illumina®

Targeted Resequencing Solutions on the MiniSeq™ System

The MiniSeq System delivers an accessible, cost-effective solution for targeted resequencing applications.

Highlights

- **Highly Focused, Manageable Studies**
  Focuses on regions of interest, generating smaller, more manageable data sets.
- **Higher Coverage Levels**
  Enables deep sequencing at high coverage levels for rare variant identification
- **Lower Costs and Smaller Data Sets**
  Reduces sequencing costs and data analysis burdens
- **Fast Sample-to-Answer Time**
  Reduces turnaround time compared to broader approaches

Figure 1: Targeted Resequencing—Targeted Resequencing on the MiniSeq System enables focused, deep sequencing for identification of rare variants.

Introduction

Scientists around the world have identified thousands of disease-associated loci using Illumina sequencing and array technology. After identifying broad genomic regions through genome-wide association studies (GWAS), whole-genome sequencing (WGS), or whole-exome sequencing (WES) studies, targeted resequencing (Figure 1) is often performed to investigate more precise target regions at greater depth.

Targeted resequencing offers several advantages: it efficiently and cost-effectively focuses the power of next-generation sequencing (NGS) on a subset of genes or genomic regions. It allows sequencing at much higher coverage levels, providing a virtually unlimited dynamic range and higher sensitivity. Targeted resequencing can reveal variants that would be too expensive or impossible to identify with WGS, PCR, or capillary electrophoresis (CE) sequencing. The ability to detect rare variants can lead to the identification of novel functional variants, facilitate biomarker discovery, or lead to the identification of clinically relevant targets for translational research.[1] Amplicon sequencing is particularly useful for the discovery of rare somatic mutations in complex samples such as cancerous tumors mixed with germline DNA.[2, 3] Whether performing a GWAS follow-up study, or profiling samples for cattle breeding or crop selection, or performing translational research, users can target regions of the genome relevant to their specific interests.

Simple, Integrated Workflow

The MiniSeq System offers the most affordable and accessible NGS workflow for targeted resequencing (Figure 2). In addition to custom panels, targeted resequencing panels can be purchased with preselected content. A wide variety of targeted sequencing library prep kits are available, including kits with probe sets focused on cancer, cardiomyopathy, inherited diseases, and more (Table 1).

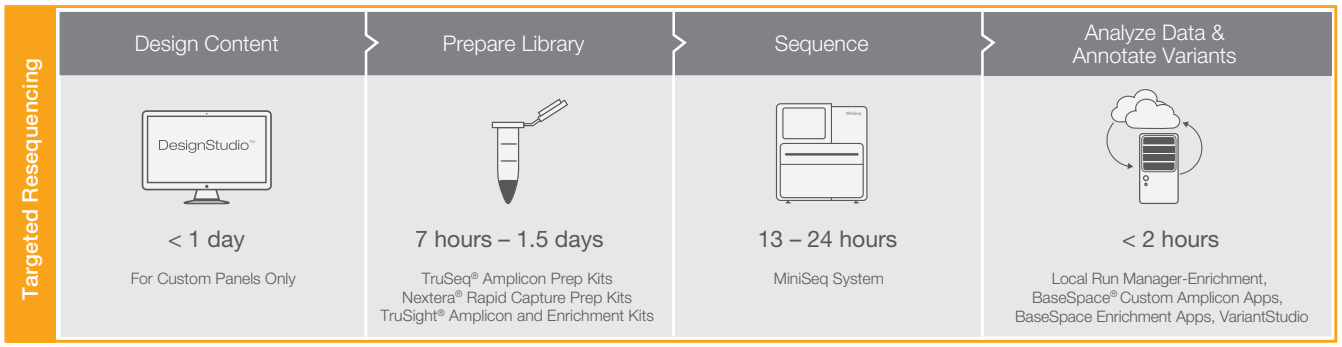


Figure 2: MiniSeq System Targeted Resequencing Workflow—The integrated workflow enables streamlined library preparation, sequencing, and data analysis, allowing cost-effective studies for a broad range of samples.

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

Library Preparation

Illumina Targeted Resequencing Methods

Illumina currently supports 2 methods for targeted resequencing—capture-based target enrichment and amplicon generation (Figure 3). With target enrichment, specific regions of interest are captured by hybridization to biotinylated probes, then isolated by magnetic pulldown. This highly multiplexed approach enables a wide range of applications for the discovery, validation, or screening of genetic variants. The second method, amplicon sequencing, involves the amplification and purification of regions of interest using highly multiplexed oligo sets.

