Polony DNA Sequencing

This unit provides protocols for performing polony DNA sequencing, a nonelectrophoretic sequencing method that eliminates in vivo cloning artifacts and affords a lower cost per base than conventional Sanger sequencing (Shendure et al., 2005). The first step in polony DNA sequencing is to construct an in vitro paired-tag library from genomic DNA (Basic Protocol 1; Fig. 7.8.1). Next, library molecules are clonally amplified on microbeads by emulsion PCR (Basic Protocol 2; Dressman et al., 2003). This clonal amplification yields polymerase colonies, or polonies, that can be sequenced. A Support Protocol describes a method used to optimize the amount of library DNA used as the ePCR template. Amplicon-bearing beads are then enriched (Basic Protocol 3) and two-dimensionally arrayed within an acrylamide gel matrix on a microscope coverglass (Basic Protocol 4). Finally, millions of short reads are generated in parallel from the microbeads via a cyclic DNA sequencing strategy that utilizes T4 DNA ligase to selectively tag each microbead with fluorescent labels that correlate with the unique nucleotide sequence present on any given bead (Basic Protocol 5).

The combination of protocols can be thought of as a modular platform, with each major step (library construction, emulsion PCR, sequence readout) independent enough to be replaced by an alternate. For example, if paired tags are not necessary to place reads (i.e., to match sequences to a reference genome sequence), single-tag libraries can be constructed (e.g., for barcode sequencing or RNA quantitation). The emulsion PCR protocol presented in this unit is a variation of that published by Dressman et al. (2003), but should not be thought of as the only applicable amplification method. Different sequencing biochemistries can also be used; in fact, during development of the polony sequencing platform, a number of alternate polymerase- and ligase-driven approaches were evaluated.

To implement the protocols presented in this unit, various instruments (e.g., fluorescent microscope, autosampler, and flow cell) and computer hardware and software are required. A discussion of the critical features of the setup is provided (see Critical Parameters and Troubleshooting), but the reader is directed to *http://arep.med.harvard.edu/Polonator* for more detailed information, including full schematics and drawings. Instructions are available there for using the hardware and software, as are binary and source codes for the software. All software is free and open-source, and all hardware is commercially available.

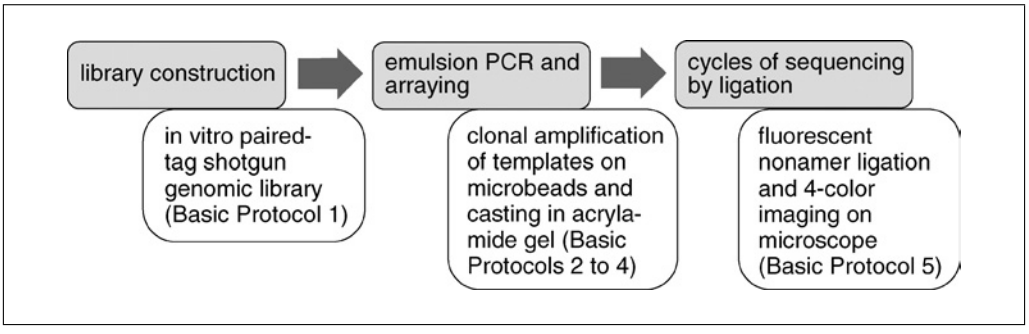

Figure 7.8.1 General procedure for polony sequencing. A paired-tag genomic library is used as template for emulsion PCR on microbeads to generate polymerase colonies (polonies). The beads are cast on a coverslip in a thin layer of polyacrylamide, and put through iterative cycles of single-base sequencing.

Contributed by Gregory J. Porreca, Jay Shendure, and George M. Church
Current Protocols in Molecular Biology (2006) 7.8.1-7.8.22

BASIC PROTOCOL 1

CONSTRUCTION OF A SHOTGUN PAIRED-TAG GENOMIC LIBRARY

Polony genomic resequencing is generally performed on a shotgun paired-tag library (Shendure et al., 2005). Each library molecule is 135 bp in length, and has two 17- to 18-bp paired genomic tags separated and flanked by common sequences (Fig. 7.8.2). The key step in this protocol is the circularization of randomly sheared and size-selected genomic DNA around a synthetic insert-linker. This links sequences that are separated on the genome by a defined distance distribution.

Materials

Genomic DNA
Buffer EB (Qiagen)
End-It DNA end repair kit (Epicentre), including:
- 10× buffer
- 10× ATP
- 10× dNTP mix
- Enzyme mix

10× PCR buffer without $MgCl_2$ (Invitrogen)
50 mM $MgCl_2$
100 mM dATP
5 U/μl *Taq* DNA polymerase
20 mg/ml glycogen
Oligonucleotides:
- 100 μM T30-T: 5′-phosphorylated-GTCGGAGGCCAAGGCGGCCGTACGT CCAACT-3′ (purified by HPLC)
- 100 μM T30-B: 5′-phosphorylated-GTTGGACGTACGGCCGCCTTGGCC TCCGACT -3′ (purified by HPLC)
- 1 mM N6 oligonucleotides: 5′-NNNN*N*N-3′ (*signifies phosphorothioate linkage; IDT)
- 100 μM FDV-B: 5′-ATCACCGACTGCCCATAGAGAGGAAAGCGGAGGC GTAGTGGTT-3′ (purified by HPLC)
- 100 μM FDV-T: 5′-AACCACTACGCCTCCGCTTTCCTCTCTATGGGCAGT CGGTGAT-3′ (purified by HPLC)
- 100 μM RDV-B: 5′-AGAGAATGAGGAACCCGGGGCAGTT-3′ (purified by HPLC)
- 100 μM RDV-T: 5′-AACTGCCCCGGGTTCCTCATTCTCT-3′ (purified by HPLC)

Quick ligation kit (NEB), including:
- 2× Quick ligation buffer
- Quick T4 DNA ligase

20 U/μl Exonuclease I (NEB)
100 U/μl Exonuclease III (NEB)
TE buffer, pH 8.0 (*APPENDIX 2*)
25 mM dNTP mix (25 mM each nucleotide)
10× RepliPhi phi29 reaction buffer (Epicentre Technologies)
100 U/μl RepliPhi phi29 polymerase (Epicentre)
10× NEBuffer 4 (NEB)
1.6 mM *S*-adenosylmethionine (SAM; diluted from 32 mM stock in 1× NEBuffer 4; NEB)
2 U/μl *Mme*I (NEB)
2000 U/μl T4 DNA ligase and 10× buffer (NEB)
40% (w/v) polyethylene glycol 8000 (PEG) in H_2O
10× NEBuffer 2 (NEB)
10 U/μl *E. coli* DNA polymerase I (NEB)
5 U/μl Platinum *Taq* DNA polymerase (Invitrogen)

Qiaquick columns (Qiaquick PCR cleanup kit; Qiagen)
NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies)
Microcon-30 column (Millipore)
Thermal cycler

Additional reagents and equipment for DNA quantitation (*APPENDIX 3D*), phenol/chloroform extraction and ethanol precipitation (*UNIT 2.1A*), and agarose and polyacryamide gel electrophoresis (*UNITS 2.5A & 2.7*)

Prepare genomic DNA for circularization

1. Shear 15 μg genomic DNA to desired size distribution.

   The Hydroshear (http://www.genomicsolutions.com) is the recommended instrument for this purpose since, when the manufacturer's instructions are followed, it produces a population of fragments having a relatively tight size distribution and ends that can be efficiently blunted. The authors typically shear DNA at a concentration of ~60 ng/μl.

2. Purify DNA on Qiaquick columns (≤10 μg per column) as per manufacturer's instructions.

   Use as few columns as possible.

3. To repair the DNA ends (i.e., blunt them following shearing), combine the DNA with the components of the End-It DNA end repair kit as follows:

   10 μg column-purified, sheared genomic DNA in buffer EB
   25 μl 10× buffer
   25 μl 10× ATP
   25 μl 10× dNTP mix
   5 μl enzyme mix
   H_2O to a final volume of 250 μl.

   Incubate 1 hr at room temperature.

