Benchmarks

Dual primer emulsion PCR for next-generation DNA sequencing

Ming Yan Xu¹, Anthony D. Aragon¹, Monica R. Mascarenas¹, Norah Torrez-Martinez¹, and Jeremy S. Edwards^{1,2,3}

¹*Molecular Genetics and Microbiology, University of New Mexico Health Sciences Center, Albuquerque, NM, USA,* ²*Chemical and Nuclear Engineering, University of New Mexico, Albuquerque, NM, USA, and* ³*Cancer Research and Treatment Center, University of New Mexico Health Sciences Center, Albuquerque, NM, USA*

BioTechniques 48:409–412 (May 2010) doi 10.2144/000113423

Keywords: next-generation sequencing; emulsion PCR; bridge PCR

Supplementary material for this article is available at www.BioTechniques.com/article/113423.

We have developed a highly sensitive single-molecule clonal amplification method called dual primer emulsion PCR (DPePCR) for next-generation DNA sequencing. The approach is similar in concept to standard emulsion PCR; however, in DPePCR both primers are attached to the beads, therefore following PCR amplification, both strands of the PCR products are attached to the beads. The ends of each strand can be freed for analysis by restriction digestion of the bridged PCR fragments, which allows efficient paired-end sequencing of fragment libraries.

The concept of studying single molecules in high throughput using emulsions is not new and has been used for single-molecule evaluation for a number of years (1) before the approach was first applied to next-generation sequencing (2). However, for emulsion technology to be applicable to next-generation sequencing, a method for capturing the contents from the droplets is required. To accomplish this, Dressman et al. (3) described an approach to amplify single DNA molecules onto beads for detection and enumeration of genetic variation. They named their technology BEAMing, for “beads, emulsion, amplification, and magnetics.” This approach has now evolved to emulsion PCR (ePCR). The DNA bound to beads generated from ePCR provides an excellent template for high-throughput sequencing, because PCR can amplify a single molecule of DNA into many clonal molecules per bead (2). Due to that fact, a number of next-generation sequencing approaches utilize emulsions and beads for DNA amplification prior to sequencing (4–8).

An alternative approach to amplify DNA for next-generation sequencing is

the bridge amplification strategy, used by Illumina and originally reported by Bing et al. (www.promega.com/geneticidproc/ussymp7proc/0726.html). Bridge amplification is a technology that uses a single aqueous compartment; however, the individual amplicons are constrained by primers bound to a solid phase that are extended and amplified (9). As the name implies, the extension product from one primer forms a bridge to the other primer.

In this manuscript, we describe a novel approach called dual primer emulsion PCR (DPePCR), which combines concepts from both emulsion PCR and bridge amplification for the generation of simple fragment libraries for paired-end next-generation sequencing. The DPePCR strategy can amplify short DNA fragments (less than ~300 bp, including genome fragment and primers) and enables sequencing of both ends of a DNA fragment. This effectively shortens library preparation time, increases the library complexity (2) when compared with the construction of a mate-paired DNA library, and will contribute to the \$1,000-genome goal.

Normal ePCR has been extensively applied to next-generation DNA sequencing (10). Most of the next-generation sequencing approaches are restricted to short read lengths (5), and therefore to optimize resequencing of a human genome, mate-paired or paired-end sequencing is important. However, the construction of mate-paired libraries for next-generation sequencing is difficult and time-consuming (10). Fortunately, fragment library construction is quite simple, and the described DPePCR strategy (Figure 1) enables paired-end sequencing of fragment libraries with essentially no modification of existing sequencing approaches.

To perform DPePCR, both forward and reverse primers are attached to 1- μ m beads (Figure 1B) (see Supplementary Materials) that are included in a modified ePCR protocol (10). Additionally, since the amplicons are confined to the droplets, the amplification efficiency is increased by including free primers in the aqueous phase (Figure 1C). After 120 PCR cycles, a single DNA fragment in an emulsion drop can be amplified effectively. After amplification, we have found the DNA to be highly stable in the double-stranded state. Denaturing conditions will denature the DNA; however, since both strands are present on the bead, the double-stranded state immediately reforms, which inhibits the ability to sequence the DNA. To overcome this issue, type IIIs recognition enzyme sites (i.e., *BceAI* and *AclI*) were placed at the ends of the amplicons being amplified immediately adjacent to the unknown sequence during library construction (Figure 1A). The DPePCR product can then be digested with restriction enzymes (i.e., *BceAI* and *AclI*), and capping adaptors are ligated to the free end of the dsDNA (Figure 1D). This provides the ability to sequence from both strands of the DPePCR using standard sequencing by ligation (SBL) (Figure 1D) (2).

