	5′-ligation adapter; PCR specific primer for tester genomic DNA	5′-GCCACCAGAAGAGCGTG TACGCCA-3′
T-dpn2–12 mer	5′-ligation linker for tester genomic DNA	5′-GATCTGGCGTAC-3′ (5′dephosphorylated)
S-dpn2–24 mer	5′-ligation adapter; PCR specific primer for subtracter genomic DNA	5′-CGGTAGTGACTCGGT TAAGATCGA-3′
S-dpn2–12 mer	5′-ligation linker for subtracter genomic DNA	5′-GATCTCGATCTT-3′ (5′-dephosphorylated)
T-hpa2–24 mer	5′-ligation adapter; PCR specific primer for tester cDNA	5′-GCCACCAGAAGAGCGTG TACGTCC-3′
T-hpa2–11mer	5′-ligation linker for tester cDNA	5′-CGGGACGTACA-3′ (5′-dephosphorylated)
S-hpa2–24 mer	5′-ligation adapter; PCR specific primer for subtracter cDNA	5′-CGGTAGTGACTCGGT TAAGATCGC-3′
S-hpa2–11 mer	5′-ligation linker for subtracter cDNA	5′-CGGCGATCTTA-3′ (5′-dephosphorylated)

To prevent the reassociation of undesired subtracter–subtracter duplexes during hybridization, the amino-groups of subtracter DNAs must be blocked or removed by chemical blocking agents before covalent modification. The blocking reaction is preferably carried out by acetylating the amino-groups of the subtracter purines to form inactive acetamido-groups [20], which are incapable of binding to the carboxyl-groups of another modified sequence, which results in single-stranding the subtracter sequences. Acetic anhydride and alkaline acetic chloride are major ingredients in the amino-blocking reagent for CHS. Because the single-stranded subtracter DNAs will no longer protect their pyrimidine bases from oxidative modification, a carboxylating agent can easily oxidize the alkene, carbonyl or sometimes methyl groups [20] of the pyrimidine bases into activating carboxyl-groups, which are able to form covalent peptide-like bonds with the activating amino-groups of nonmodified tester sequences. Hot alkaline potassium permanganate is a major ingredient in carboxylating reagents due to the reaction of nucleophilic addition. Although adenine (A), guanine (G), cytosine (C), thymine (T), and uracil (U) bases were first used in the generation of covalently modified subtracter sequences, any nucleotide or its analog capable of being incorporated and modified into nucleotide sequences may be used as well. For example, such possible substitutes could be 2′-deoxyuracil derivatives, para-toluene derivatives etc. that have the same capability of being covalently modified.

To increase subtraction efficiency, the subtracter is carboxylated on C_4 of uracil/thymine or C_5/C_6 of pyrimidines to generate sufficient affinity for peptide-bond formation with the C_6/C_2 amino-groups of the tester purines, respectively. Most frequently, the carboxylated group is generated on the C_5 of uracil and covalently bound to the C_6-amino-group of adenine. These covalent bonds cannot be broken during PCR amplification; therefore, unbound tester can be amplified with thermostable DNA polymerases like Taq DNA polymerases.

After covalent modification, the denatured and modified subtracter DNAs only covalently hybridize with the homologous tester sequences in a mild alkaline condition, resulting in an increase of binding efficiency amend more complete subtraction. The preferred medium is a heat-stable EPPS/EDTA buffer (pH 8.5) in which the blocked amino-groups of the subtracter are released. The homologous sequences are reassociated at a temperature sufficient to inhibit nonspecific hybridization, preferably between about 60–80°C, most preferably about 68–74°C. The ratio of the modified subtracter to the nonmodified tester is preferably between about 5:1 and 10:1.

