CHAPTER 6

PRIMER EXTENSION ENRICHMENT REACTION (PEER) AND OTHER METHODS FOR DIFFERENCE SCREENING

LILIA M. GANOVA-RAEVA

Centers for Disease Control and Prevention, Division of Viral Hepatitis, 1600 Clifton Rd. NE, MS A-33, Atlanta, Georgia 30329, USA
E-mail: lkg7@cdc.gov

Abstract: Our knowledge on the subject of genetic divergence is ever expanding and encapsulates a wide spectrum of research areas from the study of single nucleotide polymorphism to examination of the complex differences of host–pathogen interactions. The understanding of genetic differences is essential to our ability to address adequately human health issues caused by mono allelic genetic disorders, altered gene expression in cancers, development of drug resistance and the variety of ways organisms respond to infections and the environment. Large number of hybridization-based applications has been developed to query and find such differences. This review introduces a new method for genetic difference screening — the primer extension enrichment reaction (PEER) — presented in the context of similar subtraction-based hybridization methods. PEER is a novel approach to difference screening and is not intended to replace the existing elegant hybridization methods but to expand their scope. PEER is tailored to find unknown targets present in very low copy numbers and in the context of an imperfect genomic match. The PEER method takes advantage of the greater hybridization specificity of shorter oligonucleotides coupled with enzymatic extension specificity.

Keywords: Primary structure, genetic differences, gene expression, sequencing-by-hybridization, primer extension enrichment reaction (PEER), class II restriction endonuclease, class IIS restriction endonuclease, IIS enzyme, spot hybridization, enrichment, enrichment value, cost efficient, subtractive cloning, arbitrary primed PCR, amplification fragment length polymorphism (AFLP), serial analysis of gene expression (SAGE), genomic signature tag (GST), serial analysis of binding elements (SABE), differential analysis of restriction fragments amplification (DARFA), ligation mediated enrichment, selectively primed adaptive driver RDA, enzymatic degrading subtraction (EDS), phenol, linker capture subtraction (LCS), differential subtraction chain (DSC), cloning of deleted sequences (CODE), negative subtraction chain (NSC), SABRE, DNA enrichment by allele-specific hybridization (DEASH).

A. Buzdin and S. Lukyanov (eds.), Nucleic Acids Hybridization, 125–165.
© 2007 *Springer.*

Abbreviations: AFLP, amplification fragment length polymorphism; cDNA, complementary DNA; CODE, cloning of deleted sequences; DARFA, differential analysis of restriction fragments amplification; DD, differential display; DEASH, DNA enrichment by allele-specific hybridization; dNTP, deoxyribonucleotidetriphosphate; DSC, differential subtraction chain; ds, double stranded; EDS, enzymatic degrading subtraction; GST, genomic signature tag; LCS, linker capture subtraction; MDR, mispaired DNA rejection; mRNA, messenger RNA; NA, nucleic acid; NSC, negative subtraction chain; PCR, polymerase chain reaction; PEER, primer extension enrichment reaction; RaSH, Rapid subtraction hybridization; RDA, representational differences analysis; RFLP, restriction fragment length polymorphism; RT, reverse transcription; SABE, serial analysis of binding elements; SABRE, selective amplification via biotin- and restriction-mediated enrichment; SAGE, serial analysis of gene expression; SH, subtractive hybridization; SNP, single nucleotide polymorphism; SPAD, Selectively primed adaptive driver; ss, single stranded; SSH, suppression subtractive hybridization; UDG, uracil deglycosilase.

TABLE OF CONTENTS

1. INTRODUCTION

Identifying and isolating genetic differences without *a priori* knowledge of the primary structure, i.e. sequence, of the genetic material, is technically demanding and laborious. Genetic differences, which can be found at any level of genetic organization or gene expression, can be caused by gene rearrangements, deletions, insertions, or by the presence of genomes from extraneous organisms and can lead to disparate disease outcomes. They can be represented by a single nucleotide polymorphism (SNP) (e.g. hepatitis B virus vaccine escape mutants), single allele differences (e.g. cystic fibrosis, sickle cell anemia, Tay Sachs disease and over 4000 other genetic diseases), gene expression differences in the various cells of one organ that occur in order to execute the specific function of the tissue or organ, gene expression at different stages of an organism's development that provide for cell differentiation related to growth and development, genetic differences between organisms, sex-related differences, complex genetic disorders (e.g. Alzheimer, diabetes), and gene expression prompted by the interaction of different organisms, such as pathogen–host interactions, or in response to different environmental stimuli.

Most genetic differences can be queried by using RNA, DNA, cDNA, or total nucleic acids (NAs), depending on the experimental goal. Direct screening of large plasmid or phage libraries, the earliest approach used for identifying target NAs of unknown sequence, can be inefficient and labor-intensive [1]. Over the last two decades, several new approaches have been developed to improve the efficiency of this task. They can be divided into two broad categories: (1) subtractive approaches, such as representational differences analysis (RDA) [2] and its variants [3,4], differential subtraction chain (DSC) [5], selective amplification via biotin- and restriction-mediated enrichment (SABRE) [6], suppression subtractive hybridization (SSH) [7], differential display (DD) [8], and others [9,10]; and (2) high-throughput [11,12,13,14,15,16,17] and microarray-based [18,19] methods.

All subtractive approaches are based on molecular comparison of two specimens: (1) a "tester" – a specimen that is presumed to contain the unknown target of interest, and (2) a "driver" – a specimen that is a perfect genetic match for the tester but is not believed to contain the target [20]. The comparison is usually done by hybridizing the two specimens to each other, eliminating the hybrids they have in common and screening the final product for a subset of molecules that reflect the differences between the two. Subtractive methods are often used in molecular studies because of their relative simplicity and high efficiency.

High-throughput methods approach gene expression *in toto* and are based on miniaturization, artificial nucleic acid synthesis, and hybridization and often require the heavy use of bioinformatics for data management. In recent years, the integration of subtractive approaches into high-throughput methods [21,22,23,24] has generated efficient methodology for identifying differentially expressed sequences.

Among the subtractive techniques, RDA and the closely related SSH are the most popular and have been used successfully to recover unknown sequences and

differentially expressed genes. SSH has been used to find a number of differentially expressed RNA messages [25,26,27,28,29,23] and to identify a new calicivirus in walrus [30]. The GBV–A and –B viruses [31] as well as TT virus (TTV) [32] (all suspect viral hepatitis agents) were discovered by RDA. In practice, SSH can enrich a target gene by approximately 3×10^3 times [7,1]. Although impressive, this performance is nonetheless insufficient for the detection of an infectious agent whose genomes might be present at only a few copies in the specimen of interest [33], for example, in chronic carriers the levels of hepatitis C virus might be as low as 10^3 or less, yet the disease persists and the individual is still a potential source of infection. Such levels are below the molecular resolution of the current enrichment methods. A common limitation of all subtractive approaches is their requirement for perfect hybridization, which goes hand in hand with the need for an abundance of ideally matched drivers. Since such conditions can rarely be achieved, most subtraction methods are intrinsically biased against single-stranded, low-copy-number molecular species [31,33]. The enrichment of the target of interest is usually achieved by hybridization between long DNA fragments and, in the case of cDNA generated by random priming, a heterogeneous population of DNA fragments. If present in low numbers, these molecules have little chance to form complete hybridization products after denaturing. Another limitation of most subtraction methods is that they use the target of interest as a potential PCR template only. If present at very few copies or if it has remained single stranded, the target might fail to amplify efficiently. In addition, many subtraction approaches rely on the presence of a poly-A tail to generate starting material and consequently are not suitable when working with DNA or RNA that is not polyadenylated.

Microarray approaches tend to be costly and require some prior knowledge of the sequence of interest. The more recent sequencing-by-hybridization approaches and massive parallel sequencing methods [13,14,16], which have been proposed as alternative solutions to uncovering differences, are also expensive and might not be suitable for small sample volumes.

The primer extension enrichment reaction (PEER) method belongs to the category of subtraction techniques. It is based on two new strategies: (1) the use of tester DNA to generate both PCR primer and template, and (2) the selective inactivation of primers containing sequences common to the tester and driver to ensure preferential amplification of templates that contain sequences unique to the tester. PEER improves the sensitivity of current subtraction methods and takes direct advantage of the unknown target's unique specificity.

2. PRIMER EXTENSION ENRICHMENT REACTION

2.1 Method Outline

A general outline of PEER is presented in Figure 1. Total NA is extracted from a tester and a driver specimen and used in a modified SMART cDNA protocol (Clontech, Palo Alto, CA) to generate dsDNA with two different sets of

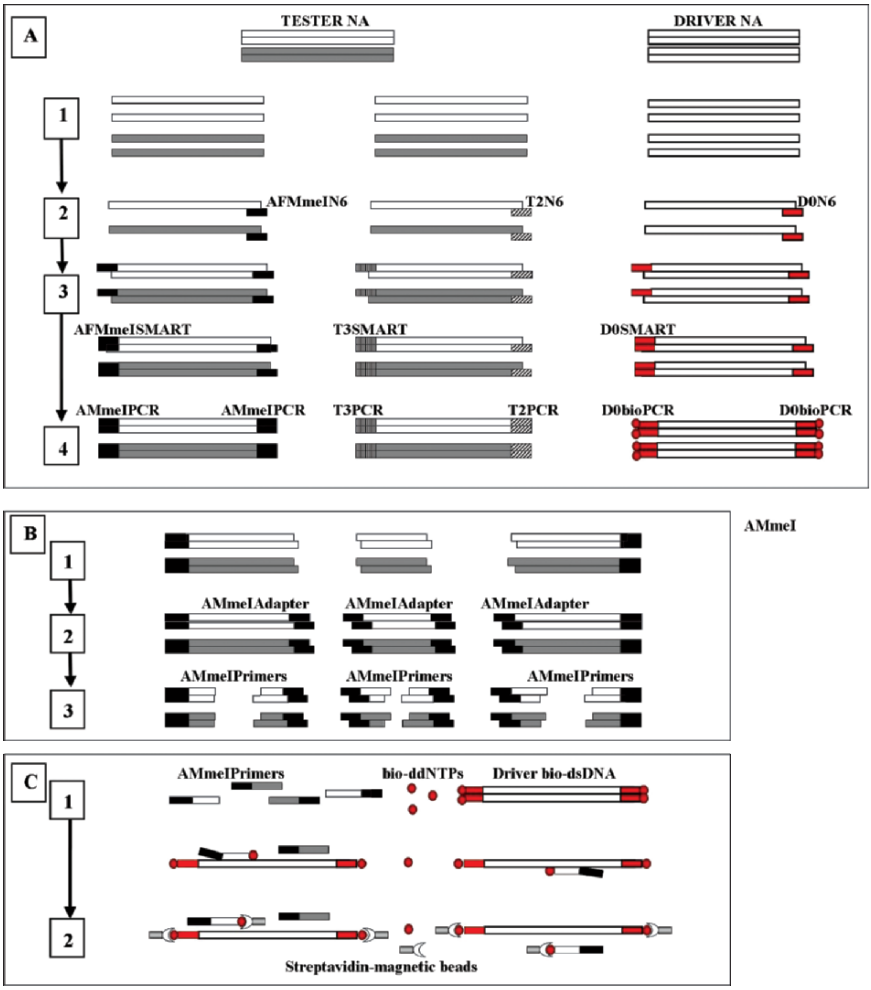

Figure 1. Primer extension enrichment reaction (PEER). Panel A: Generation of dsDNA from total nucleic acid (NA). 1. Tester NA (white and gray rectangle) is split in two aliquots and denatured; driver NA (white rectangle) is denatured as well; 2. Single strands are reverse transcribed (RT) by SuperScript RT with three different primers – AFMmeIN6* for the first Tester aliquot, T2N6 (diagonal fill rectangle) for the second aliquot, and D0N6 (shadowed rectangle) for the driver; 3. Reverse transcription switches templates and copies annealed SMART primers (SMART technology, Clontech); 4. RT products are amplified with Advantage2 Polymerase to yield Tester1 dsDNA with primers AMmeIPCR (black rectangle), Tester23 dsDNA with T3PCR (vertical fill rectangle) and T2PCR (diagonal fill rectangle), and driver bio-dsDNA with D0bioPCR biotinylated at the 5′-end (shadowed rectangle with circle) Panel B: Processing of Tester1 dsDNA. 1. DNA is cleaved by a cocktail of restriction enzymes that leave 3′-GC protruding ends; 2. Ends are treated with the Klenow fragment of DNA Polymerase I in the presence of dCTP only and then ligated to AMmeIAdapter; 3. Tagged fragments are cut to uniform size by MmeI to create multiple AMmeIPrimers. Panel C: Blocking reaction. 1. AMmeIPrimers generated from Tester1 dsDNA are extended on Driver bio-dsDNA template in the presence of biotinylated ddNTPs (red circles) and ThermoSequenase; 2. Biotinylated molecules are captured with streptavidin-coated magnetic beads (white crescent with gray bar) and removed from the reaction.

*black rectangles – primers AFMmeIN6, AFMmeISMART, AMmeIPCR, AMmeIAdapter

(*Continued*)

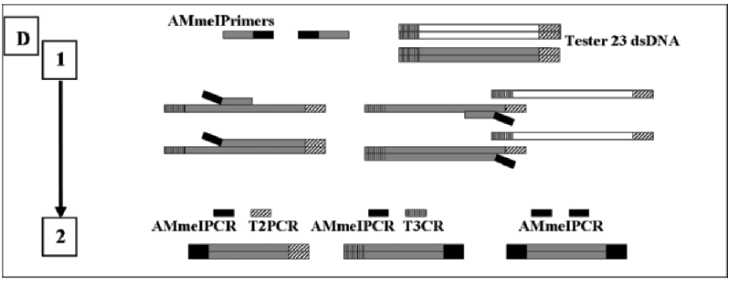

Figure 1. cont'd. Panel D: Retrieval of targets of interest from the Tester23 dsDNA. 1. Capture PCR – AMmeIPrimers that were not blocked and removed in the preceding steps are added to Tester23 dsDNA and in the presence of regular dNTP are annealed and extended to capture the targets of interest; 2. Regular PCR amplification of the capture products with different primer combinations.

primers for the tester (called Tester 1 cDNA and Tester 2 cDNA primers in Table 1) and one set for the driver (Driver cDNA primers in Table 1). The product is referred to as dsDNA to distinguish it from cDNA since it is generated from total NA instead of from RNA alone (Figure 1A). Tester 1 dsDNA material is converted into small fragments by extensive endonuclease cleavage and then tagged by ligation to a specially designed adapter. The 3′ end of the adapter incorporates a recognition site for a class IIS restriction endonuclease [34, 35]. After ligation, the fragments are cleaved with the appropriate IIS enzyme to create oligonucleotides with unique sequence at the 3′-end derived from the tester and a 5′-end derived from the adapter (Figure 1B). These adapter-tagged oligonucleotides are annealed to the driver dsDNA template and extended in the presence of biotinylated ddNTPs. All oligonucleotides that prime a reaction from the driver template can acquire biotinylated ddNTP. This event blocks any further extension and allows the removal of the biotinylated molecules from the reaction by use of streptavidin-coated magnetic beads. Primers that share driver sequences are blocked and removed leaving only primers with unique sequences that can only be found in the tester (Figure 1C). In the presence of Tester23 dsDNA and dNTPs, these oligonucleotides can prime an extension reaction from the fragments unique to the tester (target capture). This step converts the tagged primers into DNA templates suitable for PCR amplification by oligonucleotides containing only the adapter sequences or in combination with T2PCR or T3PCR oligonucleotides. The last step in PEER is a standard PCR amplification with primers containing only adapter and T2PCR/T3PCR sequences that can be used without any molarity restrictions. The final step is expected to generate collection of fragments of different size (Figure 1D).

