Ligation-mediated PCR of restriction fragments from large DNA molecules.

D R Smith

Genome Res. 1992 2: 21-27
Access the most recent version at doi:10.1101/gr.2.1.21

Ligation-mediated PCR of Restriction Fragments from Large DNA Molecules

Douglas R. Smith

Department of Human Genetics and Molecular Biology, Collaborative Research, Inc., Waltham, Massachusetts 02154

A general method is described for PCR amplification of single restriction fragments from large DNA molecules. The method involves sequence-specific ligation of synthetic oligonucleotides to ambiguous 4-base 5′ overhangs produced by type IIS restriction endonucleases. Such "adapter-tags" provide one target for primer annealing in subsequent PCR reactions. The second target for primer annealing is provided by a universal "bubble-tag" ligated to blunt ends produced with another endonuclease. The key advantage of this approach is that specific fragments can be isolated without any prior knowledge of the nucleotide sequence of the target. Using bacteriophage λ DNA as a test system, unique PCR products could be generated consistently. Conditions of temperature, ionic strength, and substrate concentration in the adapter-tag ligations—which affect sequence specificity—were found to have a major influence on the purity of PCR-generated fragments. In principle, the method permits the amplification of virtually any sequence from purified cosmid or YAC DNA using a library of only 240 adapter-tags.

Several ligation-mediated PCR techniques have been described to amplify specific fragments from various sources. Typically, the methods involve three major steps: (1) DNA fragmentation (by treatment with chemical reagents, restriction endonucleases, or other enzymes); (2) ligation of an oligonucleotide adapter; and (3) PCR using a primer that matches the adapter plus another, specific primer (or nested set of primers) that matches an adjacent genomic sequence.

Examples of this approach include: anchor PCR to amplify genomic DNA fragments,[1] vectorette PCR to amplify the ends of yeast artificial chromosome (YAC) inserts,[2] a method to isolate genomic sequences adjacent to *Not*I sites,[3] walking methods to isolate genomic sequences adjacent to cloned cDNAs,[4,5] a method to amplify fragments from microdissected chromosomes,[6] and methods to amplify genomic footprinting or DNA sequencing ladders.[7–9] In addition, several groups have used ligation-mediated PCR to amplify low-abundance cDNAs prior to cloning; this includes specific cDNAs,[10,11] random cDNAs,[12–15] and subtractive cDNAs.[16] Finally, a somewhat different ligation-dependent procedure has been used to amplify coincident sequences present in two different source mixtures.[17] Most of these techniques are designed to amplify specific sequences for which DNA sequence information is available from only one end of a desired fragment. Procedures to amplify completely unknown fragments have, as yet, only been suitable for coamplifying all molecules in a mixture.

The approach presented in this report provides a means to amplify individual, unknown fragments from any purified DNA molecule ranging from about 50 to 250 kb in size. The incorporation of multiplex tag sequences allows potential mixing and multiplex sequencing[18] of amplified products. The method takes advantage of type IIS restriction endonucleases[19] that cleave outside their recognition sequences to produce ambiguous 4-base 5′ overhangs to which specific adapters can be ligated. Currently, there are eight commercially available type IIS enzymes that produce ambiguous 4-base 5′ overhangs (*Bbs*I, *Bbv*I, *Bsa*I, *Bsm*AI, *Bsp*MI, *Esp*3I, *Fok*I, and *Sfa*NI). One of these, *Fok*I, has been used[20] for sampled sequence fingerprinting of cosmids and genomic DNAs up to 4.7 megabases (Mb) in size.

The principle of the method is illustrated in Figure 1. The target DNA is digested with a blunt-cutting enzyme, and a bubble-tag is ligated to the resulting ends. The DNA is then digested with a type IIS enzyme that leaves ambiguous 4-base 5′ overhangs, and a specific adapter-tag is ligated to the sample. If the appropriate 5′ overhang is present, the adapter-tag will be ligated. Finally, the DNA is subjected to PCR amplification using bubble and adapter-tag primers to amplify any appropriately ligated fragments. The advantage of the bubble primer[2] is that it cannot prime DNA synthesis until a complementary strand has been generated from the adapter-tag primer at the other end of the fragment. This effectively eliminates mispriming and circumvents the need for removal of excess bubble-tag.

