Published online January 29, 2004

Nucleic Acids Research, 2004, Vol. 32, No. 2 ***e21***
DOI: 10.1093/nar/gnh019

Selective DNA amplification from complex genomes using universal double-sided adapters

Matthew J. Callow, Snezana Drmanac and Radoje Drmanac*

Callida Genomics Inc., 675 Almanor Avenue, Sunnyvale, CA 94085, USA

Received October 16, 2003; Revised and Accepted December 11, 2003

ABSTRACT

There is a rapidly developing need for new technologies to amplify millions of different targets from genomic DNA for high throughput genotyping and population gene-sequencing from diverse species. Here we describe a novel approach for the specific selection and amplification of genomic DNA fragments of interest that eliminates the need for costly and time consuming synthesis and testing of potentially millions of amplicon-specific primers. This technique relies upon Type IIs restriction enzyme digestion of genomic DNA and ligation of the fragments to double-sided adapters to form closed-circular DNA molecules. The novel use of double-sided adapters, assembled through the combinatorial use of two small universal sets of oligonucleotide building blocks, provides greater selection capacity by utilizing both sides of the adapter in a sequence-specific ligation event. As demonstrated, formation of circular structures results in protection of the desired molecules from nuclease treatment and enables a level of selectivity high enough to isolate single, or multiple, predefined fragments from the human genome when digested at over five million sites. Priming sites incorporated into the adapter allows the utilization of a common pair of primers for the amplification of any adapter-captured DNA fragment of interest.

INTRODUCTION

The quest to isolate specific DNA fragments from complex genomes has centered around two major approaches. In the first, genomic fragments are captured during library construction with subsequent clone isolation using sequence-specific probes. In the second, sequence-specific oligonucleotide primers are used in a DNA amplification reaction. Existing approaches for selective amplification from genomic DNA have, therefore, often relied upon the custom generation of reagents (primers and/or probes) specific for each polymorphism, which can be costly and time consuming. In this study we describe an approach for isolating single, or multiple, targeted genomic DNA fragments generated by restriction endonucleases that cut outside of their recognition sequence (Type IIs) (1). Sequence specificity is achieved by assembling two oligonucleotides into double-sided adapters for a ligation mediated selection process.

The use of Type IIs restriction endonucleases to fragment DNA, and the capture of those fragments by ligation to adapters, has been described in the past (2,3). These studies also relied upon the incorporation of primer binding sites in the adapters to amplify the captured sequences over other sequences by standard PCR methodologies. Type II enzymes have also been applied to this goal (4); however, we are unaware of any studies that have demonstrated the isolation of specific, pre-defined fragments from the human genome when digested into millions of fragments. Sibson *et al.* (5) also recently described an approach for sorting or indexing of Type IIs restriction fragments using successive rounds of digestion and ligation to adapters with two overhanging bases. This was achieved by incorporating Type IIs restriction sites into the adapter to cut into the adjacent genomic sequence. In all, six bases were used to achieve a maximal possible enrichment of 1920-fold. The incorporation of Type IIs recognition sequences into adapters for cutting into adjacent sequences (6) has also been applied in the technique of Serial Analysis of Gene Expression (SAGE) (7).

There are several methodologies available that can reduce the complexity of genomes including AFLP (8), Alu-PCR (9,10), DOP-PCR (11–13) and PCR amplicon size selection (14,15). These techniques are limited however, because they typically select from only a portion of the genome or they co-isolate many undesired fragments. In the methodology we present here, the entire genome is accessible and single, specific DNA fragments can be isolated rapidly. Further rounds of isolation with alternative techniques may enable the selection of single fragments from complex genomes but this adds considerably to the complexity of the methods and multiplexing cannot typically be performed without the associated isolation of many unwanted fragments. Alternatively, the user is limited in some way to utilize only sequences that fall within regions of the genome that are specifically targeted by the technology such as specific restriction enzyme fragments or size ranges. DOP-PCR and Alu-PCR can amplify random regions of the genome and therefore reduce sequence complexity but they have yet to be able to isolate specific sequences from complex genomes

*To whom correspondence should be addressed. Tel: +1 408 746 4525; Fax: +1 408 746 4596; Email: rade@sbh.com
Present address:
Matthew J. Callow, Perlegen Sciences, 2021 Stierlin Ct, Mountain View, CA 94043, USA

Nucleic Acids Research, Vol. 32 No. 2 © Oxford University Press 2004; all rights reserved

without unwanted fragments and a universal oligonucleotide set. The methodology we report here is a rapid alternative to these methods for the reduction of genome complexity and for the single or multiplexed isolation of specific DNA fragments.