4. SUBTRACTIVE HYBRIDIZATION WITH COVALENTLY MODIFIED SUBTRACTERS

The present protocol describes an improved subtractive hybridization method, called the CHS assay, for finding sequences which differ between two cDNA, or genomic DNA libraries. This method is primarily designed for quickly isolating

differentially expressed genes (either up- or down-regulated), easily detecting large genomic deletions/insertions, and precisely searching chromosome-specific loci. The principle of CHS is dependent on the subtraction efficiency of covalent binding between common sequences (homologues) during PCR or cloning, resulting in no amplification of the homologues. The principle is based on: single-stranding of subtracter DNAs, covalent modification of the subtracter base structures, hybridization of the modified subtracter DNAs with other nonmodified tester DNAs to subtract covalently bound common sequences in both DNA libraries, and then amplification of remaining heterogeneous tester DNAs to quantify the differentially expressed genes in the tester. In conjunction with adapter-ligation and adapter-specific PCR amplification, very small subtracter and tester libraries can be used as starting materials for comparison.

Covalently modified subtractive hybridization provides an easy, fast, and effective isolation of desired differentially expressed sequences from either cDNA or genomic DNA libraries, following the steps shown in Figure 2: (a) providing a library of tester DNAs, which is ligated to a tester-specific adapter for selective amplification; (b) mixing the denatured tester DNAs with a library of denatured subtracter DNAs, which have been modified by chemical agents so as to covalently bond with the tester homologues to form a denatured product; (c) permitting both tester and subtracter DNAs in the denatured mixture to form double-stranded hybrid duplexes composed of hydrogen-bonded homoduplexes and covalently-bonded heteroduplexes; and (d) amplifying the hydrogen-bonded homoduplexes with tester-specific primers, thereby providing a differential DNA library enriched in DNAs unique to the tester condition. Steps b–e can be repeated on the enriched DNA library for more differential enrichment.

The utilization of covalently modified subtracter DNAs avoids several limitations of subtractive hybridization. First, during subtractive hybridization, the affinity of the subtracter to its homologous tester can be greatly enhanced by covalent modification, resulting in peptide-like binding to the tester amino-groups (Figure 3). Such covalent peptide-like binding between subtracter and tester homologues fully inhibit their amplification in PCR and therefore minimizes the needed subtractive hybridization cycle. Second, the covalently modified subtracter DNAs are single-stranded, resulting in strong binding to tester but not subtracter DNAs. Third, because the covalent binding occurs only in interstrand base pairing either between adenine–thymine (–uracil) or between guanine–cytosine, this feature significantly increases the specificity of hybridized subtraction and the sensitivity of differential sequence detection.

For generation of tester amplicon DNAs, the DNA library of experimental cells is digested by a restriction endonuclease on both ends, preferably a four-cutter restriction enzyme, and ligated to a specific 5′-adapter. This ligated DNA library called a tester-amplicon is then used to generate tester DNA by a template-dependent primer-extension reaction in the presence of a tester-specific primer, preferably using the adapters and primers listed in Table 1. On the other hand, subtracter amplicon DNAs are amplified by a similar procedure but with

a subtracter-specific adapter and primer, which share no affinity to the tester-specific ones. However, when the starting materials are abundant, the subtracter amplicon DNAs can be made by digestion with the four-cutter restriction enzyme but without the adapter ligation.

As shown in step 2 of Figure 2, blocking the activating amino-groups of the subtracter must be completed before covalent modification in order to prevent the formation of covalent bonds between subtracter and subtracter sequences. This blocking reaction is carried out by acetylating the amino-groups of subtracter purines to form inactive acetamido-groups [20], which are incapable of binding to the modified subtracter, resulting in single-stranded subtracter sequences. Acetic anhydride and alkaline acetic chloride are two preferred amino-blocking agents. Since the single-stranded subtracter cannot protect its base structures from oxidative agents, a carboxylating agent (as shown in step 3 of Figure 2) can easily oxidize the alkene, carbonyl, or sometimes methyl groups [20] of the subtracter bases to form carboxyl-groups, which are capable of forming covalent peptide-like bonds with the nonmodified amino-groups of the tester (Figure 3). Alkaline potassium permanganate and sodium cyanide/sulfuric acid mixtures are two preferred carboxylating agents based on the principles of oxidation and nucleophilic addition, respectively.

