Nucleic Acids Research, Vol. 20, No. 14 3753–3762

Rapid shotgun cloning utilizing the two base recognition endonuclease *Cvi*JI

Michael C.Fitzgerald, Peter Skowron[1], James L.Van Etten[2], Lloyd M.Smith and David A.Mead*
Chemistry Department, University of Wisconsin, Madison, WI 53706, [1]CHIMERx, Madison, WI 53704 and [2]Department of Plant Pathology, University of Nebraska, Lincoln, NE 58583, USA

Received February 27, 1992; Revised and Accepted June 23, 1992

ABSTRACT

A new approach has been developed for the rapid fragmentation and fractionation of DNA into a size suitable for shotgun cloning and sequencing. The restriction endonuclease *Cvi*JI normally cleaves the recognition sequence PuGCPy between the G and C to leave blunt ends. Atypical reaction conditions which alter the specificity of this enzyme (*Cvi*JI) yield a quasi-random distribution of DNA fragments from the small molecule pUC19 (2686 base pairs). To quantitatively evaluate the randomness of this fragmentation strategy, a *Cvi*JI** digest of pUC19 was size fractionated by a rapid gel filtration method and directly ligated, without end repair, to a lacZ minus M13 cloning vector. Sequence analysis of 76 clones showed that *Cvi*JI** restricts PyGCPy and PuGCPu, in addition to PuGCPy sites, and that new sequence data is accumulated at a rate consistent with random fragmentation. Advantages of this approach compared to sonication and agarose gel fractionation include: smaller amounts of DNA are required (0.2–0.5 μg instead of 2–5 μg), fewer steps are involved (no pre-ligation, end repair, chemical extraction, or agarose gel electrophoresis and elution are needed), and higher cloning efficiencies are obtained (*Cvi*JI** digested and column fractionated DNA transforms 3–16 times more efficiently than sonicated, end-repaired, and agarose fractionated DNA).**

INTRODUCTION

Clone banks of DNA are critical to the nucleotide sequence analysis of organisms and their genes. Depending on the circumstances, the 'perfect' library will be enriched for or unbiased against the particular genetic unit under analysis. A variety of biochemical and biophysical strategies have been utilized to construct such libraries (1). Most large scale DNA sequencing strategies depend on randomly fragmenting the target molecule into small pieces which can be subcloned into a bacteriophage like M13 (2–6). These vectors produce template DNA in a single stranded form, the optimal substrate for enzymatic sequence analysis (7). The data obtained from such cloned subfragments are combined and overlapped until approximately 80–95% of both strands are covered; after which gap filling techniques are typically utilized to complete the sequence.

Four methods are presently used to fragment large DNAs into a size suitable for enzymatic sequence analysis: DNAse I treatment (8), low pressure shearing (9), sonication (10), and digestion with restriction enzymes. Sonication, low pressure shearing, and treatment with DNAse I all break DNA randomly, and result in a collection of overlapping fragments. As these physical methods tend to shear the middle of the monomer targets a preliminary pre-ligation is necessary to equalize the representation of the DNA ends in the final library. Another drawback to these methods is the inefficiency with which the resultant jagged ends can be ligated, necessitating an enzymatic end-repair step prior to cloning. Sonication, the most commonly used method, requires relatively large amounts of DNA, results in a low transformation efficiency, and is technically difficult to automate. DNAse I requires recalibration with new batches and age, is sensitive to trace contaminants, and is somewhat variable in its digestion rate (1). Although fragmentation with restriction enzymes is attractive due to the relative abundance of sequence specificities available, it is seldom utilized. A complete restriction digest results in non-overlapping fragments and partial digests often exhibit non-uniform restriction rates. Generally, as many as four separate libraries utilizing four different restriction digests must be prepared to supply overlaps between fragments.

The steps involved in constructing a random clone library (shotgun cloning) for DNA sequencing by current methods include 1) isolating the DNA fragment, 2) ligating the DNA to itself, 3) randomly shearing the material by sonication, 4) repairing the ragged ends with a DNA polymerase or nuclease, 5) size fractionation by preparative agarose gel electrophoresis, 6) extraction with organic chemicals to re-purify the DNA, 7) ligating this product into a bacteriophage cloning vector, usually M13mp18 or 19, and 8) transforming special strains of competent *E. coli* (4). These steps are inherently difficult to automate and require large amounts of DNA, as the sonication and/or

* To whom correspondence should be addressed

fractionation steps result in low cloning efficiencies. In addition, the entire process is lengthy, typically requiring several days for a skilled researcher to complete.

As an alternative to these methods, we reasoned that a partial digest with a two or three base recognition endonuclease would cleave DNA frequently enough to be functionally random with respect to the rate at which sequence data can be accumulated from a shotgun clone bank. The restriction enzyme *Cvi*JI normally recognizes the sequence PuGCPy and cleaves between the G and C to leave blunt ends (11). Under 'relaxed' conditions (in the presence of 1mM ATP and 20 mM DTT) the specificity of *Cvi*JI can be altered to cleave the dinucleotide GC. This activity is referred to as *Cvi*JI*, for star or altered specificity. Because of the high frequency of the dinucleotide GC in all DNA (16 bp average fragment size for random DNA), the construction of quasi-random libraries should be possible by partial digestion with *Cvi*JI*. A DNA degradation method with low levels of sequence specificity should produce a smear of the target DNA when analyzed by agarose gel electrophoresis. Digestion of the small DNA molecule pUC19 under partial *Cvi*JI* conditions does not result in a non-discrete smear; rather, a number of discrete bands are found superimposed upon a light background of smearing. Atypical reaction conditions which further alter the specificity of *Cvi*JI* (termed *Cvi*JI**) were developed which result in a quasi-random distribution of DNA fragments. The work described in this paper replaces steps 2 through 6 above with a simple DNA fragmentation and fractionation strategy. The use of the restriction enzyme *Cvi*JI**, in combination with a rapid gel filtration size exclusion step, streamlines a number of aspects involved in shotgun cloning.

MATERIALS AND METHODS

Reagents

The enzymes *Cvi*JI, *Eco*RV, and calf intestinal alkaline phosphatase were purchased from CHIMERx (Madison, WI). *Bacillus sterothermophilus* DNA polymerase was obtained from BioRad Laboratories (Richmond, CA), and DNA polymerases from bacteriophage T4 and the *E. coli* Klenow fragment were obtained from United States Biochemicals (Cleveland, OH). T4 DNA ligase, *Hinc*II, and pUC19 (2) were purchased from New England Biolabs (Beverly, MA). Lambda DNA and ampicillin were from Sigma (St. Louis, MO), and X-GAL and IPTG were obtained from Gold Biotechnologies (St. Louis, MO). Sephacryl S-500 and the Sephaglass M13 miniprep kit were obtained from Pharmacia LKB (Piscataway, NJ). The 100bp ladder and 1kbp ladder molecular weight markers, as well as the competent bacteria were obtained from Life Technologies (Bethseda, MD). The lacZ minus M13 sequencing primer (5'-GGGGAAA-GCCGGAGAACGTGG-3') was synthesized on an Applied Biosystems 380B DNA synthesizer at the University of Wisconsin Biotechnology Center.