Our method, termed Universal Selective Amplification, utilizes double-sided adapters to induce selective closed-circle formation with target DNA, and exonuclease treatment to enhance selection and specificity, especially critical for multiplex amplification. Our approach for Universal Selective Amplification is designed for the combinatorial use of a relatively small and finite number of universal adapter-building oligonucleotides for specific fragment selection from millions of short genomic fragments. Furthermore, by utilizing common PCR primers, this technique will be more amenable to applications requiring multiplex amplification of DNA sequences.

There are two broad application areas for this methodology. The first area utilizes and allows for the selection of multiple fragments that share a particular fragment overhang-type, such as for sequence complexity reduction in which a specific fragment is enriched. The second area utilizes and requires a higher level of selection for isolation of single, specific fragments or multiple specific fragments in a multiplexed selection process. Applications of the methodology include detection and isolation of specific exons, genotyping of SNPs in a low sequence-complexity sample, and isolation of specific genes or regulatory elements. We present the successful amplification of three *Escherichia coli* and four human amplicons demonstrating the feasibility of the method.

MATERIALS AND METHODS

Oligonucleotides

Oligonucleotides were purchased from Qiagen-Operon (Alameda, CA) and were HPLC purified for lengths >40 bases and 2-step HPLC purified for lengths >60 bases. All other oligonucleotides were supplied de-salted. In some cases oligonucleotides were ordered phosphorylated from the manufacturer or were phosphorylated with T4 polynucleotide kinase (PNK) [New England Biolabs (NEB), Beverly, MA] by incubating at 37°C for 30 min and then 65°C for 20 min. Reaction conditions were 10 U of PNK, 1 mM ATP, 1× PNK buffer (NEB) and 250 pmol of 5′ termini in a 20 µl reaction.

Genomic DNA digestion

Escherichia coli genomic DNA was sourced from the MG1655 strain supplied from the American Type Culture Collection (Manassas, VA). DNA was isolated by lysis of bacterial cells in an SDS–proteinase K solution followed by RNAse treatment and multiple phenol/chloroform extractions before ethanol precipitation (16). Human genomic DNA was purchased from Promega (Madison, WI). The DNA (10 µg) was digested with 8 U of the restriction enzyme BbvI (NEB) for 3 h at 37°C before heat inactivation at 65°C for 20 min.

Adapter preparation

Adapters were prepared either by annealing two complete single-stranded oligonucleotides to produce 4-base, 5′ overhangs at each end or by annealing two complementary oligonucleotides that, when double stranded, produce an adapter core with 14 and 17 base 3′ overhangs. Two shorter, variable oligonucleotides were then ligated to the core with T4 DNA ligase (20 U/µl) (NEB) at 25°C to produce the 4-base 5′ overhangs. The final concentration of variable oligonucleotides to core adapter was 8 µM each to 4 µM, respectively. The final adapter was then phosphorylated and purified using a Qiaquick spin protocol (Qiagen, Valencia, CA). The DNA was collected in 40 µl of 10 mM Tris/0.1 mM EDTA pH 8 (TE).

Adapter formation test

Core adapter, adapter after ligation of overhang oligonucleotides and adapter after final phosphorylation was treated with 2.5 U of Lambda Exonuclease (Roche, Indianapolis, IN) in a reaction buffer of 67 mM glycine–KOH, pH 9.4, 2.5 mM $MgCl_2$ and 50 µg/ml BSA in 10 µl. The reaction was incubated at 37°C for 10 min and then at 75°C for 10 min.

First-round selection

Adapter (20 fmol) was ligated to 150 ng of BbvI-digested genomic DNA in a volume of 10 µl for 30 min at 25°C in the presence of 1× T4 DNA ligase buffer (NEB) and 200 U of T4 DNA ligase (NEB). The enzyme was heat inactivated at 65°C for 10 min. The ligation reaction was then treated with 1 U of Bal31 nuclease (NEB) in the presence of 1× Bal31 nuclease buffer (NEB) for 30 min at 30°C and then heat inactivated at 75°C for 10 min. The sample was diluted 10-fold with TE and 8 µl were digested with 2.5 U of NotI enzyme (NEB) in a 10 µl volume before 1 µl was used in a 30 µl PCR. PCR from Adapter A was carried out at 94°C for 3 min, denaturation at 94°C for 20 min, annealing at 66 to 62°C for five cycles and 62°C for the following 35 cycles for 30 s. Extension was at 72°C for 30 s. Reaction conditions were 1× Thermopol buffer (NEB), 200 µM dNTP, 0.2 µM primers (5′-TGAGACCACAGCCTAGACAGC and 5′-CTGCAAGGCGATTAAGTTGG) and 0.6 U/100 µl Vent exo-DNA polymerase (NEB) or 1.8 U/100 µl *Taq* (Qiagen) in 10 mM Tris–HCl pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% gelatin. If two rounds of selection were to be performed the NotI digest was omitted and the combined volume of the PCRs was 200 µl.