Following step 4 of Figure 2, tester DNAs are denatured and then hybridized with an excess amount of covalently modified subtracter at 68–72°C. The ratio of subtracter to tester DNAs is preferably in the range of about 5:1–10:1.

If the ratio is too high, enrichment of rare DNA sequences that only exist in the tester will not be obtained. If the ratio is too low, common nonspecific sequences will not be completely subtracted, and may cause false-positive contamination. The optimal ratio will vary depending on the stringency of subtractive hybridization between compared DNA libraries.

During the subtractive hybridization step, two kinds of hybrid duplexes are formed as follows: First, the tester–tester homohybrid duplexes, which consist of desired heterologous (differential) sequences only present in tester but almost absent in subtracter; And, the tester–subtracter heterohybrid duplexes which consist of common homologues present in both tester and subtracter. Because the linkage between tester and subtracter homologues is formed via covalent bonds, the resulting tester–subtracter heterohybrid duplexes cannot be amplified by PCR or vector cloning, in which each round of amplification requires the separation of the hybridized duplexes. Contrarily, the binding of tester–tester homohybrid duplexes is hydrogen-bonding (H-bond), which can be amplified by PCR or vector cloning. Therefore, the amounts of the desired differentially expressed sequences are greatly increased, whereas the contribution of common sequences will be negligible.

The subtracted tester DNAs can be subjected to another round of subtractive hybridization and amplification. The identified differential tester sequences are useful for DNA library selection assay and cloning analysis, representing the desired differential DNA sequences which are stimulated or up-regulated in the treated, mutated, infected, differentiated, or abnormal cells. By the same

token, the roles of tester and subtracter DNAs can be performed in a reverse order to isolate the suppressed or down-regulated sequences. These identified sequences can also be used to probe the full-length mRNAs or cDNAs from the tester library (if cDNA tester is used as a sample), or to locate the deleted/inserted loci in a specific chromosome by *in situ* hybridization (if genomic DNA is used). The information so obtained will provide further understanding of a variety of diseases, physiological phenomena, and genetic functions.

5. APPLICATIONS

The CHS will be very useful in the identification of genes specifically involved in development, cell differentiation, aging, and a variety of pathological disorders, such as cancer, genetic defects, autoimmune diseases, and any other disorders related to genetic malfunction. The identification of these differentially expressed genes will lead to the determination of their open-reading frames and translated polypeptide products, which may contribute to specific drug-design or therapy for regulation of these genes. Such therapeutic approaches include transcription inhibitors, monoclonal antibodies against the expressed protein, anti-sense RNA, and chemicals that can interact with the gene or its protein product to cure or alleviate related disorders. For example, the methods of the present invention can be used to screen candidate genes for gene therapy to correct inherent defects. When a defect is caused by stimulation of a specific unknown gene, the identification of this gene will help the design of antisense oligonucleotides against the gene or production of monoclonal antibodies against the corresponding protein product.

Alternatively, the CHS subtractive hybridization can also be used to screen some types of chromosomal abnormalities, such as deletion and insertion. Because genomic DNA fragments of less than 1 kb are prepared by restriction enzyme digestion before subtractive hybridization [11], the target deletion or insertion must be larger than this size for efficient amplification. The identification of these chromosomal deletions or insertions may contribute to the diagnosis or prognosis of certain virus infections, inherent problems, or developmental defects. For example, retinoblastoma (RB) gene deletion occurs in hereditary RB. If the deletion can be identified early, this information might allow therapeutic intervention to prevent the onset of RB.

