cytiva

Streamline NGS – a guide to hybrid capture target enrichment

Target enrichment is an optional step of next-generation sequencing (NGS), that stands to help you achieve greater depth and accuracy in sequencing data through the amplification and interrogation of specific disease-associated or clinically relevant aberrations or regions of interest. Several established methods can achieve this level of depth and accuracy, some of which specifically relate to your chosen NGS protocol.

cytiva.com

Traditionally, enrichment strategies used PCR or molecular inversion probes (MIPs) to target and amplify a genomic region of interest. But since the mid-2010s, more and more researchers and clinicians are opting for hybrid capture enrichment strategies because of their advantages.

Hybrid capture target enrichment allows comprehensive analysis of variant types — including single nucleotide polymorphisms (SNPs), insertion/deletions (indels), copy number variations (CNVs), and structural variations — by capturing multiple target regions in a single experiment. Capturing multiple regions is advantageous because it allows longer reads while providing greater depth and resolution than PCR and MIPs.

As such, hybrid capture enrichment enables high-resolution sequencing of specific target regions from complex biological samples and is shown to be especially powerful for monitoring rare genetic disorders, infectious diseases, and interrogating ancient DNA samples.

Magnetic beads are one tool increasingly used throughout the NGS workflow, with targeted sequencing being no exception. From DNA extraction to enrichment, magnetic beads are valuable in enabling and optimizing multiple steps in NGS sample and library preparation.

This guide discusses the importance of target enrichment in the NGS workflow and how effective hybridization-based enrichment can markedly improve downstream applications. It also provides insights on optimizing the process, including how magnetic beads could streamline your NGS workflow.

Read our scientists' guide to magnetic beads

Targeted sequencing and the benefits of target enrichment

Basic concepts

NGS technologies have revolutionized genomics, enabling advances across all life science sectors. While first-generation sequencing was limited to producing a sequence for one template per reaction, NGS can perform millions to billions of individual sequencing reactions simultaneously in a process called massively parallel sequencing (1).

Extraordinary progress continues to be made in NGS. Substantial improvements in sequencing data quality and yield have given rise to various novel sequencing approaches, some of which have proven to be more effective and adaptable than others (2). Despite the large number of unique NGS platforms commercially available, a typical NGS workflow has three common steps: sample preparation, nucleic acid sequencing, and data analysis. The specifics can vary greatly, with targeted sequencing approaches being a subset that enables analysis of specific genomic regions of interest.

Targeted sequencing approaches

The refinement of NGS has given rise to high-throughput, long-read applications such as whole-genome sequencing (WGS). With NGS, long-read sequencing can be delivered at a fraction of the cost and time needed for first-generation, or Sanger, sequencing.

The Human Genome Project cost billions of dollars, required unprecedented international collaboration, and took more than a decade to complete. Today, NGS technology can sequence a human genome at a cost of approximately $1000 in three days (3). Despite this progress, the sequencing of whole genomes or exomes has since remained at this price level for more than a decade, and the amount of data gathered continues to cause a significant bottleneck at the data analysis stage (4). The expense and data-analysis bottleneck along with known limitations in accuracy (5), makes traditional long-read sequencing unsuitable for many day-to-day applications, particularly in clinical settings where cost, speed, and accuracy are of paramount importance.

Many research and clinical NGS applications overcome the large-data burden of long-read NGS by instead interrogating small, targeted genome regions. So-called "targeted sequencing" approaches focus on specific regions within the genome. The approaches can include targets within individual or multiple genes (targeted amplicon sequencing) or entire genes of interest (targeted gene sequencing) (6).

Most clinical gene panels use a targeted gene sequencing approach to analyze genetic aberrations within genes associated with the pathogenesis of a disease or aberrations of clinical relevance. Efficient sequencing of the complete exome (all transcribed sequences) is a popular current application, but researchers and clinicians regularly focus on much smaller sets of genes or genomic regions for identifying small-scale mutations, such as in monogenetic diseases.

Targeted sequencing enables the examination of multiple samples in parallel and is especially advantageous for the detection of SNPs, CNVs, and DNA rearrangements (7). By eliminating unwanted genetic material, such as non-coding regions of the genome, it's possible to achieve greater sequencing depth and accuracy, with significantly reduced costs and data burden (8).