Table 1. Primers used in the PEER study

Primer sequence 5′–3′	Name	Function
MmeI experiments		
AATGCAGACACAGAAGGTCCAT CCGAC	AFMmeI	TESTER MmeI adapter forward
P-GGTCGGATGGACCTTCTGTGT CTGC	ARMmeIP[1]	TESTER MmeI adapter reverse
GCTGCAGACACAGAAGGTCCATC CGACNNNNNN	AFMmeIN6	TESTER 1 cDNA
GCTGCAGACACAGAAGGTCCATC CGACGGG	AFMmeISMART	TESTER 1 cDNA
CAGACACAGAAGGTCCATCCGAC	AMmeIPCR	TESTER 1 cDNA PCR
ACACTAGAGCATGCGTCAAGAG AANNNNNN	T2N6	TESTER 23 cDNA
ACACTCCAGGAGGTCAGAAACAAC GGG	T3SMART	TESTER 23 cDNA
ACACTAGAGCATGCGTCAAGAGAA	T2PCR	TESTER 23 cDNA PCR
ACACTCCAGGAGGTCAGAAACAAC	T3PCR	TESTER 23 cDNA PCR
AAGCAGTGGTATCAACGCAGAGTA NNNNNN	D0N6	DRIVER cDNA
AAGCAGTGGTATCAACGCAGAGTA CGCGGG	D0SMART[1]	DRIVER cDNA
Bio-AAGCAGTGGTATCAACGCAGA GTA	D0bioPCR	DRIVER cDNA PCR
*Bpm*I experiments		
ACACTCGAGGAGGTCTGGAGIIIIIII	PEER1*Bpm*N6	TESTER 1 cDNA
ACACTCGAGGAGGTCTGGAGGG	PEER1*Bpm*G	TESTER 1 cDNA
AACACTCGAGGAGGTCTGGAG	PEER1*Bpm*AF	TESTER *Bpm*I adapter forward
CTCCAGACCTCCTCGAGTGTG	PEER1*Bpm*AR	TESTER *Bpm*I adapter reverse
GAGCTGTGGTGAGTTGGTTGG AAIIIIIII	PEERT7N7	TESTER 78 cDNA
AAGCAGAGGCAGCATTGGAGGG	PEERT8G	TESTER 78 cDNA
AGCTGTGGTGAGTTGGTTGG	PEERT7	TESTER 78 cDNA RCR
AGCAGAGGCAGCATTGGAGG	PEERT8	TESTER 78 cDNA RCR
AAGCAGTGGTATCAACGCAGAG TAIIIIIII	D0N6	DRIVER cDNA
AAGCAGTGGTATCAACGCAGAG TACGCGGG	D0SMART[2]	DRIVER cDNA
AAGCAGTGGTATCAACGCAGAGTA	D0PCRbio	DRIVER cDNA PCR
Control primers		
AATGCAGACACAGAAGGTCCAT CCGAC*TAATACGACTCACTATAGGG*	A*T7*[3]	PEER control primer
AATGCAGACACAGAAGGTCCATCCG AC*GAAACAGCTATGACCATGAT*	A*SK*[3]	PEER control primer

I = 5-nitro indol; N = random base

[1]P indicates that the oligo was phosphorylated to improve ligation

[2]According to the SMART cDNA technology (Clontech, Palo Alto, CA)

[3]These primers are not part of PEER but were used to monitor the success of the protocols' steps using a "control" template

2.2 Discussion

The proof of the PEER concept was tested in preliminary experiments. PEER is intended to find unknown targets at unknown and potentially very low concentrations. This goal was challenged in series of experiments designed to identify the minimum amount of a target DNA present in the background of a complex mixture that could be found and "captured" using oligonucleotides that match the target sequence but were used in very low concentrations. If we could create unique primers from the unknown target DNA itself, we would have the ideally matched oligonucleotides for its amplification. However, their molarity may not exceed that of the template. Since such unfavorable primer concentrations would not promote a PCR reaction, we needed to be able to tag them with a different 5′-sequence that we knew of and that could later be supplied in excess. We synthesized oligonucleotides so that the 20 nucleotides at the 3′-terminal match the template and the remaining 5′-nucleotides cannot be found in the template; these were named "capture primers". We then conducted experiments to determine whether the target template could be amplified after "capture" of the target by only the mismatched portion of the capture oligonucleotides. The results indicated that adapter-primed reactions (i.e. PCR with primers whose sequences did not exist in the original template) yielded amplification products from as little as 0.063 amol of template (1360 copies/ml) and with as little as 4 fmol of capture primers.

To test whether a large number of primers could be successfully, specifically, and completely blocked by di-deoxytermination, we tested a variety of polymerases and a range of nucleotide concentrations using an artificial template (pB6, a fragment of WCV cloned in pTAdvantage vector) and 50 pmol each of SK and T7 generic primers. After multiple PCR rounds of extension and blocking in the presence of ddNTPs, an aliquot of the product was used in regular PCR in the presence of a fresh master mix containing conventional Taq Polymerase (Roche) and regular dNTPs (Roche). The best results, as measured by the absence of product in the reactions to which ddNTPs were added initially, were achieved with Thermo Sequenase [36]. We also observed blocking by Vent (exo-) polymerase and Taq polymerase with a ddNTP: dNTP ratio of 10:1, but Thermo Sequenase remained the enzyme of choice because it generated consistent results under all experimental conditions.

The new PEER method exploits unique target sequences by creating primers from the double-stranded material of interest and then using an intact aliquot of the material as a template for amplification. The first strand of double-stranded DNA (dsDNA) is created using total NA as a template for reverse transcription with primers that have a random hexamer at the 3′ end and the appropriate adapter sequences at the 5′ end, so that these can be used in a subsequent PCR step. The use of an enzyme such as SuperScript II Reverse Transcriptase, a derivative of M-MLV with DNA polymerase activity [37], ensures that single-stranded RNA (ssRNA), DNA, or RNA: DNA hybrids will be copied into cDNA and

enter the enrichment process. Once priming sites are generated on both ends of the fragments the product can be exponentially amplified by SMART PCR to generate dsDNA. This approach maintains the correct representation [38] of the NAs entering the protocol and, via the PCR step, supplies a renewable source of the target material. We found that SuperScript II Reverse Transcriptase generated good results for both templates tested, one from a DNA virus (hepatitis B virus – HBV) and one from an RNA virus (WCV). In both cases, we had previous knowledge of the viral titer, and the process of cDNA generation that includes a Smart PCR step did not alter it. The initial primer design (Table 1) included 5-nitro-indol instead of random bases at the 3′-end of the RT primers. The random bases were eventually favored because they provided a higher efficiency of the PCR step; additional experiments, not discussed here, showed that the 5-nitro-indol's higher affinity to itself hinders the reaction performance. The PEER protocol can be modified for use with other pre-dsDNA/cDNA procedures. DNase/RNase treatment, filtration, ultracentrifugation, gradient separation, and other procedures, may be incorporated, depending on the application.

Once generated, the double-stranded material is converted into unique primers by extensive endonuclease cleavage (Figure 1B). This strategy ensures that the primers perfectly match the unknown template. To digest the dsDNA into small fragments (i.e. create multiple "primers" from the unknown target), we tested two approaches. First, we used the unique 2.5-cutter CviJI* (cuts RGCY, and RGCR/YGCY, but not YGCR) to generate a maximum number of small fragments with one treatment [39]. CviJI has proven to be a useful tool for generating probes from low-copy-number DNA sequences by thermal-cycle labeling [40]. A drawback of using this enzyme is the formation of blunt ends, which interfere with the efficient ligation of adapters because large numbers of fragments are available to religate to themselves. This inefficiency was confirmed experimentally. As an alternative to CviJI*, we used a cocktail of four-cutter enzymes (AciI, HpaII, HinP1I, MaeII, and TaqI) that have different recognition sites (CCGC, CCGG, GCGC, ACGT, and TCGA, respectively) but leave GC-5′ overhangs. After the digestion, the DNA fragments were treated with Klenow DNA polymerase in the presence of dCTP to fill in the 5′-overhangs with one nucleotide. This step converts the self-complementary 5′-GC protrusions into 5′-C overhangs that can ligate only to the synthetic adapters as specifically designed. To convert these short DNA fragments into primers that can be recovered and used in the enrichment protocol, they were "tagged" by ligation to adapter sequences (Figure 1B).

In the context of the human genome (3.2 Gb) [41], 18 nt is the absolute minimum length (x) required for the creation of a specific oligonucleotide, calculated by the formula $Nx/4^x < 1$, where N is the size of the target. For a large viral genome (e.g. N = 100,000 nt), this minimum length is reduced to 10 bp. However, if the aim is to distinguish a viral genome of that size within the context of the human genome as a background, a minimum size of 18 bp is needed to ensure unique sequence specificity. For viral discovery, the PEER

protocol uses MmeI [42] adapters to generate primers from the double-stranded cDNA tester with a 5′-end sequence that is artificially introduced by the adapter and a 3′-end that is derived from the target cDNA sequence (Figure 1B). With minor modifications, the adapter can alternatively incorporate sites for other IIS restriction enzymes, such as *Bpm*I, *Bsg*I, and *Eco*57I, that cleave 16 nt downstream of their recognition sites, a feature that might be useful for other applications (e.g. SNP discovery). Class IIS endonucleases that leave 5′-protruding ends [34] are not suitable for PEER because they will generate self-blocking primers. The IIS cleavage allows all cDNA fragments of various lengths that have acquired adapters to be trimmed to a uniform length. This generates a population of molecules with 3′-ends that are derived directly from the target of interest and that are suitable for extension reactions at a reasonably narrow temperature range. Such approach represents a novel use of these class IIS enzymes.

The enrichment efficiency of PEER was tested on serum containing HBV with a titer 3×10^8 IU/ml and Vero monkey kidney (VMK) cell culture infected with WCV at an inoculums' size of 10^4 copies. The HBV-containing serum was diluted tenfold four times with normal human serum pool to create testers with different viral loads, i.e. 3×10^7, 3×10^6, 3×10^5, and 3×10^3 IU/ml. dsDNA fragments before the enrichment protocol and the corresponding products after the enrichment were cloned in *Escherichia.coli* libraries. Such paired libraries were generated for all tested serum dilutions. As previously established [1], to find a high titer virus (e.g. 10^8 copies/ml) within a library representing the entire human genome one needs only to search through about 100 of the clones since ~2–3% of this library should contain viral sequences; however, if the viral titer is 10^3, one needs to screen 10^7 clones. To circumvent exhaustive screening by hybridization of libraries generated from serum containing low viral titer we resorted to evaluating the copy number of targets of interest in the dsDNA by PCR and hybridization. The PCR approach, although very reliable when used on the dsDNA material prior to enrichment, cannot be applied to PEER products because they might not consist of fragments that will contain both priming sites. To assess the presence of the targets of interest in the PEER product we did spot hybridizations and, from the corresponding libraries, isolated colonies at random and sequenced them. After PEER, the serum with initial titer of 3×10^6 IU/ml was enriched 5.3×10^2 times for HBV sequences and the one with initial titer 3×10^3 IU/ml was enriched 1.3×10^4 times. The observed higher enrichment values that were obtained for the lower titer library may be explained by the fact that the tester material was being diluted with the driver pool, thus creating a population almost perfectly matched to the driver and hence achieving a greater blocking efficiency. In fact, when enriching the low titer WCV virus with a perfect driver (noninfected VMK cells) the achieved enrichment value was 4.45×10^4.

PEER is a conceptually new approach for the subtractive enrichment of complex nucleic acid mixtures and represents a novel use for both class IIS restriction enzymes [43] and di-deoxytermination. Unlike the other techniques that are

based on subtractive hybridization of long DNA molecules that are eventually used as PCR templates, PEER centers on selective blocking of short DNA fragments through hybridization and highly specific enzymatic extension, and the use of these fragments as PCR primers. In addition, because PEER was designed to create normalized starting material that is double stranded, the method does not have preference for DNA or any particular type of RNA. PEER is suitable for use with samples of limited volume and is very cost efficient, especially when compared to new high-throughput sequencing methods. We did not observe loss of integrity of the background DNA, i.e. no recombination or insertion/deletion events. We did find some primer multimers among the clones, but did not quantify them since we filtered the sequence data for background vector or primer noise prior to analysis. PEER should also allow for several rounds of enrichment, as do RDA and SSH, i.e. the final PEER product can be digested again with the GC cutter cocktail, adapter-tagged, cut with *Mme* I and blocked on the same driver or even on an alternative driver, depending on the experimental goals. We have not yet attempted such experiments.

Finally, the PEER method is flexible and can be modified for the discovery of single-nucleotide polymorphisms (SNPs) or minor differences in allele states as well as for other subtraction applications such as pathogen discovery and differential expression of genes. Although descriptions of these other uses of PEER are beyond the scope of this chapter, a review of the PEER protocol should identify steps that may be modified to increase the versatility of the technique. For example, one may use different restriction enzymes separately or in various combinations to fragment the dsDNA. The primers can be created with the use of DNAses or exonucleases after the tagging step to generate fragments with randomly distributed ends suitable for a total comparative SNP analysis of the target NAs.

Our findings demonstrate that PEER is very robust, can be applied to different targets, and can detect NAs of unknown sequence at very low concentrations. In our experiments, PEER outperformed the commercially available SSH technique [44]. The method was recently successfully applied to a variety of viruses representing different genome structures – human herpesvirus and Ectromelia virus (dsDNA), human echovirus, West Nile virus, and human respiratory syncytial virus (ssRNA), Orthoreovirus (polysegmented ds RNA), porcine circovirus (circular ssDNA) and was able to recover viral genetic material from as little as 10^3 pfu [45].

2.3 PEER Protocol

2.3.1 Nucleic acid extraction

Total NA is extracted from 100–200 µl of serum or cell culture using Masterpure complete kit (Epicenter Biotechnologies, Madison, WI) or High Pure viral nucleic acid extraction kit (Roche) and resuspended in 10 µl 10 mM Tris (pH 8–8.5).

2.3.2 Modified SMART protocol

About 5 µl of the extracted NA is reverse transcribed (RT) with SuperScript II (Invitrogen, Carlsbad, CA). Two RT reactions are performed for the tester, one using 10 pmol each primer AFMmeIN6 and AFMmeISMART, and the other using 10 pmol primers T2N6 and T3SMART. Primers D0SMART and D0N6 are used for the driver reaction. Reaction volumes and conditions are described in the SMART cDNA synthesis protocol (Clontech). After synthesis, the enzyme is heat-inactivated and the product diluted with 40 µl of TE.

2.3.3 First PCR amplification

About 10 µl of the RT product is amplified with Advantage 2 Polymerase (Clontech) as recommended in the Smart cDNA protocol and using the corresponding PCR primers (AMmeIPCR for Tester 1, T3PCR, and T2PCR for Tester 23 and D0bioPCR for the driver) in triplicate reactions under the conditions suggested by the manufacturer. The amplification parameters are 95°C/1 min, and (95°C/3 s, 68°C/3 min) × 28 cycles. The dsDNA is purified on a Qiagen PCR purification column (Qiagen, Inc., Valencia, CA) and eluted in 75 µl 10mM Tris (pH 8).

2.3.4 Digestion with restriction endonucleases

About 70 µl of the Tester 1 dsDNA are digested overnight with HpaII (MBI Fermentas Amherst, NY), HinP1I, AciI (NEB Ipswich, MA), MaeII (Roche Molecular Biochemicals, Germany), and TaqI (NEB) using 1 µl of each enzyme and TaqI buffer (NEB) at 37°C. After digestion, the enzymes are heat-inactivated, with the fragments purified through a QIAquick PCR purification kit and eluted in 55 µl 10 mM Tris (pH 8).

2.3.5 Klenow treatment

The ends of the fragments are filled in with Klenow polymerase (Roche) in the presence of dCTP for 1 h at 37°C. The enzyme is then heat-inactivated; the reaction mixture purified with QIAquick nucleotide removal column (Qiagen), and the product eluted in 50µl 10 mM Tris (pH 8).