Predesigned Targeted Sequencing Panels

Targeted sequencing panels are useful tools for analyzing specific mutations in a given gene or region of interest. Predesigned panels contain important genes or gene regions associated with a disease or phenotype, selected from publications and expert guidance. By focusing specific genes or regions, these panels conserve resources, minimize data analysis time, and decrease storage requirements. For sample screening, or variant identification, 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 several research areas including cancer, inherited disorders, cardiac conditions, and more (Table 1).

Custom Sequencing Panels and DesignStudio™

For specific regions of interest, researchers can design and order custom panels with DesignStudio. To get started with DesignStudio, simply upload a list of targets identified by GWAS, WGS, or microarray experiments. Quickly build a custom panel with up to thousands of amplicons (depending on the kit) or add new targets to a previously ordered panel. DesignStudio provides dynamic feedback to optimize target region coverage, reducing the time required to design custom projects. Custom target enrichment captures between 10 kb–62 Mb regions depending on the library prep kit parameters. Custom amplicon sequencing allows researchers to sequence 16–1536 (or more with Illumina Concierge) amplicons at a time, spanning 2.4–652.8 kb of total content, depending on the library prep kit used.

Expanded Options With Illumina Concierge

Illumina Concierge services offer additional design support and expanded features for Illumina custom targeted resequencing projects. Some custom targeted sequencing kits incorporate unique molecular identifiers for enhanced allelic detection and increased sensitivity.[4] Unique molecular identifiers allow the removal of PCR duplicates, which enables the detection of individual molecules. The TruSeq Custom Amplicon Assay is compatible with dual-strand sequencing, which eliminates false positives that can arise from deamination events during formalin fixation or from other DNA lesions. Illumina Concierge also offers the ability to design smaller amplicons (~100 bp), increasing compatibility with fragmented DNA, such as DNA from formalin-fixed, paraffin-embedded (FFPE) tissue. Contact an Illumina representative for access to Illumina Concierge services.

Sequencing on the MiniSeq System

The MiniSeq System is designed for easy, streamlined operation (Figure 4). For quick, load-and-go set up, the MiniSeq Reagent Kits provide a flow cell, wash reagents, and a single reagent cartridge preloaded with all required sequencing reagents. The reagent kits are available in Mid-Output and High-Output formats, allowing optimization of study designs based on read-length, sample number, and output requirements.

The MiniSeq System offers an intuitive touch screen interface that provides simple, step-by-step guidance through each stage of the sequencing run, including library and reagent loading, run configuration, and run monitoring. MiniSeq Control Software performs onboard image analysis, base calling, and quality scoring. Quality statistics from 1 or multiple runs can be monitored in real time using the Sequencing Analysis Viewer (SAV) software. SAV software can be used onboard the sequencing instrument, or can be accessed from any location with a Windows-based PC.

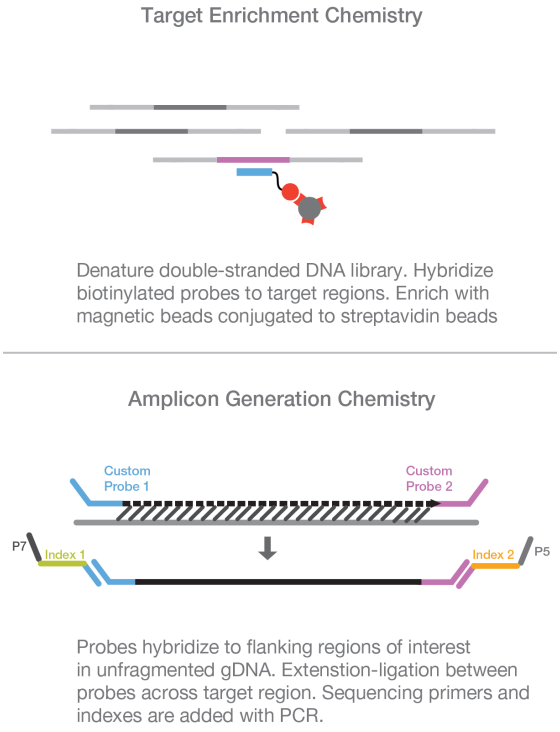


Figure 3: Targeted Resequencing Methods—Illumina offers 2 methods for targeted resequencing library preparation: targeted enrichment and amplicon generation.