Library preparation and target enrichment

The first stage in a targeted-sequencing workflow is library preparation (9). For targeted NGS applications, this process is generally achieved in two steps. First, DNA is fragmented to application-specific template lengths, end-repaired, and size selected using magnetic beads, enzymes, or electrophoresis (9, 10). Second, adapters are ligated to facilitate the attachment of the DNA fragments to solid surfaces (such as glass slides, microchips, or nanowells) for in-platform clonal amplification (3) (Fig 1).

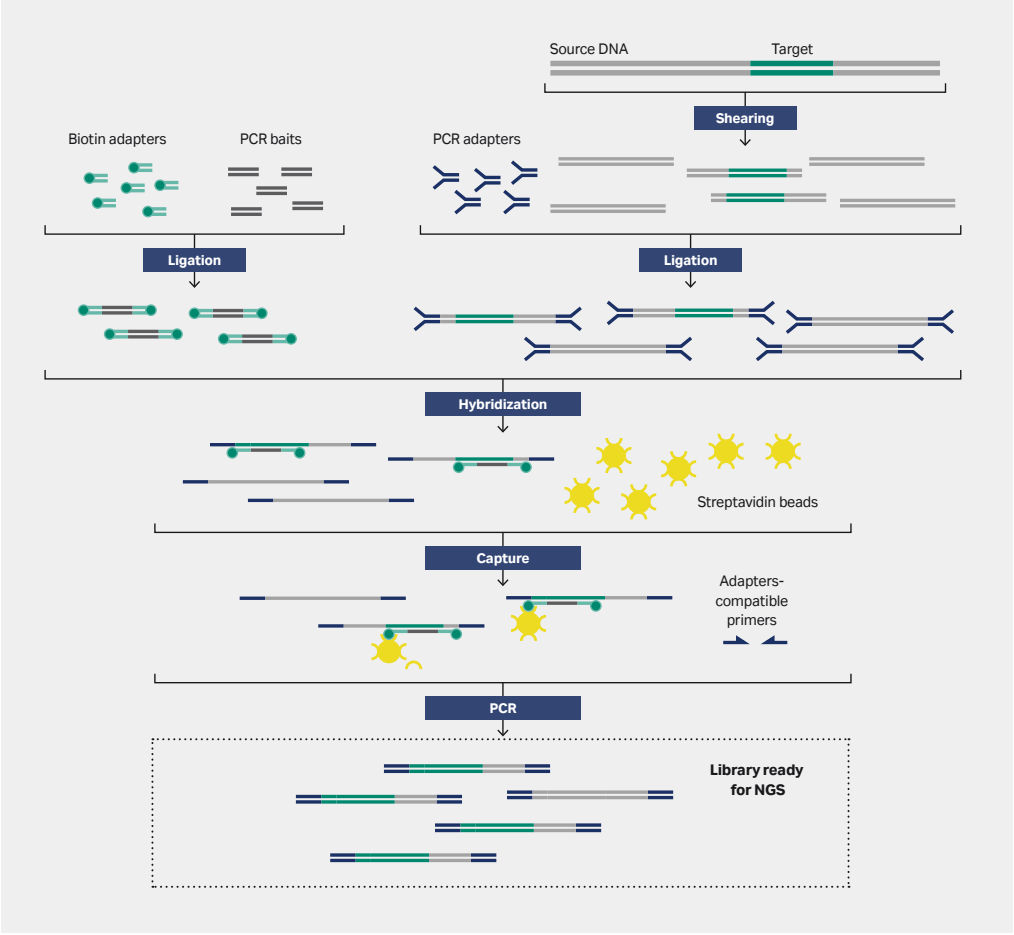

Fig 1. Library preparation and hybrid capture target enrichment steps for NGS.

The adapters can contain index bases, comprising of a sequence or "barcode" unique to the sample. After sequencing, bioinformatic analysis separates out the data by barcode so multiple samples can be processed simultaneously in a single sequencing compartment (11).

Following library preparation, the second stage in a targeted-sequencing workflow is target enrichment, which enhances the sequencing signal (Fig 1). By using a specific probe or set of probes complementary to the genomic region(s) of interest, a sequencing panel can be created and enriched with millions of copies of only the specific target region(s) of interest.

Enrichment helps provide sufficient copies of the target sequence for reliable detection by the sequencer. It also helps yield high-resolution data and drives the massively parallel sequencing power of short-read NGS. Target enrichment can be achieved via two approaches that harness either PCR or hybrid capture to create an NGS panel (10).