4. Purify DNA on a Qiaquick column per manufacturer's instructions, except elute with 90 μl buffer EB.

5. Quantitate DNA using a NanoDrop ND-1000 spectrophotometer (see *APPENDIX 3D*).

   Concentration should be ~100 ng/μl.

6. Incubate 15 min at 70°C to eliminate any residual enzyme activity.

7. Add an adenosine tail by combining the following components:

   2.2 μg end-repaired DNA
   10 μl 10× PCR buffer without $MgCl_2$
   6 μl 50 mM $MgCl_2$
   0.5 μl 100 mM dATP
   0.5 μl 5 U/μl *Taq* DNA polymerase
   H_2O to 100 μl.

   Incubate 30 min at 70°C then cool to 4°C. Transfer to ice.

   In the A-tailing reaction, Taq polymerase adds a single A to the 3′ ends of the blunted DNA in a template-independent fashion.

8. Extract DNA from the reaction mixture with phenol/chloroform and precipitate in ethanol using 20 μg glycogen as carrier (*UNIT 2.1A*). Resuspend pellet in 40 μl buffer EB.

   Glycogen should be used as a carrier in all ethanol precipitation steps in this protocol.

9. Separate the sheared, A-tailed DNA on a 6% polyacrylamide gel with TBE buffer (*UNIT 2.7*). Stain with ethidium bromide.

   For Invitrogen precast gels, mix 20 μl DNA with 5 μl of 5× loading buffer (Novex) and load across two 0.5-cm lanes.

10. Cut out the desired bands and extract DNA using the crush-and-soak method (*UNIT 2.7*). Precipitate in ethanol using glycogen as a carrier, and resuspend in 10 μl buffer EB.

    For a paired-tag library with an average intertag distance of 1 kb, excise and pool gel bands centered at 1 kb with a total width of no more than 500 bp. Exposure of the DNA to UV light during visualization should be minimized.

11. Quantitate the size-selected DNA by separating on an agarose or polyacrylamide gel (*UNIT 2.5A or 2.7*) and comparing the brightness of the smear to that of a known amount of a molecular weight standard.

    The concentration should be ~20 to 40 ng/μl.

Circularize genomic material around synthetic oligonucleotides

12. Anneal oligonucleotides T30-T and T30-B in a thermal cycler by combining equal amounts and heating to 95°C for 10 min and then slowly cooling to room temperature over the course of 1 hr.

    The final concentration of the T30 insert-linker should be 100 μM. T30 has T 3′ overhangs designed to be complementary to the A 3′ overhangs of the genomic DNA.

13. Ligate genomic DNA fragments in the presence of annealed T30 using the Quick ligation kit as follows:

    170 ng size-selected DNA (0.25 pmol at ~1 kb)
    0.8 μl 1 μM annealed T30 (0.8 pmol)
    40 μl 2× Quick ligation buffer
    4.0 μl Quick T4 DNA ligase
    H_2O to a final volume of 80 μl.

    Incubate reaction 10 min at room temperature then heat-inactivate 10 min at 65°C. Use the heat-inactivated reaction mixture directly in the next step without further purification.

    The ligation should be performed under conditions favoring formation of monomeric recombinant circles (i.e., genomic DNA–T30). To this end, T30 should be present at a three-fold molar excess to genomic fragments.

14. Eliminate all noncircularized material by exonucleolysis as follows:

    1.0 μl 20 U/μl Exonuclease I
    0.1 μl 100 U/μl Exonuclease III
    80 μl heat-inactivated reaction mix
    TE buffer, pH 8.0, to a final volume of 90 μl.

    Incubate 45 min at 37°C, and then heat-inactivate 20 min at 80°C. Use this reaction mixture directly in the next step without further purification.

Amplify circular DNA

15. Prepare the master mix for hyper-branched rolling-circle amplification of the circular DNA:

    12 μl 25 mM dNTP mix
    30 μl 10× RepliPhi phi29 reaction buffer
    15 μl 1 mM N6 oligonucleotides
    213 μl H_2O
    30 μl circularized DNA.

    Split into 6 tubes of 50 μl each.

16. Denature material by heating to 95°C for 5 min, then anneal by rapidly cooling to 4°C. Add 2.5 μl phi29 polymerase to each tube, keeping on ice. Incubate overnight at 30°C.

17. Pool amplified circular DNA into a single tube. Purify DNA using a Microcon-30 column according to the manufacturer's instructions, washing with 1 ml TE buffer, pH 8.0. Wash the membrane several times to maximize recovery.

18. Resuspend the DNA using 750 μl buffer EB, preheated to 50°C. Quantitate the DNA using a NanoDrop ND-1000 spectrophotometer.

    Concentration should be ~230 ng/μl.

Release paired tags

19. Digest the amplified DNA with *Mme*I to release the 70-bp T30-linked paired tags:

    40 μg amplified circular DNA
    100 μl 10× NEBuffer 4
    1.6 μl of 1.6 mM SAM
    60 μl *Mme*I (2 U/μl)
    H_2O to 1000 μl.

    Split into eight tubes of 125 μl each on ice, and then incubate at 37°C for 30 min.

    MmeI cuts 18 to 19 bp from its recognition site, which is 1 bp from the start of the genomic fragment. The released fragments contain the common T30 sequence flanked by two 17- to 18-bp tags of genomic DNA.

20. Immediately extract the DNA with phenol/chloroform and precipitate with ethanol using glycogen as a carrier. Resuspend pellet in a total volume of 80 μl TE buffer, pH 8.0.

21. Purify the 70-bp paired-insert library as in steps 9 and 10, resuspending precipitated DNA in 20 μl TE buffer, pH 8.0.

    The sample should be split across four 0.5-cm lanes of a precast 6% polyacrylamide gel in TBE buffer.

22. Quantitate the gel-purified DNA by electrophoresis (step 11).

    DNA concentration should be ~12 ng/μl.

Add emulsion PCR adapters

23. To perform the blunting reaction (to remove 2-nt 3′ overhangs from the *Mme*I digestion), combine the DNA with components of the End-It DNA end repair kit as follows:

    8.5 μl paired-insert library (~100 ng DNA)
    1.25 μl 10× buffer
    1.25 μl 10× ATP
    1.25 μl 10× dNTP mix
    0.25 μl enzyme mix.

    Incubate at room temperature for 45 min.

24. Put on ice and add TE buffer, pH 8.0 to bring volume to 50 µl. Extract the DNA with phenol/chloroform and precipitate in ethanol using 20 µg glycogen as carrier. Resuspend DNA in 8 µl TE, pH 8.0.

25. Prepare emulsion PCR amplification primer sequences by annealing oligonucleotides FDV-T and FDV-B and oligonucleotides RDV-T and RDV-B as in step 12.

    The final concentration of each oligonucleotide should be 50 µM. FDV and RDV are not 5′-phosphorylated, so they will not ligate to one another.

26. Prepare the ligation reaction mixture by mixing the following components on ice:

    8.0 µl blunt-ended library DNA (~100 ng, 2 pmol)
    1.0 µl 50 µM FDV (50 pmol)
    1.0 µl 50 µM RDV (50 pmol)
    2.5 µl 10× T4 DNA ligase buffer (~0.5×)
    21.1 µl 40% PEG
    12.3 µl H_2O.

27. Mix on ice, then add 2.0 µl of 2000 U/µl T4 DNA ligase and incubate overnight at 16°C to ligate primer sequences to the library molecules.

    After ligation, several species will be present: FDV-insert-FDV, FDV-insert-RDV, and RDV-insert-RDV.

28. Increase volume to 100 µl with TE buffer, pH 8.0. Extract the DNA with phenol/chloroform, precipitate in ethanol using glycogen as a carrier, and resuspend in 10 µl buffer EB.

29. Purify the 135-bp FDV-RDV paired-insert library as in steps 9 and 10, resuspending the precipitated DNA in 10 µl TE buffer, pH 8.0.

30. Assemble on ice the reaction mixture used to remove nicks in double-stranded library molecules by nick translation:

    10.0 µl library DNA
    0.5 µl 25 mM dNTP mix
    2.5 µl 10× NEBuffer 2
    1.0 µl 10 U/µl *E. coli* DNA polymerase I
    11.0 µl H_2O.