Figure 4: MiniSeq System—The MiniSeq System leverages the latest advances in SBS chemistry and an easy, integrated workflow.

Table 1: Illumina Targeted Resequencing Solutions

	Key Features/Advantages	Cumulative Target Region Size/ Number of Probes or Amplicons	DNA Input	Number of Samples per MiniSeq Run[a]
Custom Targeted Sequencing Panels				
Nextera Rapid Capture Custom	• Enrich custom content • 1.5 day library prep	0.5–15 Mb 3000–67,000 probes[b]	50 ng	1-96 samples/run[c]
TruSeq Custom Amplicon v1.5	• **FFPE compatible** • Amplify custom content	2–650 kb 16–1536 amplicons[d]	50 ng for gDNA 150 ng for FFPE	1-96 samples/run[c]
TruSeq Custom Amplicon Low Input	• **FFPE compatible** • Amplify custom content • Low DNA input amount	2–650 kb 16–1536 amplicons	10 ng for gDNA 10-50 ng for FFPE[e]	1-96 samples/run[c]
Predesigned Targeted Sequencing Panels				
TruSight® One Panel	• Targets 4813 genes associated with known clinical phenotypes • 1.5 day library prep	12 Mb	50 ng	3 samples/run
TruSight Cardio Panel	• Targets 174 genes related to 17 inherited cardiac conditions • 1.5 day library prep	244 kb	50 ng	12 samples/run
TruSight Inherited Disease Panel	• Targets 552 genes related to severe, recessive pediatric diseases • 8801 target exons	2.25 Mb ~30,000 probes	50 ng	8 samples/run
Predesigned Targeted Sequencing Panels for Cancer Research				
TruSight Tumor 15	• **FFPE compatible** • Targets 15 genes commonly mutated in solid tumors • Detect variants down to 5% allele frequency	44 kb 250 amplicons	20 ng	8 samples/run
TruSight Myeloid Sequencing Panel	• Targets 54 genes focused on somatic mutations in myeloid malignancies • Detect variants down to 5% allele frequency	141 kb 568 amplicons	50 ng	8 samples/run
TruSight Cancer Panel	• Targets 94 genes associated with a predisposition towards cancer • Detect variants down to 5% allele frequency	255 kb ~4000 probes	50 ng	24 samples/run
TruSeq Amplicon Cancer Panel	• **FFPE compatible** • Targets 48 genes with mutational hotspots in frequently mutated cancer genes	> 35 kb 212 amplicons	150 ng for gDNA 250 ng for FFPE	42 samples/run

a. With High-Output MiniSeq Sequencing Kit
b. Probe number depends on kit configuration
c. Sample throughput varies with experimental design and mean coverage
d. More amplicons available with Illumina Concierge
e. Input amount depends on QC results with TruSeq FFPE DNA Library Prep QC Kit

Selecting Sequencing Depth for Amplicon Sequencing

Sequencing coverage (sensitivity) describes the average number of reads that align to, or "cover," known reference bases. Coverage level often determines whether variant discovery can be made with a certain degree of confidence at particular base positions. At higher coverage levels, base calls can be made with a higher degree of confidence and the ability to detect rare variants increases. The following are coverage guidelines to achieve appropriate detection levels for certain studies:

- Heterozygote detection—40× coverage
- 5% variation of single-base changes and multibase deletions—1000× coverage
- 1% variation of single-base changes and multibase deletions— ≤ 5000× coverage
- Single-base indels might require additional depth

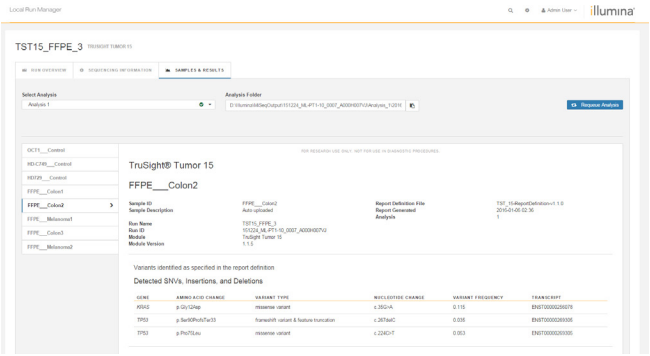


Figure 5: Local Run Manager User Interface—With Local Run Manager, runs can be set up, organized, and analyzed directly on the sequencing instrument.