DNA Fragmentation and End Repair Conditions

Fragments of pUC19 and lambda DNA, ranging from 200–2000 base pairs (bp), were generated by either sonication or restriction with the endonuclease *Cvi*JI (11). DNA was mechanically sheared by high frequency vibrations utilizing a Heat Systems Ultrasonics (Farmingdale, NY) W-375 cup horn sonicator as specified by Bankier (4). DNA was biochemically fragmented with the restriction enzyme *Cvi*JI under the following conditions: 1–5 μg of DNA was incubated at 37°C with 0.33–1.65 units of *Cvi*JI in the presence of 1 mM ATP, 20 mM DTT, 20% v/v DMSO, and a modified reaction buffer (10mM Tris-HCl pH 8.0, 10 mM $MgCl_2$, 10 mM NaCl) for various time periods. This treatment is referred to as *Cvi*JI** reaction conditions. The reaction was terminated by heating at 65°C for 15 minutes. High resolution analysis of these digests was achieved by electrophoresis on 2% agarose gels in 90mM Tris-HCl, 90mM boric acid, 2 mM EDTA, pH 8.0 (12).

The ragged ends of the sonicated DNA were rendered blunt utilizing two different end repair reactions. In one end repair reaction (ER 1) sonicated DNA was treated according to the procedure outlined by Bankier (4), where 2.0 μg of sonicated lambda DNA is combined with 10 units Klenow DNA *Pol I*, 10 units T4 DNA *Pol*, 0.1 mM dNTPs, and reaction buffer (50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 10 mM DTT). This mixture was incubated at room temperature for 30 minutes followed by heat denaturation of the enzymes at 65°C for 15 minutes. In a second end repair reaction (ER 2), an excess of the reagents and enzymes described above were utilized to ensure a more efficient conversion to blunt ends. In this reaction, 0.2 μg of the sonicated lambda DNA sample was treated under the same reaction conditions described above.

Size Fractionation

*Cvi*JI** digested and sonicated samples were size fractionated to eliminate small DNA fragments by agarose gel electrophoresis and electroelution or by spin columns packed with the size exclusion gel matrix, Sephacryl S-500. Spin columns (0.4 cm in diameter) were packed to a height of 1.3 cm by adding 1 ml of slurry and centrifuging at 2000 RPM for 5 minutes in a Beckman CPR centrifuge. The columns were rinsed 3 times with 1 ml aliquots of 100 mM Tris-HCl (pH 8.0) by centrifugation at 2000 RPM for 2 minutes. Typically, 0.2–2.0 μg of fragmented DNA in a total volume of 30 μl was applied to the column. The void volume, containing the larger DNA fragments, was recovered in the column eluant after spinning at 2000 RPM for 5 minutes. The capacity of this micro-column procedure is 2 μg of DNA. Agarose gel electrophoresis and electroeleution have been described in detail elsewhere (1,4). In these experiments, 5 μg of sample was pipetted into a 2 cm wide slot on a 1% agarose gel. Electrophoresis was halted after the bromophenol blue tracking dye had migrated 6 cm. Fragments larger than 750 bp, as judged by molecular weight markers, were separated from smaller sizes and electrophoresed onto dialysis tubing (1000 MW cutoff). The fractionated material was phenol-chloroform extracted and precipitated.

Cloning Sonicated and *Cvi*JI** Digested Lambda DNA

To compare the cloning efficiencies of sonicated and *Cvi*JI** digested nucleic acid, lambda DNA was fragmented by each method and ligated to *Hinc*II restricted and dephosphorylated pUC19. Ligation reactions were added to competent DH5αF' (*gyr*A96 *rec*A1 *rel*A1 *end*A1 *thi*-1 *hsd*R17 *sup*E44 λ^-) as specified by the manufacturer and aliquots of the transformation mixture were plated on T agar (2) containing 20 μg/ml ampicillin, 25 μl of a 2% solution of isopropylthiogalactoside (IPTG) and 25 μl of a 2% solution of 5-dibromo-4-chloro-3-indolylgalactoside (X-GAL). DNA fragmented by *Cvi*JI** digestion and sonication was cloned both before and after Sephacryl S-500 size fractionation. Sonicated lambda DNA was subjected to an end repair treatment prior to ligation.

Ligation reactions were carried out overnight at 12°C in 20

μl mixtures using the following conditions: 25 mM Tris-HCl (pH 7.8), 10 mM $MgCl_2$, 1 mM DTT, 1 mM ATP, DNA, and 2000 units of T4 DNA ligase. For unfractionated samples,10 ng of fragments and 100 ng of *Hin*cII restricted, dephosphorylated pUC19 were combined under the conditions just described. For fractionated samples, 50 ng of size-selected fragments were ligated with 100 ng of *Hin*cII restricted, dephosphorylated pUC19. This increase in fractionated DNA was determined empirically to compensate for the lower concentration of 'ends' resulting from the fractionation procedure and/or the lowered efficiency of cloning larger fragments. One tenth of the ligation reaction (2 μl) was utilized in the transformation procedure and the fraction of nonrecombinant (blue) and recombinant (white) colonies was used to calculate the efficiency of this process.

Cloning and Sequencing *Cvi*JI** Digested pUC19

pUC19 DNA was digested under *Cvi*JI** conditions and size fractionated as described above. The fractionated DNA was cloned into the *Eco*RV site of M13SPSI, a *lac*Z minus vector constructed by adding an EcoRV restriction site to wild type M13 at position 5605 (13). M13SPSI lacks a genetic cloning selection trait, therefore after ligation of the pUC19 fragments into the vector the sample was restricted with *Eco*RV to reduce the background of nonrecombinant plaques. Bacteriophage M13 plaques were picked at random and grown for 5−7 hours in 2 ml of 2×TY broth containing 20 μl of a DH5αF′ overnight culture. After centrifuging to remove the cells, single stranded phage DNA was purified using Sephaglass as specified by the manufacturer. The single stranded DNA was sequenced by the dideoxy chain termination method using a radiolabeled M13-specific primer and *Bst* DNA polymerase (14). The first 100 bases of 76 clones were read from the autoradiograph to determine the orientation and sequence of the cloned DNA fragments.

RESULTS

*Cvi*JI** Quasi-Random Fragmentation of DNA

In order to generate an overlapping population of DNA fragments suitable for shotgun cloning and sequencing, it was necessary to determine *Cvi*JI partial conditions yielding an apparently random pattern, or smear, of fragments in the appropriate size range. Initially, agarose gel electrophoresis and ethidium bromide staining of the treated DNA was utilized to assess the randomness and size distribution of the fragments. Common methods for obtaining partially restricted DNA include limiting the incubation time or limiting the amount of enzyme. pUC19 DNA incubated for limited time periods or with limiting amounts of *Cvi*JI under normal or relaxed conditions did not produce a quasi-random restriction pattern, or smear. Instead, a number of discrete bands were observed, as shown in Fig. 1, lane 3 for the *Cvi*JI* partial digestion of pUC19. Complete digests of pUC19 under normal and *Cvi*JI* buffer conditions are shown in lanes 1 and 2 respectively.