2.3.6 Adapter ligation

Double-stranded adapters are prepared by mixing the forward (AFMmeI) and reverse (ARMmeIP) adapter primers (Table 1) at equimolar ratio (200 pmol each), heating to 96°C for 5 min and slowly cooling to room temperature. 200 pmol of the adapter is ligated overnight to 45 µl of dCTP-filled-in Tester 1 fragments. The ligation products are purified to remove the T4 ligase and buffer with QIAquick nucleotide removal column and eluted in 55 µl 10 mM Tris (pH 8).

2.3.7 MmeI digestion

The ligation products are digested with 5U MmeI (NEB) for 2 hrs. The cleaved DNA is resolved in 10% polyacrylamide gel, the resulting 50 bp fragment is cut out, isolated from the gel with QIAquick gel extraction kit (Qiagen) and resuspended in 50 µl 10mM Tris (pH 8).

2.3.8 Blocking of MmeI-tagged primers

About 25 µl of the fragment is used as primer with 10 µl Driver bio-dsDNA template in the presence of 2.5 µl each ddNTPs-bio (Biotin-11-ddNTPs, NEN Life Science Products Inc., Boston, MA), 0.025 mM each dNTPs (Roche) and Thermo Sequenase (Amersham Pharmacia Biotech, Inc., Piscataway, NJ). The blocking reaction is carried out as follows: 96°C/3 min; and (95°C/2 s, 55°C/20 s, and 68°C/20 s) × 55 cycles. The product is purified with QIAquick nucleotide removal kit to remove the excess ddNTPs and eluted in 100 µl 10 mM Tris (pH 8).

2.3.9 Removal of biotinylated products

The cleaned product is heated to 95°C, and 50 µl of streptavidin-coated magnetic beads are added (SPHERO Streptavidin Magnetic Particles from Spherotech, Inc., Libertyville, IL). After 10 min incubation at >60°C, the beads are captured on a magnet rack (Qiagen) and the supernatant removed to a fresh tube, ensuring that the temperature remains above 55°C.

2.3.10 Capture reaction

About 50 µl of the supernatant (purified nonblocked primers) are used in a 100 µl capture reaction with 5 µl of the Tester 23 cDNA as template under the following conditions: 95°C/2 min; (95°C/20 s, 45°C/30 s, 72°C/2 min) × 10 cycles; (95°C/20 s, 52°C/30 s, and 72°C/2 min) × 30 cycles; and 72°C/7 min.

2.3.11 Final PCR

About 5 µl of the capture product is amplified in a 100 µl final reaction volume with primers AMmeIPCR and T2PCR, AMmeIPCR, and T3PCR or AMmeIPCR alone under the following conditions: 95°C/2 min, and (94°C/10 s, 60°C/20 s, and 72°C/90 s) × 30 cycles. The product is quantified, cloned, and sequenced.

3. OTHER SUBTRACTION AND HYBRIDIZATION BASED METHODS FOR DIFFERENCE SCREENING

Subtractive hybridization is the core approach behind most techniques described here. Differentiating and excluding/subtracting nucleic acid species by hybridization or by physical comparison can be accomplished in a creative variety of ways.

3.1 Differential Screening

Differential screening [46] enables comparison of two complete mRNA populations by probing a cDNA library generated from one of the samples with labeled fragments representing one of the RNA populations. A summary schematic representation of the approach, also known as plus/minus screening, is shown

on Figure 2. This protocol can be very thorough and informative because it theoretically allows the acquisition of a complete picture of all mRNAs involved. In actual experimental settings, however, the method has always proved to be very laborious [47]. It requires the building of a good cDNA library that has an adequate number of clones plated at manageable density and the generation of several lifts from said library, that need to have adequate and approximately equal amount of the clone/plaque material on them, so that the positive–negative hybridization calls after the screening are not biased by clone copy numbers. The hybridization itself needs to be done with limiting concentration of the probe as to distinguish differences between mRNA copy numbers in the compared populations; such conditions can be difficult to achieve. Even though many of the technical drawbacks that come with cDNA library construction are no longer a great technical challenge today, there still are obstacles that cannot be easily overcome. The major disadvantages are: the intensive labor

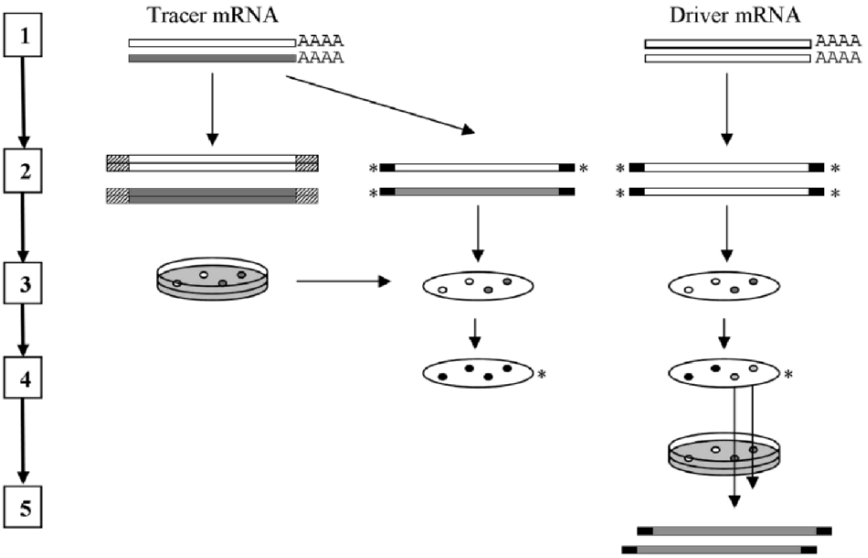

Figure 2. Differential screening or plus/minus screening (DS). 1. Total mRNA from tracer/cancer material (white and gray rectangle) and mRNA from driver/healthy material (white rectangle) are reverse transcribed in the presence of random hexamers and oligo(dT) primer; 2. cDNA is ligated to restriction site linkers (vertical diagonal fill rectangle) and the product is cloned in appropriate λ-expression vector. Aliquots of the mRNA are labeled with P^{32}-ATP (black fill rectangles*) and polynucleotide kinase, or alternatively the corresponding cDNA is synthesized in the presence of P^{32}-dNTPs; 3. Duplicate plaque lifts from the cDNA library are made on nylon membranes; 4. One set of the membranes is probed with the labeled tracer mRNA and the duplicate set is probed with the labeled driver mRNA; 5. Plaques that do not give hybridization signal when probed with the driver mRNA are isolated form the corresponding original plate used to create the lift and subject to secondary screening and further analysis.

required, poor reproducibility, bias against less abundant mRNA species, and the requirement for multiple controls and secondary screening. Successful uses of this approach include: differential expression studies of polycystic kidney disease [47], TNF-induced genes [48], and early embryonic differentiation [49].

3.2 Subtractive Hybridization

The search for differentially expressed genes [50,46] and the development of subtractive hybridization have gone hand in hand [51]. Subtractive hybridization is any hybridization applied to remove common material that can interfere with the desired outcome of a screen and may be applied to the test sample of interest, to the probe designed to look for targets of interest [52,53,54,55], or to both [56]; it is depicted here in Figure 3. The best enrichment obtained by subtractive

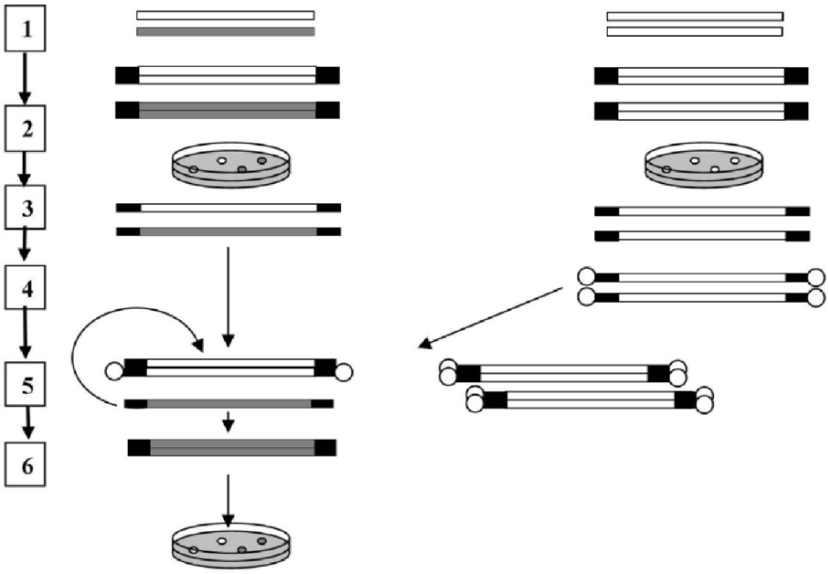

Figure 3. Subtractive Hybridization (SH). 1. Total mRNA from tracer/cancer material (white and gray rectangle) and mRNA from driver/healthy material (white rectangle) are reverse transcribed in the presence of random hexamers and oligo(dT) primer; 2. cDNA is ligated to cloning site linkers and the product is cloned in a λ-bacteriophage vector; 3. Single-stranded phage DNA is recovered by transfection of *Escherichia coli* with in parallel with a helper phage; 4. Driver phage is biotinylated; 5. Tenfold excess of the driver is mixed with the single-stranded tester material and allowed to hybridize; 6. Biotinylated homo and heteroduplex molecules are removed by avidin agarose or streptavidin and the remaining subtracted ssDNA is taken into another round of subtraction with fresh driver; or converted to dsDNA with Klenow enzyme and cloned (subtractive cloning) or labeled and used as positive hybridization probe on libraries containing the target of interest; or amplified by PCR and subsequently used in the same applications.

hybridization followed by amplification and cloning has been between 100- and 1000-fold, depending on the target [55,20]; the kinetics of the process have been described previously [57,58]. An interesting example of recent work that uses straightforward subtractive hybridization is the isolation of all messages involved in the utilization of complex polysaccharides from *Aspergillus nidulans* [59]. The approach used was to collect mRNA from fungi grown on variety of polysaccharide sources different from glucose, and to pool and clone those together without any other selective enrichment or PCR. The resulting library was then probed by negative subtractive hybridization with labeled mRNA from *Aspergillus* grown in the presence of glucose only, which resulted in the isolation of over 3000 negative clones that generated more than 2000 unique contigs. To confirm their function, the clones were arrayed and probed with cDNAs from fungi grown under different sugar conditions. Subtractive hybridization has been applied with success to uncover tumor suppressor genes [60]; to find moderately induced sequences from humans [61] and mycobacteria [62], especially when coupled with PCR; and to study monoallelic imprinted genes [63] and organ-specific gene expression [64].

3.3 Subtractive Cloning

Subtractive cloning takes the differential screening approach a step further by constructing cDNA libraries that are already enriched in differentially expressed sequences [50,46,65]. It is achieved by generating single-stranded cDNA from the material of interest and hybridizing it to excess mRNA from another cell type. The double-stranded hybrids are the genes expressed in common and the single-stranded mRNAs represent the ones unique to the cell of interest that can be further cloned and validated. This approach has the revolutionary advantage of removing the background of common genes that otherwise interferes with the desired outcome. The main disadvantages are the requirement for large quantities of mRNA (only 5% of the general RNA population) and inefficient recovery of the single-stranded material, which thus limits the availability of cDNA for cloning. The approach has been used successfully [47] to isolate a new member of the Ras super family, the T-cell antigen receptor, the murine IL-4 receptor, etc. The subtractive cloning has also seen a great benefit from coupling with methods that take advantage of poly-A tailing of most mRNAs [66], PCR [67,68], secondary hybridization [69] or a combination of these [70,71,72],

3.4 Differential Display

DD methods are not true subtraction methods but can nonetheless play a great role when choosing an approach to find unique sequence characteristics. Many of the methods described below can be used in combination with subtraction technique.

The mRNA DD [8] or arbitrarily primed PCR [73] came about as an alternative to the differential screening. It takes advantage of the power and specificity of PCR and does not require *a priori* construction of libraries. In its original design, the method was intended to display only a subset (1/12) of the original mRNA population (Figure 4). However, it retains great flexibility as to the kind of arbitrary or specific oligonucleotides used (gene-specific, gene-family specific, anchor-polyA-tail, etc.), as these can be employed in variety of combinations to create subsets of fragments that are unbiased and suitable for screening. Its advantages are the use of amplification, the ability to query with different sets of oligonucleotides, and greater reproducibility and opportunity to compare more than two specimens of interest at a time. However, regardless of the size of the subpopulation displayed, the background can mask true differences; rare transcripts can be missed in the PCR step; multiple primer sets may be required to cover adequately the entire mRNA population, and the results usually reflect uniquely induced rather than up-regulated or down-regulated

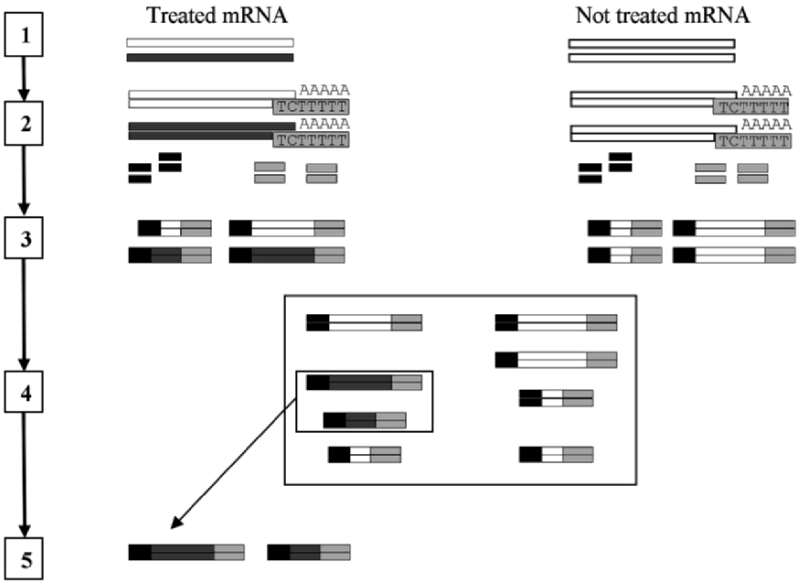

Figure 4. Differential display (DD). Generation of representations from mRNA. 1. mRNA from treated material (white and gray rectangle) and mRNA from untreated material (white rectangle) are reverse transcribed with 5′-T_{11}CA to allow anchored annealing to approximately one twelfth of the entire mRNA population; 2. cDNA is the PCR amplified with primers T_{11}CA and a 10 mer Ltk3 (black rectangles) and labeled with radionuclide (α-^{35}S) to allow visualization on polyacrylamide gel by autoradiography; 3. Fragments bellow 500 bp in size are resolved on a DNA sequencing gel; 4. Resolved fragments display the differences between the compared (could be more than two) specimens and after thorough screening the different bands are isolated form the gel; 5. Isolated fragments are reamplified, cloned, and analyzed.

gene expression. Attempts to circumvent these disadvantages include primer redesign, increasing the cDNA concentration, increasing the 5′-primer:3′-anchor ratio, and devising strategies for systematic rather than arbitrary display [74,75,76,77,9,78]. An interesting attempt to improve this method was the coapplication of the subtraction approach, generating the differential subtraction display [79] from which cDNA eluted from down-regulated gel band is amplified, biotinylated, and used in excess as "driver cDNA", and cDNA form an up-regulated band is amplified and used as "tester cDNA" without biotinylation. The two amplicons were allowed to hybridize, after which the hybrids were removed by streptavidin and the residual cDNA cloned.

3.5 AFLP, SAGE/CAGE, GSTs, and DARFA

Multiple approaches to the display of genetic differences have been described in addition to DD. All approaches are intended to provide complete and unbiased fingerprints of specific transcriptomes or developmental stages. Highlighted here are the more widely used techniques that have become the methods of choice for various targets [80,81].

