PCR-based detection of a rare linear DNA in cell culture

Sergei V. Saveliev

Department of Biochemistry, University of Wisconsin-Madison, 433 Babcock Drive, Madison, WI 53706-1544, USA

Corresponding author e-mail: saveliev@biochem.wisc.edu

ABSTRACT

The described method allows for detection of rare linear DNA fragments generated during genomic deletions. The predicted limit of the detection is one DNA molecule per 10^7 or more cells. The method is based on anchor PCR and involves gel separation of the linear DNA fragment and chromosomal DNA before amplification. The detailed chemical structure of the ends of the linear DNA can be defined with the use of additional PCR-based protocols. The method was applied to study the short-lived linear DNA generated during programmed genomic deletions in a ciliate. It can be useful in studies of spontaneous DNA deletions in cell culture or for tracking intracellular modifications at the ends of transfected DNA during gene therapy trials.

INTRODUCTION

Analysis of the reaction intermediates and products is the traditional way to study chemical reactions. This approach is difficult to apply in studies of chromosomal DNA deletions or other genomic rearrangements. Reaction intermediates such as cleaved DNA, are transient and detecting them in a cell is a daunting task. In rare cases, the intermediates can be detected by Southern blot-hybridization. For example, open DNA ends generated during meiosis at hot genomic spots in yeast were identified by Southern blot-hybridization (1). However, the DNA reaction intermediates generated in other genomic rearrangements occur at levels that are well below the detection capability of blot-hybridization.

In the ciliate *Tetrahymena thermophila*, thousands of DNA sequences are eliminated in the developing macronucleus during programmed genomic deletions (2). In a recent study of

the deletions, we focused on linear forms of the deleted DNA sequences (3). Although previous results had suggested that the deleted DNA is immediately degraded by nucleases or converted into circles by DNA ligases, we found that it is present in cells in an intact linear form. The deleted DNA is converted to circles only occasionally (4). Studies performed in other labs demonstrated that the deleted DNA is packed inside of proteinaceous subnuclear structures after deletion (5). Apparently, these capsule-like structures protect the deleted DNA from DNA metabolizing enzymes. The DNA is then transferred to the cytoplasm and rapidly digested to nucleotides without producing intermediate degraded forms. Because the DNA is quickly transported to the cytoplasm after deletion and digested, it is present in trace amounts in cells and is not detected by Southern blot-hybridization (6). Although the particular steps of the deletion reaction in *Tetrahymena* may be unique, it is possible that DNA fragments generated during spontaneous chromosomal deletions in other organisms are released in an intact linear form as well. The deleted DNA with open DNA ends is a potential problem for the genome. It can easily integrate into a new chromosomal locus causing gene inactivation or, even worse, destabilization of chromosomes. The best way to avoid these events would be to immediately isolate the deleted DNA from chromosome. A protein capsule would serve this purpose. Incidentally, this capsule will protect the DNA from DNA metabolizing enzymes. This would create a unique opportunity to isolate the deleted DNA in an intact form.

The study of the deleted DNA can be very rewarding. For example, we found a distinct pattern of cleavage sites at ends of the deleted DNA fragments. Using the information obtained during study of the deleted linear DNA, we reconstructed several steps in the genomic deletion pathway (3). This paper describes in detail the protocol for detection of rare linear deleted DNA in cell culture. As PCR is the most sensitive

technique to detect DNA, we adapted it for the detection. Our method detects the linear DNA occurring at levels orders of magnitude below the detection limit of Southern blot-hybridization. At the same time, it is selective enough to detect the deleted DNA over the background of randomly broken chromosomal DNA fragments accumulated during DNA isolation. The most critical steps of the protocol are instant inactivation of cellular enzymes after cell lysis to prevent degradation of the deleted DNA and separation of the linear deleted DNA from chromosomal DNA before PCR. In the final step, all of the tremendous amplification capability of PCR is focused on a few molecules of the deleted DNA. The detection sensitivity of the method depends only on the ability to separate linear DNA from the genomic DNA. We also developed additional PCR-based applications to determine the structure of the ends of the linear DNA.

MATERIALS AND METHODS

Cell culture

Cell cultivation was carried out as described previously (7). Programmed genomic deletions were initiated by mating two different strains of *Tetrahymena*. The cell density was $2-3 \times 10^5$ cells/ml. Cells were collected by centrifugation at 1,000xg and resuspended to the final density of $2-3 \times 10^7$ cells/ml. Cells can be frozen in liquid nitrogen at this stage and stored at -70-80°C until use.

DNA isolation

DNA was isolated according to Austerberry and Yao (8), with modifications as specified below.

Cell lysis

It is important to note that, during preparation of cell culture for mating, we washed the cells in a buffer (10 mM Tris-HCl, pH 7.4) because the culture medium decreases the efficiency of cell lysis.

The frozen cells were first thawed at room temperature with constant mixing by inversion of the tube containing the cells. A 1 ml aliquot containing $2-3 \times 10^7$ cells (concentrated from 100 ml of cell culture) was mixed with an equal volume of hot (60°C) 2x lysis buffer [1% SDS, 0.5 M EDTA, 10 mM Tris–HCl final pH 9.5, 0.2 mg/ml pronase]. The pronase can be replaced by proteinase K at the same concentration. This lysis buffer is ideal for isolation of the linear DNA in an intact form because DNA metabolizing enzymes are immediately inactivated. Highly concentrated EDTA inhibits nucleases released upon lysis because the majority of DNA nucleases are divalent metal ion-dependent and hot SDS denatures cellular

enzymes. The enzymes are then digested with pronase during 3-4 hour incubation at 60°C.

Dialysis of the lysate

The next step is DNA isolation. Normally, the best way to isolate DNA from a lysate is to precipitate it with ethanol. However, our experience showed that, because of their high concentration in the lysis mixture, SDS and EDTA co-precipitate with DNA. To remove SDS and EDTA the lysate was dialyzed against 5 liters of TE buffer (1 mM EDTA, 10 mM Tris-HCl, pH 8). As a rule, dialysis is performed in a cold room (4°C) to avoid degradation of biological substance. However, SDS quickly precipitated after transfer of the solution (in the dialysis bag) into the cold dialysis buffer and the precipitate did not leave the dialysis bag during dialysis. To solve the problem, the dialysis was initiated in a warmer (room temperature) dialysis buffer. The beaker with the buffer and lysate was then transferred to the cold room. The dialysis buffer slowly cooled down in the cold room during dialysis giving SDS enough time to leave the bag. The molecular weight cut off of the dialysis membrane was 12-14 kDa. After 4-5 hours of dialysis, the bag was transferred into fresh dialysis buffer to ensure full removal of SDS and EDTA. The second dialysis was carried out overnight in the cold room. The volume of the DNA solution increased about fivefold (from 2 ml to 10-12 ml) after dialysis. The solution was viscous because it contained chromosomal DNA fragments of high molecular weight.