In an attempt to eliminate strong restriction site preferences observed under partial restriction conditions, a series of altered reaction conditions were explored (15, 16). Conditions of high pH, low ionic strength, addition of solvents such as glycerol or dimethylsulfoxide, and/or substitution of Mn^{2+} for Mg^{2+} were systematically tested with *Cvi*JI endonuclease using the plasmid pUC19. The best enzymatic 'smearing' pattern was obtained when the ionic strength of the relaxed reaction buffer was lowered and an organic solvent was added (Fig.1, lane 4). Plasmid pUC19 partially digested under these conditions yields a relatively non-discrete smear that resembles the results obtained from sonicating this molecule, as demonstrated in Fig. 2, lanes 5 and 6. This activity is referred to as *Cvi*JI** to differentiate it from the originally characterized star activity (11). The appearance of diffuse, faint bands overlying a background smear generated from this 2686 bp molecule indicates that some weakly preferred or resistant restriction sites could bias the results of subsequent cloning experiments. Utilizing a large substrate such as lambda

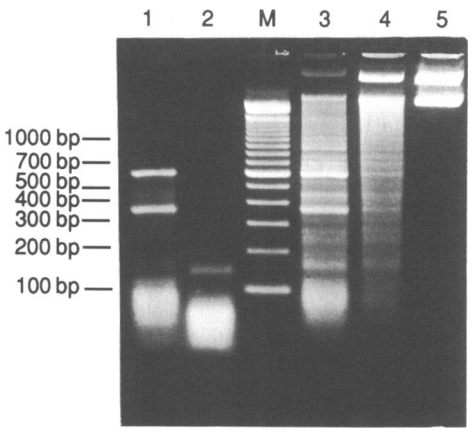

Figure 1. *Cvi*JI restriction digests of pUC19. Lane M: 100 bp DNA ladder. Lanes 1−4: pU19 (1.0 μg) was digested at 37°C in a 20μl volume for the indicated times and conditions. Lane 1: complete *Cvi*JI digest (1 unit of enzyme for 90 min in 50 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$, 50mM NaCl). Lane 2: complete *Cvi*JI* digest (1 unit of enzyme for 90 min in 50 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$, 50mM NaCl, 1 mM ATP, 20 mM DTT). Lane 3: partial *Cvi*JI* digest (0.25 units of enzyme for 30 min in 50 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$, 50mM NaCl, 1 mM ATP, 20 mM DTT). Lane 4: partial *Cvi*JI** digest (0.5 units of enzyme for 60 min in 10 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$, 10mM NaCl, 1 mM ATP, 20 mM DTT, 20% v/v dimethylsulfoxide). Lane 5: uncut pUC19 (1.0 μg).

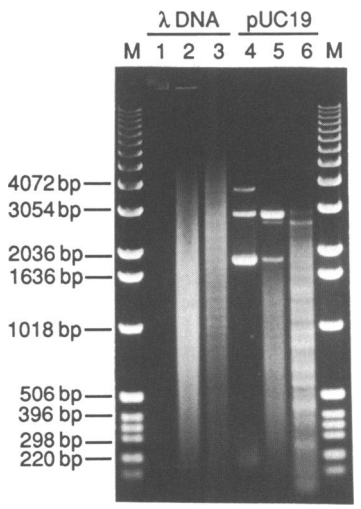

Figure 2. Comparison of sonicated versus *Cvi*JI** partially digested DNAs. Lanes M: 1 kbp DNA ladder. Lanes 1−3: untreated λ DNA (0.25 μg), sonicated λ DNA (1.0 μg), and *Cvi*JI** partially digested λ DNA (1.0 μg). Lanes 4−6: untreated pUC19 (0.25 μg), sonicated pUC19 (1.0 μg), and *Cvi*JI** partially digested pUC19 (1.0 μg).

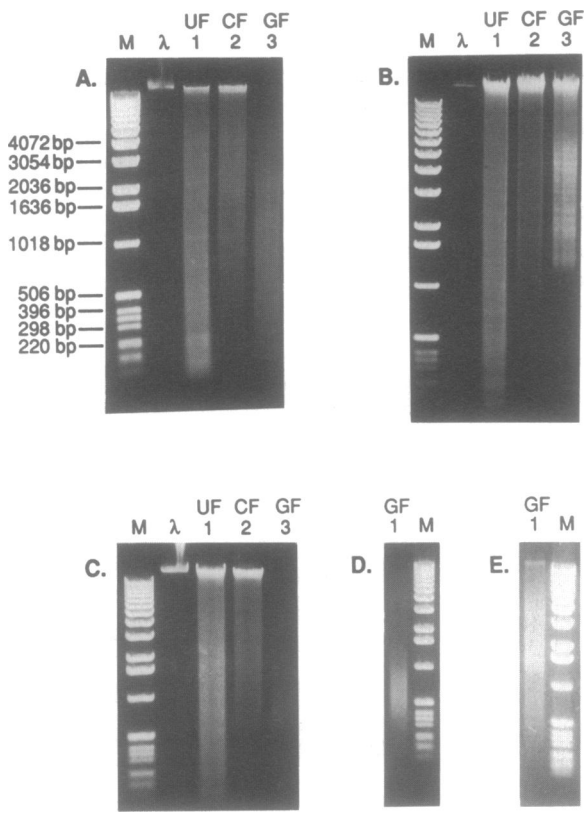

Figure 3. Size fractionation of DNA by micro-column chromatography compared to agarose gel electroelution. A.) Lane M, 1 kbp DNA ladder; lane λ, untreated λ DNA (0.25 μg); lane 1, unfractionated (UF) *Cvi*JI** partially digested λ DNA (1.0 μg); lane 2, column fractionated (CF) *Cvi*JI** partially digested λ DNA (1.0 μg); lane 3, gel fractionated (GF) *Cvi*JI** partially digested λ DNA (1.0 μg). B.−E.) Additional trials of the same treatments as in A.

Table 1. Cloning efficiencies[1] of *Cvi*JI** partially digested lambda DNA fractionated by micro-column chromatography versus agarose gel electroelution.

	Trial I		Trial II	
	Colony Phenotype			
DNA/treatment	Blue	White	Blue	White
Supercoiled pUC19	55000	<10	50000	<10
pUC19/*Hinc*II/CIAP	210	<1	320	1
pUC19/*Hinc*II/CIAP/ T4 DNA ligase	150	4	210	7
λ/*Cvi*JI** partial/CF + pUC19	140	240	210	240
λ/*Cvi*JI** partial/GFE1 + pUC19	98	49	200	18
λ/*Cvi*JI** partial/GFE2 + pUC19	82	54	95	74

[1]number of ampicillin resistant colonies/ng pUC19
CIAP: calf intestinal alkaline phosphatase
CF: column fractionated
GFE: agarose gel fractionated and electroeluted
λ: bacteriophage lambda

DNA (45 kbp) revealed essentially no banding differences between the biochemical and biophysical methods, as demonstrated in Fig. 2, lanes 2 and 3. As expected, the minor bias evident with a small molecule such as pUC19 was not detectable with a larger substrate such as lambda.

