illumina®

Benefits of NGS Targeted Resequencing

Taking science further: beyond qPCR and Sanger sequencing for somatic and germline variant detection

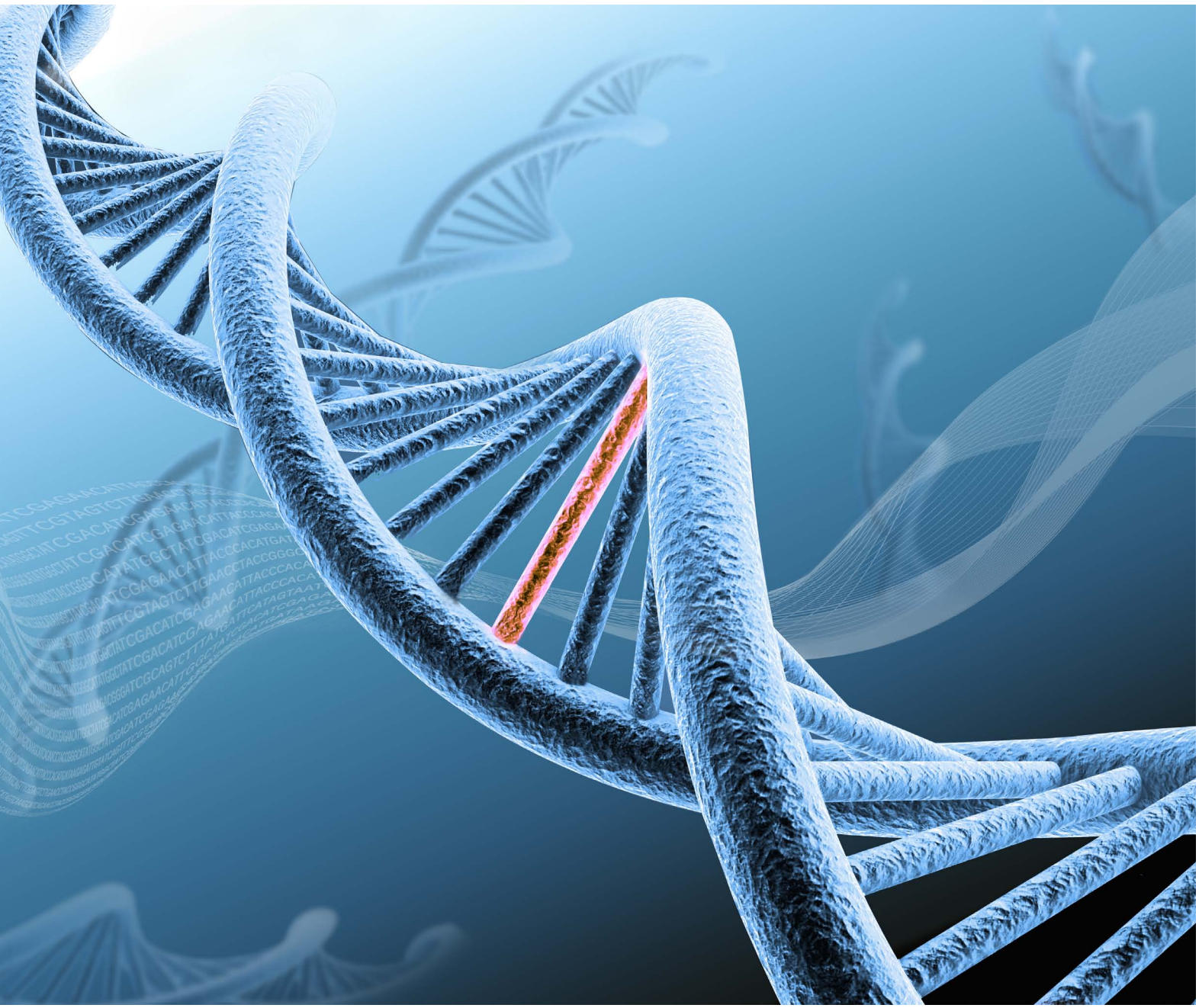

Table of Contents

Why is now the time to consider NGS targeted resequencing? 4

- The dramatic decrease in the cost of next-generation sequencing 4
- The focused power of NGS targeted resequencing 5
- Grant funding trends for NGS targeted resequencing 5

When does targeted resequencing make the most sense? 8

- Choosing between PCR, Sanger, targeted resequencing, and WGS 8

What are the benefits of targeted resequencing vs PCR and Sanger sequencing? 10

- Targeted resequencing allows deep sequencing and higher sensitivity 10
- Targeted resequencing delivers faster turnaround for clinical research 11
- Targeted resequencing provides higher discovery power/mutation resolution 12
- Next-generation sequencing enables integrated genomics 12
- Summary of targeted resequencing benefits 13

Illumina NGS targeted resequencing solutions and methods 15

- Amplicon vs. enrichment targeted resequencing 15
- Amplicon-based targeted resequencing 16
- Enrichment-based targeted resequencing 17
- Illumina targeted resequencing comprehensive workflows 18
- Library preparation with ready-to-use vs. custom panels 18
- Ready-to-use panels 18
- Library prep and array kit selector 22
- Custom panels with DesignStudio™ Software 23
- Expanded options with Illumina Concierge Services 23
- The sequencing coverage calculator 25
- User-friendly data analysis 26

Summary 27

Glossary 27

References 29

This document highlights recent publications that demonstrate the use of Illumina technologies in immunology research. To learn more about the platforms and assays cited, visit www.illumina.com.

Why is now the time to consider NGS targeted resequencing?

The dramatic decrease in the cost of next-generation sequencing

Over the past decade, molecular biologists and translational researchers have focused on the identification of somatic or germline DNA variants with strong associations to cancer, neurobiological disorders, and other genetic diseases. Quantitative PCR, Sanger sequencing (capillary electrophoresis sequencing), and microarray technology have all played an important role in these genetic studies. During the same time period, the massively parallel sequencing technology known as next-generation sequencing (NGS) has revolutionized the biological sciences. From the emergence of the Illumina Genome Analyzer in 2006 to the release of the NovaSeq™ System in 2017, the data output of NGS has outpaced Moore's law—more than doubling each year (Figure 1).[1] The rapid drop in sequencing cost and the massive increase in data output have resulted in fundamental changes to the kinds of questions scientists can ask and answer. Researchers can now analyze the complete human genome in a single sequencing experiment, and sequence thousands to tens of thousands of genomes in a single year. As Eric Lander, founding director of the Broad Institute of MIT and Harvard stated, "As costs continue to come down, we are entering a period where we are going to be able to get the complete catalog of disease genes. This will allow us to look at thousands of people and see the differences among them, to discover critical genes that cause cancer, autism, heart disease, or schizophrenia."[2]

While the drop in cost for whole-genome sequencing (WGS) is an exciting development for science, the reduction in cost and the increasing simplicity of other NGS methods, such as targeted resequencing, have made the benefits of NGS accessible to the wider research community. With the adoption of NGS on the rise, targeted resequencing is proving to be a powerful tool for somatic and germline variant detection.

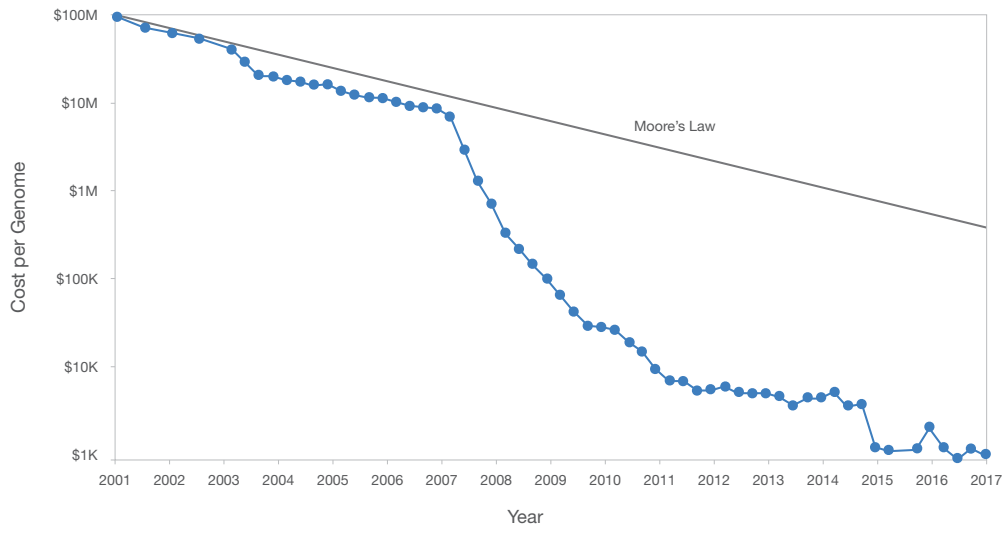

Figure 1: The dramatic reduction in sequencing cost—Moore's Law is the observation that over the history of computing hardware, the number of transistors in an integrated circuit doubles approximately every two years. Since 2001, genome sequencing costs are decreasing at a rate that outpaces Moore's Law.

See how genomics and targeted resequencing are empowering researchers around the globe: Global perspectives on the impact of next-generation sequencing.

The focused power of NGS targeted resequencing

With targeted resequencing, a subset of genes or target regions (a sequencing panel) are amplified or enriched before sequencing, efficiently and cost-effectively focusing the power of NGS. Targeted resequencing offers several significant advantages. It enables deep sequencing (sequencing at much higher coverage levels), which allows greater confidence over Sanger sequencing for calling variants or low-frequency alleles in a given region of interest.[3,4] When speed is critical to success, targeted resequencing can also provide fast turnaround times, due to a higher multiplexing capacity (ie, ability to pool libries and sequence them together), lower data analysis requirements, and the ability to sequence anywhere from a small number of genes to the entire coding region or exome.[5]

Targeted resequencing can reveal variants, such as low-frequency variants that would be more expensive or more challenging to identify with PCR or Sanger sequencing.[5,6] The ability to detect low-frequency variants can enable identification of novel functional variants, facilitate biomarker discovery, or lead to the identification of clinically relevant targets for translational research. [4,5] Targeted resequencing is particularly useful for the discovery of somatic mutations in complex samples such as cancerous tumors mixed with germline DNA.[7–9] Whether performing cancer studies, microbial genomics, agrigenomics, or molecular epidemiology, researchers can target regions of the genome relevant to their specific interests.

