How Blocking Oligos affect target enrichment performance

What is Target Capture

Next generation sequencing (NGS) is a sequencing technology where randomly fragmented genomic DNA is sequenced, assembled, and mapped for analysis. The advent of NGS-based whole genome sequencing (WGS) has greatly expanded our scope of knowledge on genetic diseases especially by accelerating the sequencing process. Nevertheless, NGS-based WGS faces many technical limitations that are overcome by innovative technologies, one of which is Target Capture.

Human genome is known to consist of 3 billion base pairs (3 Gb), and in order to achieve the average of 30 depth coverage of WGS for each sample, approximately 90 Gb of total sequencing output is required. When assessing somatic variations or sequencing from solid cancer samples or cell-free DNA (cfDNA), even higher depths are demanded from hundreds to tens of thousands for accurate analysis. But such is a challenge, as higher sequencing amount is followed by high costs and has to be supported by robust computing power for data analysis. Also, NGS-based WGS is considered not the most effective sequencing test because only a small fraction of the result data represents the variants and genes associated with human diseases.

Targeted sequencing overcomes such major limitations of WGS analysis. It targets selected sets of genes of interest out of the whole genome. While WGS demands massive sequencing output with high costs, only a small amount of sequencing is sufficient for targeted sequencing to achieve higher depth and accurate analysis with lower costs. Because of these benefits, targeted sequencing has become widely used in recent years for research and clinical diagnoses in various fields such as genetic diseases and cancer.

Celemics, Inc. has seamlessly incorporated hybridization capture method in targeted sequencing to provide Target Enrichment Kit. The hybridization capture is a sequencing method where the genomic regions of interest are specifically selected by hybridization probes. The other key factors incorporated for successful targeted sequencing are probe design technology and reagent optimization. We have developed an exceptional probe design technology targeting against GC-rich, AT-rich and homologous regions that are known to be the “hard-to-capture” regions in the market. The probes used for the targeted capture are RNA probes, which have stronger binding affinity to complementary DNA sequences than DNA probes. The Celemics proprietary technology that designs blocking oligonucleotides or blocking oligos and optimized reagents is developed for the highly reliable target enrichment assay. Celemics is providing these market leading technologies at the most affordable price.

What is the role of blocking oligos in target capture

There are several key strategies integrated into the Celemics targeted sequencing panel, one of which is blocking oligos. The blocking oligos block specific DNA sequences and the most commonly used are (i) Adapter blocking oligos and ii) Cot-1 DNA blocking oligos.

Elaborately designed blocking oligos are essential for obtaining sufficient sequencing depths and coverage to analyze the genomic regions of interest, subsequently enabling high-fidelity, cost-effective sequencing.

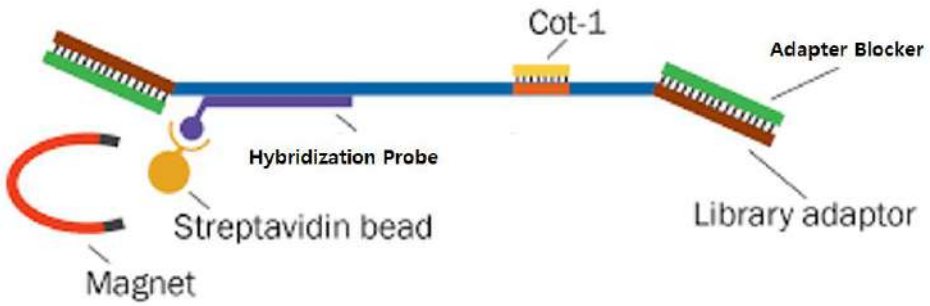


Figure 1. Schematic drawing on the mechanism of CoT-1 blocking oligos

Adapter blocking oligos

Since adapters are ligated to each DNA fragment in the library preparation process, a large portion of the total library is comprised of the adapter sequences. The blocking oligos against adapters play a crucial role in enhancing the performance of the probes. By blocking the adapter sequences, the reagent prevents daisy chains between on-target and off-target reads, and the formation of adapter dimers. A previous study states that on-target read ratio can be reduced by as much as 50% without the use of blocking oligos.1 The difference in adapter sequences by NGS platform (Illumina and Ion Torrent) can be an inconvenience to researchers. At Celemics, we design the adapter blocking oligos highly optimized to ensure the best target enrichment performance with NGS platforms.

