

Informing analysis and empowering interpretation of variants in rare disease

A use case with the TruSight™ Software Suite.

Introduction

Whole-genome sequencing (WGS) using next-generation sequencing (NGS) technology is a powerful tool for investigating variants linked to genetic disease. It provides a high-resolution, unbiased view across the entire genome to discover causative variants of inherited disorders. However, the vast amounts of data produced by WGS requires comprehensive data analysis tools that can efficiently translate the raw sequencing data into meaningful, interpretable results. To address these challenges, Illumina offers the TruSight Software Suite.

This hypothetical use case describes the workflow a laboratory analyst would follow using the TruSight Software Suite to identify variants associated with a genetic disease (Figure 1).

Case receipt and request for variant analysis

A laboratory analyst receives a new request for WGS variant analysis. The case is an example trio data set (Family 1) that consists of a male proband and unaffected parents (Table 1). After libraries are prepared from DNA samples from the proband and each parent, they are sequenced on the NovaSeq™ 6000 System. The analyst then begins working in the TruSight Software Suite to find disease-associated variants that correspond with the phenotypic features of the proband.

Table 1: Example data set case description

Case	Gender	Proband		Mother	Father
		Phenotype			
Family 1	Male	<ul style="list-style-type: none"> night blindness optic disc pallor 	Unaffected	Unaffected	

Case management

Using the Case Management Portal in the TruSight Software Suite, the lab analyst creates a new case, imports data files, and associates sequencing data files with the case. Alternatively, this information can be easily imported via an application programming interface (API). Next, the lab analyst enters pedigree information about the case (up to five individuals), including family structure, proband gender, proband phenotypic features (optional), and affected status of family members (Table 1). Real-time updates of case status are displayed in the TruSight Software Suite, allowing the analyst to monitor progress through the analysis workflow.

After sequencing is complete, the DRAGEN™ Bio-IT Platform automatically performs sequence alignment and variant calling for the proband and associated family members. Comprehensive variant calling includes small variants, copy number variants (CNVs), structural variants, mitochondrial variants, repeat expansions, and more. After reviewing information in the Quality Control (QC) Metrics tab for the proband and each family member, the lab analyst proceeds with variant triage, visualization, and interpretation (Figure 1).

Variant triage

The lab analyst begins variant triage by opening the Interpretation tab. Variant filtering can follow a custom plan or a prebuilt filter plan. Family-based filtering in the TruSight Software Suite enables comparison of the proband with other family members. Additional filter options include filtering on population frequencies from sources such as the Genome Aggregation Database (gnomAD), variant consequences, sample metrics, ClinVar pathogenicity, and more