Although the CHS assay is primarily designed for medical and biological research, the method will also be useful in pharmaceutical, agricultural, and environmental research, which involving biological systems. For example, when gene expression is compared between drug-treated and nontreated cells, the results may indicate the mechanism by which the drug acts. For another example, when the genomic DNA from disease-resistant plant cells is compared with that from disease-susceptible plant cells, the results will indicate the candidate loci for the resistance gene(s). Thus, CHS can provide a variety of information critical to understanding changes in gene expression across different genomes.

With the high efficiency of covalent modification and CHS, the labor- and time-consuming factors in subtractive hybridization assay can be reduced to the minimum. Also, the preparation of covalently modified subtracter is cheaper and more efficient than that of other modification methods. Most importantly, covalent modification can be carried out continuously with only a few changes of buffers. Taken together, these special features make CHS a fast, simple, effective, and inexpensive protocol for quickly isolating differentially expressed gene sequences of interest.

6. PROTOCOLS

6.1 Preparation of Subtracter and Tester DNA Libraries

For example, LNCaP cells, a prostate cancer cell line, were grown in Dulbecco's Modified Eagle's medium (DMEM) supplemented with 2% fetal calf serum. For a 3-day activin treatment, six dishes of experimental cells were treated with 1.5 ml of 200 $\mu g l^{-1}$ activin per day, while two dishes of control cells were not treated. On the fifth day after the first treatment, a 55% reduction in growth was observed in the experimental (tester) cells compared to the control (subtracter) cells by both microscopy and cell counting. All cells were trypsinized and total RNAs were isolated with a TRIzol reagent (GIBCO/BRL), respectively. After 1 μg of RNAs were mixed with the oligo-dT primer and heated to 65°C (10 min), a reverse transcription (RT) reaction was performed using a cDNA cycle kit (Invitrogen, CA), and all RT products (2 μg) were double-stranded with a DNA polymerase–ligase–RNase cocktail mixture [21]. About 1 μg of subtracter cDNAs was then digested by a four-cutting enzyme, such as Hpa2 (20 U for 5 h, 37°C), and ligated with a subtracter-specific primer 5′-pCGGTAGTGAC TCGGT TAAGA TCGC-3′ in the 5′-end, while tester cDNAs were ligated with a tester-specific primer 5′-pGCCACCAGAA GAGCGTGTAC GTCC-3′ in the same manner. This produced the subtracter and tester cDNA libraries, respectively.

6.2 Covalent Modification of Subtracter DNAs

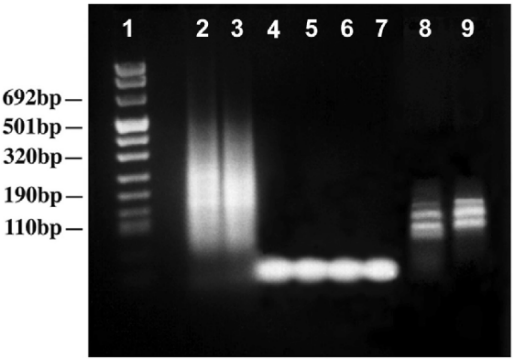
Subtracter cDNAs were diluted and amplified by the PCR with the subtracter-specific primer. During PCR, the recessed 3′-ends of the subtracter were filled by a Taq-like thermostable DNA polymerase (7 min, 72°C) with dATP (2 mM), dCTP (2 mM), dGTP (2 mM), dTTP (0.5 mM), and dUTP (3.5 mM). A 30-cycle amplification was performed (1 min, 95°C; 1 min, 72°C; 3 min, 68°C), and the amplified products, namely U-DNA sequences, were recovered by a Micropure–EZ column (Microcon) and resuspended in 10 μl of ddH_2O. About 50 μl of pure acetic anhydride was added (3 min, 94°C) into the U-DNAs of the subtracter to block the activating amino-groups by acetylation, by which the subtracter sequences also become single-stranded, and then the reaction was neutralized by 500 μl of Tris buffer (10 mM, pH 7.4). After the acetylated U-DNAs were recovered by a Micropure–EZ column and resuspended in a total of

