Optimized on-target performance

One of the most direct measurements of enrichment efficiency is analysis of the sequencing reads for percent on-target across a diverse set of genomic locations. With the SureSelect Target Enrichment System, the on-target reads for eight designs (Figure 1) is within 40 to 80 percent, which translates into enrichments of 300- to 7400-fold. Enrichment is largely dependent upon the size of the targeted region and the number of reads per sequencing run. Small-capture designs typically yield a higher level of enrichment due to the kit's efficiency in focusing sequencing reads on a smaller subset of the targeted genome.

No matter how specific the capture methodology, because of random shearing, both the targeted region and nearby DNA targets are captured. Agilent has optimized the design algorithms and protocol to limit the capture of near-target sequences. To this end, we measured the specificity of the SureSelect Target Enrichment System by analyzing sequencing reads exactly on target, within 100 base pairs (bp) of target, and within 200 bp of target. Figure 1 shows that the off-target capture rate is not substantially increased by including sequence reads within 200 bp of the target.

Read distribution and sequence coverage show ability to detect mutations

Another metric of great importance to researchers in DNA sequencing is read distribution, because it affects the ability to adequately cover genomes to identify sequence variation. There are two ways to visualize this metric:

1. Plot the distribution of actual percent of bases with a certain number of reads.
2. Plot the cumulative number of reads with at least a certain depth.

Figure 2 is a plot of a representative sample that shows both graphs. The figure demonstrates that roughly 80 percent of all sequences have at least a 20x read depth, which indicates the suitability of this method for identification of single-nucleotide polymorphisms (SNPs) with targeted resequencing. In addition, the sequence coverage is very even, showing negligible bias. In this type of scenario, one could expect approximately 95 percent or more of the targeted bases to have at least one read, 80 percent or more to have at least five reads, and 50 percent or more to have at least 20 reads, making this target enrichment method ideally suited for interrogating genomes for mutations.

Design robustness, specificity, and comprehensive sequence coverage distinguish the Agilent SureSelect Target Enrichment System from other commercially available target enrichment products. If you need to streamline DNA-sequencing research, learn more about the Agilent SureSelect Platform.
Performance of the SureSelect Target Enrichment System over a diverse set of library designs shows high target-region specificity.
Figure 2.

The SureSelect Target Enrichment System shows excellent sequence coverage, which enables reliable identification of SNPs.