The SBL sequencing strategy for DPePCR beads is identical to sequencing from standard ePCR beads (Figure 1D). The difference is that since there are two paired-end fragments, both ends can be sequenced independently from both the 3'→5' and 5'→3' directions using four different anchor primers (Figure 1D).

To validate the formation of strong double-stranded DPePCR product, the beads (following DPePCR) were treated with 0.1 M NaOH (without restriction enzyme digestion). After denaturing, a Cy5-labeled oligo (Cy5-labeled FDV2-PM; see Supplementary Material) was annealed

Figure 1. Dual primer emulsion PCR approach. (A) The DPpPCR library. 70–130-bp DNA fragments were ligated to forward and reverse adaptors FDV2-BceAI-Primer-F/R and RDV2-AclI-Primer-F/R. Restriction enzymes recognition sites were designed for *BceAI* and *AclI* at the ends of the two primers for digestion of the DPpPCR product. (B) Dual primer beads. Equal amounts of FDV2-dual biotin and RDV2-dual biotin were attached to MyOne streptavidin beads. (C) Dual primer beads, diluted DNA fragment, ePCR reagents, and both free primers were mixed with silicone oil phase. The goal was for each microdroplet to contain a single bead, single DNA molecule, and enough PCR reagents for clonal amplification. After 120 cycles, amplified DNA bridges formed. (D) Digestion and sequencing of the DPpPCR beads. After the emulsion amplification, the DPpPCR beads were digested with *BceAI* and *AclI*. The digestion left 13 bases on the forward (*BceAI*) strand and 15 bases on the reverse (*AclI*) strand of fragment DNA for sequencing. Specific anchor primers were hybridized at the end of the forward and reverse strands (individually), followed by sequencing by ligation.

to the bead-bound DNA fragment (Supplementary Material). The results indicated that the Cy5-oligo could not hybridize to the DPpPCR product (data not shown), which suggests the formation of double-stranded DPpPCR product was in a bridged conformation. Following restriction enzyme digestion, we were able to anneal the anchor primers (Supplementary Material) and successfully sequence the DNA on the beads from both ends of the fragment.

Theoretically, the DNA fragment on each bead arose from a single molecule in an emulsion and should therefore be clonal, which is the critical requirement for the DPpPCR. To validate the clonal amplification, we used DPpPCR and SBL in the G.007 Polonator (Westborough, MA, USA) to sequence a *Streptococcus pyogenes* genome fragment library; when sequenced, the clonal amplified beads showed a single color during each SBL cycle (Figure 2, A and B), which indicated the beads were clonal. Additionally, a random sampling of $\sim 10^6$ reads were mapped to the *S. pyogenes* genome (AE014074) to ensure we were sequencing paired ends. The reads were NNNNNNN from the forward primer site, and NNNNNNNN(NN)NNNNNN from the reverse priming site [note that

We Illuminate Your Life Science Research!

Life science and laboratory research professionals depend on high-quality Spectroline® ultraviolet products to make their work faster and easier. Our products are built to exacting standards, ensuring years of reliable, trouble-free service.

UV Crosslinkers

Provide accurate, error-free and super-fast membrane crosslinking and sample UV dosing.

MiniMAX™ UV Viewing Station

Space-saving mini-cabinet and UV lamp combination makes the perfect portable UV viewing station. Light weight and compact.

UV-Protective Eye and Face Wear

Specially designed to protect user against unwanted UV exposure.

Miniature UV Pencil Lamps

Double-bore quartz lamps used for wavelength calibration and emission. Custom lamps for OEM applications also available.

UV Bench and Display Lamps

Lamps with various wattages and UV wavelengths provide both high intensity and wide area of light coverage.

UV Transilluminators

Deliver superior uniformity and UV intensity for all viewing and photodocumentation needs.

AccuMAX™ Digital Meters

Advanced microprocessor-controlled units with software-driven functions. Ensure maximum accuracy and reliability in measurement of UV and visible light sources.

