



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Rapid approach to identify an unrecognized viral agent

Yuan Hu^{a,*}, Irvin Hirshfield^b

^a U.S. Food and Drug Administration, Northeast Regional Laboratory, Microbiological Sciences Branch,
158-15 Liberty Avenue, Jamaica, NY 11433, USA

^b St. John's University, Jamaica, NY 11439, USA

Received 12 October 2004; received in revised form 17 February 2005; accepted 22 February 2005

Available online 12 April 2005

Abstract

For epidemic control, rapid identification and characterization of the responsible unknown agent are crucial. To address this critical question, a method was developed for virus discovery based on a flexible nested-PCR subtraction hybridization. As a positive control, we used hepatitis C virus as a hypothetical unrecognized virus and “discover” it in the sample. Using template-switching universal long-PCR to produce large quantities of cDNA, our nested-PCR-based subtractive hybridization coupled with a single-strand deletion technology removed most of the common cDNA. Following subtraction hybridization, a cDNA library was constructed and displayed by differential reverse dot blot hybridization. This new genomic subtraction hybridization method will be ideally suited to identify rapidly any previously unrecognized viral agent.

Published by Elsevier B.V.

Keywords: Unrecognized viral agents; mRNA; cDNA; Subtraction hybridization; Long-PCR; Nested-PCR; Mung bean nuclease; cDNA library

1. Introduction

Based on past history, it is not just a hypothetical risk that many people have been infected with unrecognized viruses. By the time a new virus, such as hepatitis C virus (HCV), human immunodeficiency virus (HIV) or severe acute respiratory syndrome (SARS) associated coronavirus, is found, it is too late to save the lives of many individuals even thousands. For longer-term solutions, a powerful molecular technique is needed to rapidly identify any previously unrecognized viral agent without using virus isolation. Such a method could then be applied for investigation of future outbreaks. The most recent technologies for detecting and identifying previously unrecognized pathogens are differential display, representational difference analysis, subtractive hybridization, expression library screening, broad-range polymerase chain reaction (PCR) and serial analysis of gene expression. But

they are all time-consuming and not very reproducible approaches (Allander et al., 2001).

It is now more than 10 years since the invention of differential display (Liang, 2002; Lisitsyn et al., 1993). Subtractive hybridization, a differential display-like technology, has been an approach used to identify and isolate cDNAs of differentially expressed genes in the field of cancer research (Yen, 2000; Su et al., 1997). These methodologies all have in common that one DNA population (driver) is hybridized in excess against a second population (tester) DNA to remove sequences present in both populations. In this project, we adapted the method for mRNA analysis to rapidly detect genes from an unknown virus. For this purpose, a subtraction hybridization-like method with a single-strand deletion technology was developed, and HCV was used as a hypothetical unrecognized virus in this study.

The traditional subtractive hybridization often requires large amounts of mRNA as starting material and it is commonly a challenging problem to obtain sufficient amounts of mRNA from the virus. In this project, high-quality full length

* Corresponding author. Tel.: +1 718 340 7086; fax: +1 718 340 7058.
E-mail address: yhu@ora.fda.gov (Y. Hu).

cDNAs in large quantities were obtained by using a universal long-PCR based method (Clontech, BD, USA).

Normally, subtraction hybridization generated libraries typically contain some background clones representing non-differential genes (Liang, 2002). To overcome this problem a single-strand specific mung bean endonuclease in between the 1st and 2nd PCR was used which is ideally suited to rapidly reduce the number of false positives that could arise. More specifically it is single driver strands and linear amplified tester strands that are degraded. The double stranded products of the exponential amplification of the unique tester sequence remain unaffected. After reducing the background of the 1st PCR products, a nested PCR was performed by using the nested primers to further enrich the differentially produced sequences, which are from the tester but not from the driver.