Target enrichment approaches

More than any other stage in the NGS workflow, successful target enrichment can have the greatest impact on your results. It maximizes on-target reads and sequencing depth, while improving confidence in base-calling for variant analysis. The many benefits of target enrichment prompted an increased focus on the development of enrichment approaches. Several established approaches exist, and each has benefits and challenges.

PCR-based enrichment approaches and multiplexing

PCR-based enrichment, also referred to as amplicon-based enrichment, is one of the more common NGS enrichment techniques. It relies on complementary primers to amplify genomic regions of interest (10).

Thousands of primers are used in PCR reactions, rapidly generating millions of amplicons in just a few cycles from a small amount (nanograms) of input DNA. Once generated, the amplicons must be pooled in equimolar amounts so that the final coverage of the total region of interest is as even as possible (11).

With multiplex PCR enrichment, you can interrogate multiple target regions in parallel using multiple primer pairs. A recent enhancement of this approach is the development of microdroplet PCR enrichment in which DNA molecules are fragmented and collected into thousands of microdroplets, and independent PCR reactions are run in parallel in each microdroplet (12).

The strengths of the PCR-based approach include straightforward methodology, very low input DNA requirement, and few off-target reads. However, this approach is appropriate only for the interrogation of small target regions. With longer reads, difficulties in equimolar pooling of amplicons are common, and much greater quantities of input DNA are required (13). Other drawbacks include the high cost of library preparation reagents and the high specificity of primers, which can result in the non-detection of variants that differ too much from the designed primers (14).

Hybrid capture target enrichment (solid-phase and in-solution)

Hybridization-based target enrichment strategies begin with fragmentation of DNA into randomly sized strands, sometimes known as a "shotgun library" (15). Adapters are then ligated to the DNA library, which act as barcodes for specific sequences. Because sequencing results can be separated out by barcode at the data analysis stage, multiple samples can be run simultaneously in a single sequencing platform (10).

Following library preparation and sometimes PCR amplification of the DNA library, fragmented DNA libraries are hybridized with single-stranded DNA or RNA probes specific for the region of interest. Unbound, nonspecific molecules are washed away, and the enriched DNA is eluted for NGS (16). This DNA hybridization step can be carried out in solution or on a solid support.

In solid-phase enrichment, DNA probes are bound to a solid support, such as a glass microarray slide. During in-solution enrichment, capture is achieved in the hybridization solution with an excess of probes (16). Typically, the DNA or RNA probes are biotinylated and, as such, recovering the fragment-probe duplexes is a straightforward process using streptavidin-coated magnetic beads.

Streptavidin has an incredibly high naturally occurring affinity for biotin, and the streptavidin-biotin complex is stable and is resistant to solvents, denaturants, and high temperature. As a result, streptavidin-coated magnetic bead kits are extremely efficient at capturing hybridized DNA tagged with biotin.

The magnetic beads are washed to remove nonspecific fragments, while captured fragments are eluted and amplified by PCR to create a strong sequencing signal and can be quantified by qPCR (10). Some magnetic beads are engineered to allow on-bead PCR amplification of the target sequence, streamlining the process with fewer washing and elution steps.

Magnetic beads can enable significantly greater capture efficiency, and their use can lead to fewer artifacts reaching the sequencing stage. The result is an increase in your NGS yield while minimizing off-target reads. Using magnetic beads is also automation friendly for high-throughput sequencing applications.

Hybrid capture is best suited for the enrichment of medium to large target genome regions, because larger fragments can be captured and processed in parallel without the need for equimolar pooling of amplicons (16). For this reason, hybrid capture can be more beneficial for clinical applications where targeted sequencing of larger regions is appropriate, such as in oncological applications in which hundreds to thousands of genes are sequenced for diagnostic purposes (17). Table 1 compares some of the practicalities of PCR and hybrid capture enrichment.

Table 1. Comparison of typical PCR and hybrid capture enrichment workflows

Feature	PCR enrichment	Hybrid capture enrichment
Workflow time	Shorter	Longer
Input DNA	10–100 ng	>1 µg
Number of targets per panel	Fewer than 10 000 amplicons	Virtually unlimited
Sensitivity	Down to 5%	Down to 1%
Cost per sample	Generally lower	Can vary (in-solution hybrid capture much more economic than solid-phase)

Hybrid capture: an inside perspective

Because hybrid capture can be adapted for a wide range of NGS applications from short reads up to whole exome sequencing, it can be used in many day-to-day clinical and research NGS applications. Despite the myriad benefits of hybrid capture, this enrichment approach does have challenges.

