Agilent

Trusted Answers

Whole-Exome Sequencing with Expertly Curated Clinical Content Powered by the Genome

Key Advantages

– Enhanced identification of disease-causing genetic variants, powered by the genome, with the speed and ease of an exome sequencing workflow

– Comprehensive coverage of 41.1 Mb of human protein-coding and selected non-coding regions for optimal clinical research, curated by Dr. Madhuri Hegde, PhD, FACMG, formerly of Emory University

– Highly uniform enrichment for efficient sequencing

– Native automation on Bravo platform and Magnis system for increased productivity

– Streamlined integration with cloud-based Alissa Clinical Informatics Platform for faster time-to-answers

Introduction

The Agilent SureSelect Clinical Research Exome V4 (CRE V4) has been designed to provide comprehensive coverage of human protein coding regions, as well as clinically relevant content thoughtfully curated by world-renowned clinical molecular geneticist Dr. Madhuri Hegde, PhD, FACMG, formerly of Emory University. The SureSelect Clinical Research Exome spans over a decade of systematic enhancements as more insights from the genome become available. The SureSelect CRE V4 leverages valuable insights and discoveries from whole-genome sequencing (WGS) to create an exome that bridges the gap between whole-exome sequencing (WES) and WGS, with unique features such as curated deep intronic sites and mini-genomes. CRE V4 provides enhanced coverage of more than 6,500 disease-associated genes. In addition, CRE V4 provides excellent coverage of targeted non-coding pathogenic sites from ClinVar, HGMD, and ACMG, ancestry single nucleotide variants (SNPs), pharmacogenomic SNPs, a unique design to detect common disease-causing CNVs using a breakpoint detection approach, and other expertly curated disease-relevant content.

The SureSelect CRE V4 spans a 41.1 Mb target region of the human genome, with an efficient end-to-end design size of 48.5 Mb. CRE V4 delivers excellent enrichment performance and efficient sequencing, ensuring enhanced identification of disease-causing variants with minimal increase in sequencing requirement and without the expense, infrastructure burden, and informatic complexities of WGS.

The SureSelect CRE V4 is compatible with the streamlined Agilent SureSelect XT HS2 library preparation and target enrichment system. The CRE V4 enrichment workflow is natively supported on the Agilent Bravo automated liquid handling platform for high-throughput sample preparation and the Agilent Magnis NGS prep system for complete, walkaway automation. The sequencing data can be readily analyzed using the Agilent Alissa Reporter software for efficient variant analysis, providing a sample-to-result workflow for clinical genetics laboratories.[*]

* Detection of complex variants is currently not supported by Alissa Reporter.

Table 1. Superior utility by design. SureSelect Clinical Research Exome V4 (CRE V4) contains clinical content carefully curated by world-renowned clinical geneticist Dr. Madhuri Hegde and her team of trained clinical geneticists. Unlike conventional exome panels, CRE V4 is designed based on clinical evidence to maximize the power of exome sequencing for clinical research. It provides enhanced coverage of about 6,500 disease-associated genes, and comprehensive coverage of non-coding targets from ClinVar, HGMD, and ACMG databases, ancestry SNPs, pharmacogenomic SNPs, and other expertly curated disease-relevant content.