Second-round selection

The first round PCR (200 µl) was purified using the Qiaquick protocol (Qiagen) and isolated in 40 µl of TE. Digestion of the adapter was performed in a 10 µl reaction with 8 µl of purified PCR, 1× NEB buffer 2 and 4 U of FokI enzyme (NEB). Digestion was for 60 min at 37°C followed by inactivation at 65°C for 20 min. The digest (8 µl) was then combined with 10× T4 DNA ligase buffer (1 µl) adapter B (20 fmol, 1 µl) and 200 U of T4 DNA ligase. After incubation for 30 min at 25°C and inactivation at 65°C for 10 min the reaction was digested with 1 U of Bal31 nuclease in 1× Bal buffer (NEB). The sample was diluted 10-fold with TE and 8 µl were digested with 2.5 U of NotI enzyme (NEB) in a 10 µl volume before 1 µl was used in a 30 µl PCR at 94°C for 3 min, 94°C for 20 s, 58°C for 30 s, 72°C for 30 s over 35 cycles. Reaction conditions were 1× Thermopol buffer (NEB), 200 µM dNTP, 0.2 µM adapter B primers (5′-GACGGCTGAAATTGGTAAGG and 5′-CGGAATCAAAGCAGGATAAGG) and 0.6 U/100 µl Vent exo-DNA polymerase (NEB).

RESULTS

Universal Selective Amplification process

We selected Type IIs restriction enzymes that produce 4-base, 5′ overhangs of digested genomic DNA to fragment the genome into an estimated 32 768 (non-directional) variants defined by the sequence of the overhangs. A double-sided adapter preparation was used to capture specific fragments into closed-circular molecules that protect the fragments from subsequent exonuclease digestion (Fig. 1). For isolation of single fragments from complex mammalian genomes, a second round of selection was used by incorporating Type IIs recognition sites into the adapter that cut further into the genomic sequence of the captured fragment. The released fragment was re-captured with a new adapter and amplified by PCR utilizing common primer binding sites in the adapter. This two-step process predicts a theoretical 10^9-fold enrichment of the desired fragment resulting in a specific, predefined fragment being isolated from human genomic DNA digested to 100 million fragments (one cut every 30 bases).

Adapter assembly

Although adapters could be prepared from as few as two sets of 256 oligonucleotides that can form all possible 32 768 complementary adapters, for this study adapters were prepared using a common phosphorylated core sequence with two different (one on each side) 14 to 17 base 3′ overhangs. Two shorter, variable, 5′ overhang-generating oligonucleotides, each matching to one 14–17 base overhang were then ligated to the core DNA and the final product was phosphorylated. To confirm that adapter assembly was occurring efficiently at each stage, we developed a nuclease-sensitivity based assay by utilizing lambda exonuclease to degrade one strand of a double stranded structure from the 5′ phosphorylated end. Non-phosphorylated ends are resistant to lambda exonuclease degradation. Core adapter sequences were degraded by exonuclease treatment, however after ligation of the 5′ overhang oligonucleotides, most of the adapter was protected from degradation suggesting efficient ligation of the overhang oligonucleotides was occurring (Fig. 2). Final phosphorylation of the adapter resulted in degradation of the adapter after lambda exonuclease treatment demonstrating final phosphorylation of the adapter was also occurring efficiently.

