TECHNICAL NOTE

Analyze germline CNVs with TruSight™ Hereditary Cancer Panel

- Temperature reduction during hybridization optimizes the assay for CNV calling
- Reference normal panel data for CNV calling saves time and resources

illumina®

Introduction

While cancer itself is not an inherited disease, genes with mutations that can lead to a predisposition to a specific cancer type are inherited. In fact, 5-10% of all diagnosed cancers are linked to inherited genetic factors.[1-4] Of the observed inherited mutations, copy number variations (CNVs) are prevalent at 8.3% among pathogenic variants.[5] Germline CNVs have long been suspected of playing a role in inherited cancer types, including breast, prostate, ovary, pancreas, colon, rectum, endometrium, lung, and melanoma.[6] With this high prevalence, it is crucial that any protocol designed to screen for hereditary cancer risk is optimized for CNV detection.

To simplify germline CNV detection in hereditary cancers, Illumina is providing baseline files of a "panel of normal" samples to enable comparison CNV calling when using the TruSight Hereditary Cancer Panel on Illumina platforms. The panel of normal data are available to all TruSight Hereditary Cancer Panel users and minimize the need for customers to procure, sequence, and analyze their own "normal" samples for use when analyzing tumor DNA.

This technical note discusses the samples used to create the panel of normal data, optimized library preparation using the TruSight Hereditary Cancer Panel, sequencing on four different Illumina platforms, and data analysis using the DRAGEN™ Enrichment app (Figure 1). Information on how users can access the resulting data is also provided.

Materials and methods

Samples

A total of 72 samples were characterized for this study (Tables 1-3). Case samples have a known CNV within a gene of interest. Panel of normal samples do not contain a CNV with known pathogenicity in any of the genes of interest.

Table 1: 72 samples selected to characterize TruSight Hereditary Cancer Panel performance

Sample type	No. of samples	Source
Case samples with known truth	30	Coriell Institute[a]
Panel of normal samples	42	Coriell Institute[b]

a. Details in Table 2.
b. Details in Table 3.

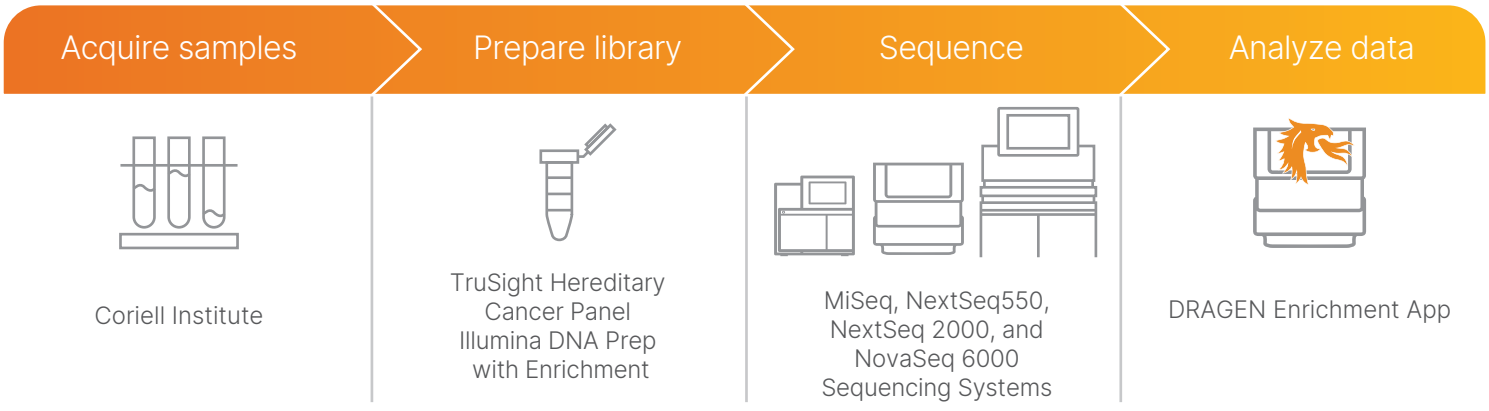


Figure 1: Overview of the workflow used to analyze samples with the TruSight Hereditary Cancer Panel.