Figure 2. Validation of Polonator DNA sequencing image. (A) The DPpPCR beads were sequenced using sequencing by ligation. Green, red, and blue beads relate to A, T, and C, respectively, in the image shown. Note that images are displayed in RGB for visualization, and the forth base is present but not shown. For example, the sequence from bead 1 is CTT; bead 2 is TAC from the three cycles shown. (B) The tetrahedral plot is used for the base calling. Each dot on the plot corresponds to a single bead at one cycle in the sequencing. The normalized fluorescence intensity in all four channels is plotted in 4-D and displayed here in a 3-D plot. The base is called by the highest intensity value. The quality of each call is defined by the distance from each bead to the centroid of the beads that have been called the same base. Since each of the beads is a clonal amplicon, the data points are clustered into four groups. The top 50% of the data from a single frame is shown in the figure. (C) The histogram shows that the average separation in the genome is ~100 bases, of a random sampling of ~20,000 reads that mapped uniquely to the *Streptococcus pyogenes* genome (AE014074). Thus, the beads are clonal amplicons of ~100-bp fragments from the *S. pyogenes* genome.

the (NN) bases was not sequenced]. To map to the *S. pyogenes* genome, the reverse priming site sequence was first mapped (to both strands of the genome) and then the forward priming site sequence was mapped to the opposite strand on the genome. Roughly, 25% of the reads mapped to the genome and the average separation between the paired ends in the genome was found to be ~100 bases (Figure 2C); this is consistent with the size range that selected in the gel purification step of the library preparation (Supplementary Material).

Additionally, as a final validation for the approach, we measured the library complexity. The complexity of the library is defined as the diversity of DNA molecules that are sequenced. For example, when a sequencing library is prepared and subjected to extensive amplification (PCR or rolling circle amplification) the complexity is reduced. The fragment library was found to have high complexity, with >99% of the reads being unique, thus improving upon the traditional mate-pair library production protocol (2,10).

In conclusion, we have presented DPpPCR, an efficient emulsion PCR strategy for paired-end next-generation sequencing. The DPpPCR concept will

be applicable to any next-generation sequencing platform using emulsion PCR.

Acknowledgments

This work was supported by the National Institutes of Health and the National Human Genome Research Initiative (NIH/NHGRI \$1,000 Genome grant no.R21HG004350, to J.S.E.). This paper is subject to the NIH Public Access Policy.

Competing interests

The authors declare no competing interests.

References

1. Tawfik, D.S. and A.D. Griffiths. 1998. Man-made cell-like compartments for molecular evolution. *Nat. Biotechnol.* 16:652-656.
2. Shendure, J., G.J. Porreca, N.B. Reppas, X. Lin, J.P. McCutcheon, A.M. Rosenbaum, M.D. Wang, K. Zhang, et al. 2005. Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* 309:1728-1732.
3. Dressman, D., H. Yan, G. Traverso, K.W. Kinzler, and B. Vogelstein. 2003. Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. *Proc. Natl. Acad. Sci. USA* 100:8817-8822.
4. Shendure, J. and H. Ji. 2008. Next-generation DNA sequencing. *Nat. Biotechnol.* 26:1135-1145.
5. Metzker, M.L. 2010. Sequencing technologies - the next generation. *Nat. Rev. Genet.* 11:31-46.
6. Ansorge, W.J. 2009. Next-generation DNA sequencing techniques. *N. Biotechnol.* 25:195-203.
7. Fullwood, M.J., C.L. Wei, E.T. Liu, and Y. Ruan. 2009. Next-generation DNA sequencing of paired-end tags (PET) for transcriptome and genome analyses. *Genome Res.* 19:521-532.
8. Mardis, E.R. 2008. Next-generation DNA sequencing methods. *Annu. Rev. Genomics Hum. Genet.* 9:387-402.
9. Pemov, A., H. Modi, D.P. Chandler, and S. Bavykin. 2005. DNA analysis with multiplex microarray-enhanced PCR. *Nucleic Acids Res.* 33:e11.
10. Porreca, G.J., J. Shendure, and G.M. Church. 2006. Polony DNA sequencing. *Curr. Protoc. Mol. Biol. Chapter 7:Unit 7.8.*

Received 2 February 2010; accepted 6 March 2010.

Address correspondence to Jeremy S. Edwards, University of New Mexico, Molecular Genetics and Microbiology, 915 Camino de Salud NE MSC08 4660, Albuquerque, NM 87131, USA. e-mail: jsedwards@salud.unm.edu