10 µl of Tris buffer (10 mM, pH 7.4), 20 µl of an alkaline potassium permanganate reagent (1 mM $KMnO_4$, 1 mM NaCl, pH 10.0) was added (3 min, 80°C) to generate carboxyl-groups on C-5/C-6 of uracil/cytosine in the subtracter which can covalently bind to the amino-groups on C-6/C-2 of adenine/guanine of the tester, respectively. The carboxylated subtracter was finally recovered by a Micropure–EZ column and resuspended in a total of 10 µl of 10 mM N-[2-hydroxyethyl]piperazine-N′-[3-propanesulfonic acid] and [ethylenedinitrilo] tetraacetic acid (EPPS/EDTA) mixture buffer (pH 8.5). The modified subtracter must be used immediately for subtractive hybridization with the tester.

6.3 Subtractive Hybridization and CHS–PCR Amplification

For hybridization, the tester DNAs (500 ng) from the experimental cells were mixed with the covalently modified subtracter (3 µg) in 10 µl of the EPPS/EDTA buffer and denatured at 94°C for 5 min. The mixture was then vortexed, added to 1 µl of 5 M NaCl to adjust salt concentration, and incubated at 70°C (16 h). The hybridized DNAs were finally diluted with 20 µl $MgCl_2$ solution (2.5 mM) and amplified by PCR with the B-specific primer. A 20-cycle amplification was performed (1 min, 95°C; 3 min, 73°C) after nick translation with *Escherichia coli* DNA polymerase 1 plus T_4 DNA polymerase 3:1 mixture (5 min, 37°C without dNTPs; 35 min, 37°C with dNTPs), and the resulting products were phenol-extracted, isopropanol-precipitated, and resuspended in 15 µl of 10 mM Tris buffer for display on a 3% agarose electrophoresis gel (Figure 4, upper panel). The DNA bands shown on the gel were excised and recovered by a gel-extraction kit (Qiagen) to give the final difference products, and then further purified by a 4% nondenatured polyacryamide gel. The processes of subtractive hybridization with modified subtracter and selective amplification was repeated until clear bands are observed on the gel. Both cDNAs and genomic DNAs were processed in the same way as mentioned above.

As shown in the Figure 4 (upper panel), the subtracter DNA sequences were amplified with subtracter-specific primer (lane 2), the tester DNA sequences were amplified with tester-specific primer (lane 3), the subtracter DNAs were amplified with tester-specific primer (lane 4), the tester DNAs were amplified with subtracter-specific primer (lane 5), the subtracter DNAs were self-subtracted by modified subtracter U-DNA and amplified with subtracter-specific primer (lane 6), the tester DNAs were self-subtracted by modified tester U-DNA and amplified with tester-specific primer (lane 7), the subtracter DNAs were subtracted by modified tester U-DNA sequences and amplified with subtracter-specific primer (lane 8), and the tester DNAs were subtracted by modified subtracter U-DNA sequences and amplified with tester-specific primer (lane 9). The self-subtraction of subtracter to subtracter (lane 6) and tester to tester (lane 7) shows complete elimination of all sequences, while the mutual subtraction between tester and subtracter (lanes 8 and 9) presents different final results on the electrophoresis gel, indicating different gene expression in



CHS probes	Gene (size)	Homology	Change % (σ)	Function
Down-regulated				
probe 1 (LC2)	Myosin-like (1.1 kb)	99%	–47.6 (1.46)*	cytoskeleton
probe 2 (LC3)	CD168 (2.8 kb)	95%	–55.9 (1.54)*	cytoskeleton
probe 3 (LC8)	novel (2.0 kb)		–64.9 (3.09)**	?
probe 4 (LC9)	Helicase motif-like (1.3 kb)	95%	–60.5 (4.66)*	replication
probe 5 (LC12)	Pax2 (3.7 kb)	97%	–77.0 (2.37)**	proliferation
probe 6 (LC13)	eIf-4A1 (1.7 kb)	100%	–53.5 (0.00)*	translation
Up-regulated				
probe 7 (LT1)	novel (0.8 kb)		+728 (1.53)**	?
probe 8 (LT6)	rBub1-like (1.8 kb)	100%	+265 (4.38)**	spindle lesion apoptosis
probe 9 (LT11)	p53 (1.7 kb)	100%	+213 (5.35)**	G1 arrest