Table 2: Case samples with known truth from the Coriell Institute

Sample	Event	Gene	Sample	Event	Gene
NA10401	DUP	chr2	NA11570	DEL	*APC*
NA12606	DUP	13q11.q21.2	NA20539	DEL	*PALB2*
NA04127	DUP	3p26.3p21.31	HG03857	DEL	*PALB2*
NA07216	DUP	3pter.q21	NA08039	DUP	*PALB2*
NA00343	DEL	rs11940551	NA09711	DUP	*BRCA2*
NA01555	DUP	*BRCA2*	NA07045	DEL	*ATM*
NA02718	DEL	*BRCA2*	NA08618	DUP	*ATM*
NA07150	DUP	*BRCA2*	NA13451	DEL	*MSH2*
NA03330	DUP	*BRCA2*		DEL	*EPCAM*
NA14238	DEL	*APC*		DEL	*MSH6*
NA14234	DEL	*APC*	NA07503	DUP	*EPCAM*
NA11571	DEL	*APC*		DEL	*MSH6*
Case samples with pseudogene events					
HG00343	DEL	*CHEK2*	NA14626	DUP	*BRCA1*
NA07106	DUP	*CHEK2*	NA18949	DEL	*BRCA1*
NA05047	DUP	*BMPR1A*	NA18540	DUP/DEL	*PMS2*
NA20125	DUP	*BMPR1A*	HG00451	DUP	*PMS2*
	DUP	*PTEN*	HG01503	DUP	*PMS2*

DEL, deletion; DUP, duplication

Table 3: Samples from the Coriell Institute used to generate the panel of normal data

NA02250	NA11410	NA12144	NA14090	NA21677	NA24143	NA24694
NA07078	NA11602	NA12154	NA14091	NA21730	NA24149	NA24695
NA07414	NA11630	NA12155	NA14170	NA21781	NA24385	NA25504
NA09373	NA12006	NA12872	NA16533	NA21833	NA24465	NA25581
NA10080	NA12043	NA12874	NA21070	NA23249	NA24590	NA25582
NA11409	NA12044	NA12878	NA21660	NA23251	NA24631	NA25591

Library preparation

Libraries were prepared using 50 ng DNA, the TruSight Hereditary Cancer Panel (Illumina, Catalog no. 20029551), and the Illumina DNA Prep with Enrichment workflow (Illumina, Catalog no. 20025524). The TruSight Hereditary Cancer Panel is a targeted sequencing panel designed to assess germline mutations. The panel was designed in collaboration with experts in cancer genomics and contains 10,341 probes that target 113 genes and 125 single nucleotide polymorphisms (SNPs) for identification purposes and polygenic risk scoring.[7] Illumina DNA Prep with Enrichment takes advantage of on-bead tagmentation for a fast, streamlined workflow.[8]

Protocols were followed as stated in the Illumina DNA Prep with Enrichment reference guide[9] with one exception: hybridization was performed at 58°C and 62°C (temperature recommended in the protocol).

Sequencing

Sequencing was performed on four different sequencing systems, the MiSeq™, NextSeq™ 550, NextSeq 2000, and NovaSeq™ 6000 sequencing systems using a read length of 2 × 150 bp at 100×–500× coverage (Table 4). Two replicates were run for each sample.

300× coverage is recommended; however, internal testing with high-quality samples demonstrates good senstivity with 100× coverage. Customers should validate coverage according to sample quality.

Analysis

Sequencing data were analyzed using the DRAGEN Enrichment App v3.10[10] accessible on BaseSpace™ Sequence Hub following the DRAGEN CNV pipeline.[11] DRAGEN Baseline Builder was used with default settings. The app performs small variant calling, CNV calling, and structural variant calling, among other applications.