Grant funding trends for NGS targeted resequencing

With the declining cost of sequencing overall, and improvements in targeted resequencing methods, targeted resequencing is accelerating the pace of research and driving high impact publications. From 2004 to 2018, the number of publications with targeted resequencing has grown from a handful of studies to over 800 publications (Figure 2).

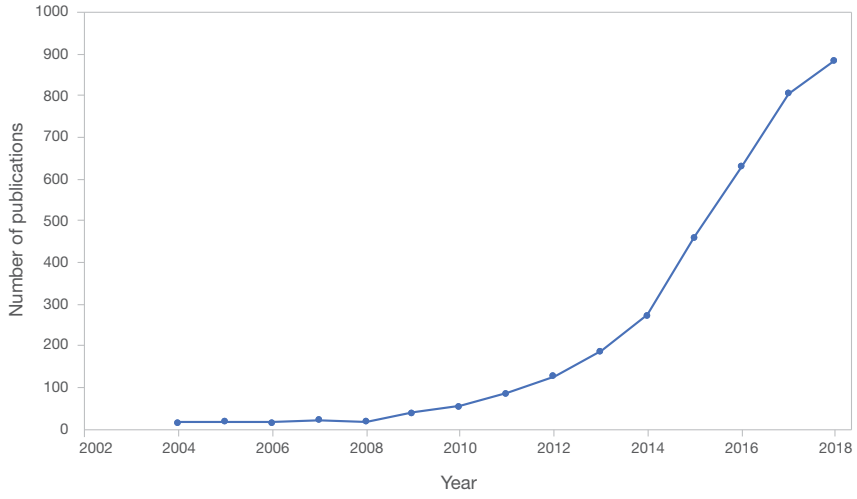

Figure 2: Publications with targeted resequencing are increasing—In PubMed, the number of times in which the keywords for targeted resequencing (TGRS) have been mentioned in titles and abstracts of papers has grown from ~10 to > 800 in the last 10 years. Keywords include: “targeted sequencing” OR “targeted resequencing” OR “TGRS” OR “amplicon sequencing” OR “enrichment sequencing.”

As the impact of targeted resequencing continues to grow, grant funding trends for targeted resequencing are also on the rise (Figure 3). Data from the National Institutes of Health (NIH) show that funding for research including NGS targeted resequencing has steadily increased each year and is 99.6% higher in 2018 than in 2004. These trends show that studies using NGS targeted resequencing methods are benefiting from increased funding and resulting in more publications every year.

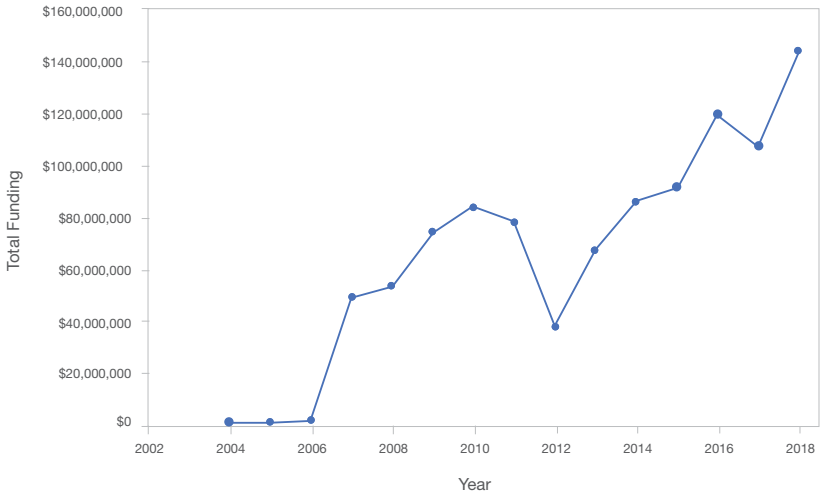

Figure 3: Grant funding for targeted resequencing on the rise—NIH funding for studies with targeted resequencing has seen a steep growth rate between 2004 and 2018.

When does targeted resequencing make the most sense?

Choosing between PCR, Sanger sequencing, targeted resequencing, and WGS

Deciding when to use PCR, Sanger sequencing, targeted resequencing, or broader NGS approaches such as exome sequencing or WGS depends on a combination of factors. These include the number of samples, the total amount of sequence in the target regions, budgetary considerations, and ultimately, the overall goals of the research study. Sanger sequencing and PCR are typically good choices when the number of target regions is low (1–20 target regions), and when the study aims are limited to screening or identification of known targets or variants (Figure 4). At the opposite end of the sequencing spectrum, exome sequencing and WGS are excellent methods for comprehensive genetic analysis and variant discovery. Targeted NGS offers a balanced choice between these options that supports both screening and variant discovery study designs. It is the most cost-effective approach for the sequencing of tens to thousands of genes with a high number of samples (Figure 5).[5,6,10] With the ability to sequence multiple genes across multiple samples simultaneously, targeted resequencing methods save time and resources compared to traditional iterative methods.[5,6,10] Moreover, targeted NGS methods preserve precious sample material by requiring lower DNA input—an important consideration for applications such as analysis of forensic DNA samples, single-cell samples, or cancer biopsy samples (Figure 5).[5,12]

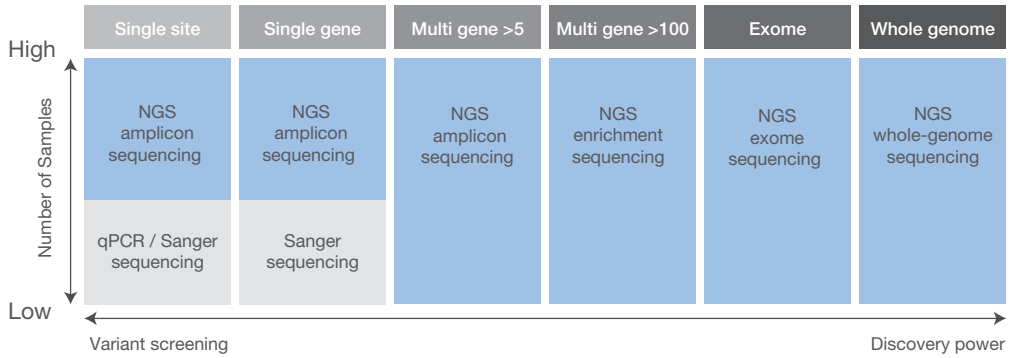

Figure 4: Options for WGS, exome sequencing, targeted resequencing, Sanger sequencing, and PCR—For variant screening studies where the sample number is high, NGS amplicon sequencing is more efficient and cost-effective. For discovery-related applications, any NGS approach will provide higher discovery power compared to PCR/Sanger sequencing.

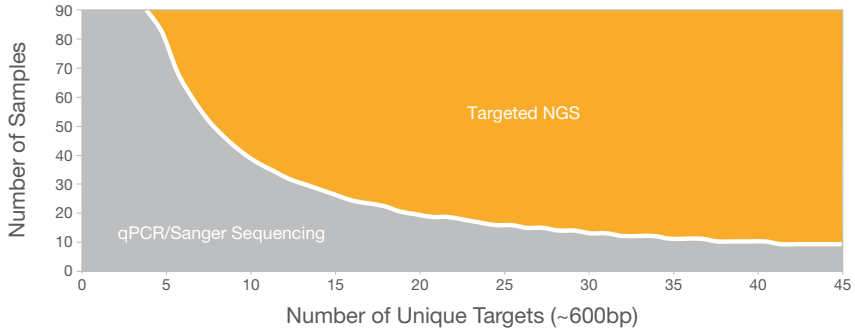

Figure 5: Cost effectiveness for targeted resequencing vs PCR and Sanger sequencing—The area above the line represents higher cost-effectiveness with targeted DNA sequencing compared to Sanger sequencing or qPCR.

What are the benefits of targeted resequencing vs PCR and Sanger sequencing?

Targeted resequencing allows deep sequencing and higher sensitivity

As with all targeted NGS methods, the targeted resequencing approach provides a rapid and cost-effective alternative to single-gene testing. However, the most powerful aspect of targeted resequencing may be the ability to sequence a given region of interest at much higher coverage levels compared to WGS or Sanger sequencing. WGS is typically performed at 30×–75× coverage while targeted NGS enables sequencing depths of 5000× or higher (Figure 6). From a research perspective, deep sequencing translates into higher sensitivity (ability to detect low-frequency variants). Sanger sequencing can provide a limit of detection down to 20% allele frequency while NGS targeted resequencing provides a limit of detection down to 1%.[13,14] Deep sequencing and the resulting high-sensitivity can be critical for certain kinds of studies, such as the detection of low-frequency subclones within heterogenous tumors or the detection of somatic mosaicism.[3,15,16] Several studies have shown that Sanger sequencing does not have the sensitivity required to detect mosaic variants consistently.[3, 15]
Dr. Saumya Jamuar, cofounder of Global Gene Corporation, used amplicon sequencing on the MiSeq™ System to identify rare mosaic mutations that lead to neural malformations. Dr. Jamuar discussed targeted resequencing vs. Sanger sequencing for somatic variant detection:

> Actually, there weren't many advantages to using Sanger sequencing for variant detection. It was a lengthy process because Sanger sequencing proceeds gene by gene. With NGS, all we had to do was put all our genes of interest together in one panel, one reaction, and we would get the results. Sanger sequencing has a limit of detection of about 15%. This isn't low enough to detect somatic mosaicism."
>
> Dr. Saumya Jamuar, Genetics Service Consultant at KK Women and Children's Hospital in Singapore and cofounder of Global Gene Corporation

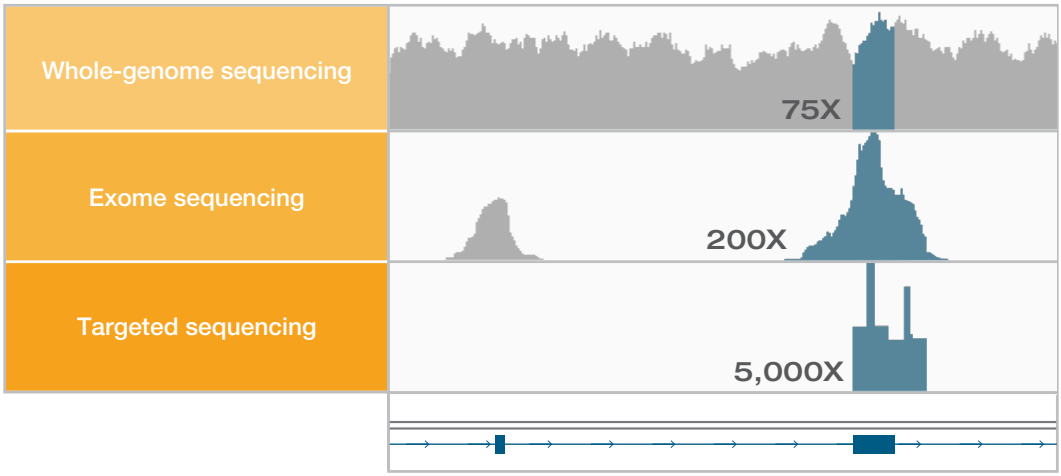

Figure 6: Sequence regions of interest at higher resolution—With more reads focused on fewer regions, targeted resequencing achieves deeper coverage and higher resolution compared to WGS or exome sequencing.

