Reciprocal subtraction differential RNA display: An efficient and rapid procedure for isolating differentially expressed gene sequences

(differential gene expression/subtraction display/reverse Northern blotting/progression elevated genes/progression suppressed genes)

DONG-CHUL KANG, RAPHAEL LAFRANCE, ZAO-ZHONG SU, AND PAUL B. FISHER*

Departments of Neurosurgery, Pathology and Urology, Herbert Irving Comprehensive Cancer Center, Columbia University, College of Physicians and Surgeons, 630 West 168th Street, New York, NY 10032

ABSTRACT A reciprocal subtraction differential RNA display (RSDD) approach has been developed that permits the rapid and efficient identification and cloning of both abundant and rare differentially expressed genes. RSDD comprises reciprocal subtraction of cDNA libraries followed by differential RNA display. The RSDD strategy was applied to analyze the gene expression alterations resulting during cancer progression as adenovirus-transformed rodent cells developed an aggressive transformed state, as documented by elevated anchorage-independence and enhanced *in vivo* oncogenesis in nude mice. This approach resulted in the identification and cloning of both known and a high proportion (>65%) of unknown sequences, including cDNAs displaying elevated expression as a function of progression (progression-elevated gene) and cDNAs displaying suppressed expression as a function of progression (progression-suppressed gene). Sixteen differentially expressed genes, including five unknown progression-elevated genes and six unknown progression-suppressed genes, have been characterized. The RSDD scheme should find wide application for the effective detection and isolation of differentially expressed genes.

Changes in gene expression are important determinants of normal cellular physiology, including cell cycle regulation, differentiation, and development, and they directly contribute to abnormal cellular physiology, including developmental anomalies, aberrant programs of differentiation, and cancer (1–4). In these contexts, the identification, cloning, and characterization of differentially expressed genes should provide relevant and important insights into the molecular determinants of processes such as growth, development, aging, differentiation, and cancer. A number of procedures can be used to identify and clone differentially expressed genes. These include subtractive hybridization (5–10), differential RNA display (DDRT-PCR) (1, 2, 11, 12), RNA fingerprinting by arbitrarily primed PCR (13, 14), representational difference analysis (15), serial analysis of gene expression (16, 17), electronic subtraction (18, 19), and combinatorial gene matrix analysis (20).

The DDRT-PCR approach developed by Liang and Pardee (11) has gained wide popularity in analyzing and cloning differentially expressed genes. DDRT-PCR is a powerful methodology in which a vast number of mRNA species (>20,000, if no redundancy occurs) can be analyzed with only a small quantity of RNA ($\approx$5 μg) (11). DDRT-PCR is often the method of choice when the RNA source is limiting, such as tissue biopsies. A direct advantage of DDRT-PCR is the ability to identify and isolate both up- and down-regulated differentially expressed genes in the same reaction. Furthermore, the DDRT-PCR technique permits the display of multiple samples in the same gel, which is useful in defining specific diagnostic alterations in RNA species and for temporally analyzing gene expression changes. However, the DDRT-PCR technique is not problem-free (21). Difficulties encountered when using standard DDRT-PCR include a high incidence of false positives and redundant gene identification, poor reproducibility, biased gene display, and lack of functional information about the cloned cDNA. Furthermore, poor separation can mask differentially expressed genes of low abundance under the intense signals generated by highly expressed genes. The generation of false positives and redundancy can be highly problematic, resulting in an inordinate expenditure of resources to confirm appropriate differential expression and uniqueness of the isolated cDNAs. The cDNAs must be isolated from the gels in pure form (contamination of bands with multiple sequences complicates clone identification), reamplified, placed in an appropriate cloning vector, analyzed for authentic differential expression, and, finally, sequenced. These limitations of the standard DDRT-PCR approaches emphasize the need for improvements in this procedure to more efficiently and selectively identify differentially expressed genes.

Subtractive hybridization, in which hybridization between tester and driver is followed by selective removal of common gene products, enriches for unique gene products in the tester cDNA population and reduces the abundance of common cDNAs (7). A subtracted cDNA library can be analyzed to identify and clone differentially expressed genes by randomly picking colonies or by differential screening (22–24). Although subtractive hybridization has been used successfully to clone a number of differentially expressed genes (5, 6, 8, 9), this approach is labor-intensive and does not result in isolation of the full spectrum of genes displaying altered expression (7, 18).

In principle, DDRT-PCR performed with subtracted RNA or cDNA samples should provide a powerful strategy to clone up- and down-regulated gene products. This approach should combine the benefits of both techniques, resulting in the enrichment of unique sequences and a reduction or elimination of common sequences. This scheme also should result in a consistent reduction in band complexity on a display gel, thereby permitting a clearer separation of cDNAs, resulting in fewer false positive reactions. Additionally, it should be possible to use fewer primer sets for reverse transcription and PCR reactions to analyze the complete spectrum of differentially expressed genes. Of particular importance for gene identification and isolation, rare gene

---

products that are masked by strong common gene products should be displayed by using subtraction hybridization in combination with DDRT-PCR. In addition, the DDRT-PCR approach with subtractive libraries also could prove valuable for efficiently screening subtracted cDNA libraries for differentially expressed genes. However, even though subtraction hybridization plus DDRT-PCR appears attractive for the reasons indicated above, a previous attempt to use this approach has proven of only marginal success in consistently reducing the complexity of the signals generated, compared with the standard DDRT-PCR scheme (25).