RNA digestion

DNA is not yet ready for loading onto a gel. Approximately 3/4 of the nucleic acids in the solution is RNA, which must be removed to prevent interference with DNA fragments during electrophoresis. Before digesting the RNA, the solution was cleared of proteinase. Proteinase self-digested after several hours of incubation at 37°C. However, a sufficient amount of proteinase remained, capable of digesting the RNAse. The proteinase was removed by two consequent extractions with 3 volumes of chloroform: isoamyl ethanol (24:1). Aqueous phase and chloroform were separated by centrifugation at 4,000xg. To remove the traces of chloroform, the DNA solution was dialyzed against TE in the cold room for 6 hours. Finally, nucleic acids were precipitated with 3 volumes of ethanol in the presence of 2M ammonium acetate at room temperature. Several inversions of the solution provided adequate mixing of DNA/RNA solution and ethanol. The concentration of nucleic acids was high at this stage and it efficiently precipitated during the mixing. The nucleic acids can be stored indefinitely at 4°C in the ethanol-precipitated form.

In the next step, the nucleic acids were collected by centrifugation at 14,000xg for 15 min. The pellet was washed

in 80% ethanol to remove salt and redissolved in 1 ml of TE. The RNA was digested with 100 µg/ml DNAse free RNAse A at 37°C for 2 h. RNAse was removed from the solution by extraction with chloroform: isoamyl ethanol and DNA was precipitated with ethanol in the presence of ammonium acetate as described above. It is imperative that the salt is removed from the pellet by washing it with 80% ethanol. Otherwise, it will impact mobility of DNA during electrophoresis. The yield of DNA was about 150 µg from 2-3x10^7 cells.

Purification of linear forms of deleted DNA using an agarose gel

The isolated DNA was redissolved in 200 µl of TE, mixed with 100 µl of loading buffer (0.25% Bromphenol blue, 0.25% Xylene Cyanol, 20% Ficoll) and loaded onto a 1.5% horizontal agarose gel. The DNA must be fragmented by pipetting from a narrow tip before loading onto the gel. The solution contains high molecular weight DNA molecules at this stage. As a result, the solution is viscous, and it can leak on top of the gel after loading. Pipetting the solution through a narrow tip reduces the size of the DNA fragments to 20-40 kbp. Also, DNA and loading buffer should be properly mixed. The concentration of DNA is relatively high and it does not mix easily with loading buffer. The mix should be slowly loaded to the bottom of the well to allow for complete filling-in of the well. In our studies, the predicted size of the linear deleted DNA was 0.6 and 0.9 kbp (Fig. 1) and the 1.5% gel was the most suitable for resolving DNA fragments in that range. A gel of lower percentage should be used if the expected size of the deleted DNA is bigger. A 9 mm thick gel with 1 mm x 5 mm slots provided the best resolution of the DNA. The amount of the loaded DNA was 20 µl of the solution per lane (about 10 µg DNA, equivalent to the DNA extracted from about 2 million cells). A 100 bp DNA ladder (Promega) was used as a DNA marker. The electrophoresis was run at 4 °C and 3 V/cm. When electrophoresis was complete, the gel was stained with ethidium bromide to visualize the DNA markers. The deleted linear DNA is invisible after staining of the gel because of its low abundance. Hence, the DNA extraction from the gel was done "blind," guided only by the molecular size markers in another gel lane. The gel lanes were cut into slices. The each slice contained DNA fragments of a particular size. Gel slices containing DNA fragments of identical size were pooled. The DNA was extracted from the slices using the QIAquick gel extraction kit and columns (Qiagen). The final volume of the each gel fraction after the column purification was 300 µl (in TE buffer). The steps are shown in Fig. 1.

Screening of the gel fractions for linear forms of the deleted DNA

PCR is required to detect the deleted DNA in the gel extract because the DNA is present in trace amounts. In a standard PCR protocol, two primers anneal to the targeted DNA sequence. Then, the segment of DNA between the primers is amplified by PCR. In the experiments described here, regular

PCR generated a high background of false signals in all fractions because the gel fractions were contaminated with randomly broken chromosomal DNA fragments (data not shown). Although the amount of contaminating DNA fragments was low, PCR amplified the contaminants at a frequency that competed with the signals generated by the rare DNA species we were trying to detect.

To solve this problem, the standard PCR protocol was replaced with anchor PCR (Fig. 1). In anchor PCR, a DNA oligonucleotide is attached to an end of the linear DNA. This oligonucleotide is called the "anchor." The DNA is then amplified by PCR. The anchor oligonucleotide is used as a second PCR primer. From our earlier studies, it was clear that the developmentally programmed deletions in *Tetrahymena* generate open DNA ends with 4 nt 5' extensions at the ends of the deleted sequences (9). Therefore, we designed a linker, which was compatible with the anticipated end structure at the end of the deleted DNA. The linker consisted of two oligonucleotides. One of them was an anchor oligonucleotide. It was attached to an end of the DNA by ligation (Fig. 1). The protocol is a modification of ligation-mediated PCR developed by Mueller and Word in 1989 (10). If the structure of the DNA ends of the linear DNA being investigated is unknown, there are additional variations of anchor PCR that can be applied for detection (discussed in Results and Discussion).