Escherichia coli genomic fragment isolation

The isolation of specific fragments from a complex mixture of fragments was first tested on the 4.6 Mb *E.coli* genome which, when digested with BbvI, produces an estimated 18 000 DNA fragments with variable 4-base, 5′ overhangs. Three fragments were selected of 100, 150 and 200 bp in size from three random regions of the published genome and adapters were designed and prepared for ligation with the digested genomic DNA (Table 1). During the ligation process undesirable events are also likely to be occurring such as inter- and intra-fragment ligation, ligation of an adapter to just one end of DNA fragments and free adapter remaining un-ligated. To eliminate these structures the ligation mix was treated with Bal31 nuclease to destroy all linear molecules but preserve circular molecules. PCR using primer binding sites in the adapter and DNA polymerase was then used to amplify all fragments

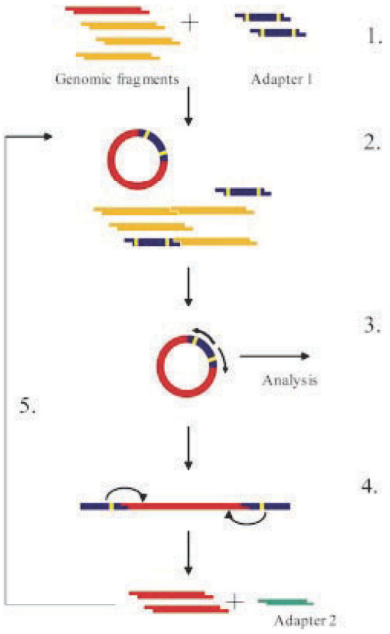

Figure 1. Schematic representation of Universal Selective Amplification from the human genome. (1) Digestion of DNA with Type IIs restriction enzyme and ligation to adapter. The desired fragment (red) has overhangs complementary to the adapter. (2) Multiple structures are formed in a ligation reaction, but the desired fragment is captured into a closed circular molecule. (3) Digestion of linear molecules and mismatch structures with exonuclease preserves and enriches the circular molecules for PCR amplification. (4) Digestion with a second Type IIs restriction enzyme (recognition sites in yellow) releases the genomic insert and presents new 4-base overhangs. (5) A new adapter complementary to the new overhangs is introduced and the process (ligation, exonuclease digestion and PCR) is repeated. For selective amplification from the *E.coli* genome only the first three steps are used.

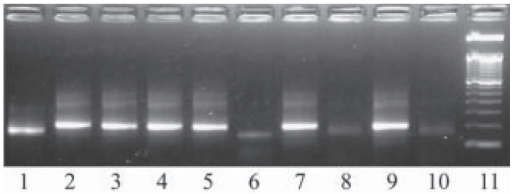

Figure 2. Gel analysis of the effects of lambda exonuclease treatment on adapter assembly. Lanes 1 and 6, phosphorylated core adapter without overhang oligonucleotides; lanes 2 and 7, Adapter 1 without phosphorylation; lanes 3 and 8, Adapter 1 with phosphorylation; lanes 4 and 9, Adapter 2 without phosphorylation; lanes 5 and 10, Adapter 2 with phosphorylation; lanes 1–5, without lambda exonuclease treatment; lanes 6–10, with lambda exonuclease treatment; lane 11, 50 bp DNA marker (Promega). DNA (1.25 pmol) was loaded onto 3% agarose gels and stained with Gelstar stain (Cambrex).

captured in closed circles. This was performed either directly from the circular target or on a target digested with NotI, to linearize the circular molecule. All three selected fragments were successfully amplified from *E.coli* genomic DNA when single adapters were included in the ligation (Fig. 3). Each adapter selects for 1 of 32 768 fragment variants and so with 18 000 fragments generated from the *E.coli* genome the

Table 1. Enzyme recognition and cleavage sites for genomic fragments selected from the *E.coli* and human genomic sequences

Genomic fragment	Location	Left flanking sequence	Right flanking sequence
E.coli 100 bp	4229991–4230091	GACCGGGATATC**GCTGC**...	...TCAATGCGTTTT**GCTGC**
E.coli 150 bp	1976452–1976602	CAACGGAGGGGG**GCTGC**...	...**GCAGC**TCCACCGATTTT
E.coli 200 bp	4224653–4224853	**GCAGC**ACGATGATTACA...	...CACCATCTGGGA**GCTGC**
Human 125 bp	Chr19: 50040675–50040800	**GCAGC**ACTATCCACAACAGCA...	...**GCAGC**AAAAGCAAACAT
Human 262 bp	Chr19: 50152102–50152364	**GCAGC**CCTGTGGTCACCAGGA...	...GGGGACTCACCACCTT**GCTGC**
Human 318 bp	Chr19: 51678472–51678790	**GCAGC**CCACACAGGGCCATTGG...	...**GCAGC**ACACCACCACGC
Human 499 bp	Chr19: 51876594–51877093	CTATCCCGTGGA**GCTGC**...	...**GCAGC**GTGGTGCTGGAT