MATERIALS AND METHODS

Oligonucleotides

A set of seven oligonucleotide adapter-tags and corresponding tag-complements (that double as PCR primers) were

2:21–27©1992 by Cold Spring Harbor Laboratory Press ISSN 1054-9803/92 $3.00

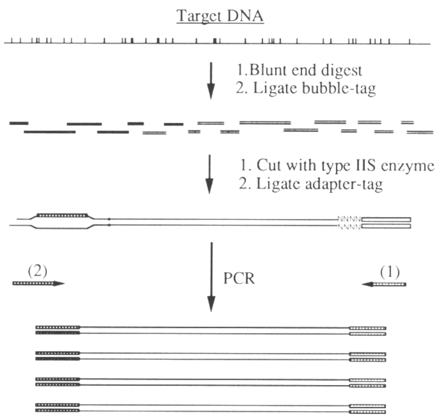

FIGURE 1 Principle of the method. The target DNA is digested with a frequent blunt cutter, and a bubble-tag is ligated to the resulting ends. The DNA is then digested with a type IIS endonuclease that leaves 4-base 5′ overhangs, and one of 240 specific adapter-tags is ligated under stringent conditions. The reaction is stopped, and specific fragments with attached bubble and adapter tags are amplified by PCR.

TABLE 1 Adapter-tags with Corresponding Tag-Complements[a]

Name	Sequence
GAAA-1E	5'pGAAACCCCCAATAAAATCATACTA GGGGGTTATTTTAGTATGAT 5' (tag1E complement)
TAAA-3E	5'pTAAACTAACAACAAACCTTACTAC GATTGTTGTTTGGAATGATG 5' (tag3P complement)
CAAA-4E	5'pCAAACAACACCCATCCACTAAACT GTTGTGGGTAGGTGATTTGA 5' (tag4E complement)
GTAA-10P	5'pGTAACCTAATCATCAATATACTCA GGATTAGTAGTTATATGAGT 5' (tag10P complement)
CGTT-11E	5'pCGTTCCACCTTTATACTTCTTACA GGTGGAAATATGAAGAATGT 5' (tag11E complement)
CCCC-13E	5'pCCCCCAAAAACTAATTCCCAAAAA GTTTTTGATTAAGGGTTTTT 5' (tag13E complement)
TGAA-13P	5'pTGAACAACTTACTCTACACCCCTT GTTGAATGAGATGTGGGGAA 5' (tag13P complement)
CCAA-21P	5'pCCAACTCCACACCACCATCTTTTT GAGGTGTGGTGGTAGAAAAA 5' (tag21P complement)

[a]The names given are used throughout the text to refer to double-stranded adapter-tags, as illustrated. The tag-complements double as PCR primers.

used in these experiments (see Table 1). All synthetic oligonucleotides were purchased from Operon Technologies (Alameda, CA). The tag sequences correspond to multiplex tags present in the PLEX vectors developed by Church and Kieffer Higgins[18] or by Milligen. They are designated by a number and a letter (E or P) indicating the PLEX vector they are derived from and location relative to the cloning site (*Eco*RI or *Pst*I side).

Additional synthetic oligonucleotides included a duplex bubble-tag with an internal noncomplementary region, and a corresponding PCR primer that can only prime DNA synthesis after a first strand complementary to the G-rich strand of the bubble-tag has been generated. The sequences of the bubble-tag oligonucleotides and primer are shown below:

5′pTCCCTTCTTCTCCCAAAAAAAA
AAAAAAAACCTCCTTCCTCTCTTC

5′GAGAGGAAGGAGGTGTTGGTAGT
TGTTTTGGGAGAAGAAGGGA

5′GTGTTGGTAGTTGTTTTGG
(primer)

The specific sequence of the bubble primer was chosen to be free of secondary structure and to be compatible with the tag complements in PCR reactions; the selection was aided by OLIGO computer software.[21] All oligonucleotides were purified by electrophoresis in denaturing polyacrylamide gels[22] before use. One strand of the bubble and adapter-tags (as indicated in Table 1) was phosphorylated using polynucleotide kinase[22] before gel purification. The bubble-tag oligos and adapter-tags were annealed at 100 μM in 0.1 M NaCl for 30 min at 37°C.

Ligation Reactions

To prepare bubble-tagged samples, bacteriophage λ DNA was digested with an excess of *Rsa*I or *Hae*III, phenol-extracted, and precipitated. Five micrograms of the resulting DNA was ligated in a 100-μl reaction with 370 pmoles of annealed bubble-tag (10-fold molar excess over *Rsa*I ends). The reaction included: 4000 units of T4 ligase (NEB; large excess to ensure complete blunt-end ligation), 50 mM Tris-acetate (pH 7.6), 10 mM MgOAc, 20 mM dithiothreitol, and 100 μM ATP (standard ligation buffer). Ligation products were purified by phenol extraction and ethanol precipitation and then resuspended in water. A sample of the ligated DNA was electrophoresed in a 2% agarose gel next to an unligated digest to visualize the reaction products. A similar pattern of bands was observed, except the ligated fragments were shifted in accordance with the length of the bubble-tag (46 bp at each end = 92 bp).