Detection of rare linear deleted DNA in Tetrahymena cell culture
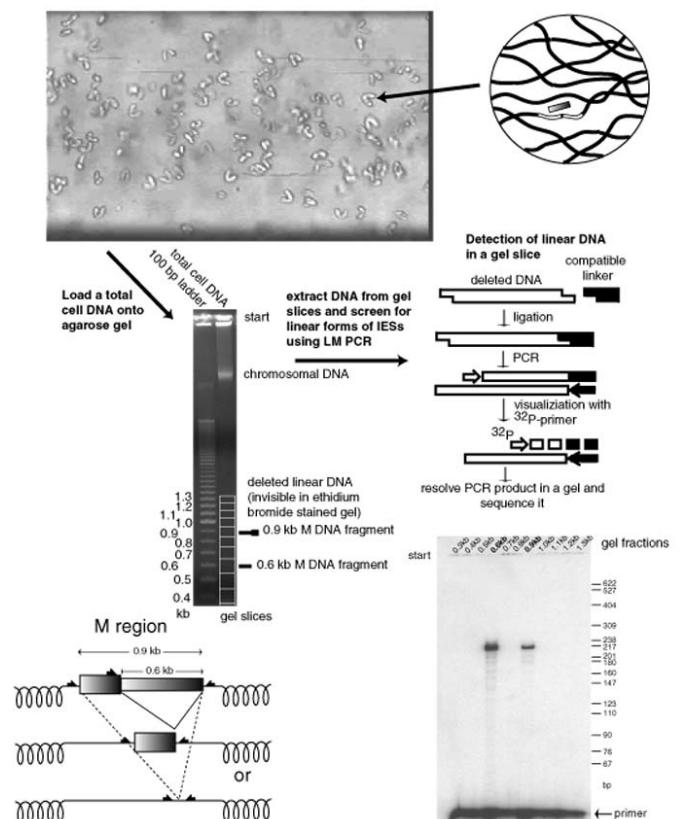
Fig. 1: The protocol for detection of linear forms of deleted DNA. The top panel shows a culture of *Tetrahymena thermophila* mating cells. During mating, the parental somatic

nuclei (macronuclei) are destroyed and new ones develop from copies of the germ line micronuclei. The process entails an elaborate set of DNA rearrangements, including the deletion of 6,000 short DNA fragments from the genome, called IES (internal eliminated sequence) elements. Because the linear deleted DNA product is short-lived only a minor population of cells contains it at any given moment. In accordance with our estimation one cell only (a random cell is indicated by the arrow for the purpose of demonstration) from the cell population shown on the panel contains the deleted DNA. Our work focused on an IES element called the M region (shown at the bottom). Two alternative deletions of 0.6 and 0.9 kbp occur in the M region. The deleted sequences share the same right boundary.

The detection protocol includes several steps. After cell lysis and purification, the total cellular DNA is loaded onto an agarose gel. The linear form of the deleted DNA is invisible after staining the gel with ethidium bromide due to its low abundance. The expected positions of the 0.6 and 0.9 kbp deleted DNA fragment are marked at the right. The gel is cut into slices designated by rectangles. The DNA is extracted from the gel slices, subjected to ligation-mediated PCR as shown (details in the text) and visualized in a ^{32}P-extension step. The gel with ^{32}P-labeled PCR products is shown at the bottom. The ^{32}P-extension primer anneals about 190 nt upstream from the right end of 0.6 and 0.9 kbp deleted DNA fragments. Therefore, both of the amplified DNA fragments are visualized as a 220 nt (extended sequence plus length of ^{32}P-labeled primer) ^{32}P-extended product. The linear DNA was found in the gel fractions containing 0.6 and 0.9 kbp fragments as expected (shown). In addition to the full-size linear DNA, we detected minor bands below the 0.6 and 0.9 kbp fragments. These bands mark sites of preliminary termination of the polymerase during ^{32}P-primer extension rather than partially degraded fragments of linear DNA.

The detection protocol included ligation, amplification and visualization steps (Fig. 1). Ligation of the linker to the linear DNA was carried out at 16°C for 4 h in 15 µl of mixture containing 0.5 µl of the extract from the gel fraction containing the linear DNA (equivalent of the DNA isolated from about 40,000 cells), 0.01 Weiss unit of T4 DNA ligase (Promega)/µl, 0.3 µM of the compatible linker, 30 mM Tris–HCl pH 7.8, 10 mM MgCl$_2$, 10 mM dithiothreitol (DTT) and 1 mM ATP. A 5 µl aliquot of the ligated mixture was mixed with 45 µl of PCR buffer (0.45 µM each of anchor oligonucleotide and amplification primer, 0.22 mM each of dGTP, dATP, dTTP, and dCTP, 2.78 mM MgCl$_2$, 0.055 M KCl, 0.011 M Tris-HCl (pH 9.0), 0.11% Triton X-100, 11% glycerol). The mixture was incubated at 94°C for 2 min and 55°C * for 2 min, followed by addition of 0.4 µl of *Taq* DNA polymerase (5 U/µl)(Promega) (*all of the primers used in our studies were designed in such a way that 55°C was the optimal temperature for the annealing step in all of the PCR reactions). Our experiments showed that completion of hot step with incubation at 55°C improved reproducibility of the amplification. PCR primers annealed with DNA during incubation at 55°C. This ensured that *Taq*

polymerase extended properly annealed primers during the short period between addition of the polymerase in the reaction mixture and beginning of PCR. The mixture was then incubated at 94°C for 1 min. Then the reaction immediately proceeded into the PCR cycle. PCR was run in an MJ Research PTC-200 Peltier Thermal Cycler for 25 cycles. This cycler was used in every protocol below. Mineral oil was not required because evaporation of the solution in PCR tubes was prevented in this cycler. Each cycle consisted of 10 s at 94°C, 45 s at 55°C and 1 min at 72°C. Cycling was concluded with a final extension at 72°C for 7 min.

Trials indicated that 0.01 Weiss U/µl was the optimal concentration of the ligase in the reaction. Higher concentration led to increased accumulation of nonspecific PCR products in the protocol because the ligase, at high concentration, forced nonspecific ligation of the linker with randomly broken chromosomal DNA fragments that were present as contaminants in the extract (data not shown). An increase of the ligation time above 4 hours is not recommended. Commercially available ligases contain trace amounts of nucleases, which are not detectable if nanogram or higher amounts of DNA are ligated. The linear DNA is present in the reaction at much lower concentration and the nucleases can pose a threat to the integrity of the linear DNA if the reaction is incubated for a longer period.

