Nucleic Acids Research, 1995, Vol. 23, No. 20 4127–4133

Isolation of YAC insert sequences by representational difference analysis

Mieke Schutte, Luis T. da Costa, Christopher A. Moskaluk, Ester Rozenblum, Xiaoping Guan[1], Pieter J. de Jong[1], Michael Bittner[2], Paul S. Meltzer[2], Jeffrey M. Trent[2] and Scott E. Kern*

Departments of Pathology and Oncology, The Johns Hopkins Medical Institutions, Baltimore, MD 21205-2196, USA, [1]Human Genetics Department, Roswell Park Cancer Institute, Buffalo, NY 14263, USA and [2]Laboratory of Cancer Genetics, National Center of Human Genome Research, The National Institutes of Health, Bethesda, MD 20892, USA

Received June 27, 1995; Revised and Accepted September 6, 1995

ABSTRACT

We present a method for the isolation of YAC insert sequences by representational difference analysis (RDA). To achieve maximal representation of the sequences, the amplicons were generated from a *MboI* digestion product. RDA was performed using a 970 kb insert YAC clone. After two rounds of re-association and selective amplification 92% of the difference product represented sequences derived from the YAC insert. Twenty insert-specific sequence-tagged sites were readily defined. The difference product was also successfully used to isolate microsatellite markers, to identify clones from a human PAC library and as a chromosome painting probe in fluorescence *in situ* hybridization.

INTRODUCTION

A detailed physical map of a large genomic locus is commonly produced as part of a positional gene cloning effort. Often one starts with a YAC contig encompassing the region of interest. Once YAC end markers have been isolated, a region is subdivided into subregions. At this point the definition of sequence-tagged sites (STSs) from the YAC insert sequences is useful for further confirmation of the contig, the generation of additional contigs using other cloning vectors or high-resolution analysis of the region in test samples.

STSs from YAC inserts can be obtained by a variety of approaches. Total genomic DNA of the YAC-containing yeast clone can be subcloned and screened for insert-derived subclones with a probe of highly repetitive DNA elements from the same species. Alternatively, the YAC can be isolated by preparative pulsed field gel electrophoresis (PFGE), followed by direct subcloning or by subcloning after a 'whole-genome' PCR amplification using adaptor primers. Specific amplification of human or murine 'inter-Alu' sequences using primers specific for the consensus sequence of Alu elements is less tedious, but does not generate evenly spaced STSs due to the clustering of Alu elements in the genome. Also, Alu-derived STSs need not represent single copy sites within the genome. To some extent all methods suffer from the concurrent isolation of yeast DNA sequences.

Here we describe an alternative method for the isolation of YAC insert sequences by representational difference analysis (RDA) (1) of a YAC of interest and an irrelevant YAC from the same library. A related approach, termed marker addition through subtraction (MATS), has recently been described (2). RDA differs from MATS, and from other subtractive DNA hybridization methods, by the efficient exploitation of kinetic enrichment for target sequences. Target sequences are those present in one genome (the tester) but absent in another (the driver). The enrichment for target sequences in other methods is limited theoretically by the ratio of the two genomes in the hybridization reactions, but in practice frequently by the inefficiency of hybridization of complex DNA populations. In RDA the second order kinetics of self-association produces improvements in the enrichment for target sequences in the second and in additional rounds of hybridization and selection. Target sequences that have been enriched n times relative to the non-target sequences in the initial enrichment will self-associate n^2 times more rapidly relative to the non-target sequences in the subsequent hybridization reaction. In general, the potential of kinetic enrichment exists any time an enrichment step is followed by re-annealing and selection for duplexes of complementary tester sequences (tester–tester homoduplexes). In RDA, tester–tester homoduplexes are selected upon each hybridization reaction, thus maximizing the benefits of the kinetic component in the enrichment procedure.

Difference cloning of genomic DNA was first reported by Lamar and Palmer (3) and subsequently with variations on this theme by others (4,5). In their method re-annealed tester–tester homoduplexes were selectively isolated by cloning of the homoduplexes into plasmid vectors. Straus and Ausubel (6) and Wieland *et al.* (7) introduced the use of multiple rounds of

* To whom correspondence should be addressed

hybridization and selection, combined with a 'whole genome' PCR amplification (8) of the final difference products. Although not discussed in their report, the Straus and Ausubel procedure probably took advantage of kinetic enrichment to isolate the difference of two yeast clones (6). Wieland *et al.* achieved an enrichment of 100- to 1000-fold in a complex human genome after three rounds of hybridization and selection (7). They acknowledged the potential use of kinetic enrichments, but suggested that stochastic biases during the PCR amplification may present an obstacle to the implementation of the kinetic component. Practical obstacles in their procedure may have involved incomplete re-annealing reactions due to the high complexity of the human genome and an absence of selection for the re-annealed tester–tester homoduplexes. Lisitsyn *et al.* (1) solved these problems with the RDA technique, through a reduction of the genomic complexity by the use of representations of the genomes, combined with selective PCR amplification of only the re-annealed tester–tester homoduplexes after each hybridization reaction. They showed that combined subtractive and kinetic enrichment of the RDA can enrich the difference of two complex genomes over 1 000 000-fold after three rounds of hybridization and selection.