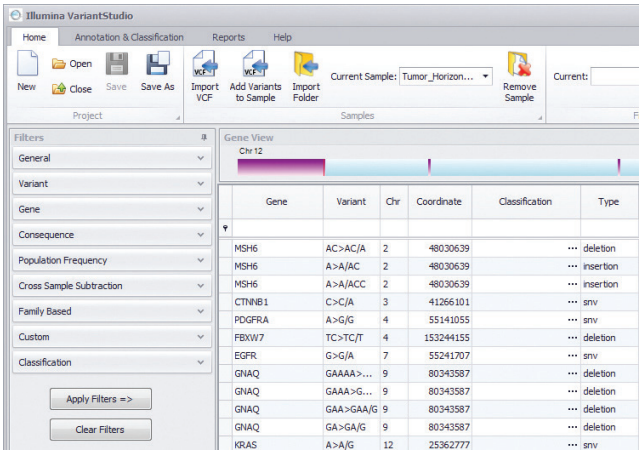


Figure 6: VariantStudio User Interface—Quickly identify, classify, and report disease-relevant variants with Illumina VariantStudio annotation software.

Simplified Bioinformatics

Data analysis with the MiniSeq System requires no informatics expertise or command-line experience. The MiniSeq System features Local Run Manager software, an on-instrument system for creating a run, monitoring status, and analyzing sequencing data (Figure 5). With Local Run Manager, on-instrument data analysis can be automatically performed upon completion of the sequencing run. The data analysis modules generate simple reports for a wide range of sequencing applications. The modular design allows users to install and update individual analysis modules as needed.

In addition, sequencing data generated with the MiniSeq System can be instantly transferred, stored, and analyzed in the BaseSpace computing environment (cloud-based or onsite). BaseSpace Targeted Resequencing Software Apps provide expert-preferred data analysis tools packaged in an intuitive, click-and-go user interface designed for informatics novices (Figure 5). These Apps deliver optimized pipelines that support a range of common sequencing data analysis needs such as alignment, variant calling, and more. For enrichment workflows, the BaseSpace Isaac™ Enrichment App[5] aligns targeted sequence reads with the ultrafast Isaac Aligner[6] and performs variant calling with the Starling Variant Caller.[6] For amplicon workflows, the TruSeq Amplicon App[7] performs a banded Smith-Waterman alignment and enables variant calling with the genome analysis toolkit (GATK 1.6),[8] Isaac Variant Caller,[6] or the Illumina-developed Somatic Variant Caller.[9]

For downstream analysis, the Illumina VariantStudio analysis software enables identification and classification of disease-relevant variants as well as generation of structured, detailed reports (Figure 6). Additionally, BaseSpace Apps generate output files that can be directly input into a broad range of data analysis tools. The BaseSpace Environment includes a growing community of developers who use and provide software tools for visualization, analysis, and sharing. This NGS ecosystem provides one of the largest collections of commercial and open-source analysis tools currently available.

NGS Targeted Resequencing vs Traditional Technologies

While traditional methods, such as CE-based sequencing and PCR can be used to interrogate specific regions of interest, NGS targeted resequencing provides the most cost-effective approach to sequencing the broadest regions of interest with the highest sensitivity (Table 2).

Table 2: Comparison of CE Sequencing, q/RT-PCR, and NGS Targeted Resequencing

	CE Sequencing	q/RT-PCR	Targeted Resequencing
Benefits	• Cost-effective sequencing for small stretches[a] of DNA sequence • Quick and simple workflow • Current gold standard in sequencing	• High sensitivity[b] • Quick and simple workflow • Capital equipment already placed in most labs	• Higher sequencing depth enables higher sensitivity (down to 1%)[b] • Higher discovery power (screen hundreds of genes simultaneously) • Higher mutation resolution (nucleotide identity can be determined) • Produce more data with the same amount of input DNA[d] • Higher sample throughput with sample multiplexing
Challenges	• Low sensitivity (down to 20%)[b] • Low discovery power • Not as cost-effective for large stretches[c] of DNA sequence • Low scalability due to increasing sample input requirements	• Can only interrogate a limited set of mutations • Virtually no discovery power • Limited mutation resolution • Low scalability due to increasing sample input requirements	• Not as cost-effective for sequencing small stretches[a] of DNA sequence • Not as time-effective for sequencing small stretches[a] of DNA sequence

a. small stretches = less than ~15-20 amplicons
b. sensitivity = allele frequency limit of detection
c. large stretches = more than ~15-20 amplicons
d. 10 ng DNA will produce ~1 kb with CE sequencing or ~300 kb with targeted resequencing (250 bp amplicon length × 1536 amplicons with TruSeq Custom Amplicon workflow)