    Incubate 30 min at 16°C.

31. Increase volume to 100 µl with TE buffer, pH 8.0. Extract the DNA with phenol/chloroform, precipitate in ethanol using glycogen as a carrier, and resuspend in 10 µl TE buffer, pH 8.0.

Amplify FDV/RDV-adapted library

32. Prepare the master mix for PCR amplification (for eight tubes of 50 µl each):

    50 µl 10× PCR buffer without $MgCl_2$
    4.0 µl 25 mM dNTP mix
    15 µl 50 mM $MgCl_2$
    426 µl H_2O
    1.0 µl 100 µM FDV-T
    1.0 µl 100 µM RDV-T
    2.0 µl 5 U/µl Platinum *Taq* DNA polymerase.

33. Split the master mix into eight tubes of 50 µl each, and add 0.5 µl FDV/RDV-adapted library material to each tube.

34. Carry out PCR using the following amplification cycles:

Initial step:	2 min	94°C	(denaturation)
12 cycles:	30 sec	94°C	(denaturation)
	30 sec	55°C	(annealing)
	90 sec	72°C	(extension).

The number of cycles is kept low to minimize PCR amplification bias. Due to suppression PCR effects (Siebert et al., 1995), FDV/FDV- and RDV/RDV-adapted molecules are not amplified.

35. Purify DNA using Qiaquick columns according to the manufacturer's instructions for PCR cleanup.

36. Purify the amplified 135-bp FDV-RDV paired-insert library as in steps 9 and 10, resuspending the precipitated DNA in 10 µl TE buffer, pH 8.0.

When excising the 135-bp library band, be careful to avoid contamination by marker DNA, if used.

37. Quantitate the final DNA library by electrophoresis (step 11).

The concentration should be ~2 ng/µl.

EMULSION PCR OF PAIRED-TAG LIBRARY ON MICROBEADS

BASIC PROTOCOL 2

Emulsion PCR (ePCR) is used to generate clonally amplified DNA fragments suitable for polony sequencing. First, a biotinylated PCR primer is immobilized on streptavidin-coated beads. Next, billions of microreactors are formed by the emulsification of a PCR mix in mineral oil (Dressman et al., 2003). When one bead (bearing an immobilized PCR primer) and one template molecule are trapped in a compartment, amplification will proceed and result in tens of thousands of copies of the template molecule attached to the bead. The result is a complex mixture of clonally amplified templates (polonies) anchored to beads. Each bead bears numerous copies of the same template, but different beads bear different templates.

Materials

MyOne C1 1-µm paramagnetic, streptavidin-coated beads (Dynal)
Bind and wash (B&W) buffer (see recipe)
Oligonucleotides:
  1 mM PR1-F-2BIO: 5′-dual-biotin-CCACTACGCCTCCGCTTTCCTCTCTAT GGGCAGTCGGTGAT-3′
  2 mM PR1-R: 5′-CTGCCCCGGGTTCCTCATTCTCT-3′
  10 µM PR1-3LF: 5′-CCTCTCTATGGGCAGTCGGTGAT-3′
TE buffer, pH 8.0 (*APPENDIX 2*)
Light mineral oil (Sigma)
10% Span 80 (see recipe)
Tween 80
Triton X-100
10× PCR buffer without $MgCl_2$ (Invitrogen)
50 mM $MgCl_2$
25 mM dNTP mix (25 mM each nucleotide)
5 U/µl Platinum *Taq* DNA polymerase (Invitrogen)
Template DNA (library DNA at appropriate concentration; see Basic Protocol 1 and Support Protocol)

NXS buffer (see recipe)
0.1 M NaOH

1.5-ml microcentrifuge tubes
Magnetic particle concentrator (MPC; Dynal)
2-ml cryogenic vials (Corning no. 430661)
Stir bar, flea-size (VWR no. 58948-353)
Magnetic stirrer (closed-loop; VWR)
200-µl eight-tube PCR strips
Thermal cycler

Bind forward PCR primer to microbeads

1. To 30 µl MyOne beads in a 1.5-ml microcentrifuge tube, add 30 µl B&W buffer.
2. Mix and then remove all liquid using a magnetic particle concentrator (MPC).

   Vortex to mix in all steps except those where enzyme activity is required. In those cases mix gently by flicking or rotating the tube.

3. Wash two times in 60 µl B&W buffer using an MPC, and resuspend in 60 µl B&W buffer.
4. Add 1.2 µl PR1-F-2BIO, mix, and incubate 20 min at room temperature.

   Mix periodically by pipetting to prevent settling of beads.

   PR1-F-2BIO is synthesized with a dual biotin at the 5′ end to tightly bind the streptavidin-coated microbeads. Dual biotin is necessary to ensure the template remains on the bead for the duration of the sequencing run.

5. Remove all liquid using an MPC, and then wash twice in 100 µl B&W buffer and once in 100 µl TE buffer, pH 8.0.
6. Remove all liquid using an MPC and resuspend in 60 µl TE buffer.

   Concentration should now be ~5 × 10^9 beads/ml.

Perform emulsion PCR reaction

7. Prepare the oil phase by assembling six 1.5-ml microcentrifuge tubes containing the following:

   545 µl light mineral oil
   450 µl 10% Span 80
   4.0 µl Tween 80
   0.5 µl Triton X-100.

   Vortex each tube individually for at least 20 sec to thoroughly mix all components.

   Watch the pipet tips carefully to ensure that viscous reagents are delivered into the tube. Use reverse pipetting (see http://www.rainin.com/pdf/rainin_classic_manual.pdf) for all oil-phase components.

8. Prepare the aqueous phase by combining the following:

   96 µl 10× PCR buffer without $MgCl_2$
   359 µl 50 mM $MgCl_2$
   135 µl 25 mM dNTP mix
   12.0 µl 2 mM PR1-R
   4.8 µl 10 µM PR1-3LF
   60 µl microbeads from step 6
   246 µl H_2O
   54 µl 5 U/µl Platinum *Taq* DNA polymerase
   1.0 µl template DNA.

9. Place twelve 2-ml cryogenic vials in the centers of three magnetic stirrers set to 1400 rpm.
10. Place a flea-size stir bar in each vial.
11. Add 400 µl oil phase to each vial.
12. Add 75 µl aqueous phase, dropwise over 1 min, to each vial.

    Be sure each drop of liquid hits the bottom of the tube, not the walls.

13. Stir for 30 min using a magnetic stirrer.
14. Pipet the contents of each cryogenic vial into a separate eight-tube PCR strip (50 µl per tube) for a total of twelve strips, or 96 tubes.

    Aspirate and dispense slowly to minimize the amount of emulsion that sticks to the walls of the pipet tip.

15. Carry out PCR using the following amplification cycles:

Initial step:	2 min	94°C	(denaturation)
120 cycles:	15 sec	94°C	(denaturation)
	30 sec	57°C	(annealing)
	75 sec	70°C	(extension)
Finish:	2 min	72°C	
Hold:	indefinitely	4°C.	

16. Pool contents of each strip of eight 200-µl tubes into a separate 1.5-ml microcentrifuge tube (twelve tubes). Immediately proceed with recovery.

Recover beads from emulsion

17. Add 800 µl NXS buffer to each tube.
18. Vortex each tube individually at maximum speed for at least 20 sec.
19. Centrifuge all tubes 90 sec at 16,110 × *g*, room temperature.
20. Aspirate all but 100 µl supernatant (being sure to remove all of the top oil phase).
21. Repeat steps 17 to 20 two more times.
22. Pellet beads in tubes using an MPC, and aspirate all liquid.
23. Resuspend each pellet in 25 µl TE buffer, pH 8.0, and transfer all to a single new tube. Mix by pipetting up and down.

    All twelve tubes should be pooled into this one new tube. Try to minimize the amount of oil and detergent carried over into the new tube from the walls of the old tube.