A novel DNA fingerprinting technique displaying the amplification fragment length polymorphism (AFLP) of total digests of genomic DNAs [82] was based on the restriction fragment length polymorphism technique (RFLP) [83], supplemented and enhanced with PCR with specific primers that provided the ability to amplify only selected subsets of the genome. AFLP has proven indispensable for microbial typing [84,85].

A display method that directly addresses the specific fingerprint of gene expression at the sequence level is serial analysis of gene expression (SAGE) [86]. It is founded on two principles: (1) a short sequence tag contains adequate information to uniquely identify a transcript if it is isolated from a defined position of this transcript, and (2) concatenation of such tags in a single clone will provide instant, rich, and manageable sequence information about the particular transcriptome (Figure 5). The method has shown very good reproducibility ($R^2 = 0.96$) [81] and has found wide application [87,88,89]. A drawback of this method, however, is the fact that frequently (in about 25–30% of the instances) the generated tags are not long enough to be unambiguously assigned to a transcript. SAGE has also been modified to query the 5′-end of transcripts along with the corresponding promoter elements to generate a technique named CAGE [90] and to provide another method for identification and quantitative analysis of genomic DNAs called genomic signature tags (GSTs) [91]. The principle of SAGE has been combined with subtractive hybridization and immunoprecipitation in another recently developed technique called serial analysis of binding elements (SABE), targeted to study gene regulation in humans [92].

One of the newest additions to the DD technologies, called differential analysis of restriction fragments amplification (DARFA) [85], is amenable to both complete transcriptome analysis and DNA fingerprinting. The technique is open,

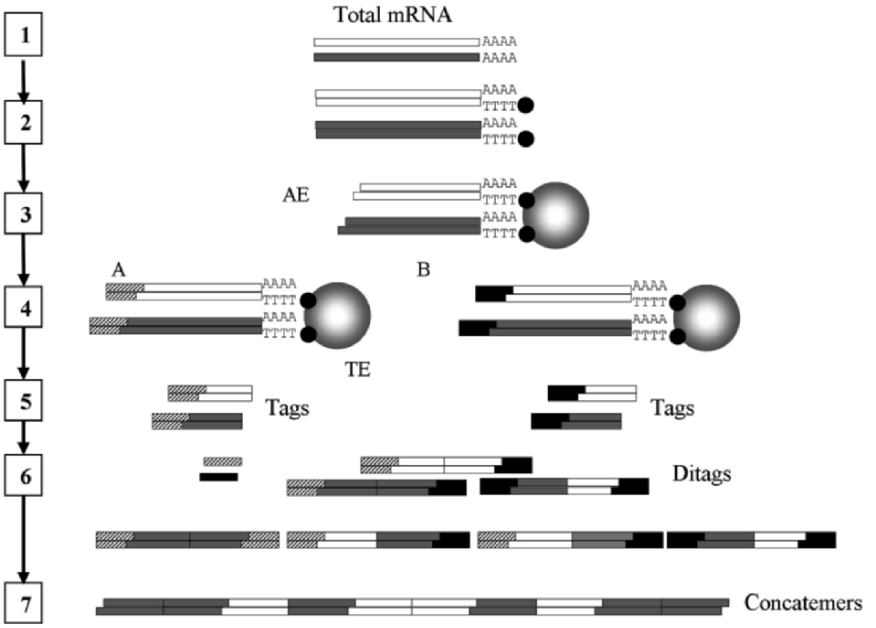

Figure 5. Serial analysis of gene expression (SAGE). 1. Total mRNA from the target material (white and gray rectangle) is reverse transcribed with 5′-bio-oligo(dT) primer (TTTT black-fill circle); 2. cDNA is cleaved with a four-cutter restriction endonuclease-anchoring enzyme (AE), which is expected to cut most transcripts more than once; 3. Fragments are bound to streptavidin beads and divided in two aliquots; 4. Fragments are ligated to different adapters (diagonal stripe rectangles – A, black rectangles – B) containing a site for an IIS endonuclease and then digested with this IIS-tagging enzyme (TE); 5. Released tags are ligated to create ditags; 6. Ditags are PCR amplified with primers A and B and the primers are removed by cleavage with the AE enzyme; 7. Fragments are ligated into concatemers, cloned, and analyzed.

i.e. has no requirement for prior sequence information and claims to be able to display the entire transcriptome due to the characteristic of Hpy188III to generate 120 subpopulations based on the 2 nt 5′-overhang sequence combinations (Figure 6). This approach is intended to ensure that every one restriction fragment will acquire adapter and thus PCR sites and a 4 bp combination identifier of the subpopulation, but the method is no less cumbersome than other display approaches and requires large amounts of starting material. The starting material may also be subject to a preliminary PCR step to provide adequate amount for the downstream procedures, which may diminish the accuracy of the method's representation.

3.6 Representational Differences Analysis

Representation differences analysis (RDA) builds successfully on the subtractive hybridization approach and was designed to find small differences between the sequences of two DNA populations [2,93]. The approach employs a combination

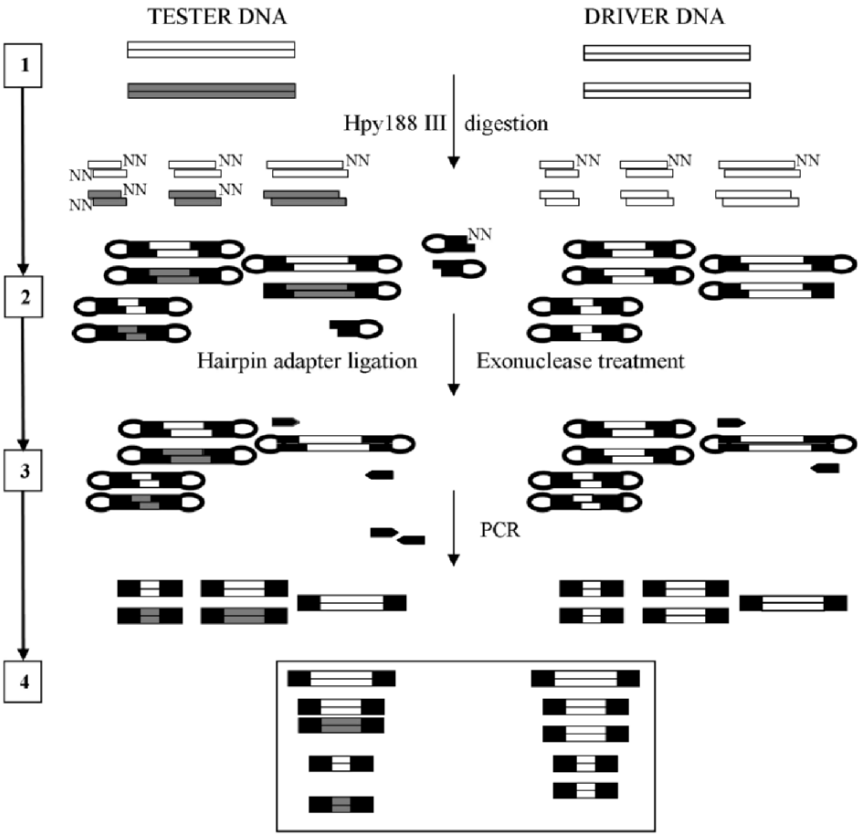

Figure 6. Differential analysis of restriction fragments amplification (DARFA). 1. Generation representations from total DNA. Tester (white and gray rectangle) and driver (white rectangle) DNA or cDNA are digested with Hpy188III restriction endonuclease (TCNNGA) to reduce the complexity of the starting material and create a pool of 120 subpopulations differing by their two nucleotide 5′-overhang sequences; 2. Fragments are ligated to hairpin adapters (black staggered rectangles with loop); 3. Products are treated with exonucleases to remove the unligated fragments and hairpin adapters and amplified by PCR with primers derived from the first 18–3′-nucleotides of the adapters (black-filled directional rectangle); 4. Amplicons are separated on 5% denaturing acrylamide gel and the difference bands isolated for further analysis.

of subtractive and kinetic enrichment of PCR amplicons. DNA "representations" are created by cleavage of DNA by a fairly infrequent restriction endonuclease, ligation to oligonucleotide adapters, and PCR amplification. This generates a representative subset of the genome whose complexity is reduced by 10–50 times depending on the endonuclease of choice. The adapters of the tester representation product are removed by cleavage and replaced by a different new set of adapters. The tester then is allowed to hybridize to the driver (subtraction step, where the tester/driver heterodimers will represent the identical sequences),

and the ends are filled in. The resulting material is not subject to physical separation but amplified with the second-adapter oligo alone (Figure 7). This represents the kinetic enrichment step where tester sequences gain amplification advantage because they alone have two priming sites and thus will be amplified exponentially. The approach has great merit and its best applications are the generation of restriction fragment polymorphism probes and difference cloning. Many variations to improve the method's range have been suggested, including

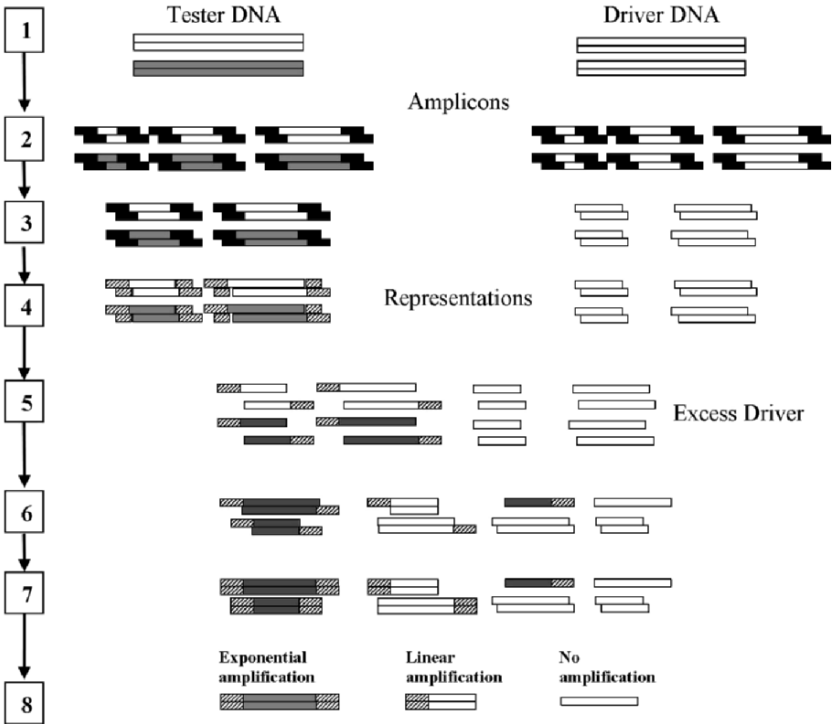

Figure 7. Representational differences analysis (RDA). Generation representations from total DNA: 1. Tester DNA (white and gray rectangle) in two and driver DNA (white rectangle) are digested with restriction endonuclease (Bam HI, Bgl II and Hind III) to reduce the complexity of the starting material; 2. Fragments are ligated to Adapters (black rectangles); 3. Products are amplified for 20 cycles by PCR with adapter primers thus creating representation of molecules smaller than 1 kb; 4. All PCR adapters are removed by cleavage and the tester is ligated to a different set of dephosphorylated adapters (upward diagonal fill rectangles). Subtraction and Kinetic Enrichment; 5. The tester representation is mixed with excess driver and denatured; 6. Mixture is allowed to rehybridize and if the amount of driver is adequate the only hybrid molecules with adapters on both strands will be the ones that are specific to the tester (subtraction); 7. Hybrids are treated with Taq polymerase in the presence of all dNTPs to fill in the 3′-ends; 8. PCR amplification with the second set adapter primers generates exponential amplification only form templates with priming sites on both strands (kinetic enrichment). Products may be additionally treated with mung bean nuclease to remove the single-stranded material and reamplified.

a variety of restriction endonucleases and different adapters to make it usable for the comparison of more that two genomes from environmental samples [94]. Rapid subtraction hybridization (RaSH) was introduced [95] to provide a streamlined RDA-based cloning. RDA has been used successfully not only on DNA but also on cDNA representations [96,97,98], and applications such as the discovery of GBV-A and GBV-B [99,100], and TTV [32], cloning of apoptosis-related genes [101], iron-regulated gene expression in bacteria [3,102], identification of new tumor suppressor genes [103], changes of expression in malignant formations [104] and identification of developmentally regulated genes [105]. RDA served as a base concept for a creative alternative PCR subtraction technique termed ligation mediated enrichment (Limes) [106]. The later was specifically designed to address the problem of genetic backgrounds that have high repeats content and uses Taq DNA ligase to join only perfectly matched ends, thus creating amplifiable templates only from perfect hybrids.

3.7 SPAD–RDA

Selectively primed adaptive driver (SPAD)–RDA is a fairly novel adaptation of the subtractive hybridization and differs from RDA by an alternative approach to the generation of the tester and the driver and by the optimization of the driver material in the course of subtraction [4] (Figure 8).This improvement potentially circumvents the recognized drawbacks of any PCR-based subtraction, that is if the complexity of the starting material is at too high or the target at too low a concentration, the enrichment will be ineffective. In essence, SPAD–RDA combines RDA [2] with the use of selective primers, as for AFLP [82], and driver-control subtraction, as for SABRE [6]. The approach leads to improved recovery of viral sequences on a greatly reduced background but does not address the low-copy-number problem.

3.8 Enzymatic Degradation Subtractions (EDS, LCS, DSC, NSC, UDG/USA, and CODE)

Many adaptations of subtraction hybridization (SH) PCR approaches have been proposed in attempt to circumvent one of SH's most problematic areas – the background created by incomplete driver/tester hybridization. One large category of adaptations consists of methods that employ one or multiple enzymatic degradation steps to disable the unwanted templates.

The enzymatic degrading subtraction (EDS) [107] is one of the earliest proposed alternatives for the construction of subtractive libraries from PCR-amplified cDNA. With EDS, the tester DNA is blocked by thionucleotide incorporation that renders it resistant to exonucleases III treatment; the rate of tester/driver hybridization is accelerated by phenol–emulsion reassociation [108], and the driver cDNA and hybrid molecules are enzymaticaly removed by digestion with exonucleases III and VII rather than by physical partitioning.

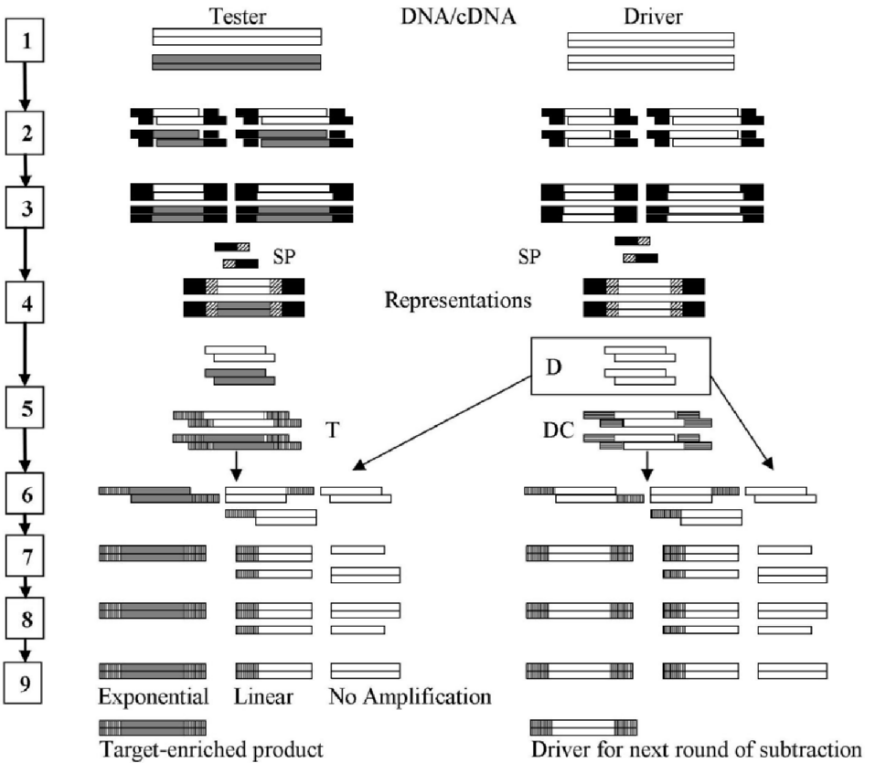

Figure 8. Selectively primed adaptive driver–RDA (SPAD–RDA). 1. Tester DNA (white and gray rectangle) and driver DNA (white rectangle) are digested with restriction endonuclease SauIIIA to reduce the complexity of the starting material; 2. Fragments are ligated to 5′-dephosphorylated adapters (black rectangles) and after column purification the ends are filled in; 3. Products are amplified by PCR with selective primers (SP) to create representations; 4. SP sequences are removed by cleavage; 5. Tester (T) is ligated to a different set of dephosphorylated adapters (vertical fill rectangles) and so is a portion of the driver to create a driver control (DC); 6. T and excess adapter-free driver (D) are mixed, denatured, and reanneled (subtraction); the same is done in parallel with a D/DC mix (driver control subtraction); 7. 3′-ends of hybrids are filled in; 8. PCR amplification with the second set adapter primers generates exponential amplification only from templates with priming sites on both strands (kinetic enrichment); 9. Product is treated with mung bean nuclease to remove the single-stranded material and reamplified.