The intensity and duration of sonic treatment affects the size distribution of the resulting DNA fragments. The results obtained from the sonication of lambda and pUC19 samples (Fig. 2) were obtained from three 20 second pulses at a power setting of 60 watts. Sonicated and *Cvi*JI** generated smears are similar, although the size distribution of fragments is consistently greater with *Cvi*JI** fragmentation. This result favors the cloning of larger inserts, which should facilitate the efficiency of end-closure strategies (5). The size distribution of the DNA fragmented by *Cvi*JI** is controlled by incubation time and amount of enzyme. An excess of enzyme or a long incubation time will completely digest pUC19 DNA, resulting in fragments which range in size from several bp to approximately 150 bp (Fig. 1, lanes 1 and 2). The results shown in Fig. 2 were obtained by incubating pUC19 for 40 minutes and lambda DNA for 60 minutes with 0.33 units of *Cvi*JI/μg substrate. The efficiencies of the two methods for randomly fragmenting DNA were quantitatively analyzed for molecular cloning purposes, as described below.

Rapid DNA Size Fractionation Utilizing Spin Column Chromatography

The amount of data obtained by the shotgun sequencing approach is substantially increased if those fragments less than 500bp are eliminated prior to the cloning step. Small fragments yield only a portion of the data which can be collected from polyacrylamide gel based separations and thus they lower the efficiency of this strategy. Agarose gel electrophoresis and electroelution is commonly used to size fractionate DNA prior to shotgun cloning (4). Approximately three hours are required to prepare the agarose gel, electrophorese the sample, electroelute fragments larger than 500 bp, perform the phenol-chloroform extractions, and precipitate the resulting material. The results of 5 out of 9 independent trials size fractionating *Cvi*JI** fragmented lambda DNA by this method are shown in Fig. 3A−E.

Small DNA fragments can also be removed by passing the sample through a short column of Sephacryl S-500 (see methods). Approximately 15 minutes are needed to prepare the column and 5 minutes to fractionate the DNA by this method. The results of three out of nine such trials are shown in Fig. 3A−C. The efficiency of eliminating small DNA fragments (<500 bp) by spin column chromatography appears high, and the reproducibility was excellent. This result is in contrast to the agarose gel electrophoresis and electroelution data presented in Fig. 3A−E. Nine replicate trials of this method yielded nine differently sized products, regardless of the source of the agarose. Both methods yielded 30−40% recoveries as measured by UV spectrophotometry. To quantitate the relative efficiencies of the two fractionation methods, the lambda DNA size fractionated in Fig. 3A lanes 2 and 3, and 3B lane 3 were analyzed for cloning efficiency and insert size, as described below.

Cloning Efficiencies of Gel Elution and Chromatography Fractionation Methods

The efficacy of size selection was quantified by two criteria: 1) comparing the relative cloning efficiency of *Cvi*JI** partially digested lambda DNA fragments fractionated either by agarose gel electrophoresis and electroelution or micro-column chromatography, and 2) determining the size distribution of the resulting cloned inserts. To reduce potential variables, large quantities of the cloning vector and ligation cocktail were prepared, ligation reactions and transformation of competent *E. coli* were performed on the same day, numerous redundant controls were performed, and all cloning experiments were repeated twice. The cloning efficiencies reported are the average of triplicate platings of each ligation reaction. The concentration

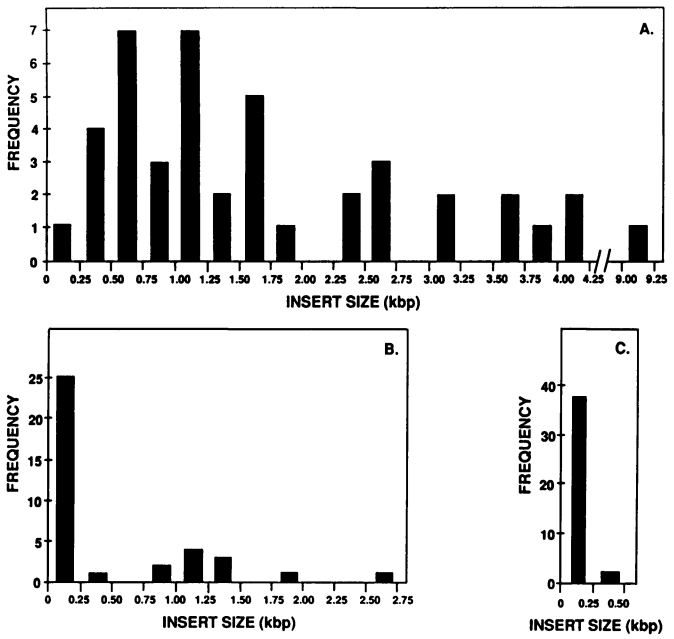

Figure 4. Size distribution of fragments cloned after fractionation by micro-column chromatography or agarose gel electroelution. A.) A *Cvi*JI** partial digest of 2μg of λ DNA was size fractionated on a 4 mm by 13 mm column of sephacryl S-500 at 2,000×*g* for 5 minutes. The void volume was directly ligated to linear, dephosphorylated pUC19 and 43 resulting clones were analyzed for insert size. The DNA for this experiment is the same as that shown in Fig. 2A, lane 2. B.) A *Cvi*JI** partial digest of 5 μg of λ DNA was size fractionated by agarose gel electroelution. The eluted DNA was phenol extracted and ligated to linear, dephosphorylated pUC19 and the resulting 40 clones were analyzed for insert size. The DNA for this experiment is the same as that shown in Fig. 2A, lane 3. C) Same as in B, except the DNA for this experiment came from Fig. 2B, lane 3.

Table 2. Cloning efficiencies[1] of *Cvi*JI** partially digested lambda DNA versus sonication.

DNA/treatment	Trial I Colony Phenotype Blue	Trial I White	Trial II Blue	Trial II White
A. Unfractionated Samples				
Supercoiled pUC19	30000	<10	16000	<10
pUC19/*Hin*cII/CIAP	150	<1	31	1
pUC19/*Hin*cII/CIAP/ T4 DNA ligase	100	<1	15	1
λ/*Alu*I + pUC19	200	400	73	250
λ/*Cvi*JI** Partial + pUC19	100	160	97	340
λ/Sonicated + pUC19	–	–	11	29
λ/Sonicated/ER 1 + pUC19	17	10	10	44
λ/Sonicated/ER 2 + pUC19	–	–	40	100
B. Fractionated Samples				
Supercoiled pUC19	35000	<10	12000	<10
pUC19/*Hin*cII/CIAP	30	<1	180	<1
pUC19/*Hin*cII/CIAP/ T4 DNA ligase	60	<1	10	<1
λ/*Alu*I + pUC19	28	23	33	48
λ/*Cvi*JI** Partial + pUC19	31	90	36	68
λ/Sonicated + pUC19	20	6	99	19
λ/Sonicated/ER 1 + pUC19	27	32	40	19
λ/Sonicated/ER 2 + pUC19	–	–	25	63

[1]number of ampicillin resistant colonies/ng pUC19.
CIAP: calf intestinal alkaline phosphatase
ER 1 and 2: end repair reactions detailed in methods section
λ: bacteriophage lambda

of the fractionated material was checked spectrophotometrically so that 50 ng was added to all ligation reactions. This material was ligated to *Hin*cII digested and dephosphorylated pUC19. This cloning vector was chosen because it permits a simple blue to white visual assay to indicate whether a DNA fragment was cloned (2).