We presently describe a reciprocal subtraction differential RNA display (RSDD) approach that efficiently and consistently reduces the complexity of DDRT-PCR and results in the identification and cloning of genes displaying anticipated differential expression. The model used for RSDD was an adenovirus-transformed rat embryo cell line, E11, that acquires an aggressive oncogenic progression phenotype when injected into athymic nude mice and reestablished in cell culture (E11-NMT) (6, 26, 27). Injection of E11 cells into nude mice results in tumors in 100% of animals with a tumor latency time of ≈35–40 days whereas E11-NMT cells form tumors in 100% of nude mice with a tumor latency time of 15–20 days (6, 26, 27). Additionally, E11 cells form colonies in agar with an efficiency of ≈3% whereas E11-NMT display an agar cloning efficiency of >30% (6, 26, 27). The increased tumorigenicity and enhanced anchorage independence phenotypes are key indicators of tumor progression in the E11/E11-NMT model system (6, 26, 27). RSDD has resulted in the identification and cloning of genes displaying elevated expression in progressed tumor cells [progression-elevated gene (PEGen)] and suppressed expression in progressed tumor cells [progression-suppressed gene (PSGen)].

MATERIALS AND METHODS

RNA Isolation and cDNA Library Construction. Total RNA from E11 and E11-NMT cells was isolated by the guanidinium isothiocyanate/CsCl centrifugation procedure, and Poly(A)$^+$ RNA was purified with oligo(dT) cellulose chromatography (5). Two λ-ZAP cDNA libraries from E11 and E11-NMT mRNAs were constructed with λ-ZAP cDNA library Kits (Stratagene) following the manufacturer's protocol. Reciprocal subtraction between E11 and E11-NMT libraries was performed, and two subtracted cDNA libraries (E11 minus E11-NMT and E11-NMT minus E11) were constructed as described (5, 6). Plasmid cDNA libraries from the subtracted λ-ZAP cDNA libraries were obtained by *in vivo* excision following the manufacturer's protocol (Stratagene), and the plasmids were isolated with Qiagen columns (Qiagen, Chatsworth, CA).

RSDD Methodology. The purified plasmids of reciprocally subtracted cDNA libraries were subjected directly to differential display as in Liang *et al.* (28) with minor modifications. The plasmids of reciprocally subtracted cDNA libraries were PCR-amplified with the combination of three single-anchor 3′ primers (T$_{13}$A, T$_{13}$C, or T$_{13}$G) and 18 arbitrary 5′ 10-mer primers obtained from Operon Technologies (Alameda, CA. OPA 1–20 except OPA1 and 3). The 20-μl PCR reaction consisted of 10 mM Tris·HCl (pH 8.4), 50 mM KCl, 1.5 mM MgCl$_2$, 2 μM dNTP, 0.2 μM 5′ arbitrary primer, 1 μM 3′ anchor primer, 50 ng of plasmid of a subtracted library, 10 μCi α-^{35}S-dATP (3,000 Ci/mmol from Amersham), and 1 unit of *Taq* DNA polymerase (GIBCO/BRL). The parameters of PCR were 30 sec at 95°C, 40 cycles of 30 sec at 95°C, 2 min at 40°C, and 30 sec at 72°C, and an additional 5 min at 72°C. After the cycling, 10 μl of 95% formamide, 0.05% bromophenol blue, and 0.05% xylene cyanol were added to each PCR reaction. The mixture was heated at 95°C for 2 min and was separated in a 5% denaturing DNA sequencing gel maintained at 50°C. PCR reactions of plasmids from each subtracted library in a primer set were run side by side. Differentially amplified bands from plasmids of each subtracted library were marked with an

18-G needle through the film and were cut out with a razor. The gel slice was put in 100 μl of 10 mM Tris and 1 mM EDTA (TE; pH 8.0) and was incubated at 4°C overnight. After the incubation, the mixture was boiled for 5 min and was microcentrifuged for two min. The supernatant was collected and stored at −20°C until reamplification. The band extract was reamplified with the same cycling parameters in a 50-μl reaction consisting of 10 mM Tris·HCl (pH 8.4), 50 mM KCl, 1.5 mM MgCl$_2$, 20 μM dNTP, 0.2 μM 5′ arbitrary primer, 1 μM 3′ anchor primer, 5 μl of band extract, and 2.5 units of *Taq* DNA polymerase (GIBCO/BRL).

Reverse Northern Blotting Procedure. Differential expression of the reamplified DNA fragment was scrutinized by reverse Northern analysis and Northern blot analyses. In reverse Northern analysis, after confirmation in a 1% agarose gel, the reamplified DNA fragment (10 μl of PCR reaction) was mixed with 90 μl of TE and was spotted on a positively charged Nylon membrane (Boehringer Mannheim) with a 96-well vacuum manifold. The membrane was soaked with denaturing and neutralizing solution successively, and the spotted DNA was crosslinked to the membrane with a UV crosslinker (Stratagene). ^{32}P-labeled first strand cDNA was prepared by reverse transcription of total RNA. After heating at 70°C for 10 min and quenching on ice for two min, 0.4 μM each T$_{13}$A, T$_{13}$G, and T$_{13}$C and 10 μg total RNA mixture was added with 50 mM Tris·HCl (pH 8.3), 75 mM KCl, 3 mM MgCl$_2$, 10 mM DTT, 0.5 mM dATP, 0.5 mM dGTP, 0.5 mM dTTP, 0.02 mM dCTP, 0.5 μl of RNase inhibitor (GIBCO/BRL), 100 μCi dCTP (3,000 Ci/mmol from Amersham), and 200 units of Superscript RT II (GIBCO/BRL) in a final 25-μl reaction. The reaction mixture was incubated at 42°C for 1 hr and at 37°C for 30 min after addition of 2 μl of RNase H (10 units, GIBCO/BRL). The membrane was hybridized at 42°C overnight in a 50% formamide hybridization solution. The hybridized membrane was washed at room temperature for 15 min with 2× standard saline citrate containing 0.1% SDS twice and at 55°C for at least 1 hr with 0.1× standard saline citrate containing 0.1% SDS, successively. The membrane was probed with the ^{32}P-labeled cDNA of E11, was stripped off, and was probed with ^{32}P-labeled cDNA of E11-NMT. The signal intensity of each spot was normalized against that of glyceraldehyde-3-phosphate dehydrogenase and was compared between E11 and E11-NMT. Reamplified DNA fragments displaying differential expression levels ≥1.8-fold higher between the two cell types were selected and analyzed by Northern blot analysis.