	SureSelect Clinical Research Exome V4	Vendor T Exome 2.0
Target Size	41.1 Mb	36.5 Mb
Design Size	48.5 Mb*	43.2 Mb
Coding Content		
CCDS	Release 22	Yes
GENCODE	V31	Yes
RefSeq	Release 95 - NM	Yes
Non-Coding Content		
ClinVar	All P and LP variants (2022), including variants >100 bp	P and LP variants < ~100 bp
HGMD	DM with AF < 0.5% (2021)	–
ACMG 73 genes	All P and LP variants (2022), including variants > 100 bp	–
Additional Content		
Enhanced coverage	> 6,500 disease-associated genes and sex chromosomes	–
TERT gene	Promoter region	Promoter region
Mitochondrial genome	All 37 mitochondrial genes (16 kb)	–
Ancestry SNPs	1,500+ ancestry informative SNPs	–
PGx SNPs	100+ well-documented variants spanning pharmacogenomic genes like *CYP2C19*, *CYP2D6*, and *TPMT*	–
Structural variants	20+ events, including Boland Inversion, HBA/HBB deletions, GALC 30 kb, EPCAM/MSH2 and GREM1/SCG5 deletion	–
Mini-genomes	41 complete genes, including *DMD* gene	–
Repeat disorders	Probes included for 43 loci (rule out approach)	–
Sample tracking SNPs	AI 24 optimized SNPs from Pengelly *et al*.	Pengelly *et al*. SNPs

* 48.5 Mb merged probes size, 51.0 Mb including sequences with ≥95% homology

A

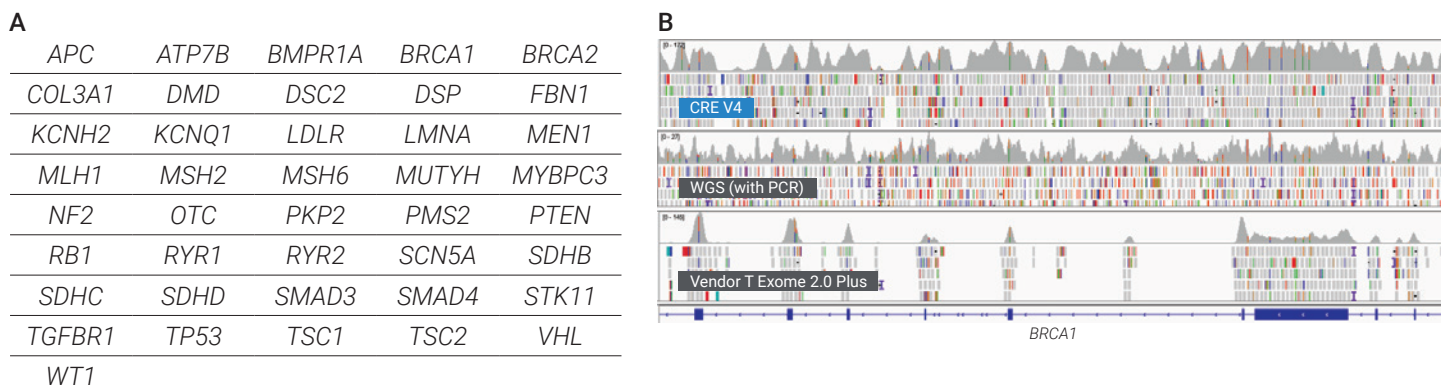|				
APC	*ATP7B*	*BMPR1A*	*BRCA1*	*BRCA2*
COL3A1	*DMD*	*DSC2*	*DSP*	*FBN1*
KCNH2	*KCNQ1*	*LDLR*	*LMNA*	*MEN1*
MLH1	*MSH2*	*MSH6*	*MUTYH*	*MYBPC3*
NF2	*OTC*	*PKP2*	*PMS2*	*PTEN*
RB1	*RYR1*	*RYR2*	*SCN5A*	*SDHB*
SDHC	*SDHD*	*SMAD3*	*SMAD4*	*STK11*
TGFBR1	*TP53*	*TSC1*	*TSC2*	*VHL*
WT1				

B

BRCA1

Figure 1. Bridging the gap between exome sequencing and genome sequencing. A) SureSelect Clinical Research Exome V4 (CRE V4) expands coverage beyond exonic regions to include 41 mini-genomes to enable discovery of novel and complex genomic variants for rare genetic disorders, such as Duchenne muscular dystrophy and Marfan syndrome. CRE V4 provides near complete gene coverage of 40 frequently mutated genes, as well as full gene coverage of the *DMD* gene. B) The coverage of reads across the mini-genomes is comprehensive, similar to the coverage obtained from whole-genome sequencing.