Targeted resequencing delivers faster turnaround for clinical research

Clinical or translational studies require tests that are not only accurate, but also rapid and cost-effective. Tests must be accurate because they can lead to the development of divergent diagnosis, risk stratification, and drug treatments.[18] Speed is also a critical factor when assessing clinical samples for translational research. While Sanger sequencing has been considered the gold standard in clinical research for many years, the process can be time consuming and inefficient when multiple gene tests and many samples are required (Figure 7).[5,10,11] Furthermore, limited tissue availability can make iterative Sanger gene testing nearly impossible.[5,12] Targeted resequencing is becoming more prevalent with clinical and translational researchers because it effectively addresses many of these challenges.[18,19] A study from the University Hospital of Cologne Germany comparing Sanger sequencing and targeted resequencing methods found that the multiplexing capacity of targeted resequencing led to reduced turnaround times for analysis of tumor samples.[5] In addition, the study authors found that targeted resequencing required significantly less input material and proved to be a more cost-effective approach, compared to the Sanger method.

Figure 7: Case study: targeted resequencing supports faster turnaround times—A translational researcher investigating the genetics of epilepsy wants to interrogate 3 genes with a total of 94 possible variants across 10 samples (ATP1A2 = 34 targets, ATP13A2 = 47 targets, and CDKL5 = 13 targets). With 1 reaction per variant, Sanger sequencing would require a total of 20 reaction plates (10 samples x 1 plate x 2 reactions (forward and reverse) = 20 x 96-well plates. Assuming 1 full-time employee preps 4 plates per day, the Sanger method would require 5 days. In contrast, using targeted resequencing with sample multiplexing (Illumina amplicon sequencing on the MiniSeq™ System), the same study can be performed in 1.5 days. Multiplexing combined with the ability to sequence thousands of targets simultaneously enables a faster turnaround time compared to other technologies.

Dr. John Robinson, research manager of the Blackburn Cardiovascular Genetics Laboratory at the Robarts Research Institute, uses a custom targeted resequencing panel (Nextera™ Rapid Capture Enrichment) on the MiSeq™ System to profile genetic variants associated with dyslipidemias and related metabolic disorders. Dr. Robinson discussed his experience with targeted resequencing vs. Sanger sequencing for the identification of variants associated with lipid metabolism:

> We found that an individual subject might take a month to Sanger sequence before we found the causative mutation, and that's with all the Sanger costs and the costs for the labor, PCR, etc. With targeted resequencing, within a 2-week period we get 24, 700 kb sequencing data files that have been annotated and categorized on rareness and pathogenicity. That's 24 people every 2 weeks. So the timeframe is just impossible to compare to a Sanger 1000-base-pair-at-a-time sequencing process."
>
> Dr. John Robinson, research manager of the Blackburn Cardiovascular Genetics Laboratory at the Robarts Research Institute

Targeted resequencing provides higher discovery power and mutation resolution

While PCR and Sanger sequencing offer quick, familiar workflows, targeted resequencing has higher discovery power (ability to identify novel variants), and high variant resolution (single base resolution).[3,4,10,11] Depending on the research goals, PCR, Sanger sequencing, and targeted resequencing provide different levels of information. PCR can indicate whether a variant is present, but cannot provide single-base resolution. Sanger sequencing can cost-effectively identify specific variants with single-base resolution for a small number of genes or target regions. In contrast to qPCR and Sanger sequencing, targeted resequencing can identify variants across thousands of target regions, down to single-base resolution, in a single experiment (Figure 8).[17] Furthermore, due to the larger scale of sequence interrogation and targeted sequence capture methods, the possibility of detecting novel variants is much higher.[11]

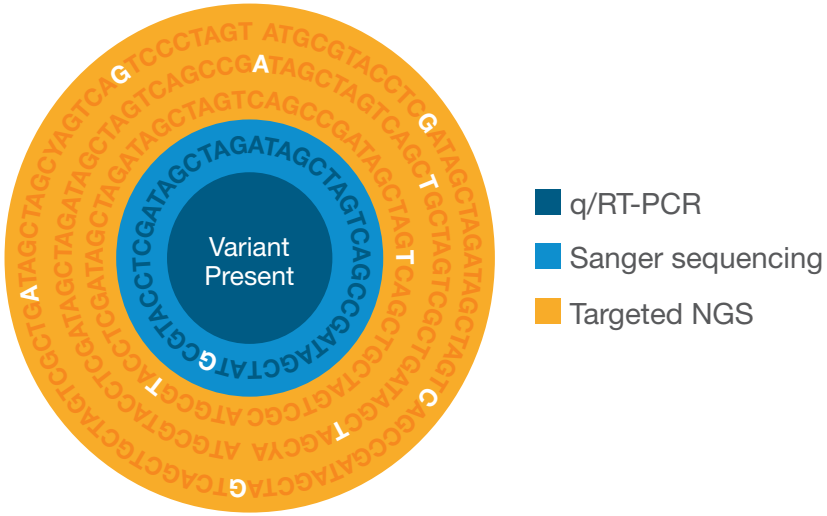

Figure 8: Targeted resequencing delivers higher discovery power and highest mutation resolution—Targeted resequencing can provide single base resolution across hundreds to thousands of target regions.

Next-generation sequencing enables integrated genomics

While the detection of somatic and germline variants has provided important insights into the genetic basis of disease, many researchers are now leveraging integrated genomics analysis to obtain a multifaceted view of biological processes and disease pathogenesis. In a research study from the Netherlands Cancer Institute, comprehensive genomic, transcriptomic, and proteomic analysis was used to identify three genetic subtypes within a large cohort of invasive lobular carcinoma (ILC) histological samples.[20] The subtypes demonstrated differences in mRNA profiles, hormone characterization, and significant differences in prognostic indicators. For cutaneous melanoma, data analysis suggests that integrating multiple types of genomic data leads to prognostic models with an improved prediction performance.[21] For investigators seeking to integrate multiple genome-wide techniques, NGS technology provides a single platform for genomics, transcriptomics, and epigenetics research.

> We definitively showed that NGS was a better alternative than Sanger sequencing for detecting somatic mutations.
>
> Dr. Saumya Jamuar, Genetics Service Consultant at KK Women and Children's Hospital in Singapore and cofounder of Global Gene Corporation

Summary of targeted resequencing benefits

With readily available predesigned targeted sequencing panels, easy online tools for custom panel design, and the emergence of personal sequencers such as the iSeq™ 100 System, targeted resequencing has become simpler and more accessible than ever before. The overall cost of NGS sequencing and targeted resequencing continues to fall and easy bioinformatics solutions are removing the previous challenges associated with NGS data analysis. Multiple types of library preparation kits for targeted resequencing have expanded the choices available to scientists and researchers. These include various panels with preselected content, and the ability to target specific regions of interest with online methods for designing and ordering custom sequencing panels. With significant advantages compared to PCR and Sanger sequencing, more and more biologists and translational researchers are choosing NGS targeted resequencing for their variant detection studies (Table 1).

Table 1: Comparison of PCR, Sanger sequencing, and targeted resequencing

	Benefits	Challenges
PCR	• High sensitivity[a] • Familiar workflow • Capital equipment already available in most labs	• Can only interrogate a limited set of mutations • Virtually no discovery power • Limited mutation resolution • Low scalability due to increasing sample input requirements
Sanger Sequencing	• Cost-effective sequencing low numbers of targets (1–20 targets) • Familiar workflow • Current gold standard in sequencing	• Low sensitivity (down to 20%) • Low discovery power • Not as cost-effective for high numbers of targets (> 1–20 targets) • Low scalability due to increasing sample input requirements
Targeted Resequencing	• Higher sequencing depth enables higher sensitivity (down to 1%) • Higher discovery power[b] • Higher mutation resolution[c] • Produce more data with the same amount of input DNA[d] • Higher sample throughput with sample multiplexing	• Not as cost-effective for sequencing low numbers of • targets (1–20 targets) • Not as time-effective for sequencing low numbers of targets (1–20 targets)

a. sensitivity = allele frequency limit of detection
b. discovery power = ability to identify novel variants
c. mutation resolution = identification of mutations from large chromosomal rearrangements down to single-nucleotide variants.
d. 10 ng DNA will produce ~1 kb with CE sequencing or ~300 kb with targeted resequencing (250 bp amplicon length × 1536 amplicons with an AmpliSeq for Illumina workflow)

Illumina NGS targeted resequencing solutions and methods

Amplicon vs. enrichment targeted resequencing

Illumina currently supports two methods for targeted resequencing: amplicon sequencing and enrichment sequencing. Amplicon sequencing, involves amplification and purification of regions of interest using primer-mediated, highly multiplexed oligo sets. With enrichment sequencing, regions of interest are captured by sequence-specific hybridization probes. Amplicon sequencing is a more affordable, rapid workflow than enrichment sequencing, supports small to mid-size panels, and is ideal for ready-to-use panels such as oncology panels or human transcriptome panels. Enrichment sequencing enables larger gene content panels (including whole human exome panels) and more comprehensive profiling for all variant types, including copy number variants, splice variants, and *de novo* discovery. Both of these highly multiplexed approaches enable a wide range of applications for the discovery, validation, or screening of genetic variants (Figure 9).