In our experiments, PCR generated high amount of nonspecific products that looked like a smear after analysis of the products in agarose gel. Thousands of DNA sequences are deleted during programmed genomic deletions in *Tetrahymena* (2), and many of them are likely to have lengths comparable to the ones we were interested in isolating. In addition, thousands of DNA fragments are generated during apoptotic degradation of nuclei accompanying meiosis in the same conjugating *Tetrahymena* cells. The linker ligated to some of these nonspecific DNA fragments. This led to the generation of nonspecific PCR products that outnumbered the specific products hundredfold (data not shown). To visualize the amplified linear DNA in the mixture with nonspecific PCR products, the ^{32}P-extension technique was employed (Fig. 1). A ^{32}P-labeled oligonucleotide annealed with the amplified linear DNA. In the next step, the oligonucleotide was extended with use of DNA polymerase and the extended product was identified by PAGE analysis. The ^{32}P-extension step increased the detection sensitivity by 100 fold in comparison with visualization by ethidium bromide. Only the specific PCR product was detected. To prepare PCR products for ^{32}P-extension, they were precipitated with 3 volumes of ethanol in the presence of 2M ammonium acetate. The concentration of DNA was extremely low, so glycogen was added to the final concentration of 100 µg/ml. Glycogen helped to fully precipitate trace amounts of DNA. DNA precipitation was additionally stimulated by exposure of the solution to a low temperature (-70°C) for 2 h. The DNA was collected by centrifugation at 14,000xg for 15 min in the cold room. The DNA precipitated in the form of a tiny dot at the bottom of the tube. The supernatant was slowly and carefully removed from the tube with a pipette to preserve the pellet. The pellet was washed with 80% ethanol to remove salt and collected again by centrifugation at 14,000xg at 5 min

in the cold room. Again, the supernatant was removed with the great care.

The oligonucleotide used in ^{32}P-extension was labeled with T4 DNA kinase and [γ^{32}P]ATP according to a protocol provided by Amersham. Amersham Redivue [γ^{32}P]ATP is the most suitable reagent for this purpose. The solution contains inert red dye and the radionucleotide can be visually tracked down during operation steps.

The precipitate was resuspended in 20 μl of extension buffer [40 pmol/ml [^{32}P]-labeled visualization primer, 0.2 mM each dGTP, dATP, dTTP and dCTP, 2.5 mM MgCl$_2$, 50 mM KCl, 10 mM Tris-HCl (pH 9.0), 0.1% Triton X-100, 40 U/ml of Sequencing grade *Taq* DNA polymerase]. The tube with the reaction mixture was transferred to the cycler, preheated to 94^{0}C. Extension was performed for 5 cycles. Each cycle consisted of 10 s at 94°C, 45 s at 55°C and 1 min at 72°C. After extension was complete, the tube was placed on ice. The reaction mixture was mixed with 10 μl of 7.5 M ammonium acetate and 7.5 μl of distilled water. The PCR product was precipitated with 3 volumes (112.5 μl) of ethanol at -70°C for 2 h. The precipitation was required to remove glycerol from the solution, which would otherwise impact the mobility of the PCR product in the polyacrylamide gel in the next step. Glycogen was not added because it remained in the solution from the previous step. The PCR products were collected by centrifugation. Again, the pellet was handled with the great care. The pellet was washed with 80% ethanol and redissolved in 15 μl TE. A portion of the mixture (5 μl) was mixed with 3 μl of sequencing stop solution (11), boiled for a few minutes to denature the DNA and analyzed by electrophoresis in a 6% PAGE (sequencing gel, 0.33 mm thick) under denaturing conditions (11). The gel was preheated at 35-40 W for 1 h to avoid renaturation of the DNA strands during electrophoresis. After electrophoresis (at 35 W), the extended products were visualized by exposing the gel with X-ray film. The amplified linear DNA was studied by sequencing, as described in **Sequencing of the PCR products** below.

Determining of the structure of DNA ends

Mapping of 5' termini

The protocol for mapping the 5' termini at ends of the linear DNA is similar to the ligation-mediated PCR protocol described above. The additional steps are denaturation and extension (Fig. 2). The method allows for the detection of the 5' termini irrespective of the end structure of the DNA fragment. The yield of the PCR product in this protocol was particularly low and additional PCR with nested primers was frequently required to bring amplification to a detectable level. Low specificity of primer annealing during the extension step and/or poor ligation of the blunt-end linker with the 5' terminus may account for the decreased efficiency of the protocol. Also, some of the ends generated by *Taq* polymerase in the extension step were probably sticky. The polymerase adds an extra 3' nontemplate-directed nucleotide during DNA synthesis at a

low efficiency (12). The generated DNA end is sticky and does not ligate with the blunt-end linker. To compensate for decreased efficiency of the detection, the amount of the gel extract per reaction was increased to 5 μl.

The protocol was performed as follows. A 5 μl aliquot of the gel extract containing the linear DNA (equivalent of the DNA isolated from about 400,000 cells) was mixed with 0.1 pmol of the extension primer. The reaction volume was brought to 50 μl by adding water and MgCl$_2$ to 2.5 mM, dGTP, dATP, dCTP and dTTP, each to 0.2 mM, KCl to 50 mM, Tris-HCl pH 9.0 to 10 mM, Triton X-100 to 0.1% and glycerol to 10%. The DNA was denatured by incubation of the mixture at 94°C for 2 min. The annealing of the extension primer to the linear DNA was performed at 55°C for 20 min. The annealed primer was then extended with one unit of Sequencing grade *Taq* polymerase for 7 min at 72°C. The polymerase generated a blunt end at the 5' terminus. *Taq* polymerase was removed by treatment with 3 volumes of chloroform: isoamyl ethanol (twice). DNA was precipitated with ethanol in the presence of ammonium acetate and glycogen, then washed with 80% ethanol. DNA was dissolved in 10 μl TE. Ligation was carried out at 16°C overnight in 15 μl of mixture containing 10 μl of the extended DNA, 0.2 Weiss U/μl of T4 DNA ligase (Promega), 1 μM of the blunt-end linker, 30 mM Tris-HCl pH 7.8, 10 mM MgCl$_2$, 10 mM dithiothreitol (DTT) and 1 mM ATP. The concentrations of the ligase and the linker as well as the incubation period were increased in this protocol to compensate for the poor ligation efficiency between the blunt DNA end and the linker. A 5 μl aliquot of the ligation mixture was amplified with the anchor primer and the amplification primer in 35 PCR cycles as above. After the amplification was complete, 45 μl of the PCR mixture were precipitated with ethanol, ammonium acetate and glycogen, washed with 80% ethanol and visualized with [^{32}P]-extension primer. 5 μl of the PCR mixture was saved for nested PCR. The extended products were analyzed in 6% PAGE under denaturating conditions.