The BbvI recognition site is highlighted in bold and the 4-base 5′ overhang sequence generated is underlined. The position of the second 4-base 5′ overhang generated by digestion from the adapter incorporated recognition site is double underlined. The orientation of the cut site relative to the recognition site appears to alternate depending upon which strand contains the recognition sequence GCAGC. Genomic locations are indicated for the position of the cut sites. *Escherichia coli* genomic sequence was obtained from NC_000913.

chance of selecting an additional fragment to the one targeted is 0.56. Multiplexing of the adapters in which all three were combined into the one ligation also demonstrated the successful amplification of the three fragments. Further optimization of adapter concentration and ligation conditions may help to minimize non-specific bands.

Human genomic fragment isolation

To amplify four specific fragments from human genomic DNA of 125, 262, 318 and 499 bp in size (Table 1), two rounds of selection were used. The first round of capture was performed on BbvI-digested human genomic DNA using the four adapters either separately or in a pool of four adapters (Fig. 4A). If it is assumed the genome is fragmented into 6 million fragments as expected statistically for the BbvI digest, then an adapter that is 1 of 32 768 variants will select about 188 unique fragments. However, other structures are likely to form such as two or more genomic fragments ligated and captured into the adapter circle—but their frequency will be lower. Figure 4A shows that after amplification, DNA of a range of sizes was produced but on average the distribution is centered around 500 bp. This distribution pattern is a result of the frequency of cutting and also the polymerase extension time.

A single round of selection would provide an ~32 000-fold purification of genomic fragments and could provide a means for partial or whole genome amplification; however, the amplification of specific fragments from the human genome would require two rounds of selection. After the first round PCR with primer set A, the products were digested with FokI enzyme which recognizes a 5-base sequence incorporated in the adapter but cuts into the captured genomic DNA. The released DNA was re-ligated with an adapter specific for the new overhangs generated in the genomic region of the first-capture DNA. After ligation, the DNA was again digested with Bal31 nuclease before amplification with alternate PCR primers to those used in the first round. When a single adapter was used in each ligation for both the first round and the second rounds of selection, a single band was produced for the 125, 262, 318 and 499 bp products (Fig. 4B). When all four adapters were combined for the first step ligation and a second set of four were combined for the second step ligation all fragments were detected on the gel, except for the 499 bp fragment. We believe the failure to visualize the 499 bp fragment in the multiplex reaction may be due to sequence features of the fragment that are leading to a competitive disadvantage during the PCR phase. The addition of high GC favorable components such as DMSO in addition to altering primer, magnesium and enzyme concentrations also failed to promote the appearance of the band.

DISCUSSION

We have described a technique for the specific amplification of DNA fragments from a complex genome using a small, universal set of adapter-forming oligonucleotides. In the simplest construction of adapters, one of 256 variants for one strand and one of 256 variants for the other strand could be annealed, prior to ligation of the adapter to the genomic fragment. Alternatively, a pre-ligation of short overhang oligonucleotides to a common core adapter could be performed to generate the adapter as we have demonstrated here. There are 256 possible combinations of the bases at each end of each adapter resulting in 65 536 possible adapters (functionally this equates to 32 768 unique adapters because of directional reversibility of the adapter). In this case a total of 512 adapter oligonucleotides are all that is required to produce all 32 768 possible variations of adapter overhangs. This effectively results in an adapter which, when ligated with a fragmented genome of 6 million pieces, could capture in the order of 188 unique fragments in a single round of selection. The actual number and average size of genomic fragments may, however, be controlled by performing digests with multiple enzymes.

Many of the alternative techniques for sequence complexity reduction currently available have found usefulness with small bacterial genomes because of their inherent lower complexity. For fragment selection on this scale our approach would be as simple, and more rapid than most techniques currently in use. In our method, selection of a fragment from a bacterial genome would only require a genomic digestion, ligation, a short nuclease digestion and one PCR to achieve the isolation of any fragment from the genome. Achieving single fragment isolation with other techniques is typically much more difficult because they are unable to achieve such a degree of selection in as few steps.