Results

Lowering hybridization temperature

The recommended hybridization and wash temperature for the TruSight Hereditary Cancer Panel is 62°C. Lowering this temperature to 58°C resulted in higher specificity for the panel of normal samples and higher sensitivity for the case samples (Table 4). To determine specificity, the panel of normal samples was analyzed using the DRAGEN CNV pipeline and measured by the formula, (number of exons called)/(number of exons in the panel) (Figure 2). To examine sensitivity, the case sample was analyzed using the DRAGEN CNV pipeline to determine the number of exons missed from the known truth sample. If the known truth was positioned on a pseudogene, samples were categorized separately (Figure 3, Table 5).

Table 4: High-accuracy CNV calling at 58°C

Platform	Case sample coverage	Panel of normal sample coverage	Sensitivity (Case samples)	Specificity (Panel of normal samples)
MiSeq system	100×	400×	99.54	99.99
	200×	400×	99.69	99.97
	300×	400×	99.84	99.96
	400×	400×	99.69	99.98
NextSeq 2000 system	500×	500×	99.46	99.98
NextSeq 550 system	500×	500×	99.46	99.97
NovaSeq 6000 system	500×	500×	99.62	99.97

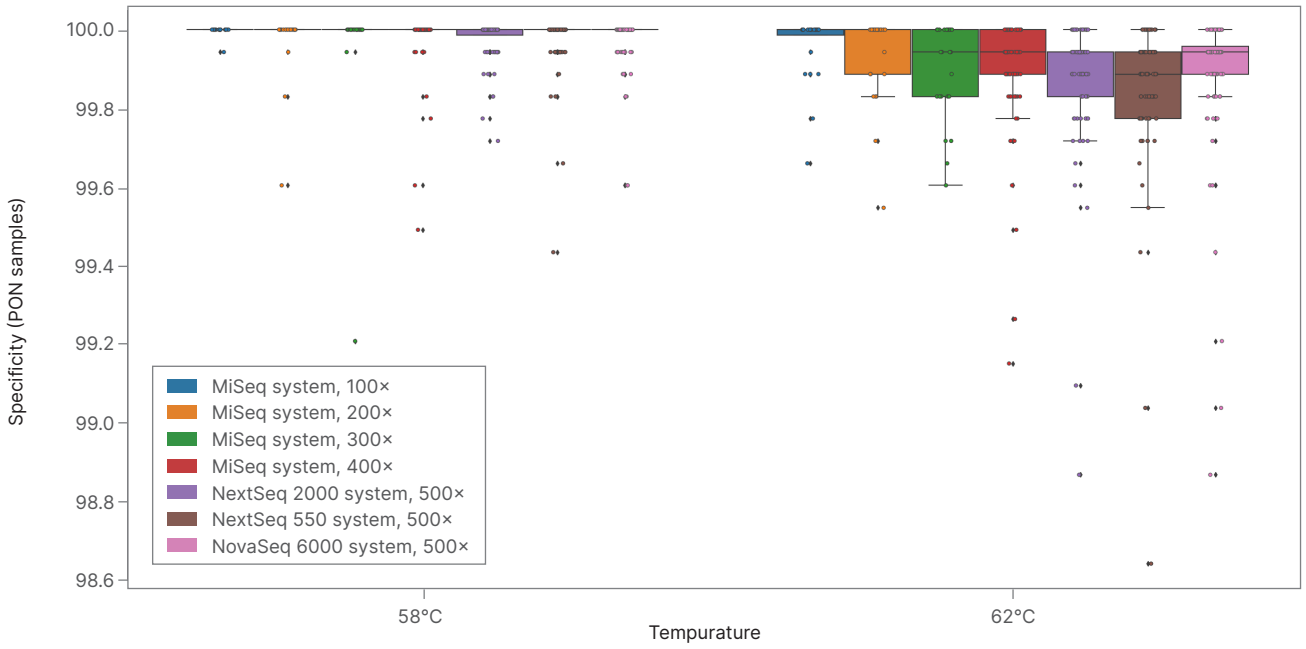


Figure 2: Higher specificity of panel of normal observed at 58°C—Panel of normal samples were hybridized at 58°C and 62°C and sequenced on four different Illumina sequencing systems in duplicate.