If no PCR product was visualized in the gel after extension with [^{32}P]-labeled extension primer, the PCR product was reamplified with nested primers. A 0.5 μl aliquot from the saved 5 μl of the PCR reaction was transferred into the second PCR mixture. Amplification was performed as described for the first PCR reaction except that the amplification primer was replaced with the nested primer, which was positioned closer to the end of the linear DNA (Fig. 2). The number of cycles was decreased to 30 to lower the yield of nonspecific PCR products, which inevitably accumulated during the PCR reaction. PCR products were precipitated with ethanol, ammonium acetate and glycogen, washed with 80% ethanol and visualized with [^{32}P]-extension primer. The products were analyzed in the polyacrylamide gel and the 5' termini were mapped by sequencing as described in **Sequencing of the PCR products** below.

Mapping of 3' termini

In principle, an anchor oligonucleotide can be attached to the 3' terminus of the DNA either by ligation, with RNA ligase (13), or in a reaction involving terminal deoxynucleotidyl transferase (TdT) (14). Our attempts to detect 3' termini using RNA ligase were unsuccessful (data not shown). In contrast, the TdT-mediated protocol was a reliable technique. The TdT-mediated protocol is illustrated in Fig. 2. Note that the anchor primer is attached in the different way than in the ligation-mediated protocol. However, the oligonucleotide is still the anchor primer because it is attached to the terminus. A disadvantage of the TdT-mediated protocol is that, in the first step, TdT synthesizes a polynucleotide tail at 3' termini of all DNA molecules present in the solution including contaminating DNA. PCR generates an enormous amount of nonspecific products. The specific PCR product is rarely detectable in this mix even after ^{32}P-extension step (data not shown). The second PCR reaction with the nested primer is required to bring the specific PCR product to a detectable level (Fig. 2). Although inefficiency might be expected, the method is robust and more sensitive than the ligation-mediated PCR.

One third (100 µl) of the gel extract containing the linear DNA was used in the detection protocol. The preliminary experiments indicated that the gel extract prepared using the Qiagen column contained an unidentified activity that inhibited TdT. The activity was removed by treatment of the extract with 2 volumes of chloroform:isoamyl ethanol. After the treatment with chlorhorm:isoamyl ethanol, the DNA was precipitated with ethanol, ammonium acetate and glycogen, then washed with 80% ethanol. The pellet was redissolved in 100 µl TE. The solution was brought to 125 µl by adding dGTP to 0.01 mM, cacodylate buffer (pH 6.8) to 100 mM, $CoCl_2$ to 1 mM and DTT to 0.1 mM. Cacodylate buffer and $CoCl_2$ were supplied with TdT by Promega. The reaction was initiated by adding TdT to the final concentration of 0.2 U/µl. The mixture was incubated at 37°C for 30 min. TdT generated a poly(dG) tail of 50-100 nt (data not shown). DNA was precipitated with ethanol and ammonium acetate. Glycogen was not added because it was present from the previous step. The pellet was washed with 80% ethanol and redissolved in 100 µl TE. The DNA was then amplified in the reaction using the two types of anchor primers. The first anchor oligonucleotide had poly(dC) tail. The tail annealed with the poly(dG)tail generated at 3' termini of the DNA. *Taq* polymerase extended the primer. The extension generated a DNA strand with the anchor oligonucleotide at one end. This strand was then amplified with use of the anchor primer without poly(dC) tail and the amplification primer. The reaction was performed as follows. 0.5 µl of the solution of the TdT-treated DNA (equivalent of the DNA isolated from about 40,000 cells) was mixed with 50 µl of the PCR mixture containing 0.2 mM each of dGTP, dATP, dTTP and dCTP, 2.5 mM $MgCl^2$, 10 mM Tris–HCl pH 9.0, 0.1% Triton X-100, 10% glycerol and 0.002 µM anchor primer with poly(dC) tail. The mixture was incubated at 94°C for 2 min and the anchor primer was annealed with the DNA at 50°C for 30 min. Then, 2 µl of the 10 µM solution of the anchor primer without the poly(dC) tail, 2 µl of 10 µM solution of amplification primer and 0.4 µl of Taq DNA polymerase (5 U/µl) were added directly to the 50°C solution.

At this point, the only primer annealed to the DNA was the anchor primer with the poly(dC) tail. Two other primers, the anchor primer without poly(dC) tail and the amplification primer, were not bound to the DNA yet. After extension of the anchor primer with the poly(dC) tail (72°C for 3 min) the mixture immediately proceeded for 35 PCR cycles because the PCR primers were already present in the mixture. Each PCR cycle consisted of 15 s at 94°C, 45 s at 55°C and 1 min at 72°C. Amplification was concluded with a final extension at 72°C for 7 min. Because the PCR products were rarely detected after this step, the mixture was reamplified as follows. 0.5 µl from the first PCR mixture was transferred into the second PCR mixture. The second PCR mixture contained the nested amplification primer and the anchor primer without poly(dC) tail. The other components were as in the first PCR reaction. The number of cycles was decreased to 30 to lower the yield of nonspecific PCR products. PCR products were precipitated, visualized with the [^{32}P]-extension primer and the 3' termini were mapped by sequencing as described below. Because of the length variability of the poly(dG) tail in the individual PCR molecules, the PCR product appeared as a smear after gel electrophoresis (Fig. 2).
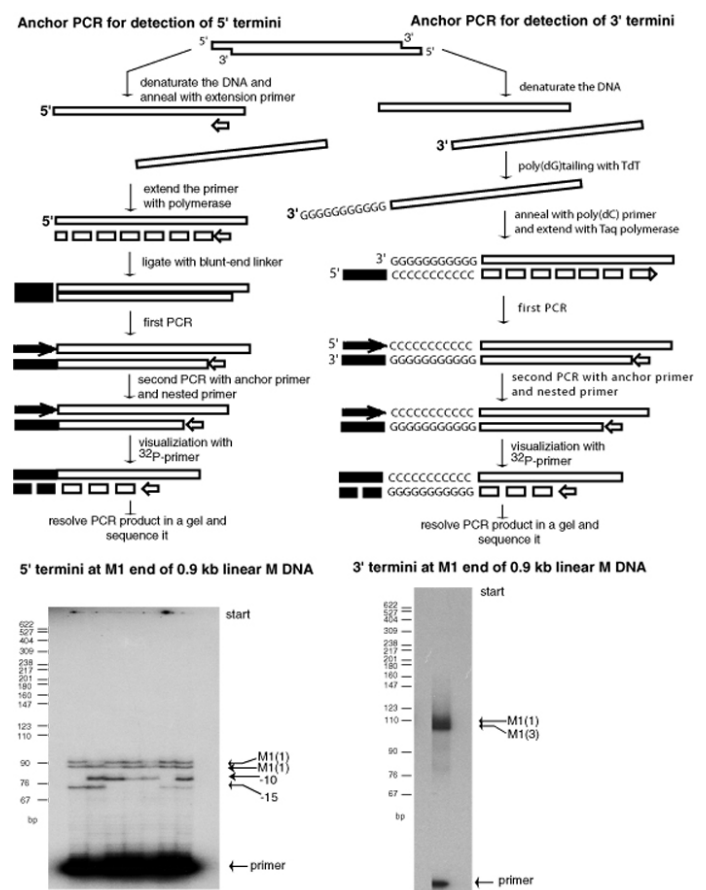
Fig. 2: The protocols for mapping the 5' and 3' termini at the ends of the linear DNA. DNA denaturation is required in both protocols to expose the termini (details in the text). The anchor oligonucleotide is attached to the 5' terminus by ligation. Attachment of the oligonucleotide to the 3' terminus occurs in two steps – generation of poly(dG) tail and annealing