Northern Blot Analysis. In Northern blot analysis, 10 μg of total RNA from E11 and E11-NMT cells were run side-by-side in a 1% agarose gel with formaldehyde and were transferred to a positively charged Nylon membrane. Reamplification reaction (5 μl) was ^{32}P-labeled with a multiprime labeling kit (Boehringer Mannheim) used to probe the membrane as described above. DNA fragments expressed differentially between E11 and E11-NMT in Northern blot analyses were cloned into the *Eco*RV site of the pZEro-2.1 cloning vector (Invitrogen) and were sequenced.

To confirm differential expression, the cloned cDNA fragment was released by *Eco*RI–*Xho*I, was ^{32}P-labeled, and was used to probe Northern blots as described above. Samples of RNAs from various E11 and E11-NMT derivatives displaying either a progressed or suppressed progression phenotype, based on nude mice tumorigenesis and soft agar cloning assays, were analyzed. These included E11, E11-NMT, CREF × E11-NMT F1 and F2 somatic cell hybrids (suppressed progression phenotype), CREF × E11-NMT R1 and R2 somatic cell hybrids (progression phenotype), E11 × E11-NMT A6 somatic cell hybrid (suppressed progression phenotype), E11 × E11-NMT A6TD tumor-derived somatic cell hybrid (progression phenotype), E11 × E11-NMT 3b somatic cell hybrid (suppressed progression phenotype), E11 × E11-NMT IIa (progression phenotype), E11-NMT AZA B1 and C1 5-azacytidine treated Ell-NMT clones (suppressed progression phenotype), E11-Ras R12 clone containing the Ha-*ras* oncogene (progression phenotype), and E11-HPV E6/E7 clone containing the human papilloma virus-18 E6 and E7 gene region

(progression phenotype). Differential expression of the PEGen and PSGen genes in the various cell types was confirmed by using ^{32}P-labeled probes and Northern hybridization analysis. After reconfirmation of differential expression, the plasmids containing the differentially expressed DNA fragments were sequenced by the dideoxy sequencing procedure.

RESULTS AND DISCUSSION

Subtraction hybridization provides a direct means of enriching for unique cDNA species and eliminating common sequences between complex genomes (7, 18). DDRT-PCR is a proven methodology for the rapid identification and cloning of differentially expressed sequences between cell types (1, 2, 29). In principle, subtraction hybridization combined with DDRT-PCR should reduce band complexity, which often obscures the identification of differentially expressed genes and generates false positive signals (21, 30). RSDD has been used to analyze genes differentially expressed during transformation progression (Fig. 1). Differential RNA display was performed directly with reciprocally subtracted cDNA plasmid libraries (E11 minus E11-NMT and E11-NMT minus E11) that had not been subjected to PCR. Three single anchored oligo dT 3′ primers were used for subsequent amplification before display. To further streamline the DDRT-PCR procedure, reamplified cDNAs identified by using RSDD were analyzed by using the reverse Northern blotting procedure (31, 32). cDNAs displaying differential expression by reverse Northern blotting subsequently were confirmed for true differential expression by Northern blot analysis.

The differential RNA display pattern of E11 and E11-NMT cells by using standard DDRT-PCR and RSDD is shown in Fig. 2 *Left*. The differential RNA display pattern of RSDD is significantly less complex than that of DDRT-PCR. These experiments demonstrate that subtractive hybridization before differential RNA display is effective in simplifying display patterns, permitting the efficient identification of differentially expressed cDNAs. Because RSDD substantially reduced the number of bands displayed, single anchor oligo dT primers, which can increase band numbers, were used successfully in subsequent applications of the RSDD approach (Fig. 2 *Right*). By using RSDD, 234 differentially displayed cDNAs in the E11/E11-NMT tumor progression model system were isolated.
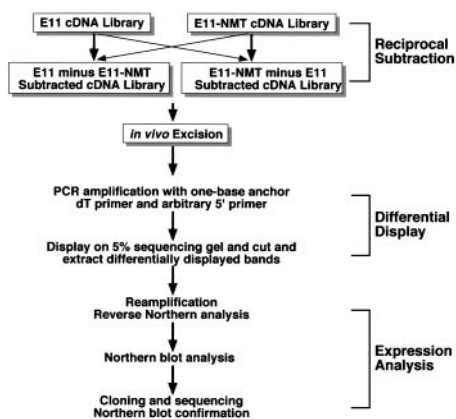