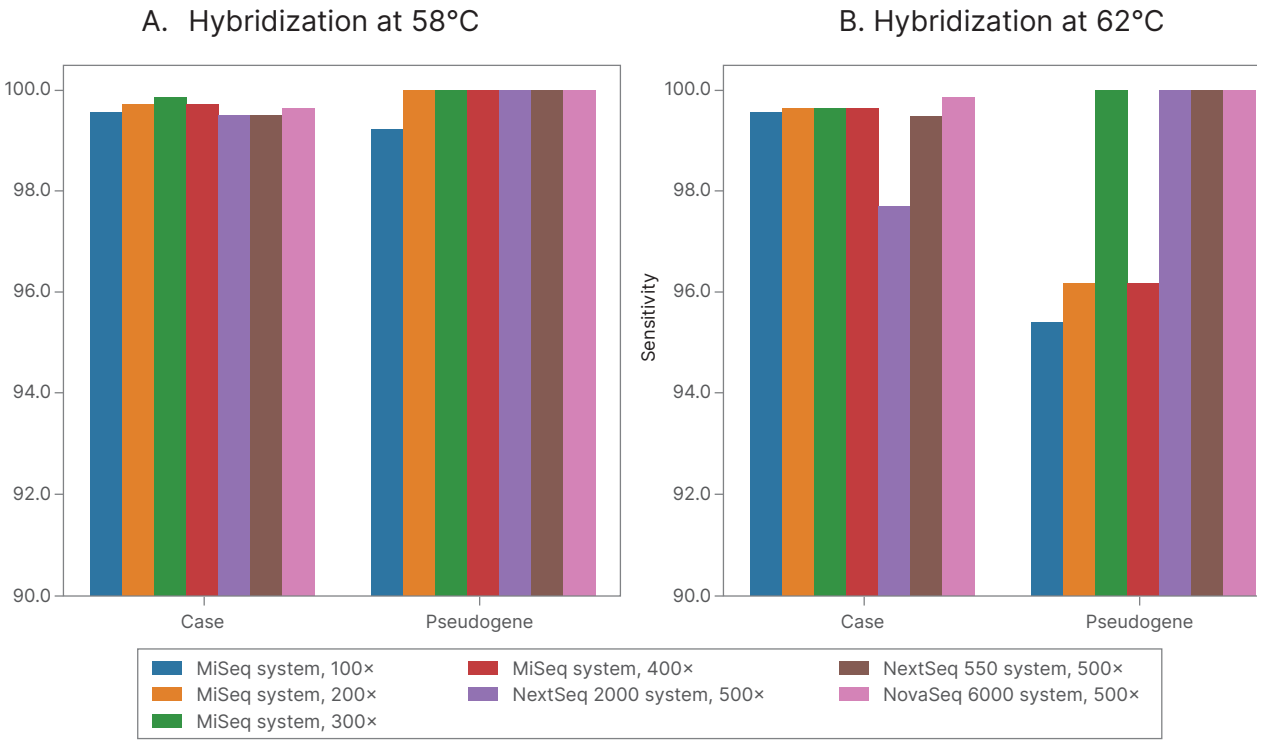


Figure 3: Higher sensitivity of case samples observed at 58°C—Exons counts in case samples were compared to those of known truth samples. Samples included in psedogene preformance testing are described in Table 5.

Table 5: Samples included in pseudogene performance testing

Sample	Gene	Transcript	Affected exons
HG00343	*CHEK2*	NM_007194.3	11-15
NA07106	*CHEK2*	NM_007194.3	11-15
NA05047	*BMPR1A*	NM_004329.2	12-13
NA20125	*BMPR1A*	NM_004329.2	12-13
	PTEN	NM_000314.4	9
NA14626	*BRCA1*	NM_007294.3	2
NA18949	*BRCA1*	NM_007294.3	2
NA18540	*PMS2*	NM_000535.6	6-15
HG00451	*PMS2*	NM_000535.6	1-12
HG01503	*PMS2*	NM_000535.6	1-10

Accessing the panel of normal data

The panel of normal data, generated using validated samples, provides researchers with a known normal data set for comparison with tumor samples without the need to spend time and resources acquiring, sequencing, and analyzing this sample set. Data can be downloaded from BaseSpace Sequence Hub (Figure 4) and used as a reference in CNV studies using the DRAGEN Enrichment app for analysis (Figure 5).