RDA was described for the isolation of the difference between two representative fractions of <10% of the original complex (mammalian) genomes (1). Representations, or 'amplicons', were generated by a PCR-based size selection of DNA fragments from both genomes upon digestion with a restriction endonuclease having a 6 bp recognition site. For the isolation of YAC insert sequences we modified the RDA procedure by generating amplicons upon digestion with *Mbo*I restriction endonuclease. The recognition site of *Mbo*I comprises 4 bp, generating short restriction fragments (9) that are amplifiable by conventional PCR. Thus the generated amplicons should represent maximal coverage of the genome of the original YAC-containing yeast clone, while the complexity remains below that used in RDAs performed on more complex genomes.

MATERIALS AND METHODS

Generation of amplicons

A 970 kb insert YAC clone (886d8) from the Généthon human YAC library (10), localizing to the *DPC/BRCA2* locus at chromosome 13q12 (11,12), was used as the tester (containing the target sequences, i.e. the YAC insert). An irrelevant YAC clone (934d4) from the same library, localizing to chromosome 14, was used as the driver. RDA was performed after modification of the original protocol (1). Genomic yeast DNA (1 µg) from both tester and driver was digested with the restriction endonuclease *Mbo*I (Gibco-BRL, Gaithersberg, MD) for 1 h at 37°C, according the manufacturer's recommendations. Following a phenol:chloroform (PC) (1:1, pH 8) extraction and an ethanol precipitation (13) with three washes with 70% ethanol, the restriction fragments were ligated to 0.5 nmol each of 12mer and 24mer unphosphorylated RBamHI oligonucleotides (Table 1) in 30 µl T4 DNA ligase buffer (New England Biolabs, Beverly, MA). T4 DNA ligase (400 NEB U) was added after an annealing step during which the mixture was gradually cooled from 50 to 10°C for 1 h and ligation was performed overnight at 16°C. An aliquot of 10 ng RBamHI-ligated DNA fragments was amplified by PCR in 67 mM Tris–HCl, pH 8.8, 4 mM $MgCl_2$, 16 mM $(NH_4)_2SO_4$, 10 mM β-mercaptoethanol, 100 µg/ml acetylated bovine serum albumin, 200 µM each dNTP, 1 µM RBamHI 24mer primer and 20 U Taq polymerase (Gibco-BRL) in a final reaction volume of 200 µl. The enzyme was pre-diluted in 1× PCR buffer and added after 3 min incubation at 72°C and the primer was added after another 3 min at 72°C. Amplification curves were constructed for both tester and driver DNAs, in which the templates were amplified using 1 min at 94°C and 3 min at 72°C for 10, 12, 14 and 16 cycles respectively, followed by a final extension of 5 min at 72°C in an OmniGene thermocycler (Hybaid, Middlesex, UK). PCR products were purified again by PC extraction, ethanol-precipitated with one 70% ethanol wash and 10% of each product was analyzed on a 1.5 % agarose gel. The number of cycles that generated the highest amount of product but did not show smearing of the product toward the well (indicative of single-stranded DNA) was chosen for further amplifications. PCRs to generate tester and driver amplicons were then performed with the above conditions for the optimal number of cycles. To ensure the production of double-stranded products an additional cycle of amplification and final extension was performed after addition of another 20 U Taq polymerase and 0.5 nmol primers (8). Nearly 10 µg amplicon was generated per 200 µl PCR reaction. The tester amplicon was size-selected by excision of DNA fragments larger than 200 bp upon electrophoresis on a 1.5% low melting point agarose gel (Sigma A3038, St Louis, MO) and subsequently purified by Qiagen chromatography (Qiagen, Chatsworth, CA). Isolated agarose blocks were mixed with 0.8 vol. 0.5 M MOPS buffer, 1 vol. 5 M NaCl and 5.2 vol. H_2O, fragmented by vortexing and melted at 70°C for 10 min. The warm mixture was loaded onto an equilibrated Q20 column, washed with 6 ml Qiagen wash solution and eluted with 800 µl Qiagen elution buffer. DNA was precipitated by addition of 650 µl isopropanol and subsequent centrifugation for 10 min at room temperature, followed by two 70% ethanol washes. DNA recovery was quantified by electrophoresis in a 1% agarose gel for 3 min at 100 V. The size-selected tester amplicon (20 ng) was re-amplified as above for 6, 8, 10 and 12 cycles respectively plus an additional cycle with fresh reagents and analyzed on a 1.5% agarose gel. Amplicons from the optimal PCR reaction were purified by PC extraction and ethanol-precipitated with one 70% ethanol wash. RBamHI primers were removed by *Mbo*I digestion of 5 µg tester amplicon and 120 µg driver amplicon, followed by PC extraction and a standard isopropanol precipitation (13) with three 70% ethanol washes. Amplicons were resuspended in TE and quantified by gel electrophoresis as above.