Demonstrated Workflow: TruSeq Custom Amplicon Low Input

Custom Design

A known set of gene targets, listed in Excel format, was uploaded to DesignStudio. After uploading the gene list, the following parameters were selected:

- **Assay Version**—TruSeq Custom Amplicon Low Input
  Rationale: To leverage the low input feature of the TruSeq Custom Amplicon Low Input Kit.
- **Variant Source**—dbSNP
  Rationale: The most relevant SNP source for this assay design, with all world-populations selected.
- **Amplicon Length**—175 bp
  Rationale: Demonstrate performance with highly fragmented FFPE DNA.

DesignStudio generated a probe set of 144 amplicons with 100% of the submitted targets covered (Figure 7). After reviewing the undesignable gaps in the UCSC Genome Browser,* we approved the design and ordered the probe set through DesignStudio.

Library Preparation

The TruSeq Custom Amplicon Low Input library preparation method is bead-based, utilizing a hyb-extension-ligation process (Figure 3). The optional TruSeq FFPE QC companion kit (Cat No. FC-131-9999) is recommended to assess FFPE sample quality and to provide DNA input amount recommendations. Libraries were prepared according to the TruSeq Custom Amplicon Low Input Sample Prep Guide.10 The TruSeq Custom Amplicon libraries were generated from 3 tumor-normal paired samples consisting of highly degraded FFPE lung, stomach, and rectal tissue samples. Each sample library was prepared from 10-50 ng of total input FFPE DNA depending on the TruSeq FFPE QC Kit assessment results and input recommendations. Library QC revealed that all libraries had sufficient yield for cluster generation and sequencing on the MiniSeq System. Each tumor-normal pair was prepped in duplicate, and all 12 samples were pooled together for the MiniSeq System run.

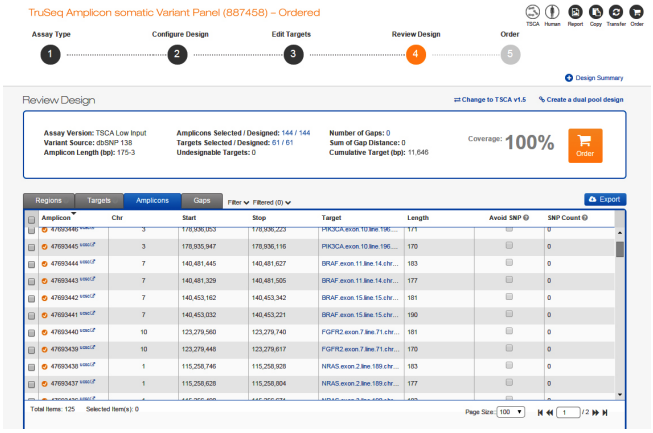


Figure 7: DesignStudio User Interface—DesignStudio screen shot showing designed targets at the Review Design step. Users can view the coverage gaps, if any, by linking out to the UCSC Genome Browser. The final step in DesignStudio allows the user to order the probe set and/or save the design for future use.

