NimbleGen Sequence Capture Technology Enables Targeted Sequencing

TECHNICAL NOTE

Roche NimbleGen has developed NimbleGen Sequence Capture technology that enables targeted sequencing of thousands of exons or contiguous genomic loci of up to 5Mb in a single experiment. The microarray hybridization-based NimbleGen Sequence Capture technology has considerable cost, throughput, and quality advantages when compared to PCR.

Introduction

Recent technological advances in sequencing systems have rapidly increased raw sequence output. However, these next-generation sequencing systems do not currently have the throughput to sequence the whole human genome cost-effectively. Thus, they require that the complexity of genomic DNA samples be reduced to a manageable subset prior to sequencing. The prevailing method for complexity reduction has been the preparation of amplicons by parallel, multiplex, or long range PCR amplification. These PCR methods have severe cost and performance limitations when scaled to the level required to take full advantage of the capacity of currently available sequencing systems. As a result of these limitations, the bottleneck for sequencing projects has shifted to sample preparation.

To address this sample preparation bottleneck, Roche NimbleGen has developed the microarray hybridization-based NimbleGen Sequence Capture technology that utilizes high-density oligonucleotide microarrays as a programmable genomic selection device to allow targeted sequencing of subsets of the genome. These genome subsets can be exons, disease-associated regions, quantitative trait loci, promoters and enhancers, and other targeted regions. The revolutionary and simple workflow of NimbleGen Sequence Capture technology enables isolating megabase regions in as little as one week and eliminates the cost, labor, and infrastructure required for large-scale PCR experiments.

Sequence Capture Advantages

- NimbleGen Sequence Capture technology greatly simplifies and streamlines the process of isolating target regions for high-throughput sequencing. A single microarray experiment using this technology will select up to 5Mb of genomic sequences.
- NimbleGen Maskless Array Synthesis (MAS) technology and array design pipeline allow arrays to be easily produced for any desired region. This ability eliminates the time and labor required to design specific PCR primer pairs for each target region as well as the high cost of PCR primers when many regions are targeted.
- An optimized protocol can be applied to capture any target region. The labor intensive and time-consuming step of optimizing parameters for long range PCRs is completely eliminated.
- The hybridization and elution instruments are easy to set up. There is no need to invest in expensive infrastructure for high-throughput PCR experiments.

