illumina®

APPLICATION NOTE

An integrated exome solution for variant analysis in challenging genes: *HBA1/2* and *SMN1/2* genotyping

Enabled by Illumina DNA Prep with Exome 2.5 Enrichment, custom spike-in content, and DRAGEN™ analysis

Reduce reliance on orthogonal assays

Enhance workflow efficiency and cost-effectiveness

Access more data by sequencing the whole exome

Introduction

In 2021, the American College of Medical Genetics and Genomics (ACMG) provided recommendations to steward equitable carrier screening through accessible next-generation sequencing (NGS) technology.[1] ACMG developed a list of 113 "Tier 3" genes that should be offered for all individuals expecting or planning a pregnancy to inform reproductive risk. Economic evaluations have demonstrated cost-effectiveness for carrier testing panels of up to 300 genes.[2,3] Genes meeting relevant criteria vary based on regional carrier frequency, and shifting demographics require workflow flexibility that allows updated gene lists over time.

Whole-exome sequencing (WES) can provide a flexible solution for NGS applications that require comprehensive gene evaluation at low cost, like carrier screening research. However, for some challenging genes, orthogonal assay techniques may be required to detect or verify the genotype. Alpha thalassemia and spinal muscular atrophy (SMA) are autosomal recessive genetic conditions that result from deleterious changes in the *HBA1/2*[4] and *SMN1/2*[5] genes, respectively. Both *HBA* and *SMN* occur in regions of the genome that are difficult to map accurately due to high homology with paralogous genes. To achieve the required sensitivity, these genes are typically assessed using dedicated protocols and incur additional costs to a research program.

Illumina DNA Prep with Exome 2.5 Enrichment delivers economical human WES results with outstanding performance and data quality.[6] Combined with novel targeted callers in DRAGEN Secondary Analysis v4.4 and a supplemental Illumina CS/PGx Custom Enrichment Research Panel, Illumina DNA Prep with Exome 2.5 Enrichment now enables a differentiated NGS workflow for carrier gene research (Figure 1). Here, we show that variant analysis for challenging genes, like *HBA* and *SMN* copy number detection, can be achieved with this integrated WES solution to enhance efficiency and cost-effectiveness and further improve data accessibility.

Methods

Illumina CS/PGx Custom Enrichment Research Panel

The Illumina CS/PGx Custom Enrichment Research Panel was developed along with DRAGEN targeted callers for WES to enable detection of challenging gene variants relevant for carrier screening and pharmacogenomic