24. Pellet the beads using an MPC, aspirate all liquid, and add 250 µl TE buffer, pH 8.0. Vortex briefly, then pulse centrifuge to pull the liquid into the bottom of the tube. Repeat two more times.
25. Pellet beads using an MPC, and aspirate all liquid.
26. Remove the nonbiotinylated strand from the beads by incubating in 100 µl of 0.1 M NaOH for 20 min at room temperature. Periodically mix by pipetting to prevent beads from settling.
27. Pellet beads using an MPC, and remove all liquid.
28. Wash once with 100 µl 0.1 M NaOH and three times with 150 µl TE buffer. Mix each wash by vortexing to ensure removal of residual NaOH.
29. Pellet beads using an MPC, aspirate all liquid, and resuspend in 60 µl TE buffer.

ENRICHMENT FOR AMPLICON-BEARING BEADS

The result of ePCR is a population of beads, most (~80%) of which do not bear an amplified template. In order to maximize the number of amplicon-bearing beads on the sequencing array, an enrichment step is required. Fluorescence-activated cell sorting can be employed to select for these beads by hybridizing a fluorescently labeled probe specific to amplicons. Alternately, the centrifugation technique presented below can be used for more rapid and cost-effective enrichment.

In this protocol, the dense paramagnetic ePCR beads are incubated with low-density polystyrene beads bearing an oligonucleotide complementary to a common amplicon sequence. Amplicon-bearing beads hybridize to these capture beads, while non-amplicon-bearing beads do not. The density differential between captured and uncaptured ePCR beads is exploited for enrichment by centrifugation through a viscous 60% glycerol solution; the former remain in the supernatant while the latter form a pellet.

Materials

Spherotech particles (3-μm streptavidin-coated polystyrene particles; Spherotech no. SVP-30-5)
Bind and wash (B&W) buffer (see recipe)
1 mM PR1-BIOXL:
5′-biotinTEG-CGTACCCCGCTTGGTCTTTCTCCCGTACCCCGCTTGG TCTTTCTCCCTGCCCCGGGTTCCTCATTCTCT-3′
TE buffer, pH 8.0 (*APPENDIX 2*)
ePCR beads (Basic Protocol 2)
Glycerol
0.1 M NaOH
Magnetic particle concentrator (MPC; Dynal)
1.5-ml microcentrifuge tubes

Bind capture oligonucleotide to capture beads

1. Centrifuge 400 μl Spherotech particles for 30 sec at 16,110 × *g*, room temperature.
2. Remove all liquid, and wash pellet with 400 μl B&W buffer.
3. Centrifuge for 30 sec at 16,110 × *g*.
4. Remove all liquid, and resuspend in 400 μl B&W buffer.
5. Add 8 μl of 1 mM PR1-BIOXL and incubate at room temperature for 20 min.
6. Wash three times in 800 μl TE buffer, pH 8.0, pelleting each time by centrifugation. Resuspend final pellet in 80 μl B&W buffer.

Anneal capture beads to emulsion PCR beads

7. Add 60 μl ePCR beads to 80 μl capture beads and incubate at 56°C for 10 min.
8. During the incubation, prepare 1 ml of fresh 60% (v/v) glycerol in H_2O. Vortex to mix thoroughly.

   Must be made fresh each day.
9. Add 150 μl of 60% glycerol to each of four 1.5-ml microcentrifuge tubes.
10. Layer 35 μl bead mixture from step 7 onto the top of the glycerol solution in each of the four tubes.

    Be careful to pipet gently so the bead solution does not mix with glycerol solution.

Separate bead populations by centrifugation

11. Centrifuge tubes for 1 min at 16,110 × *g*.

    The amplicon-bearing ePCR beads are hybridized to the complementary sequences carried on the low-density polystyrene capture beads and will remain in the supernatant; the ePCR beads without amplified sequences will form a pellet in the bottom of the tube.

12. Transfer all the supernatant to four new microcentrifuge tubes, being careful not to disturb the pellets.

    If necessary, wash the walls of the tubes with the supernatant to collect any beads bound there.

13. To each recovered supernatant tube, add 1 ml H_2O and vortex until thoroughly mixed.

14. Centrifuge for 3 min at 16,110 × *g*, room temperature, to pellet all beads.

15. Aspirate, being careful to not disturb the pellet, and then place tubes on an MPC to hold the pellet firmly in place. Remove any remaining liquid.

16. Separate ePCR beads from capture beads by adding 120 µl of 0.1 M NaOH and incubating 10 min at room temperature to denature the hybridized DNA strands.

17. Using the MPC, collect the amplicon-bearing ePCR beads into a pellet. Remove and discard all liquid.

    The liquid contains the released capture beads.

18. Wash the pellet twice by adding 120 µl of 0.1 M NaOH, vortexing briefly, centrifuging briefly to pull the liquid to the bottom of the tube, and collecting the pellet using the MPC.

19. Resuspend in 6.5 µl TE buffer, pH 8.0.

CASTING A POLONY BEAD ARRAY

BASIC PROTOCOL 4

The amplified, enriched beads are cast in a two-dimensional array to fix their positions during the subsequent sequencing reactions by mixing with acrylamide and pouring into a shallow mold formed by a Teflon-masked microscope slide. The coverslip placed on top of the gel is treated with Bind Silane, which covalently attaches the gel to the glass coverslip. The acrylamide gel mixture is formulated such that it takes 30 min or more to polymerize, during which time the beads settle into a single monolayer before their positions become fixed. This allows efficient imaging of all beads in the gel, since they all reside in the same focal plane of the optics used for imaging.

Once polymerized, the gel is approximately 30 microns thick, while the beads are only 1 micron in diameter. To ensure the DNA-bearing beads are accessible to reagents for enzymatic sequencing reactions, the gel is cast such that the beads will be fixed at the exposed surface (i.e., opposite the side attached to the Bind Silane–treated glass).

Materials

1% Triton X-100 in H_2O
Glacial acetic acid
Bind Silane (Promega)
100% ethanol
Enriched ePCR beads in TE buffer, pH 8.0 (Basic Protocol 3)
40% acrylamide/bisacrylamide (19:1) solution
5% TEMED
0.5% ammonium persulfate (APS)
Wash 1 (see recipe)

Round coverslips, 40 mm diameter, no. 1.5 (Bioptechs no. 40-1313-0319)
1000-ml plastic beaker
Shaker
Vacuum desiccator
Teflon-masked microscope slides (Erie Scientific, no. ER-203W)

Treat coverslips with Bind Silane

1. Wash round coverslips in 1% Triton X-100 for 20 min with gentle shaking.
2. Prepare silane solution in a 1000-ml plastic beaker:

   500 ml H_2O
   110 µl glacial acetic acid
   2 ml Bind Silane.

   Stir for 15 min until Bind Silane droplets have completely disappeared.
3. Immerse coverslip in silane solution and incubate 1 hr at room temperature with gentle shaking.
4. Wash thoroughly with H_2O, then rinse with 100% ethanol.
5. Allow to dry completely in vacuum desiccator. Store coverslips up to 1 month under vacuum.

Cast enriched ePCR beads

6. Prepare bead/acrylamide mixture:

   6.5 µl enriched ePCR beads in TE buffer
   1.25 µl 40% acrylamide/bisacrylamide solution (19:1)
   0.5 µl 5% TEMED
   0.5 µl H_2O.

   Mix by pipetting, then add 0.75 µl of 0.5% APS.
7. Immediately pipet 9.5 µl onto a Teflon-masked microscope slide and cover with a silane-coated coverslip. Allow to polymerize for 45 min.

   The coverslip should spread the liquid across the entire surface of the circular well on the Teflon-masked slide. During polymerization, the beads will settle to the surface of the gel closest to the microscope slide.
8. Invert the slide/coverslip stack and remove microscope slide from gel by gently lifting with a razor blade near the edge of the coverslip.

   The gel should remain attached to the coverslip.
9. Immerse the coverslip, with polony bead array attached, in wash 1 until ready for use, up to ~1 month at 4°C.