Table 1. Primer sequences

Primer	Sequence (5′→3′)
RBam24	AGC ACT CTC CAG CCT CTC ACC GAG
RBam12	GAT CCT CGG TGA
JBam24	ACC GAC GTC GAC TAT CCA TGA ACG
JBam12	GAT CCG TTC ATG
NBam24	AGG CAA CTG TGC TAT CCG AGG GAG
NBam12	GAT CCT CCC TCG
T3-20	ACG CCA AGC TCG GAA TTA AC
T3-24	ACC CTC ACT AAA GGG AAC AAA AGC
T7-20	AAA CGA CGG CCA GTG AAT TG
T7 promoter	TAA TAC GAC TCA CTA TAG GG

Comparative DNA hybridization

For the first round of subtraction the tester amplicon was ligated to JBamHI primers (Table 1) as above. Tester–JBamHI amplicon (500 ng) and driver amplicon (40 μg) were mixed in a volume of 200 μl TE, followed by PC extraction and isopropanol precipitation with two 70% ethanol washes and resuspended in ~3.5 μl 3× EE buffer (30 mM EPPS, 3 mM EDTA, pH 8.0; EPPS, Sigma E9502) to a final volume of 4 μl. The solution was overlaid with a drop of mineral oil and denatured at 98°C for 5 min. Pre-warmed 5 M NaCl (1 μl) was added, thoroughly mixed by pipetting and the DNA allowed to hybridize for 20 h at 67°C. The sample was diluted to a final volume of 500 μl with TE and 50 μl was amplified for 10 cycles of PCR as above, except that JBamHI 24mers were used and that Taq polymerase was added after 3 min at 85 instead of 72°C, to reduce priming mediated by duplexes of near-identical repetitive elements. PCR products were purified by PC extraction, ethanol-precipitated with one 70% ethanol wash and subsequently treated with mung bean nuclease for 30 min at 30°C in a final volume of 40 μl mung bean nuclease buffer containing 20 U enzyme (New England Biolabs). The reaction was stopped by addition of 160 μl 50 mM Tris, pH 8.9, and heat-inactivated for 5 min at 98°C. An aliquot (30 μl) of this solution was amplified by PCR as above for 10, 12, 14 and 16 cycles respectively plus the additional cycle with fresh reagents. The difference products were purified by PC extraction, ethanol-precipitated with one 70% ethanol wash and analyzed on a 10% polyacrylamide gel.

For the second round of subtraction the difference product of round 1 was digested and re-ligated to NBamHI primers (Table 1). For the second hybridization reaction 50 ng NBamHI-ligated round 1 difference product was mixed with 40 μg driver amplicon. Hybridization and selection were performed as in the first round.

A detailed protocol of this modified RDA procedure is available from the authors.

Analysis of the difference product

The RDA difference products of both rounds 1 and 2 were digested with *Mbo*I, ligated into *Bam*HI-digested and dephosphorylated pBluescript II plasmid vector (Stratagene, La Jolla, CA) and used to transform the DH10B strain of *Escherichia coli* by electroporation. Direct colony amplification was performed to screen for and to isolate inserts. Single colonies were resuspended in 20 μl TE. Half of this suspension was used to inoculate a culture and the other half was amplified as described above, but with 1 μM each of the T3-20 and T7-20 primers (Table 1) and 2.5 U Taq polymerase in a final volume of 50 μl. Incubation for 5 min at 94°C was followed by cooling to 78°C and the addition of Taq polymerase. PCR was performed for 35 cycles of 15 s at 94°C, 1 min at 62°C and 1 min at 72°C, followed by a final extension of 5 min at 72°C. PCR-amplified inserts were purified by PC extraction, ethanol precipitation and one 70% ethanol wash and used directly to prepare probes labeled by random priming for Southern blots containing 300 ng/lane tester and driver amplicon DNAs. Inserts were sequenced either directly after PCR amplification and purification or from plasmid DNA by the SequiTherm cycle sequencing kit (Epicentre Technologies, Madison, WI) using the T3-24 or T7 promoter primers (Table 1). STS primer sets (20mers) were designed from the isolated sequences and were analyzed by PCR on templates of total yeast DNAs from the tester and driver YAC clones, as well as on DNA from a monochromosomal somatic cell hybrid containing human chromosome 13 (Coriell Cell Repositories, Camden, NJ). PCR was performed as above, but with an annealing temperature of 58°C in a reaction volume of 15 μl.

Isolation of microsatellite markers

Bacterial colonies containing the cloned RDA round 2 difference product were screened with a 30mer CA repeat oligonucleotide. The oligonucleotide was labeled with [γ-^{32}P]ATP using T4 polynucleotide kinase (Gibco-BRL) and hybridized overnight at 42°C to membranes lifted from the plated colonies (14). Washes were performed using 3× SSC, 0.1% SDS at 55°C. Positive clones were sequenced from plasmids by cycle sequencing using the T3-24 or T7 promoter primers (Table 1).