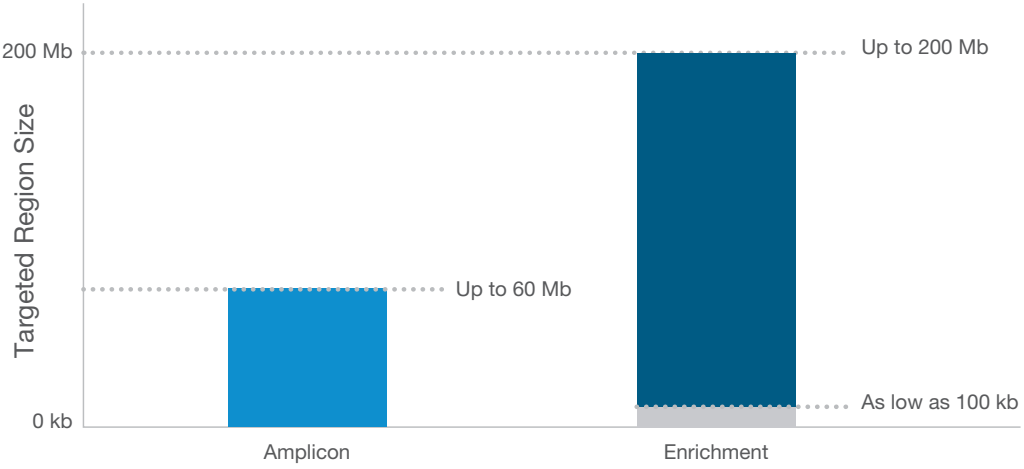

Figure 9: Comparison of targeted region size between amplicon and enrichment methods—Enrichment sequencing methods enable capture of much larger total content than amplicon generation.

Amplicon-based targeted resequencing

Illumina amplicon sequencing involves a multiplexed PCR approach, amplifying the predefined targeted regions from genomic DNA (Figure 10). This library prep uses two primer pools for multiplex PCR to amplify and target regions of interest. Using index adapters, libraries are indexed, further amplified, and combined in preparation for sequencing.

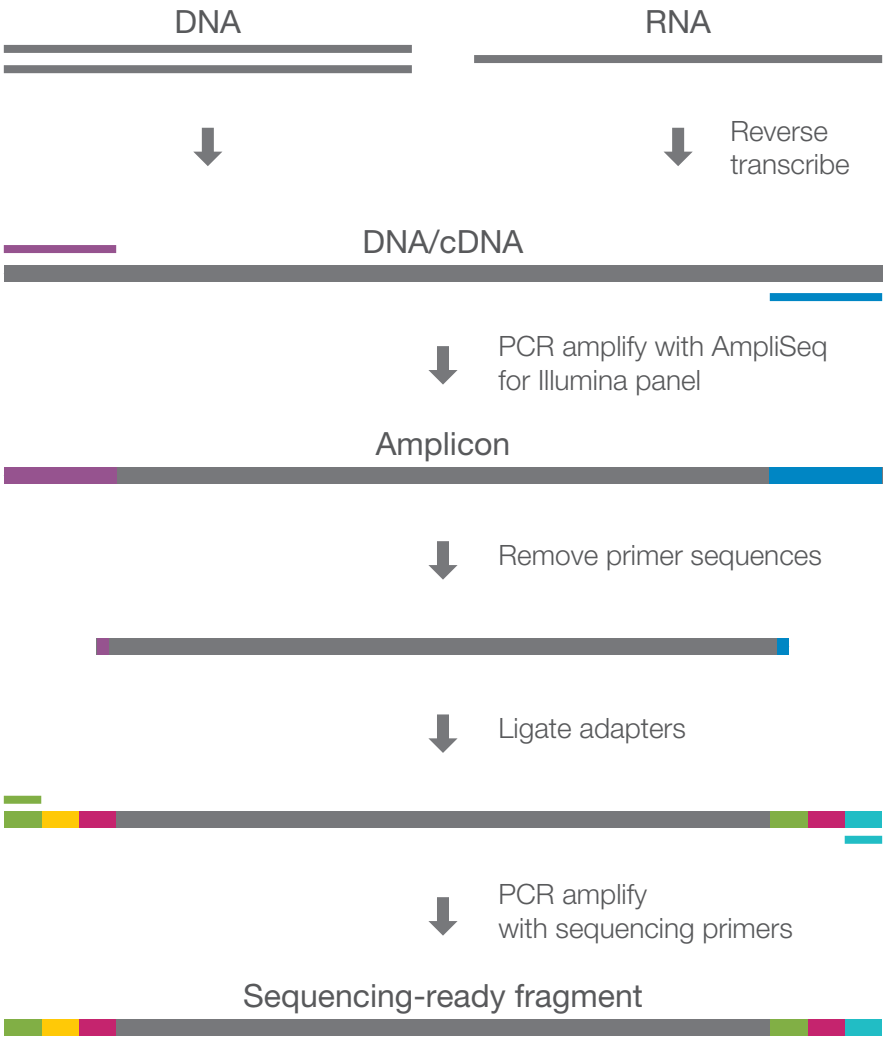

Figure 10: Amplicon-based library preparation—Amplicon sequencing uses multiplex PCR to amplify the target of interest.

To learn more about AmpliSeq™ for Illumina from ease-of-use operations to data analysis solutions, see the AmpliSeq for Illumina videos:

Bringing Together Two Leaders: Illumina and ThermoFisher Scientific

Sequencing, Amplified: AmpliSeq for Illumina

AmpliSeq for Illumina and Custom Content Creation

AmpliSeq for Illumina - Data Analysis with BaseSpace™ Sequence Hub

AmpliSeq for Illumina - Data Analysis with Local Run Manager

AmpliSeq for Illumina - Tertiary Analysis with BaseSpace Variant Interpreter

Enrichment-based targeted resequencing

Enrichment-based library prep begins with Nextera tagmentation, which converts input genomic DNA into adapter-tagged libraries without the need for mechanical shearing (Figure 11). Next, libraries are denatured and biotin-labeled probes specific to targeted regions are used for hybridization. The pool is enriched for regions of interest by adding streptavidin-coated beads that bind to the biotinylated probes. DNA fragments bound to the streptavidin-coated beads via biotinylated probes are magnetically pulled down from the solution. The enriched DNA fragments are then eluted from the beads.

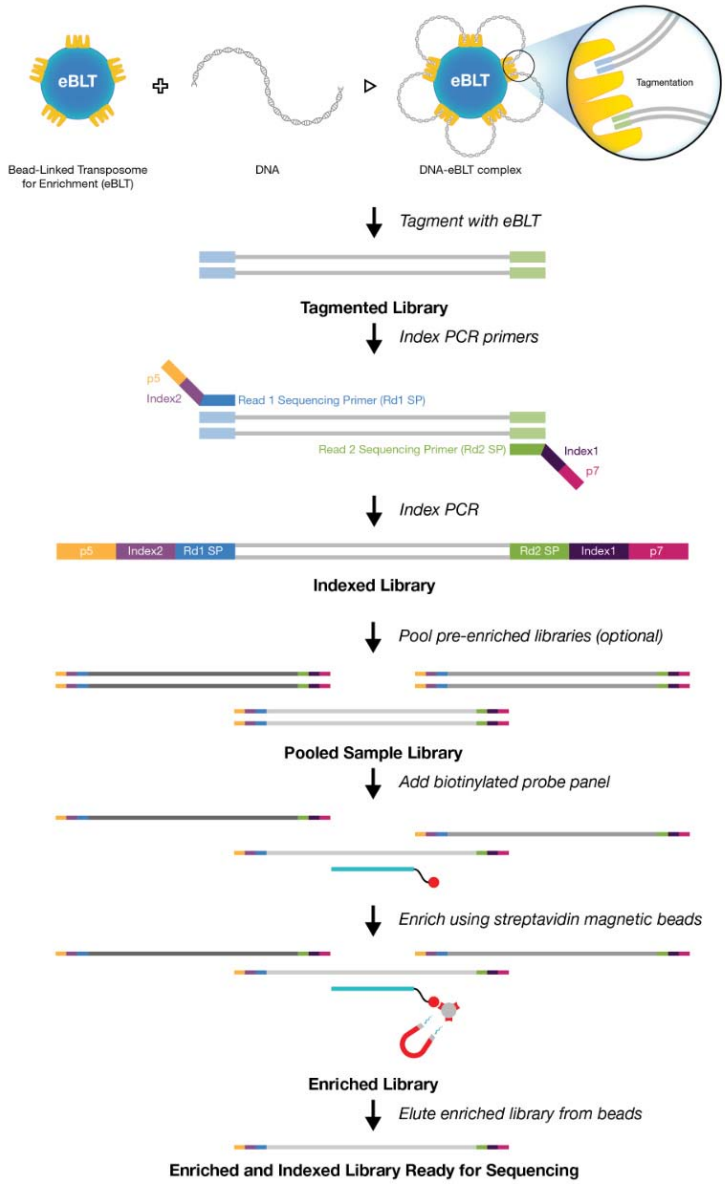

Figure 11: Enrichment-based library preparation—This method harnesses Nextera technology to provide a streamlined protocol that combines library preparation and enrichment steps.

To learn more about Nextera™ Flex for Enrichment, see the Nextera Flex video:

Nextera Flex for Enrichment: Simple, Fast, and Flexible

Illumina targeted resequencing comprehensive workflows

Illumina targeted resequencing workflows offer DNA-to-data solutions that include content design (for custom panels only), library preparation, sequencing, and data analysis (Figure 12).

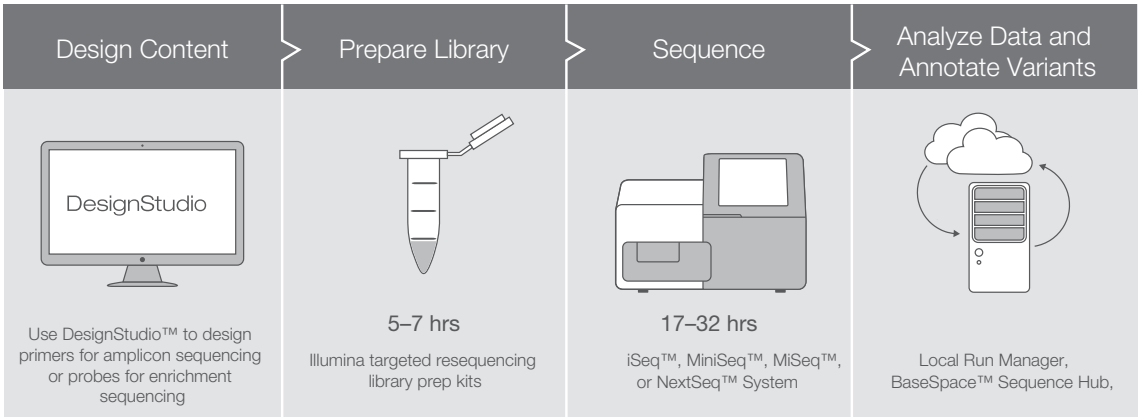

Figure 12: Targeted resequencing workflow—The integrated workflow begins with selection of a predesigned panel or design of custom content, followed by streamlined library preparation, sequencing, and data analysis. Targeted resequencing enables cost-effective studies for a broad range of samples and applications.