A summary of the cloning efficiencies calculated from two independent trials is given in Table 1. These trials represent repeated experiments in which lambda DNA fragments generated by *Cvi*JI** partial digestion were ligated to *Hin*cII linearized, dephosphorylated pUC19 and transformed into DH5αF′ competent cells. The first three rows in Table 1 show controls performed to establish a baseline to better evaluate the various treatments. Supercoiled pUC19 transforms *E. coli* 10 times more efficiently than the *Hin*cII restricted plasmid (data not presented) and 150–260 times more efficiently than the *Hin*cII restricted and dephosphorylated plasmid. The number of blue and white colonies which resulted from transforming *Hin*cII cut and dephosphorylated pUC19 was determined both before and after treatment with T4 DNA ligase in order to differentiate these background events from cloning inserts. The background of blue colonies (which presumably represent the uncut and/or non-dephosphorylated population of molecules) averaged 0.4%, compared to supercoiled plasmid. The background of white colonies (which presumably results from contaminating nucleases in the enzyme treatments or genomic DNA in the plasmid preparations) after *Hin*cII restriction, dephosphorylation, and ligation of pUC19 averaged 0.014% as compared to the supercoiled plasmid.

The number of white colonies obtained when micro-column fractionated DNA was cloned into pUC19 was 240/ng vector in both trials (Table 1). The efficiency of cloning gel fractionated and electroeluted DNA ranged from 18–74 white colonies/ng vector. The data show that column fractionated DNA results in three to thirteen times the number of white colonies, and presumably recombinant inserts, as gel fractionated and electroeluted DNA. The size distribution of the inserts present in these white colonies is depicted in Figure 4. A total of 43 random clones obtained from micro-column chromatography fractionation were analyzed for insert size (Fig. 4A). Most of these inserts were larger than 500 bp (37/43 or 86%), 11.6% (5/43) were smaller than 500 bp, and one clone (2.3%) was smaller than 250 bp. The average insert size was 1630 bp. These results are in contrast to those obtained by agarose gel fractionation (Fig. 4B and C). In the first trial (Fig. 4B) most of the inserts were smaller than 500 bp (26/37 or 70.3%) and only 29.7% (11/37) were larger than 500 bp in size. In the second trial (Fig. 4C) all of the inserts (40 total) were smaller than 500 bp. This last result was surprising in view of the apparently efficient size selection achieved (see Fig. 3B, lane 3).

Fragmentation Methods And Molecular Cloning Efficiencies

The efficacy of the methods was quantified by comparing the cloning efficiency of lambda DNA fragments generated either by sonication or *Cvi*JI** partial digestion. To reduce potential cloning differences based on size preference, the size distribution of the DNA generated by these two methods was closely matched. Other experimental details were designed to reduce potential variables, as described above. Certain variables were unavoidable, however. For example, the sonicated DNA

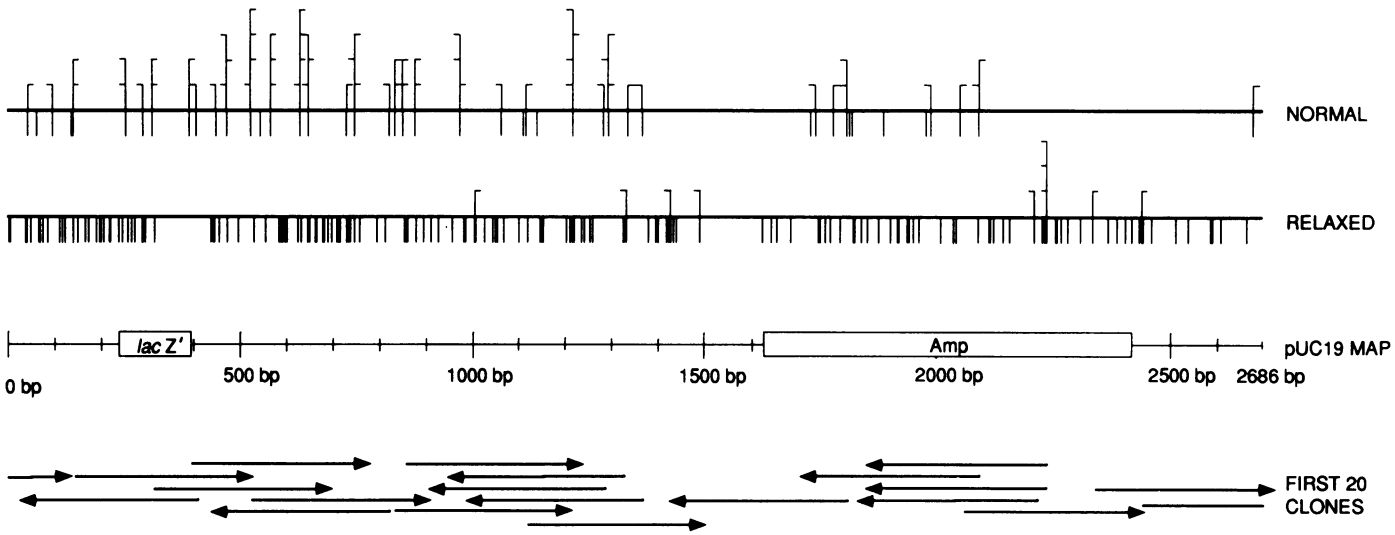

Figure 5. Distribution of *Cvi*JI sites in pUC19. A linear map of pUC19 showing the relative position of the lacZ′ gene (α peptide of B-galactosidase gene) and ampicillin resistance gene (Amp) is indicated for reference. The marks extending beneath the top line (labeled normal) show the relative position of the 45 normal *Cvi*JI sites (PuGCPy) present in pUC19. The marks above the line are the cleavage sites found from sequencing the *Cvi*JI** partial library. The height of the line indicates the number of clones obtained from cleavage at that site, and the orientation of the flag designates the right or left orientation of the respective clone. The marks extending beneath the second line (labeled relaxed) show the relative positions of the 160 *Cvi*JI* sites (GC) present in pUC19. Those marks above the line were found from sequencing the *Cvi*JI** partial library. The bottom portion of the figure shows the relative position and orientation of the first 20 clones sequenced, assuming a 350 bp read per clone.

fragments required an enzymatic step to repair the ragged ends prior to ligation (see methods section), whereas the *Cvi*JI** digests were heat denatured and directly ligated to *Hin*cII digested pUC19.

A summary of the cloning efficiencies calculated from two independent trials is given in Table 2A and B. These trials represent repeated experiments in which two identical sets of lambda DNA fragments generated by *Alu*I complete digestion, *Cvi*JI** partial digestion, or sonication were each ligated to *Hin*cII linearized, dephosphorylated pUC19 and transformed into DH5αF′ competent cells. The cloning efficiencies reported are the average of triplicate platings of each ligation reaction. In case the Sephacryl S-500 size fractionation step introduced inhibitors of ligation or transformation or resulted in differences attributable to the size of the material, the sonicated and *Cvi*JI** digested samples were ligated with pUC19 both prior to (A) and after (B) the fractionation steps. The first three rows in Table 2A and B are controls performed to establish a baseline to better evaluate the various treatments. These data show that supercoiled pUC19 transforms *E. coli* 200−1000 times more efficiently than the *Hin*cII restricted and dephosphorylated plasmid. Without this dephosphorylation step, the cloning efficiency is 10% that of the supercoiled molecule (data not presented). The background of blue colonies averaged 0.5% in these experiments, compared to supercoiled plasmid, while the background of white colonies averaged 0.005%.