Screening of a human PAC library

A difference product probe was prepared by removing the NBamHI adaptors from the RDA round 2 difference product by digestion with *Mbo*I enzyme and purification by PC extraction and isopropanol precipitation with one 70% ethanol wash. An Alu PCR probe was generated by amplification of genomic DNA from the YAC-containing yeast clone as described (15), using the PDJ34 Alu primer (16). The DNA fragments were labeled with [α-^{32}P]dCTP by random priming, pre-annealed with sheared human placental DNA and hybridized overnight at 65°C to PAC library filters (17). Washes were performed using 40 mM $NaHPO_4$, pH 7.2, 1 mM Na_2EDTA, 1% SDS at 65°C (18). PCR library screening was performed under the above-described conditions for STS primer sets.

Fluorescence hybridization *in situ* (FISH)

The RDA round 2 difference product was digested with *Mbo*I enzyme to remove the NBamHI adaptors and purified by PC extraction and isopropanol precipitation with one 70% ethanol wash. The difference product was labeled by an additional round of PCR amplification under the above-described conditions for RDA PCRs, with the substitution of the rhodamine-conjugated nucleotide SpectrumOrange dUTP (Vysis, Framingham, MA) for 14% of the dTTP in the reaction. A chromosome 13-specific painting probe (19) was labeled with biotin-16-dUTP (Boehringer Mannheim, Indianapolis, IN). The hybridization signal of the probe was detected using two layers of FITC-conjugated avidin (Vector Laboratories, Burlingame, CA) and amplified with one layer of anti-avidin antibody (Vector Labs). Two-color FISH of the difference product probe and the chromosome 13 painting probe was carried out as described (20).

RESULTS

RDA and non-polymorphic STSs

YAC insert sequences were isolated by RDA (1). The RDA procedure is diagrammed in Figure 1. RDA was performed on a YAC of interest (the tester) localizing to the *DPC/BRCA2* region on human chromosome 13q12 (11,12) and an irrelevant YAC from the Généthon library (the driver). The YAC-containing yeast clones were assumed to be identical except for the YAC inserts (the target sequence). Figure 2 illustrates the RDA

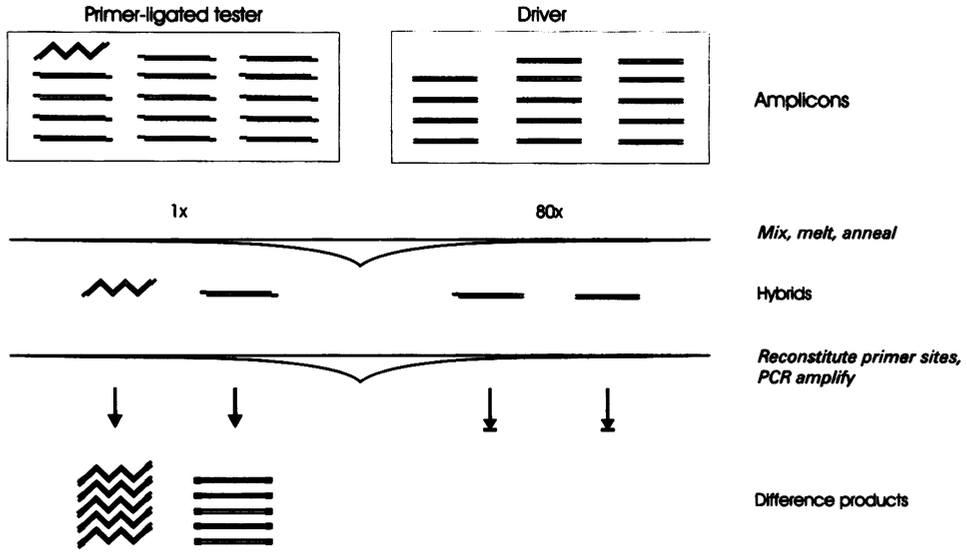

Figure 1. Schematic representation of the RDA procedure, after generation of the amplicons. Straight lines represent non-target DNA fragments, present in both tester and driver genomes; wavy lines represent target DNA fragments. Primer-ligated tester amplicon is mixed with an 80-fold excess of driver amplicon. Upon hybridization self-associated tester–tester homoduplexes are selected by exponential PCR amplification. The details of the procedure are described in Materials and Methods and in Lisitsyn *et al.* (1).

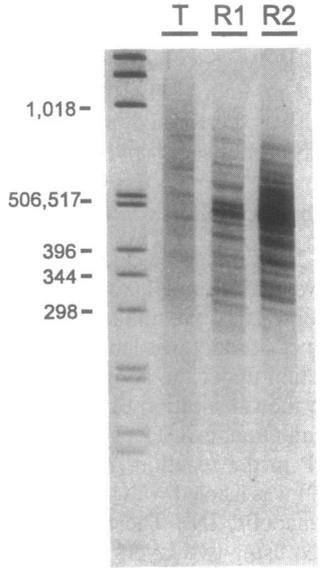

Figure 2. RDA of a YAC of interest and an irrelevant YAC from the same library. Lane T, amplicon of tester DNA after size selection. Lanes R1 and R2, difference product after the first and second rounds of hybridization and selection respectively. Sizes (in bp) of a 1 kb DNA ladder (Gibco-BRL) are indicated on the left.

difference products. Both round 1 and round 2 difference products produced a discrete pattern of bands.