BASIC PROTOCOL 5

DNA SEQUENCING BY LIGATION WITH DEGENERATE FLUORESCENT NONAMERS

An anchor primer is first hybridized to a known sequence, or priming site, within the single-stranded template. This primer is synthesized with deoxyuridine in place of thymidine. A ligation reaction is then performed with a pool of fully degenerate, fluorescently labeled nonanucleotides. During the ligation reaction, each bead becomes tagged with a fluorophore that indicates the identity of the base present at the position being interrogated. After imaging the array to determine which fluorophore labels each bead, the anchor-primer/fluorescent-nonamer complex is removed from the template by reacting

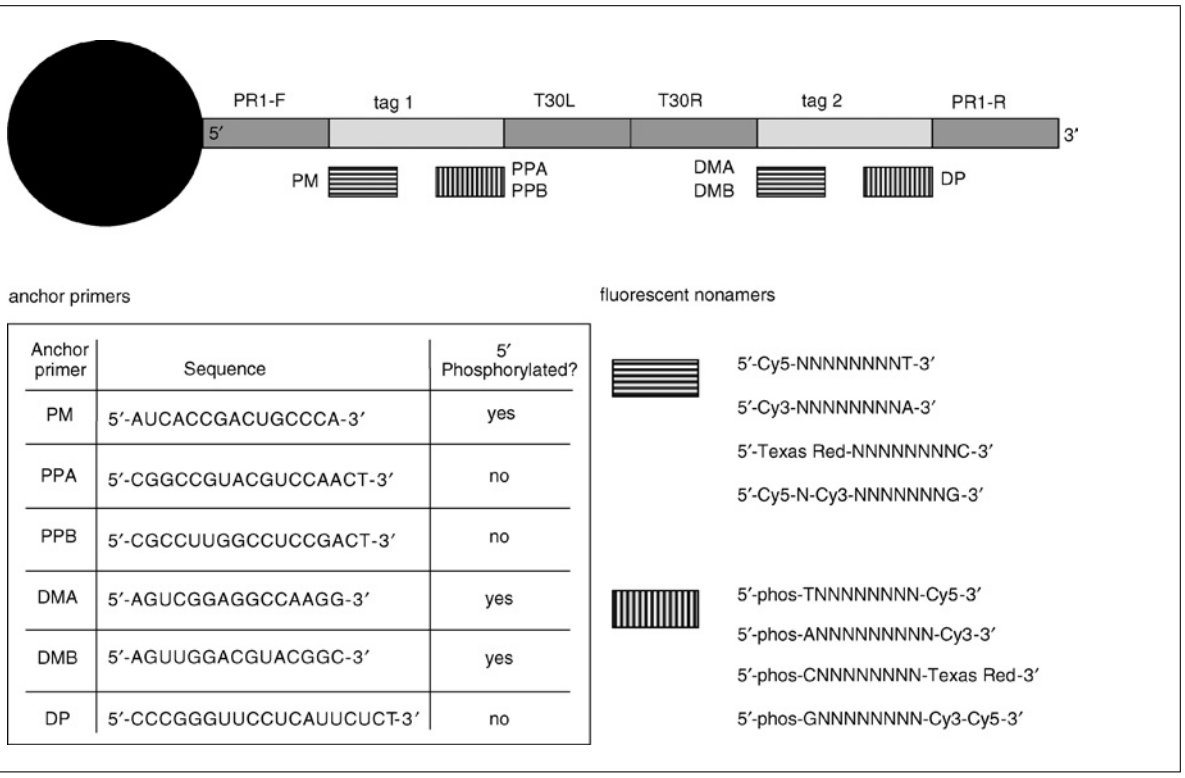

Figure 7.8.2 Schematic of the 135-bp bead-bound template and sequences of oligonucleotides used in sequencing reactions (Basic Protocol 5). For each reaction, the combination of anchor primer(s) and fluorescent nonamers is specified by the position in the tag being sequenced. All anchor primers should be synthesized with deoxyuridine to allow enzymatic removal of the primers. Fluorescent nonamers should be synthesized by Integrated DNA Technologies (IDT) with fully degenerate positions hand mixed to ensure equal ratios of each nucleotide at each position. The T30 sequence is shown as T30L and T30R to emphasize that the sequencing primers do not bind the full length of T30. Abbreviations: DMA and DMB, distal minus A and B strands; DP, distal plus; PM, proximal minus; PPA and PPB, proximal plus A and B strands; PR1-F and PR1-R, forward and reverse PR1 primers.

with the uracil-specific excision reagent (USER) enzyme. The USER enzyme excises all deoxyuridines from the anchor primer, resulting in short fragments that dissociate at moderate temperatures (<60°C) due to their low melting points (T_m), and leaving the bead-bound template ready for another position to be interrogated.

The combination of degenerate nonamer pool and anchor primer specifies which position in the template the reaction will interrogate (see Fig. 7.8.2). Whether the anchor primer or the nonanucleotides should be synthesized with a 5′ phosphate is dependent on which position in the template is being queried with a given combination of anchor primer and nonamer. When sequencing the 5′ side of a tag, the anchor is phosphorylated, and when sequencing the 3′ side, the nonamers are phosphorylated.

At the start of a sequencing run, all 3′ ends of unextended forward ePCR primers and templates are capped by the addition of dideoxynucleotides using terminal deoxytransferase. This capping reaction is necessary to prevent nonspecific ligation of labeled nonamers to bead-bound ePCR primers and templates.

This protocol describes the ligation and stripping steps required to query the base at a specific position in an iterative fashion. When using the autosampler and flow cell system, multiple solutions can be prepared for one or several cycles and used with a computer program preset with the appropriate times, temperatures, and volumes indicated in the protocol steps. Imaging and analysis procedures are discussed in Critical Parameters and Troubleshooting.

Materials

Bead array with template DNA (Basic Protocol 4)
5× tailing buffer (Invitrogen)
1.25 mM ddNTP mix (1.25 mM each dideoxynucleotide)
15 U/μl terminal deoxytransferase (recombinant; Invitrogen)
Wash 1 (see recipe)
6× SSPE (*APPENDIX 2*)/0.01% Triton X-100
Anchor primer (see Fig. 7.8.2)
2000 U/μl T4 DNA ligase and 10× buffer (NEB)
100 μM degenerate fluorescent nonamer mix (Integrated DNA Technologies)
TE buffer, pH 8.0 (*APPENDIX 2*)
1 U/μl USER enzyme mix (NEB)

Automated fluorescent microscope with flow cell (see Table 7.8.1 for parts list)
Autosampler (see Table 7.8.1)
1-ml disposable syringe
0.45-μm pore size, 4-mm cellulose acetate syringe filter (VWR)

Perform dideoxynucleotide capping reaction

1. Prepare the capping reaction mix:

   114 μl H_2O
   30 μl 5× tailing buffer
   4 μl 1.25 mM ddNTP mix
   2 μl 15 U/μl terminal deoxytransferase.

2. Place the bead array with template DNA on a coverslip into the flow cell mounted on the microscope stage. Prime fluid lines with wash 1.

3. Add the reaction mix to the appropriate well of the autosampler. Incubate 60 min at 35°C.

Hybridize anchor primer to array

4. Prepare the hybridization mix by combining:

   150 μl 6× SSPE with 0.01% Triton X-100
   1.5 μl 1 mM anchor primer.

5. Add hybridization mix to the appropriate well of the autosampler. Incubate 5 min at 56°C, then 2 min at 42°C.

Ligate fluorescent nonamers to visualize complexes

6. Prepare ligation reaction on ice:

   15 μl 10× T4 DNA ligase buffer
   12.0 μl 100 μM degenerate fluorescent nonamer mix
   3.0 μl 2000 U/μl T4 DNA ligase
   120 μl H_2O.

   When preparing multiple ligation reaction mixtures (for sequential cycles) at once, be sure they are consumed within 6 hr, since T4 DNA ligase is unstable once mixed with reaction buffer. In addition, only one mix containing 5′-phosphorylated nonamers should be prepared, and it should be used in the first sequencing cycle performed.

7. Aspirate the solution into a 1-ml syringe, attach a 0.45-μm filter to the Luer syringe adapter, and pass the contents through the filter into the appropriate well of the autosampler. Incubate 30 min at 32°C.