After the subtracted cDNA library was obtained, it was cloned directly into a pGEM-T cloning vector, and plasmid DNAs from 96 random selected clones were analyzed. Differential reverse dot blot hybridization was carried out to determine whether the unique subtracted tester products represented the genomic region of the hypothetical unknown virus, in this case HCV. The hybridized clones were sequenced and sequences corresponding to the hypothetical unknown virus, HCV, were recovered. The inclusion of mung bean endonuclease treatment in the protocol precludes the need for multiple rounds of subtraction.

In recent years, methods for PCR amplification of long DNA fragments have advanced significantly. With intrinsic terminal transferase activity of a reverse transcriptase (RT) (Clontech, 2002; Petalidis et al., 2003) long PCR is capable of generating a full-length cDNA library from a small quantity of RNA. Additionally, the combined use of single-strand deletion technology in subtraction hybridization may eliminate or shorten steps required by traditional methods. This method will improve the ability to respond rapidly to any infectious disease outbreak and provide added value in the event of a bioterrorist event.

2. Material and methods

2.1. Experimental design, an overview

(1) The hypothetical unknown virus (HCV) was considered as an unknown mRNA in our study. (2) Unknown virus mRNA sequences present in the specimen were considered as tester (tester = HCV infected cell line). (3) The same type of specimen obtained from the same environment was considered as driver (driver = non-infected cell line). (4) Large quantities of cDNA were produced by universal long-PCR. (5) Our method is ideally suited to subtract out the most abundant common mRNA sequences present in both populations (tester and driver). (6) The method focused on a small number of unknown virus mRNA sequences present in the specimen and then enriched it by stoichiometry and kinetics. (7) A

single-strand deletion technology limited the background from common sequences. (8) A subtraction cDNA library was constructed and displayed by differential reverse hybridization. (9) DNAs were sequenced for further characterization.

2.2. HCV infected cell lines

The human MOLT-4 acute lymphoblastic leukemia T cell line was obtained originally from the American Type Culture Collection (ATCC, Rockville, MD, USA; Lohmann et al., 1999; Wellnitz et al., 2002), and was maintained in RPMI 1640 ATCC modified medium (ATCC, Rockville, MD, USA) supplemented with 5% fetal bovine serum (ATCC, Rockville, MD, USA) in 75-cm² culture flasks at approximately 3×10^5 to 2×10^6 cells/ml. In vitro infection of the cell line was performed by using HCV RNA-positive serum from chronic HCV carriers. The cells were subcultured regularly with a split of 1:10 every week (Hu et al., 2003).

2.3. Total RNA preparation

Total cellular RNA was isolated from cell lines by a single-step extraction method with an acid guanidinium thiocyanate–phenol–chloroform mixture (Biotecx Laboratories Inc., Houston, TX, USA). The total RNA remained exclusively in the aqueous phase while proteins and DNA were extracted into an organic phase and interphase. The total RNA was precipitated from the aqueous phase by addition of an equal volume of 100% isopropanol, washed with 75% ethanol and solubilized in RNase free water. During RNA extraction, special care was taken to avoid cross-contamination by separating pre- and post-PCR work areas and using aerosol-proofed tips. After extraction, a small aliquot of the RNA product was run on a gel to determine its quality.

2.4. mRNA preparation

Once the isolation of total RNA was complete, high quality mRNA was prepared from it with a Poly (A)⁺ Tract mRNA Isolation System III (Promega, Madison, WI, USA) following the manufacturer's instructions. Briefly, 500 μ l of total RNA were heated at 65 °C for 10 min. 3 μ l of biotinylated-oligo (dT) probe and 13 μ l of 20 \times SSC were added and incubated at room temperature until completely cooled. The entire mixture was added to the tube containing the streptavidin paramagnetic particles and incubated at room temperature for 10 min. The mixture of mRNA-biotinylated oligo (dT) probe-streptavidin paramagnetic particles was captured on a magnetic stand and washed for four times with 0.1 \times SSC. mRNA was eluted from the streptavidin paramagnetic particles with 100 μ l of RNase-free water, and then used as a template for cDNA synthesis.