Library preparation with ready-to-use vs. custom panels

Gene panels for targeted sequencing can be purchased with read-to-use (predesigned) content (Table 2, Table 4) or they can be custom designed to include genomic regions of interest (Table 3, Table 5).

Ready-to-use panels

Targeted resequencing panels are useful tools for analyzing specific variants in a given gene or region of interest. Predesigned panels contain important genes or gene regions associated with a disease or phenotype. Content is selected by key opinion leaders and from curation of the scientific literature. By far the greatest advantage of working with predesigned content is that the content has already been tested—saving significant time and cost. By focusing on a subset of the genome, these panels also minimize data analysis time and decrease storage requirements. For sample screening or variant discovery, multiple genes can be assessed across many samples in parallel, saving time and reducing costs associated with running separate, iterative assays. Predesigned panels are available for many research areas including cancer, recessive pediatric onset diseases, cardiac conditions, and more (Table 2, Table 4).

Key advantages of ready-to-use panels:

- Sequence key genes or regions of interest to high depth (500–1000× or higher), allowing identification of low-frequency variants
- Save significant time because content has already been selected and tested
- Provide a cost-effective solution to study disease-related genes
- Deliver accurate, easy-to-interpret results, identifying variants at low allele frequencies (down to 1%)
- Enable confident identification of causative or inherited mutations in a single assay

Table 2: Targeted sequencing panels with amplicon-based chemistry

Products	Key Features/Advantages	Total No. of Amplicons, Cumulative Target Size	DNA or RNA Input	Recommended Sequencing Systems
AmpliSeq for Illumina Ready-to-Use Targeted Panels				
AmpliSeq for Illumina ready-to-use targeted sequencing panels provide a fast and cost-effective way to analyze hundreds to thousands of genes without the delays associated with designing and testing amplicons and target selections. Most panels are compatible with samples derived from formalin-fixed, paraffin-embedded (FFPE) tissue. For additional details, visit the AmpliSeq for Illumina Ready-to-Use page.				
AmpliSeq for Illumina BRCA Panel	• Targets exonic and flanking intronic regions of the *BRCA1* and *BRCA2* • Detect variants down to 5% allele frequency • FFPE compatible	265 amplicons 22 kb	1–100 ng DNA	iSeq 100 System (12–96 samples/run) MiniSeq System (24–80 samples/run) MiSeq System (3–96 samples/run)
AmpliSeq for Illumina Cancer Hotspot Panel v2	• Targets hotspot regions of 50 genes associated with pan-cancer and solid tumors • Detect variants down to 5% allele frequency • Compatible with FFPE and blood samples	207 amplicons 22 kb	1–100 ng DNA	iSeq 100 System (16 samples/run) MiniSeq System (32–96 samples/run) MiSeq System (4–96 samples/run)
AmpliSeq for Illumina Comprehensive Cancer Panel	• Full exon coverage of 409 genes associated with pan-cancer and solid tumors • Detect variants down to 5% allele frequency • Compatible with FFPE and blood samples	15,992 amplicons 1.7 Mb	1–100 ng DNA	NextSeq™ 550 System (4–12 samples/run)
AmpliSeq for Illumina Comprehensive Cancer Panel v3	• Targets 161 unique cancer-associated genes in solid tumors • Detect variants down to 5% allele frequency • Compatible with FFPE and blood samples	4648 amplicons	1–100 ng DNA, RNA	MiniSeq System (3 samples/run) MiSeq System (3 samples/run) NextSeq 550 System (16–48 samples/run)
AmpliSeq for Illumina Focus Panel	• Targets biomarkers across 52 genes relevant to solid tumors • Detect variants down to 5% allele frequency • Compatible with FFPE and blood samples	553 amplicons 55 kb	1–100 ng DNA, RNA	iSeq 100 System (8 samples/run) MiniSeq System (16–48 samples/run) MiSeq System (8–48 samples/run)
AmpliSeq for Illumina Childhood Cancer Panel	• Targets 203 genes associated with childhood and young adult cancers • Identify multiple variant types including hotspots, SNVs, indels CNVs, and gene fusions • Compatible with blood, bone marrow, and FFPE tissue	4770 amplicons	10–20 ng DNA, RNA	MiniSeq System (1–25 samples/run) MiSeq System (3–25 samples/run) NextSeq 550 System (27–96 samples/run)
AmpliSeq for Illumina Immune Repertoire Plus, TCR beta Panel	• Provides comprehensive coverage of TCR beta chains, including CDR1, CDR2, and CDR3 • Compatible with blood, fresh/frozen tissue, FACS cells	variable	10–1000 ng RNA	MiniSeq System MiSeq System NextSeq 550 System
AmpliSeq for Illumina TCR beta-SR Panel	• Obtain comprehensive coverage of TCR beta chain sequences • Compatible with FFPE, whole blood, fresh/frozen tissue, or fluorescence-activated cell sorting (FACS) cells	variable	25–1000 ng RNA	MiniSeq System MiSeq System NextSeq 550 System
AmpliSeq for Illumina Immune Response Panel	• Targets 395 genes associated with immune response • Compatible with FFPE and blood samples	42 kb	1–100 ng RNA	MiniSeq System (8–24 samples/run) MiSeq System (16–24 samples/run) NextSeq 550 System (96 samples/run)
AmpliSeq for Illumina Myeloid Panel	• Targets biomarkers across 69 genes associated with hematologic cancers • Detect variants down to 5% allele frequency • Compatible with blood, bone marrow tissue	1226 amplicons ~191 kb	10–20 ng DNA, RNA	MiniSeq System (4–12 samples/run) MiSeq System (7–12 samples/run)
AmpliSeq for Illumina Transcriptome Human Gene Expression Panel	• Measures expression levels of > 20,000 human RefSeq genes • Compatible with FFPE, blood samples	20,802 amplicons 2.2 Mb	1–100 ng RNA	MiniSeq System (3 samples/run) MiSeq System (3 samples/run) NextSeq 550 System (12–40 samples/run)

AmpliSeq for Illumina Community Panels

AmpliSeq for Illumina Community Panels are made-to-order targeted sequencing panels. Illumina provides the designs for customer convenience; however, Illumina Community Panels do not have associated performance metrics. For additional details, visit the Illumina Community Panels page.

- Cardiac Arrhythmias and Cardiomyopathy Research Panel
- Deafness Research Panel v2
- Antimicrobial Resistance Research Panel
- Autism Research Panel
- BRCA Plus, Extended Hereditary Breast and Ovarian Research Panel
- Cardiovascular Research Panel
- Colon and Lung Research Panel v2
- Dementia Research Gene Panel
- Dermatology Research Panel v2
- Dysmorphia-Dysplasia Research Panel v2
- Ebola Research Panel
- Gastrointestinal Research Panel v2
- Hearing Loss Research Panel v1
- Hematology Research Panel
- Inborn Errors of Metabolism Research Panel v2
- Inherited Cancer Research Panel
- Long Non-Coding RNA Research Panel
- Neurological Research Panel
- Noonan Research Panel
- Ophthalmic Research Panel
- Primary Immune Deficiency Research Panel v2
- Pulmonary Research Panel v2
- RNA Fusion Lung Cancer Research Panel
- RNA Inflammation Response Research Panel
- RNA MAPK Pathway Research Panel
- RNA Stem Cell Research Panel
- RNA WNT Signaling Pathway Research Panel
- Renal Research Panel v2
- TB Research Panel
- Endocrine Research Panel v2
- Epilepsy Research Panel

Table 3: Custom sequencing panels with amplicon-based chemistry

Products	Key Features/Advantages	Total No. of Amplicons, Cumulative Target Size	DNA or RNA Input	Recommended Sequencing Systems
AmpliSeq for Illumina Custom Panels				
With AmpliSeq for Illumina Custom Panels, On-Demand Panels, and TruSeq™ Custom Sequencing Panels, content can be easily designed for specific areas of interest. Submit regions of interest through DesignStudio™, a free web-based assay design tool and receive personalized panel content. For more, visit the AmpliSeq for Illumina Custom Panels page and the AmpliSeq On-Demand page.				
AmpliSeq for Illumina Custom DNA Panels	• Select content from any supported species including human, mouse, tomato, rice, and more • Compatible with FFPE samples	12–12,288 amplicons, up to 5 Mb	1–100 ng DNA	iSeq 100 System MiniSeq System MiSeq System NextSeq 550 System
AmpliSeq for Illumina On-Demand Panels	• Select from a catalog of more than 5000 genes with known content relevant for inherited disease research • All content has been pretested with NextSeq 550 System runs • Compatible with FFPE samples	24–15,000 amplicons	1–100 ng DNA	iSeq 100 System MiniSeq System MiSeq System NextSeq 550 System
AmpliSeq for Illumina Custom RNA Panels	• Choose from a catalog of over 20,000 RefSeq genes • Measure gene expression in 12-1200 targets in a single assay • Compatible with FFPE samples	12–1200 amplicons	1–100 ng RNA	iSeq 100 System MiniSeq System MiSeq System NextSeq 550 System