A comparison of the data from unfractionated versus fractionated samples in Table 2A and B reveals a general decline in the number of white and blue colonies obtained after sizing. This decrease is primarily due to the fact that cloning efficiencies are dependent upon the size of the fragment, favoring smaller fragments and thus giving higher efficiencies for the unfractionated material. This is illustrated by comparing the efficiency of cloning unfractionated and fractionated lambda DNA which was completely restricted with *Alu*I. This four base recognition endonuclease produces blunt ends and restricts lambda (48,502 bp) at 143 sites. Only 25 of the resulting 144 fragments (17%) are larger than 500 bp. The number of white colonies obtained when unfractionated lambda DNA completely restricted with *Alu*I was cloned into pUC19 ranged from 250−400/ng vector versus 23−48/ng vector for the fractionated material. This ten fold decrease was only noticed for the lambda *Alu*I digests, and probably reflects the large portion of small molecular weight fragments (approximately 75%) which is excluded from the fractionated ligation reactions.

The number of white colonies obtained when unfractionated *Cvi*JI** treated lambda DNA was cloned into pUC19 ranged from 160−340/ng vector, versus 68−90 white colonies/ng vector if the same material was fractionated. Unfractionated lambda DNA completely digested with *Alu*I results in cloning efficiencies very similar to unfractionated *Cvi*JI** treated DNA. Sonicated lambda DNA is a poor substrate for ligation, compared to *Cvi*JI** treatment, as indicated by the roughly ten fold reduced cloning efficiencies. Enzymatic repair of the ragged ends produced by sonication results in an increased cloning efficiency. Using conditions described for the first end repair treatment (ER 1), 10−44 (fractionated) and 19−32 (unfractionated) white colonies/ng vector were observed. However, ER 1 conditions (4) may not be optimal, as an alternate end repair reaction (ER 2) resulted in greater numbers of white colonies (63 and 100/ng vector for fractionated and unfractionated DNA, respectively). In this reaction, a ten fold excess of reagents and enzymes were utilized to repair the sonicated DNA, which apparently improved the efficiency of cloning such molecules by two to three fold. It is not surprising that the efficiency of rendering the sonicated ends of DNA blunt is correlated with their cloning efficiency. The data collected from multiple cloning trials in Table 2A and B shows that *Cvi*JI** partial digestion results in three to sixteen

Table 3. Distribution of cloned *CviJI*** partially digested pUC19 sites.

Classification Group	Recognition Sequence	NGCN Site Distribution in pUC19 (%)	Cloned *CviJI*** Distribution (%)	Pu/Py Structure
Normal (N)	A C GC G T	AGCC 9 (4.4) GGCC 11 (5.4) GGCT 10 (4.9) AGCT 15 (7.3) 45 22.0)	13 (17.1) 16 (21.1) 12 (15.8) 25 (32.9) 66 (86.9)	PuPuPyPy
Relaxed (R_1)	C C GC T T	CGCC 11 (5.4) TGCC 12 (5.9) TGCT 10 (4.9) CGCT 22 (10.7) 55 (26.9)	0 2 (2.6) 1 (1.3) 2 (2.6) 5 (6.5)	PyPuPyPy
Relaxed (R_2)	A A GC G G	AGCA 16 (7.3) GGCA 8 (3.9) AGCG 11 (5.4) GGCG 22 (10.7) 57 (27.8)	1 (1.3) 0 0 4 (5.2) 5 (6.6)	PuPuPyPu
Relaxed (R_3)	C A GC T G	CGCA 10 (4.9) TGCA 13 (6.3) CGCG 10 (4.9) TGCG 15 (7.3) 48 (23.4)	0 0 0 0 0	PyPuPyPu

times the number of white colonies than sonicated ER 1 treated DNA. Even with an optimal end repair reaction for the sonicated fragments, DNA treated with *CviJI*** yielded three times more white colonies.

Analysis of *CviJI*** Fragmentation For Shotgun Cloning and Sequencing

The ability of *CviJI*** partial digestion to create uniformly representative clone libraries for DNA sequencing was tested on pUC19 DNA. *CviJI*** treated pUC19 was ligated to a *lacZ* minus M13 vector for nucleotide analysis. The first 100 bases of 76 randomly chosen clones were sequenced to determine which *CviJI* recognition site was utilized, the orientation of each insert and how effectively the cloned fragments covered the entire molecule (Fig. 5). The positions of the 45 normal *CviJI* sites (PuGCPy) in pUC19 are indicated beneath the line labeled 'Normal' in the figure. Similarly, the 160 *CviJI** sites (GC) are indicated beneath the line labeled 'Relaxed' in Fig. 5. The marks above these lines indicate the *CviJI*** pUC19 sites which were found in the set of 76 random clones sequenced here. The frequency of cloning a particular site is indicated by the height of the line, and the left or right orientation of each clone is also indicated at the top of each mark. There are a total of 205 *CviJI* and *CviJI** sites in pUC19.

The data presented in Fig. 5 demonstrate that under *CviJI*** partial conditions normal *CviJI* sites are preferentially restricted over relaxed (*CviJI**) sites. Of the 76 clones that were analyzed only 13%, or 1 in 7, had sequence junctions corresponding to a relaxed *CviJI** site. Thirty-five of the forty-five possible normal restriction sites were cloned, compared to eight of the possible one hundred sixty relaxed sites. If the enzyme had shown no preference for normal or relaxed sites under the *CviJI*** partial conditions utilized here, then 78% of the sequence junctions analyzed should have been generated by cleavage at a relaxed *CviJI** site. It is interesting to note that the relaxed *CviJI** restriction sites that were found appear to be clustered in two regions of the plasmid that are deficient in normal *CviJI* sites. In addition, the combined distribution of the normal and relaxed sites which were restricted to generate the 76 clones appears to be quasi-random. That is, the longest gap between cloned restriction sites was no greater than 250 bp and no one particular site is over-utilized.

A detailed analysis of the distribution of *CviJI*** sequence junctions found from cloning pUC19 is presented in Table 3. The GC sites in pUC19 can be divided into four classes based on their flanking Pu/Py structure. The fraction of GC sites observed in pUC19 which belong to each classification is roughly equal (22.0−27.8%). A striking difference was found between the observed distribution in pUC19 of normal and relaxed (R1, R2, R3) *CviJI* recognition sites and the distribution revealed by shotgun cloning and sequence analysis of *CviJI*** treated DNA. While most of the sites cleaved by this treatment were found to be PuGCPy (87%), or 'normal' restriction sites, a significant fraction of the cleavage occured at PyGCPy (6.5%) and PuGCPu (6.6%) sites, considering the short incubation times and limiting enzyme concentrations. The latter two categories of sites, and presumably the PyGCPu sites as well, are completely restricted under 'relaxed' conditions, provided an excess of enzyme is present and sufficient time is allowed (see Fig. 1 and reference 11). Despite the low frequency of relaxed cleavage, *CviJI*** treatment results in a relatively even distibution of breakage points across the length of the molecule (see Fig. 5). *CviJI*** cleavage at relaxed sites appears to be important in 'filling gaps' left by normal restriction.