2.5. Template-switching cDNA amplification

The template-switching activity of moloney murine leukemia virus (MMLV) reverse transcriptase with a mod-

ified oligo(dT) primer was used to synthesize a full-length first-strand cDNA by using the SMARTTM PCR cDNA Synthesis Kit as follows (Clontech Laboratories, Palo Alto, CA, USA).

2.5.1. 1st strand cDNA synthesis

mRNA was mixed with 7 μ l of 12 μ M 3' modified oligo (dT) primer (5'-AAG CAG TGG TAT CAA CGC AGA GTA CT₃₀N₋₁N-3') and 7 μ l of 12 μ M template switching primer (5'-AAG CAG TGG TAT CAA CGC AGA GTA CGC GGG-3'). The reaction mixture was incubated at 65 °C for 2 min and the temperature was then reduced to 42 °C. The following reagents were then added, 20 μ l 5 \times 1st strand buffer (250 mM Tris-HCl pH 8.3, 375 mM KCl and 30 mM MgCl₂), 2 μ l 100 mM DTT, 10 μ l 50 \times 10 mM dNTPs, 100 units RNase inhibitor (Applied Biosystems, Foster City, CA, USA) and 5 μ l MMLV reverse transcriptase. The reaction mixture was incubated at 42 °C for 30 min in a Perkin-Elmer 9700 DNA Thermal Cycler (Applied Biosystems, Foster City, CA, USA). The reaction was stopped by adding 2 μ l of 0.5 M EDTA.

2.5.2. Column chromatography purification

In order to remove unincorporated nucleotides and small cDNA unused fragments, cDNA products were purified by using the NucleoSpin Extraction Spin Column according to the manufacturer's instructions (Clontech Laboratories, Palo Alto, CA, USA).

2.5.3. Double-stranded cDNA amplification by long-PCR

20 μ l of the 1st strand cDNA was then used as template for the second strand synthesis. The following reagents were added, 62 μ l dH₂O, 10 μ l PCR Buffer (BD Clontech), 2 μ l 10 mM dNTPs, 2 μ l PCR primer (5'-AAG CAG TGG TAT CAA CGC AGA GT-3') and 2 μ l polymerase mix, and the reaction mixture was subjected to the following cycling: 95 °C for 1 min and then 25 cycles of 95 °C for 5 s, 65 °C for 5 s and 68 °C for 6 min. The reaction was stopped by adding 2 μ l of 0.5 M EDTA and the cDNAs extracted by phenol:chloroform:isoamyl alcohol (25:24:1) mix and purified by column chromatography as above. A small aliquot of the PCR product was run on a gel to determine the quality of the full-length cDNAs synthesized and a high molecular weight smear was expected.

2.6. Produce blunt ended cDNA fragments

Because long sequences will be amplified less efficiently than shorter sequences, the tester and driver cDNAs were digested with *Rsa* I, a four-base-cutting restriction enzyme that yields blunt ends (RestrictionMapper, 2004). The purified cDNA was incubated with 10 units *Rsa* I and buffer in a total of 350 μ l at 37 °C for 3 h. The reaction was stopped by adding 8 μ l of 0.5 M EDTA. After digestion, a small aliquot of the *Rsa* I digested cDNA product was run on a gel to confirm the digestion quality.

2.7. Unphosphorylated adaptor ligation

The tester cDNA was subdivided into two portions. Each blunt ended cDNA fragment was ligated with a different unphosphorylated adaptor. (adaptor 1: 5'-CTA ATA CGA CTC ACT ATA GGG CTC GAG CGG CCG CCC GGG CAG GTA CCT GCC CGG-3', adaptor 2: 5'-CTA ATA CGA CTC ACT ATA GGG CAG CGT GGT CGC GGC CGA GGT ACC TGC GGC CG-3') (Clontech BD, Palo Alto, CA, USA) The ends of the adaptors do not have a phosphate group, so only one strand of each adaptor attaches to the 5' ends of the cDNA.