* A link to the UCSC Genome Browser is provided from DesignStudio

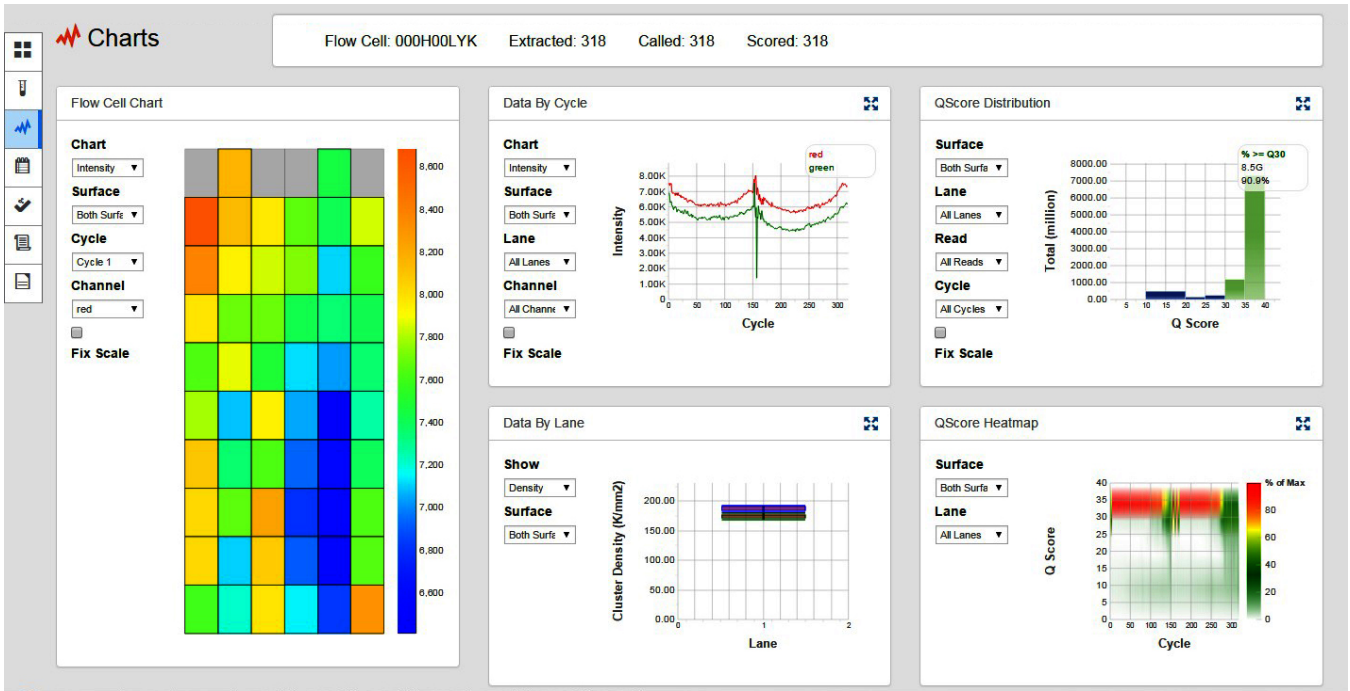


Figure 8: Run Monitoring in BaseSpace—Run progress was monitored in real time with BaseSpace run progress charts. Data by Cycle, Q-score Distribution, Data by Lane, Tile by Tile metrics, and more were viewed during the run.

Sequencing on the MiniSeq System

The pooled libraries were loaded onto the MiniSeq instrument along with the reagent cartridge and flow cell. Automated cluster generation and a 2 × 150 read length run were set up with Local Run Manager and performed without further user intervention. The sequence run took approximately 24 hours. Run progress was monitored (Figure 8) and final run metrics were generated for review on BaseSpace.

Data Analysis

Image analysis and base calling were performed on the MiniSeq System. Demultiplexing, alignment, and variant calling were performed with the BaseSpace TruSeq Amplicon App. Finally, variant filtering and annotation were performed with VariantStudio (accessible via BaseSpace). Summary tables were generated to report on-target %, coverage uniformity, and additional variant calling statistics (Figure 9). With this demonstrated workflow, 93.28% on-target coverage (average of Read 1 and Read 2 percent aligned reads) and 94.3% coverage uniformity were achieved across all 6 highly degraded FFPE samples.