The primary goal of this effort was to determine the efficacy of these methods for rapid shotgun cloning and sequencing. For the purposes of this study, only 100 bases of sequence data were acquired per clone. However, if 350 bases of sequence had been determined from each clone then the entire sequence of pUC19 would have been assembled from the overlap of the first 20 clones

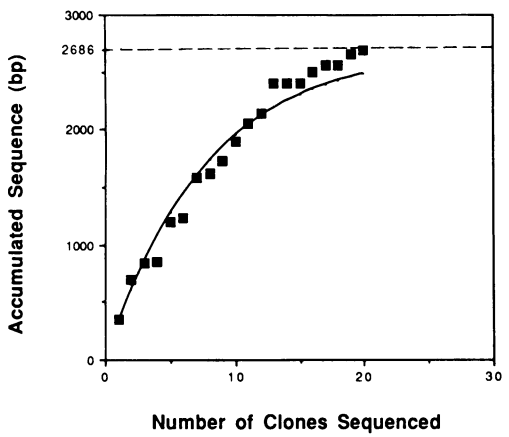

Figure 6. Rate of sequence accumulation by *Cvi*JI** shotgun cloning and sequencing. The points represent a plot of the total amount of determined pUC19 sequence versus the total number of clones sequenced. The horizontal dashed line demarcates the 2686 bp length of pUC19. The smooth curve represents a continuous plot of the discrete function $S(N)=NLe^{-cs}[((e^{cs}-1)/c)+(1-s)]$. See text for details.

(Fig. 5). In this sequencing simulation 75% of pUC19 would have been sequenced at least 2 times from the first 20 clones. The highest degree of overfold sequencing would have been 6, and only involved 2.2% of the DNA. Fig. 5 also shows that most of the 1× sequencing coverage occurred in a region of the plasmid with a very low density of normal and relaxed *Cvi*JI restriction sites. Most of the single coverage occurs in a 240 bp region of the plasmid between 1490 bp and 1730 bp where there are only 4 *Cvi*JI relaxed sites. It should also be noted that by the 27th randomly picked clone most of this region would have been covered a second time (data not presented).

Shotgun sequencing strategies are efficient for accumulating the first 80–95% of the sequence data. However, the random nature of the methods means that the rate at which new sequence is accumulated decreases as more clones are analyzed. In Fig. 6 the total amount of unique pUC19 sequence accumulated was plotted as a function of the number of clones sequenced. The theoretical accumulation curve expected for a process in which sequence information is acquired in a totally random fashion is also shown. The smooth curve is a continuous plot of the discrete function S(N) where

$$S(N)=NLe^{-c\sigma}[((e^{c\sigma}-1)/c+(1-\sigma)]$$

This equation is based upon the results developed by Lander and Waterman (17) for the progress of contig generation in genetic mapping. In the equation: N is the number of clones sequenced, L is the length of clone insert in bp, c is the redundancy of coverage or LN/G (where G is length of fragment being sequenced), and $\sigma = 1-\Theta$, where Θ is the fraction of length that two clones must share. The curve in Fig. 6 was calculated with G = 2686, L = 350 bases, and $\sigma = 1$. The plotted points lie close to the theoretical curve, and it thus appears that the sequence of pUC19 was accumulated in a random fashion utilizing CviJI** fragmentation and column fractionation. This statement does not imply that CviJI recognizes or cleaves DNA randomly, a feat which would be difficult for a sequence specific endonuclease.

Shotgun Cloning Utilizing 200 ng of Lambda DNA

Generally, 2–5 μg of DNA are needed for the sonication and agarose gel fractionation method of shotgun cloning in order to provide the several hundred colonies or plaques required for sequence analysis (4). A ten fold reduction in the amount of substrate required would greatly simplify the construction of such libraries, especially from large genomes (6). The efficiency of constructing a large shotgun library from nanogram amounts of substrate was tested utilizing 200 ng of *Cvi*JI** digested lambda DNA. This material was column fractionated as described previously. In this case, 1/2 of the column eluant (15 μl containing 50 ng of DNA) was ligated to 100 ng of *Hinc*II digested and dephosphorylated pUC19. The cloning efficiencies of the control DNAs were similar to those reported in Tables 1 and 2. The 50 ng cloning experiment yielded 230 white colonies in one trial and 410 white colonies per ligation reaction in a second trial. Thus, it should be possible to routinely construct useful quasi-random shotgun libraries from as little as 0.2–0.5 μg of starting material.

DISCUSSION

We have developed a new method for generating quasi-random shotgun libraries from nanogram amounts of DNA. The use of the two/three base recognition endonuclease *Cvi*JI, in conjunction with a simple fractionation method, produces libraries equivalent in final form to those generated by sonication and agarose gel electroelution. However, fewer steps, a shorter time period, and significantly less substrate is required for this newly developed method compared to conventional procedures. These protocols should be useful in small and large sequencing projects. Current sequencing paradigms require the generation of a new template for each 350–500 nucleotides sequenced. Sequencing both strands of the human genome would require at least 12 million templates 500 nucleotides long, assuming no overlap between templates. A random approach, such as shotgun sequencing, would require 30 to 50 million templates, assuming the entire genome were randomly subcloned. As many as 250,000 libraries may be needed to generate the requisite templates from a subcloned and ordered array of this genome, depending on the the type of vector utilized, and the degree of overlap between such clones. The ability to generate shotgun libraries in a semi-automated, microtiter plate format would greatly simplify such large scale projects.

The development of methods for cloning large DNA molecules in yeast artifical chromosomes (18) or in bacteriophage P1 derived vectors (19) simplifies the subdivision and analysis of very large genomes. However, the large size of the resulting subclones (100–1000 kbp) presents additional challenges for subsequent sequencing efforts. Recent experience gained from sequencing a 134 kbp genome by random shotgun cloning directly into a bacteriophage M13 vector eliminates numerous intermediate stages of subcloning, mapping, and overlapping such clones (6). An order of magnitude reduction in the amount of DNA required for shotgun cloning would substantially simplify efforts to directly sequence 100,000 bp sized molecules and beyond.

The ability to generate an overlapping population of randomly fragmented DNA molecules is considered essential for minimizing the closure of nucleotide sequence gaps by the shotgun cloning method. The use of a very frequent cutting restriction enzyme, such as *Cvi*JI, is an approach which has not

been utilized. Unusual reaction conditions were found which resulted in the quasi-random restriction of pUC19 and lambda DNA, as judged by the degree of smearing observed. The randomness of this *Cvi*JI** reaction was quantified by sequence analysis of 76 such partially fragmented pUC19 subclones. The analysis showed that *Cvi*JI** partial digestion is a hybrid reaction which combines the three base recognition specifity of *Cvi*JI with the two base recognition specifity of *Cvi*JI*. Interestingly, most of the 'relaxed' cleavage observed under *Cvi*JI** conditions (limiting enzyme and time) occurred in those portions of the sequence which were deficient in 'normal' restriction sites. Whatever the mechanistic reasons for this phenomenon, *Cvi*JI** treatment produces a relatively even size distribution of DNA fragments, permitting sequence information to be accumulated in a statistically random fashion.