2.8. First round of hybridization

After evaluation of the ligation efficiency, an excess of driver cDNA was added to each sample of tester cDNA. The samples were then heat denatured at 98 °C for 2 min and allowed to hybridize at 68 °C for 8 h.

2.9. Second round of hybridization

During the second hybridization, the two primary hybridization products from the 1st hybridization were mixed together without denaturing. Fresh denatured driver cDNA was added to the above mix. The second hybridization was performed at 68 °C for 16 h.

2.10. Fill in the ends

0.5 μ l of DNA polymerase, 2.5 μ l of 10 \times PCR buffer and 0.5 μ l of 10 μ M dNTP were added to the 2nd hybridization product to fill in the ends of each molecule to create a double-stranded primer sequence annealing site for PCR. Only the molecules that are unique to the tester amplicon and reanneal with each other will form double-stranded molecules with PCR priming sequences on each end that are able to amplify exponentially.

2.11. Selective PCR amplification

The entire population of molecules was then subjected to PCR to amplify the desired differentially expressed sequences in a total 25 μ l reaction mixture with PCR buffer, dNTP, polymerase as described above and only one PCR primer. (PCR primer: 5'-CTA ATA CGA CTC CAT ATA GGG C-3'') (Clontech BD, Palo Alto, CA, USA). The mixture was subjected to 27 cycles of PCR in the DNA Thermal Cycler: denaturation at 94 °C for 10 s, annealing at 66 °C for 30 s and extension at 72 °C for 1.5 min.

2.12. Mung bean nuclease treatment

Mung bean nuclease is well known to cleave DNA at regions of single-strandedness (Kuvbachieva and Goffinet, 2002). In order to eliminate potential false positive results, before doing the nested PCR, the 1st PCR products were treated

by mung bean nuclease (Promega, Madison, WI, USA) with 20 μ l of 10 \times reaction buffer (150 mM sodium acetate pH5.0, 250 mM ZnCl₂, 25% v/v glycerol) and 10 units of mung bean nuclease at 37 °C for 30 min in a total of 100 μ l. After treatment with mung bean nuclease, the mix was extracted once with phenol, twice with chloroform and ethanol-precipitated with NH₄OAc. The pellet was dissolved in 100 μ l 10 mM Tris, pH 8.0 containing 0.1 mM EDTA, pH 8.0 (TE buffer).

2.13. Nested PCR

In order to further reduce any background PCR products and to further and even more enrich specifically for unique tester cDNA, a secondary PCR amplification was performed within a total 25 μ l reaction with 10 μ l of mung bean nuclease treated 1st PCR product, 2.5 μ l of 10 \times PCR buffer, 0.5 μ l of 10 mM dNTP, 0.5 μ l of 50 \times polymerase mix and two nested-PCR primers. (Nested-PCR primer 1: 5'-TCG AGC GGC CGC CCG GGC AGG T-3'' and nested-PCR primer 2: 5'-AGC GTG GTC GCG GCC GAG GT-3') (Clontech BD, Palo Alto, CA, USA). The mixture was subjected to 10 cycles of PCR in the 9700 DNA Thermal Cycler, denaturing at 94 °C for 10 s, annealing at 68 °C for 30 s and extension at 72 °C for 1.5 min.

2.14. A one step strategy for constructing a subtracted cDNA library

After the nested-PCR amplification, subtracted cDNAs were inserted directly into a T/A cloning vector (Promega, Madison, WI, USA). Molecular cloning was performed by ligating subtracted cDNA and pGEM-T cloning plasmid vector with T4 DNA ligase and ATP at 4 °C overnight. The chimera was introduced into JM109 high efficiency competent cells by transformation with heat shock at 42 °C. The cDNA insert was detected by its interruption of the coding sequence of β -galactosidase by blue/white color screening on LB/Ampicillin/IPTG/X-Gal indicator plates, and 96 positive clones were randomly selected for further work.