of the poly(dC)-anchor oligonucleotide. The first PCR amplification rarely yields a detectable PCR product. The second PCR reaction with an anchor and nested primers brings the PCR product to a detectable level. The linear form of the deleted M region DNA turned out to be a mixture of individual DNA fragments, which have essentially the same size but differ slightly in the position of the terminal residue at the end of one strand. Four 5' termini found at the M1 end of the 0.9 kbp M fragment are shown. PCR products generated in the TdT-mediated protocol appear as a smear. The smearing effect is caused by the fluctuating length of the poly(dG) tail. Sequencing recovered two 3' termini at the M1 end of the 0.9 kbp linear DNA.

Sequencing of the PCR products

Sequencing is required to map the DNA termini. The technique is relatively easy. There is no need to clone the PCR product. Sequencing of short PCR products can be tricky, generating patterns of bands in the sequencing gel that are difficult to read. However, the method is readily tuned by modification of reaction conditions such as DNA concentration and primer annealing temperature.

The protocol includes isolation of the ^{32}P-labeled PCR product, reamplification of the product, gel purification of the reamplified PCR product and, finally, sequencing of the product. The steps were performed as follows: a 10 µl aliquot of the ^{32}P-labeled product from the final extension reaction of any of the detection protocols, was mixed with 4 µl of sequencing stop solution (11) and loaded onto a 6% PAGE gel (0.33 mm thick) under denaturing conditions (11). After electrophoresis, the glass plates holding the gel were separated, such that the gel remained stuck to one of the plates. The gel together with the supporting glass plate was coated with Saran wrap to prevent the gel from drying and contact with X-ray film. Exposure of the gel to the X-ray film should not extend for more than a few hours to avoid diffusion of the PCR product in the gel. The X-ray film with the image of the ^{32}P-labeled PCR product was then used to map the position of the PCR product in the gel and the product was cut out of the gel. The volume of the gel piece containing the ^{32}P-labeled PCR product was, in average, 30-50 µl. The PCR product was extracted from the gel piece with 100 µl TE with rigorous agitation on the shaker at 37°C for several hours. The extract contained a high concentration of urea extracted with the DNA from the gel. Urea was removed by precipitation of the PCR product with 3 volumes of ethanol in the presence of 2 M ammonium acetate and 100 µg/ml glycogen at -70°C for 2 h. The PCR product was collected by centrifugation, washed in 80% ethanol and redissolved in 15 µl TE.

A 5 µl aliquot of the solution was mixed with 45 µl of PCR buffer (0.45 µM each of anchor oligonucleotide and cold (unlabeled) primer used in the ^{32}P-extension step, 0.22 mM

each of dGTP, dATP, dTTP, and dCTP, 2.78 mM $MgCl_2$, 0.055 M KCl, 0.011 M Tris-HCl (pH 9.0), 0.11% Triton X-100, 11% glycerol). The mixture was incubated at 94^{0}C for 2 min and 55^{0}C for 2 min, followed by addition of 0.4 µl of *Taq* DNA polymerase (5 U/µl)(Promega). The mixture was then incubated at 94^{0}C for 1 min. The PCR product was amplified in 35 cycles. Each cycle consisted of 10 s at 94°C, 45 s at 55°C and 1 min at 72°C. Cycling was concluded with a final extension at 72°C for 7 min. The reamplified PCR product was precipitated with ethanol, ammonium acetate and glycogen and washed. The pellet was dissolved in 10 µl TE, mixed with 5 µl of loading buffer (0.25% Bromphenol blue, 0.25% Xylene Cyanol, 20% Ficoll) and loaded onto a native 8% polyacrylamide gel (11). Electrophoresis was required to remove unused dNTPs and PCR primers, which otherwise impact the sequencing reaction. A 1 mm thick gel was most suited for the gel purification. DNA markers were required to identify the specific PCR product in the gel as nonspecific PCR products were frequently generated during reamplification. A 100 bp DNA ladder (Promega) was used as a DNA marker. The reamplified PCR product was abundant and it was readily visualized by staining with ethidium bromide. The product was cut out of the gel. DNA was extracted from the gel with 2-3 volumes of TE with rigorous agitation on the shaker at 37°C for several hours. The extracted product was treated with 2 volumes of chloroform: isoamyl ethanol to remove contaminants and precipitated with ethanol, ammonium acetate and glycogen as above. The pellet was washed with 80% ethanol and dissolved in 12 µl TE. 2 µl of the solution was used in the cycling sequencing protocol with the Thermo Sequenase kit (Amersham). The termini were mapped as the junctions between the anchor oligonucleotide and the end of the linear deleted DNA.

As it was mentioned earlier, *Taq* polymerase sometimes adds an extra nucleotide. Thus, the extra nucleotide could theoretically be incorporated in the junction between DNA end and anchor oligonucleotide in the 5' termini detection protocol. It is important to note that the described protocol is error-free, and extra nucleotides were never found at 5' termini of the DNA fragments in our studies.

Anchor oligonucleotides

The anchor oligonucleotides used in our studies were derived from bacterial DNA sequences to avoid any homology with the cellular DNA. Otherwise, the anchor oligonucleotide would anneal with the DNA and the PCR would generate nonspecific products.