Two major applications of this technology would be for high-throughput production of specific DNA amplicons for genotyping and mutation/polymorphism discovery. Traditionally, this would require PCR with locus-specific primers and a large investment in the design, synthesis and testing of large number of primers. With Universal Selective

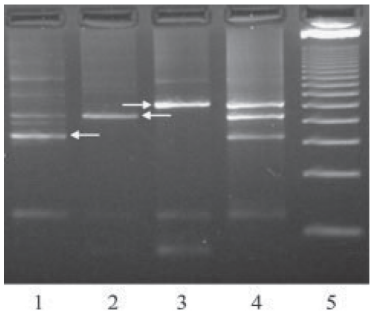

Figure 3. Isolation and amplification of three fragments from the *E.coli* genome. Adapters with common primer binding sites (primer set A) were prepared with 5′ overhangs specific for fragments of 100, 150 and 200 bp in size from the *E.coli* genome and ligated to BbvI-digested genomic DNA. The ligations were treated with Bal31 exonuclease before NotI digestion and PCR amplification using primer set A and *Taq* polymerase as described in the Materials and Methods. Lane 1, 100 bp; lane 2, 150 bp; lane 3, 200 bp; lane 4, mix of 100, 150 and 200 bp adapters in the ligation; lane 5, 50 bp marker (Promega). PCRs (2 µl) were electrophoresed on 4% Nusieve agarose (FMC) and visualized with Gelstar stain. Expected bands (arrowed) run at higher molecular weights than the genomic fragments due to primer and adapter sequences that are co-amplified.

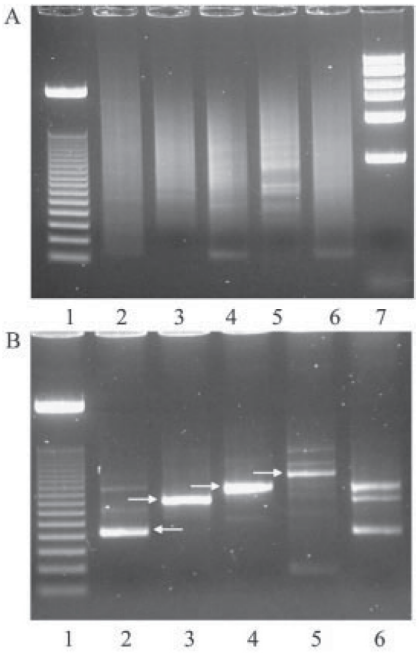

Figure 4. (**A**) First-round amplification of four fragments from the human genome. Adapters with common primer binding sites (primer set A) were prepared with 5′ overhangs specific for fragments of 125, 262, 318 and 499 bp in size from the human genome and ligated to BbvI-digested human genomic DNA. The ligations were treated with Bal31 exonuclease before PCR amplification using Vent polymerase and primer set A as described in the Materials and Methods. Lane 1, 50 bp marker (Promega); lane 2, 125 bp adapter; lane 3, 262 bp adapter; lane 4, 318 bp adapter; lane 5, 499 bp adapter; lane 6, mix of 125, 262, 318 and 499 bp adapters in the ligation; lane 7, 1 kb marker (NEB). Qiaquick (Qiagen) purified PCRs (3 µl) were electrophoresed on 3% agarose and visualized with Gelstar stain (Cambrex). (**B**) Second-round amplification of four fragments from the human genome. Adapters with common primer binding sites (primer set B) were prepared with 5′ overhangs specific for the second overhangs generated from fragments of 125, 262, 318 and 499 bp in size from the human genome and ligated to the FokI digested DNA. The ligations were treated with Bal31 exonuclease before PCR amplification using Vent polymerase and primer set B as described in the Materials and Methods. Lane 1, 50 bp marker (Promega); lane 2, 125 bp; lane 3, 262 bp; lane 4, 318 bp; lane 5, 499 bp; lane 6, mix of 125, 262, 318 and 499 bp adapters in the ligation, PCRs (2 µl) were electrophoresed on 3% agarose and visualized with Gelstar stain. Expected bands (arrowed) run at higher molecular weights than the genomic fragments due to primer and adapter sequences that are co-amplified.

Amplification, as we have described here, even high throughput, single-plex reactions would be more feasible. PCR conditions can be standardized through the use of common primers to allow parallel amplification of specific loci, without the need for designing primers and optimizing conditions for each locus. Furthermore, greater cost savings could be realized by assembling all necessary adapters from a universal set of less than 1000 oligonucleotides. The ability to perform multiplex isolations of fragments, by combining many pre-annealed adapter oligonucleotides in the one reaction, is also possible because a common primer set can be employed to amplify the captured DNA. Double-sided adapters would be critical for isolating tens of DNA fragments in one reaction because a mixture of all required pairs of earlier proposed single-sided adapters would produce an exponential number of undesired adapter combinations that would capture an excess of unwanted DNA. Due to removal of linear non-targeted DNA our process may also have advantages in producing highly defined low complexity genomic fractions comprised of hundreds or thousands of DNA fragments scattered over the genome by using one, multiple or degenerate adapters.