2.15. Generation of reverse hybridization probes

Multiple digoxigenin labeled probes were prepared from tester and driver double stranded cDNA by long-PCR with digoxigenin-11-uridine-5'-triphosphate (Roche, Indianapolis, IN, USA). The long-PCR reaction and *Rsa* I digestion were performed as described above.

2.16. Analysis of subtracted library by differential reverse dot blot hybridization

DNA from the 96 randomly selected clones from the subtracted library were prepared by using Minipreps DNA Purification System (Promega, Madison, WI, USA) and blotted on a nylon membrane (Intergen, Purchase, NJ, USA). Potentially interesting sequences were detected by solid phase reverse

hybridization with digoxigenin labeled probes. The hybridization with DIG Easy Hyb solution (Roche, Indianapolis, IN, USA) was performed at a non-stringent temperature of 42 °C, which ensures the probe would bind to any clone containing a similar sequence, but a high-stringency wash was used at 58 °C to decrease the hybridization background. Duplicate membranes were prepared and hybridized. Clones corresponding to unknown viral cDNA (HCV) hybridized only with the tester probes, and not with the driver probes.

2.17. Sequence identification

The DNAs from selected clones were subjected to a cycle sequencing reaction with 3.2 pmol of an unlabeled T7 promoter primer (5'-TAATACGACTCACTATAGGG-3') (Qiagen Operon, Alameda, CA, USA) and a terminator reaction mixture according to the manufacturer's instructions (Applied Biosystems, Foster City, CA, USA). The cycle sequencing reaction consisted of 25 cycles in the 9700 DNA Thermal Cycler, and each cycle consisted of denaturation at 96 °C for 10 s, annealing at 50 °C for 5 s and extension at 60 °C for 4 min. After 25 cycles, the fluorescent extension products were purified by a simple isopropanol precipitation step. DNA sequence data were analyzed by the Chromas program (Helensvale, Queensland 4212, Australia), the GenBank sequence database and the Blast program from the National Center for Biotechnology Information (NCBI).

2.18. Control system

It was very important to have appropriate controls in this study in order to validate the procedure. A human placenta RNA control (Clontech BD, Palo Alto, CA, USA) was used in parallel with tester and driver RNA. Also, all the PCR reactions were always carried on the same PCR machine because optimal parameters may vary with different thermal cyclers. All the agarose gels were SYBR Green I stained (Molecular Probes, Eugene, OR, USA).

3. Results

3.1. cDNA analysis

The research started with poly (A)⁺ RNA from HCV infected and non-infected cell lines. The traditional reverse transcriptase cannot always transcribe the entire mRNA sequence. By using RNase H⁻ point mutant MMLV reverse transcriptase, the enzyme can reach the 5' end of the mRNA and specially bind an oligonucleotide to the 5' end of the 7-methylguanosine cap structure allowing synthesis of the full-length cDNA (Clontech Laboratories, Palo Alto, CA, USA). A large quantity of full length cDNA was produced by long PCR. Our results confirmed that long target full length cDNA was end-to-end amplifiable directly from the 1st cDNA strand. The cDNA was then digested with *Rsa*I which created blunt ended cDNA fragments.

3.2. Subtraction hybridization analysis

The kinetic theory behind subtraction hybridization and the hybridization association rate are very important in this study. Basically, subtraction hybridization is a subtractive DNA enrichment technique. These methodologies all have in common that one DNA population (the driver) is hybridized in excess against a second population (the tester) to remove sequences present in both populations, thereby enriching target sequences unique to the tester. In theory, under standard conditions, subtraction hybridization has second order kinetics of hybridization ($C/C_0 = 1/1 + KC_0t$) (where K is the second-order hybridization rate constant, C is the surface concentration of cDNA, and C_0 is the initial concentration of cDNA in solution) (Miner et al., 1995; Towery et al., 2001; Wang and Drlica, 2003). During the hybridization reaction, re-annealing is faster for the more abundant molecules. Concentrations of high (driver cDNA) and low (tester cDNA) become roughly equal, and based on this theory, tester cDNA is significantly enriched.