The blunt-end/bubble-tagged DNA was digested with a 10-fold excess of type IIS enzymes *Bsp*MI and *Sfa*NI (NEB) in the buffer recommended by the supplier. After phenol extraction and ethanol precipitation, the DNA was dissolved in water at 10 ng/μl (0.33 fmole/μl of a unique tetranucleotide end). Adapter-tag ligations were carried out with 50 ng of this prepared DNA under various conditions as described in the next section. Optimal conditions for adapter-tag ligations were: 50 ng DNA (1.65 fmoles of a unique end), standard ligation buffer including 150 mM NaOAc, 0.165 fmoles preannealed adapter-tag, and 240 units T4 DNA ligase for 16 hr at 37°C.

Labeled ligation products were produced using 5′-^{32}P-labeled adapter-tags (phosphorylated with polynucleotide kinase and [c-^{32}P]ATP, 6000 Ci/mmole). For these experiments, the λ DNA target was digested with an excess of *Sfa*NI and treated with calf intestinal phosphatase (Boehringer Mannheim) before ligation. Ligation experiments using labeled adapter-tags were performed in standard buffer with 20 ng of λ DNA. After ligation, samples were precipitated, dissolved in formamide loading dye, and heated at 90–100°C for 2 min before loading onto a denaturing acrylamide gel.

PCR

PCR reactions were carried out in a Perkin-Elmer Gene Amp 9600 instrument. One-fifth (4 μl) of an adapter-tag ligation was added to a 40-μl PCR reaction (standard conditions: 50 mM KCl, 10 mM Tris-Cl, pH 8.5, 1.5 mM $MgCl_2$, 200 μM dNTPs, 200 nM primers). The reactions were set up on ice, then subjected to the following temperature regime: 95°C for 2 min; (94°C for 15 sec; 50°C for 1 min; 72°C for 30 sec) repeated 30 times; 72°C for 3 min.

Gel Electrophoresis

All gels were run in 50 mM Tris-borate, 1.25 mM EDTA. PCR products were analyzed in 2% agarose gels (SeaPlaque, or equal mixture of GTG/NuSieve) run at 5–7 volts/cm. Loading dye (1×) for visualization of ligation and PCR products contained 14% glycerol, 0.01% bromophenol blue, 10 mM sodium–EDTA (pH 8), and 5% phenol. Acrylamide gels for the analysis of ^{32}P-labeled ligation products contained 5% acrylamide, 0.25% BIS, and 50% urea, and were run for about 3 hr at 30 volts/cm. Gels for the purification of oligonucleotides were similar, but were 1.5 mm thick (instead of 0.4 mm) and contained 10% acrylamide, 0.5% BIS. Loading dye (1–2×) for acrylamide gels contained 98% formamide, 0.005% bromophenol blue, and 0.005% xylene cyanol. The oligonucleotides were visualized by UV shadowing.[22]

RESULTS

Bacteriophage λ DNA was chosen as a model template for this study because of its large size and known DNA sequence. Initially, the nucleotide sequence was examined to locate the cleavage sites for seven type IIS and two blunt cutting enzymes. The adapter-tags described above were selected (from an arbitrary set of 18) on the basis of this analysis. The adapter-tags were chosen to produce unique ligation products between 100 and 800 bp from fragments generated with two type IIS and two blunt cutting enzymes.

Ligation Efficiency and Specificity

The ability to generate unique amplified fragments by the approach outlined above requires high sequence specificity in the adapter-tag ligations. T4 DNA ligase is known to ligate mismatched ends,[23,24] and the extent of mismatch ligation has been shown to depend on the sequence and concentration of mismatched termini.[25] Thus, in the presence of equimolar matching ends, less mismatched product is formed than in the absence of a complementary substrate.[25] The effects of temperature, ionic strength, and ligase concentration on ligation specificity have also been investigated.[26,27] These studies determined that high ionic strength markedly improves specificity, while ligase concentration has little effect.

In view of these studies, several experiments were conducted to determine the effects of ionic strength and tag concentration on the specificity of adapter-tag ligations. Initially, these experiments were done using ^{32}P-labeled adapter-tags to permit direct visualization of ligation products. The tags were ligated to *Sfa*NI-cut λ DNA that had been dephosphorylated to prevent self-ligation of *Sfa*NI fragments. The results of one such experiment are shown in Figure 2. Adapter-tag CCAA-21P was ligated to *Sfa*NI-cut λ DNA at a 24-fold excess. The reaction products were denatured and electrophoresed in a denaturing acrylamide gel. Ligation reactions were conducted at four temperatures (16°C, 22°C, 30°, and 37°C) and four concentrations of sodium acetate (0, 100 mM, 150 mM, and 200 mM). The expected 163-nucleotide ligation product was observed in all cases. The most efficient ligation occurred at 22°C in the absence of salt, although nonspecific side products were readily apparent. Higher salt concentrations inhibited ligation but diminished the background products significantly (Fig. 2). Sodium acetate was used because in a number of related experiments it was found to be less inhibitory than sodium chloride (not shown).