* n = 3, $p < 0.01$
** n = 4, $p < 0.01$

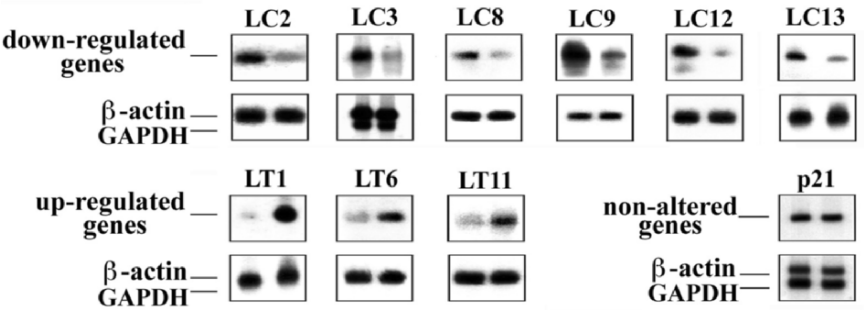


Figure 4. Identification of differentially expressed genes in human prostate cancer cells. LNCaP, after activin treatment.

the tester and the subtracter. The misuse of PCR primer (lanes 4 and 5) causes no amplification due to the specific affinity of the primer for its own adapter. Thus, based on the results of Figure 4, the CHS assay is sensitive and specific enough to subtract all homologous DNAs and distinguish the differential gene transcripts between two strands of DNA libraries after PCR amplification. Compared to RDA, the CHS assay has the advantages of low background and high efficiency, and usually the same results were obtained after one round of subtractive hybridization.

Sequence results for the final differentially expressed genes are shown in Figure 4, middle panel. The p53 gene (LT11) was previously known to be up-regulated in the activin-treated LNCaP cells. The known down-regulated genes (LC2, 3, 9, 12, 13) in the upper lane are related to cellular physiological functions, while the known up-regulated genes (LT6, 11) are involved in either cell-cycle regulation or apoptosis or both. All genes listed are transcriptionally altered by at least twofold. The size of each identified gene transcript is deduced from individual Northern blots, and the homology shown here indicates the sequence homology between the identified fragment and its deduced gene, rather than the entire identified sequence. Figure 4, bottom panel, shows an autoradiogram of positive Northern blots hybridized to the final differentially expressed genes displayed in CHS. The upper row (LC2–LC13) indicates six down-regulated genes mainly present in untreated LNCaP cells but not in the activin-treated cells, while the lower row (LT1, LT6, and LT11) shows three up-regulated genes significantly increased after activin treatment. The Northern blot of p21 is a negative control for activin-induced transcriptional alteration in LNCaP cells.

6.4 Covalent Binding Efficiency and Subtractive Stringency of CHS

To confirm the binding efficiency and subtraction efficacy of covalently modified subtracter, we used an apoptosin fragment as a target tester homologue, sharing about 300 base nucleotides with 70% homology to a subtracter sequence. Equal amounts of the tester fragment and subtracter were mixed, denatured, and subjected to DNA nuclease digestion, with or without hybridization. Hybridization was performed at 94°C for 3 min and then 70°C for 16 h in EEx3 buffer (30 mM EPPS, pH 8.5 at 20°C; 3 mM EDTA). Nuclease digestion was performed with a mixture of DNase I and nuclease S1 (50 U each, Roche) at 25°C for 10 min in 1x NS1 buffer (0.2 M NaCl, 50 mM sodium acetate, pH 4.5; 1 mM $ZnSO_4$, 0.5% glycerol). The results were electrophoresed on a 2% agarose gel as shown in Figure 5, left panel, showing lane 1, double-stranded apoptosin DNA fragments (200 ng); lane 2, single-stranded antisense subtracter (100 ng); lane 3, hybridization of the targeted DNA fragment and subtracter after nuclease digestion; lane 4, same as lane 3 but without digestion; and lane 5, hybridization of the targeted DNA to the subtracter (100 ng each) after nuclease digestion.

