Establishment of Monitoring System to Detect Single Copy DNA Included in One Genome but not in Another Using Representational Difference Analysis

Hiroaki MITA*, Fumio ITOH, Hideki KAKIUCHI, Minoru TOYOTA,
Yuji HINODA, and Kohzoh IMAI

First Department of Internal Medicine, Sapporo Medical University, School of Medicine, S-1, W-16, Chuo-ku, Sapporo 060, Japan.

Key words: Gene subtraction, Adapter

Abstract

Polymerase chain reaction (PCR)-coupled subtractive procedure, representational difference analysis (RDA), is an efficient method to find the differences between two complex genomes. RDA has been applied to detect genetic lesions in cancer, the identification of unknown pathogens from the genomes, and the isolation of polymorphic markers. However, characterization of various clones obtained by RDA is time consuming and laborious work, and it is of great importance to monitor whether RDA really works. To establish a monitoring system to detect single copy target DNA, we studied whether RDA could detect four fragments of non-human DNA which were added in one genome but not in another. We were able to successfully detect the target DNAs which were mixed at the ratio of single and ten copies per haploid genome using RDA with some modification of the original protocol. We confirmed that RDA was sensitive and effective enough to detect such genetic lesions as occurred in cancer cells. The target DNA used in this model could be utilized as a positive control in other applications of RDA.

INTRODUCTION

A variety of genetic lesions are found in tumors, including deletion, gene amplification, rearrangement, point mutation and acquisition of viral genomes (1, 2). In particular, frequent losses of both alleles at given locus or losses of one allele with functional inactivation of the other have been detected in many tumor types. These genetic lesions have been found to be hallmarks of the presence of tumor suppressor genes. Various time-consuming methods have been

* To whom requests for reprints should be addressed

applied to determine the nature of the genetic changes that occur in cancers, but recently a new method, called representational difference analysis (RDA) was designed by Lisitsyn *et al.* (3) for analyzing the differences between complex but highly related genomes.

RDA belongs to the general class of DNA subtractive methodologies. These methodologies all have in common that one DNA population (the "driver") is hybridized in excess against a second population (the "tester") to remove common hybridizing sequences, thereby enriching "target" sequences unique to the tester.

RDA combines two elements: (i) representation, (ii) subtractive and kinetic enrichment. Thus, the procedure is carried out in two stages.

The first stage is comprised of making representations of DNA populations that are called "amplicons"; DNA cleaved with relatively infrequent cutting restriction endonucleases (e. g. BamHI) is ligated to oligonucleotide adapters and amplified by polymerase chain reaction (PCR). The result is similar to size fraction, since low molecular size fragments below 1 kbp are mainly amplified. Only a subset of the whole genome is represented. However, when different restriction endonuclease are chosen, sets of amplicons that scan the genome can be made.

The second stage is comprised of the reiterative hybridization/selection steps. Prior to the hybridization/selection step, only tester amplicon is fitted with a new pair of defined oligonucleotides at their 5' ends. After reannealing amplicons of the tester and driver (competitive hybridization), the mixture of molecules is treated with DNA polymerase, adding the complement of the defined oligonucleotides to both 3' ends of only self reannealed tester amplicons. When the defined oligonucleotide is used as primer in the PCR of the mixture, only such molecules can participate in exponential amplification selectively because of the following reasons; (i) Molecules of tester amplicon that reanneal to the excess of driver amplicon are subtracted out since the heteroduplexes they form with driver amplicon have the primer complement on one end only, thus leading to inefficient, linear amplification of only one strand (subtractive enrichment). (ii) Abundant sequences contained in tester amplicon will reanneal faster than less abundant ones (kinetic enrichment) so that sequences that are enriched or amplified become even more enriched. (iii) PCR amplification increases the yield so that the hybridization/selection step can be reiterated, or the product can be cloned and analyzed. Three rounds of hybridization/selection are employed to achieve effective purification of the difference products.

In this way, probes marking loci of genetic lesions in cancer including losses, amplifications and rearrangements have been detected (4). Further, exogenous

DNA sequences such as viruses (5, 6) and polymorphic markers among different strains (3, 8) can be isolated easily. However, Lisitsyn et al. showed that there were cases where the number of probes obtained by RDA was relatively small (3).