Coupled Ligation and PCR

The ligation assay described above suggested that ionic strength might be useful for minimizing misligation of adapter-tags. However, the assay was not sensitive enough to detect all of the background products. Therefore, coupled ligation and PCR reactions were performed to define conditions under which unique fragments could be obtained.

An extensive series of optimization experiments were carried out in which ionic strength and adapter-tag concentrations were independently varied in the ligation reactions. In addition, the effect of changing the annealing temperature in the PCR reactions was investigated. The conclusions reached from these experiments are as follows. (1) The addition of salt to the adapter-tag ligations improved purity and yield of PCR products, although high concentrations were inhibitory. (2) The annealing temperature in the PCR reactions (in the range of 45–55°C) did not have a major effect on product yield or purity. (3) The use of submolar adapter-tag/target ratios generally improved product purity. Optimal ligation conditions as determined by this study are given in Materials and Methods.

Purified λ DNA was cut with *Rsa*I or *Hae*III, ligated to a 10-fold excess of preannealed bubble-tag, and digested with type IIS enzymes. The samples were then ligated to pre-annealed adapter-tags, and subjected to PCR using the tag complements and bubble primer. Initial experiments confirmed that PCR products of the expected size were consistently produced with all adapter-tags tested (a total of six; results not shown). Experiments performed with varying molar ratios (0.1- to 100-fold) of adapter-tags with respect to λ-DNA also produced the expected products with the cleanest results occurring at the lower adapter-tag concentrations (results not shown). Therefore, another experiment was performed in which the adapter-tag concentrations were dropped even further, from 0.1× to 0.001× relative to the λ DNA. The results

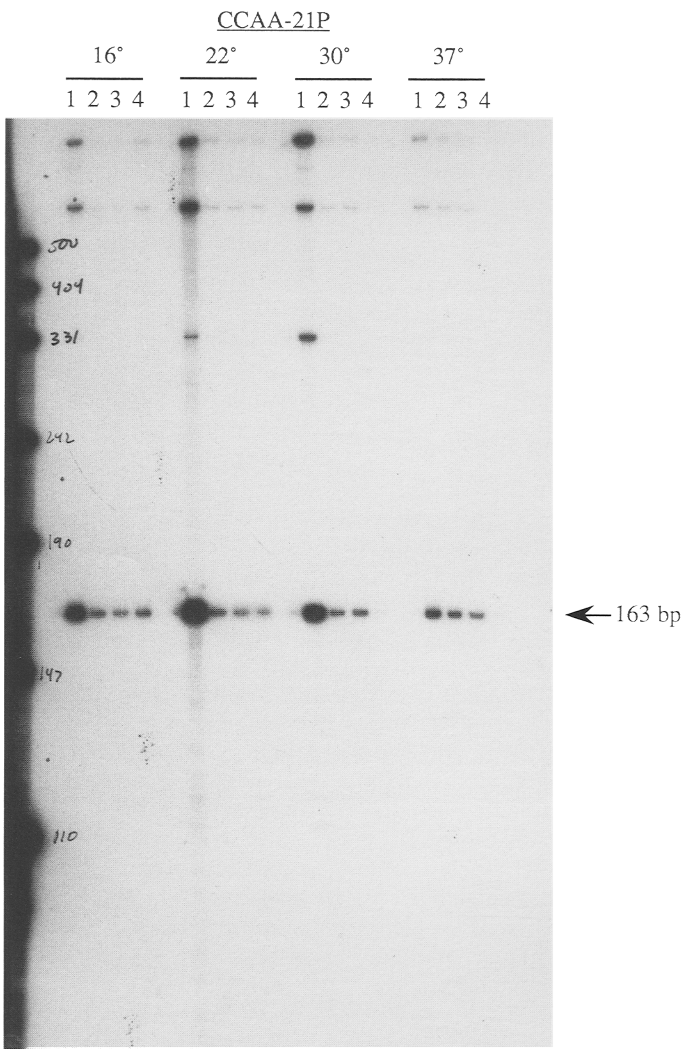

FIGURE 2 Adapter-tag ligation. The adapter-tag CCAA-21P was phosphorylated with ^{32}P and ligated to *Sfa*NI-cut λ DNA (phosphatase-treated) at the indicated temperatures and at one of four salt concentrations: (lanes *1*) no salt; (lanes *2*) 100 mM NaOAc; (lanes *3*) 150 mM NaOAc; (lanes *4*) 200 mM NaOAc (other conditions are described in Materials and Methods). The expected product size was 163 bp; marked by arrow. The pair of fragments visible in all lanes at the top of the gel are presumed to correspond to double-stranded material. The size markers are end-labeled pUC19 *Msp*I fragments.

of this experiment are shown in Figure 3, and are described in detail below.