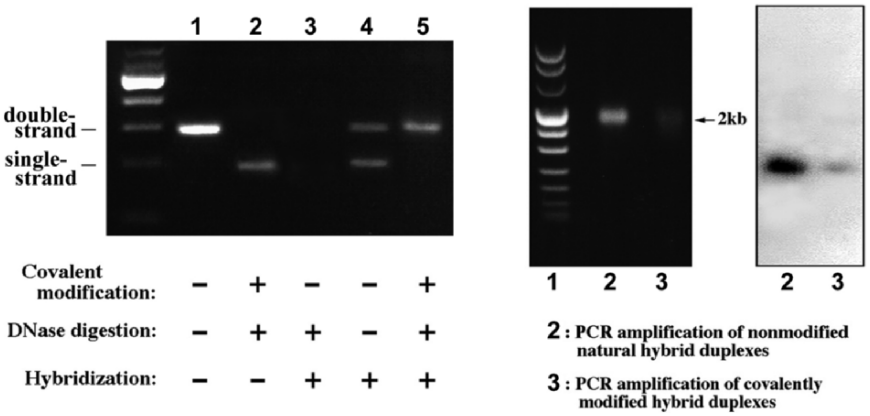


Figure 5. Detection of genomic deletion in retinoblastoma cells, Y-79.

The affinity of the subtracter for its homologous tester sequence is greatly enhanced by the covalent modification. As shown in Figure 5, left panel, the subtracter provides 100% binding efficiency (lane 5) compared to 53% in traditional probes (lane 4). Also, the interaction between subtracter and subtracter is prevented by acetylation of the amino-group of its purines, resulting in high binding efficiency between the subtracter and its targeted tester. Moreover, because of the covalent modification, the modified structures of the subtracter are highly resistant to nuclease digestion (lane 2), even after binding with the targeted sequences (lane 5). Such selective covalent bonding fully inhibits the functional activity of the targeted gene. Since covalently bound hybrid duplexes cannot be separated in cells, any enzymatic activity requiring single-stranded nucleotide templates will be effectively shut down. It has been shown that even a PCR cannot be performed through the covalently bound hybrid duplexes [22].

6.5 Identification of Genomic Deletion Using CHS

We also used the CHS assay to screen the genomic deletion in Y-79 cells and successfully identified a genomic fragment existing in the chromosome of normal retina cells but not in that of Y-79 cells (Figure 5). Y-79, a RB cell, has been known to contain an RB gene-deletion in its genome [23]. As a model of genomic subtraction by CHS, the genomic DNAs of normal retina and Y-79 cells were isolated by the IsoQuick nucleic acid extraction kit (Microprobe), respectively, restricted with Hpa2, and ligated to T-hpa-adapter and S-hpa-adapter, respectively, to give the tester (normal retina cells) and subtracter (Y-79). The sizes of restricted genomic DNAs were about 1–3 kb, which can be efficiently amplified by PCR. The uridine-analog was incorporated and covalently modified in subtracter genomic DNAs as described in Section 6.2. The subtractive hybridization and selective

amplification were performed as described in Section 6.3. The resulting differential genomic DNA sequence(s) of the tester were fractionized on a 1% agarose gel and confirmed by Southern blot analysis as shown in Figure 5, right panel. A signal was detected on the Southern blots of normal DNAs but not Y-79 DNAs, indicating an at least 2 kb deletion in RB exon 2 of the Y-79 genome.