Table 2. Comprehensive design coverage of clinically relevant database content. SureSelect Clinical Research Exome V4 provides complete design coverage of protein coding content from CCDS, GENCODE, and RefSeq, and near complete coverage of non-coding, disease-associated content from ACMG, ClinVar, and HGMD databases.

	SureSelect CRE V4	Vendor ID Exome V2	Vendor R Exome	Vendor T Exome 2.0	Vendor I Exome 2.0 Plus
Target Size	41.1 Mb	34.1 Mb	35.8 Mb	36.5 Mb	37.5 Mb
Design Size	48.5 Mb	42.3 Mb	43.0 Mb	42.1 Mb	
Coding Content		% Database Covered (theoretical)			
CCDS Release 22	100%	99.8 %	> 99.9%	97.6%	> 99.9%
GENCODE V31	100%	97.2%	99.6%	97.2%	99.6%
RefSeq Release 95	100%	99.7%	> 99.9%	97.3%	99.7%
Non-Coding Content					
ACMG 73 genes variants	> 99.9%	43.2%	70.7%	78.0%	78.0%
ClinVar P/LP variants	100%	12.8%	64.5%	98.9%	98.9%
HGMD DM variants*	100%	10.3%	37.8%	53.9%	53.9%

HGMD non-coding DM variants with allele frequency < 0.5%

Sequence Efficiently with Robust Exome Enrichment Performance

A

Enrichment Specificity

B

Breadth of Coverage

C

Estimated Number of Samples Per Flow Cell

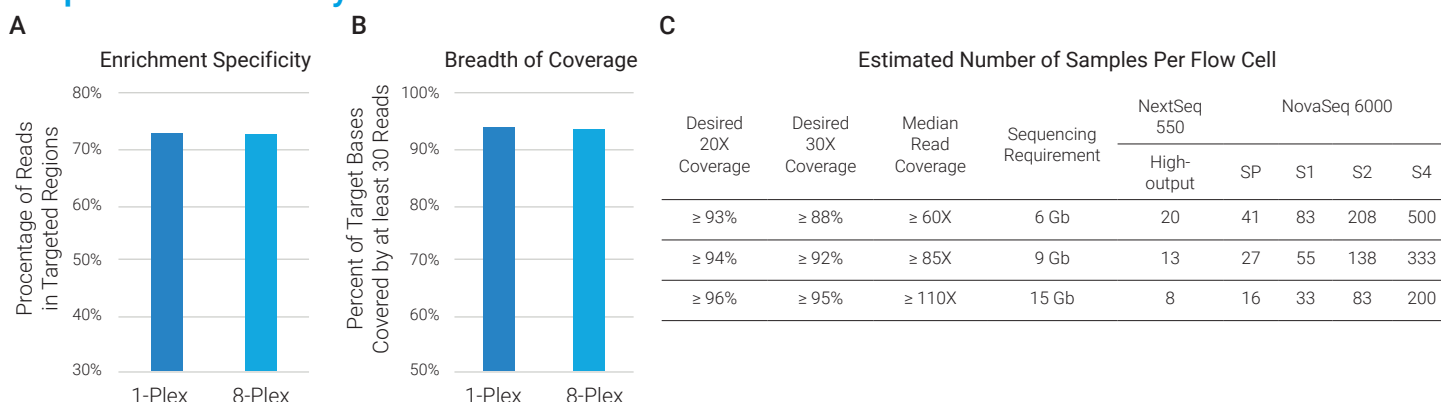| Desired 20X Coverage	Desired 30X Coverage	Median Read Coverage	Sequencing Requirement	NextSeq 550 High-output	NovaSeq 6000 SP	S1	S2	S4
≥ 93%	≥ 88%	≥ 60X	6 Gb	20	41	83	208	500
≥ 94%	≥ 92%	≥ 85X	9 Gb	13	27	55	138	333
≥ 96%	≥ 95%	≥ 110X	15 Gb	8	16	33	83	200

