Amplicon Sequencing Protocol for Genome Targeting

Protocol by: Lisa Kanizay and Tom Jacobs, Parrott Lab

Supplies:

- Any PCR-grade genomic DNA, CTAB is fine
- PCR and Gel purification kits (we use Zymo, I imagine Qiagen or some other ones would work fine)
- KAPA HiFi HotStart ReadyMix (KAPA KK2601) or other high-fidelity enzyme
- KAPA Illumina library quantification kit (KK4903 and KK4953) or other qPCR library-quantification reagents

Primers and Primer Design Considerations:

First PCR: here you need to order gene-specific primers with portions of Illumina adapters on them. We use TruSeq-style Illumina adapters. Design is modified from (Faircloth and Glenn, 2012).

Illumina adapter to add onto the 5' end of your forward gene specific primer:

5' - AACTCTTTTCCTACACGACGCTCTTCCGATCT -3'

Illumina adapter to add onto the 5' end of your reverse gene specific primer:

5' - CTGGAGTTTCAGACGTGCTCTTCCGATC -3'

Second PCR: uses primers to the portions of adapters you added in the first PCR (blue and red) and adds the barcodes and the rest of the tags.

Final F primer:

5' - AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC -3'

Barcode R primer:

5' - CAAGCAGAAGACGGCATACGAGATNNNNNNGTGACTGGAGTTTCAGACGT -3'

You need to add a 6nt barcode to the R primer where the Ns are. Here are 10 barcodes (correspond to tags 0-9):

TATACT
CAACAA
GTTGTT
TCGGTT
AGTATT
TTAATT
TCTTGT
GAGTGT
AATGGT
GGCGGT
Note: To de-multiplex with Miseq reporter (or similar software) the barcodes will need to be reverse complemented in the sample sheet.

Protocol:

1. Amplify the first PCR with the tailed primers using a high-fidelity enzyme mix. We like to use KAPA HiFi2x Mastermix since it is very efficient. For each sample I do a 10µl reaction with 1µl DNA (I never quantify), 0.3µl of each primer at 10µM, 2µl master mix and 3.4µl water. Cycling conditions will depend on the enzyme mix used, for KAPA it requires a minimum annealing temp of 60°C. I try to keep the cycles low, mid to high 20's, but some amplicons need more to get enough product. I suggest trying a few reactions first, and see how many cycles you need. Check 4µl of the reaction on a gel to make sure it amplified, faint bands are fine.

2. The first PCR products are diluted and pooled. Pool across amplicons, so that in pool 1 will be event 1 amplicon A, event 1 amplicon B, event 1 amplicon C... etc. It is OK to have different numbers of amplicons in a pool. Based on the intensity of the gels approximate how much PCR product to add. If the bands are strong, 1µl, weak 2µl, barely visible 3 or 4µl. This part is not terribly accurate, but it seems to work fine. The PCR products are diluted into 100µl of water. If there are more than 7 or 8 amplicons, try 200µl or more.

3. The second PCR uses the pooled samples as a template. One index primer per pool. Dilute the index primers to 1µM to make pipetting easier (need to add 3µl to the mix). The Final F primer is used at normal conditions. PCR is similar as above, 1µl of pool as template. I only do 10 cycles of PCR for this reaction. Check 4µl on a gel. There should be a band for each amplicon in the pool (good check for cross contamination). Then pool the remaining PCR products together (everything is indexed), column purify, and elute in a small volume (~20µl). Load everything in a 1.5% agarose, TAE + cytidine gel. Primer dimers will likely be present, so a higher agarose concentration is important. However, too much agarose and the gel extraction is a pain, so I find the 1.5% a good compromise. Then just run the gel, cut out the correct range, purify and nanodrop. The library is then ready to go. We quantify the libraries with the KAPA quant kit; it has been more reliable than the Agilent chips.

Other comments and notes

Gene-specific primer design:

I prefer to have the forward primer 50-200 bp away from the target site. I then have a single-end sequencing run to read ~50 bp past the target site (if possible). You can purchase a Miseq 2x150 kit and tell it to run a single read for 300 bases. I mainly do this because for this type of analysis because a single read is sufficient, read1 is higher quality than read2, and the sequencer runs faster. If only a reverse primer can be designed within this 50-200bp range, just switch the tails on the gene-specific primers for the first PCR.
PCR products from gene-specific primers should be 200-700 bp in length. Too small, and it’s harder to separate the products from primer-dimers. Too big, and the amplicons are less efficient during PCR and have more difficulty binding the flowcell (products over ~1.2kb will not bind the flowcell).

Pooling:

After the first PCR, samples can be pooled across amplicons. For example, for the second PCR, the first tube will contain: event 1, amplicon a; event 1, amplicon b; event 1, amplicon c; etc. After the second PCR, all samples can then be pooled. After sequencing, the Miseq reporter will separate the reads by index, and then the reads can be separate by gene-specific primer.

Adapter design:

This is based on the Truseq Illumina primer design with a single indexing read of 6-nt (Illumina calls this the low-throughput design). The index read is up to you, and there are 10-nt designs available. Also, this method could be adapted to a dual-indexing scheme.

We have tried mixing this sequencing library with a dual-indexed library and have had difficulty analyzing the results. Generally, Illumina does not recommend mixing single-index and dual-index libraries. My suggestion would be to run this library with another single-end, 6-nt index library.

Sequencing:

If this library is to be run by itself, it should be done on a Miseq. This is technically a low-diversity library, and the Nextseq500 and Hiseq cannot handle low diversity as well as a Miseq. It is also important to use at least a 10% spike-in of PhiX control (FC-110-3001) to get better quality scores.

This library can also be spiked into other sequencing runs. I typically aim for a few thousand to tens of thousands reads per amplicon/event. So depending on the number of amplicons/events you have, a million reads on a Hiseq or Nextseq500 is no big deal. I bet other users wouldn’t mind you spiking in some sample (for low cost, or free). We have had difficulty spiking this library design into a dual-indexed read, so be sure to only mix with a read that has the same index design.