In our study, two different adaptors were ligated to tester cDNA molecules. After the first hybridization, the reaction significantly enriched unique sequences found in the tester cDNA due to the second order kinetics of hybridization as described above. After the second hybridization, the reaction generated three major molecular types. (1) cDNA without adaptor to cDNA without adaptor: this would be the most probable combination of re-annealing due to the excess of unligated driver cDNA. This cDNA would not be amplified during PCR due to the absence of ligated adaptor. (2) cDNA with adaptor from tester to cDNA without adaptor from driver: this would be the second most probable combination due to the excess of cDNA without adaptor. The ligated cDNA would amplify during PCR but its product would not be amplifiable. Thus there would be a linear amplification (3) unique cDNA with adaptor 1 will only re-anneal to unique cDNA with adaptor 2 from tester. This would be the least likely combination, occurring only with unique sequences found in the tester cDNA. This cDNA would be amplified exponentially by PCR as a primer binding site is present on both strands and all newly amplified strands also contain primer sequences. The target cDNA can be enriched by repeated cycles of PCR and nested-PCR.

3.3. Selective amplification analysis

Only the tester/tester duplexes which have two different adaptors with primer sequences on both ends were amplified exponentially. These are the equalized, differentially expressed sequences. But during PCR, other molecules were missing primer annealing sites, and thus could not be amplified as discussed above.

3.4. Mung bean nuclease treatment analysis

In order to further eliminate a false positive background, the traditional subtraction hybridization protocol was

modified. The 1st PCR product was treated by a specific mung bean nuclease before nested-PCR amplification. The driver strands, linear amplified tester strands and other single-strands were degraded, but the selective amplified products remained unaffected. As a control, an experiment was carried out without mung bean nuclease treatment. No significant positive clone was found in the first round of subtraction hybridization.

3.5. Subtracted cDNA library analysis

After the transformation of competent *E.coli* cells and blue–white selection on IPTG plates for transformed bacteria, 96 white colonies were individually randomly picked on ampicillin plates. The chimeric DNAs were directly extracted by using the Minipreps DNA Purification System without traditional amplification and deletion.

3.6. Clone analysis

In the present study, 96 selected cDNA clones were used for the experiments to detect all potentially interesting sequences by solid phase reverse hybridization. Clones considered as corresponding to an unknown viral cDNA were those that hybridized only with the tester probe, e.g. clone A5 and C12 (Fig. 1). Those clones that hybridized to both tester and driver probes were excluded. Duplicate membranes were prepared for hybridization and similar results were obtained.

3.7. Evaluation of subtractive efficiency

In the standard protocols for subtractive hybridization, it is necessary to have multiple rounds to enrich the differential sequences. In our study, using long-PCR to produce a large amount of cDNA in conjunction with mung bean nuclease treatment to decrease the background, it was possible to detect the unknown virus cDNA with one round of subtractive hybridization.

3.8. DNA sequence analysis

After reverse dot blot hybridization, potentially positive clones were directly sequenced by using ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kits version 2 and the ABI Prism 310 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's instructions.