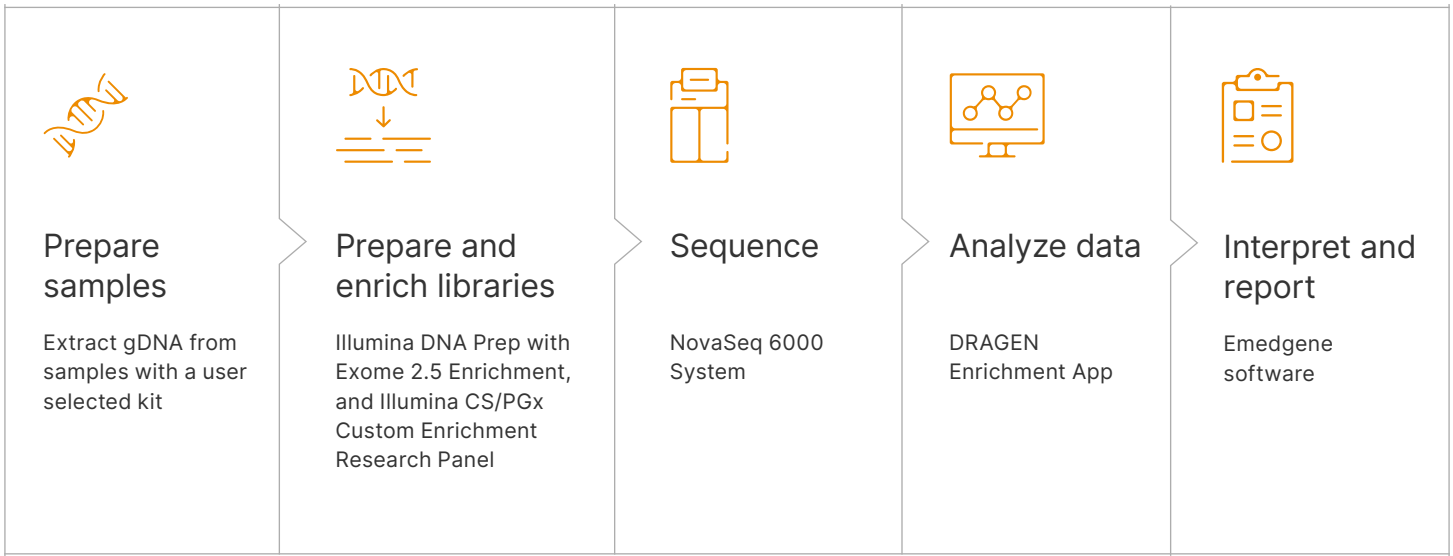

Figure 1: Custom enrichment workflow for carrier gene research
Illumina supports an end-to-end workflow for WES. Extracted gDNA (or DNA from blood or saliva following direct extraction protocols) is processed into sequencing libraries with Illumina DNA Prep with Exome 2.5 Enrichment and Illumina CS/PGx Custom Enrichment Research Panel. Qualified Methods are available on a range of automation platforms. Sequence according to scale and throughput needs on Illumina Systems. Accurate, rapid secondary analysis and variant calling is performed with the DRAGEN Enrichment app. For genetic diseases, Emedgene enables intuitive interpretation and reporting.

research (Table 1). The CS/PGx panel targets 359 kb of the genome, representing less than 1% of the Exome 2.5 panel size and thus having minimal impact on the overall coverage of exome panel targets.

Table 1: Targets with enhanced coverage with the Illumina CS/PGx Custom Enrichment Research Panel

CYP2B6/7	*CYP2D6/7*	*CYP21A2*
CYP2A6/7	*FXN*	*GBA/GBAP1*
HBA1/2	*NEB*	*RPGR*
SMN1/2	*STRC*	*TNXB*

The predesigned Illumina CS/PGx Custom Enrichment Research Panel can be ordered through the DesignStudio™ custom design portal as an Illumina Custom Enrichment Panel V2 (Illumina, Catalog nos. 20073952, 20073953, and 20111339) as shown (Figure 2). After entering a "Design Name," go to the next screen to select "Add from Existing Design" followed by Illumina Panel. Select the design "Illumina CS/PGx Custom Enrichment Research Panel," then check the boxes for all regions before "Add Designs" (Figure 3). For this workflow, it is not recommended to make changes to this design. Select "Cancel" to exit the view and proceed to review the design.

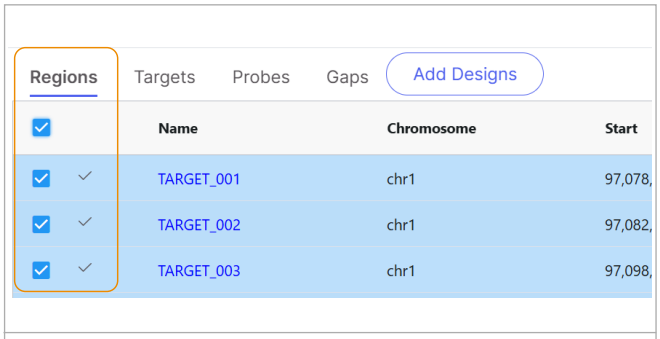

Figure 3: Select panel design for Illumina CS/PGx Custom Enrichment Research Panel
After selecting "Illumina CS/PGx Custom Enrichment Research Panel" (Figure 2) you will come to this screen and select the check the boxes for all "Regions" (orange rectangle).