Rounds 1 and 2 difference products were cloned into a plasmid vector and inserts were analyzed by using them as probes on Southern blots containing tester and driver amplicons. Six of 10 randomly selected fragments derived from the round 1 difference product hybridized with tester DNA only, indicating that they represented true subtraction products. The remainder hybridized to both tester and driver DNA, indicating that they represented either yeast and/or vector sequences that escaped the RDA subtraction or cross-hybridizing human sequences. With a target sequence of 970 kb in a yeast genome background of ~14 Mb this represented an enrichment during the first round of hybridization and selection of at least 9-fold.

From the round 2 difference product 25 randomly selected fragments were analyzed by Southern blot hybridization. Only one was found to have escaped the RDA subtraction, representing a non-target sequence. Subsequently the nucleotide sequences of eight of these fragments and of an additional 14 randomly selected fragments were determined. All 22 fragments were unique in their DNA sequence. STS primers were designed from these sequences and analyzed by PCR amplification of total yeast DNA of both tester and driver YAC clones and of DNA from a monochromosomal somatic cell hybrid containing human chromosome 13. Of the STSs that were already known to be true subtraction products based on the Southern blot analyses, all eight allowed amplification from DNAs of the tester and of the hybrid cell line, but not of the driver. Of the remaining 14 STSs 12 were true subtraction products and two had escaped the RDA subtraction as they amplified both tester and driver DNAs and not the hybrid cell line DNA. Altogether 39 fragments of round 2 difference product were analyzed, 36 (92%) of which represented true subtraction products.

Microsatellite markers

About 800 bacterial colonies containing cloned RDA round 2 difference product were screened with a $(CA)_{15}$ dinucleotide repeat oligonucleotide. Twelve CA-positive clones were identified, representing two unique fragments that both localized to human chromosome 13. In a contrasting approach we screened ~9000 phage plaques that contained subcloned genomic yeast

DNA from the same YAC clone using the $(CA)_{15}$ oligonucleotide. From 19 CA-positive phage clones we were able to determine the nucleotide sequences of 12 clones. These represented seven unique fragments, two of which localized to human chromosome 13, and five were yeast-derived sequences.

With an average insert size of 3 kb the phage mini-library represented a 1.8-fold redundancy of the YAC insert, whereas the difference cloning represented only a 0.3-fold redundancy. The RDA allowed a comparable number of markers to be found among a far less redundant library than the phage mini-library. This suggested difference cloning as a potentially more efficient approach for the isolation of microsatellite markers.

A considerable problem with the isolation of microsatellite markers from a phage library of subclones lies in the sequence analysis. The sequences are often determined by use of a $(CA)_{15}$ and a $(AC)_{15}$ primer with degenerate 3′-ends lacking cytosine or adenosine respectively. Despite this 3′ specificity, however, there still remains a likelihood of primer annealing to various sites on the dinucleotide repeat. Sequences from seven of the 19 phage clones could not be read. Microsatellite marker isolation from the difference cloning overcame these problems, since the lower library complexity readily allowed use of a smaller insert size. The complete insert sequence could then be determined by two sequencing reactions, using the universal primer sites of the vector.

Identification of human PAC clones

About 1.2×10^5 PAC clones containing human genomic sequences with an average length of 100 kb were screened by filter hybridization for clones containing inserts derived from the same region as the YAC insert. Thirty one PAC clones were identified with the use of the RDA round 2 difference product as a probe, consistent with the 3-fold genome redundancy of the PAC library (17). Twenty three PAC clones were identified when using an Alu PCR amplification product as a probe, 21 of which were included in the set of 31 clones that had been identified by the RDA difference product.

The PAC library was also screened by PCR using eight STS primer pairs that localized to an ~400 kb subregion of the 970 kb YAC insert (12, 21). These STS primers identified 17 clones in the PAC library. All 17 PAC clones had been identified by the RDA difference product. The Alu PCR-generated probe had identified 14 of these 17 PAC clones.

Fluorescence hybridization *in situ* (FISH)

The pancreatic carcinoma cell line COLO357 had been reported to contain the chromosomal translocation t(8;13)(p11;q12) and a modal chromosome number of 53, with numerous unidentifiable or partially identified marker chromosomes (22). In order to determine whether the reported break point was contained in the YAC 886d8 insert, localizing to the same chromosomal band (13q12) and encompassing the *DPC* region (12), we used the red-labeled RDA round 2 difference product and a green-labeled composite whole chromosome painting probe for chromosome 13 as probes to metaphase spreads of COLO357 cells. Figure 3 documents the presence of multiple copies of chromosome 13, all of which encompass YAC 886d8 and illustrates the complex involvement of chromosome 13 in marker chromosomes of this cell line. In addition to an apparently normal chromosome 13, three additional structural alterations of 13 were identified: ider(?)(13qter→13q10::?::13q10→13qter)X2; dic(13;13)(q34;p10); add(13p10).

The FISH result indicated that the translocation break point of COLO357 was located centromeric to the *DPC* region, as the chromosome 13 paint extended beyond the YAC insert paint. Comparison with the hybridization signals generated by a 'whole genome' amplification method for the chromosome 13 paint suggested a very well-preserved representation of the YAC insert in the RDA-generated probe. Hybridization of the RDA-generated probe to repetitive sequences, as usually observed at the acrocentric chromosomal arms or the telomeric (sub)regions, was undetectable.