The benefits of hybrid capture enrichment

The chief benefit of hybrid capture over PCR enrichment is its ability to capture and interrogate large target regions (up to a human whole-exome) in a single experiment in a more rapid and streamlined manner. The human exome, which is around 30 Mb in size, would take approximately 6000 separate PCR reactions to enrich. Each of these reactions would require specific primers and optimization and input DNA of around 120 µg. With hybrid capture, the human exome can be enriched in a single experiment in about three days with just 3 µg of input DNA (13).

Hybrid capture works particularly well for genotyping and rare-variant detection. Because capture with hybridization probes does not require PCR primer design, it is less likely to miss mutations and performs better with respect to sequence complexity.

Hybrid capture's capacity for mutation discovery and its ability to capture large target regions with greater efficiency makes the approach particularly suited to discovering mutations in rare Mendelian disorders such as hereditary hearing loss (18), mitochondrial diseases (19), and cancer-related mutations (17, 20). Its ability to capture sequences from complex biological samples has also made hybrid capture a powerful tool in the study of ancient DNA samples (21).

Solid-phase versus in-solution hybrid capture

The first commercial hybrid capture target enrichment kits developed involved enrichment on a solid array. These kits had substantial advantages over PCR enrichment approaches. Solid-phase hybrid capture is faster and less laborious than PCR and is more suitable for longer reads. Additionally, PCR target-enriched experiments require a large amount of input DNA (10–15 µg) to generate enough DNA library, irrespective of the target sequence size.

However, solid-phase hybrid capture has disadvantages. It requires expensive hardware such as a hybridization station, and it is limited by the number of samples that can be run in parallel (11).

In-solution hybrid capture technologies were developed to overcome some of the challenges of solid-phase platforms. In addition to removing the need for solid support, a primary difference between in-solution and solid-phase hybrid capture is the ratio of library fragments to probes. In-solution uses an excess of probes over library DNA, whereas solid-phase uses an excess of DNA library over probes. Thus, substantially less fragmented DNA library is required for in-solution hybrid capture, which drastically reduces the amount of input DNA necessary for sequencing and improves capture success (22).

In short-read targeted sequencing, the uniformity and specificity of sequences obtained from an in-solution capture experiment are usually higher compared to solid-phase capture. Therefore, for short-read applications, in-solution hybrid capture can deliver greater sequence coverage of the target regions from a similar yield of sequences.

For whole-exome captures, both solution and solid-phase techniques are found to perform equivalently (13). That said, in-solution hybrid capture does not require the specialized hardware that solid-array capture requires. It can instead be performed using 96-well plates and a thermal cycler, making in-solution hybrid-capture generally more accessible.

Optimizing hybrid capture enrichment outcomes

To make the most of your hybrid-capture enrichment protocols, consider the following:

- **Fully calibrate your lab instruments:** Calibration is often overlooked, but even very small differences in temperature (+/- 2°C) can have a significant influence on hybridization reactions and can affect your on-target read percentage and capture uniformity.
- **Check the seal on your reaction plate or tube:** Hybridization buffers and reagents are prone to evaporation. Ensure that your reaction vessels are tightly sealed or you may experience capture failure.
- **Avoid using outside wells of microwell plates:** The so-called "edge effect" can occur in many microwell plate-based assays. The outside wells are more prone to evaporation. Avoid using these wells as test wells to improve replicability and to reduce the chances of capture failure.
- **Consider increasing hybridization incubation time:** Increased time can improve capture rates and boost overall panel performance.
- **Regularly vortex during hybridization and bead capture:** Vortexing helps maintain optimum efficiency, homogeneity, and capture kinetics.
- **Make sure streptavidin-coated beads are in suspension during washes:** Vortexing can ensure washes are as efficient as possible and can improve data quality.

Streptavidin-coated magnetic beads and hybrid capture

Most commercial in-solution hybrid capture kits and protocols use biotinylated probes and streptavidin-coated magnetic beads for efficient cleanup and capture of hybridized DNA for sequencing. Although magnetic beads are a staple in hybrid capture, a vast number of beads are commercially available, and not all beads are created equal.

Magnetic bead cleanup

Streptavidin's high affinity for biotin is known as the strongest noncovalent biological interaction (23). This highly selective and stable interaction makes streptavidin-biotin ideal for hybrid capture enrichment. By tagging hybridization probes with biotin, hybridized DNA can be captured easily using streptavidin-coated magnetic beads. These beads are then purified by magnetic pull-down, and the target DNA is eluted ready for library preparation and sequencing. Figure 2 illustrates this process.