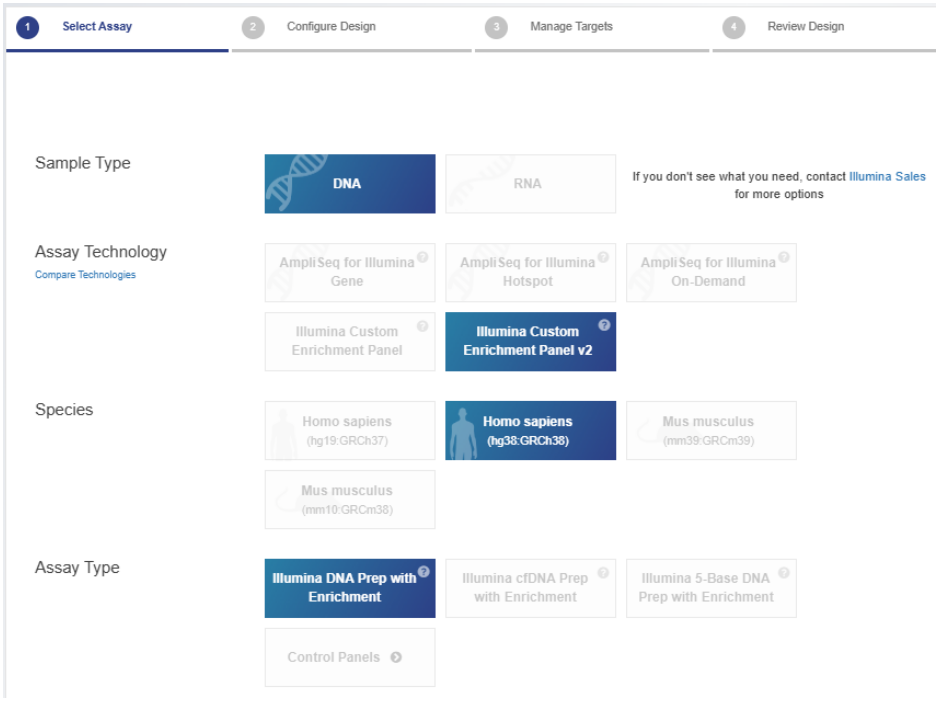

Figure 2: Ordering the Illumina CS/PGx Custom Enrichment Research Panel through the DesignStudio custom design portal
The blue rectangles show the correct selections in the DesignStudio custom design portal: Sample Type, DNA; Assay Technology, Illumina Custom Enrichment Panel v2; Species, Homo sapiens (hg38:GRCh38); and Assay Type, Illumina DNA Prep with Enrichment.

Library preparation

Cell lines (Coriell) were profiled using Illumina DNA Prep with Exome 2.5 Enrichment, (S) Tagmentation Set B (Illumina, Catalog no. 20077595), Illumina DNA Prep with Exome 2.5 Enrichment, (S) Tagmentation Set D (Illumina, Catalog no. 20077596), and Illumina CS/PGx Custom Enrichment Research Panel following the instructions in the Illumina DNA Prep with Exome 2.5 Enrichment reference guide (Document no. 1000000157112). Prior to overnight hybridization, libraries were pooled by volume for 12 libraries per hybridization reaction. For each hybridization reaction, 4 µl of Twist BioScience for Illumina Exome 2.5 Panel and 4 µl of Illumina CS/PGx Custom Enrichment Research Panel were added. All cell line library preparations were performed manually by three different operators.

An independent cohort of clinical research sample libraries, partially blinded (464/1099 blinded for *HBA*, and 665/1117 blinded for *SMN*) and partially prospective, were prepared using an automated system with Illumina DNA Prep with Exome 2.5 Enrichment and the Illumina CS/PGx Custom Enrichment Research Panel (1140 data sets from 1009 unique samples) and profiled by Juno Genetics in Spain.

Sequencing

Prepared cell line libraries were sequenced using the NovaSeq™ 6000 S4 Reagent Kit v1.5 (300 cycles) (Illumina, Catalog no. 20028312) and the NovaSeq 6000 System (Illumina, Catalog no. 20012850).

DNA extracted from whole blood was used for human research samples. Samples were sequenced using the NovaSeq 6000 S2 Reagent Kit v1.5 (300 cycles) (Illumina, Catalog no. 20028314) or NovaSeq 6000 S4 Reagent Kit v1.5 (300 cycles) (Illumina, Catalog no. 20028312) with 96 samples or 192 samples, respectively, and sequenced by Juno Genetics in Spain on the NovaSeq 6000 System. Per sample mean target coverage was 80–100×.

Analysis

Copy number detection for *HBA* and *SMN* were performed using DRAGEN secondary analysis v4.4 with targeted calling using an in-run panel of normals (PON). In-run, or within batch, depth normalization is another feature in DRAGEN secondary analysis for enrichment workflows that uses target-read counts of other samples in the same sequencing run to normalize depth of coverage and form a baseline from which to call copy number accurately. At least 30 samples per run are needed to use in-run normalization for *HBA* and *SMN* copy number calling. Samples with highly divergent depth profiles were removed before analysis. Calls were evaluated for concordance with either DRAGEN targeted caller results on matched WGS data or from quantitative PCR (qPCR) and multiplex ligation-dependent probe amplification (MLPA), if available.

Customers can implement the targeted calling workflow using DRAGEN 4.4 on-premises, or in the cloud using BaseSpace™ Sequence Hub and Illumina Connected Analytics (Table 2). A combined BED file is available in these environments, or can be downloaded from the Illumina DNA Prep with Exome 2.5 Enrichment product files support page. More details on this analysis can be found in the DRAGEN secondary analysis v4.4 product guide.