3.9. Data analysis

Clones containing tester specific fragments were recovered in our new subtraction hybridization. In the subtracted library, 16 unique tester specific fragments were detected from 96 randomly selected clones. Sequence analysis of the 16 selected clones showed that 25% of them corresponded to known HCV genes. Thus of the 96 random selected clones only four inserts could be ascribed to the HCV genome. The

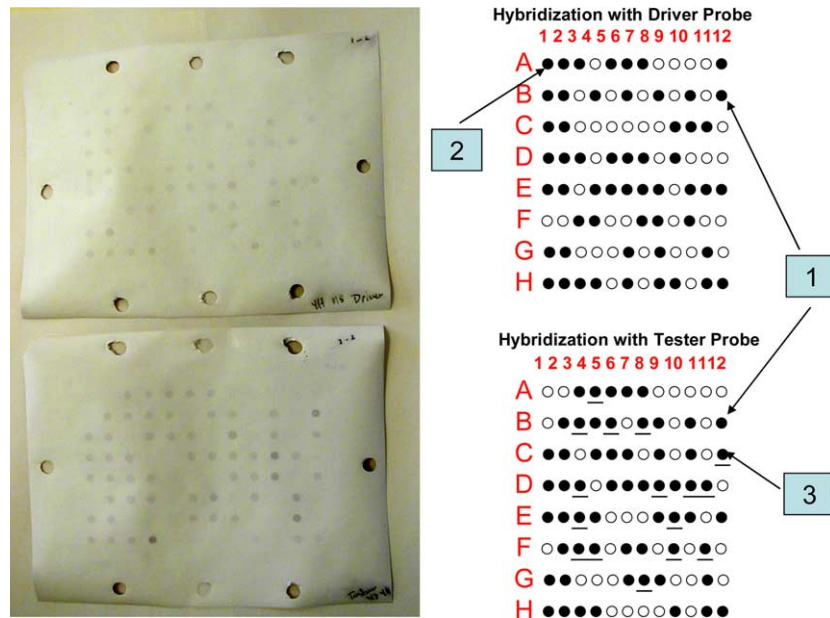


Fig. 1. Subtraction hybridization results. (1) Corresponds to cDNAs from both the tester and driver samples, (2) corresponds to cDNAs from the driver samples, (3) corresponds to cDNAs from the tester samples.

remaining clones demonstrated no significant matches with entries in the databases. They may be non-specific, chimeric fragments and incompletely digested fragments.

4. Discussion

Many human diseases of unknown etiology may be caused by unrecognized viruses, and their detection and characterization remains a major challenge for virologists and clinicians. For investigation of an outbreak caused by an unknown virus, its rapid identification and characterization must be a high priority. Once the relationship between a potential new infectious agent and a known virus is established by sequencing, appropriate culture conditions, serologic tests, nucleic acid detection and even therapeutic strategies can be rapidly developed.

Molecular analysis of unknown viruses has proceeded slowly because, to date, a highly reproducible test model has not been developed. Most of the current methods are labor-intensive and cumbersome. A problem is that there is an extremely small quantity of the viral mRNA present in comparison to the host's mRNA. Although traditional subtractive hybridization methods have been successful in some cases, they require a large quantity of mRNA as starting material and several rounds of hybridization (Kellam and Weiss, 2001) which are not well suited for the identification of unrecognized and uncharacterized viral agents. Obtaining a full-length cDNA is one of the most important, and often one of the most difficult steps in subtraction hybridization. To prepare full-length cDNA libraries for rapid discovery of a new virus still remains a challenge in the field. In this study, we used successfully the template-switching activity

of MMLV reverse transcriptase with a modified oligo (dT) primer to synthesize a full-length first-strand cDNA as starting material. In order to prepare a large quantity of cDNA, we successfully used long-PCR to generate sufficient quantities of both tester and driver cDNA samples before subtraction hybridization.

This method removed most of the common cDNA by nested-PCR-based subtractive hybridization and the single-strand deletion technology reduced the number of false positives. These steps overcome the technical limitations of traditional subtraction methods. We also showed that the long-PCR cDNA products can be used directly as hybridization probes for screening the subtracted cDNA library.