The use of a single primer pair for the amplification of multiple fragments in the one reaction, could however, introduce other issues that impact upon PCR-based amplification. Since each amplicon has the same primer binding site, competition between newly formed product and primers, as in any PCR, may lead to reduced yields of any one fragment in later cycles. In addition, if the rate of full-length product formation is dramatically different for some amplicons this may give a competitive advantage to more efficient amplicons that out-compete the less efficient amplicons over time. These factors may underlie our observations for one fragment; a GC-rich amplicon of 499 bp that appeared to be suppressed by the amplification of other amplicons with the same primer binding sites in a multiplex reaction. Other possibilities for this suppression may include differences in target concentration (post-ligation) as a result of incomplete digestion or differences in ligation efficiencies. To overcome some of these issues for large multiplexing experiments, alternatives such as isothermal amplification (17) or the selection of shorter fragments that are less influenced by sequence factors may prove more effective. Isothermal amplification in particular should result in less influence from primer competition for the annealing target as occurs during each cycle of the PCR. Although the 499 bp fragment was not visualized on the gel it may be present below the detection limit of the stain. Complete suppression would appear to be unlikely since during the first round of selection with a single adapter several hundred unique sequences were amplified and from these, the 499 bp fragment could be isolated on the second round of selection, suggesting the first round of selection was amplifying the 499 bp fragment within a multiplex of hundreds of fragments.

Although we have chosen exonuclease digestion to remove unwanted fragments and adapter sequences in this study, other

options are feasible such as a biotinylated, fully degenerated blocking adapter that ligates to unused overhang sequences, which can then be removed by streptavidin coated beads. Used in an effective concentration, the blocking adapter can also reduce formation of undesired circles comprised of multiple non-targeted DNA fragments with or without an adapter. The Universal Selective Amplification procedure could also be further simplified by possibly eliminating PCR after ligating the first adapter. In addition, cutting circular DNA and ligation of a second adapter may be performed in the same reaction if two different adapter cores are used. The first adapter would use a core that has no priming sites but has two Type IIs restriction enzyme recognition sequences, and the second adapter would use a core that has priming sites but has no restriction enzyme recognition sequences.

There are at least three commercially available Type IIs enzymes that produce 4-base overhangs, and are not sensitive to CpG methylation (BbvI, FokI and BspMI). There are also a number of alternative Type IIs enzymes available should the user choose not to use a 4-base overhang. Another alternative for increasing the number of enzymes available (by avoiding CpG methylation issues) is to treat the genomic DNA with one of a number of whole genome amplification methods available to generate non-methylated DNA. An additional six enzymes that produce 4-base overhangs would be available in this case.

Our adapter assembly process, based on universal complete libraries of oligonucleotide building blocks, allows other advanced features to be engineered into the adapters such as internal informative cleavage sites that expose an overhang complementary to an overhang from the genomic sequence after cleavage of the formed circles. Having two additional pairs of 256-oligonucleotide libraries allows one to make adapters with 16 informative bases in one assembly reaction that would replace pairs of standard double-sided adapters designed for two rounds of selection of a DNA fragment. In this case two specific internal 4mers would be incorporated, in addition to the two specific 4-base overhangs. In other envisioned improvements, by designing several adapter cores with different assembly overhangs, coupled with corresponding libraries of building blocks, multiple double-sided adapters could be assembled simultaneously. This could occur during adapter ligation to genomic DNA to produce a multiplexed DNA amplification without prior pre-assembly of each individual adapter in its own separate assembly reaction.

In conclusion, we have described a strategy for the amplification of specific genomic DNA fragments that utilizes a small, but universal set of oligonucleotides. The novel features of this approach include the ability to achieve greater selection specificity through the use of double-sided adapters and DNA circle formation, and to such a degree that amplification of selected fragments from the human genome is possible. Due to high selectivity, any specific sequence can be isolated using millions of fragments obtained by mixing independent Type IIs restriction enzyme digests prior to adapter ligation or by performing multiple digests on the same DNA. The adapter ligation technique is also advantageous in efficient amplification of various DNA sequences because it uses only universal primers under the same optimized conditions, as indicated by the successful preparation of all seven selected amplicons on the first attempt. The procedure we have described, although requires a few preparatory reactions before the intended DNA amplification, is one that is cost-effective by utilizing a relatively small set of oligonucleotides and one that can be easily standardized to allow rapid and parallel amplification of required genomic DNA fragments for high through-put genomic analysis in various species.