The utility of EDS has been demonstrated by constructing a subtractive library enriched for cDNAs differentially expressed in adult rat brains [107].

Shortly after EDS, a new modification of SH was introduced [109], which could potentially achieve the cloning of differentially expressed genes by linker capture subtraction (LCS) where the tester and driver are digested, ligated to a linker, and amplified. The linker is removed from the driver and after hybridization the products are subjected to mung bean nuclease action to remove ssDNA (Figure 9).

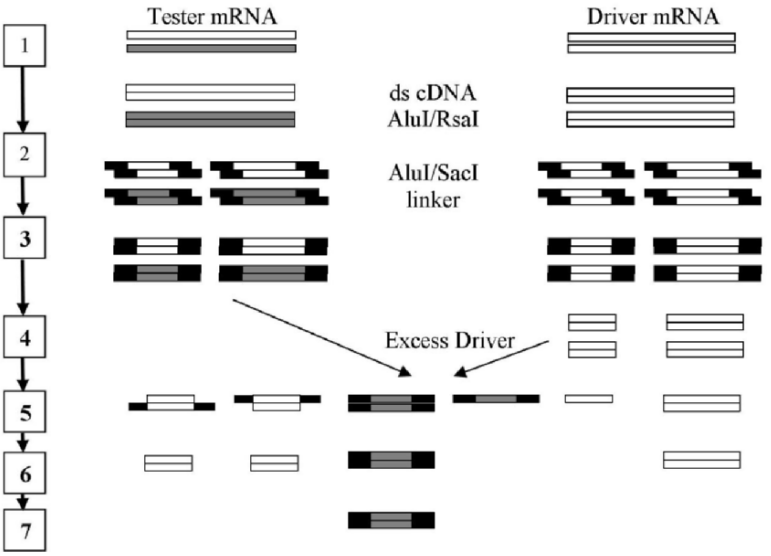

Figure 9. Linker capture subtraction (LSC). Panel A. 1. Tester mRNA (white and gray rectangle) and driver mRNA (white rectangle) are converted to double-stranded cDNA and digested with restriction endonucleases AluI and RsaII; 2. Fragments are ligated to linkers that contain AluI/SacI restriction site (black staggered rectangles); 3. Products are amplified for 20 cycles by PCR with adapter primers derived from the top strand of the linkers; 4. Linkers removed from the driver PCR products by cleavage; 5. Tester product is mixed with excess driver and denatured. Mixture is allowed to rehybridize and if the amount of driver is adequate the only hybrid molecules with adapters on both strands will be the ones that are specific to the tester (subtraction); 6. Hybrids are treated with mung bean nuclease to remove all ssDNA; 7. PCR amplification with the same linker primers generates exponential amplification only form templates with priming sites on both strands. Product may reenter the process at step 5 for additional rounds of enrichment.

The subtraction step can be repeated several times and, unlike RDA, where the enrichment is mostly kinetic at the final PCR step, the target is selected by the specifically preserved priming sites. This idea is appealing as it seems to simplify the RDA approach. Like RDA, LCS takes advantage of the improved hybridization kinetics of nucleic acid mixtures with reduced complexity with each round of subtraction providing better levels of subtraction.

DSC is another approach that is very similar to LSC [5] and is also based on the alternative "negative" amplification strategy. The main principle is that if tester sequences have counterparts in the driver, these can be rendered unamplifiable to leave the desired unique sequences available for amplification. This availability is achieved by digestion of separate driver and tester pools with a restriction endonuclease followed by ligation of different adapters to provide unique PCR priming sites. The adapter-tagged fragments are then amplified and the adapter sequences are removed from the driver by digestion. The products

are then left to hybridize and the resulting single-stranded ends are digested away with mung bean nuclease from the tester molecules that have found a homologue in the driver (Figure 10). The result is the conversion of the tester sequences that have counterparts in the driver to new driver molecules with the tester population depleted and the driver population enriched and available for further rounds of subtraction with an exponentially increasing amount of seemingly appropriate driver.

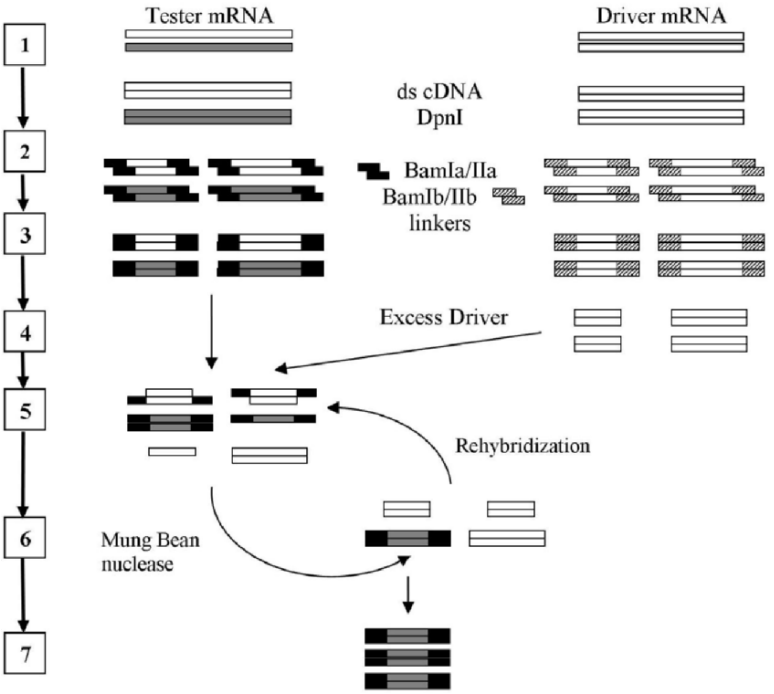

Figure 10. Differential subtraction chain (DSC). Panel A: 1. Tester mRNA (white and gray rectangle) and driver mRNA (white rectangle) are converted to double-stranded cDNA and digested with restriction endonuclease DpnI; 2. Fragments are ligated to two separate sets of primer/adapters (black staggered rectangles and diagonal fill staggered rectangles) that contain BamHI restriction site; 3. Products are amplified by PCR with adapter primers derived from the top strand of the adapters; 4. linkers removed from the driver PCR products by cleavage; 5. Tester product is mixed with excess driver and denatured. Mixture is allowed to hybridize and if the amount of driver is adequate the only hybrid molecules with adapters on both strands will be the ones that are specific to the tester (subtraction); 6. Hybrids are treated with mung bean nuclease to remove all ss DNA; this also removes all single-stranded adapter sequences found in tested/driver heterohybrids homologous molecules, thus creating additional driver fragments that can reenter the enrichment step (rehybridization); 7. PCR amplification with the same linker primers generates exponential amplification only form templates with priming sites on both strands. Product could re-enter the process at step 5 for additional rounds of enrichment.

Potential drawbacks of this approach is the possibility that the sequence of interest might contain insertions or deletions that will allow the formation of a hybrid and from this point on such sequence can be lost for the enrichment process. Another problem might be the "destructive" nature of the approach itself, since rare products may not reanneal easily and thus if they remain in single-stranded form, will be removed from further selection. The bias against single-stranded species is a common shortfall to all subtraction approaches. Still, the method is quite insightful and has been used successfully in cancer studies for the identification of a several differentially expressed sequences [110,111].

Cloning of deleted sequences (CODE) is yet another alternative of the classic subtraction method, influenced by DSC and designed to address cloning of deleted sequences, i.e. of sequences present in the tester but no longer in the driver [112]. This procedure combines (1) the use of restriction endonucleases to reduce the complexity of the sample; (2) utilization of dUTPs and Uracil deglycosilase (UDG) to remove unwanted driver DNA after the hybridization step; and (3) a biotinylated primer to rescue the fragments of interest.

A recent review on SH [113] that briefly highlights all major current subtraction strategies also gives a good account of how a combination of UDG and single-strand-specific nuclease treatment can eliminate unwanted tester/driver and driver/driver hybrids (Figure 11).

Another addition to the approaches aimed at reducing or eliminating subtraction background created by cross-hybridization is a technique called mispaired DNA rejection (MDR) [114]. MDR also takes advantage of the abilities of common mismatch repair enzymes (i.e. mung bean nuclease is able to attack single-stranded loops in double-stranded structures and the Surveyor nuclease recognizes and cleaves mispaired structures within DNA duplexes). The method is elegantly designed to test the ability of the treatment to remove from the final subtraction products such as repetitive clones and chimera clones. A very useful application of MDR is also the recovery of highly conserved sequences between different genomes.

The DSC method has evolved into negative subtraction chain NSC [115], another SH-based approach that takes advantage of DSC idea and employs improved adapters to remove the background after a round of classic SH. Another recent modification of SH, coupled with template switching universal long PCR [116] and single-strand deletion has been proposed as a quick method to identify unknown viral agents [117].

3.9 Suppression Subtraction Hybridization

Subtractive hybridization methods have proven valuable for the identification of significant genes from multiple systems representing the growth and differentiation of cells [118,119,46,120]. Most of the methods are quite tedious and involve complicated protocols that require great accuracy, attention to detail, and often

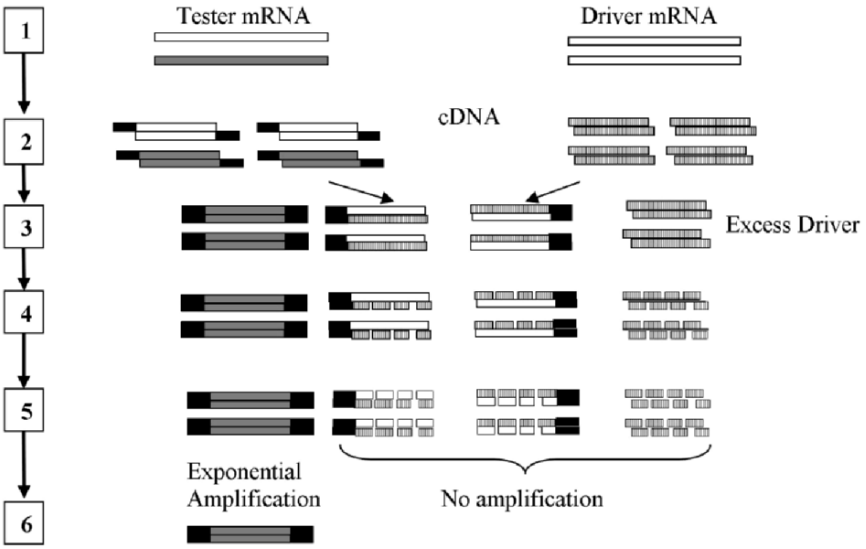

Figure 11. Uracil–DNA subtraction assay (USA). 1. Tester cDNA (white and gray rectangle) is ligated to 5′-adapters (black rectangles) and driver cDNA (vertical stripe rectangle) is synthesized from mRNA (white rectangles) in the presence of Uracil; 2. Tester is mixed with excess of the driver, the mixture melted, and allowed to reanneal; 3. Hybrids are treated with Klenow and then with uracil–DNA glycosidase to generate nicks in all molecules containing driver fragments; 4. Hybrid mixture is then treated by singe-strand-specific nuclease digestion that generates short unamplifiable fragments from the driver homohybrids and the tester–driver heterohybrids; 5. Product is then amplified by the tester adapter primers to generate tester-specific amplicons; 6. The first difference product could be used in the same scheme for another round of subtraction, back in step 3 to generate a second difference product.

substantial amounts of starting material to ensure that no rare molecular species of importance will be omitted. One of SSH's central features is the utilization of PCR suppression [121]. PCR suppression is based on the premise that if long inverted terminal repeats are attached to the ends of DNA fragments, they will form stable panhandle structures at the end of each denaturation and annealing cycle; in the presence of PCR primers that have the sequence of these repeats, no amplification will occur. SSH has another feature of the method is that it normalizes or equalizes the sequence abundance in the target cDNA population by splitting the target tester material in two aliquots and ligating them to different sets of adapters with inverted repeats. The normalization occurs as the more abundant species readily reassociate with tester or homologous driver sequences; the rare species remain single stranded. The two testers are then hybridized to excess driver and then mixed together and extra denatured driver added in a second hybridization step. This creates tester heterohybrids from the difference material that are the only molecules available for PCR (Figure 12). SSH is capable of enriching the target ~1000-fold after

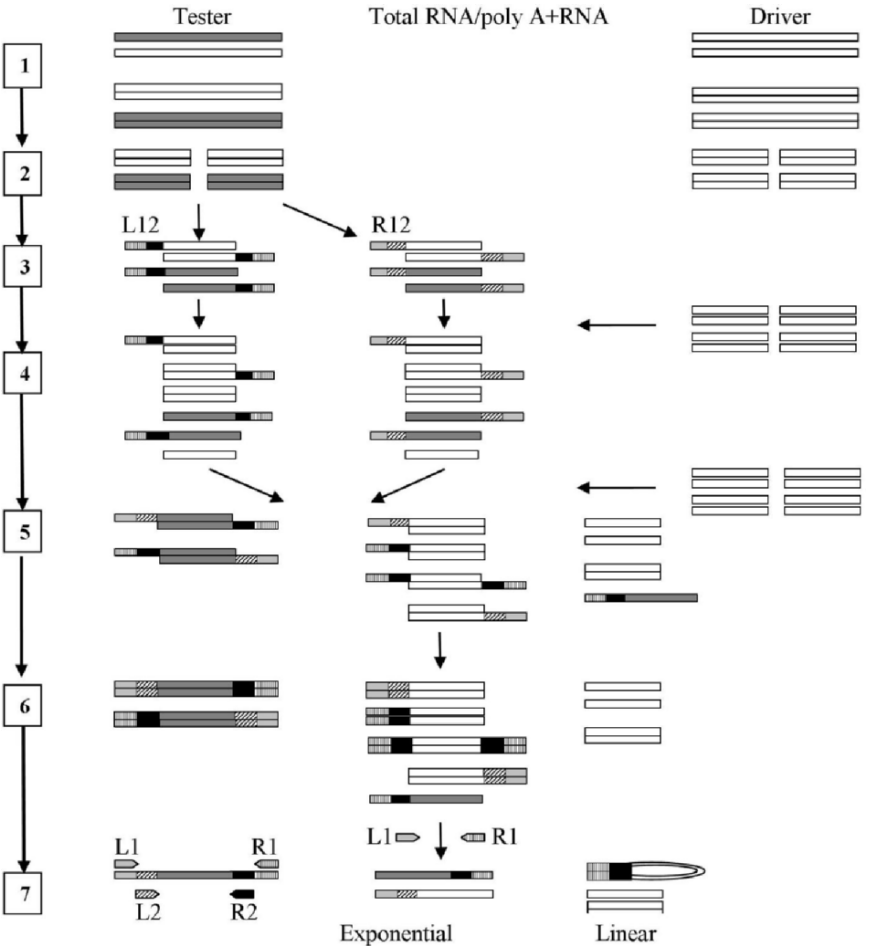

Figure 12. Suppression subtraction h ybridization (SSH). 1. Total RNA or PolyA + RNA from the tester (white and gray rectangle) and the driver (white rectangle) are converted independently to double-stranded cDNA by conventional or SMART RT–PCR; 2. All cDNAs are digested with RsaI; 3. Tester is divided in two aliquots and each is ligated to different set of nonphosphorylated adapters (L12 and R12) 4. First hybridization step. Both tester aliquots are subtracted independently with excess of the driver; 5. Second hybridization step. Products from the first step are mixed and without further denaturation are supplemented with fresh denatured driver and the mixture is allowed to rehybridize; 6. Ends of the hybrids are filled in the 3′-ends; 7. Product is amplified by two rounds of PCR with nested primers L1/R1 followed by L2/R2.