Table 2: Analysis options using DRAGEN secondary analysis v4.4

Analysis option	Inputs
DRAGEN Germline Enrichment from BCL Convert App on BaseSpace Sequence Hub	BCL files from a new planned run or existing sequencing run
DRAGEN Enrichment App on BaseSpace Sequence Hub	FASTQ, BAM, or CRAM format
DRAGEN Germline Enrichment App on Illumina Connected Analytics	FASTQ, BAM, or CRAM format

Results

Cell lines

Within the cell line cohort, 5.9% of samples were removed prior to analysis due to highly divergent depth profiles. Copy number genotype calls for *HBA1* and *HBA2* (*HBA*) (726 data sets from 203 unique samples) and *SMN1* and *SMN2* (*SMN*) (739 data sets from 216 unique samples) were evaluated for concordance with whole-genome sequencing data.

HBA calls derived from WES were 100% (710/710) concordant with orthogonal method calls across nine genotypes. Only 2.2% of genotypes were not called (Figure 4A). *SMN1* copy number calls were 98.1% (717/731) concordant with the orthogonal methods. In terms of SMA status, calls were 100% concordant (missed calls only occurred for genotypes with copy number ≥ 2, which does not have a deleterious effect) (Figure 4B). The *SMN2* copy number calls from WES were 99.3% (726/731) concordant with orthogonal methods (Figure 4C). The no-call rate for *SMN1/2* copy numbers was 1.1% in this cohort.

Human subject research samples

For the human subject research samples, all samples passed QC and were evaluated for accurate variant calling. Matched results from multiplex ligation-dependent probe amplification (MLPA) profiling for the *HBA* locus with qPCR and MLPA profiling results for *SMN1* copy number were used to evaluate the accuracy of the WES-based calls. This sample cohort included nine genotypes for *HBA*, two of which differed from those evaluated in the cell line cohort. The concordance between WES results with matched MLPA or qPCR results for *HBA* was 100% (1099/1099) (Figure 5A) with a 0% no-call rate. The concordance for SMN1 copy number and SMA status was 99.8% (1113/1115) with a 0.3% no-call rate (Figure 5B).

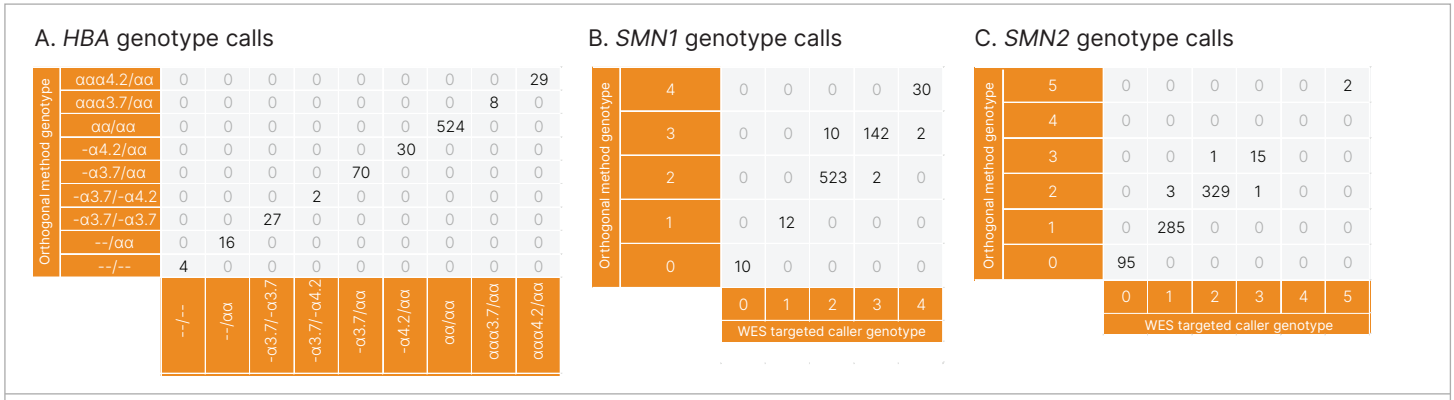

Figure 4: Cell line genotyping performance analysis
(A) Concordance analysis results between the WES-based *HBA* caller and WGS-based *HBA* caller on matched data sets. (B) Concordance analysis results between *SMN1* copy number calls obtained from the WES-based *SMN* caller and those obtained from the WGS-based *SMN* caller on matched data sets. (C) Concordance analysis results between *SMN2* copy number calls obtained from the WES-based *SMN* caller and those obtained from the WGS-based *SMN* caller on matched data sets.