Summary

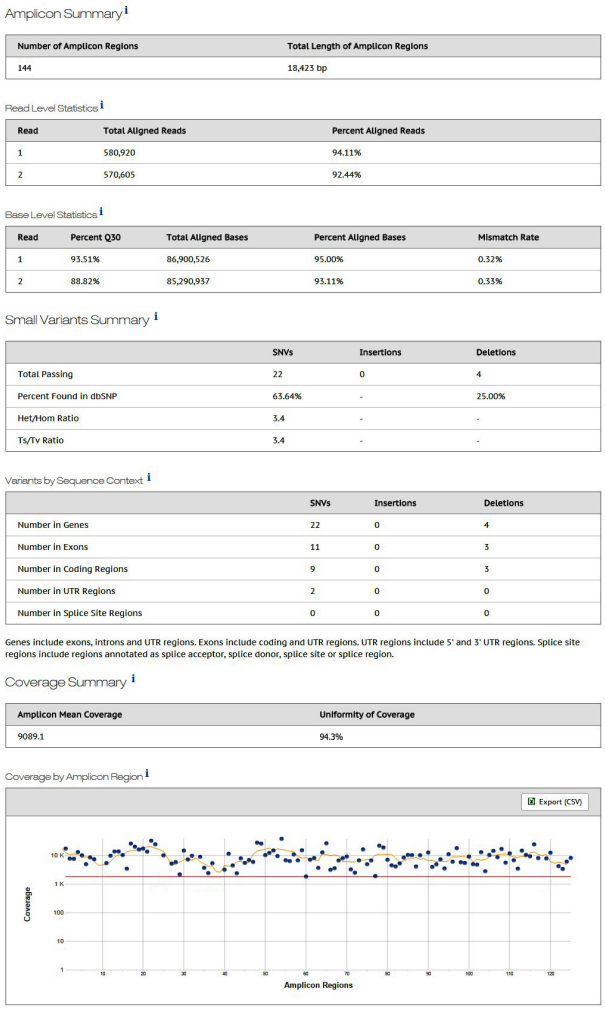
The MiniSeq System Targeted Resequencing Solution offers a highly sensitive and accurate method for analyzing specific genes or regions of interest. By harnessing the broad dynamic range of NGS sequencing, researchers can obtain more sensitive and accurate measurements for specific genes or regions of interest. Whether looking for the speed of a fixed panel or the flexibility of a custom panel, the MiniSeq System Targeted Resequencing Solution delivers high-quality NGS data in a more accessible, cost-effective platform.

Learn More

For more on DesignStudio, go to: www.illumina.com/informatics/research/experimental-design/designstudio.html.

To learn more about targeted gene panels, visit: www.illumina.com/techniques/sequencing/dna-sequencing/targeted-resequencing/targeted-panels.html.

For more on amplicon sequencing, go to: www.illumina.com/techniques/sequencing/dna-sequencing/targeted-resequencing/amplicon-sequencing.html.



Amplicon Summary

Number of Amplicon Regions	Total Length of Amplicon Regions
144	18,423 bp

Read Level Statistics

Read	Total Aligned Reads	Percent Aligned Reads
1	580,920	94.11%
2	570,605	92.44%

Base Level Statistics

Read	Percent Q30	Total Aligned Bases	Percent Aligned Bases	Mismatch Rate
1	93.51%	86,900,526	95.00%	0.32%
2	88.82%	85,290,937	93.11%	0.33%

Small Variants Summary

	SNVs	Insertions	Deletions
Total Passing	22	0	4
Percent Found in dbSNP	63.64%	-	25.00%
Het/Hom Ratio	3.4	-	-
Ts/Tv Ratio	3.4	-	-

Variants by Sequence Context

	SNVs	Insertions	Deletions
Number in Genes	22	0	4
Number in Exons	11	0	3
Number in Coding Regions	9	0	3
Number in UTR Regions	2	0	0
Number in Splice Site Regions	0	0	0

Genes include exons, introns and UTR regions. Exons include coding and UTR regions. UTR regions include 5' and 3' UTR regions. Splice site regions include regions annotated as splice acceptor, splice donor, splice site or splice region.

Coverage Summary

Amplicon Mean Coverage	Uniformity of Coverage
9089.1	94.3%

Coverage by Amplicon Region

Figure 9: Targeted Resequencing Data Analysis in the BaseSpace Cloud—The TruSeq Amplicon App in BaseSpace simplifies data analysis, delivering results in an intuitive format. Metrics for aligned read percentage, variant calls, and coverage uniformity are shown here for the MiniSeq System sequencing run.

References

1. Rivas MA, Beaudoin M, Gardet A, et al. Deep resequencing of GWAS loci identifies independent rare variants associated with inflammatory bowel disease. *Nat Genet.* 2011;43:1066-73.
2. McEllistrem MC. Genetic diversity of the pneumococcal capsule: implications for molecular-based serotyping. *Future Microbiol*. 2009;4:857-865.
3. Lo YMD, Chiu RWK. Next-generation sequencing of plasma/serum DNA: an emerging research and molecular diagnostic tool. *Clin Chem*. 2009;55:607-608.
4. Kivioja T, Vähärautio A, Karlsson K, et al. Counting absolute numbers of molecules using unique molecular identifiers. *Nat Methods*. 2011;9:72-74.
5. BaseSpace Isaac Enrichment App (www.illumina.com/informatics/research/sequencing-data-analysis-management/basespace/basespace-apps/isaac-enrichment-1253252.html). Accessed 15 Dec 2015.
6. Raczy C, Petrovski R, Saunders CT, et al. Isaac: ultrafast whole-genome secondary analysis on Illumina sequencing platforms. Bioinformatics. 2013;29:2041-2043.
7. BaseSpace TruSeq Amplicon App (www.illumina.com/informatics/research/sequencing-data-analysis-management/basespace/basespace-apps/truseq-amplicon-2005003.html). Accessed 04 January 2016.
8. Genome Analysis Toolkit (GATK) (www.broadinstitute.org/gatk/).
9. Somatic Variant Caller (www.illumina.com/documents/products/technotes/technote_somatic_variant_caller.pdf). Accessed 06 January 2016.
10. TruSeq Custom Amplicon Low Input Library Prep Reference Guide (support.illumina.com/downloads/truseq-custom-amplicon-low-input-library-prep-reference-guide-1000000002191.html). Accessed 30 December 2015.