In conclusion, a subtraction hybridization procedure using long PCR full-length cDNA preparation and single-strand deletion technology was successfully developed in this project. Using an HCV infected cell line as tester and an uninfected cell line as driver, several HCV cDNA fragments were recovered. The results indicate that our subtractive hybridization method is a high sensitivity, low background, rapid, simple and reliable protocol. For emerging diseases, such as SARS, and defense against bioterrorist attacks, rapid identification and characterization of the responsible agent are crucial first steps. This efficient technique represents a valuable tool for the discovery of unrecognized and uncharacterized viral agents present in low copy number and highly complex backgrounds.

Acknowledgment

We thank Dr. Tao Hong (Hackensack University Medical Center, NJ, USA) for critically reading the manuscript.

References

- Allander, T.S., Emerson, S., Engle, R., Purcell, R., Bukh, J., 2001. A virus discovery method incorporating DNase treatment and its application to the identification of two bovine parvovirus species. *Proc. Natl. Acad. Sci. U.S.A.* 98 (20), 11609–11614.
- Clontech, 2002. PCR-Select cDNA Subtraction Kit and Super SMART PCR cDNA Synthesis Kit Technical Manual. <http://www.clontech.com/archive/JAN02UPD/SmartPCR.shtml>.
- Hu, Y., Shahidi, A., Park, S., Guilfoyle, D., Hirshfield, I., 2003. Detection of extrahepatic HCV replication by a novel highly sensitive single tube nested-PCR. *Am. J. Clin. Pathol.* 119 (1), 95–100.
- Kellam, P., Weiss, R.A., 2001. How technology drives pathogen discovery. *Microbiol. Today* 29, 16–18.
- Kuvbachieva, A., Goffinet, A., 2002. A modification of representational difference analysis, with application to the cloning of a candidate in the reelin signaling pathway. *BMC Mol. Biol.* 3 (6), 1–5.
- Liang, P., 2002. A decade of differential display. *Bio. Techniques* 33 (2), 338–346.
- Lisitsyn, N., Lisitsyn, N., Wigler, M., 1993. Cloning the differences between two complex genomes. *Science* 259, 946–951.
- Lohmann, V., Korner, F., Koch, J., Herian, U., Theilmann, L., Bartenschlager, R., 1999. Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science* 285, 110–113.
- Miner, J., Cecchini, E., Dominy, P.J., 1995. A kinetic model for subtractive hybridization. *Nucl. Acid Res.* 23, 176–187.
- Petalidis, L., Bhattacharyya, S., Morris, G.A., Collins, V.P., Freeman, T.C., Lyons, P.A., 2003. Global amplification of mRNA by template-switching PCR: linearity and application to microarray analysis. *Nucl. Acids Res.* 31 (22), 142–149.
- RestrictionMapper, 2004. Maps sites for restriction enzymes (restriction endonucleases) in DNA sequences. <http://www.restrictionmapper.org/>.
- Su, Z., Shi, Y., Fisher, P., 1997. Subtraction hybridization identifies a transformation progression-associated gene PEG-3 with sequence homology to a growth arrest and DNA damage-inducible gene. *Proc. Natl. Sci.* 94, 9125–9130.
- Towery, R.B., Fawcett, N.C., Zhang, P., Evans, J.A., 2001. Genomic DNA hybridizes with the same rate constant on the QCM biosensor as in homogeneous solution. *Biosens. Bioelectron.* 16, 1–8.
- Wang, J.Y., Drlica, K., 2003. Modeling hybridization kinetics. *Math. Biosci.* 183, 37–47.
- Wellnitz, S., Klumpp, B., Barth, H., Susumu, I., Depla, E., Dubuisson, J., Baumert, T., 2002. Binding of Hepatitis C virus-like particles derived from infectious clone H77C to defined human cell lines. *J. Virol.* 76 (3), 1181–1193.
- Yen, V., 2000. Study of genomic alterations in human breast tumors by RDA. *MIT Biol. Undergrad. (BUG) J.* 3, 77–83.