only one round of subtraction [122,123,124,125,126]. SSH is an improvement on the subtraction approaches [7] and has found a wide application [127,128,129,130]. SSH is a clean technique with truly good performance; however, SSH-generated libraries contain some background clones that do not represent differentially expressed transcripts. SSH is not ideally suited when dealing with transcripts that

are only moderately (2–4-folds) enhanced, and generally requires a perfect driver match for optimal performance. SSH could miss unique sequences that are present at very low copy numbers and have remained single stranded and thus had not acquired adapters. The problem posed by false positive clones has been tackled by the mirror orientation selection (MOS) approach [131], which deals with redundant molecules that have evaded the hybridization subtraction step and have remained in the process. The rationale behind MOS is that such molecules have only one orientation relative to the adapters, while each genuine product of the enrichment process has been generated by both orientations of the molecule; hence, if one of the adapters is removed enzymatically, the subtracted product is allowed to denature and reanneal again, so the background molecules will remain single stranded with only one priming site which can allow only their linear amplification. SSH has served as a basis for ligation-mediated suppression PCR [132,133,134] that has been used successfully for genome walking into unknown sequence areas. SSH or the combination of SSH and MOS have been used successfully for detection of differentially expressed genes in many systems [135,136], including work identifying diversity in an environmental genome [112], early mammalian embryonic development studies where the starting material can be very scarce and some transcripts extremely rare [25,33]; disease studies [26,137,138,139,140,29,23,141,142,143]; pharmacogenomics [144]; infectious agents studies [84,122]; and virus–host interaction studies [145,146,147,22,148].

3.10 Selective Amplification Via Biotin and Restriction-Mediated Enrichment

SABRE is an approach that uses the selective enrichment principle of RDA, combines it with the use of biotin–streptavidin affinity and restriction enzyme site reconstitution to achieve purification of the desired tester homohybrid population [6]. It was designed to take advantage of restriction-mediated reduction of the complexity of the starting material and biotin- and restriction-mediated recovery of tester homohybrids. The protocol demonstrates the ability to detect moderately rare (representing ~0.03% of the total) mRNA species and 2–10-fold elevation in their expression levels (Figure 13). In addition to the tester/driver subtraction, the driver cDNA is subtracted in a parallel control experiment with another batch of the driver DNA amplified with tested adapters. The product of this subtraction control reaction is then used for a second round of subtraction of the tester material to ensure that any differences that may arise form PCR irregularities due to the adapter do not contribute to a background accumulation of false positives. This thoughtful approach uses virtually identical primers for the generation of the tester and the driver, thereby ensuring that both representations will be comparable, allows for multiple subtraction rounds, and, using two separate elements (both streptavidin capture and restriction], ensures the selection of tester-derived molecules [149]. A possible drawback of this method is bias against low-copy tester material that might remain single

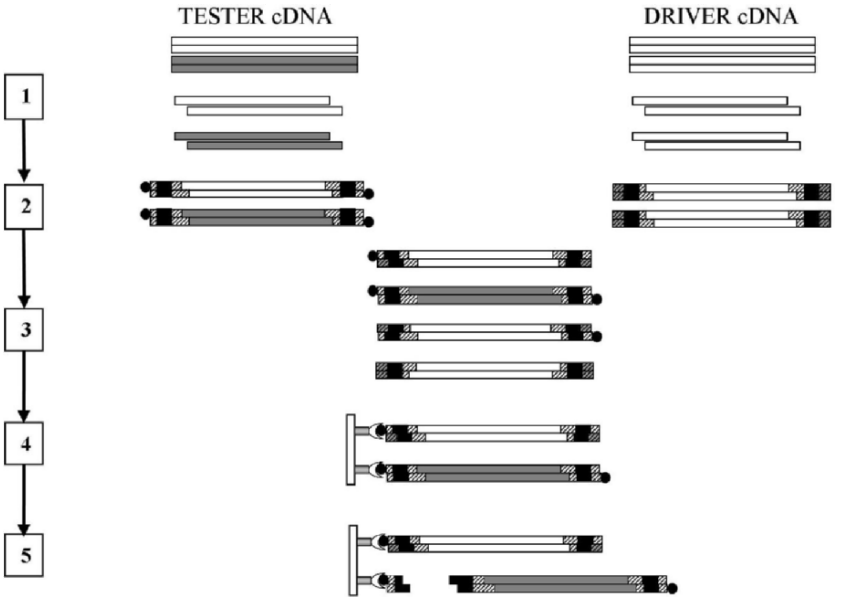

Figure 13. Selective amplification via biotin and restriction-mediated enrichment (SABRE). 1. Tester (white and gray rectangle) and Driver (white rectangle) double-stranded cDNA are digested exhaustively with restriction endonuclease MboI; 2. MboI-linker-adapters are ligated to both tester and driver. Tester adapters are biotinylated and have a functional BamHI site; the driver adapters are not biotinylated and lack the correct BamHI site. Tester and driver may be are amplified with virtually identical PCR oligonucleotides designed to anneal; 3. Products are mixed, denatured, and hybridized in the presence of 30-fold excess driver by phenol–salt emulsion, followed by digestion with S1 nuclease to remove single-stranded moieties; 4. Resulting products are captured by streptavidin-coated paramagnetic beads and the tester homohybrids are release specifically by BamHI digestion; Steps 2–5 could be repeated to amplify the differences between the compared molecules.

stranded due to incomplete hybridization. Moreover, some species might be lost due to preferential amplification in the PCR step that supplies the starting material.

3.11 DNA Enrichment by Allele-Specific Hybridization

DNA enrichment by allele-specific hybridization (DEASH) is an interesting and important approach to subtraction application that detects small differences that are not the result of differential gene or the presence of extraneous genetic material is [150]. Such small differences might arise from low-frequency base substitutions, haplotypes, SNPs, sequence variants, and recombinant molecules that represent rare mutations or pathological recombination events. dsDNA that contains different alleles is mixed with biotinylated allele-specific oligonucleotide directed to a chosen variant and a nonbiotinylated competitor complementary to the other allele (Figure 14). The specificity of the hybridization

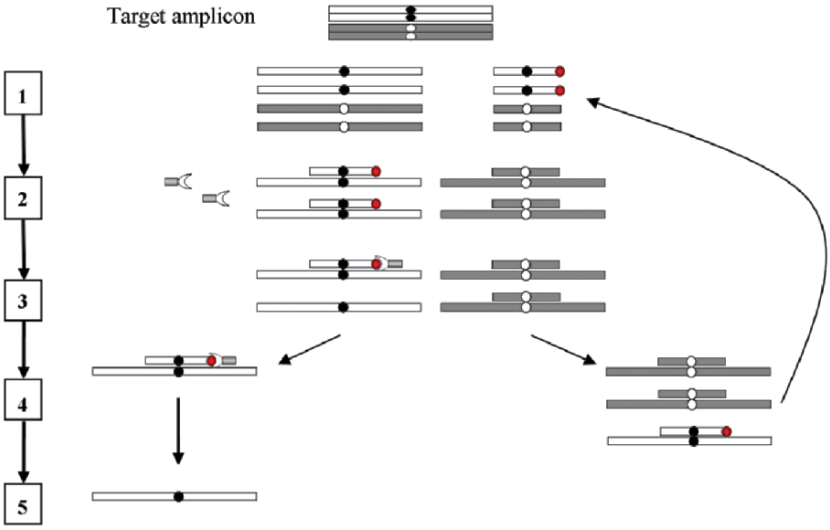

Figure 14. DNA Enrichment by allele-specific hybridization (DEASH). 1. Target amplicon (DNA, cDNA, or PCR representation in white and gray rectangle) that contains mixture of alleles differing by a chosen base substitution (white circle and black circle) is mixed with biotinylated allele-specific oligonucleotide (bio-ASO, white rectangle with black and grey circle) and the corresponding competitor ASO (gray rectangle with white circle) and denatured; 2. Streptavidin-coated magnetic beads (white crescent with gray bar) are added to the mix to retrieve the target amplicons; 3. Magnetic beads are recovered form the mixture with a magnet; 4. The biotin-captured amplicons/alleles are eluted from the beads; 5. Remaining nucleic acid mix can enter another round of enrichment with bio-ASO.

might in so doing, be improved. The hybrids are captured on streptavidin-coated paramagnetic beads and then thermally eluted in low salt buffer. The protocol can be repeated for several rounds and has the potential to detect rare variants. Its accuracy is unbiased by amplification since the separation occurs prior to PCR. It is, however, not suitable for detection of unknown changes since it requires preliminary information about the sequence of interest in order to design the specific sequence enrichment probes.

3.12 Methods Combining the use of SSH and Microarrays

The microarray approach was introduced more than 10 years ago [151]. In essence the microarray approach is a visualized subtractive hybridization method with an added internal control that does not physically remove common molecular species and can query any desired set of messages depending on the experimental goal. This method became particularly powerful with the improvement and standardization of array printing and of microhybridization methods, and with the acquisition of complete genome sequences of many medically and

industrially important organisms, including the human genome [152,41]. The method still requires substantial investment, and even though it was quickly commercialized, it has generated consistency and reproducibility issues. Another important factor that restricted the use of such data acquisition method relates to data analysis, because the ability to acquire data is exceeding the computational power required to extract comprehensive information [153]. Fortunately, there have been rapid improvements in the data analysis fields [139,154,155,156,157,158]. A combination of SSH and microarray approach [24] is certainly a streamlined way to differential gene expression profiling that can ensure that the data could be complete and manageable at the same time. This has already been done in breast cancer studies [21], virus–host interaction studies [22], and plant genetics [159]. The microarray approach to gene discovery has been compared to SSH [160] and DD [161], and the assessment was that these methods can be used as an alternative and/or complimentary transcript profiling tool, especially when the targets are new genes and transcripts of low abundance.

Finally, several methods have been developed to address the issue of quality control of the subtractive process, and the integrity and quality of the starting material. Regardless of the downstream protocol (SH, RDA, SSH, DEASH, etc.,) an essential requirement for the success of any method designed to look at differences is to provide an unbiased representation of the starting material. A number of control approaches [162,163,2,164,165,116,1,166,167,168,134] have been developed to ensure the quality of the starting material are found in references, and they may be applied discriminately depending on the application and the target of interest.

3.13 Conclusions

PEER is a new member of the subtraction hybridization-based methods designed to query genetic differences. The major advantages of PEER are that it takes direct advantage of the unique specificity of the target's sequence by using it both as a template and a primer. The blocking of common primers is an extra step that adds hybridization and enzymatic specificity to the selection of oligonucleotides that enter the final amplification process. PEER is versatile and not restricted to DNA or RNA starting material and does not require excess of driver material since the hybridization/exclusion step is done by PCR. The sensitivity and applicability of the method to both RNA and DNA targets was recently demonstrated with a variety of viruses representing different genome structures – human herpesvirus and Ectromelia virus (dsDNA), human echovirus, West Nile virus, and human respiratory syncytial virus (ssRNA), Orthoreovirus (polysegmented ds RNA), porcine circovirus (circular ssDNA) [45].

Most existing subtraction methods suffer from the inability to completely remove the background of common or highly repetitive sequences, are limited in the recovery of rare molecular species, and may omit single-stranded molecules.

Generally more advantageous are the hybridization methods that have their sensitivity and specificity enhanced by enzymatic steps and that could be coupled with good controls for the background. Standardized "chip" assays that combine subtraction and microarray approaches will clearly dominate in the near future, however, methods that are based on RDA, SSH, or DD have become "classical" and are widely used, have good versatility and general reproducibility, have been well standardizes and cannot be dismissed when screening for genetic differences at any level of complexity.

REFERENCES

1. Muerhoff AS, Leary TP, Desai SM, Mushahwar IK (1997) Amplification and subtraction methods and their application to the discovery of novel human viruses. J Med Virol 53:96–103
2. Lisitsyn N, Lisitsyn N, Wigler M (1993) Cloning the differences between two complex genomes. Science 259:946–951
3. Allander T, Emerson SU, Engle RE, Purcell RH, 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 USA 98:11609–11614
4. Birkenmeyer LG, Leary TP, Muerhoff AS, Dawson GJ, Mushahwar IK, Desai SM (2003) Selectively primed adaptive driver RDA (SPAD-RDA): an improved method for subtractive hybridization. J Med Virol 71:150–159
5. Luo J-H, Puc JA, Slosberg ED, Yao Y, Bruce JN, WrightJr TC, Becich MJ, Parsons R (1999) Differntial subtraction chain a method for identifying differences in genomic DNA and mRNA. Nucleic Acids Res 27:e24
6. Lavery DJ, Lopez-Molina L, Fleury-Olela F, Schibler U (1997) Selective amplification via biotin- and restriction-mediated enrichment (SABRE) a novel selective amplification procedure for detection of differentially expressed mRNAs. Proc Natl Acad Sci USA 94:6831–6836
7. Diatchenko L, Lau YF, Campbell AP, Chenchik A, Moqadam 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
8. Liang P, Pardee AB (1992) Differential display of eucaryotic messanger RNA by means of the polymerase chain reaction. Science 257:967–971
9. Matz MV, Lukyanov SA (1998) Different strategies of differential display: areas of application. Nucleic Acids Res 26:5537–5543
10. Shaw-Smith CJ, Coffey AJ, Huckle E, Durham J, Campbell EA, Freeman TC, Walters JR, Bentley DR (2000) Improved method for detecting differentially expressed genes using cDNA indexing. Biotechniques 28:958–964
11. Agaton C, Unneberg P, Sievertzon M, Holmberg A, Ehn M, Larsson M, Odeberg J, Uhlen M, Lundeberg J (2002) Gene expression analysis by signature pyrosequencing. Gene 289:31–39
12. Bishop R, Shah T, Pelle R, Hoyle D, Pearson T, Haines L, Brass A, Hulme H, Graham SP, Taracha EL, Kanga S, Lu C, Hass B, Wortman J, White O, Gardner MJ, Nene V, de Villiers EP (2005) Analysis of the transcriptome of the protozoan *Theileria parva* using MPSS reveals that the majority of genes are transcriptionally active in the schizont stage. Nucleic Acids Res 33:5503–5511
13. Brenner S, Johnson M, Bridgham J, Golda G, Lloyd DH, Johnson D, Luo S, McCurdy S, Foy M, Ewan M, Bridgham J, Golda G, Lloyd DH, Johnson D, Luo S, McCurdy S, Foy M, Ewan M, Roth R, George D, Eletr S, Albrecht G, Vermaas E, Williams SR, Moon K, Burcham T, Pallas M, DuBridge RB, Kirchner J, Fearon K, Mao J, Corcoran K (2000) Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. Nat Biotechnol 18:630–634