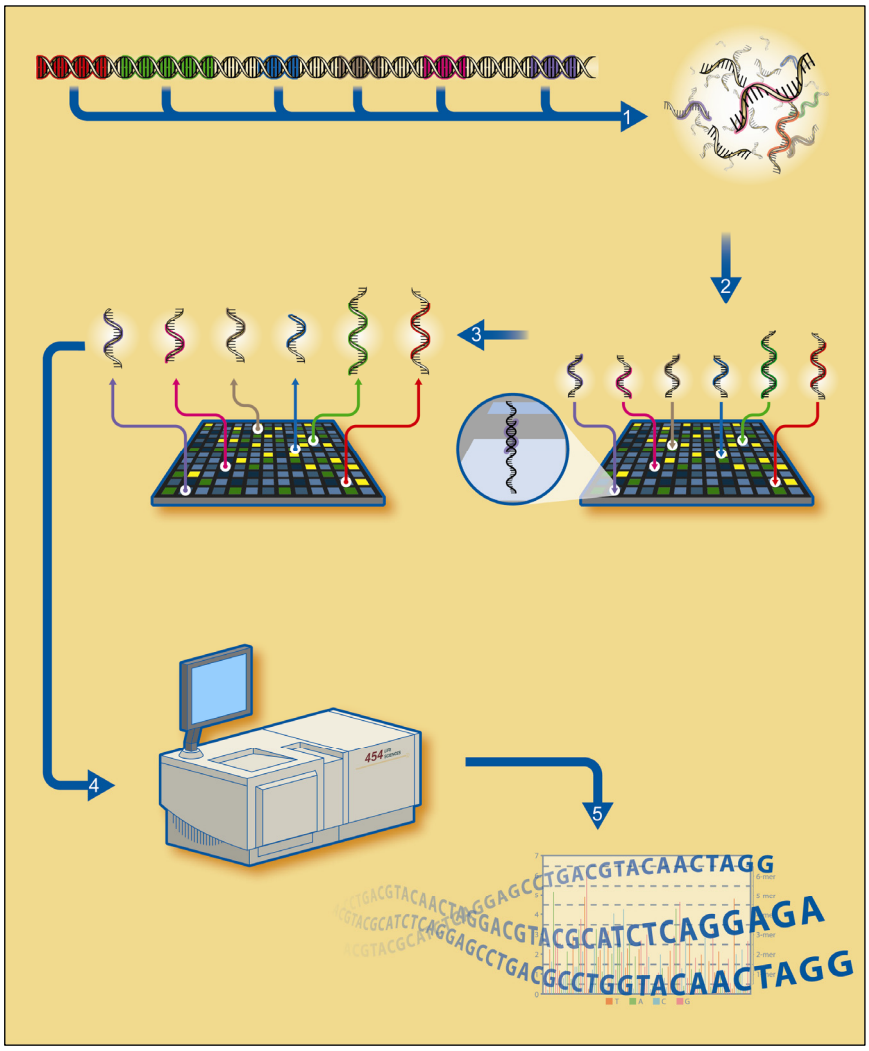


Figure 1. The NimbleGen Sequence Capture Workflow

1. *The genomic DNA is sonicated into small fragments, and linkers are added to both ends.*
2. *The genomic DNA sample is hybridized to the array designed and manufactured per customer specifications by Roche NimbleGen. After stringent washing, the genomic fraction bound to the array is greatly enriched for regions targeted by the array's probes.*
3. *An elution step is performed to release the selected DNA fragments from the array.*
4. *The eluted DNA fragments are amplified using the linker sequence if needed. The selected DNA is then ready to be sequenced using a high-throughput sequencing system, like the 454 Genome Sequencer FLX.*
5. *The high-throughput sequencer produces ~100Mb raw reads, most of which are from the targeted regions.*

Sequence Capture Performance

The protocols for the NimbleGen Sequence Capture technology are in the process of being refined and optimized for benchmark performance parameters. The performance data shown here are derived from the first peer-reviewed publication (1) of the technology and, as such, do not represent the fully optimized protocols that are under development. The commercial release of microarrays and services for NimbleGen Sequence Capture technology are expected to substantially exceed the performance shown here in terms of sequence capacity, coverage, and ease of use. To receive the latest updates on the technology's performance, we encourage you to subscribe to NimbleGen Sequence Capture news at www.nimblegen.com/seqcap.

- With a single-array experiment covering 5Mb target regions, followed by a single 454 Genome Sequencer FLX run producing ~100Mb total sequencing data, the majority of sequencing reads represented selected target regions (typically >70%). While the coverage typically depends on the composition of the target regions, ~8X median coverage can be achieved for exon-sized regions (Figure 2), and ~18X median coverage can be achieved for a single 5Mb contiguous genomic region (Figure 3). The sequence coverage will increase for smaller cumulative target region sizes, if the same amount of sequencing runs is performed. For example, an experiment covering a 500kb contiguous target region followed by a single 454 Genome Sequencer FLX run yielded >90X median coverage. In research requiring small regions, NimbleGen Sequence Capture technology combined with 454 sequencing utilizing only a portion (1/2 to 1/16) of the picotiter plate will generate sufficient reads for sequencing applications.
- Although arrays are typically designed based on the reference genome, NimbleGen Sequence Capture technology does not bias against discovery of unknown variants. As shown in Figure 4, in an experiment designed to compare the performance of this technology versus long range PCR, almost all the variants were captured with the same fidelity as PCR. In the 70kb region targeted by both methods, 98 SNPs were identified by both, and 9 and 5 rare variants were identified by NimbleGen Sequence Capture technology and long range PCR, respectively. In addition, 22 variants in repeats were detected only by long range PCR because no probes on the array were designed for these repetitive regions.