FIG. 1. Schematic outline of the RSDD protocol. This scheme incorporates three steps: reciprocal subtraction of cDNA libraries, differential display of *in vivo* excised cDNAs, and expression analysis by reverse Northern and standard Northern blotting. For the present application of RSDD, reciprocal subtraction hybridization was performed by using libraries constructed from E11 and E11-NMT cells, i.e., E11 minus E11-NMT and E11-NMT minus E11. Differentially expressed cDNAs identified on gels by using differential RNA display were isolated, reamplified, and analyzed for expression by reverse Northern blotting. To confirm differential expression, cDNAs were analyzed by using Northern blotting.
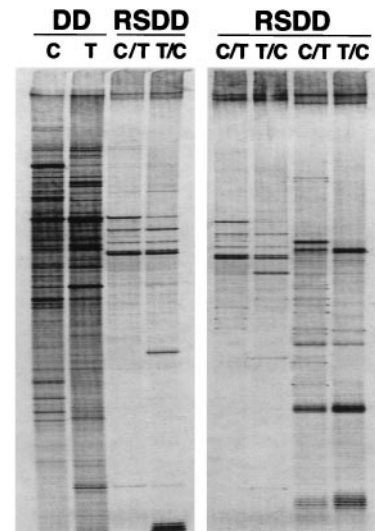

FIG. 2. Identification of differentially expressed sequence tags by using RSDD. (*Left*) Differential RNA display pattern of conventional DDRT-PCR with RNA from E11 (C) and E11-NMT (T) cells and an RSDD analysis of reciprocally subtracted E11 minus E11-NMT (C/T) and E11-NMT minus E11 (T/C) cDNA libraries. (*Right*) Representative RSDD patterns using different sets of primers.

Hakvoort *et al.* (25) used a reciprocal subtraction approach to analyze gene expression changes resulting during liver regeneration after 70% hepatectomy, i.e., normal liver subtracted from partially hepatectomized regenerating liver and vice versa. Although some bands displayed apparent enrichment, the complexity of the display pattern did not show appreciable simplification. In contrast, RSDD results in a clearer delineation and simplification of differentially expressed amplified bands (Fig. 2). Although conceptually similar, RSDD is significantly more effective than the subtraction plus DDRT-PCR approach described by Hakvoort *et al.* (25). The reasons for the improved efficiency of RSDD versus the Hakvoort *et al.* (25) approach are not known. One possibility is that the differences between the experimental approaches may reflect the subtraction hybridization strategies used. The experimental design of Hakvoort *et al.* (25) is based on the subtraction procedure described by Wang and Brown (33). This approach uses multiple rounds of PCR-amplification before each round of subtractive hybridization. In contrast, RSDD involves a single round of reciprocal subtraction without intermediate amplification (5, 6). In this respect, the complicated display pattern observed by Hakvoort *et al.* (25) even after three or four rounds of subtraction might result from reduced subtraction efficiency, PCR artifacts, or a combination of these problems. Increasing the number of reactions by using two-bp anchored oligo dT primers did not reduce the complexity of displayed bands (25). In these contexts, a critical component for the successful application of RSDD involves the use of an appropriate subtraction hybridization protocol, which can reduce cDNA complexity efficiently and can generate stable populations of cDNAs for analysis.

Previous studies demonstrate that different gene cloning strategies, including DDRT-PCR, subtraction hybridization, and electronic display, identify distinct subsets of differentially expressed genes (18). These results suggest that a single approach for gene identification may not identify the complete spectrum of differentially expressed genes. Similarly, RSDD and DDRT-PCR do not resolve the same differentially expressed bands (Fig. 2, and data not shown). Unique bands identified in DDRT-PCR that were expressed differentially when analyzed by Northern blotting were not the same as those found by using RSDD and vice versa (data not shown). Moreover, random isolation and analysis of differentially expressed cDNAs derived from the E11-NMT

Genetics: Kang *et al.*

Proc. Natl. Acad. Sci. USA 95 (1998)   13791

subtracted cDNA library resulted in no overlap with sequences obtained by using DDRT-PCR or RSDD (data not shown). These results are not surprising because, as indicated above, subtraction hybridization and differential RNA display identified distinct differentially expressed genes (18). Apparently, specific differentially expressed genes are lost during subtraction hybridization and differential RNA display of subtracted cDNAs. On the basis of these considerations, it may be necessary to use multiple gene discovery approaches to identify and clone the complete spectrum of differentially expressed genes.

DDRT-PCR can generate large numbers of differentially displayed bands, making subsequent analysis both labor intensive and a daunting challenge. To reduce these limitations of DDRT-PCR, RSDD has been used in combination with reverse Northern blot analysis of isolated cDNAs (31, 32). Gel extracted cDNA fragments were reamplified, were dot-blotted on Nylon membranes, and successively were probed with reverse transcribed ^{32}P-cDNA from E11 or E11-NMT RNAs (Fig. 3). Signals were detected in 181 of 234 reamplified bands (77%).

The signal intensities of the various cDNAs in reverse Northern analysis were quantified and normalized against that of glyceraldehyde-3-phosphate dehydrogenase, which remained unchanged in E11 and E11-NMT cells. Progression-elevated gene 3 (6) was used as an additional control to verify increased expression in E11-NMT versus E11 cells. In the reverse Northern blot analyses, progression-elevated gene 3 levels were 4-fold higher in E11-NMT than in E11 cells, which coincided with Northern blotting results, thereby demonstrating the concordance of reverse Northern blotting and Northern blotting assays. A $\geq$1.8-fold differential cut-off (after normalization for glyceraldehyde-3-phosphate dehydrogenase expression) was used to identify and isolate cDNA bands displaying modified expression in E11 versus E11-NMT cells. This resulted in the identification of 7 cDNAs with higher expression in E11 versus E11-NMT cells and 65 cDNAs with elevated expression in E11-NMT versus E11 cells. These results suggest that tumor progression in E11-NMT cells correlates with increased expression of a large number of genes whereas only a smaller subset of genes display decreased expression.