Figure 2. Cost-effective sequencing from excellent exome enrichment performance. A) SureSelect Clinical Research Exome V4 (CRE V4) provides high on-target rate, indicative of a highly specific enrichment of the thoughtfully curated content. B) CRE V4 also delivers excellent sequencing coverage as measured by the percentage of targeted bases with at least 30 reads using 9 Gb of sequencing. C) CRE V4's outstanding enrichment performance and data quality, combined with its efficient design, ensures highly economical exome sequencing, allowing optimal number of sample to be sequenced together on a sequencing run. Exome-enriched libraries were generated from 1-plex and 8-plex captures using the SureSelect XT HS2 DNA Target Enrichment system and 200 ng of genomic DNA from HapMap samples as input. The libraries were sequenced on an Illumina HiSeq 4000 or NovaSeq 6000 at 2 x 150 bp.

Achieve Superior Coverage of Disease-Associated Targets

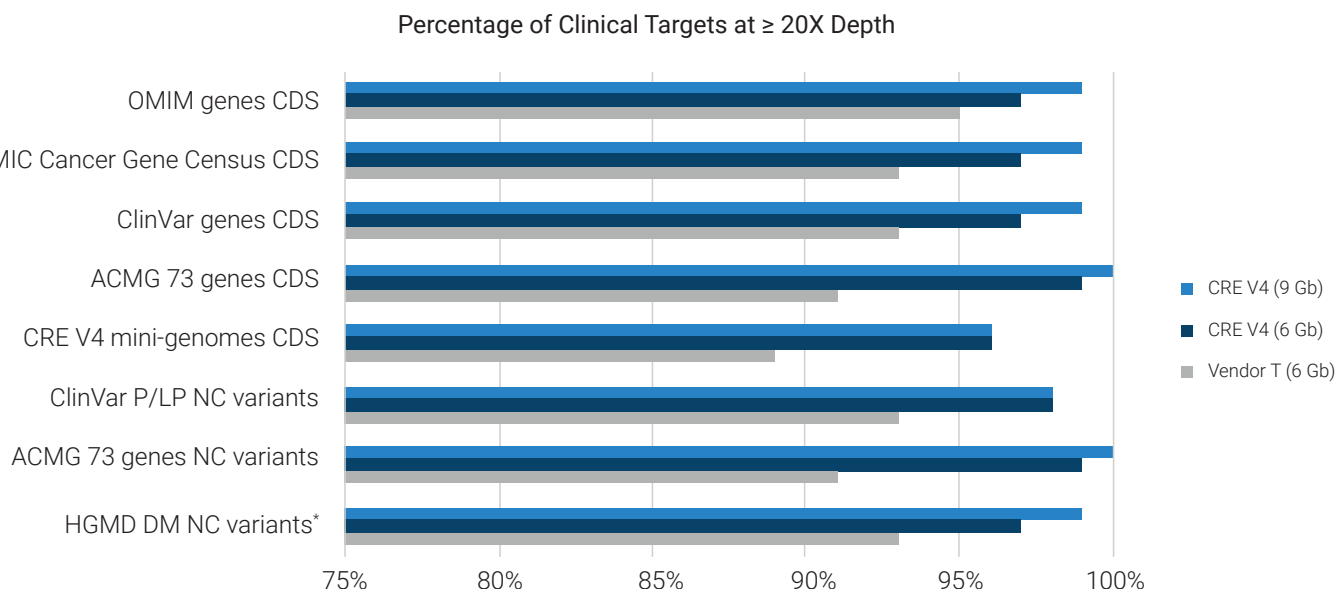## Percentage of Clinical Targets at ≥ 20X Depth