   The nonamer solutions contain large aggregates of fluorescent dye. Filtering is essential to ensure these aggregates are not introduced into the flow cell, since they will gradually accumulate by sticking to the gel, interfere with imaging and data extraction, and poison subsequent reactions.

   Incubation temperature and duration can be optimized for variations in signal. If signal is low, a slightly longer time or higher temperature can be used; if signal is very high, cycle time can be decreased by decreasing incubation time appropriately. Although increased concentrations of nonamer may also drive the reaction to completion faster, it should be noted that cost will increase and fidelity may be affected.

8. Image using fluorescence microscopy to determine the nonamer ligated to each bead and a single position of one tag.

   The automated microscope used for polony sequencing includes a number of add-on components both for imaging (e.g., motorized stage, piezo objective positioner, illumination wheels) and for fluidics (e.g., flow cell assembly, autosampler, digital I/O modules). Complete details are beyond the scope of this unit. All major components are listed in Table 7.8.1, and a discussion is provided in Critical Parameters and Troubleshooting. More detailed information on assembly (including fittings, tubing, and more) can be found at http://arep.med.harvard.edu/Polonator.

Strip fluorescent complexes from the beads

9. Prepare stripping reaction mix by combining:

   300 μl TE buffer, pH 8.0
   6 μl USER enzyme.

   Stripping is performed by reacting the array twice with the stripping solution, hence the volume is double that of a standard reaction. This can be adjusted according to the specific performance observed. If stripping is observed to be complete after one round of reaction and heating, the second stripping can be eliminated.

10. Add wash 1 and the stripping mix to the appropriate wells of the autosampler. Perform stripping protocol as follows:

    incubate at 37°C for 5 min
    incubate at 56°C for 1 min
    wash
    incubate at 37°C for 5 min
    wash.

    The autosampler software for stripping implements this protocol by default.

Query next position

11. Repeat steps 4 to 8 using a different combination of either anchor primer or fluorescent nonamer or both for up to 26 cycles, stripping the fluorescent complexes as in steps 9 and 10 between sequencing cycles.

    The read length limit of T4 DNA ligase is 6 or 7 bp from the ligation junction, for a total of 13 bp per tag.

SUPPORT PROTOCOL

TITRATION OF TEMPLATE FOR CLONAL AMPLIFICATION BY ePCR

Before performing emulsion PCR on a new sample for the first time, the appropriate amount of template must be determined empirically. Since emulsion PCR generates clonally amplified beads by a limiting dilution of template, one can be sure most beads are clonal by finding the template concentration at which approximately 15% to 20% of beads bear an amplicon. Generally, the correct amount of template is somewhere near 1 fmol per reaction. Sample four different concentrations and, if necessary, interpolate to determine the final amount. For example, given that the concentration of the library material is known, amounts of 5 fmol, 1 fmol, 0.2 fmol, and 0.05 fmol per reaction should be tried. Somewhere in this range, a linear increase in amplified beads will be observed with a linear increase in template.

After amplification of the four template samples, bead arrays are cast on microscope slides instead of coverslips (see Basic Protocol 4) for ease of handling. Amplified beads are identified by hybridizing a fluorescently labeled probe complementary to a sequence present in the central (T30) sequence. The percent of amplified beads can then be determined by fluorescence microscopy.

Additional Materials *(also see Basic Protocols 2 and 4)*

Fluorescently labeled probes:
100 μM T30-P2-Cy5-A: 5′-Cy5-AGUUGGACGUACGGCC-3′
100 μM T30-P2-Cy5-B: 5′-Cy5-AGUCGGAGGUCAAGGC-3′
Heat block or slide thermal cycler
Microscope for bright-field and fluorescence microscopy, with CCD camera and filter set for Cy5

Amplify template dilutions

1. Prepare four emulsion PCR reactions as in Basic Protocol 2, using four different amounts of template DNA in the aqueous phase (step 8). Scale down each reaction twelve-fold, so that each template amount is prepared in a single cryogenic vial (step 9) and amplified in a single eight-tube strip (step 14).

   This will result in four strips of beads, one per template amount tested. Be sure to scale down the amount of DNA used because each reaction is one-twelfth the size of a standard emulsion PCR reaction.

2. When recovering the beads from the emulsion, be sure to not combine the four strips as in step 23.

   The result will be 4 tubes of beads.

3. Resuspend each tube in 6.5 μl TE buffer, pH 8.0, and proceed to labeling of amplicons without performing enrichment.

Cast bead arrays

4. Cast each sample as a separate bead array on a Teflon-coated microscope slide. Use the same procedure as in Basic Protocol 4, but treat the microscope slide with Bind Silane instead of the coverslip, and invert the stack during polymerization so that the beads are immobilized next to the coverslip instead of the slide.

Hybridize fluorescent probe

5. Prepare the hybridization mixture:

   100 μl of 6× SSPE with 0.01% Triton X-100
   1.0 μl of 100 μM T30-P2-Cy5-A (5′-Cy5-AGUUGGACGUACGGCC-3′)
   1.0 μl of 100 μM T30-P2-Cy5-B (5′-Cy5-AGUCGGAGGUCAAGGC-3′)

6. Dry the Teflon surface of the slide with a Kimwipe, being careful to not touch the gel.
7. Pipet the hybridization mixture onto the gel.

   The mixture will stay contained on top of the gel if the slide was dried properly.
8. Place the slide face up on a heat block or slide thermal cycler and incubate at 56°C for 5 min.
9. Wash for 5 min in fresh wash 1 to remove free fluorescent probe.

Identify amplified beads

10. Place an array under a fluorescence microscope and acquire an image under transmitted bright-field illumination.

    Magnification of 20× is sufficient for the pixel size of most CCD cameras. Long working distance optics with relatively low NA are sufficient when imaging bead arrays at low density (~20,000 beads per field of view). When higher densities are desired (~150,000 to 200,000 per field of view) high-NA dry optics should be used.

    Depending on the particular optics used, the beads will appear as black circles or as white points surrounded by dark halos.
11. Acquire a fluorescence image at the same position using appropriate excitation and emission filters for Cy5.
12. Superimpose the two images in false color and determine the percentage of beads that are labeled.
13. Calculate the amount of library material to use in each ePCR based on the percentage of beads that bear labeled amplicons.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see ***APPENDIX 2****; for suppliers, see* ***APPENDIX 4****.*

Bind and wash buffer

5 mM Tris·Cl, pH 7.5 (*APPENDIX 2*)
0.5 mM EDTA, pH 8.0 (*APPENDIX 2*)
1.0 M NaCl
Store up to 6 months at room temperature

NXS buffer

10 mM Tris·Cl, pH 7.5 (*APPENDIX 2*)
1 mM EDTA, pH 8.0 (*APPENDIX 2*)
100 mM NaCl
1% Triton X-100
1% (w/v) SDS
Mix ingredients thoroughly
Store up to 6 months at room temperature

Span 80, 10% (v/v)

10% (v/v) Span 80 in light mineral oil

Measure ingredients by positive-displacement syringe or pipet. Vortex to mix thoroughly and incubate under vacuum to remove air bubbles. Store up to 1 week at room temperature.

Wash 1

10 mM Tris·Cl, pH 7.5 (*APPENDIX 2*)
50 mM KCl
2 mM EDTA, pH 8.0 (*APPENDIX 2*)
0.01% (v/v) Triton X-100
Mix ingredients thoroughly
Store up to 6 months at room temperature

COMMENTARY

Background Information

Since the dideoxy Sanger method (Sanger et al., 1977) was introduced almost 30 years ago, DNA sequencing has become a standard (albeit expensive) tool in the biologist's repertoire. With the list of reference genome sequences growing, attention is shifting toward the study of how a reference genome differs from derivative experimental genomes of interest (e.g., adaptive evolution of bacterial strains, drug resistance/sensitivity). In many cases, to fully enable these studies, the cost per quality base of DNA sequence must be lower than that obtained with state-of-the-art Sanger sequencing.