Table 4: Targeted sequencing panels with enrichment-based chemistry

Products	Key Features/Advantages	Total No. of Probes, Cumulative Target Size	DNA or RNA Input	Recommended Sequencing Systems
Nextera Flex for Enrichment Ready-to-Use Sequencing Panels				
Nextera Flex for Enrichment library preparation kits can be used with a range of ready-to-use panels. Nextera Flex Enrichment kits combined with Illumina panels are 85% faster than standard Illumina library prep and enrichment workflows. They enhances library preparation efficiency with integrated protocols for blood and saliva. To learn more, visit the Nextera Flex for Enrichment page.				
• Illumina Exome Panel • TruSight One Panel • TruSight One Expanded Panel	• TruSight Cancer Panel • TruSight Cardio Panel • Third-party oligo vendors			
TruSeq and TruSight Ready-to-Use Sequencing Panels				
TruSeq and TruSight Sequencing Panels with enrichment-based chemistry are ready-to-use targeted sequencing panels that provide a fast and cost-effective way to analyze specific regions of interest without the delays associated with designing and testing probes and target regions.				
TruSight Oncology 500 Panel	• Targeted selection of DNA from 523 and RNA from 55 cancer-related genes • Measures tumor mutation burden (TMB) and microsatellite instability (MSI) biomarkers • Compatible with FFPE samples	1.94 Mb	40 ng DNA 40 ng RNA	NextSeq 550 System NovaSeq System SP
TruSight RNA Fusion Panel	• Targets 507 genes implicated in cancer, including solid tumors, sarcomas, and hematological malignancies • Compatible with FFPE samples	~21,283 probes	10–20 ng RNA	MiniSeq System MiSeq System NextSeq 550 System NovaSeq System SP
TruSight RNA Pan-Cancer Panel	• Targets 1385 genes implicated in cancer pathways including solid tumors, soft tissue cancers, and hematological malignancies • Compatible with FFPE samples	~57,010 probes	10–20 ng RNA	MiniSeq System MiSeq System NextSeq 550 System NovaSeq System SP
TruSight One Sequencing Panel	• Targets 4811 disease-associated genes • Detects targets with disease associations identified in the Human Gene Mutation Database, the Online Mendelian Inheritance in Man catalog, and GeneTests.org	~62,000 probes 12 Mb	50 ng DNA	MiniSeq System MiSeq System NextSeq 550 System HiSeq™ Systems NovaSeq System SP
TruSight One Expanded Sequencing Panel	• Targets 6704 genes • Includes ~1900 additional genes with new disease associations in the reference databases	~86,000 probes 16.5 Mb	50 ng DNA	NextSeq 550 System HiSeq Systems NovaSeq System SP
TruSight Cardio Sequencing Panel	• Targets 174 genes related to 17 inherited cardiac conditions • Cardiac conditions include cardiomyopathies, arrhythmias, aortopathies, and more	0.572 Mb	50 ng DNA	MiSeq System NextSeq 550 System NovaSeq System SP
TruSeq DNA Exome Panel	• Targets over 200,000 coding exons exons • Covers more than 99% of exomic content from RefSeq, ENSEMBL, and GENCODE • Compatible with FFPE samples	45 Mb	100 ng DNA	MiSeq System NextSeq 550 System HiSeq Systems NovaSeq System SP
TruSight Tumor 170 Panel	• Targets 170 genes associated with solid tumors • Detect variants down to 5% allele frequency • Compatible with FFPE samples	533 kb DNA 358 kb RNA	40 ng DNA 40 ng RNA	NextSeq 550 System HiSeq Systems NovaSeq System SP
TruSight Cancer Sequencing Panel	• Targets 94 genes associated with both common and rare cancers • The panel includes 284 SNPs suspected to be associated with cancer	~4000 probes 255 kb	50 ng DNA	MiSeq System NextSeq 550 System HiSeq Systems NovaSeq System SP
TruSeq Neurodegeneration Panel	• Targets 118 genes associated with neurodegenerative diseases • Targets include exons, introns, untranslated regions, and promoters • Covers Alzheimer's, Parkinson's, Amyotrophic Lateral Sclerosis disease and others	43,600 probes 8.7 Mb	≥ 50 ng DNA	MiniSeq System MiSeq System NextSeq 550 System NovaSeq System SP
TruSeq Targeted RNA Expression Panel	• Choose from 400,000+ pre-designed targeted RNA-Seq assays • For human, rat, or mouse species • Compatible with FFPE tissue and low-input samples	12–1000 targets	50 ng DNA	MiniSeq System MiSeq System NextSeq 550 System NovaSeq System SP

Table 5: Custom sequencing panels with enrichment-based chemistry

Products	Key Features/Advantages	Total No. of Probes, Cumulative Target Size	DNA or RNA Input	Recommended Sequencing Systems
Nextera Flex for Enrichment Custom Panels				
With Nextera Flex for Enrichment Custom Panels, content can be easily designed for specific areas of interest. Submit regions of interest through DesignStudio, a free web-based design tool and receive personalized panel content. For additional details, visit the Nextera Flex for Enrichment page.				
Nextera Flex for Enrichment + Illumina Custom Enrichment Panel	• A time-saving solution that is 85% faster than standard Illumina library prep • Select content from the genome of any species including human, mouse, tomato, rice, hamster, chicken, soybean, and more • Compatible with FFPE tissue, blood, and saliva	2000–67,000 probes[a]	10–1000 ng DNA	iSeq System MiniSeq System MiSeq System NextSeq 550 System HiSeq Systems NovaSeq System SP

a. Greater range of probes for third-party panels (from 500 probes to whole exome).

For more information on Illumina Targeted Resequencing Panels, visit the Targeted Panels Page.

Library prep and array kit selector

To find the right library prep kit, researchers can use the Library Prep and Array Kit Selector. This web-based, interactive tool provides a step-by-step process to choose an application-specific library prep kit. The tool helps with kit selection based on starting material, quality of starting material, area of interest, technique, and more (Figure 13).

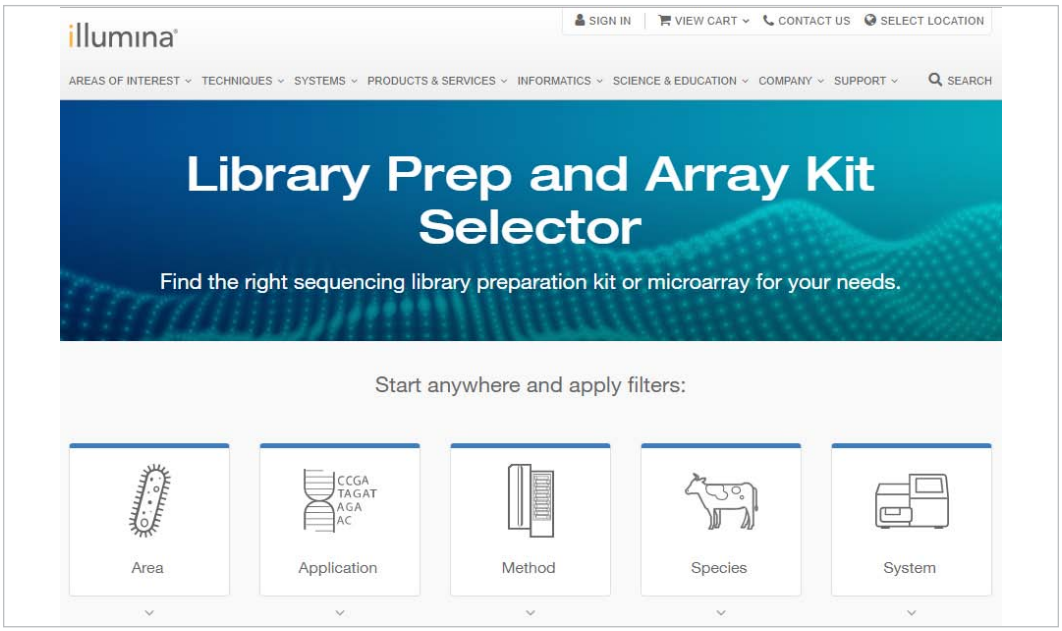

Figure 13: Library prep and array kit selector—This web-based tool helps researchers determine the best kit for their needs based on project type, starting material, and application.

Custom panels with DesignStudio™ Software

Due to the focused nature of ready-to-use targeted gene panels, content may not cover all targets of interest for every researcher. Custom sequencing panels enable gene or target content to be customized based on the research interests of the investigator using DesignStudio Software (Figure 14).

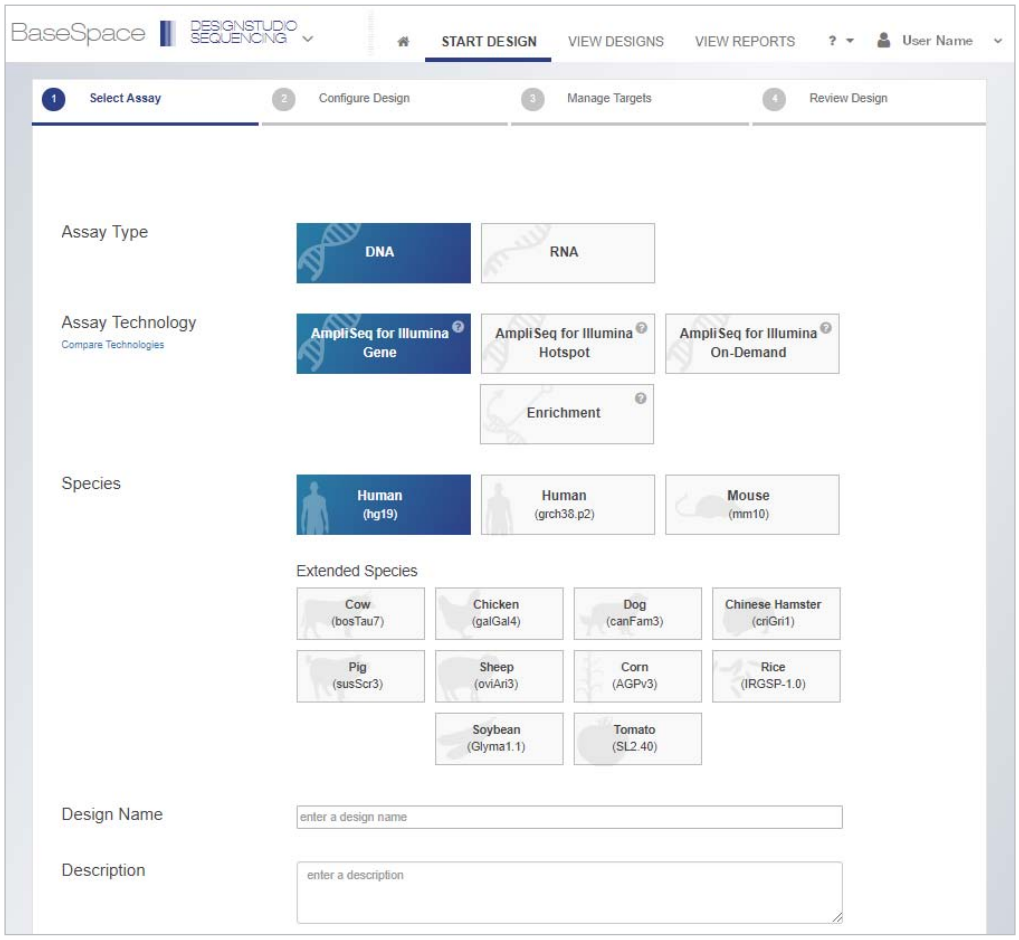

Figure 14: DesignStudio interface—Researchers can use DesignStudio Software to visualize targeted genomic regions and attempted amplicons to assess design coverage and more.