In the present study, we performed a model experiment for stringent tests of RDA. DNA from human cancer cell line cleaved with BamHI restriction endonuclease was used as tester and driver. Restriction endonuclease fragments of non-human DNA were added only to tester as target. We examined whether RDA could detect single copy of target DNA included in one genome but not in another.

MATERIALS AND METHODS

Cell culture and DNA samples

Human gastric adenocarcinoma cell line MKN28 was obtained from the Japanese Cancer Research Resources Bank. This cell line was routinely grown at 37°C with 5% CO_2 in RPMI1640 supplemented with 10% fetal bovine serum. Genomic DNA was extracted as described previously (7).

Preparation of tester and driver

Four clones of BamHI restriction endonuclease fragments of rat genomic DNA which were cloned to pBluescript II (Stratagene), randomly selected but confirmed that they were not included in the human genome (8), were used to prepare target DNA fragments. Inserts were digested with BamHI (TOYOBO), phenol extracted, ethanol precipitated as described (7), dissolved in TE and mixed. The mixture was prepared as the target in which each rat fragment were equimolar.

Genomic DNA from MKN28 was digested with BamHI, phenol extracted, ethanol precipitated and dissolved in TE. To prepare the tester amplicon, the mixture of rat DNA fragments were added to digested MKN28 DNA. The same DNA without rat DNA fragments were used to prepare the driver. The target was added to the tester in two manners: as single and ten copies per haploid genome of MKN28.

Generation of representation

RDA procedure was performed after modification (9) of the original protocol (3, 10). Sequences of oligonucleotides used for RDA are shown in Table 1. The RBam adapter was made by mixing 50 μl (100 μM) each of RBam24 primer and RBam12 primer, incubated at 65°C for 2min, and cooled down slowly to room temperature. Three μl (0.5 μg) of tester and driver were ligated to the adapter

Table 1 Primer sequences

Primer	Sequence (5'→ 3')
RBam24	AGC ACT CTC CAG CCT CTC ACC GAG
RBam12	GAT CCT CGG TGA
JBam 24	ACC GAC GTC GAC TAT CCA TGA ACG
JBam 12	GAT CCG TTC ATG
NBam24	AGG CAA CTG TGC TAT CCG AGG GAG
NBam12	GAT CCT CCC TCG

in a mixture containing: 10 μl (50 μM) of RBam adapter, 3 μl of 10× ligase buffer (New England Biolabs), 1 μl of T4 ligase (New England Biolabs), and 13 μl of DW for 12-16 hr at 16°C.

PCR reactions were set up to generate tester and driver "amplicons" as representations. PCR reactions were carried out in the mixture containing 3 μl of the ligation mix; 40 μl of 10× PCR buffer [670 mM Tris-HCl, pH 8.8 at 20°C, 40 mM $MgCl_2$, 160 mM $(NH_4)_2SO_4$, 100 mM β-mercaptoethanol, bovine serum albumin (1 mg/ml)]; 48 μl of dNTPs solution (2.5 mM each of dATP, dGTP, dCTP and dTTP); 3 μl (15 U) of Taq DNA polymerase (Amersham); 298 μl of DW. The 5' ends of 12mer were filled in at 72°C for 5 min. Eight μl (50 μM) of RBam24 primer was added to the PCR mixture at 95°C of the first cycle, and 20 cycles of amplification were performed (1 min, 95°C; 3 min 72°C). PCR products were phenol extracted, ethanol precipitated and resuspended in TE. The RBam adapters were digested with BamHI and removed with a cDNA spun column (Pharmacia) and ethanol precipitated. As for the tester amplicon, 0.5 μg of DNA was ligated to the JBam adapter as described above.

Competitive hybridization and selective amplification

For the first round of competitive hybridization, 0.5 μg (30 μl) of JBam adapter-ligated tester amplicon was mixed with 40 μg of driver amplicon. The mixture was phenol extracted, ethanol precipitated and resuspended in 4 μl of 3× EE buffer [30 mM EPPS (Sigma), pH 8.0 at 20°C; 3 mM EDTA]. The solution was overlaid with mineral oil and DNA denatured at 96°C for 10 min. The salt concentration was adjusted with 1 μl of 5 M NaCl and the sample was allowed to anneal at 67°C for 20 hr. The hybridized DNA was diluted with 45 μl of 1 M NaCl at 67°C, and 5 μl of the aliquot was used for PCR reaction. PCR mixture prepared as described above, except that JBam24 primer was used, was incubated for 3 min at 85°C to reduce priming mediated by duplexes of near-identical repetitive elements, and amplified for 10 cycles of PCR (95°C, 1 min; 70°C, 3

min). The PCR product was treated with mung bean nuclease (New England Biolabs) for 30 min at 30°C in a final volume of 40 μl mung bean nuclease buffer containing 20 U enzyme. The mixture was phenol extracted, ethanol precipitated and dissolved in 50 μl of TE. This solution was amplified as above for 20 cycles.