In general, the reaction products produced using low adapter-tag concentrations had very low background with the best results occurring at the lowest tag concentration (Fig. 3, sets 1 and 4: corresponding to adapter-tags CAAA-4E and CCCC-13E, respectively). A third tag (GTAA-10E) gave low PCR yield at all concentrations. Surprisingly, tag CAAA-4E, which worked well in set 1 using λ DNA that had been cleaved with *Hae*III prior to bubble ligation, gave low yields and extraneous products with *Rsa*I/bubble λ DNA (Fig. 3, set 2). In this case, higher adapter-tag concentrations produced better results (not shown).

Rescue of Known Fragments

A modified version of the technique, using two type IIS ends, was tested for PCR rescue of a specific *Sfa*NI λ DNA fragment with known ends. The fragment was selected arbitrarily from 168 possible choices, and was 1130 bp long with 5′TTTC and 5′AACG overhangs. *Sfa*NI-cut λ DNA was ligated with adapter-tags GAAA-1E and CGTT-11E in standard buffer at 37°C with 150 mM NaCl and 1 mM spermidine. The ligation reaction was stopped by heating at 68°C for 20 min, and an amount containing 5 ng of DNA was added to a PCR reaction. The PCR was carried out in a Techne thermocycler for 30 cycles: 94°C for 1 min; 60°C for 1 min; 72°C for 1.5 min followed by 3 min at 72°C. The results are shown in Figure 4; the expected 1170-bp product is clearly evident.

DISCUSSION

This study demonstrates the applicability of ligation-mediated PCR to amplifying individual type IIS restriction fragments from large DNA molecules. The key advantage of this approach is that amplified fragments can be generated without any knowledge of the nucleotide sequence of the target molecule. The fragments thus isolated will derive from random locations within the target. However, if the nucleotide sequence is known, specific fragments can also be rescued using two specific adapter-tags instead of newly synthesized primers.

The successful application of the method requires high sequence specificity in the adapter-tag ligations because, with the use of a universal bubble-tag, the specificity of PCR amplification is completely determined by the adapter-tag. Factors that are known to improve ligation specificity, such as high ionic strength and limiting substrate concentrations, are effective in decreasing background in the PCR reactions. However, it is important to emphasize that the best PCR results were obtained in conjunction with *very poor* ligation conditions, so specificity appears to be much more critical than ligation efficiency for successful amplification (compare the ligation conditions used in Figs. 2 and 3).

Some adapter-tags gave better results than others (for instance, tag GTAA-10P always produced a weak signal, whereas tags GAAA-1E and CCCC-13E worked well; Fig. 3 and unpublished data). However, these experiments did not address whether this resulted from differences in ligation efficiency or some other factor. It is interesting to note that tag CAAA-4E worked very efficiently with *Hae*III/bubble *Bsp*MI-cut target DNA and very poorly with *Rsa*I/bubble *Bsp*MI-cut target (Fig. 3). Several factors could account for this difference, including inadvertant sample contamination, or inefficient *Rsa*I cutting, bubble ligation, or *Bsp*MI digestion of the latter target preparation.

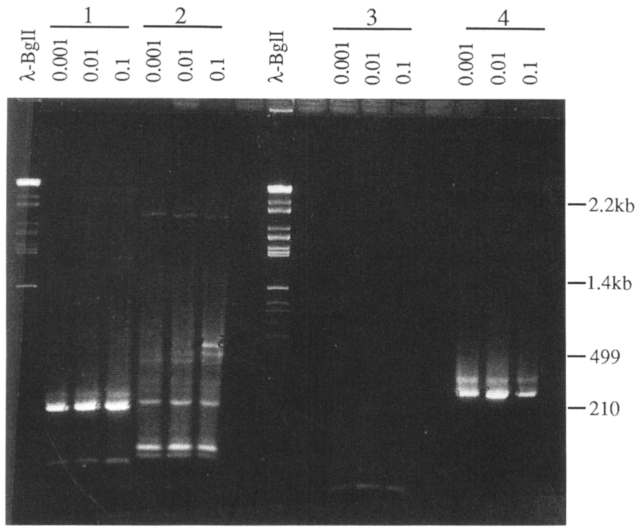

FIGURE 3 Adapter-tag PCR; ligation with submolar tag ratios. Tag ligations were performed at the indicated molar ratios of adapter-tag to target end. The samples are as follows: (set number, bubble-tag end, type IIS end, adapter-tag, expected product size); 1. *Hae*III, *Bsp*MI, CAAA-4E, 208 bp; 2. *Rsa*I, *Bsp*MI, CAAA-4E, 465 bp; 3. *Hae*III, *Sfa*NI, GTAA-10P, 225 bp; 4. *Rsa*I, *Sfa*NI, CCCC-13E, 286 bp. λ *Bgl*I fragments were used as size markers.