Expanded options with Illumina Concierge Services

Illumina Concierge Services offers additional design support and expanded features for Illumina custom targeted resequencing projects. For example Concierge Services offers assistance with designing smaller amplicons (~100 bp) and increasing compatibility with fragmented DNA, such as DNA purified from FFPE tissue. Illumina Concierge Services can also be used to assist with Nextera Flex for Enrichment and Illumina Custom Enrichment projects. Contact an Illumina representative to inquire about Illumina Concierge Services.

Sequencing chemistry and sequencing systems

After library preparation, the libraries are injected into a flow cell and captured by surface-bound oligos complimentary to library adapters. Each library fragment is then amplified into distinct, clonal clusters through bridge amplification (eg, cluster generation). After cluster generation is complete, templates are ready for sequencing. More than 90% of the sequencing data worldwide is generated by Illumina SBS chemistry.[23] This reversible, terminator–based method detects single bases as they are incorporated into growing DNA strands and enables the parallel sequencing of millions of DNA fragments. Illumina SBS chemistry employs natural competition among all four labeled nucleotides, which reduces incorporation bias and allows more robust sequencing of repetitive regions and homopolymers.

Illumina sequencing systems offer user-friendly, intuitive interfaces for easy run setup and operation at every sequencing scale. The iSeq 100 and MiniSeq Systems are ideal for amplicon, targeted RNA, small RNA, and targeted gene panel sequencing. The MiSeq System supports all these methods as well as small genome sequencing. Compared to all sequencers in the Illumina benchtop portfolio, the NextSeq 550 System offers the highest output and maximum reads per run, enabling exome, transcriptome, and targeted resequencing applications (Figure 15). The NextSeq 550, HiSeq Series, and NovaSeq 6000 Systems support higher output applications such as whole-genome sequencing, exome sequencing, and targeted enrichment (Figure 16).

To learn more about sequencing by synthesis and view a detailed animation of SBS chemistry, visit the SBS Technology page.

	iSeq 100 System	MiniSeq System	MiSeq System	NextSeq 550 System
Max Output	1.2 Gb	7.5 Gb	15 Gb	120 Gb
Max Read Number	8 million	50 million	50 million	800 million

Figure 15: Illumina portfolio of benchtop sequencers for amplicon sequencing

	NextSeq 550 System	HiSeq Series	NovaSeq 6000 System
Max Output	120 Gb	1500 Gb	3000 Gb
Max Read Number	800 million	10 billion	20 billion

Figure 16: Illumina portfolio of sequencers for enrichment sequencing

The sequencing coverage calculator

The sequencing coverage calculator is a web-based tool that helps researchers determine the reagents and sequencing runs needed to arrive at the desired coverage for a particular experiment (Figure 17). Based on the specific application, input parameters, and the sequencing systems chosen, the calculator writes a table containing the number of lanes or flow cells needed for the desired coverage. Results can be downloaded in a comma-separated values (CSV) file that can be shared or edited in Excel.

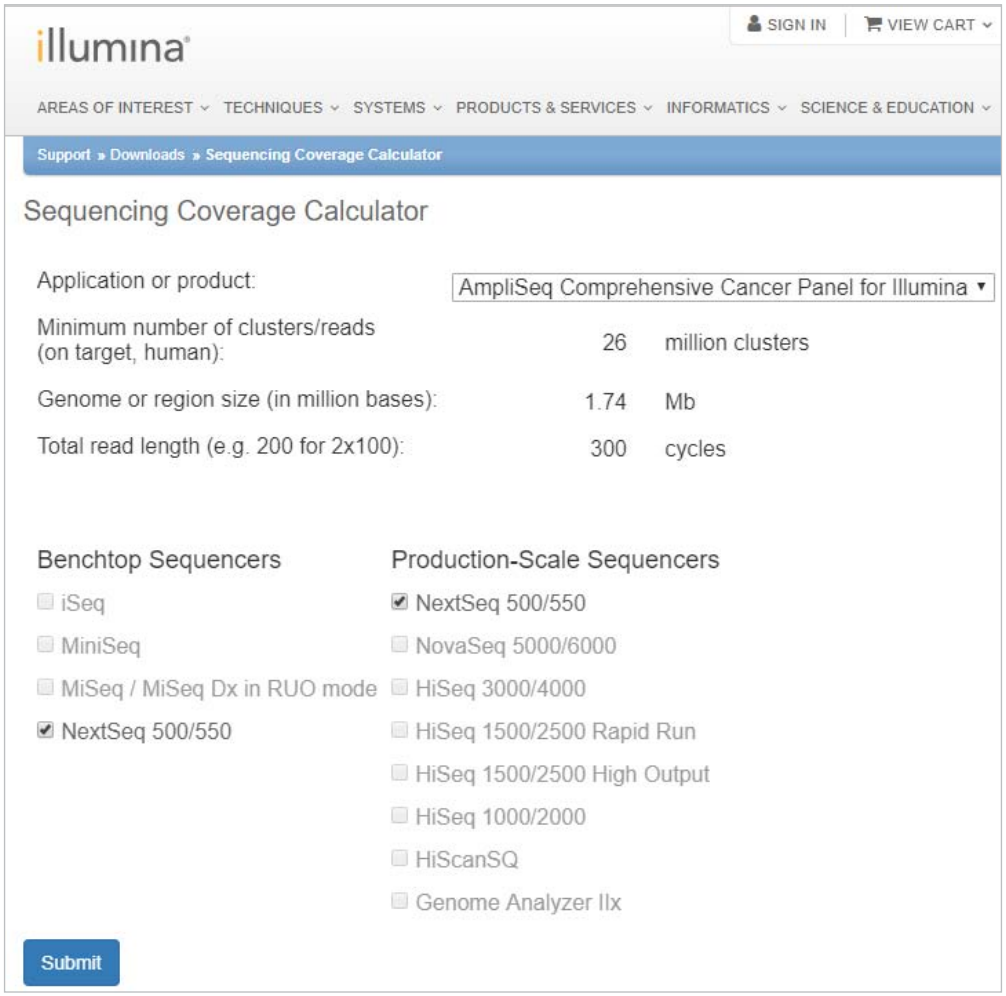

Figure 17: Sequencing coverage calculator—This online tool enables researchers to determine reagents and sequencing runs needed for a particular coverage, before starting an experiment.

User-friendly data analysis

Illumina sequencing systems include a conneciton to BaseSpace™ Sequence Hub, the Illumina genomics computing environment for sequencing data analysis and management. Labs can easily and securely store, analyze, archive, and share sequencing data. Researchers can simplify and accelerate informatics with push-button tools. Labs can also set up and monitor their sequencing runs in real time on any Illumina instrument connected to the BaseSpace Sequence Hub to experience a full sample-to-answer solution (Figure 18).

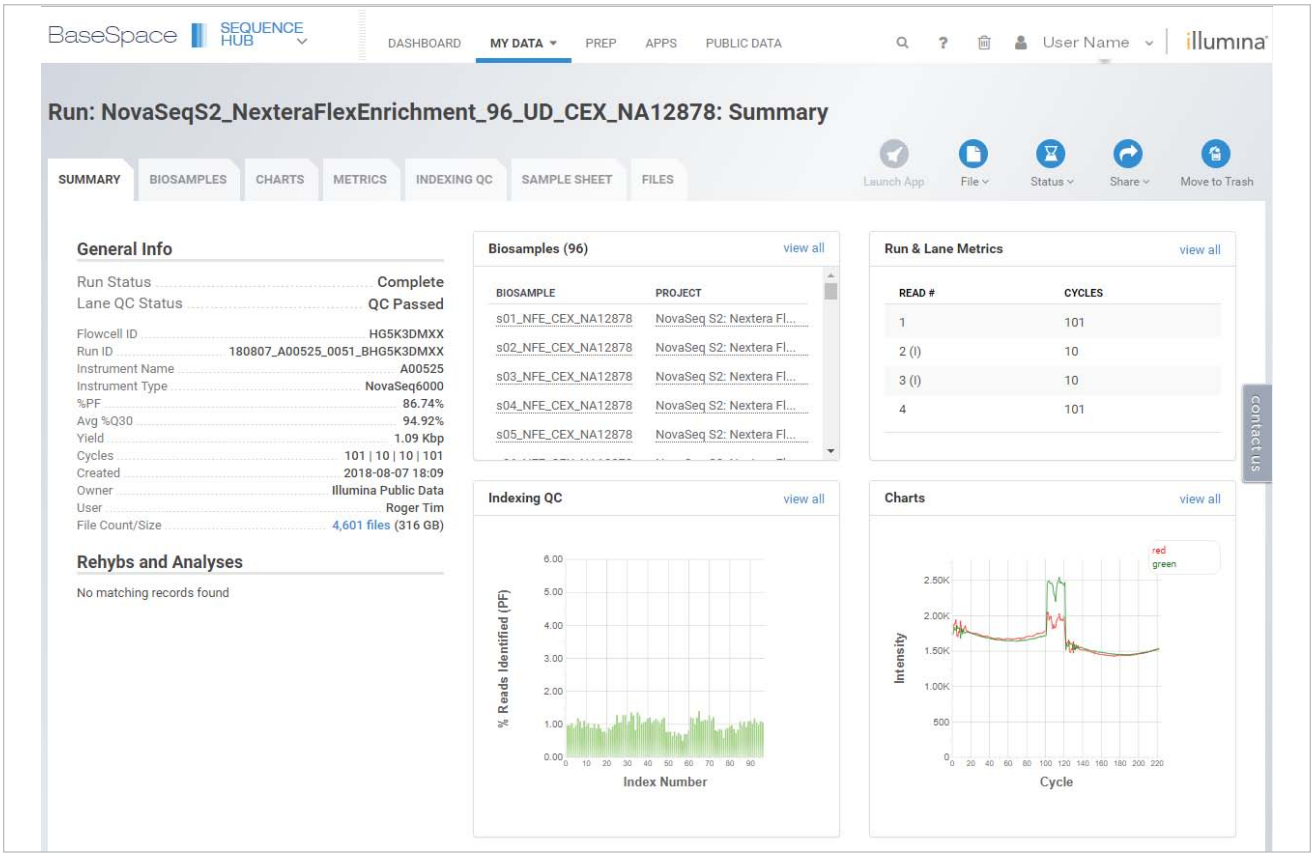

Figure 18: BaseSpace Sequence Hub—The BaseSpace Sequence Hub Interface supports real-time, cycle-by-cycle monitoring. The Charts view shows data by lane and by cycle, with Q-score distribution and heat map features.