14. Brenner S, Williams SR, Vermaas EH, Storck T, Moon K, McCollum C, Mao JI, Luo S, Kirchner JJ, EletrSEletr S, DuBridge RB, Burcham T, Albrecht G (2000) *In vitro* cloning of complex mixtures of DNA on microbeads: physical separation of differentially expressed cDNAs. Proc Natl Acad Sci USA 97:1665–1670
15. Chen JJ, WuR, Yang PC, Huang JY, Sher YP, Han MH, Kao WC, Lee PJ, Chiu TF, Chang F, Chu YW, Wu CW, Peck K (1998) Profiling expression patterns and isolating differentially expressed genes by cDNA microarray system with colorimetry detection. Genomics 51:313–324
16. Drmanac S, Drmanac R (1994) Processing of cDNA and genomic kilobase-size clones for massive screening mapping and sequencing by hybridization. Biotechniques 17:328–9–332–6
17. Leamon JH, Lee WL, Tartaro KR, Lanza JR, Sarkis GJ, deWinter AD, Berka J, Lohman KL (2003) A massively parallel PicoTiterPlate based platform for discrete picoliter-scale polymerase chain reactions. Electrophoresis 24:3769–3777
18. Schena M, Heller RA, Theriault TP, Konrad K, Lachenmeier E, Davis RW (1999) Microarrays: biotechnology's discovery platform for functional genomics. Trends Biotechnol 16:217–218
19. Wang D, Urisman A, LiuYT, SpringerM, Ksiazek TG, Erdman DD, Mardis ER, Hickenbotham M, Magrini V, Eldred J, Latreille JP, Wilson RK, Ganem D, DeRisi JL (2003) Viral discovery and sequence recovery using DNA microarrays. PLoS Biol 1:E2
20. Wieland I, Bolger G, Asouline G, Wigler M (1990) A method for difference cloning: gene amplification following subtractive hybridization. Proc Natl Acad Sci USA 87:2720–2724
21. Beck MT, Holle L, Chen WY (2001) Combination of PCR subtraction and cDNA microarray for differential gene expression profiling. Biotechniques 31:782–4, 786
22. Munir S, Singh S, Kaur K, Kapur V (2004) Suppression subtractive hybridization coupled with microarray analysis to examine differential expression of genes in virus infected cells. Biol Proceed Online 6:94–104
23. Pan YS, Lee YS, Lee YL, Lee WC, Hsieh SY (2006) Differentially profiling the low-expression transcriptomes of human hepatoma using a novel SSH/microarray approach. BMC Genomics 7:131
24. Yang GP, Ross DT, Kuang WW, Brown PO, Weigel RJ (1999) Combining SSH and cDNA microarrays for rapid identification of differentially expressed genes. Nucleic Acids Res 27:1517–1523
25. Bui LC, Leandri RD, Renard JP, Duranthon V (2005) SSH adequacy to preimplantation mammalian development: scarce specific transcripts cloning despite irregular normalization BMC Genomics 6:155
26. Fallsehr C, Zapletal C, Kremer M, Demir R, von Knebel DM, Klar E (2005) Identification of differentially expressed genes after partial rat liver ischemia/reperfusion by suppression subtractive hybridization. World J Gastroenterol 11:1303–1316
27. Galbraith EA, Antonopoulos DA, White BA (2004) Suppressive subtractive hybridization as a tool for identifying genetic diversity in an environmental metagenome: the rumen as a model. Environ Microbiol 6:928–937
28. Liu YB, Wei ZX, Li L, Li HS, Chen H, Li XW (2003) Construction and analysis of SSH cDNA library of human vascular endothelial cells related to gastrocarcinoma. World J Gastroenterol 9:2419–2423
29. Nishizuka S, Tsujimoto H, Stanbridge EJ (2001) Detection of differentially expressed genes in HeLa x fibroblast hybrids using subtractive suppression hybridization. Cancer Res 61:4536–4540
30. Ganova-Raeva L, Smith AW, Fields H, Khudyakov Y (2004) New Calicivirus isolated from walrus. Virus Res 102:207–213
31. Birkenmeyer LG, Desai SM, Muerhoff AS, Leary TP, Simons JN, Montes CC, Mushahwar IK (1998) Isolation of a GB virus-related genome from a chimpanzee. J Med Virol 56:44–51
32. Nishizawa T, Okamoto H, Konishi K, Yoshizawa H, Miyakawa Y, Mayumi M (1997) A novel DNA virus (TTV) associated with elevated transaminase levels in posttransfusion hepatitis of unknown etiology. Biochem Biophys Res Commun 241:92–97
33. Ji W, right MB, Cai L, Flament, Lindpaintner K (2002) Efficacy of SSH PCR in isolating differentially expressed genes. BMC Genomics 3:12

34. Szybalsky W, Kim SC, Hasan N, Podhajska AJ (1991) Class-IIS restriction enzymes – a review. Gene 100:13–26
35. Boyd AC, Charles IG, Keyte JW, Brammar WJ (1986) Isolation and computer-aided characterization of MmeI a type II restriction endonuclease from *Methylophilus methylotrophus*. Nucleic Acids Res 14:5255–5274
36. Tabor S, Richardson CC (1995) A single residue in DNA polymerase of the *E. coli* DNA polymerase I family is critical for distinguishing between deoxy and dideoxyribonucleotides. Proc Natl Acad Sci USA 92:6339–6343
37. Kotewicz ML, Sampson CM, D'Alessio JM, Gerard GF (1988) Isolation of cloned Moloney murine leukemia virus reverse transcriptase lacking ribonuclease H activity. Nucleic Acids Res 16:265–277
38. Seth D, Gorrell MD, McGuinness PH, Leo MA, Lieber CS, McCaughan GW, Haber PS (2003) SMART amplification maintains representation of relative gene expression: quantitative validation by real time PCR and application to studies of alcoholic liver disease in primate. J Biochem Biophys Methods 55:53–66
39. Swaminathan N, George D, McMaster K, Szablebski J, VanEtten JL, Mead DA (1994) Restriction generated oligonucleotides utilizing the two base recognition endonuclease CviJI*. Nucleic Acids Res 22:1470–1475
40. Swaminathan N, McMaster K, Skowron PM, Mead DA (1998) Thermal cycle labeling: Zeptomole detection sensitivity and microgram probe amplification using CviJI* restriction-generated oligonucleotides. Anal Biochem 255:133–141
41. Wright FA, Lemon WJ, Zhao WD, Sears R, Zhuo D, Wang JP, Yang HY, Baer T, Stredney D, Spitzner J, Stutz A, Krahe R, Yuan B (2001) A draft annotation and overview of the human genome. Genome Biol 2(7):1–25
42. Tucholski J, Skowron PM, Podhajska AJ (1995) MmeI a class-IIS restriction endonuclease: purification and characterization. Gene 157:87–92
43. Armengaud J, Jouanneau Y(1993) Addition of a class IIS enzyme site in the mutagenic primer to improve two-step PCR-based targeted mutagenesis. Nucleic Acids Res 21:4424–4425
44. Ganova-Raeva L, Zhang X, Cao F, Fields H, Khudyakov Y (2006) Primer extension enrichment reaction (PEER): a new subtraction method for identification of genetic differences between biological specimens. Nucleic Acids Res 34:e76
45. Biagini P, de Lamballerie X, de Micco P (2007) Effective detection of highly divergent viral genomes in infected cell lines using a new subtraction strategy (primer extension enrichment reaction – PEER). J Virol Methods 139(1):106–110
46. Sargent TD (1983) Differential gene expression in the gastrula of *Xenopus laevis*. Science 222:135–139
47. Maser RL (1995) Analysis of differential gene expression in the kidney by differential cDNA screening subtractive cloning and mRNA differential display. Semin Nephrol 15:29–42
48. Lee MT, Kaushansky K, Ralph P, Ladner MB (1990) Differential expression of M-CSF G-CSF and GM-CSF by human monocytes. J Leukoc Biol 47:275–282
49. Rothstein JL, Johnson D, Jessee J, Skowronski J, DeLoia JA, Solter D, Knowles BB (1993) Construction of primary and subtracted cDNA libraries from early embryos. Methods Enzymol 225:587–610
50. Sargent T (1987) Isolation of differentially expressed genes. Methods Enzymol 152:423–432
51. Ermolaeva OD, Sverdlov ED (1996) Subtractive hybridization a technique for extraction of DNA sequences distinguishing two closely related genomes: critical analysis. Genet Anal 13:49–58
52. Davis MM (1984) Cell-type-specific cDNA probes and the murine I region: the localization and orientation of Ad alpha. Proc Natl Acad Sci USA 81:2194–2198
53. Rubenstein JL, Brice AE, Ciaranello RD, Denney D, Porteus MH, Usdin TB (1990) Subtractive hybridization system using single-stranded phagemids with directional inserts. Nucleic Acids Res 18:4833–4842

54. Schweinfest CW, Henderson KW, Gu JR, Kottaridis SD, Besbeas S, Panotopoulou E, Papas TS (1990) Subtraction hybridization cDNA libraries from colon carcinoma and hepatic cancer. Genet Anal Tech Appl 7:64–70
55. Scott MR, Westphal K-H, Rigby PWJ (1983) Activation of mouse genes in transformed cells. Cell 34:557–567
56. Reynet C, Kahn CR (1993) Rad: a member of the Ras family overexpressed in muscle of type II diabetic humans. Science 262;1441–1444
57. Ermolaeva OD, Lukyanov SA, Sverdlov ED (1996) The mathematical model of subtractive hybridization and its practical application. Proc Int Conf Intell Syst Mol Biol 4:52–58
58. Milner JJ, Cecchini E, Dominy PJ (1995) A kinetic model for subtractive hybridization. Nucleic Acids Res 23:176–187
59. Ray A, Macwana S, Ayoubi P, Hall LT, Prade R, Mort AJ (2004) Negative subtraction hybridization: an efficient method to isolate large numbers of condition-specific cDNAs. BMC Genomics 5:22
60. Lee SW (1991) Positive selection of candidate tumor-suppressor genes by subtractive hybridization. Proc Natl Acad Sci USA 88:2825–2829
61. Konietzko U, Kuhl D (1998) A subtractive hybridisation method for the enrichment of moderately induced sequences. Nucleic Acids Res 26:1359–1361
62. Li MS, Monahan IM, Waddell SJ, Mangan JA, Martin SL, Everett MJ, Butcher PD (2001) cDNA-RNA subtractive hybridization reveals increased expression of mycocerosic acid synthase in intracellular *Mycobacterium bovis* BCG. Microbiology 147:2293–2305
63. Ishino F, Kuroiwa Y, Miyoshi N, Kobayashi S, Kohda T, Kaneko-Ishino T (2001) Subtraction-hybridization method for the identification of imprinted genes. Methods Mol Biol 181:101–112
64. Sharma S, Chang JT, Della NG, Campochiaro PA, Zack DJ (2002) Identification of novel bovine RPE and retinal genes by subtractive hybridization. Mol Vis 8:251–258
65. Zimmermann CR (1980) Molecular cloning and selection of genes regulated in Aspergillus development. Cell 21:709–715
66. Hara E, Kato T, Nakada S, Sekiya S, Oda K (1991) Subtractive cDNA cloning using oligo(dT)30-latex and PCR: isolation of cDNA clones specific to undifferentiated human embryonal carcinoma cells. Nucleic Acids Res 19:7097–7104
67. Kohchi T, Fujishige K, Ohyama K (1995) Construction of an equalized cDNA library from *Arabidopsis thaliana*. Plant J 8:771–776
68. Lopez-Fernandez LA, del Mazo J (1993) Construction of subtractive cDNA libraries from limited amounts of mRNA and multiple cycles of subtraction. Biotechnique 15:654–659
69. Laveder P, De Pitta C, Toppo S, Valle G, Lanfranchi G (2002) A two-step strategy for constructing specifically self-subtracted cDNA libraries. Nucleic Acids Res 30:e38
70. Ying SY, Lin SL (2003) Subtractive cloning of differential genes using RNA-PCR. Methods Mol Biol 221:253–259
71. Zhumabayeva B, Chang C, McKinley J, Diatchenko L, Siebert PD (2001) Generation of full-length cDNA libraries enriched for differentially expressed genes for functional genomics. Biotechniques 30:512–520
72. Zhumabayeva B, Chang C, McKinley J, Diatchenko L, Siebert PD (2003) Generation of full-length cDNA libraries enriched for differentially expressed genes. Methods Mol Biol 221:223–237
73. Welsh J, Chada K, Dalal SS, Cheng R, Ralph D, McClelland M (1992) Arbitrarily primed PCR fingerprinting of RNA. Nucleic Acids Res 20:4965–4970
74. Fuchs B, Zhang K, Bolander ME, Sarkar G (2000) Differential mRNA fingerprinting by preferential amplification of coding sequences. Gene 258:155–163
75. Kato K (1995) Description of the entire mRNA population by a 3′ end cDNA fragment generated by class IIS restriction enzymes. Nucleic Acids Res 23:3685–3690
76. Kato K (1996) RNA fingerprinting by molecular indexing. Nucleic Acids Res 24:394–395
77. Matz M, Usman N, Shagin D, Bogdanova E, Lukyanov S (1997) Ordered differential display: a simple method for systematic comparison of gene expression profiles. Nucleic Acids Res 25:2541–2542

78. Zabeau M, Vos P (1993) Selective restriction fragment amplification: a general method for DNA fingerprinting. Eur Patent Appl Publ #0534858A1
79. Pardinas JR, Combates NJ, Prouty SM, Stenn KS, Parimoo S (1998) Differential subtraction display: a unified approach for isolation of cDNAs from differentially expressed genes. Anal Biochem 257:161–168
80. SAGE (2006) Current technologies and applications. Horizon Bioscience, Norwich, England
81. Dinel S, Bolduc C, Belleau P, Boivin A, Yoshioka M, Calvo E, Piedboeuf B, Snyder EE, Labrie F, St Amand J (2005) Reproducibility bioinformatic analysis and power of the SAGE method to evaluate changes in transcriptome. Nucleic Acids Res 33:e26
82. Vos P, Hogers R, Bleeker M, Reijans M van de LT, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414
83. Botstein D (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314–331
84. Akopyants NS, Fradkov A, Diatchenko L, Hill JE, Siebert PD, Lukyanov SA, Sverdlov ED, Berg DE (1998) PCR-based subtractive hybridization and differences in gene content among strains of *Helicobacter pylori*. Proc Natl Acad Sci USA 95:13108–13113
85. Kim S, Park S, Choung S, Park HO, Choi YC (2005) DARFA: a novel technique for studying differential gene expression and bacterial comparative genomics. Biochem Biophys Res Commun 336:168–174
86. Velculescu VE, Zhang L, Vogelstein B, Kinzler KW (1995) Serial analysis of gene expression. Science 270:484–487
87. Gnatenko DV, Dunn JJ, McCorkle SR, Weissmann D, Perrotta PL, Bahou WF (2003) Transcript profiling of human platelets using microarray and serial analysis of gene expression. Blood 101:2285–2293
88. Huang ZG, Ran ZH, Lu W, Xiao SD (2006) Analysis of gene expression profile in colon cancer using the Cancer Genome Anatomy Project and RNA interference. Chin J Dig Dis 7:97–102
89. Wang X, Zhang C, Zhang L, Wang X, Xu S (2006) High-throughput assay of DNA methylation based on methylation-specific primer and SAGE. Biochem Biophys Res Commun 341:749–754
90. Shiraki T (2003) Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. Proc Natl Acad Sci USA 100:15776–15781
91. Dunn JJ, McCorkle SR, Praissman LA, Hind G, Van Der LD, Bahou WF, Gnatenko DV, Krause MK (2002) Genomic signature tags (GSTs): a system for profiling genomic DNA. Genome Res 12:1756–1765
92. Chen J, Sadowski I (2005) Identification of the mismatch repair genes PMS2 and MLH1 as p53 target genes by using serial analysis of binding elements. Proc Natl Acad Sci USA 102:4813–4818
93. Lisitsyn NA (1995) Representational difference analysis: finding the differences between genomes. Trends Genet 11:303–307
94. Felske A (2002) Streamlined representational difference analysis for comprehensive studies of numerous genomes. J Microbiol Methods 50:305–311
95. Jiang H (2000) RaSH a rapid subtraction hybridization approach for identifying and cloning differentially expressed genes. Proc Natl Acad Sci USA 97:12684–12689
96. Bowler LD (2004) Representational difference analysis of cDNA. Methods Mol Med 94:49–66
97. Hubank M, Schatz DG (1999) cDNA representational difference analysis: a sensitive and flexible method for identification of differentially expressed genes. Methods Enzymol 303:325–349
98. Pastorian K, Hawel L III, Byus CV (2000) Optimization of cDNA representational difference analysis for the identification of differentially expressed mRNAs. Anal Biochem 283:89–98
99. Muerhoff AS, Leary TP, Simons JN, Pilot-Matias TJ, Dawson GJ, Erker JC, Chalmers ML, Schlauder GG, Desai SM, Mushahwar IK (1995) Genomic organization of GB viruses A and B: two new members of the Flaviviridae associated with GB agent hepatitis. J Virol 69:5621–5630