A problem frequently encountered in DDRT-PCR, which is reduced but still can occur in RSDD, is the isolation of multiple cDNA species from what appears to be a single amplified band. When this occurs, these multiple species can produce spurious results when analyzed by reverse Northern blot analyses. For example, if two distinct species are isolated, one displaying

modified expression and a second not displaying modified expression, an accurate estimate of differential expression will not be obtained by reverse Northern blot analysis. In this case, a number of potential false positives generated by using reverse Northern blot analyses may, in reality, not be false positives but instead may represent multiple cDNAs. By performing single strand conformational polymorphism or reverse Northern blot analyses by using cloned cDNA populations (34, 35), this problem can be ameliorated.

The expression pattern of representative RSDD-derived cDNAs in E11 versus E11-NMT and in a more expanded E11/E11-NMT progression cell culture series is shown in Figs. 4 and 5, respectively. Reverse Northern blot results correlated well with Northern blots using E11 and E11-NMT ($\approx$75% concordance) or a larger panel of cells differentially displaying the progression phenotype, including progression-negative E11, CREF × E11-NMT F1 and F2, E11 × E11-NMT A6, E11 × E11-NMT 3b, E11-NMT Aza B1 and Aza C1 cells, and progression-positive E11-NMT, CREF × E11-NMT R1 and R2, E11 × E11-NMT A6TD, E11 × E11-NMT IIa, E11-Ras R12, and E11-HPV E6/E7 cells. Sequence analysis of the various PEGen cDNAs identified both known and unknown genes (Table 1). Of 10 PEGen cDNAs, 5 (50%) were classified as novel sequences because no matches were found in current DNA databases. Novel PEGen cDNAs include PEGen 13, 14, 24, 28, and 32. Known PEGen genes included PEGen 7 (human papilloma virus-16 early region 1 binding protein; HPV16 E1BP), PEGen 8 (phosphofructokinase kinase C; PFK-C), PEGen 21 (a fibroblast growth factor-4 inducible gene; FIN-14), PEGen 26 [poly(ADP-ribose) polymerase] and PEGen 30 (rat esp1 homology). In the case of the PSGen cDNAs, six of six (100%) were novel, including PSGen 12, 13, 26, 27, 28, and 29 (Table 1).

PEGen 7 is expressed at $\approx$4-fold higher levels in E11-NMT than in E11 cells. PEGen 7 is $\approx$90% homologous to 16E1-BP, a cDNA encoding a protein identified by using the yeast two-hybrid assay that interacts with human papillomavirus type 16 E1 protein (36). 16E1-BP encodes a 432-aa protein of unknown function but does contain an ATPase signature motif (Gly-X$_4$-Gly consensus ATP binding motif at amino acids 179–186). 16E1-BP appears to be a form of TRIP13, a protein previously shown (36) to bind thyroid hormone receptor in yeast two-hybrid assays. The role of PEGen 7/16E1-BP in the progression phenotype in the E11/E11-NMT progression model is not known. Additional studies
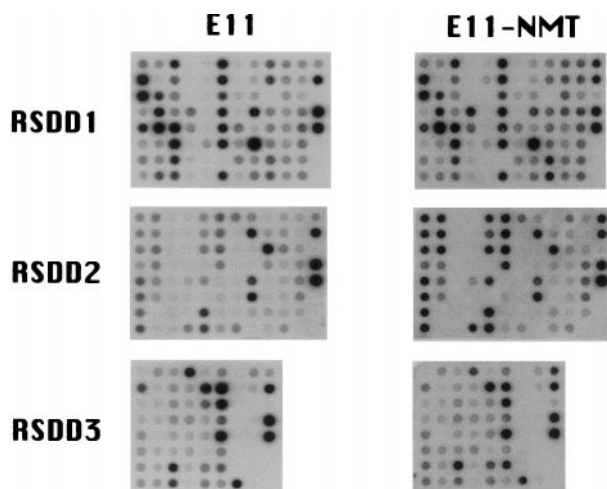

FIG. 3. Reverse Northern blot analysis of differentially expressed sequence tags identified by RSDD. Differentially expressed sequence tags obtained from RSDD were dot-blotted onto Nylon membranes and were probed with ^{32}P-cDNA reverse transcribed from RNA samples of E11 and E11-NMT cells.
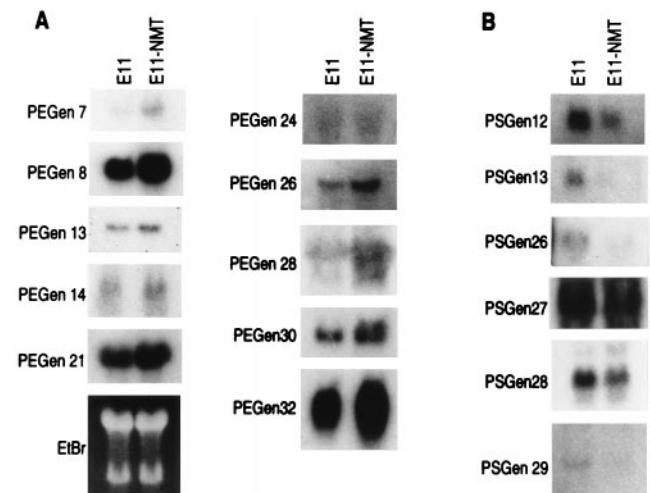