Figure 3. Deeper coverage of clinically relevant content. SureSelect Clinical Research Exome V4 (CRE V4) delivers higher observed average percent coverage at ≥ 20X of disease-associated targets from OMIM, COSMIC, ClinVar, ACMG, and HGMD, exome-enriched libraries were generated from 1-plex captures using CRE V4 with the SureSelect XT HS2 DNA Target Enrichment system, or using Exome 2.0 Plus and associated library prep and overnight hybridization reagents from Vendor T. The libraries were sequenced on an Illumina HiSeq 4000 or NovaSeq 6000 using 2 x 150 bp reads. Samples were downsampled to 9 Gb and 6 Gb for analysis. 6 Gb of sequencing is equivalent to 145X raw coverage by target size for CRE V4 and 160X for Vendor T. CDS: coding sequence; NC: non-coding.

HGMD non-coding DM variants with allele frequency <0.5%

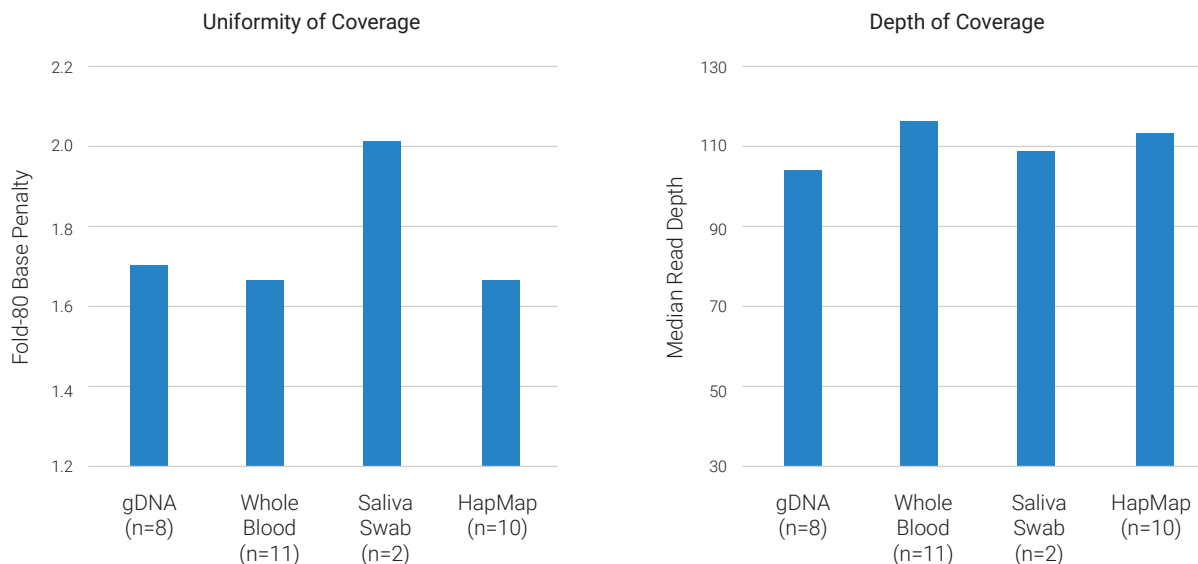# Generate Reliable Results from Challenging Clinical Samples

Figure 4. High-quality exome enrichment data from clinical samples. SureSelect Clinical Research Exome V4 (CRE V4) provides excellent and consistent enrichment performance from clinical sample types. CRE V4-enriched libraries were generated from HapMap sample and collection of clinical samples, including whole blood, saliva swabs, and genomic DNA (gDNA) from various samples types, that were previously characterized using WGS and WES. The libraries were sequenced on an Illumina NovaSeq 6000 using 2 x 150 bp reads. All samples were downsampled to 15 Gb for analysis. The saliva swab and gDNA samples exhibited higher variability compared to other samples, resulting in sightly higher fold-80 base penalty and lower median read depth.