Ordering Information

Sequencing System	Catalog No.
MiniSeq System	SY-420-1001
Sequencing Kits	
MiniSeq High Output Kit (75 Cycles)	FC-420-1001
MiniSeq High Output Kit (150 Cycles)	FC-420-1002
MiniSeq High Output Kit (300 Cycles)	FC-420-1003
MiniSeq Mid Output Kit (300 Cycles)	FC-420-1004
Custom Targeted Sequencing Kits	
Nextera® Rapid Capture Custom Kits (48 samples)	FC-140-1007
Nextera® Rapid Capture Custom Kits (96 samples)	FC-140-1008
Nextera® Rapid Capture Custom Kits (288 samples)	FC-140-1009
TruSeq Custom Amplicon v1.5 (96 samples)	FC-130-1001
TruSeq Custom Amplicon Low Input (96 samples)	FC-134-2001
TruSeq Custom Amplicon Low Input (16 samples)	FC-134-2002
TruSeq FFPE DNA Library Prep QC Kit	FC-121-9999
TruSeq Custom Amplicon Index Kit (96 indexes, 384 samples)	FC-130-1003
TruSeq Index Plate Fixture Kit	FC-130-1005
TruSeq Index Plate Fixture and Collar Kit (2 each)	FC-130-1007
Predesigned Targeted Sequencing Kits	
TruSight One (9 samples)	FC-141-1006
TruSight One (36 samples)	FC-141-1007
TruSight Cardio (12 samples)	FC-141-1010
TruSight Cardio (48 samples)	FC-141-1011
TruSight Inherited Disease Panel (4 enrichments)	FC-121-0205
Predesigned Targeted Sequencing Kits for Cancer	
TruSight Tumor 15 Includes library preparation consumables, oligos, and indexes sufficient for 24 samples	OP-101-1002
TruSight Tumor 15 MiniSeq Kit Includes library preparation panel and 3 MiniSeq High Output Kits (300 Cycles), sufficient for 24 samples	20005610
TruSight Myeloid Sequencing Panel (96 samples)	FC-130-1010
TruSight Cancer Panel Includes oligos sufficient for 4 enrichments and up to 48 samples	FC-121-0202
TruSight Cancer MiniSeq Bundle Includes oligos, library prep panel, and 2 MiniSeq High Output Kits (300 Cycles), sufficient for 48 samples	20005612
TruSeq Amplicon Cancer Panel (96 samples)	FC-130-1008

Maximize Performance and Productivity with Illumina Services, Training, and Consulting

Illumina service and support teams provide a full suite of expedient, customized solutions from initial trainings, to instrument support, and ongoing NGS education. Our support offerings include:

Illumina Professional Care Services Packs

Illumina offers Professional Care Services Packs - allotments of points that can be redeemed for discounted Illumina Professional Services. Benefits include:

- One-time Investment - no need for additional, postsale expenditures
- Risk Mitigation - bank points for unanticipated future services
- Savings - cost-effective versus a la carte pricing

Professional Care Services

Product Care Services

- Tiered Instrument Service Plans + Add-On Services
- Instrument Compliance Services
- Instrument On-Demand Services

Illumina University Training

- Instructor-Led Training at Your Chosen Facility
- Instructor-Led Training at an Illumina Training Center
- Online Courses and Webinars

Illumina Consulting

- Proof-of-Concept Services for instrument and library preparation testing
- Concierge Services for design assistance and product optimization

For more on Illumina support offerings, visit: www.illumina.com/services/instrument-services-training.html

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.

 Pub. No. 770-2015-047 Current as of 21 May 2016

illumina®