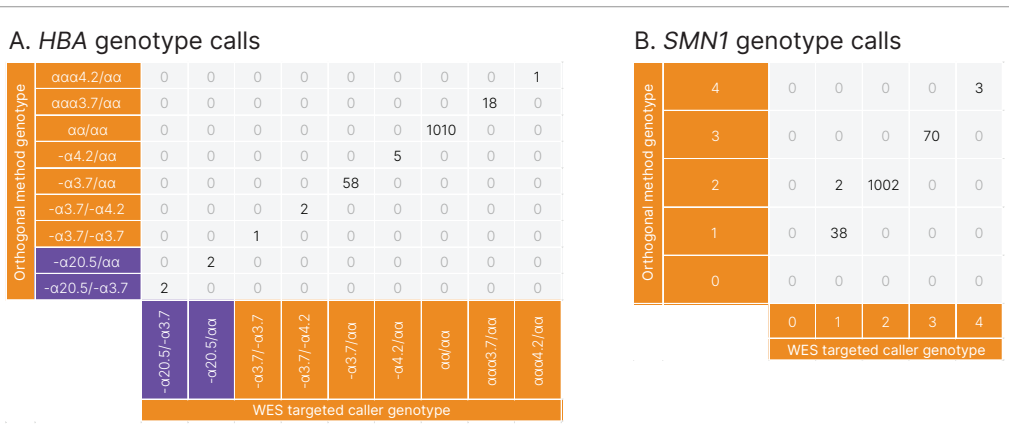

Figure 5: Cell line genotyping performance analysis
(A) Concordance analysis results between the WES-based *HBA* caller and WGS-based *HBA* caller on matched data sets. Unique genotypes analyzed in this cohort are highlighted in purple. (B) Concordance analysis results between *SMN1* copy number calls obtained from the WES-based *SMN* caller and those obtained from the WGS-based *SMN* caller on matched data sets. (C) Concordance analysis results between *SMN2* copy number calls obtained from the WES-based *SMN* caller and those obtained from the WGS-based *SMN* caller on matched data sets.

Conclusion

Comprehensive carrier gene research using NGS commonly relies on orthogonal methods to analyze challenging gene variants, driving up costs, and limiting accessibility. The Illumina DNA Prep with Exome 2.5 Enrichment and Illumina CS/PGx Custom Enrichment Research Panel, combined with novel WES-based *HBA* and *SMN* targeted callers implemented in DRAGEN secondary analysis, offer a reliable approach for profiling challenging *HBA* and *SMN* genes within an integrated NGS workflow. By proactively targeting relevant regions beyond *HBA1/2* and *SMN1/2*, this workflow can use future DRAGEN targeted callers developed for WES with minimal or no changes to wet lab workflows. Emedgene software is also available to help labs perform additional analysis such as variant interpretation and research report generation.

Learn more →

Illumina DNA Prep with Exome 2.5 Enrichment

Illumina sequencing platforms

DRAGEN secondary analysis

DRAGEN Enrichment application

Emedgene software

References

1. Gregg AR, Aarabi M, Klugman S, et al. Screening for autosomal recessive and X-linked conditions during pregnancy and preconception: a practice resource of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 2021;23(10):1793-1806. doi:10.1038/s41436-021-01203-z
2. Schofield D, Lee E, Parmar J, et al. Economic evaluation of population-based, expanded reproductive carrier screening for genetic diseases in Australia. *Genet Med*. 2023;25(5):100813. doi:10.1016/j.gim.2023.100813
3. Beauchamp KA, Johansen Taber KA, Muzzey D. Clinical impact and cost-effectiveness of a 176-condition expanded carrier screen. *Genet Med*. 2022;24(4):968. doi:10.1016/j.gim.2022.02.011
4. Farashi S, Harteveld CL. Molecular basis of α-thalassemia. *Blood Cells Mol Dis*. 2018;70:43-53. doi:10.1016/j.bcmd.2017.09.004
5. Kolb SJ, Kissel JT. Spinal Muscular Atrophy. *Neurol Clin*. 2015;33(4):831-846. doi:10.1016/j.ncl.2015.07.004
6. Illumina. Illumina DNA Prep with Exome 2.5 Enrichment data sheet. illumina.com/content/dam/illumina/gcs/assembled-assets/marketing-literature/illumina-dna-exome-enrich-m-gl-00002/illumina-dna-exome-enrich-data-sheet-m-gl-00002.pdf. Published 2024. Accessed March 2, 2026.

illumina®

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

M-GL-03631 v1.0