Table 3. Robust analytical performance for SNV and indel detection. SureSelect Clinical Research Exome V4 (CRE V4) delivers excellent sensitivity, specificity, and positive predictive value (PPV) for single nucleotide variant (SNV) and insertion deletion (indel) detection. CRE V4-enriched libraries were generated from HapMap samples and sequenced on an Illumina NovaSeq 6000 using 2 x 150 bp reads. All samples were downsampled to 15 Gb for analysis.

	NA12878	NA24143	NA24149	NA24385	NA24631	NA24695
SNV						
Sensitivity	97.6%	97.5%	97.7%	97.6%	97.2%	97.3%
Specificity	100%	100%	100%	100%	100%	100%
PPV	98.8%	99.0%	99.0%	98.8%	99.0%	98.9%
Indel						
Sensitivity	90.4%	88.7%	89.6%	87.7%	93.1%	90.0%
Specificity	100%	100%	100%	100%	100%	100%
PPV	95.6%	96.2%	95.7%	96.2%	99.1%	96.9%

Table 4. Concordant detection of variants in clinical samples. SureSelect Clinical Research Exome V4 (CRE V4) can detect SNVs and CNVs previously identified using WGS and WES. CRE V4-enriched libraries were generated from clinical samples (whole blood and extracted gDNA from various samples types) and sequenced on an Illumina NovaSeq 6000 using 2 x 150 bp reads. All samples were downsampled to 15 Gb, and variant analysis was carried out using Alissa Reporter. Previously confirmed SNVs (not shown) and CNVs in the clinical sample were detected.

Sample Type	CNV Region	CNV Genes	CNV Size	CNV Type
gDNA	16q22.1	188 genes	20.28 Mb	Gain
	16p13.3	51 genes	1.05 Mb	Loss
gDNA	14q13.3	GALC	22.1 kb	Loss
gDNA	2p21	MSH2, KCNK12	124.2 kb	Loss
Whole Blood	Xp21.1	DMD	102.9 kb	Loss
gDNA	16q24.3	DPEP1, CHMP1A, SPATA33, CDK10, SPATA2L, VPS9D1, VPS9D1-AS1, ZNF276, FANCA	157.3 kb	Loss
Whole Blood	8q21.11-q21.12	GDAP1	218.2 kb	Gain

Increase Productivity with a Sample-to-Report Workflow Powered by Automation and Clinical Informatics Software

Sample and Library Preparation	Sequencing	Variant Analysis and Reporting

Library Preparation and Exome Enrichment

SureSelect XT HS2 library prep and target enrichment kits SureSelect Clinical Research Exome V4

Magnis SureSelect XT HS2 Clinical Research Exome V4

Automation

Magnis system for walkaway automation and Bravo platform for high-throughput automation

Sample QC

TapeStation and Femto Pulse systems for sample QC

Available Across Multiple Sequencing Platforms

SureSelect is optimized for Illumina sequencers, and is compatible with short-read sequencers from MGI/Complete Genomics, Element Biosciences, and Pacific Biosciences

Alissa Reporter

Seamless variant analysis soltuion with simultaneous SNV/indel and exon-level CNV detection