Shotgun cloning with *Cvi*JI** digested DNA is efficient because the resulting fragments are blunt ended. Other methods currently used to randomly fragment DNA, including sonication, DNAse I treatment, and low pressure shearing, leave ragged ends which must be converted to blunt ends for efficient vector ligation. Other than a heat denaturation step to inactivate the endonuclease, no additional treatments are required for cloning *Cvi*JI** restricted DNA. In addition, the pre-ligation step required to equalize representation of the ends of a DNA molecule prior to sonication or DNAse I treatment should not be necessary with *Cvi*JI** fragmentation. *Cvi*JI* will cleave its cognate recognition site very close to the ends of a linear molecule, as judged by the very small fragments resulting from complete digestion of pUC19 (see Fig. 1).

The overall efficiency of shotgun cloning depends not only on the fragmentation process, but also upon the size fractionation procedure used to remove small DNA fragments. In this study, the efficiency of cloning agarose gel fractionated DNA was found to be unexpectedly variable. Numerous experiments produced an erratic distribution of sized material and the resulting cloned inserts were uniformly small (70% < 500 bp in one trial, 100% < 500 bp in another). Although this method could have been optimized, we chose instead to develop a simple and rapid micro-column fractionation method, which resulted in three to thirteen times more transformants than agarose gel fractionation. More importantly, the size distribution of the cloned inserts from column fractionated DNA was skewed toward larger fragments (88% > 500 bp). Micro-column fractionation also eliminates the chemical extraction steps required for agarose fractionated DNA. After the target DNA has been column fractionated, no further treatments are required for cloning. Combining *Cvi*JI** partial restriction with micro-column fractionation permits the construction of useful libraries from as little as 200 ng of substrate, an order of magnitude less material than recommended for sonication/end-repair and agarose gel fractionation procedures.

The *Cvi*JI** reaction represents a unique alternative for controlling the partial digestion of DNA, a technique which is fundamental to the construction of genomic libraries (20) and restriction site mapping of recombinant clones (21). Partial DNA digests are notably variable and are strongly dependent on the concentration and purity of the DNA, the amount of enzyme used, the incubation time, and the batch of enzyme. Partial digestions can also be variable with respect to the rate at which a particular recognition sequence is cleaved throughout the substrate. Optimal reaction conditions which render such partial digests independent of one or more of these variables enables a more precise control of the final end product. Several controlling schemes have been implemented, including: the addition of a constant amount of carrier DNA (22), the use of limiting amounts of Mg^{2+} (23), ultraviolet irradiation (24), and the combination of a restriction enzyme and a sequence complementary DNA methylase (25). Utilizing three different batches of *Cvi*JI, and three different DNA templates from five separate preparations, we obtained a uniform *Cvi*JI** partial digestion pattern that was primarily time dependent when a constant ratio of 0.3 units of enzyme per μg of DNA was used.

The rate at which a particular restriction site is cleaved at different locations in a substrate is variable for many endonucleases (14). We have optimized reaction conditions for *Cvi*JI which reduce the site preferences of this enzyme during partial digestion substantially (see Fig.1, lanes 3 and 4). Normally, 'star' reaction conditions result in cleavage at new sites. The use of star reaction conditions here (DMSO and lowered ionic strength) to affect the partial digestion activity of *Cvi*JI* does not result in an altered restriction site cleavage as assayed by sequencing 76 cases of this treatment. Instead, the relative rate of cleavage of individual sites appears to be more uniform under these conditions. We have observed a 3–5 fold increase in the rate of normal *Cvi*JI restriction with the standard buffer and DMSO, which further substantiates this idea. All of these results suggest that under the appropriate reaction conditions *Cvi*JI may prove useful for a number of other applications, such as high resolution restriction mapping and fingerprinting, diagnostic restriction of small PCR fragments, and construction of genomic DNA libraries.

ACKNOWLEDGEMENTS

We would like to thank Dwight Burbank for assistance in cell and virus culturing, Karolyn Morrison and Rose Geraci for technical assistance, and Joseph Heitman for directing our search for the 'two cutter' to *Cvi*JI.

REFERENCES

1. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) Molecular Cloning: A Laboratory Manual, Second Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
2. Messing,J. (1983) *Methods in Enzymol.*, **101**, 20–78.
3. Baer,R., Bankier,A.T., Biggin,M.D., Deininger,P.L., Farrell,P.J., Gibson,T.J., Hatfull,G., Hudson,G.S., Satchwell,S.C., Seguin,C., Tuffnell,P.S. and Barrell,B.G. (1984) *Nature*, **310**, 207–211.
4. Bankier,A.T., Weston,K.M. and Barrell,B.G. (1987) *Methods in Enzymol.*, **155**, 51–93.
5. Edwards,A., Voss,H., Rice,P., Civitello,A., Stegemann,J., Schwager,C., Zimmermann,J., Erfle,H., Caskey,C.T. and Ansorge,W. (1990) *Genomics*, **6**, 593–608.
6. Davison,A.J. (1991) *DNA Sequence-J. DNA Sequencing and Mapping*, **1**, 389–394.
7. Sanger,F., Nichlen,S. and Coulson,R.A. (1977) *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 5463–5467.
8. Anderson,S. (1981) *Nucleic Acids Res.*, **9**, 3015–3027.
9. Schriefer,L.A., Gebauer,B.K., Qiu,L.Q.Q., Waterston,R.H. and Wilson,R.K. (1990) *Nucleic Acids Res.*, **18**, 7455.
10. Deininger,P.L. (1983) *Anal. Biochem.*, **129**, 216–223.
11. Xia,Y., Burbank,D.E., Uher,L., Rabussay,D. and Van Etten,J.L. (1987) *Nucleic Acids Res.*, **15**, 6075–6090.
12. Peacock, A.C. and Dingman, C.W. (1967) *Biochemistry*, **6**, 1818–1827.
13. Uzgiris,J., Mead,D.A. and Smith,L.M. (unpublished).
14. Mead,D.A., McClary,J.A., Luckey,J.A., Kostichka,A.J., Witney,F.R. and Smith,L.M. (1991) *BioTechniques*, **11**, 76–87.

15. Brooks,J.E. (1987) *Methods in Enzymol.*, **152**, 113–129.
16. Barany,F. (1988) *Gene*, **65**, 149–165.
17. Lander,E.S. and Waterman,M.S. (1988) *Genomics*, **2**, 231–239.
18. Burke, D.T. Carle, G.F., and Olson, M.V. (1987) *Science* 236, 806–812.
19. Sternberg, N. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 103–107.
20. Maniatis,T., Hardison,R.C., Lacey,E., Lauer,J., O'Connell,C., Quon,D., Sim,G.K. and Efstratiadis,A. (1978) *Cell*, **15**, 687–701.
21. Smith, H.O. and Birnstiel, M.L. 1976. *Nucleic Acids Res.*, **3**, 2387–2398
22. Kohara, Y., Akiyama, A. and Isono, K. (1987) *Cell* , **50**, 495–508.
23. Albertson,H.M., Paslier,D.L., Abderrahim,H., Dausset,J., Cann,H. and Cohen,D. (1989) *Nucleic Acids Res.*, **17**, 808.
24. Whittaker,P.A. and Southern,E.A. (1986) *Gene*, **41**, 129–134.
25. Hoheisel,J.D., Nizetic,D. and Lehrach,H. (1989) *Nucleic Acids Res.*, **17**, 9571–9582.