In some experiments, there was a significant level of generalized background. This may have been caused by the weak homology between the 3′ end of the bubble primer and the top strand of the bubble-tag. It is possible that this could be reduced by using a different bubble primer that hybridizes to a slightly different region of the bubble-tag.

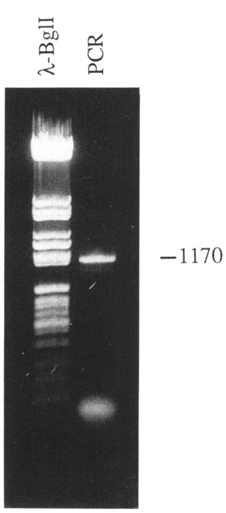

FIGURE 4 Adapter-tag fragment rescue. Two adapter-tags, GAAA-1E and CGTT-11E, were ligated with *Sfa*NI-cut λ DNA, and the resulting product amplified as described in Results. (Lane *1*) λ *Bgl*I size standards; (lane *2*) PCR products (the major product at 1170 bp corresponds to the expected size).

Some variability in product yield with different adapter-tags could arise from differences in priming efficiency during PCR amplification. This could be eliminated by placing a universal primer in the adapter-tags. However, variability in ligation efficiency of different 4-base sequences is hard to circumvent. Such variability is apparent because several reactions with different adapter-tags, ligated and amplified in parallel with identical target DNA, produced very different results (unpublished results). If significant differences in ligation efficiency do exist, it might be necessary to optimize ligation conditions separately for each adapter-tag. It is even possible that differences in the tag sequences juxtaposed to each 4-base adapter end could affect the ligation efficiency. Clearly, additional studies will be required before the method can be successfully applied on a large scale.

Frequency of Single "Hits" in Large Molecules

Eight type IIS endonucleases that generate ambiguous 4-base 5′ overhangs are currently available (NEB catalog). These include enzymes with 5 bp and 6 bp recognition sequences; all are asymmetric. Calculations of cutting frequencies based on nearest neighbor frequencies in human DNA[28] can be readily made. These values (Fig. 5) agree well with the predicted cutting frequencies in the 73-kb sequence of the β-globin region in GenBank. Based on the calculated values, the average occurrence of specific 5′ termini can be estimated statistically using the Poisson distribution. The results of such calculations for two representative enzymes, *Bsp*MI and *Fok*I, are shown in Figure 5. The tables can be interpreted as follows. For *Bsp*MI, a representative 200-kb DNA molecule will contain about 128 cleavage sites, or 256 tetranucleotide termini. In this case, ligations with each of the 256 possible 4-base adapter-tags will be nonproductive in 36% of cases, will yield single fragments in 37% of cases, and will yield two or more fragments in 27% of cases.

Possible Applications

This ligase-mediated PCR technique was originally conceived as a way to generate representative sequencing templates from large molecules, such as amplified YACs,[29] without subcloning. Using the optimized protocols described here, it should indeed be possible to generate large numbers of unique amplified fragments. The use of many different adapter-tags permits sample mixing for multiplex sequencing protocols to allow efficient chemical sequencing of these products.[18] The availability of several different type IIS enzymes together with a complete set of adapter-tags (including all 240 possible nonpalindromic 4-base 5′ overhangs) should permit virtually any DNA sequence to be amplified from purified cosmid or YAC DNA. Currently, however, the technique has only been applied using λ DNA.

The theoretical feasibility of using the method for genomic sequencing was tested by performing a mock sequencing project on the 73-kb sequence of the β-globin locus in GenBank. The position and size of gaps that would have been obtained using a set of three type IIS enzymes (*Sfa*NI, *Bsp*MI, *Bbs*I) and four blunt-cutters (*Hae*III, *Rsa*I, *Ssp*I, *Dra*I) were catalogued. Imaginary sequence runs of 350 nucleotides were done from each end of each possible double-digest restriction fragment with one type IIS end and one blunt end. Most parts of the sequence were covered in both directions; 80 regions were covered in only one direction, and 20 gaps averaging 225