Agilent Tertiary Analysis Partners

Streamlined variant interpretation and reporing powered by phenotype-driven prioritization and curated databases
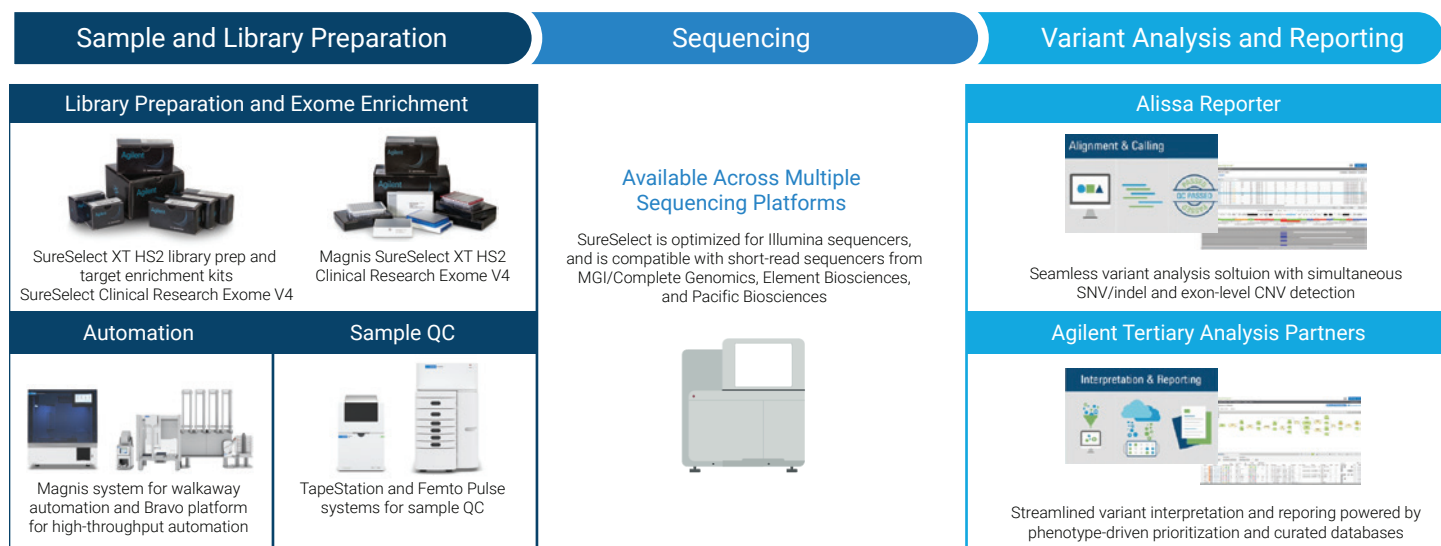
Figure 5. Sample-to-report workflow from a single vendor. The SureSelect Clinical Research Exome V4 (CRE V4) can be paired with Bravo and Magnis automation and the Alissa Clinical Informatics platform to achieve operational efficiency. CRE V4 is readily automated on the Bravo Automated Liquid Handling platform for high-throughput applications, requiring only 60 minutes of hands-on time to process up to 96 samples. The Magnis SureSelectXT HS2 CRE V4 kit contains consumables and pre-aliquoted reagents for use on the Magnis NGS Prep system. The instrument requires only 10 minutes of hands-on time to set up. From there, the Magnis system delivers eight exome-enriched libraries in less than nine hours without further operator intervention. Sequencing can be performed on an Illumina sequencer. Alternatively, sequencing can be carried out on a MGI DNBSEQ sequencer, Element Biosciences AVITI benchtop sequencer, or Pacific Biosciences Onso sequencing system with the use of a platform-specific library conversion kit. Alissa Reporter provides rapid variant analysis and user-friendly interface, enabling molecular laboratories to efficiently analyze the samples.

Ordering Information

Product Description*	16 rxns	96 rxns	96 rxns Auto
SureSelect XT HS Clinical Research Exome V4	5280-0020	5280-0021	5280-0022
Product Description*	2 hybs	12 hybs	12 hybs Auto
SureSelect XT PreCap Clinical Research Exome V4	5280-0029	5280-0030	5280-0031
Product Description	32 rxns	96 rxns	
Magnis SureSelect XT HS2 DNA CRE V4 ILM	G9775A	G9775B	
Product Description			
Magnis NGS Prep system	Contact sales		
Bravo NGS Workstation (Option A)	Contact sales		
Bravo NGS Workstation (Option B)	Contact sales		
Alissa Reporter	Contact sales		

The SureSelect XT HS probe sets are compatible with fast hybridization, and the SureSelect XT probe sets are compatible with overnight hybridization.

www.agilent.com/genomics

Agilent
Trusted Answers