FIG. 4. Differential expression of representative PEGen and PSGen genes identified by RSDD and reverse Northern blotting. Northern blots of E11 and E11-NMT RNA samples were probed with radiolabeled (^{32}P) expressed sequence tags identified by RSDD and reverse Northern blotting. Equal loading of E11 and E11-NMT RNA is demonstrated by ethidium bromide (EtBr) staining.
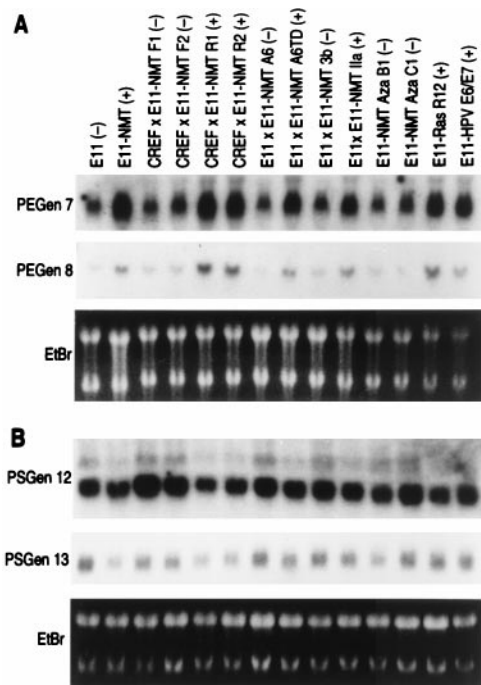
FIG. 5. Differential expression of representative PEGen and PS-Gen genes identified by RSDD and reverse Northern blotting in a large panel of rodent cells displaying differences in transformation progression. Northern blots of cells displaying various stages of transformation progression were probed with radiolabeled (^{32}P) expressed sequence tags identified by RSDD and reverse Northern blotting. The cell types used include unprogressed E11 (−), CREF × E11-NMT F1 (−) and CREF × E11-NMT F2 (−) somatic cell hybrids, E11 × E11-NMT A6 (−) somatic cell hybrid, E11 × E11-NMT 3b (−) somatic cell hybrid, and E11-NMT AZA B1 (−) and E11-NMT AZA C1 (−) 5-azacytidine-treated E11-NMT clones; and progressed E11-NMT (+), CREF × E11-NMT R1 (+) and CREF × E11-NMT R2 (+) somatic cell hybrids, E11 × E11-NMT A6TD (+) nude mouse tumor derived somatic cell hybrid, E11 × E11-NMT IIa (+), E11-Ras R12 (+) and E11-HPV E6/E7 (+); an E11 clone transformed by the E6 and E7 region of HPV-18. Equal loading of RNAs is demonstrated by ethidium bromide (EtBr) staining.

are necessary to determine whether this gene change is associative or causative of transformation progression.

PEGen 8 is expressed at ≈3- to 4-fold higher levels in E11-NMT than in E11 cells. PEGen 8 shows 100% homology to rat phosphofructokinase C (PFK-C) (37). PFK catalyzes the rate-limiting and committed step in glycolysis, the conversion of fructose 6-phosphate to fructose 1,6-biphosphate. Three subunit isozymes of PFK have been identified that form homo- and heterotetramers with differing catalytic and allosteric properties. PFK-M is specific for cardiac and skeletal muscle, PFK-L is expressed in many tissues but is most abundant in the liver, and PFK-C is expressed in several brain regions and the anterior pituitary but not in liver, skeletal muscle, or several other human tissues. The cDNA of PFK-C isolated from a rat hypothalamic cDNA library is 2,643 bp and encodes a protein of 765 amino acids (37). In a recent study, Sanchez-Martinez and Aragon (38) demonstrated that PFK-C is the predominant form of PFK in ascites tumor cells (obtained from a transplantable mouse carcinoma of mammary origin) whereas PFK-L is most abundant in the normal mammary gland. These results suggest the interesting possibility that PFK-C might contribute to the malignant nature of specific target cells. The role of PEGen 8/PFK-C in progression in the E11/E11-NMT model remains to be determined.

PEGen 21 is expressed at ≈3- to 4-fold higher levels in E11-NMT than in E11 cells. PEGen 21 displays ≈94% homology with the fibroblast growth factor-4 inducible gene FIN-14 (39). FIN-14 is a novel cDNA of unknown function that hybridizes with

Table 1. PEGen and PSGen genes isolated by using RSDD

Nomenclature*	Identity†	Homology,‡ %
PEGen 7	Human HPV16 E1BP	90
PEGen 8	PFK-C	100
PEGen 13	Unknown	Novel
PEGen 14	Unknown	Novel
PEGen 21	Murine FIN 14	94
PEGen 24	Unknown	Novel
PEGen 26	PARP	100
PEGen 28	Unknown	Novel
PEGen 30	Rat esp1	98
PEGen 32	Unknown	Novel
PSGen 12	Unknown	Novel
PSGen 13	Unknown	Novel
PSGen 26	Unknown	Novel
PSGen 27	Unknown	Novel
PSGen 28	Unknown	Novel
PSGen 29	Unknown	Novel

*PEGen are progression-elevated genes that display elevated expression in E11-NMT versus E11 cells. PSGen are progression suppressed genes that display elevated expression in E11 versus E11-NMT cells.
†Sequences have been compared with reported genes in various DNA data bases (including GenBank and the European Molecular Biology Laboratory), and identifications with known genes are indicated. Genes without homology to currently reported genes are indicated as unknown.
‡Percentage homology with known sequences, either human, rat, or mouse, is indicated. Where no homology exists, the cDNA is considered novel.