DISCUSSION

We present an alternative method for the isolation of YAC insert sequences which employs RDA. The isolated difference product after two rounds of hybridization and selection contained a high representation of the YAC insert, and only 8% were non-target sequences that had escaped subtraction. The difference product was successfully used for the isolation of non-polymorphic STSs and of microsatellite markers, for the identification of PAC clones in the region of the YAC insert and as a chromosome painting probe in FISH.

RDA has been described for the isolation of the difference between two nearly identical complex genomes (1). For successful enrichment, amplicons are generated that represent <10% of the original mammalian genomes. For the isolation of YAC insert sequences, however, there is no necessity to reduce the genomic complexity, since the yeast genome is two orders of magnitude less complex than the mammalian genome. Genomic DNA of the YAC-containing yeast clones was digested with restriction endonuclease *Mbo*I, generating (in the human genome) DNA fragments with an average length of ~300 bp (9). The vast majority of these (~96%) are smaller than 1 kb (9) and thus amplifiable by conventional PCR. Amplicons thus generated should represent a nearly full coverage of the YAC insert and yeast genome sequences.

The presence of a significant fraction of small DNA fragments may present a problem in the enrichment procedure. Beyond a certain threshold the smaller fragments are presumably favored in the hybridizations, for as yet unknown reasons. Of note, the artifacts described by Straus and Ausubel likely refer to the same phenomenon (6). Size-selection of the tester amplicon for fragments larger than 200 bp successfully abolished selective enrichment of the smaller DNA fragments. We estimate that after this size-selection the amplicon may represent ~50% of the original genome (9). A precise estimate of the representation after RDA is difficult to make, but the intense signal observed in FISH when the round 2 difference product was used as a chromosome painting probe to metaphase spreads (Fig. 3) suggested that a high representation is preserved.

The use of *Mbo*I and the concurrent size-selection of the tester amplicon are modifications of the original RDA protocol (1) that we have introduced for the specific purpose of isolation of YAC insert sequences. Additionally, we have introduced a high-stringency incubation directly following the hybridization reactions. Preceding reconstitution of the primer annealing sites on the re-annealed DNA molecules, we incubate the hybridization mixture for 3 min at 85°C. At this temperature incompletely