Two hundred ng of the difference product was digested with BamHI and ligated to NBam adapter. For the second round of competitive hybridization reaction 100 ng of NBam adapter-ligated difference product of round 1 was mixed with 40 μg of driver amplicon. For the third round, the difference product of round 2 was digested and ligated to the JBam adapter, and 2ng of JBam adapter-ligated difference product of round 2 was mixed with 40 μg of driver amplicon. Competitive hybridization and selective amplification were performed as described in the first round.

Southern Blot analysis

One μg of tester amplicon, difference products of round 1, 2, 3 which were each obtained from two RDA processes using single and ten copies of target, and driver amplicon were run through 1% agarose gel electrophoresis and blotted onto nylon filter (Hybond-N^+: Amersham). Probe DNAs were prepared by purifying above 4 fragments of rat DNAs, and then by labeling them with [α-^{32}P] dCTP by random DNA labeling (Random DNA labeling kit: Takara). Prehybridization and hybridization were performed as described (7).

RESULTS

Amplification of the target DNA fragments by PCR with ligation of the RBam adapter

Four clones of BamHI digested fragments of rat DNA [fragment A: 450 bp, B: 1100 bp, C: 550 bp, D: 310 bp] were used as target DNA (Fig. 1A). Preliminary to RDA, the fragments were ligated to the RBam adapter and we examined whether they could be amplified by adapter-ligated PCR. Each of the adapter-ligated fragments were efficiently amplified by 20 cycles of PCR as shown in Fig. 1B. As regards extra bands which were not shown in Fig. 1A, it was suspected that the fragments ligated in tandem and were amplified, judging from the sizes. DNA fragment from pBluescript II, in which rat genomic DNAs were cloned, were also amplified.

Detection of single copy DNA included in one genome but not in another using RDA

Genomic DNA from human gastric adenocarcinoma cell line MKN28 were

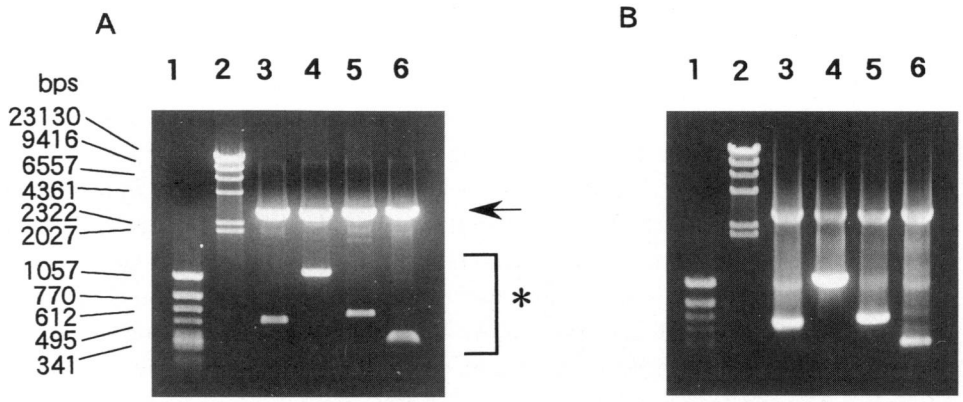

Fig. 1 Four clones of rat genomic DNA fragments which were digested from pBluescript II and used as target in RDA. $\phi \times 174$-HincII and λ-HindIII DNA size markers (lane 1, 2 in Fig. 1A and lane 1, 2 in Fig. 1B) are indicated in base pairs. (A) Agarose gel electrophoresis of rat fragment A (lane 3 : 450 bp), B (lane 4 : 1100 bp), C (lane 5 : 550 bp), D (lane 6 : 310 bp) and pBluescript II (lane 3-6 : 2961 bp) were digested with BamHI. Mixed aliquot (10 μg) of rat fragments (indicated by asterisk) and pBluescript II (indicated by arrow) were electrophoresed in a 1% agarose gel. (B) Agarose gel electrophoresis of rat fragments which were amplified by RBam adapter- ligated PCR, and corresponded to each lane of Fig. 1A. Mixed aliquots (10 ng) of each rat fragments and pBluescript were ligated to RBam adapter and amplified by PCR. 1/10 volume (10 μ1) of the PCR mixture were electrophoresed in a 1% agarose gel.