Polony sequencing aims to address this need by leveraging highly multiplex processes during both sample preparation and sequencing itself (Shendure et al., 2005). This allows an effective decrease in reagent volumes per reaction from the microliter scale to the femtoliter scale, while also dramatically reducing data acquisition time per base. These reductions in time and materials translate to a lower cost per base and, as an added bonus, higher throughput. As a cyclic array method (for a more thorough discussion, see Shendure et al., 2004), sequence is generated base-by-base through successive cycles of ligation and imaging. Each cycle of ligation and imaging sequences a single base in the template. Each 17- to 18-bp genomic tag in the sequencing template yields 13 bp of sequence when queried by T4 DNA ligase (with the central 4 to 5 bp of the tag remaining unsequenced). The result of a sequencing run is one or more million discontiguous reads of 26 bp each. While each is much shorter than a Sanger read, for many applications this length is sufficient.

Polony sequencing derives its name from the sequencing of DNA molecules that are clonally amplified in or on a solid matrix. Each amplified clone represents a polymerase colony or "polony." Polonies can be amplified from template molecules immobilized in a polyacrylamide matrix (Mitra and Church, 1999) or on beads (Diehl et al., 2006). In this unit, polonies are generated from a genomic paired-tag library on magnetic beads and are then enriched by hybridization to a common sequence. The enriched beads are cast in an array in a polyacrylamide matrix, where they are sequenced.

The focus of this unit is genomic resequencing. By using a reference genome as a scaffold onto which reads are placed and interrogated for differences, the genome in question can be effectively resequenced using much shorter reads than if a reference did not exist. Additionally, paired tags allow for placement of reads too short to otherwise span repetitive regions. All that is required is that a read-pair be unique enough, given its intertag distance distribution, that it places to only one location on the reference genome. Shendure et al. (2005) demonstrated that 90% coverage of the *E. coli* genome is feasible with paired 13-bp reads having an intertag distance distribution of 1 kb $\pm$ 250 bp. In addition to genomic resequencing, polony technology in general has been used for haplotyping (Mitra et al., 2003; Turner et al., 2006; Zhang et al., 2006) and splicing analysis (Zhu et al., 2003).

Critical Parameters and Troubleshooting

Paired-tag library

The key to successful execution of the library protocol is to achieve yields at each step at least as high as indicated in the protocols. In particular, gel purifications, ethanol precipitations, and column purifications must be reproducibly efficient. DNA quantification should be performed as indicated in the protocol, because NanoDrop readings are not reliable in the authors' hands following ethanol precipitation.

The most critical step in the paired-tag library protocol is the circularization of the genomic fragment around the synthetic linker-insert. Shendure et al. (2005) found this to be the rate-limiting step in terms of the number of unique sequencing templates that were

generated in the library construction. This library protocol is inefficient from a DNA mass standpoint because it represents the product of a number of inefficient individual steps. Additionally, the release of a pair of 18-bp tags from each fragment following circularization results in a nominal mass loss of ~96% for 1000-bp fragments (linearly worse for longer fragments). This loss necessitates the hyperbranched rolling-circle amplification performed after circularization.

Emulsion PCR

When performing emulsion PCR, it is critical not to deviate in any way from the protocol provided until it works to the point of producing clonal, strongly amplified beads from stable emulsions. Once reliable, the method can be used as a control in further optimizations. Inconsistent stability from one batch to the next is usually due to either inconsistent measuring of reagents (especially detergents) or using oil phase components (e.g., 10% Span 80) that were not freshly prepared.

Reverse amplification primer (nonbiotinylated) is present at a very high concentration in the reaction. This primer must be HPLC-purified to prevent formation of primer-dimer product on beads (which may also bear library molecule amplicons). Although primer-dimer formation is also a function of sequence, the PR1 sequences yield especially strong ePCR amplicons relative to several other primer pairs.

The interested reader is directed to excellent papers (Dressman et al., 2003; Diehl et al., 2005, 2006; Li et al., 2006) which cover the technique of emulsion PCR on beads in more detail.

Casting a polony bead array

When preparing Bind Silane–treated slides or coverslips, be sure to use Bind Silane which has not reached its expiration date. Use Erie Teflon-masked slides as the gel molds. Their Teflon coating is unusually hydrophobic and uniform in surface flatness, both of which are necessary to keep the gel from leaking out the top of the circular well before polymerization.

Be certain that no bubbles are present in the gel after it has been spread between the slide and the coverslip. A small bubble noticed in the gel after lowering the coverslip onto the slide can be removed by gently lifting the edge of the coverslip slightly with the edge of a razor blade to move the bubble to the edge of the gel and slowly lowering the coverslip back onto the slide. If this process is performed carefully, the gel will not spread outside the Teflon well.

A jig can be used to hold the coverslip in the correct position relative to the Teflon well of the microscope slide. A 0.5-in.-thick block of polyethylene can be machined with a shallow pocket that will constrain movement of both the slide and the coverslip. This ensures that the array always forms in the center of the coverslip, and thus is always properly placed in the center of the flow cell chamber.

Software, hardware, and sequence analysis

To perform polony sequencing using the protocols presented here, repetitive cycles of sequencing biochemistry and imaging are performed until the desired read length is reached (e.g., 26 cycles for 26 bp). Then, fluorescence data must be extracted from the ~50,000 2-MB images generated by the instrument, normalized, and converted to sequence. In the case of genomic resequencing, these short reads must be placed onto (or matched to) a reference genome to determine which positions differ. The process is summarized below.

The polony sequencing instrument is an inverted epifluorescence microscope which has motorized, encoded *x*, *y*, and *z* axes, as well as motorized fluorescence excitation and emission wavelength selection. The motorized features are controlled by a series of C++ binary files (available with source codes at *http://arep.med.harvard.edu/Polonator*). Attached to the microscope is the fluidics/temperature control system, which performs the biochemistry in the flow cell attached to the automated stage. It is controlled by a series of Matlab scripts, also available at *http://arep.med.harvard.edu/Polonator*.

A list of the major components necessary to build a duplicate of the instrument used by the authors is shown in Table 7.8.1; a full parts list (including fittings, tubing, and details for assembly) is available at *http://arep.med.harvard.edu/Polonator*. In building an instrument, the recommendation is to make no substitutions for parts, even when the particular components seem to have similar performance characteristics. Subtle features present in each component listed are in many cases necessary for the system to function as expected. For example, Nikon is the only microscope manufacturer that manufactures a 20× plan apochromatic objective with both a phase contrast ring (necessary for imaging densely packed bead arrays) and a 1.0-mm working distance (necessary for adequate clearance around the flow cell).

Table 7.8.1 Parts List for Automated Microscope Used in Polony Sequencing

Quantity	Part no.	Manufacturer	Description
1	86022-SPR	Chroma	Cy3/Cy5 polychroic filter w/matched individual exciters/emitters
1	41004*	Chroma	Custom TxRed filter set (S572/23x, Q595LP, S630/60m)
2	CFL	Chroma	Empty filter cube for TE2000
1	91002-02	Hamamatsu	Camera, 1000x1000, electron multiplied, 30 fps readout
1	500-H117E1T4	Prior	Stage, motorized encoded *x,y* for TE2000; 1-mm screw pitch
1	500-H30XY-E	Prior	Controller for *x,y* stage w/encoder connectors
1	500-CS152V2	Prior	Joystick for manual *x,y* control
1	LB-LS/OF30	Sutter	Lambda light source, 300 W xenon
1	LLG	Sutter	Liquid light guide for LS, 2 meters
1	LG-N27	Sutter	Adapter to attach light guide to TE2000 (no epi arm)
1	LB10-3	Sutter	Control unit for 3 wheels and 2 shutters
1	LB10-NWIQ	Sutter	10-position wheel for 25-mm filters with smart shutter
1	LB10-NEW	Sutter	10-position wheel for 25-mm filters - setup for emission
1	10-N27-EM	Sutter	Adapter to attach wheel to side port of TE2000
1	IQ25-SA	Sutter	Smart shutter for transmitted light
1	10-N27-EC	Sutter	Adapter to attach transmitted light shutter to Nikon HMX-II
1	O661301	Sutter	Extra xenon bulb, 300 W
1			Pifoc 400-µm travel piezo positioner w/controller
1	MEA51001	Nikon	TE2000-E2 inverted microscope
1	MEF55010	Nikon	T-HUBC hub controller
1	MEF55000	Nikon	T-RP remote control pad
1	MPF52061	Nikon	Universal power supply, 110-240V
1	MAE15001	Nikon	Lamphouse HMX2
1	MEE59900	Nikon	T-DH dia-illuminator, 100W
1	MAE15002	Nikon	Collector lens for QH100
1	MBF13240	Nikon	Lamp socket F/HMX 100W QI
1	MXA20425	Nikon	FXA-HMX2 adpt f/halogen
1	84125	Nikon	Microphot-FXA L.L. 12V-100W BU
1	91141-IN	Nikon	T-BP R100 optical path prism
1	MEF42252	Nikon	TE2-PS 100W power supply

continued

Table 7.8.1 Parts List for Automated Microscope Used in Polony Sequencing, *continued*