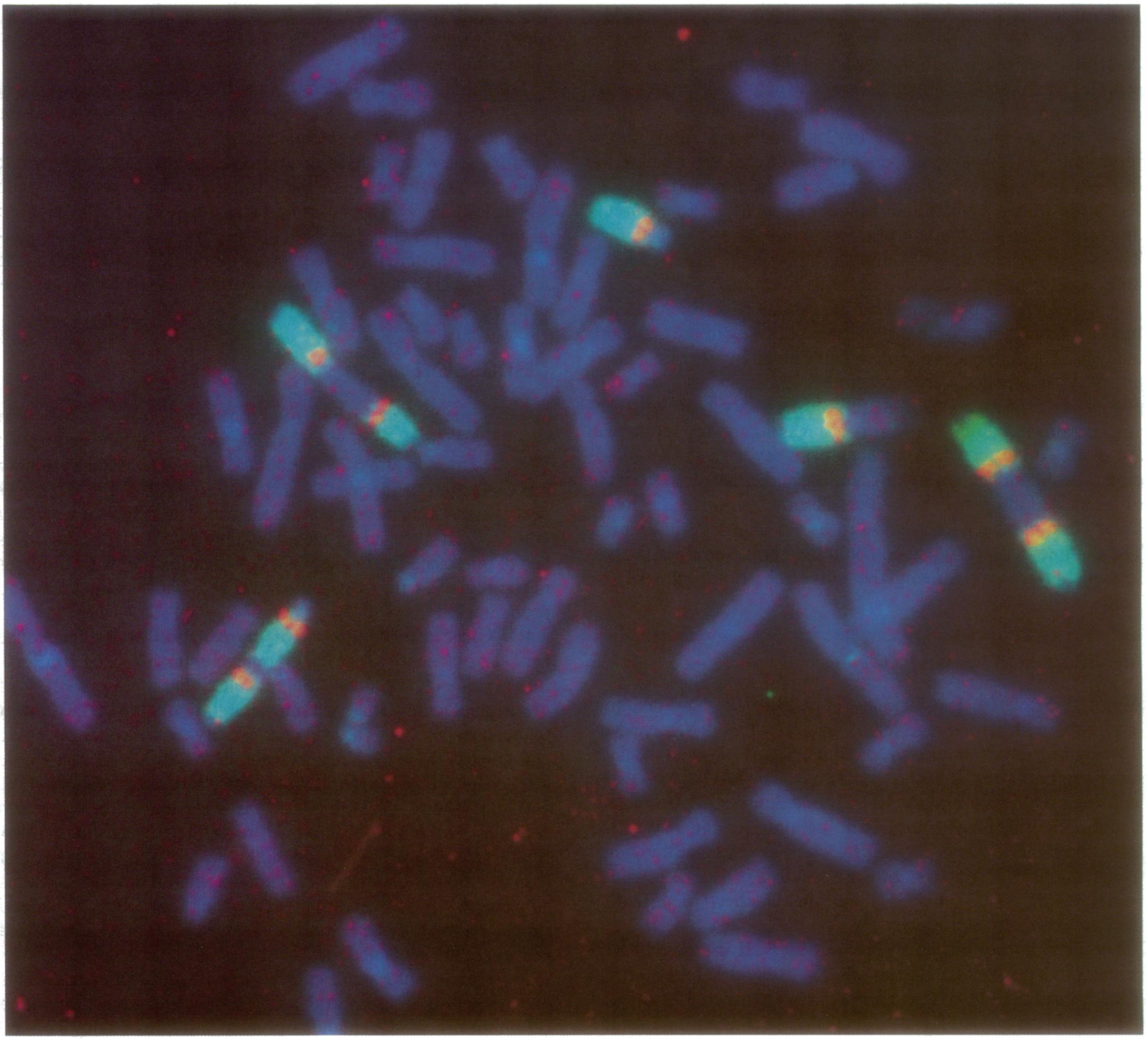

Figure 3. Metaphase chromosome spread from the COLO357 pancreatic carcinoma cell line. Two-color FISH was performed using the RDA difference product probe (in red) and a chromosome 13 painting probe (in green). The cell line contains a variety of rearrangements of chromosome 13q, all of which can be seen to be centromeric to the *DPC* region, which is contained within the YAC insert.

matched DNA sequences are presumably denatured and thus are no longer selected for in the difference product (M.Schutte, unpublished data). Indeed, upon analysis of over 100 cloned difference fragments from five RDAs we rarely identified repetitive sequences (M.Schutte, L.T.da Costa and S.E.Kern, unpublished data). This presumption is also supported by the absence of significant hybridization to acrocentric chromosomal arms or telomeric (sub)regions in FISH when using the RDA difference product as a probe.

Apart from these modifications, the presented protocol is essentially as described by Lisitsyn *et al.* (1). Some aspects of the procedure deserve attention. DNA is quantified at several steps in the RDA procedure. Accurate quantification becomes especially important for the hybridization reaction, since this reaction is driven by the DNA concentration (C_0t: $C_0 \times t$, where C_0 is the DNA concentration at the beginning of the hybridization reaction and t is the time of re-annealing). In essence, the amount of driver DNA in the hybridization reaction is determined by the limits of solubility of DNA (~10 μg/μl). We find the quantification of complex mixtures of DNA fragments most reliable by comparison with a DNA standard upon gel electrophoresis.

It is inherent in PCR technology that particular fragments are favored during the amplification reaction. To minimize such stochastic biases, tight control of the PCR is crucial. Amplification of a mixture of DNA fragments may generate high amounts of PCR product, inevitably exhausting the primers or polymerase.

At several points in the procedure we perform parallel PCRs with a range of cycle numbers. The optimal number of cycles will vary with each experiment, mainly due to inevitable variations in template input. The optimal number of cycles is determined as the PCR that generates the maximum amount of PCR product without showing smearing of the product towards the well. Smearing represents complexes of single-stranded DNA molecules with retarded gel mobilities, indicative of exhaustion of a PCR reagent. Optimal performance of PCR can minimize stochastic biases in the amplification reaction and thus add to the reproducibility of the RDA procedure. We previously found that two 'parallel' RDAs provided comparable results (12).

Each of the currently available techniques for the isolation of YAC insert sequences have their specific problems, suffering from contamination by yeast-derived sequences, isolated sequences that are not evenly spaced or represented or a procedural requirement for specialized skills or apparatus. Using RDA we found that after two rounds of hybridization and selection the difference product contained few yeast-derived sequences and no detectable repetitive sequences. For an experienced molecular biologist the generation of the amplicons takes a few days and each round of comparative DNA hybridization (as described in Materials and Methods) takes ~4 days, about half of which is hands-on time. Also, the RDA procedure is amenable to novel application. For example, the isolation of sequences from a consensus region of several YACs should be favored when pooled DNAs of these YAC clones are used as the tester. Since the non-consensus sequences would be present in a less than equimolar ratio, the kinetic component will be selectively impaired for these sequences.

Recently a similar procedure for the isolation of YAC insert sequences, termed MATS, was described (2). The percentage of true subtraction products using MATS varied from 60 to 75%. MATS uses one round of subtractive hybridization and the results were not significantly different from our results after one round of RDA (60%). The RDA procedure as used here may have an advantage over MATS in that ligation to adaptor primers is done before, rather than following, the hybridization reaction. Using RDA there is thus only a single step of the enrichment procedure wherein a low amount of template DNA might present technical difficulties, that being at the initial generation of amplicons from the genomic DNA. Also, RDA allows the ready manipulation of a high-stringency incubation following the hybridization reaction, which might reduce the selection for repetitive sequences in the subsequent PCR amplification (discussed above). Lastly, RDA becomes especially powerful in the second round of hybridization and selection. During the second round target enrichment is both subtractive and kinetic. It is the optimal exploitation of the kinetic enrichment that distinguishes RDA from previously described subtractive hybridization methods. We feel that the enrichment achieved here, where the target sequences represent 92% of the difference product, may be close to the practical limits of comparative genomic DNA hybridization procedures.