digested with BamHI. Single or ten copies of target DNA per haploid genome were added to the digested DNA to prepare the tester (Fig. 2A, lane 3, 4). To prepare the driver, the same DNA was used without the target. Tester and driver amplicons were prepared by PCR after ligation of the RBam adapter. These two tester and one driver amplicons showed almost the same smear bands in the range of about 500 bp to 2 kbp (Fig. 2B, lane 5-7). RDA was performed using a combination of driver amplicon and either tester amplicon.

After three round of RDA, difference products were obtained as shown in Fig. 2B, lane 9-14. The products from RDA using tester containing single copy per haploid genome were electrophoresed in agarose gel (Fig. 2B, lane 9-11). The product of 1st round RDA showed the smear band including ambiguous bands, the product from 2nd round RDA showed indistinct bands which corresponded to fragment B and C, and the 3rd product showed three distinct bands which corresponded to fragments A, B, C, and an indistinct band which corresponded to fragment D.

On the other hand, 1st round product from RDA using tester containing ten copies of target (Fig. 2B, lane 12) showed almost the same pattern as that of the

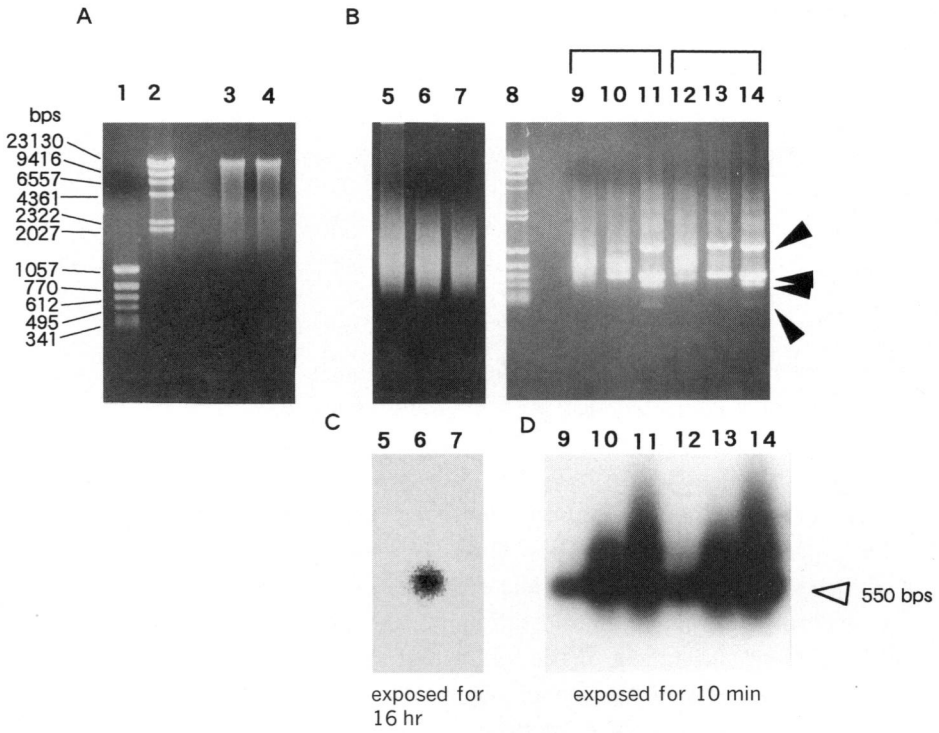

Fig. 2 Amplification of target DNA by RDA. ϕ×174-HincII, λ-HindIII DNA size markers (lane 1, 2 in Fig. 2A) , and mixture of those (lane 8 in Fig. 2B) are indicated in base pairs. (A) . Agarose gel electrophoresis of testers which were digested DNA of MKN28 with BamHI and contained target. Aliquots (200ng) of tester containing a single copy of target (lane 3) and ten copies of target (lane 4) were electrophoresed in a 1% agarose gel. (B) Agarose gel electrophoresis of tester amplicons (lane 5 and 6), driver amplicon (lane 7) and each difference product of 1st (lane 9 and 12) , 2nd (lane 10 and 13) and 3rd (lane 11 and 14) round. The difference products obtained from RDA with tester containing a single copy of target (lane 5) were placed in lanes 9 to 11, and those obtained from RDA with tester containing ten copies of target (lane 6) were placed in lanes 12 to 14. Aliquots (1 μg) of those were electrophoresed in a 1% agarose gel. The amplified rat fragments in difference products are indicated by arrowheads. Southern blot of two tester amplicons, driver amplicon (C), difference products from 1st, 2nd, and 3rd round of RDA (D). The blots from Fig. 2B were probed with rat fragment C (550 bp).