	BspMI		s = 1560	
Target (kb)	n	ƒ(0)	ƒ(1)	ƒ(>1)
50	64	0.78	0.19	0.03
100	128	0.61	0.30	0.09
150	192	0.47	0.35	0.17
200	256	0.36	0.37	0.27
250	320	0.28	0.35	0.36

	FokI		s = 527	
Target (kb)	n	ƒ(0)	ƒ(1)	ƒ(>1)
50	189	0.48	0.35	0.17
100	380	0.23	0.34	0.44
150	570	0.11	0.24	0.65
200	760	0.05	0.15	0.80
250	950	0.02	0.09	0.89

ℓ = target size (bp) $\qquad f(x) = \frac{m^x e^{-m}}{x!}$

s = Average fragment size (bp)

n = Average number of 4 base termini = $\frac{2\ell}{s}$

m = Average frequency of a particular 4 base end = $\frac{n}{256}$

FIGURE 5 Statistics of type IIS cleavage. Calculations were made on the basis of the Poisson distribution using the parameters shown. The tables predict the probability of finding 0, 1, or more than 1 copy of a typical 4-base end in target molecules of various sizes, as described in the text.

bp (accounting for 6% of the sequence) were evident. Most of these gaps could have been easily bridged in one step by primer walking. However, the task of generating the templates would have required 80 bubble ligations and 2880 adapter-tag ligations with 100% success in the PCR reactions. The limited experimental data presented in this paper suggests that an actual success rate would be lower. Nevertheless, the approach might still be useful.

The use of two different adapter-tags to rescue fragments with known ends could be useful in large-scale shotgun sequencing projects for second-strand sequencing or gap filling in the final stages of a project (after four- to fivefold coverage has been achieved). The advantage of this approach is that a *total* of only 270 oligonucleotides would need to be synthesized to amplify and sequence any desired region. Individual type IIS fragments with one specific pair of ends would occur, on average, at a frequency similar to that of a 17-mer, about once every 10^{10} nucleotides. In practice, it would be convenient to divide up the set of 240 adapter-tags into a number of sets with shared tag sequences. This simplifies the task of designing compatible tag/primer sequences. With 30 tags (each attached to 8 adapters), there is only about a 3% chance that two randomly selected oligonucleotides would share the same tag sequence. If this happened, another fragment with different ends produced with another type IIS enzyme could be amplified.

Another possible application of the technique could be in fingerprinting purified cosmids, YACs, or even bacterial genomes (perhaps using mixtures of tags). Such fingerprints could have very high information content. Finally, it is possible that pools of tagged PCR products from the ends of heterologous DNA segments cloned in YACs or cosmids could be employed for multiplex chromosome walking in clone libraries arrayed in high density grids.

ACKNOWLEDGMENTS

I am grateful to John Tagliamonte for technical assistance and George Church for helpful comments. This work was funded by SBIR contract N43-DK-0-2259 from the NIDDK.

REFERENCES

1. Roux, K.H. and P. Dhanarajan. 1990. A strategy for single site PCR amplification of ds DNA. Priming digested clones or genomic DNA from an anchor-modified restriction site and a short internal sequence. *BioTechniques* 48–57.
2. Riley, J., R. Butler, D. Ogilvie, R. Finniear, D. Jenner, S. Powell, R. Anand, J.C. Smith, and A.F. Markham. 1990. A novel, rapid method for the isolation of terminal sequences form yeast artificial chromsomes. *Nucleic Acids Res.* **18:** 2887–2890.
3. Patel, K., R. Cox, J. Shipley, F. Kiely, K. Fraser, D.R. Cox, H. Lehrach, and D. Sheer. 1991. A novel and rapid method for isolating sequences adjacent to rare cutting sites and their use in physical mapping. *Nucleic Acids Res.* **19:** 4371–4375.
4. Fors, L., R.A. Saavedra, and L. Hood. 1990. Cloning of the shark Po promoter using a genomic walking technique based on the polymerase chain reaction. *Nucleic Acids Res.* **18:** 2793–2799.
5. Shyamala, V. and G.F.-L. Ames. 1989. Genome walking by single-specific-primer polymerase chain reaction. *Gene* **84:** 1–8.
6. Saunders, R.D.C., D.M. Glover, M. Ashburner, I. Siden-Kiamos, C. Louis, M. Monastirioti, C. Savakis, and F. Kafatos. 1989. PCR amplification of DNA microdissected from a single polytene chromosomes band: A comparison with conventional microcloning. *Nucleic Acids Res.* **17:** 9027–9037.
7. Pfeifer, G.P., S.D. Steigerwald, P.R. Mueller, B. Wold, and A.D. Riggs. 1989. Genomic sequencing and methylation analysis by ligation mediated PCR. *Science* **246:** 810–813.
8. Mueller, P.R. and B. Wold. 1989. In vivo footprinting of a muscle specific enhancer by ligation mediated PCR. *Science* **246:** 780–786.
9. Steigerwald, S.D., G.P. Pfeifer, and A.D. Riggs. 1990. Ligation-mediated PCR improves the sensitivity of methylation analysis by restriction enzymes and detection of specific DNA strand breaks. *Nucleic Acids Res.* **18:** 1435–1439.
10. Huang, C. and B.D. Stollar. 1991. Construction of representative immunoglobulin variable region cDNA libraries from human peripheral blood lymphocytes without in vitro stimulation. *J. Immunol. Methods* **141:** 227–236.
11. Chaudhary, V.K., J.K. Batra, M.G. Gallo, M.C. Willingham, D.J. Fitzgerald, and I. Pastan. 1990. A rapid method of cloning functional variable-region antibody genes in Escherichia coli as single-chain immunotoxins. *Proc. Natl. Acad. Sci.* **87:** 1066–1070.
12. Collasius, M., H. Puchta, S. Schlenker, and G. Valet. 1991. Analysis of unknown