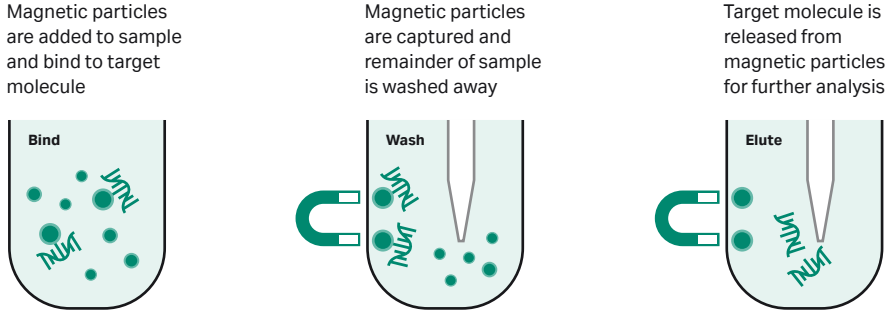

Fig 2. Magnetic bead capture of hybridized DNA following hybrid capture enrichment.

Despite the highly selective interaction, some unhybridized DNA can sometimes bind to streptavidin-coated beads, which then acts as a contaminant in sequencing and complicates analysis. Working with chilled buffers, increasing the number of wash steps, and performing washes at room temperature can all help minimize the possibility of unhybridized DNA carrying over.

Optimized magnetic beads for hybrid capture

Streptavidin-coated magnetic bead kits are extremely efficient at capturing hybridized DNA tagged with biotin. However, the costs associated with many commercially available beads can be high, elution steps can be challenging to get right, and some beads have different levels of yield and nonspecific binding. These challenges can be particularly frustrating because they potentially lead to inconsistencies in the amplified target sequence and poor performance in downstream applications.

A new generation of magnetic beads specifically optimized for NGS target enrichment, however, aims to overcome some of these potential pitfalls. Developed with optimized characteristics and adhering to tighter quality standards, these magnetic beads can potentially provide improved capture efficiency for a streamlined and reliable hybrid capture workflow.

The bead surface is engineered for optimized streptavidin levels, so capture efficiency is maximized. Nonspecific binding is reduced resulting in fewer off-target reads and improved replicability of NGS results.

Because they are optimized specifically for target enrichment, this new generation of magnetic beads maintains high colloidal stability in hybridization buffers and is designed to withstand high-temperatures, which allows on-bead PCR amplification of the captured target sequence. The result is a more streamlined protocol that is more automation-friendly for high-throughput applications.

By enabling rapid liquid-phase reaction kinetics through optimal streptavidin load, these new generation magnetic beads make it possible to achieve direct, rapid, and reproducible isolation of biotinylated molecules with high yield and low nonspecific binding. As these beads become more common, commercial kits will be able to deliver improved operational efficiency, leading to higher quality and more replicable data in downstream applications.

Conclusions

Target enrichment has become a powerful tool for the detection of disease-causing genomic variants and is a highly effective way of reducing sequencing costs while minimizing the time needed for sequencing data analysis. However, it could be advantageous to consider the inherent pros and cons of both PCR and hybrid-capture approaches when choosing a target enrichment strategy for your NGS workflow.

Because hybrid capture can be adapted to a wide range of NGS applications, it can often be more applicable than PCR for day-to-day clinical and research NGS. Advancements in the hybrid capture approaches including the development of in-solution hybrid capture and improved magnetic-bead technologies, have vastly improved the efficiency and sequencing depth of downstream NGS applications.

If you are considering adopting hybrid capture or bead-based enrichment into your NGS research, getting support and guidance from Cytiva can help you to adopt and adapt to these technologies smoothly.

Looking to discuss your magnetic bead protocols or NGS workflows?
Contact a product specialist at Cytiva