REFERENCES

1. Davis et al. (1987) Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell 51:987–1000
2. Lamar et al. (1984) Y-encoded, species-specific DNA in mice: evidence that the Y chromosome exists in two polymorphic forms in inbred strains. Cell 37:171–177
3. Maddon et al. (1985) The isolation and nucleotide sequence of a cDNA encoding the T cell surface protein T4: A new member of the immunoglobulin gene family. Cell 42:93–104
4. Littman et al. (1985) The isolation and sequence of the gene encoding T8: a molecule defining functional classes of T lymphocytes. Cell 40:237–246
5. Chang et al. (1989) Cloning and expression of a gamma-interferon-inducible gene in monocytes: a new member of a cytokine gene family. Int Immunol 1:388–397
6. Nussbaum et al. (1987) Isolation of anonymous DNA sequences from within a submicroscopic X chromosomal deletion in a patient with choroideremia, deafness, and mental retardation. Proc Natl Acad Sci USA 84:6521–6525
7. Kunkel et al. (1985) Specific cloning of DNA fragments absent from the DNA of a male patient with an X chromosome deletion. Proc Natl Acad Sci USA 82:4778–4782
8. Wang et al. (1991) A gene expression screen. Proc Natl Acad Sci USA 88:11505–11509
9. Coochini et al. (1993) Identification of genes up-regulated in differentiating nicotania glauca pith tissue, using an improved method for construction a subtractive cDNA library. Nucleic Acids Res 21:5742–5747
10. Wicland et al. (1990) A method for differential cloning; Gene amplification following subtractive hybridization. Proc Natl Acad Sci USA 87:2720–2724
11. Lisitsyn N, Wigler M (1993) Cloning the differences between two complex genomes. Science 259:946–951
12. Hubank M, Schatz DG (1994) Identifying differences in mRNA expression by representational difference analysis of cDNA. Nucleic Acids Res 22:5640–5648
13. Straus et al. (1990) Genomic subtraction for cloning DNA corresponding to deletion mutations. Proc Natl Acad Sci USA 87:1889–1893
14. Duguid et al. (1990) Library subtraction of in vitro cDNA libraries to identify differentially expressed genes in scapic infection. Nucleic Acids Res 18:2789–2792
15. Lin SL, Ying SY (1999) Differentially expressed genes in activin-induced apoptotic LNCaP cells. Biochem Biophys Res Commun 257:187–192
16. Bjourson et al. (1992) Combined subtraction hybridization and polymerase chain reaction amplification procedure for isolation of strain-specific rhizobium DNA sequences. Appl Environ Microbiol 58:2296–2301
17. Hartley JA, Berardini M, Ponti M, Gibson NW, Thompson AS, Thurston DE, Hoey BM, Butler J (1991) DNA cross-linking and sequence selectivity of aziridinylbenzoquinones. Biochemistry 30:11719–11724
18. Hampson IN, Pope L, Cowling GJ, Dexter TM (1992) Chemical crosslinking subtraction; A new method for the generation of subtractive hybridization probes. Nucleic Acids Res 20:2899
19. Ying SY, Lin SL (1999) High performance subtractive hybridization of cDNAs by covalent bonding between specific complementary nucleotides. BioTechniques 26, 966–979
20. Solomons et al. (1996) Organic chemistry, 6th edn. Wiley, New York, pp 693, 803–804

21. Ueli et al. (1983) A simple and very efficient method for generating cDNA libraries. Gene 25:263–269
22. Ying SY, Chuong CM, Lin SL (1999) Suppression of activin-induced apoptosisby a novel antisense strategy. Biochem Biophys Res Commun 265:669–673
23. Lee, EY, Bookstein R, Young LJ, Lin CJ, Rosenfeld MG, Lee WH (1988) Molecular mechanism of retinoblastoma gene inactivation in retinoblastoma cell line Y79. Proc Natl Acad Sci USA 85:6017–6021