DNA sequences by polymerase chain reaction (PCR) using a single specific primer and a standardized adaptor. *J. Virol. Methods* **32:** 115–119.

13. Ko, M.S.H. 1990. An "equalized cDNA library" by the reassociation of short double-stranded cDNAs. *Nucleic Acids Res.* **18:** 5705-5711.
14. Ko, M.S.H., S.B.H. Ko, N. Takahashi, K. Nishiguchi, and K. Abe. 1990. Unbiased amplification of a highly complex mixture of DNA fragments by "lone-linker"-tagged PCR. *Nucleic Acids Res.* **18:** 4293–4294.
15. Akowitz, A. and L. Manuelidis. 1989. A novel cDNA/PCR strategy for efficient cloning of small amounts of undefined RNA. *Gene* **81:** 295–306.
16. Timblin, C., J. Battey, and W.M. Kuehl. 1990. Application for PCR technology to subtractive cDNA cloning: Identification of genes expressed specifically in murine plasmacytoma cells. *Nucleic Acids Res.* **18:** 1587–1593.
17. Brookes, A.J. and D.J. Porteous. 1991. Coincident sequence cloning. *Nucleic Acids Res.* **19:** 2609–2613.
18. Church, G.M. and S. Kieffer-Higgins. 1988. Multiplex DNA sequencing. *Science* **240:** 185–188.
19. Szybalski, W. 1985. Universal restriction endonucleases: Designing novel cleavage specificities by combining adapter olidodeoxynucleotide and enzyme moieties. *Gene* **40:** 169–173.
20. Brenner, S. and K.J. Livak. 1989. DNA fingerprinting by sampled sequencing. *Proc. Natl. Acad. Sci.* **86:** 8902–8906.
21. Rychlik, W. and R.E. Rhoads. 1989. A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and *in vitro* amplification. *Nucleic Acids Res.* **17:** 8543–8551.
22. Asubel, F.M., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl. 1987. *Current protocols in molecular biology*. John Wiley and Sons, New York.
23. Tsiapalis, C.M. and S.A. Narang. 1970. On fidelity of phage T4-induced polynucleotide ligase in joining of chemically synthesized deoxyribonucleotides. *Biochem. Biophys. Res. Commun.* **39:** 631–636.
24. Goffin, C., V. Bailly, and W.G. Verly. 1987. Nicks 3′ or 5′ to AP sites or to mispaired bases, and one-nucleotide gaps can be sealed by T4 DNA liage. *Nucleic Acids Res.* **15:** 8755–8771.
25. Wiaderkiewicz, R. and A. Ruiz-Carrillo. 1987. Mismatch and blunt to protruding-end joining by DNA ligases. *Nucleic Acids Res.* **15:** 7831–7848.
26. Wu, D.Y. and R.B. Wallace. 1989. Specificity of the nick-closing activity of bacteriophage T4 DNA ligase. *Gene* **76:** 245–254.
27. Landegren, U., R. Kaiser, J. Sanders, and L. Hood. 1988. A ligase-mediated gene detection technique. *Science* **241:** 1077-1080.
28. Drmanac, R., N. Petrovic, V. Glisin, and R. Crkvenjakov. 1986. A calculation of fragment lengths obtainable from human DNA with 78 restriction enzymes: An aid for cloning and mapping. *Nucleic Acids Res.* **14:** 4691–4692.
29. Smith, D.R., A.P. Smyth, and D.T. Moir. 1990. Amplification of large artificial chromosomes. *Proc. Natl. Acad. Sci.* **87:** 8242–8246.

Received January 24, 1992; accepted in revised form March 17, 1992.