a 4.5-kilobase mRNA that is induced 4-fold in NIH 3T3 mouse cells after treatment with FGF-4. The induction of FIN-14 occurs late (18 hr) after treatment with FGF-4 and does not occur when cells are treated for 18 hr with FGF-4 in the presence of cycloheximide (39). These results confirm that FIN-14 encodes a late-inducible gene. Moreover, nuclear run-on assays document that FIN-14 is activated transcriptionally in NIH 3T3 cells after growth factor stimulation. Tissue distribution studies indicate expression of a single mRNA species in the kidney, with low levels of expression observed in several other tissues, including testis and thymus. Mouse embryogenesis studies indicate that FIN-14 expression occurs constitutively in mouse embryos between days 10.5 and 15.5. Unlike NIH 3T3, FIN-14 was expressed constitutively in PC12 cells, and its level did not vary appreciably in response to growth factor stimulation. The role of PEGen 21/FIN-14 in progression in E11/E11-NMT model system is not currently known.

PEGen 26 is expressed at ≈3- to 4-fold higher levels in E11-NMT than in E11 cells. This cDNA is identical to rat poly(ADP-ribose) polymerase (PARP) (40). PARP contributes to the ability of eukaryotic cells to contend with both environmental and endogenous genotoxic agents (41). PARP is a nuclear enzyme that binds to DNA breaks and then catalyzes the covalent modification of acceptor proteins with poly(ADP-ribose) (40, 41). PARP activity contributes to the recovery of proliferating cells from DNA damage and to the maintenance of genomic stability, which may be regulated by effects on chromatin structure, DNA base-excision repair, and cell cycle regulation (40, 41). The role of PEGen 26/PARP in mediating the progression phenotype is not currently known. However, because cancer is a progressive disease characterized by the accumulation of genetic alterations in the evolving tumor (6), it is tempting to speculate that overexpression of PEGen 26/PARP in E11-NMT may facilitate the ability of these aggressive cancer cells to maintain genomic stability during cancer progression. In this context, PEGen 26/PARP may be an integral component of progression. This hypothesis is readily testable.

PEGen 30 is expressed at 2- to 3-fold higher levels in E11-NMT than in E11 cells. This cDNA displays ≈98.5% homology to rat esp1 (42). Rat esp1 encodes a 24-kDa nuclear protein which is the

rat homologue of *Drosophila Enhancer of split.*, a gene involved in ventral ectodermal development in *Drosophila* (42). PEGen 30 appears to be a homologue of esp1 because the message detected in E11 and E11-NMT cells (≈4 kilobases) is larger in size than the reported esp1 transcript (1.3 kilobases) (42). The role of PEGen 30/esp1 in tumor progression in E11/E11-NMT model system remains to be determined.

The PSGen cDNAs, 12, 13, 26, 27, 28, and 29, consist of sequences without homology to those presently reported in various DNA databases. Expression of PSGen-12 and PSGen-13 cDNAs is ≈3- to 4-fold higher in E11 versus E11-NMT cells (Fig. 4). It is not currently known whether these genes simply correlate with or functionally regulate the progression phenotype. The identification of full length cDNAs for PSGen-12 and PSGen-13, as well as the other novel PSGen and PEGen cDNAs, is in progress, and, once isolated, experiments should be able to be conducted to define directly the role of these progression-related genes in cancer progression.

We presently demonstrate that a modified gene-identification and gene-cloning technique, RSDD, can identify differentially expressed cDNAs efficiently. As predicted, subtractive hybridization before differential RNA display greatly reduces band complexity, a problem encountered in standard DDRT-PCR in which RNA samples are analyzed directly without subtraction. Unlike a previous report using subtracted cDNAs processed through successive rounds of PCR (25, 43), common bands were eliminated by using reciprocally subtracted cDNA libraries that had not been processed by using PCR. In addition to subtraction hybridization, the discovery of differentially expressed genes was streamlined further by using reverse Northern blot analyses with isolated cDNAs. With 3 single anchored oligo dT primers and 18 arbitrary 5′ primers, 72 bands were identified that displayed differential expression by using reverse Northern blot analysis. Currently, 38 cDNA species have been analyzed by Northern blotting, and 31 (≈82%) displayed differential expression in E11 versus E11-NMT cells. Sequence analysis of the cloned cDNA fragments revealed 16 different genes, including 11 novel genes not reported in recent DNA databases. RSDD represents a method of choice either as a more efficient and less time-consuming modification of the differential RNA display strategy or as a screening methodology for identifying differentially expressed genes in reciprocally subtracted cDNA libraries. Moreover, the ability of RSDD to identify differentially expressed genes that are dissimilar to those recognized by using standard DDRT-PCR or subtraction hybridization indicates that this approach should be a valuable adjunct to these approaches in identifying and cloning differentially expressed genes occurring between complex genomes and resulting from changes in cellular physiology.