The linker used to detect linear forms of the M region deleted DNA was prepared by annealing the I and JM3MIC oligonucleotides: I, 5'-GGCTCGGACCGTGGCTAGCATTAGT-3'; JM3MIC, 5'-ATAAACTAATGCTAG-3'. The first four nucleotides of the JM3MIC oligonucleotide was complementary to the 4'

extension at linear form of the M region deleted DNA. The anchor linker used in the protocol for detection of 5' termini was prepared by annealing the I and J oligonucleotides. The J oligonucleotide was 5'- ACTAATGCTAG-3'. The J oligonucleotide was four nucleotides shorter than the JM3MIC oligonucleotide and the generated linker was blunt-ended.

The anchor oligonucleotide A(C) used in the annealing-extension step of the 3' detection protocol was 5'-GATCTCATGCTGGAGTTCTTCGCCAAGT(C)$_{16}$-3', and the anchor oligonucleotide A used as a primer in the PCR amplification step was 5'-GATCTCATGCTGGAGTTCTTCGCC-3'.

RESULTS AND DISCUSSION

A protocol for the detection of rare linear forms of deleted DNA should comply with several requirements. First, the protocol should allow for the isolation of the deleted DNA in an intact form. An enormous amount of DNA nucleases are released during cell lysis. The nicks introduced by the nucleases will be subsequently mistaken for DNA breaks generated during DNA deletion reaction. Therefore, cellular enzymes should be immediately inactivated upon cell lysis. Second, the method should feature outstanding sensitivity. Finally, the detection should be highly selective. The goal is to detect rare linear DNA molecules in the presence of a vast excess of chromosome-integrated homologous copies. The protocol described here meets all of these criteria.

Cellular enzymes were instantly inactivated during the cell lysis. PCR, the most sensitive DNA detection technique, was adapted in the protocol. The linear and chromosomal DNA were separated by gel electrophoresis and the linear DNA was amplified under conditions, which allowed for amplification of the linear DNA over chromosome-integrated copies. To improve sensitivity and specificity of the detection even more, we completed the detection protocol with the ^{32}P-extension step.

The method is very sensitive. With some assumptions, calculations may serve to provide an estimate of the detection limits of the method. Since conjugating cells were synchronized, the majority of the cells were simultaneously engaged in DNA deletions. However, the deleted DNA was present in trace amounts in the cells because the DNA was immediately transported outside of the nuclei and digested. Removal of the deleted DNA occurred so quickly that an extremely sensitive method was required to detect this short-lived DNA. As already mentioned, Southern blot hybridization failed to detect free forms of the deleted DNA. The level of sensitivity of Southern blot hybridization procedures employed in our studies was 1 DNA molecule per several hundred cells (data not shown). To develop the argument, we will set the upper limit of occurrence of the deleted DNA at one DNA molecule per 1000 cells. In the TdT-mediated protocol, the deleted DNA was detected in the extract prepared from about 40,000 cells. Therefore, according to our upper limit, the signal originated from 40 or less cells. The actual number of the cells containing the linear DNA in the reaction mixture is difficult to predict. During the period of genomic deletions, a mating cell contains 4 to 8 copies of each chromosome (6). If the deleted DNA is synchronously excised from all 4 (or 8) copies of a chromosome and all the copies of the deleted DNA are retained in the nucleus at the moment of cell lysis, then just a few cells can supply multiple copies of the deleted DNA. In some of the trials, three different 3' termini at one end of the linear DNA were detected. Two of them occurred at frequency level about half of the third one (3). In other words, the DNA extract used in the detection protocol contained no more than 10 molecules of either minor form of the deleted DNA and 20 molecules of the major form. This analysis suggested that the protocol detected a signal originating from no more than 10 linear DNA molecules.

To further investigate the sensitivity of the protocol, the cell culture was analyzed several hours after programmed genomic deletions were completed. Because the cell culture could not be completely synchronized, the deletions were still under way in a minor population of the cells, which entered in the conjugation with the delay. Indeed, microscopic analysis demonstrated that less than 0.5% of the cells were still engaged in mating (data not shown). Surprisingly, the deleted DNA was still detected in the culture. The signal was very weak. The gel with the ^{32}P-labelled PCR product was exposed to X-ray film for two weeks to detect the signal. Besides, the signal was detected only after the amount of cells per the detection protocol was increased to 10,000,000 cells. Although the correct number was difficult to calculate, it suggested that a few or even one molecule of the linear DNA was detected in the population of 10,000,000 cells.

The technique can be useful to study sporadic chromosomal deletions in cell culture. Sporadic deletions in other organisms are extremely rare (5×10^{-7} in some bacteria (15) and lower) and a much bigger volume of cell culture should be applied in the protocol to find the deleted DNA fragment. On the positive side, the deleted DNA in other organisms is not a subject to programmed removal from nucleus and following degradation and it might reside in the nucleus in a stable form for hours. Besides, the cells from other organisms are not as heavily contaminated with the fragmented chromosomal DNA as mating ciliate cells. Thus, the protocol will have less potential to yield nonspecific PCR products. The additional problem in detecting the deleted DNA in *Tetrahymena* was that the organism contains 45 copies of each chromosome (2). In other words, one thousand *Tetrahymena* cells are equivalent to 45 thousand cells with a haploid genome. The high ploidy of the genome in *Tetrahymena* caused dilution of the deleted DNA in

our experiment. This problem will not complicate the detection of deleted DNA in bacteria and other organisms, most of which have a haploid or diploid genome.

Information obtained during study of linear form of the deleted DNA can help in understanding of the DNA deletion mechanism. Indeed, the linear form of the deleted DNA generated during programmed genomic deletions in *Tetrahymena* turned out to be a mixture of individual DNA fragments, which had essentially the same size but differed slightly in the position of the terminal residue at the end of one strand. The pattern of the DNA breaks indicated the new type of the deletion reaction in *Tetrahymena*, which was not demonstrated in other organisms. Using the information obtained in the study of the deleted DNA, we reconstructed the mechanism of the deletions (3).

The protocol can be readily applied to studies involving human cells. The amount of DNA per nucleus is approximately the same in *Tetrahymena* and in human cells: 10 and 7 pg, respectively (16,17). The indicated amount of DNA in *Tetrahymena* was determined for vegetative cells. In our experiments, the actual amount of DNA in *Tetrahymena* cells can be somewhat lower because of degradation of old macronuclei in mating cells (2). Even after adjustment, however, the amount of DNA per nuclei in *Tetrahymena* mating cells and human cells is still comparable. Therefore, quantification parameters like the amount of cells and DNA do not need to be adjusted.