Quantity	Part no.	Manufacturer	Description
2	79035	Nikon	Power cord
1	91155	Nikon	Null modem cable
1	MEP51300	Nikon	T-N6 sextuple nosepiece
1	MBN11710	Nikon	Filter 45mm NCB11
1	MEL3000	Nikon	Diaphot condenser turret
1	MEL36200	Nikon	Diaphot condenser Lwd lens
1	MEH41200	Nikon	Te-C Lwd Ph2 module
1	MEV51100	Nikon	T-FLMC motorized cassette holder
1	MRD30200	Nikon	Cfi plan apo dm 20× objective
1	060319-2-0303-20	Bioptechs	Temperature controller, no alarm
1	060319-2-03-30	Bioptechs	FCS2 chamber only, 30-mm aperture
1	060319-2-1242HG	Bioptechs	Low dead volume top, 3-in. (0.020-in. i.d.) stainless-steel tubing, drilled
1	060319-2-0049	Bioptechs	FCS2 connector assembly, modified for slide only
5	130119-5-292	Bioptechs	Specialty microaqueduct slide - bus bars, 2 holes, no grooves
20	1907-250	Bioptechs	Shiny gasket, harvard die 452458
1	40-1313-0319	Bioptechs	40-mm round coverslips, 250/pk
1	A	HMS machine shop	Fluidics bulkhead[a]
2	B	HMS machine shop	Flow cell support arms[a]
1	C	HMS machine shop	A/D I/O module[a]
1	D	HMS machine shop	Round coverslip caddy[a]
1	AL719-P	Alcott Chromatography	Harvard autosampler
1	3000XL	Cavro	Syringe pump: RS232 1/4-28 3-port, 500-µl

[a]Drawings and schematics available at *http://arep.med.harvard.edu/Polonator/2006*.

Neither of these specifications is advertised as important by most vendors, but for this particular application, both are indeed important.

At the start of the instrument run, a C++ binary file is executed on the computer attached to the microscope to determine the proper focal position for each raster location. In this way, during each subsequent cycle of imaging, the microscope can quickly acquire images from all 500 to 700 raster locations without the need to refocus.

At the conclusion of the sequencing run, images from each cycle must be aligned to each other such that bead pixels from one image correspond to the same bead-pixels in all other images from that same raster location. Bright-field images are then thresholded and used to define the image pixels which coincide with beads. The mean intensity of each bead in each image is computed and saved to disk. These fluorescence values are normalized and converted to four-dimensional unit vectors. When plotted in four-dimensional space, the

data for each cycle will naturally form four clusters corresponding to the four base identities (A, C, G, or T). Bases are called by assigning each bead to the base identity of the cluster to which it is nearest by Euclidean distance. The quality score of each call is the Euclidean distance from the bead to the center of its assigned cluster.

To determine which sequenced bases differ from bases in the reference genome, the polony reads must be placed on the reference sequence while allowing for mismatches. The output of this algorithm is the list of input polony reads annotated with the genomic position to which each places, if any. The critical constraint is that a given read only place to one location on the reference genome, such that bases in the experimental genome can be unambiguously assigned. Once a list of placements has been computed, the experimental genome sequence can be easily derived by sorting the reads by location and reporting the call at each location. When more than one base is called at a given location (e.g., from two different reads), quality scores of each raw base call can be used to resolve the ambiguity.

Anticipated Results

The raw data produced by each sequencing run is ~50,000 2-MB images. These are processed automatically by software available at *http://arep.med.harvard.edu/Polonator* to yield several million 26-bp reads, each annotated with quality scores for each raw base. However, the number of reads produced depends on the number of microbeads arrayed on the coverglass and the fraction of those beads that bear useable amplicons.

Time Considerations

Library construction takes ~1 week if long incubations are performed overnight. The protocol can be stopped, if necessary, after any purification step. Titration of template concentration for emulsion PCR takes 2 days. Emulsion PCR takes part of two days. In experienced hands, two plates of emulsion PCR can be prepared, recovered, and enriched in 2 days by one person. The sequencing run takes 3 hours to set up. Each sequencing cycle takes 1.75 hr, for a total of 49 hours if performed continuously.

Literature Cited

Diehl, F., Li, M., Dressman, D., He, Y., Shen, D., Szabo, S., Diaz, L.A. Jr., Goodman, S.N., David, K.A., Juhl, H., Kinzler, K.W., and Vogelstein, B. 2005. Detection and quantification of mutations in the plasma of patients with colorectal tumors. *Proc. Natl. Acad. Sci. U.S.A.* 102:16368-16373.

Diehl, F., Li, M., He, Y., Kinsler, K.W., Vogelstein, B., and Dressman, D. 2006. BEAMing: Single-molecule PCR on microparticles in water-in-oil emulsions. *Nat. Methods* 3:551-559.

Dressman, D., Yan, H., Traverso, G., Kinzler, K.W., and Vogelstein, B. 2003. Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. *Proc. Natl. Acad. Sci. U.S.A.* 100:8817-8822.

Li, M., Diehl, F., Dressman, D., Vogelstein, B., and Kinzler, K.W. 2006. BEAMing up for detection and quantification of rare sequence variants. *Nat. Methods* 3:95-97.

Mitra, R.D. and Church, G.M. 1999. In situ localized amplification and contact replication of many individual DNA molecules. *Nucleic Acids Res.* 37:e34.

Mitra, R.D., Butty, V.L., Shendure, J., Benjamin, R., Williams, B.R., Housman, D.E., and Church, G.M. 2003. Digital genotyping and haplotyping with polymerase colonies. *Proc. Natl. Acad. Sci. U.S.A.* 100:5926-5931.

Sanger, F., Nicklen, S., and Coulson, A.R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* 74:5463-5467.

Shendure, J., Mitra, R.D., Varma, C., and Church, G.M. 2004. Advanced sequencing technologies: Methods and goals. *Nat. Rev. Genet.* 5:335-344.

Shendure, J., Porreca, G.J., Reppas, N.B., Lin, X., McCutcheon, J.P., Rosenbaum, A.M., Wang, M.D., Zhang, K., Mitra, R.D., and Church, G.M. 2005. Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* 309:1728-1732.

Siebert, P.D., Chenchik, A., Kellogg, D.E., Lukyanov, K.A., and Lukyanov, S.A. 1995. An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Res.* 23:1087-1088.

Turner, D.J., Shendure, J., Porreca, G., Church, G., Green, P., Tyler-Smith, C., and Hurles, M.E. 2006. Assaying chromosomal inversions by single-molecule haplotyping. *Nat. Methods* 3:439-445.

Zhang, K., Zhu, J., Shendure, J., Porreca, G.J., Aach, J.D., Mitra, R.D., and Church, G.M. 2006. Long-range polony haplotyping of individual human chromosome molecules. *Nat Genet.* 38:382-387.

Zhu, J., Shendure, J., Mitra, R.D., and Church, G.M. 2003. Single molecule profiling of alternative pre-mRNA splicing. *Science* 301:836-838 .

Contributed by Gregory J. Porreca, Jay Shendure, and George M. Church
Harvard Medical School
Boston, Massachusetts