



## In Sequence

The Inside Read on Genome Sequencing

April 06, 2010

# NimbleGen Array Capture Outperforms Two Target-Enrichment Methods in ABRF Research Group Comparison

**ROCHE NIMBLEGEN'S ARRAY CAPTURE** method outperformed both Agilent's SureSelect and Febit's HybSelect in most categories in a recent comparison of the three target-enrichment methods for Illumina sequencing.

According to the study, conducted since last summer by the Association of Biomolecular Resource Facilities' DNA Sequencing Research Group and presented last month at the ABRF annual meeting in Sacramento, Calif., Roche NimbleGen's sensitivity, specificity, coverage, and overall capture quality were better than those of the other two methods.

However, experiments using the Agilent and Febit platforms might be repeated because of technical issues.

The goal of the evaluation was "to provide an unbiased analysis of the technologies," something small core labs might be unable to do on their own due to cost constraints, said Anoja Perera, lab manager of the Molecular Biology Facility at the Stowers Institute for Medical Research in Kansas City, Mo., who coordinated the study.

The research group chose two array-based platforms for its evaluation — Roche NimbleGen's Sequence Capture arrays and Febit's HybSelect — and one in-solution method, Agilent's SureSelect.

They did not include RainDance Technologies' microdroplet PCR platform because at the time they designed the study last summer, that method could only capture a

few hundred kilobases of DNA per experiment, according to Perera.

The group also left out other enrichment methods, such as Roche NimbleGen's SeqCap EZ in-solution product, which was not available at the time, and Olink Genomics' Selector technology, of which the researchers were not aware last summer.

The researchers selected as their sample Craig Venter's DNA, also known as HuRef, which is available from the Coriell cell repository. They chose two types of regions as their targets: a 2-megabase continuous region and 31 genes that cover about 1.5 megabases in total.

The genes, which ranged in length from 2 kilobases to 400 kilobases, differed in exon number, GC content, amount of repetitive sequence, and number of transcripts.

After Roche NimbleGen, Agilent Technologies, and Febit received the DNA sample, target list, and an Illumina paired-end sample preparation kit, the researchers asked them to design and perform a capture experiment in duplicate at their respective sites. They were also asked to return the DNA libraries for sequencing at the Stowers Institute and at the DNA core at the University of Michigan.

After checking the sample quality on an Agilent 2100 Bioanalyzer chip, these labs sequenced the samples on an Illumina GAII sequencer, each in a single paired-end run. The sequence data were analyzed independently by a team at the Stowers Institute and one at Tufts University.

The researchers found that the NimbleGen arrays captured more than 90 percent of the target bases with at least 1-fold coverage, whereas both Agilent and Febit captured only between 40 and 50 percent of the targets.

Looking at only the 2.1-megabase continuous region, they found that NimbleGen covered more than 90 percent of the target, Agilent around 70 percent, and Febit around 40 percent.



www.genomeweb.com  
The GenomeWeb Intelligence Network



April 06, 2010

## NimbleGen ...

continued from page 1

On specificity, more than 80 percent of reads from the NimbleGen capture mapped on target, compared to about 60 percent of reads from the Agilent capture and less than 5 percent of the reads from Febit's capture.

Overall depth of coverage and uniformity of coverage were also better for NimbleGen than for the other two methods. The amount of input DNA required was similar for all three methods, and Agilent and NimbleGen both showed good reproducibility.

Based on information from the companies, both Agilent's SureSelect and NimbleGen's Sequence Capture cost about \$1,000 per sample, whereas Febit costs more than \$2,000 per sample, the study noted. But unlike NimbleGen and Febit, Agilent's SureSelect does not require an instrument purchase to perform the method in house, Perera said, so Agilent scored best in terms of per-sample cost. The DSRG also found that Agilent's in-solution method is more scalable than the array-based methods.

But the results presented at the ABRF meeting are not the final word, and both Agilent and Febit might repeat the experiments because of problems with the original results.

For example, Agilent mistakenly based its probe design on an earlier human genome build — hg18 instead of hg19 — "so the study is truly not an apple-to-apple comparison," Fred Ernani, senior product manager for emerging genomic applications at Agilent, told *In Sequence*. "We are currently making the proper design and will repeat the study properly for ABRF."

Ernani said he expects the performance of the SureSelect method will be "more comparable" to the NimbleGen array platform after that.

The DSRG also believes that errors with Febit's capture and library preparation affected the results. For example, the Febit sample showed "a very large primer-dimer peak that affected the sequencing," Perera told *In Sequence*. She added that the results might have benefited from a size-selection step prior to sequencing.

Also, while some genes and genome sections were captured well, "for there to be such a high percentage of sequences that were off target, approximately 97 percent, there has to be a reason," she said.

According to Peer Stähler, Febit's chief scientific officer, the comparison did not address the "sweet spot" of the current version of HybSelect, which is designed to capture smaller targets — ideally on the order of 200 kilobases — rather than megabases.

He also said that Febit's platform "performs better" in conjunction with the SOLiD platform, and the company does not currently promote HybSelect for use with the Illumina sequencing platform.

Stähler told *In Sequence* that another round of evaluation "does make sense provided [the DSRG is] interested in the sweet spot we currently have: small but many, and using the Life Technologies SOLiD system, if available to them."

Roche NimbleGen's senior product manager for enrichment, Xinmin Zhang, told *In Sequence* that the study's results for its arrays "are consistent with and similar to" results from other customers.

Platform comparisons conducted by customers, he said, also found that NimbleGen's capture arrays score high on sensitivity and uniformity, but that specificity depends on the target region and is "more similar across platforms" than in the DSRG study.

He said that according to studies by NimbleGen and others, the arrays' high sensitivity "is directly correlated" to the company's repeat masking method, which allows the firm to place probes in more regions of the genome than other tools, such as RepeatMasker.

He attributed NimbleGen's high specificity at least in part to the hybridization and washing conditions used. One reason that the platform had the best uniformity, he suggested, is that the company uses "a proprietary design algorithm" for its probes that is based on empirical data, rather than "a simple tiling design" where probes are evenly distributed across the target region.

He also said that NimbleGen's capture arrays have more probes — between 385,000 and 2.1 million — than Agilent's SureSelect, which uses up to 55,000 long probes, which he said affords more design flexibility. Febit's Geniom Biochips currently include 120,000 probes in eight separate channels.