CoT-1 DNA blocking oligos

The hybridization-based targeted sequencing enables comprehensive analysis across a wide range of targeted regions with higher uniformity compared to PCR amplicon method. More unique

reads are secured by removal of the duplicate reads generated by PCR. Non-specific binding of hybridization probes to off-target regions may be a concern and needs to be minimized in the hybridization process. For example, if the probes are designed to target a region embedded in the repeated sequences, which are frequently found in intron, CoT-1 DNA block oligos is used to block the repeated sequences prior to the hybridization process, increasing on-target ratio and enhancing sensitivity.

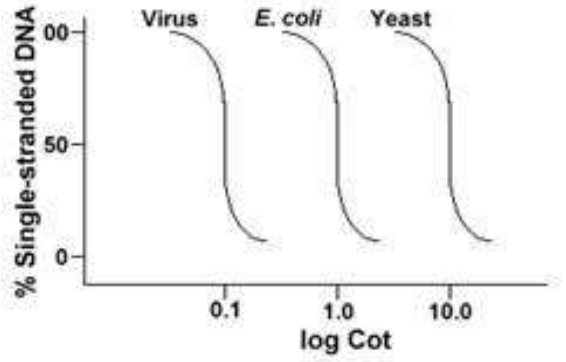
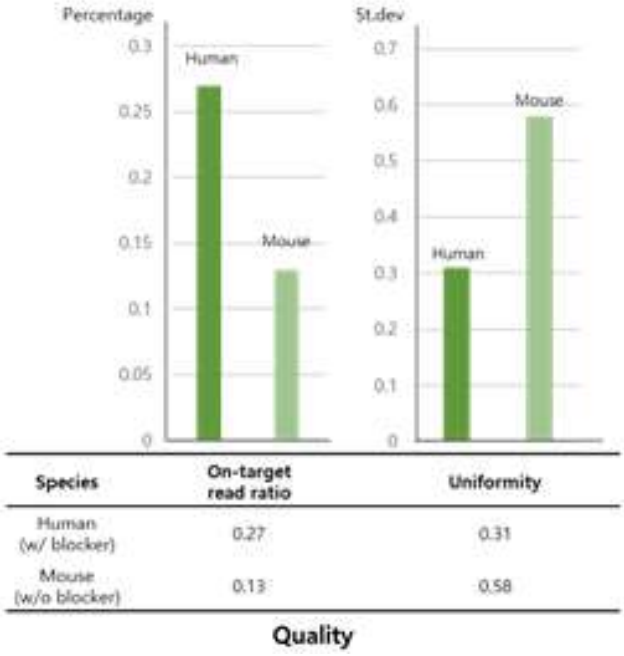


Figure 2. CoT curves of different systems such as virus, E.coli, and yeast[3,4]

The CoT-1 DNA, which has been used for microarray screening and FISH assay, is enriched for repeated and non-coding sequences commonly found in gDNA. Celemics has employed CoT-1 DNA in NGS to block nonspecific hybridization of the probes. The main binding site for CoT-1 DNA is the repetitive sequences increasing on-target read ratio in NGS analysis. A recent study elaborates on the positive effect of CoT-1 DNA, supporting increased on-target read ratio and decreased off-target read ratio.[2]



Species	On-target read ratio	Uniformity
Human (w/ blocker)	0.27	0.31
Mouse (w/o blocker)	0.13	0.58

Quality

Figure 3. Performance comparison chart between the panels with and without the blocking oligos

To achieve high performance of target enrichment, it is essential that CoT-1 DNA is designed according to the sequence complexity and the level of frequency and abundance of repetitive sequences which vary by species (Figure 2). CoT-1 DNA blocking oligos that are not designed species-specifically result in reduced on-target read ratio and uniformity (Figure 3). The sequencing data with lower on-target read ratio indicates more off-target reads that are to be discarded in the process of BI analysis, limiting the number of samples in a single NGS run thereby producing ineffective data results with high overall costs.

Conclusion

The main business of Celemics, Inc. is NGS-based Target Enrichment Kit in which proprietary design technologies for hybridization probe and blocking oligos are incorporated. We have optimized the hybridization reagents and the noise removal

technology to provide the most effective hybridization of the probes to the sample DNA with high on-target ratio. The optimization process also includes the blocking oligos. The reagents are highly optimized to each NGS platform of Illumina and Ion Torrent, and designed species-specifically to produce data with high quality.

We have provided hundreds of products for genetic diseases including solid cancer, hereditary cancer, and kits for liquid biopsy and cell line QC, to researchers from over 10 different countries.

We have recently developed and provided a target enrichment kit specifically designed for capturing African-Swine Fever Virus. The successful development of the kit was enabled by the exceptional design technology at Celemics with which the species-specific blocking oligos are elaborately designed and manufactured.