ACKNOWLEDGEMENTS

This work was funded by The National Institute of Standards and Technology, Advanced Technology Program Project ID: 200-00-4467A

REFERENCES

1. Roberts,R.J., Vincze,T., Posfai,J. and Macelis,D. (2003) REBASE: restriction enzymes and methyltransferases. *Nucleic Acids Res.*, **31**, 418–420.
2. Smith,D.R. (1992) Ligation-mediated PCR of restriction fragments from large DNA molecules. *PCR Methods Appl.*, **2**, 21–27.
3. Unrau,P. and Deugau,K.V. (1994) Non-cloning amplification of specific DNA fragments from whole genomic DNA digests using DNA 'indexers'. *Gene*, **145**, 163–169.
4. Guilfoyle,R.A., Leeck,C.L., Kroening,K.D., Smith,L.M. and Guo,Z. (1997) Ligation-mediated PCR amplification of specific fragments from a class-II restriction endonuclease total digest. *Nucleic Acids Res.*, **25**, 1854–1858.
5. Sibson,D.R. and Gibbs,F.E. (2001) Molecular indexing of human genomic DNA. *Nucleic Acids Res.*, **29**, e95.
6. Kim,S.C., Podhajska,A.J. and Szybalski,W. (1988) Cleaving DNA at any predetermined site with adapter-primers and class-IIS restriction enzymes. *Science*, **240**, 504–506.
7. Velculescu,V.E., Zhang,L., Vogelstein,B. and Kinzler,K.W. (1995) Serial analysis of gene expression. *Science*, **270**, 484–487.
8. Vos,P., Hogers,R., Bleeker,M., Reijans,M., van de Lee,T., Hornes,M., Frijters,A., Pot,J., Peleman,J., Kuiper,M. *et al.* (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.*, **23**, 4407–4414.
9. Nelson,D.L., Ledbetter,S.A., Corbo,L., Victoria,M.F., Ramirez-Solis,R., Webster,T.D., Ledbetter,D.H. and Caskey,C.T. (1989) Alu polymerase chain reaction: a method for rapid isolation of human-specific sequences from complex DNA sources. *Proc. Natl Acad. Sci. USA*, **86**, 6686–6690.
10. Ledbetter,S.A., Nelson,D.L., Warren,S.T. and Ledbetter,D.H. (1990) Rapid isolation of DNA probes within specific chromosome regions by interspersed repetitive sequence polymerase chain reaction. *Genomics*, **6**, 475–481.
11. Telenius,H., Carter,N.P., Bebb,C.E., Nordenskjold,M., Ponder,B.A. and Tunnacliffe,A. (1992) Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer. *Genomics*, **13**, 718–725.
12. Cheung,V.G. and Nelson,S.F. (1996) Whole genome amplification using a degenerate oligonucleotide primer allows hundreds of genotypes to be performed on less than one nanogram of genomic DNA. *Proc. Natl Acad. Sci. USA*, **93**, 14676–14679.
13. Jordan,B., Charest,A., Dowd,J.F., Blumenstiel,J.P., Yeh Rf,R.F., Osman,A., Housman,D.E. and Landers,J.E. (2002) Genome complexity reduction for SNP genotyping analysis. *Proc. Natl Acad. Sci. USA*, **99**, 2942–2947.
14. Dong,S., Wang,E., Hsie,L., Cao,Y., Chen,X. and Gingeras,T.R. (2001) Flexible use of high-density oligonucleotide arrays for single-nucleotide polymorphism discovery and validation. *Genome Res.*, **11**, 1418–1424.
15. Kennedy,G.C., Matsuzaki,H., Dong,S., Liu,W.M., Huang,J., Liu,G., Su,X., Cao,M., Chen,W., Zhang,J. *et al.* (2003) Large-scale genotyping of complex DNA. *Nat. Biotechnol.*, **21**, 1233–1237.
16. Russell,D.W. and Sambrook,J. (2001) *Molecular Cloning: A Laboratory Manual*. 3rd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
17. Dean,F.B., Nelson,J.R., Giesler,T.L. and Lasken,R.S. (2001) Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Res.*, **11**, 1095–1099.