Certain considerations should be taken into account if the technique were used for detection of sporadic genomic deletions. First of all, the technique can be applied to study deletions with known breakage sites only. Boundaries of the deletion should be mapped with precision of several hundreds of base pairs before the protocol is applied. In the next step, the internal region of the deleted DNA will be used to design the PCR primer. The technique will detect DNA termini located up to 1 kbp downstream of the PCR primer. Detection of the more remote DNA breaks was unreliable in our studies (data not shown). The protocol will detect the deleted DNA even if the deletion breakpoints show heterogeneity. Anchor PCR will amplify all the variable breakpoints in one reaction mixture. Fig. 2 demonstrates amplification of several breakpoints scattered by the deletion boundary in the range of 15 bps in one PCR reaction. Identical results were obtained in the study of breakpoints scattered in the range of 600 bp (data not shown).

In our studies, the detection was based on ligation of linear DNA to a compatible linker. We used this straightforward approach because we had already established the structure of one end of the deleted DNA under investigation. However, if the structure of the ends of a linear DNA is unknown, alternative anchor PCR protocols based on attachment of an anchor oligonucleotide at either 5' or 3' termini on linear DNA

can be applied. The advantage of these alternative protocols is that the DNA is labeled with anchor oligonucleotide irrespective of the end structure. As already indicated, the TdT-mediated PCR developed for amplification of 3' termini provides the most sensitive and reliable detection.

Only the phosphorylated 5' termini and hydroxylated 3' termini can be detected in anchor PCR protocols. Fortunately, the vast majority of DNA endonucleases generate 5' phosphorylated and 3' hydroxylated termini at DNA ends. However, some of them generate 5' hydroxylated and 3' phosphorylated termini. If the latter nucleases are involved in a DNA deletion reaction the deleted DNA fragment will not be detected. The DNA termini should be modified before the detection protocols are applied. The 5' hydroxylated termini at the ends of the DNA can be phosphorylated with kinase. Similarly, 3' phosphorylated termini can be dephosphorylated with dephosphorylase. Both reactions are technically simple and reliable (11).

The technique can also be applied in gene therapy trials to study modifications at ends of transfected DNA, which the DNA undergoes before integration into the genome.

ACKNOWLEDGMENTS

REFERENCES

1. Paques F, Haber JE. Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae. *Microbiology & Molecular Biology Reviews*. 1999; 63(2):349-404.
2. Yao M-C. Site-specific chromosome breakage and DNA deletions in ciliates In *Mobile DNA* (ed. D.E.Berg and M.M.Howe), pp. 715-734. American Society for Microbiology, Washington, D.C., 1989.
3. Saveliev SV, Cox MM. Product analysis illuminates the final steps of IES deletion in *Tetrahymena thermophila*. *EMBO J*. 2001; 20(12):3251-3261.
4. Saveliev SV, Cox MM. The fate of deleted DNA produced during programmed genomic deletion events in *Tetrahymena thermophila*. *Nucleic Acids Res*. 1994; 22:5695-5701.
5. Coyne RS, Nikiforov MA, Smothers JF, Allis CD, Yao M-C. Parental Expression of the chromodomain protein Pdd1 is required for completion of programmed DNA elimination and nuclear differentiation. *Molecular Cell*. 1999; 4:865-872.
6. Austerberry CF, Allis CD, Yao M-C. Specific DNA rearrangements in synchronously developing nuclei of *Tetrahymena*. *Proc. Natl. Acad. Sci. USA* 1984; 81:7383-7387.

7.   Bruns PJ, Brussard TB. Pair formation in *Tetrahymena pyriformis*, an inducible developmental system. *Journal of Experimental Zoology*. 1974; 188(3):337-344.

8.   Austerberry CF, Yao M-C. Nucleotide sequence structure and consistency of a developmentally regulated DNA deletion in *Tetrahymena thermophila*. *Mol. Cell Biol.* 1987; 7:435-443.

9.   Saveliev SV, Cox MM. Developmentally programmed DNA deletion in *Tetrahymena thermophila* by a transposition-like reaction pathway. *Embo J.* 1996; 15(11):2858-2869.

10. Mueller PR, Wold B. *In vivo* footprinting of a muscle specific enhancer by ligation mediated PCR. *Science.* 1989; 246:780-786.

11. Sambrook J, Russell DW. Molecular Cloning: A laboratory manual. 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York 2001.

12. Clark JM. Novel non-template nucleotide addition reactions catalyzed by prokaryotic and eucaryotic DNA polymerases. *Nucleic Acids Res.* 1988; 16:9677-9686.

13. Dumas JB, Edwards M, Delort J, Mallet J. Oligonucleotide ligation to single-stranded cDNAs: a new tool for cloning 5' ends of mRNAs and for constructing cDNA libraries by *in vitro* amplification. *Nucleic Acids Res.* 1991; 19(19):5227-5232.

14. Loh EY, Elliott JF, Cwirla S, Lanier LL, Davis MM. Polymerase chain reaction with single-sided specificity: analysis of T cell receptor delta chain. *Science.* 1989; 243:217-220.

15. Bach S, Buchrieser C, Prentice M, Guiyoule A, Msadek T, Carniel E. The high-pathogenicity island of **Yersinia** enterocolitica Ye8081 undergoes low-frequency deletion but not precise excision, suggesting recent stabilization in the genome. [Journal Article] *Infection & Immunity.* 1999; 67(10):5091-5099.

16. Seyfert HM, Cleffmann G. Mean macronuclear DNA contents are variable in the ciliate *Tetrahymena*. *Journal of Cell Science*. 1982; 58:211-223.

17. Mikel UV. Absolute DNA values from Feulgen microspectrophotometric measurements and quantitative electron microscopy: A comparison of two species. *Analytical & Quantitative Cytology & Histology*. 1987; 9(1):13-16.

PROTOCOLS

1. Lyse cells in hot SDS/EDTA

2. Dialyze the lysate to remove SDS and decrease concentration of EDTA

3. Digest RNA with RNAse A

4. Separate the linear DNA and chromosomal DNA in an agarose gel

5. Detect gel fraction with the linear DNA using anchor PCR

6. Sequence the PCR product to map ends of the linear DNA

7. Reconstruct mechanism of the deletion based on structure of the breaks