Reference CNV analysis data were generated using optimal samples and Illumina DNA Prep with Enrichment, with hybridization occurring at 58°C. For accurate data comparison, tumor samples need to be of matching quality and sequenced under the same conditions.

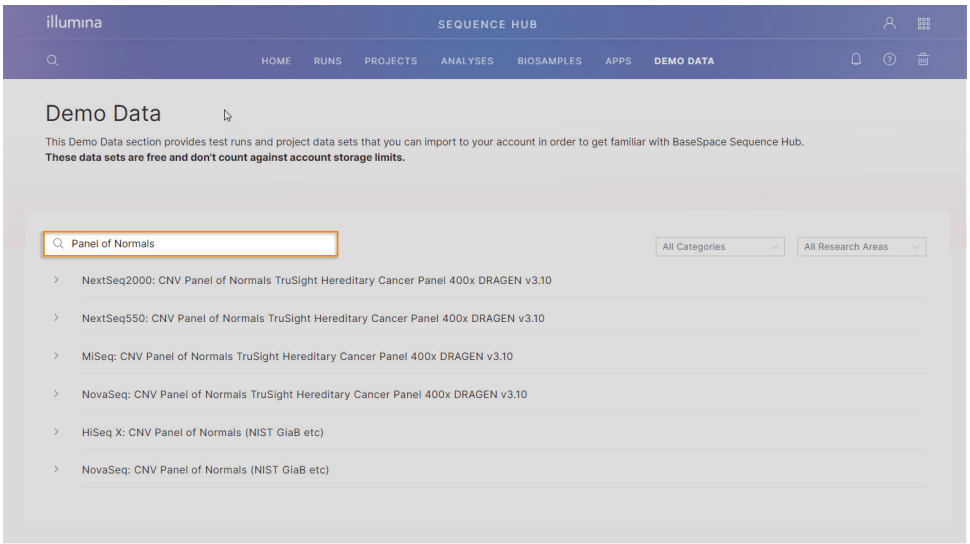
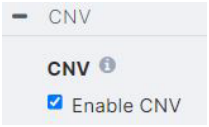
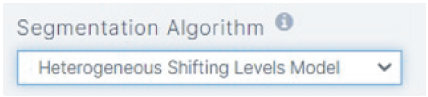


Figure 4: Accessing the panel of normal data—Data are available in BaseSpace Sequence Hub at basespace.illumina.com/datacentral. In the Demo data screen, type "Panel of Normals" into the search bar (orange box). Choose the data set generated on the Illumina sequencing system in use. Import the project into the DRAGEN Enrichment app.

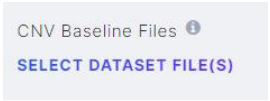
1. Enable CNVs.



2. Select the segmentation algorithm.



3. Select the baseline panel of normal files imported from BaseSpace Sequence Hub. There are 42 files in each platform folder.



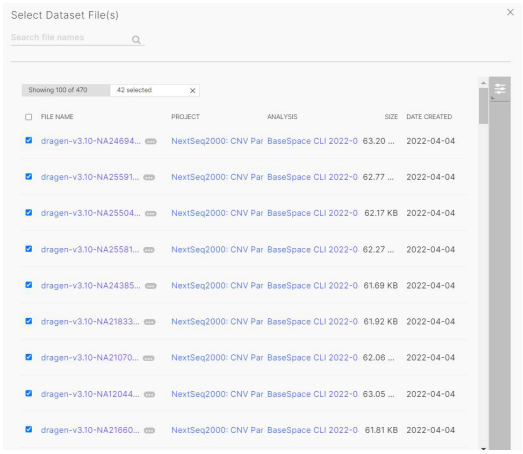


Figure 5: Configuring the DRAGEN Enrichment app for use with the panel of normal data—In the app, select input sample for analysis. Files are decompressed for analysis. Samples then go through the mapping and aligning process to the reference, followed by position sorting, and then through variant calling. Users can then configure the app for CNV analysis. An output VCF, gVCF, and/or BAM file is created after variant calling. Multiple VCF files will be generated if CNV/SV are enabled.