former. However, the product from 2nd round RDA (Fig. 2B, lane 13) showed distinct bands corresponding to fragment B and C, and product from 3rd round (Fig. 2B, lane 14) made three distinct bands which corresponded to fragments A, B, C, and an indistinct band which corresponded to fragment D.

To confirm whether RDA products obtained were rat DNA fragments, two

tester amplicons, driver amplicon, and each difference product were analyzed by Southern blot analysis using four rat fragments as probes. Representative results from Southern blot analysis probed with fragment C are shown in Fig. 2C and 2D. The difference products show much stronger signal than tester amplicons, and the products of later round showed stronger signals than those of earlier rounds.

These results indicated that we were able to recover only single copy of target which existed in haploid genome, very effectively after three rounds of RDA.

DISCUSSION

Here, we demonstrated that single and ten copies per haploid genome added only to the tester could be recovered by RDA effectively. The difference products using tester which contained single copy of target DNA per haploid genome were enriched and concentrated more slowly than those obtained from tester which contained ten copies of target, however, they were sufficiently enriched in the third round (Fig. 2B, Fig. 2D). Three cycles of hybridization/selection were proved to be necessary for the sufficient recovery of target DNA.

In general, subtractive hybridization techniques reported previously do not achieve sufficient enrichment of the target that exist only in tester partly because of the high complexity of the human genome, which prevents effectively complete hybridization (3). Even when subtractive steps are repeated, target sequences are enriched only 100 to 1000 times (11). This enrichment is insufficient for more common situations in which the magnitude of enrichment require is 10^6. On the other hand, in RDA, if the target is equimolar with respect to the tester (that is, single copy), the target will be enriched by about 10^5 at the end of the second round, and at the end of the third round by more than 10^{10} (3).

Theoretically, single copy per haploid of target DNA contained in tester could surely be recovered. However, RDA utilizes multiple steps of PCR, ligation and digestion. Therefore, the chance of contamination would always pursue. Additionally, in the step of competitive hybridization, the mixture of tester amplicon (or difference product) and driver amplicon must be kept at high temperature (67°C) to make and maintain stringent annealing. If these complex processes are not carried out correctly, the efficiency of subtraction will be decreased, and more effort would be required to analyze difference products. To reduce the possibility of getting false positive (e. g. background DNA fragments such as repetitive sequences), control experiment should be performed in parallel. The DNA fragments from rat genomic DNA used in this test were amplified sufficiently by adapter-ligated PCR, and will be useful as positive controls in our other applications of RDA.

RDA can be applied to detect genetic lesions in cancer in two different ways: using tumor DNA as driver and normal DNA from the same patient as tester, or in the reverse way. In either case, driver must be relatively free (at least 95%) of tester sequences. Tester DNA can include significant quantities of driver, up to 90%. (10)

When tumor DNA is used as driver, RDA yields probes that detect loss of heterozygosity (LOH), homozygous deletion, and hemizygous loss (loss of Y chromosome or deletion of other sequences present on one chromosome only in normal cells). The percentage of probes of each type generally depends on the total length of each loss, which may comprise up to 10% of genome for LOH, up to 1% for loss of Y chromosome (probes of this type can be eliminated by using female normal/tumor DNA pairs), and up to 0.3% for homozygous deletions (10). However, not more than 3% of fragments in the human genome are polymorphic, and this leads to a strong bias toward the cloning homo- and hemizygous losses. According to Lisitsyn *et al.* (3) , based on RDA of colon cancer and renal cell carcinoma in female DNA pairs, probes detecting LOH comprise around 70% of difference products, and probes detecting homozygous and/or hemizygous losses contribute 30%. On the other hand, Kern et al. presume that the efficiency ratio for the identification by RDA of detected fragments (comprising those within a homozygous deletion versus those within a site of simple LOH) will be 50 : 1 in their way of calculation (12).