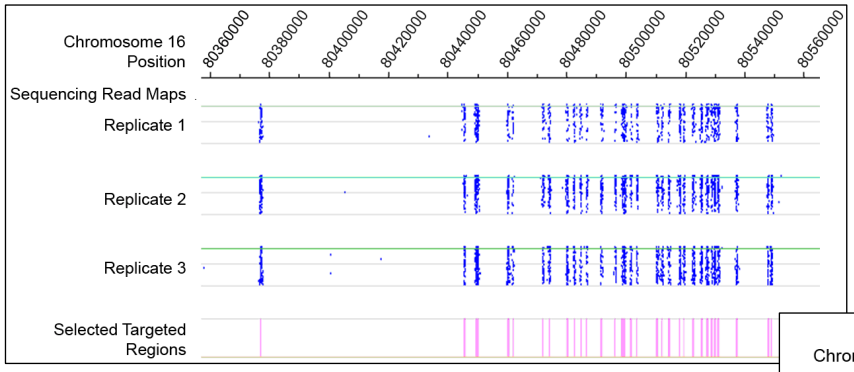
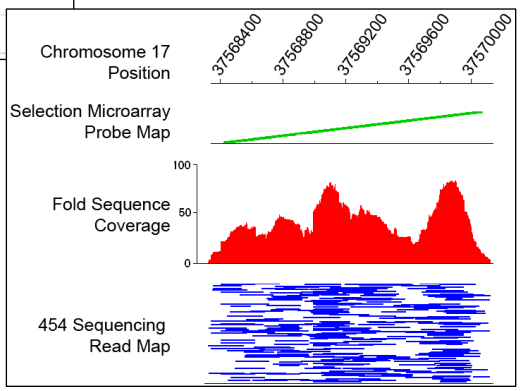


Figure 2 (above). Sequencing Read Maps of ~190kb of Chromosome 16

Sequencing capture read maps depict ~190kb of chromosome 16 from three replicates of sequence capture experiments of human exons.

Figure 3 (below). Sequencing Read Map of 2kb of Chromosome 17

A sequencing read map shows 2kb of chromosome 17 from a microarray selection of a 2Mb contiguous region that contains the BRCA1 gene.



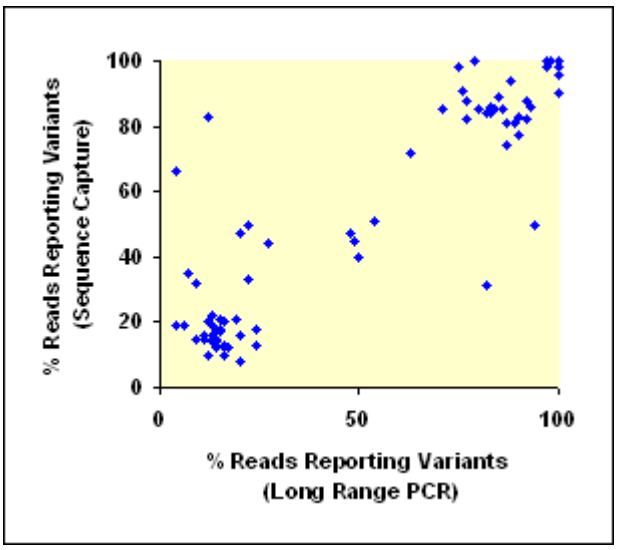


Figure 4. NimbleGen Sequence Capture Technology Showed No Bias against Novel Variants
In this experiment, genomic DNA from a mixed cell population was used, representing both common and rare SNP variants. A 200kb region surrounding the EGFR gene was captured using NimbleGen Sequence Capture technology, and a 70kb region out of the 200kb region was amplified using long range PCR. DNA samples derived from target regions using both methods were sequenced separately with the 454 Genome Sequencer FLX instrument, and SNP discovery was performed on both data sets. The percentages of 454 sequencer reads that report variants, either from PCR or NimbleGen Sequence Capture technology, were plotted for each of the SNPs detected by both methods.

Accessing Sequence Capture Technology

Roche NimbleGen will launch its NimbleGen Sequence Capture technology as a service in early 2008. Researchers will be able to send DNA samples to Roche NimbleGen and specify regions of interest. Roche NimbleGen will capture those regions and return the enriched DNA to the researcher. The DNA will be in a form that is ready for high-throughput sequencing.

Later in 2008, Roche NimbleGen will launch a kit including all necessary reagents, enabling researchers to perform their own sequence capture experiments.

References

1. Albert TJ, et al. Direct selection of human genomic loci by microarray hybridization. *Nature Methods* 2007 Nov; 4(11):903-5.
2. Okou DT, et al. Microarray-based genomic selection for high-throughput resequencing. *Nature Methods* 2007 Nov; 4(11):907-9.
3. Hodges E, et al. Genome-wide in situ exon capture for selective resequencing. *Nature Genetics* 2007 Dec; 39(12):1522-7.

For More Information

To learn more about NimbleGen Sequence Capture technology or to subscribe to Sequence Capture news for updates on this technology, contact Roche NimbleGen or refer to our web site at www.nimblegen.com/seqcap.

Toll-free in US: (877) NimbleGen / (877) 646-2534
(608) 218-7600
www.nimblegen.com

Current as of 14 February 2008

Part No. TN-SEQ0100