> BaseSpace Apps make complex analysis easy for less-experienced bioinformaticians and enable them to trace the analysis steps that were executed."
>
> Dr. Raffaele Calogero, Assoc. Prof , Dept of Molecular Biotechnology and Health Science at the University of Torino

Summary

Labs worldwide are taking advantage of NGS-based targeted resequencing to interrogate somatic and germline variants with a lower cost, faster turnaround time, and lower sample input requirements compared to Sanger sequencing and PCR-based genotyping. With targeted resequencing, researchers can focus their time, expenses, and data storage resources on the most impactful regions of the genome for their area of research. Because targeted resequencing assesses a predefined set of target regions, it enables deep sequencing and higher sensitivity for calling low-frequency variants. For rapid adoption, predesigned targeted resequencing panels featuring expert-selected content allow researchers to avoid the time and expense associated with designing and testing their own panels. Targeted resequencing is a powerful tool that is helping researchers understand the pathogenesis of complex diseases, identify disease-resistant genes in agriculturally important crops, and may one day lead to more effective treatment options for future generations.

Glossary

coverage level: The average number of sequenced bases that align to each base of the reference DNA. For example, a whole genome sequenced at 30× coverage means that, on average, each base in the genome was sequenced 30 times.

deep sequencing: Sequencing to high coverage levels. For example, WGS is typically performed to 30×–75× coverage while targeted NGS enables sequencing depths of 5000× or higher.

discovery power: The ability to identify novel variants.

exome sequencing: An NGS method where only the protein-coding regions (exons) and untranslated regions (UTRs) in a genome are sequenced. The human exome consists of roughly 180,000 exons and constitutes about 2% of the total human genome.

library preparation: A molecular biology protocol that converts a genomic DNA sample (or cDNA sample) into a sequencing library, which can then be sequenced on an NGS instrument. The first step in library preparation is random fragmentation of the DNA sample, followed by ligation of 5' and 3' adapters to each DNA fragment. Alternatively, “tagmentation” combines the fragmentation and ligation reactions into a single step and greatly increases the efficiency of the library preparation process.

multiplexing: An NGS process where unique short DNA sequences are added to each DNA fragment during library preparation. The unique sequences allow many libraries to be pooled together and sequenced simultaneously. Sequencing reads from pooled libraries are identified and sorted computationally before final data analysis. Library multiplexing is a useful technique when working with small genomes or targeting genomic regions of interest. Multiplexing can exponentially increase the number of samples analyzed in a single run, without drastically increasing run cost or run time.

mutation resolution: The size of mutation, in base pairs, a technology is able to detect. For example, karyotyping provides a mutation resolution of 5-10 Mb, while array comparative genomic hybridization provides “higher resolution” by detecting mutations down to 50 kb. NGS techniques provide the highest possible mutation resolution because they can provide single-base pair variant detection (detect the presence of a mutation) and nucleotide identification (detect the identity of a mutation).

read: The process of next-generation DNA sequencing involves using sophisticated instruments to determine the sequence of a DNA or RNA sample. In general terms, a sequence "read" refers to the data string of A,T, C, and G bases corresponding to the sample DNA. With Illumina technology, millions of reads are generated in a single sequencing run.

Sanger sequencing: The sequencing method, also known as capillary electrophoresis sequencing, developed in 1977 by Frederick Sanger. It involves a method of DNA sequencing based on the selective incorporation of chain-terminating dideoxynucleotides by DNA polymerase during *in vitro* DNA replication.

sensitivity: In sequencing, the ability to detect low-frequency, rare variants.

sequencing by synthesis (SBS): SBS technology uses fluorescently labeled nucleotides to sequence tens of millions of sequence templates in parallel. During each sequencing cycle, a single labeled dNTP is added to the nucleic acid chain. The nucleotide label serves as a "reversible terminator" for polymerization: after dNTP incorporation, the label is identified through laser excitation and imaging, then enzymatically cleaved to allow the next round of incorporation. As all four reversible terminator-bound dNTPs (A, C, T, G) are present, natural competition minimizes incorporation bias. Base calls are made directly from signal intensity measurements during each cycle, which greatly reduces raw error rates compared to other technologies. The result is highly accurate base-by-base sequencing that eliminates sequence-context-specific errors, enabling robust base calling across the genome, including repetitive sequence regions and within homopolymers.

sequencing panel: A subset of genes or genomic regions of interest in a targeted resequencing study. The sequencing panel (target regions) can be amplified or enriched using sequence-specific probe sets. The sequencing panel, which represents a smaller subset of the whole genome, can be sequenced for a fraction of the time and cost compared to broader sequencing approaches. Targeted resequencing also enables deeper sequencing of the genomic regions of interest.

somatic mosaicism: Somatic mosaicism occurs when the somatic cells of the body are of more than one genotype. Mosaicism can result from a single fertilized egg cell, due to mitotic errors at first or later cleavages. It can also arise from a mutation during development, which is then propagated to only a subset of the adult cells.

target region: A specific sequence of the genome, identified as a region of interest, due to possible involvement in or association with biological development, pathogenesis, or other area of study of interest to the investigator. The sequence can be a gene, a gene segment, a gene fusion, a promotor region, part of an intron or exon, or any stretch of sequence of interest to the investigator.

targeted resequencing: A subset of genes or regions of the genome are isolated and selectively enriched or amplified before sequencing. Targeted approaches using NGS allow researchers to focus time, expenses, and data analysis on specific areas of interest. Such targeted analysis can include the exome (the protein-coding portion of the genome), specific genes of interest (custom content), targets within genes, or mitochondrial DNA.

References

1. Reprinted from NIH DNA Sequencing Costs Data. Accessed Feb 2019.
2. Fallows J. (2013) When will genomics cure cancer? A conversation with Eric S. Lander. The Atlantic. Accessed Feb 2019.
3. Jamuar SS, Lam AT, Kircher M, et al. Somatic mutations in cerebral cortical malformations. *N Engl J Med.* 2014; 371:733–43.
4. Rivas MA, Beaudoin M, Gardet A, et al. Deep resequencing of GWAS loci identifies independent low-frequency variants associated with inflammatory bowel disease. *Nat Genet.* 2011; 43:1066–73.
5. König K, Peifer M, Fassunke J, et al. Implementation of Amplicon Parallel Sequencing Leads to Improvement of Diagnosis and Therapy of Lung Cancer Patients. *J Thorac Oncol.* 2015; 10:1049–57.
6. Shokralla S, Porter TM, Gibson JF, et al. Massively parallel multiplex DNA sequencing for specimen identification using an Illumina MiSeq platform. *Sci Rep.* 2015; 5: 9687.
7. Rehm HL. Disease-targeted sequencing: a cornerstone in the clinic. *Nat Rev Genet.* 2013; 14:295–300.
8. McEllistrem MC. Genetic diversity of the pneumococcal capsule: implications for molecular-based serotyping. *Future Microbiol.* 2009; 4:857–865.
9. Lo YMD, Chiu RWK. Next-generation sequencing of plasma/serum DNA: an emerging research and molecular diagnostic tool. *Clin Chem.* 2009; 55:607–608.
10. Klee EW, Hoppman-Chaney NL, Ferber MJ. Expanding DNA diagnostic panel testing: is more better? *Expert Rev Mol Diagn.* 2011; 11:703e709.
11. Myllykangas S and Hanlee JP. Targeted deep resequencing of the human cancer genome using next-generation technologies. *Biotechnol Genet Eng Rev*. 2010; 27: 135–158.
12. Kerick M, Isau M, Timmermann B, et al. Targeted high throughput sequencing in clinical cancer settings: formaldehyde fixed-paraffin embedded (FFPE) tumor tissues, input amount and tumor heterogeneity. *BMC Med Genomics.* 2011; 4:68.
13. Shendrue J, and Hanlee J. Next-generation DNA sequencing. *Nat Biotech.* 2008; 26:1135–1145.
14. Schuster SC. Next-generation sequencing transforms today's biology. *Nat Methods.* 2008; 5:16–18.
15. Lindhurst MJ, Sapp JC, Teer JK, Johnston JJ, et al. A mosaic activating mutation in AKT1 associated with the Proteus syndrome. *N Engl J Med.* 2011; 365:611–9.
16. Rohlin A, Wernersson J, Engwall Y, Wiklund L, Björk J, Nordling M. Parallel sequencing used in detection of mosaic mutations: comparison with four diagnostic DNA screening techniques. *Hum Mutat.* 2009; 30:1012–20.
17. Illumina (2017). TruSight™ One Sequencing Panel Data Sheet. Accessed Feb 2019.
18. Chang F, Li MM. Clinical application of amplicon-based next-generation sequencing in cancer. *Cancer Genet.* 2013; 206:413–9.
19. Kou T, Kanai M, Matsumoto S, Okuno Y, Muto M. The possibility of clinical sequencing in the management of cancer. *Jpn J Clin Oncol.* 2016; 46:399–406.
20. Michaut M, Chin SF, Majewski I, et al. Integration of genomic, transcriptomic and proteomic data identifies two biologically distinct subtypes of invasive lobular breast cancer. *Sci Rep.* 2016;6:18517.
21. Jiang Y, Shi X, Zhao Q, et al. Integrated analysis of multidimensional omics data on cutaneous melanoma prognosis. *Genomics.* 2016;107: 223-230.
22. Data calculations on file. Illumina, Inc. 2015.

Industry-Leading Solutions

A Community of Support

From library prep, arrays, and sequencing to informatics, Illumina genomic solutions empower researchers and clinical researchers across the globe to find the answers they seek.

When you join the Illumina community, you become part of a dynamic scientific movement that includes thousands of researchers and industry thought leaders. Throughout the year, we host user group meetings, symposia, consortia, online forums, and other initiatives—all designed to bring the best minds together to share ideas and advance science.

In addition to on-site training, ongoing support, and phone consults, we offer webinars and courses at various Illumina locations. We're here with all the resources you need to accelerate progress.

Contact your Illumina representative to learn more or visit
www.illumina.com/targeted-resequencing

A global genomics leader, Illumina provides complete next-generation sequencing workflow solutions to the basic and translational research communities. Illumina technology is responsible for generating more than 90% of the world's sequencing data.* Through collaborative innovation, Illumina is fueling groundbreaking advancements in the fields of oncology, reproductive health, genetic disease, microbiology, and agriculture.

*Data calculations on file. Illumina, Inc., 2017.

Illumina • 1.800.809.4566 toll-free (US) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com

For Research Use Only. Not for use in diagnostic procedures.

illumina®