Summary

The TruSight Hereditary Cancer Panel enables researchers to access an expert-defined content set for analyzing variation within genes previously linked with a predisposition towards cancer. Data demonstrate that reducing the hybridization temperature during library preparation with Illumina DNA Prep with Enrichment optimizes capture of CNV regions. To assist with CNV analysis, panel of normal data, available of BaseSpace Sequence Hub, can be used as a comparison for data generated using high-quality tumor samples.

Learn more

TruSight Hereditary Cancer Panel, illumina.com/TruSightHereditaryCancer

Illumina DNA Prep with Enrichment, illumina.com/products/by-type/sequencing-kits/library-prep-kits/nextera-flex-enrichment.html

Illumina sequencing systems, illumina.com/systems

DRAGEN Enrichment app, illumina.com/products/by-type/informatics-products/basespace-sequence-hub/apps/dragen-enrichment.html

References

1. American Cancer Society. Family Cancer Syndromes. https://www.cancer.org/healthy/cancer-causes/genetics/family-cancer-syndromes.html. Updated September 14, 2022. Accessed October 19, 2022.
2. Tsaousis GN, Papadopoulou E, Apessos A, et al. Analysis of hereditary cancer syndromes by using a panel of genes: novel and multiple pathogenic mutations. *BMC Cancer*. 2019;19(1):535. Published 2019 Jun 3. doi:10.1186/s12885-019-5756-4
3. Nagy R, Sweet K, Eng C. Highly penetrant hereditary cancer syndromes. *Oncogene*. 2004;23(38):6445-6470. doi:10.1038/sj.onc.1207714

4. Garber JE, Offit K. Hereditary cancer predisposition syndromes. *J Clin Oncol*. 2005;23(2):276-292. doi:10.1200/JCO.2005.10.042
5. Truty R, Paul J, Kennemer M, et al. Prevalence and properties of intragenic copy-number variation in Mendelian disease genes. *Genet Med*. 2019;21(1):114-123. doi:10.1038/s41436-018-0033-5
6. Kumaran M, Cass CE, Graham K, et al. Germline copy number variations are associated with breast cancer risk and prognosis. *Sci Rep*. 2017;7(1):14621. Published 2017 Nov 7. doi:10.1038/s41598-017-14799-7
7. Illumina. TruSight Hereditary Cancer Panel Data Sheet. https://science-docs.illumina.com/documents/LibraryPrep/trusight-hereditary-cancer-data-sheet-1170-2019-012/Content/Source/Library-Prep/TruSight/Hereditary-Cancer/trusight-hereditary-cancer-data-sheet-1170-2019-012/trusight-hereditary-cancer-data-sheet-1170-2019-012.html. Accessed October 20, 2022.
8. Illumina. Illumina DNA Prep with Enrichment Data Sheet. https://www.illumina.com/content/dam/illumina/gcs/assembled-assets/marketing-literature/illumina-dna-prep-for-enrichment-770-2020-010/illumina-dna-prep-enrichment-data-sheet-770-2020-010.pdf. Accessed October 20, 2022.
9. Illumina. Illumina DNA Prep with Enrichment Reference Guide. https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/illumina_prep/illumina-dna-prep-with-enrichment-reference-guide-1000000048041-07.pdf. Accessed October 20, 2022.
10. Illumina. DRAGEN Enrichment. https://www.illumina.com/products/by-type/informatics-products/basespace-sequence-hub/apps/dragen-enrichment.html. Accessed October 20, 2022.
11. Illumina. Copy Number Variant Calling. https://support.illumina.com/content/dam/illumina-support/help/Illumina_DRAGEN_Bio_IT_Platform_v3_7_1000000141465/Content/SW/Informatics/Dragen/CopyNumVariantCalling_fDG_dtSW.htm. Accessed October 24, 2022.

illumina®

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

M-GL-01341 v1.0