100. Simons JN, Pilot-Matias TJ, Leary TP, Dawson GJ, Desai SM, Schlauder GG, Muerhoff AS, Erker JC, Buijk SL, Chalmers ML (1995) Identification of two flavivirus-like genomes in the GB hepatitis agent. Proc Natl Acad Sci USA 92:3401–3405
101. Hubank M, Bryntesson F, Regan J, Schatz DG (2004) Cloning of apoptosis-related genes by representational difference analysis of cDNA. Methods Mol Biol 282:255–273
102. Bowler LD, Hubank M, Spratt BG (1999) Representational difference analysis of cDNA for the detection of differential gene expression in bacteria: development using a model of iron-regulated gene expression in *Neisseria meningitidis*. Microbiology 145:3529–3537
103. Hollestelle A, Schutte M (2005) Representational difference analysis as a tool in the search for new tumor suppressor genes. Methods Mol Med 103:143–159
104. Boukerche H, Su ZZ, Kang DC, Fisher PB (2004) Identification and cloning of genes displaying elevated expression as a consequence of metastatic progression in human melanoma cells by rapid subtraction hybridization. Gene 343:191–201
105. Wada J, Kumar A, Ota K, Wallner EI, Batlle DC, Kanwar YS (1997) Representational difference analysis of cDNA of genes expressed in embryonic kidney. Kidney Int 51:1629–1638
106. Hansen-Hagge TE, Trefzer U, zu Reventlow AS, Kaltoft K, Sterry W (2001) Identification of sample-specific sequences in mammalian cDNA and genomic DNA by the novel ligation-mediated subtraction (Limes). Nucleic Acids Res 29: E20
107. Zeng J, Gorski RA, Hamer D (1994) Differential cDNA cloning by enzymatic degrading subtraction (EDS). Nucleic Acids Res 22:4381–4385
108. Kohne DE (1977) Room temperature method for increasing the rate of DNA reassociation by many thousandfold: the phenol emulsion reassociation technique. Biochemistry 16:5329–5341
109. Yang M (1996) Cloning differentially expressed genes by linker capture subtraction. Anal Biochem 237:109–114
110. Lin F (2001) Myopodin a synaptopodin homologue is frequently deleted in invasive prostate cancers. Am J Pathol 159:1603–1612
111. Yu YP (2001) Identification of a novel gene with increasing rate of suppression in high grade prostate cancers. Am J Pathol 158:19–24
112. Li J, Wang F, Kashuba V, Wahlestedt C, Zabarovsky ER (2001) Cloning of deleted sequences (CODE): A genomic subtraction method for enriching and cloning deleted sequences. Biotechniques 31:788,790,792–788,790,793
113. Lin TY, Ying SY (2003) Subtractive hybridization for the identification of differentially expressed genes using uracil-dNA glycosylase and mung-bean nuclease. Methods Mol Biol 221:239–251
114. Chalaya T, Gogvadze E, Buzdin A, Kovalskaya E, Sverdlov ED (2004) Improving specificity of DNA hybridization-based methods. Nucleic Acids Res 32:e130
115. Li L, Techel D, Gretz N, Hildebrandt A (2005) A novel transcriptome subtraction method for the detection of differentially expressed genes in highly complex eukaryotes. Nucleic Acids Res 33:e136
116. Matz M, Shagin D, Bogdanova E, Britanova O, Lukyanov S, Diatchenko L, Chenchik A (1999) Amplification of cDNA ends based on template-switching effect and step-out PCR. Nucleic Acids Res 27:1558–1560
117. Hu Y, Hirshfield I (2005) Rapid approach to identify an unrecognized viral agent. J Virol Methods 127:80–86
118. Hedrick SM (1984) Isolation of cDNA clones encoding T cell-specific membrane-associated proteins. Nature 308:149–153
119. Hubank M (1994) Identifying differences in mRNA expression by representational difference analysis of cDNA. Nucleic acids research 22:5640–5648
120. Wang Z (1991) A gene expression screen. Proc Natl Acad Sci USA 88:11505–11509
121. Lukyanov KA (1995) Inverted terminal repeats permit the average length of amplified DNA fragments to be regulated during preparation of cDNA libraries by polymerase chain reaction. Anal Biochem 229:198–202

122. Chua KB, Wang LF, Lam SK, Crameri G, Yu M, Wise T, Boyle D, Hyatt AD, Eaton BT (2001) Tioman virus a novel paramyxovirus isolated from fruit bats in Malaysia. Virology 283:215–229
123. DeShazer D (2004) Genomic diversity of *Burkholderia pseudomallei* clinical isolates: subtractive hybridization reveals a *Burkholderia mallei*-specific prophage in *B. pseudomallei* 1026b. J Bacteriol 186:3938–3950
124. Diatchenko L, Lukyanov S, Lau YF, Siebert PD (1999) Suppression subtractive hybridization: a versatile method for identifying differentially expressed genes. Methods Enzymol 303:349–380
125. Diatchenko L, Slade GD, Nackley AG, Bhalang K, Sigurdsson A, Belfer I, Goldman D, Xu K, Shabalina SA, Shagin D, Max MB, Makarov SS, Maixner W (2005) Genetic basis for individual variations in pain perception and the development of a chronic pain condition. Hum Mol Genet 14:135–143
126. Gurskaya NG, Diatchenko L, Chenchik A, Siebert PD, Khaspekov GL, Lukyanov KA, Vagner LL, Ermolaeva OD, Lukyanov SA, Sverdlov ED (1996) Equalizing cDNA subtraction based on selective suppression of polymerase chain reaction: cloning of Jurkat cell transcripts induced by phytohemaglutinin and phorbol 12-myristate 13-acetate. Anal Biochem 240:90–97
127. Bonaldo MF, Lennon G, Soares MB (1996) Normalization and subtraction: two approaches to facilitate gene discovery. Genome Res 6:791–806
128. Chenchik A (1996) Full-length cDNA cloning and determination of mRNA 5′ and 3′ ends by amplification of adaptor-ligated cDNA. Biotechniques 21:526–534
129. Lukyanov K, Diatchenko L, Chenchik A, Nanisetti A, Siebert P, Usman N, Matz M, Lukyanov S (1997) Construction of cDNA libraries from small amounts of total RNA using the suppression PCR effect. Biochem Biophys Res Commun 230:285–288
130. Siebert PD (1995) An improved PCR method for walking in uncloned genomic DNA. Nucleic Acids Res 23:1087–1088
131. Rebrikov DV, Britanova OV, Gurskaya NG, Lukyanov KA, Tarabykin VS, Lukyanov SA (2000) Mirror orientation selection (MOS): a method for eliminating false positive clones from libraries generated by suppression subtractive hybridization. Nucleic Acids Res 28:E90
132. Jeung JU, Cho SK, Shin JS (2005) A partial-complementary adapter for an improved and simplified ligation-mediated suppression PCR technique. J Biochem Biophys Methods 64:110–120
133. Strauss C, Mussgnug JH, Kruse O (2001) Ligation-mediated suppression-PCR as a powerful tool to analyse nuclear gene sequences in the green alga *Chlamydomonas reinhardtii*. Photosynth Res 70:311–320
134. Yueqing C, Zhengbo H, Zhongkang W, Youping Y, Guoxiong P, Yuxian X (2006) Hybridization monitor: A method for identifying differences between complex genomes. J Microbiol Methods 64:305–315
135. Boengler K, Pipp F, Schaper W, Deindl (2003) Rapid identification of differentially expressed genes by combination of SSH and MOS. Lab Invest 83:759–761
136. Rebrikov DV, Desai SM, Siebert PD, Lukyanov SA (2004) Suppression subtractive hybridization. Methods Mol Biol 258:107–134
137. Iguchi K, Takahashi Y, Kaneto Y, Kubota M, Usui S, Hirano K (2005) Identification of differentially expressed genes in hepatic HepG2 cells treated with acetaminophen using suppression subtractive hybridization. Biol Pharm Bull 28:1148–1153
138. Larose M, Bouchard C, Chagnon YC (2001) A new gene related to human obesity identified by suppression subtractive hybridization. Int J Obes Relat Metab Disord 25:770–776
139. Beissbarth T (2006) Interpreting experimental results using gene ontologies. Methods Enzymol 411:340–352
140. Malyala A, Pattee P, Nagalla SR, Kelly MJ, Ronnekleiv OK (2004) Suppression subtractive hybridization and microarray identification of estrogen-regulated hypothalamic genes. Neurochem Res 29:1189–1200
141. Rebrikov D, Desai S, Kogan YN, Thornton AM, Diatchenko L (2002) Subtractive cloning: new genes for studying inflammatory disorders. Ann Periodontol 7:17–28

142. Shackel NA, McGuinness PH, Abbott CA, Gorrell MD, McCaughan GW (2003) Novel differential gene expression in human cirrhosis detected by suppression subtractive hybridization. Hepatology 38:577–588
143. Yokota N, Mainprize TG, Taylor MD, Kohata T, Loreto M, Ueda S Dura W, Grajkowska W, Kuo JS, Rutka JT (2004) Identification of differentially expressed and developmentally regulated genes in medulloblastoma using suppression subtraction hybridization. Oncogene 23:3444–3453
144. Wang X, Feuerstein GZ (2000) Suppression subtractive hybridization: application in the discovery of novel pharmacological targets. Pharmacogenomics 1:101–108
145. Bai GQ, Liu Y, Cheng J, Zhang SL, Yue YF, Huang YP, Zhang LY (2005) Transactivating effect of complete S protein of hepatitis B virus and cloning of genes transactivated by complete S protein using suppression subtractive hybridization technique. World J Gastroenterol 11:3893–3898
146. He N, Qin Q, Xu X (2005) Differential profile of genes expressed in hemocytes of White Spot Syndrome Virus-resistant shrimp (*Penaeus japonicus*) by combining suppression subtractive hybridization and differential hybridization. Antiviral Res 66:39–45
147. Kiss C, Nishikawa J, Dieckmann A, Takada K, Klein G, Szekely L (2003) Improved subtractive suppression hybridization combined with high density cDNA array screening identifies differentially expressed viral and cellular genes. J Virol Methods 107:195–203
148. Yin J, Chen MF, Finkel TH (2004) Differential gene expression during HIV-1 infection analyzed by suppression subtractive hybridization. AIDS 18:587–596
149. Schibler U, Rifat D, Lavery DJ (2001) The Isolation of differentially expressed mRNA sequences by selective amplification via biotin and restriction-mediated enrichment. Methods 24:3–14
150. Jeffreys AJ, May CA (2003) DNA enrichment by allele-specific hybridization (DEASH): a novel method for haplotyping and for detecting low-frequency base substitutional variants and recombinant DNA molecules. Genome Res 13:2316–2324
151. Schena M, Shalon D, Davis RW, Brown PO (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270:467–470
152. Schena M, Shalon D, Heller R, Chai A, Brown PO, Davis RW (1996) Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. Proc Natl Acad Sci USA 93:10614–10619
153. Mutch DM, Berger A, Mansourian R, Rytz A, Roberts MA (2002) The limit fold change model: a practical approach for selecting differentially expressed genes from microarray data. BMC Bioinformatics 3:17
154. Gresham D (2006) Genome-wide detection of polymorphisms at nucleotide resolution with a single DNA microarray. Science 311:1932–1936
155. Roussel EE, Gingra MM-C (2006) High-throughput gene expression profiling—a work-in-progress with great potential for proteomics. Curr Opin Drug Discov Devel 9:332–338
156. Rubinstein R, Simon I (2005) MILANO–custom annotation of microarray results using automatic literature searches. BMC Bioinformatics 6:12
157. Semeiks JR, Rizki A, Bissell MJ, Mian IS (2006) Ensemble attribute profile clustering: discovering and characterizing groups of genes with similar patterns of biological features. BMC Bioinformatics 7:147
158. Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF Zeeberg B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T, Max SI, Wang J, Hsieh F, Diatchenko L, Marusina K, Farmer AA, Rubin GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL, Scheetz TE, Brownstein MJ, Usdin TB, Toshiyuki S, Carninci P, Prange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mullahy SJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JA, Gunaratne PH, Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW, Villalon DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J, Helton E, Ketteman M, Madan A, Rodrigues S, Sanchez A, Whiting M, Madan A, Young AC, Shevchenko Y,

Bouffard GG, Blakesley RW, Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J, Schmutz J, Myers RM, Butterfield YS, Krzywinski MI, Skalska U, Smailus DE, Schnerch A, Schein JE, Jones SJ, Marra MA (2002) Generation and initial analysis of more than 15000 full-length human and mouse cDNA sequences. Proc Natl Acad Sci USA 99:16899–16903

159. van den Berg N, Crampton BG, Hein I, Birch PR, Berger DK (2004) High-throughput screening of suppression subtractive hybridization cDNA libraries using DNA microarray analysis. Biotechniques 37:818–824
160. Cao W, Epstein C, Liu H, DeLoughery C, Ge N, Lin J, Diao R, Cao H, Long F, Zhang X, Zhang X, Chen Y, Wright PS, Busch S, Wenck M, Wong K, Saltzman AG, Tang Z, Liu L, Zilberstein A (2004) Comparing gene discovery from Affymetrix GeneChip microarrays and Clontech PCR-select cDNA subtraction: a case study. BMC Genomics 5:26
161. Rajeevan MS, Ranamukhaarachchi DG, Vernon SD, Unger ER (2001) Use of real-time quantitative PCR to validate the results of cDNA array and differential display PCR technologies. Methods 25:443–451
162. Gastel JA, Sutter TR (1996) A control system for cDNA enrichment reactions. Biotechniques 20:870–875
163. Gonzalez JM, Portillo MC, Saiz-Jimenez C (2005) Multiple displacement amplification as a pre-polymerase chain reaction (pre-PCR) to process difficult to amplify samples and low copy number sequences from natural environments. Environ Microbiol 7:1024–1028
164. Lukyanov KA, Matz MV, Bogdanova EA, Gurskaya NG, Lukyanov SA (1996) Molecule by molecule PCR amplification of complex DNA mixtures for direct sequencing: an approach to in vitro cloning. Nucleic Acids Res 24:2194–2195
165. Makrigiorgos GM, Chakrabarti S, Zhang Y, Kaur M, Price BD (2002) A PCR-based amplification method retaining the quantitative difference between two complex genomes. Nat Biotechnol 20:936–939
166. Reyes GR, Kim JP (1991) Sequence-independent single-primer amplification (SISPA) of complex DNA populations. Mol Cell Probes 5:473–481
167. Soares MB, Bonaldo MF, 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
168. Ying SY (2004) Complementary DNA libraries: an overview. Mol Biotechnol 27:245–252