ACKNOWLEDGEMENTS

This work was supported by NIH SPORE grant in Gastrointestinal Cancer CA62924, by the Dutch Cancer Society KWF (MS) and by JNICT Scholarship BD1508/91-ID (LT da C). XG and PJ de J are supported by the Cancer Center Support grant at Roswell Park Cancer Institute CA-16056, by Department of Energy grant DE-FG02-94ER61883 and by NIH grant 1R01RG01165-01. SEK is a McDonnell Foundation Scholar.

REFERENCES

1 Lisitsyn,N., Lisitsyn,N. and Wigler,M. (1993) *Science*, **259**, 946–951.
2 Chen,H., Pulido,J.C. and Duyk,G.M. (1995) *Genomics*, **25**, 1–8.
3 Lamar,E.E. and Palmer,E. (1984) *Cell*, **37**, 171–177.
4 Kunkel,L.M., Monaco,A.P., Middlesworth,W., Ochs,H.D. and Latt,S.A. (1985) *Cell*, **82**, 4778–4782.
5 Nussbaum,R.L., Lesko,J.G., Lewis,R.A., Ledbetter,S.A. and Ledbetter,D.H. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 6521–6525.
6 Straus,D. and Ausubel,F.M. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 1889–1893.
7 Wieland,I., Bolger,G., Asouline,G. and Wigler,M. (1990) *Proc. Natl. Acad. Sci. USA* 87, 2720–2724.
8 Kinzler,K.W. and Vogelstein,B. (1989) *Nucleic Acids Res.*, **17**, 3645–3653.
9 Bishop,D.R., Williamson,J.A. and Skolnick,M.H. (1983) *Am. J. Hum. Genet.*, **35**, 795–815.
10 Cohen,D., Chumakov,I. and Weissenbach,J. (1993) *Nature*, **366**, 698–701.
11 Wooster,R., Neuhausen,S.L., Mangion,J., Quirk,Y., Ford,D., Collins,N., Nguyen,K., Seal,S., Tran,T., Averill,D., Fields,P., Marshall,G., Narod,S., Lenoir,G.M., Lynch,H., Feunteun,J., Devilee,P., Cornelisse,C.J., Menko,F.H., Daly,P.A., Ormiston,W., McManus,R., Pye,C., Lewis,C.M., Cannon-Albright,L.A., Peto,J., Ponder,B.A.J., Skolnick,M.H., Easton,D.F., Goldgar,D.E. and Stratton,M.R. (1994) *Science*, **265**, 2088–2090.
12 Schutte,M., da Costa,L.T., Hahn,S.A., Moskaluk,C., Hoque,A.T.M.S., Rozenblum,E., Weinstein,C.L., Bittner,M., Meltzer,P.S., Trent,J.M., Yeo,C.J., Hruban,R.H. and Kern,S.E. (1995) *Proc. Nat. Acad. Sci. USA*, **92**, 5950–5954.
13 Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
14 Fearon,E.R., Cho,K.R., Nigro,J.M., Kern,S.E., Simons,J.W., Ruppert,J.M., Hamilton,S.R., Preisinger,A.P., Thomas,G., Kinsler,K.W. and Vogelstein,B. (1990) *Science*, **247**, 49–56.
15 Aslanidis,C. and de Jong,P.J. (1990) *Nucleic Acids Res.*, **18**, 6069–6074.
16 Shutler,G., Korneluk,R.G., Tsilfidis,C., Mahadevan,M., Bailly,J., Smeets,H., Jansen,G., Wieringa,B., Lohman,F., Aslanidis,C. and de Jong,P.J. (1992) *Genomics*, **13**, 518–525.
17 Ioannou,P.A., Amemiya,C.T., Garnes,J., Kroisel,P.M., Shizuya,H., Chen,C., Baltzer,M.A. and de Jong,P.J. (1994) *Nature Genet.*, **6**, 84–89.
18 Church,G.M. and Gilbert,W. (1991) *Proc. Natl. Acad. Sci. USA*, **81**, 1991–1995.
19 Guan,X.-Y., Meltzer,P.S. and Trent,J.M. (1994) *Genomics*, **22**, 101–107.
20 Pinkel,D., Landegent,J., Collins,C., Fuscoe,J., Segraves,R., Lucas,J. and Gray,J. (1988) *Proc. Natl. Acad Sci. USA*, **85**, 9138–9142.
21 Schutte, M., Rozenblum, E., Moskaluk, C.A., Guan, X., Hoque, A.T.M.S., Hahn, S.A., da Costa, L.T., de Jong, P.J. and Kern, S.E. (1995) *Cancer Res.*, in press.
22 Morgan,R.T., Woods,L.K., Moore,G.E., Quinn,L.A., McGravan,L. and Gordon,S.G. (1980) *Int. J. Cancer*, **25**, 591–598.