References

1. McCombie, W, R, McPherson, J, D, Mardis, E, R, Next-Generation Sequencing Technologies. *Cold Spring Harb. Perspect. Med*. 2019;9(11).
2. Slatko, B, E, Gardner, A, F, Ausubel, F, M, Overview of Next-Generation Sequencing Technologies. *Curr. Protoc. mol. biol.* 2018;122(1).
3. Kumar, K, R, Cowley, M, J, Davis, R, L, Next-Generation Sequencing and Emerging Technologies. *Semin. Thromb. Hemost.* 2019;45(7):661–673.
4. Amarasinghe, S, L, Su, S, Dong, X, Zappia, L, Ritchie, M, E, Gouil, Q, Opportunities and challenges in long-read sequencing data analysis. *Genome biol.* 2020;21(1):30.
5. Liu, L, Li, Y, Li, S, et al. Comparison of next-generation sequencing systems. *J. Biomed. Biotechnol.* 2012; 251364.
6. van Dijk, E, L, Auger, H, Jaszczyszyn, Y, Thermes, C, Ten years of next-generation sequencing technology. *Trends Genet*. 2014;30(9):418–426.
7. Bewicke-Copley, F, Arjun-Kumar, E, Palladino, G, Korfi, K, Wang, J, Applications and analysis of targeted genomic sequencing in cancer studies. *Comput. Struct. Biotechnol. J.* 2019;17:1348–1359.
8. Gulilat, M, Lamb, T, Teft, W, A, et al. Targeted next generation sequencing as a tool for precision medicine. *BMC Med. Genom.* 2019;12(1):81.
9. van Dijk, E, L, Jaszczyszyn, Y, Thermes, C, Library preparation methods for next-generation sequencing: tone down the bias. *Exp. Cell Res.* 2014;322(1):12–20.
10. Kozarewa, I, Armisen, J, Gardner, A, F, Slatko, B, E, Hendrickson, C, L, Overview of Target Enrichment Strategies. *Curr. Protoc. Mol. Biol.* 2015;112(7.21.1–7.21.23).
11. Mamanova, L, Coffey, A, J, Scott, C, E, et al. Target-enrichment strategies for next-generation sequencing. *Nat. Methods*. 2010;7(2):111–118.
12. Tewhey, R., Warner, J. B., Nakano, M, et al. Microdroplet-based PCR enrichment for large-scale targeted sequencing. *Nat. Biotech.* 2009;27(11):1025–1031.
13. Samorodnitsky, E, Jewell, B, M, Hagopian, R, et al. Evaluation of Hybridization Capture Versus Amplicon-Based Methods for Whole-Exome Sequencing. *Hum. Mutat*. 2015;36(9):903–914.
14. Altmüller, J, Budde, B, S, Nürnberg, P, Enrichment of target sequences for next-generation sequencing applications in research and diagnostics. *Biol. Chem.* 2014;395(2):231–237.
15. Albert, T, J, Molla, M, N, Muzny, D, M, et al. Direct selection of human genomic loci by microarray hybridization. *Nat. Methods.* 2007;4(11):903–905.
16. Kiialainen, A, Karlberg, O, Ahlford, A, Sigurdsson, S, Lindblad-Toh, K, Syvänen, A, C, Performance of microarray and liquid-based capture methods for target enrichment for massively parallel sequencing and SNP discovery. *PloS one*, 2011;6(2).
17. Shen, W, Shan, B, Liang, S, et al. Hybrid capture-based genomic profiling of circulating tumor DNA from patients with advanced ovarian cancer. *Pathol. Oncol. Res.* 2021;27.
18. Shearer, A. E., DeLuca, A. P., Hildebrand, M. S, et al. Comprehensive genetic testing for hereditary hearing loss using massively parallel sequencing. *PNAS USA*. 2010;107(49):21104–21109.
19. Calvo, S. E., Compton, A. G., Hershman, S, G, et al. Molecular diagnosis of infantile mitochondrial disease with targeted next-generation sequencing. *Sci. Transl. Med*. 2012;4(118).
20. Rozenblum, A, B, Ilouze, M, Dudnik, E, Dvir, A, Soussan-Gutman, L, Geva, S, Peled, N, Clinical impact of hybrid capture-based next-generation sequencing on changes in treatment decisions in lung cancer. *J. Thorac. Oncol.* 2017;12(2):258–268.
21. Gaudin, M, Desnues, C, Hybrid capture-based next generation sequencing and its application to human infectious diseases. *Front. Microbiol.* 2018;9.
22. Gnirke, A, Melnikov, A, Maguire, J, et al. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat. Biotechnol.* 2009;27(2):182–189.
23. Chivers, C, E, Koner, A, L, Lowe, E, D, Howarth, M. (2011). How the biotin-streptavidin interaction was made even stronger: investigation via crystallography and a chimeric tetramer. *Biochem. J*. 2011;435(1):55–63.

cytiva

cytiva.com/sequencing

For local office contact information, visit cytiva.com/contact

CY29105-09Sep22-WP