1. Watson, J. B. & Margulies, J. E. (1993) *Dev. Neurosci.* **15,** 77–86.
2. Winkles, J. A. (1998) *Prog. Nucleic Acid Res. Mol. Biol.* **58,** 41–78.
3. Fisher, P. B., ed. (1990) *Model Cell Culture Systems for Studying Differentiation: Mechanisms of Differentiation* (CRC, Boca Raton, FL), Vol. 1.
4. Fisher, P. B., ed. (1990) *Modulation of Differentiation by Exogenous Agents: Mechanisms of Differentiation* (CRC, Boca Raton, FL), Vol. 2.
5. Jiang, H. & Fisher, P. B. (1993) *Mol. Cell. Differ.* **1,** 285–299.
6. Su, Z.-Z., Shi, Y. & Fisher, P. B. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 9125–9130.
7. Sagerström, C. G., Sun, B. I. & Sive, H. L. (1997) *Annu. Rev. Biochem.* **66,** 751–783.
8. Jiang, H., Lin, J., Su, Z.-Z., Herlyn, M., Kerbel, R. S., Weissman, B. E., Welch, D. R. & Fisher, P. B. (1995) *Oncogene* **10,** 1855–1864.
9. Jiang, H., Lin, J. J., Su, Z.-Z., Goldstein, N. I. & Fisher, P. B. (1995) *Oncogene* **11,** 2477–2486.
10. Jiang, H., Su, Z.-Z., Lin, J. J., Goldstein, N. I., Young, C. S. H. & Fisher, P. B. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 9160–9165.
11. Liang, P. & Pardee, A. B. (1992) *Science* **257,** 967–971.
12. Shen, R., Su, Z.-Z., Olsson, C. A. & Fisher, P. B. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 6778–6782.
13. McClelland, M., Mathieu-Daude, F. & Welsh, J. (1995) *Trends Genet.* **11,** 242–246.
14. Ralph, D., McClelland, M. & Welsh, J. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 10710–10714.
15. Hubank, M. & Schatz, D. G. (1994) *Nucleic Acids Res.* **22,** 5640–5648.
16. Velculescu, V. E., Zhang, L., Vogelstein, B. & Kinzler, K. W. (1995) *Science* **270,** 484–487.
17. Zhang, L., Zhou, W., Velculescu, V. E., Kern, S. E., Hruban, R. H., Hamilton, S. R., Vogelstein, B. & Kinzler, K. W. (1997) *Science* **276,** 1268–1272.
18. Wan, J. S., Sharp, S. J., Poirier, G. M.-C., Wagaman, P. C., Chambers, J., Pyati, J., Hom, Y.-l., Galindo, J. E., Huvar, A., Peterson, P. A., *et al.* (1996) *Nat. Biotechnol.* **14,** 1685–1691.
19. Adams, M. D., Kerlavage, A. R., Fields, C. & Venter, J. C. (1993) *Nat. Genet.* **4,** 256–267.
20. Schena, M., Shalon, D., Davis, R. W. & Brown, P. O. (1995) *Science* **270,** 467–470.
21. Debouck, C. (1995) *Curr. Opin. Biotechnol.* **6,** 597–599.
22. Rangnekar, V. V., Waheed, S. & Rangnekar, V. M. (1992) *J. Biol. Chem.* **267,** 6240–6248.
23. Wong, B., Park, C. G. & Choi, Y. (1997) *Semin. Immunol.* **9,** 7–16.
24. Maser, R. L. & Calvet, J. P. (1995) *Semin. Nephrol.* **15,** 29–42.
25. Hakvoort, T. B., Leegwater, A. C., Michiels, F. A., Chamuleau, R. A. & Lamers, W. H. (1994) *Nucleic Acids Res.* **22,** 878–879.
26. Reddy, P. G., Su, Z.-Z. & Fisher, P. B. (1993) in *Chromosome and Genetic Analysis: Methods in Molecular Genetics*, ed. Adolph, K. W. (Academic, Orlando, FL), Vol. 1, pp. 68–102.
27. Babiss, L. E., Zimmer, S. G. & Fisher, P. B. (1985) *Science* **228,** 1099–1101.
28. Liang, P., Bauer, D., Averboukh, L., Warthoe, P., Rohrwild, M., Muller, H., Strauss, M. & Pardee, A. B. (1995) *Methods Enzymol.* **254,** 304–321.
29. Liang, P. & Pardee, A. B. (1995) *Curr. Opin. Immunol.* **7,** 274–280.
30. Averboukh, L., Douglas, S. A., Zhao, S., Lowe, K., Maher, J. & Pardee, A. B. (1996) *BioTechniques* **20,** 918–921.
31. Zhang, H., Zhang, R. & Liang, P. (1997) *Methods Mol. Biol.* **85,** 87–93.
32. Zhao, S., Ooi, S. L., Yang, F. C. & Pardee, A. B. (1996) *BioTechniques* **20,** 400–404.
33. Wang, Z. & Brown, D. D. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 11505–11509.
34. Zhang, H., Zhang, R. & Liang, P. (1996) *Nucleic Acids Res.* **24,** 2454–2455.
35. Mathieu-Daude, F., Cheng, R., Welsh, J. & McClelland, M. (1996) *Nucleic Acids Res.* **24,** 1504–1507.
36. Yasugi, T., Vidal, M., Saka, H., Howley, P. M. & Benson, J. D. (1997) *J. Virol.* **71,** 5941–5951.
37. Gekakis, N., Johnson, R. C., Jerkins, A., Mains, R. E. & Sul, H. S. (1994) *J. Biol. Chem.* **269,** 3348–3355.
38. Sanchez-Martinez, C. & Aragon, J. J. (1997) *FEBS Lett.* **409,** 86–90.
39. Guthridge, M. A., Seldin, M. & Basilico, C. (1997) *Oncogene* **12,** 1267–1278.
40. Beneke, S., Meyer, R. & Burkle, A. (1997) *Biochem. Mol. Biol. Intl.* **43** (4), 755–761.
41. de Murcia, G. & de Murcia, J. M. (1994) *Trends Biochem. Sci.* **19,** 172–176.
42. Schmidt, C. J. & Sladek, T. E. (1993) *J. Biol. Chem.* **268,** 25681–25686.
43. Wu, C. G., Hakvoort, T. B., Lamers, W. H. & Chamuleau, R. A. (1996) *Biochimic. Biophysic. Acta* **1315,** 169–175.