When tumor DNA is used as tester, RDA might detect small restriction fragments acquired by the tumor as a consequence of mutations, genomic rearrangements, and the presence of DNA pathogens. It is also possible to clone highly amplified sequences since these cannot be completely subtracted by driver (normal DNA) and become enriched due to kinetic factors (4). Genomic rearrangement always creates a bridging fragment that is not present in the normal DNA. The probability that this fragment will be present in the amplicon is around 10% (10). Thus it is neccessary to use several different representations to have a high probability of finding a particular bridging fragment. One can calculate that for a six base cutter, roughly only 1 in 5000 point mutation will create a new amplifable fragment in a tumor. This is enough of an effect to be a nuisance.

Hence it follows that RDA can be a powerful method for detecting genetic lesions of cancer. However, we should carefully consider the experimental design as follows: (i) samples selected to prepare tester and driver should be as pure as possible. (ii) preparation of samples to bring out more valuable information, should be undertaken.

ACKNOWLEDGEMENT

We are grateful to the Japanese Cancer Research Resources Bank for providing cell lines. This work was financed by grants for Cancer Research from the Ministry of Education, Science, Sports and Culture.

REFERENCES

1. SOLOMON E, BORROW J, GODDARD AD. Chromosome aberrations and cancer. Science 1991, 254: 1153-1160.
2. LASKO D, CAVENEE W, NORDENSKJOLD M. Loss of constitutional heterozygosity in human cancer. Annu Rev Genet 1991, 25: 281-314.
3. LISITSYN N, LISITSYN N, WIGLER M. Cloning the differences between two complex genomes. Science 1993, 259: 946-951.
4. LISITSYN N, LISITSYN N, DALBAGNI G, BARKER P, SANCHEZ CA, GNARRA J, LINEHAN WM, REID BJ, WIGLER M. Comparative genomic analysis of tumors: detection of DNA losses and amplification. Proc Natl Acad Sci USA 1995, 92: 151-155.
5. CHANG Y, CESARMAN E, PESSIN MS, LEE F, CULPEPPER J, KNOWLES DM, MOORE PS. Identification of herpesvirus like DNA sequences in AIDS-associated Kaposi's sarcoma. Science 1994, 266: 1865-1869.
6. SIMONS JN, PILOT MATIAS TJ, LEARY TP, DAWSON GJ, DESAI SM, SCHLAUDER GG, MUERHOFF AS, ERKER JC, BUIJK SL, CHALMERS ML. Identification of two flavivirus-like genomes in the GB hepatitis agent. Proc Natl Acad Sci USA 1995, 92: 3401-3405.
7. AUSUBEL FA, BRENT R, KINGSTON RE, MOORE DD, SEIDMAN JG, SMITH JA, STRUHL K. (eds.) Current Protocols in Molecular Biology. John Wiley & Sons, New York, 1996.
8. TOYOTA M, CANZIAN F, USHIJIMA T, HOSOYA Y, KURAMOTO T, SERIKAWA T, IMAI K, SUGIMURA T, NAGAO M. A rat genetic map constructed by representational difference analysis markers with suitability for large-scale typing. Proc Natl Acad Sci USA 1996, 93: 3914-3919.
9. SCHUTTE M, da COSTA LT, MOSKALUK CA, ROZENBLUM E, GUAN X, de JONG PJ, BITTNER M, MELTZER PS, TRENT JM, KERN SE. Isolation of YAC insert sequences by representational difference analysis. Nucleic Acids Res 1995, 23: 4127-4133.
10. LISITSYN N, WIGLER M. Representational difference analysis in detection of genetic lesions in cancer. Methods Enzymol 1995, 254: 291-304.
11. WIELAND I, BOLGER G, ASOULINE G, WIGLER M. A method for difference cloning: gene amplification following subtractive hybridization. Proc Natl Acad Sci USA 1990, 87: 2720-2724.
12. SCHUTTE M, da COSTA LT, HAHN SA, MOSKALUK C, HOQUE AT, ROZENBLUM E, WEINSTEIN CL, BITTNER M, MELTZER PS, TRENT JM, KERN SE. Identification by representational difference analysis of a homozygous deletion in pancreatic carcinoma that lies within the BRCA2 region. Proc Natl Acad Sci USA 1995, 92: 5950-5954.