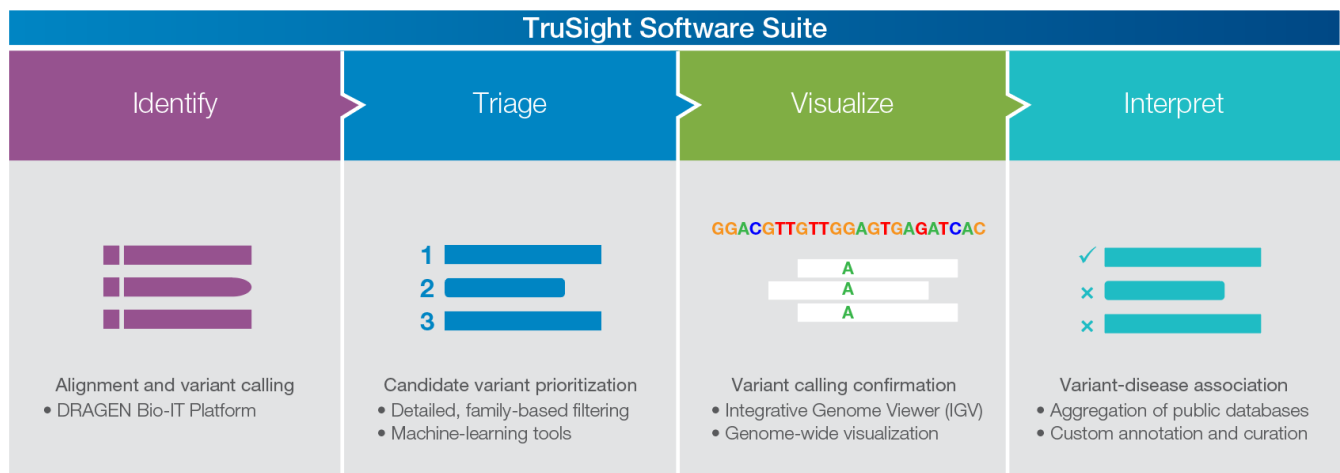


Figure 1: Variant analysis in the TruSight Software Suite—After the DRAGEN platform performs alignment and variant calling, users proceed with triaging, visualizing, and interpreting variants.

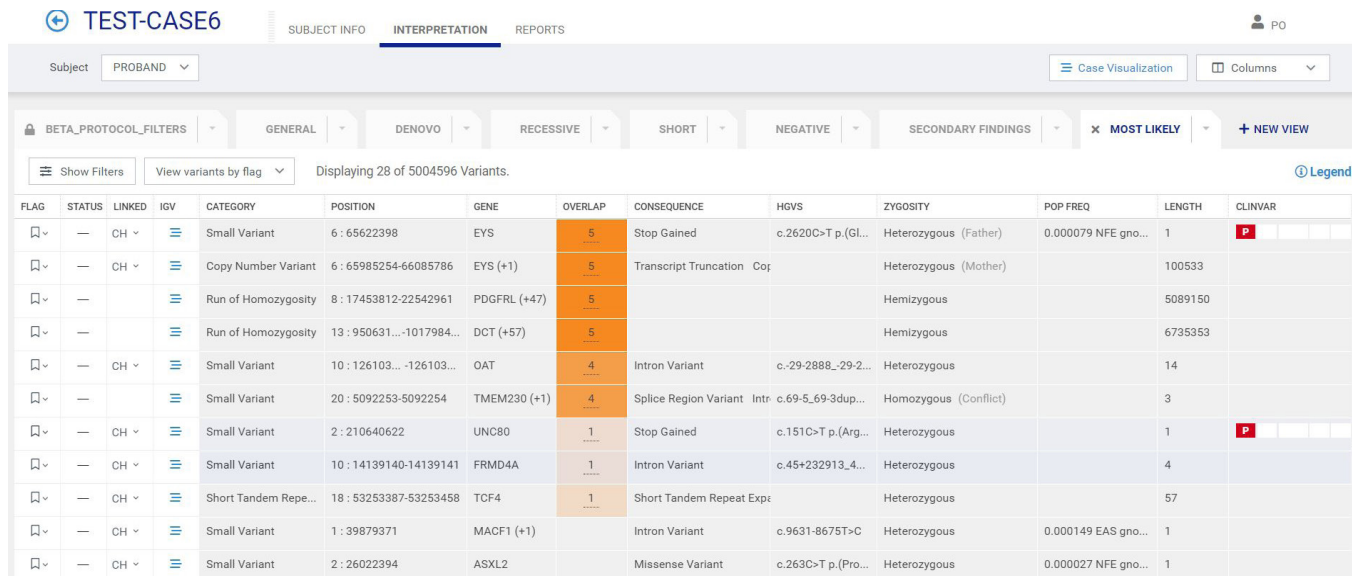


Figure 2: Variant filtering and prioritization—The lab analyst can perform family-based variant filtering for identification of inherited and *de novo* variants. As shown, the variant grid shows information on category of variant, chromosomal position, gene affected, overlap (number of overlapped phenotypes for the variant), consequence of the variant, population frequency of the variant (if known), and more.

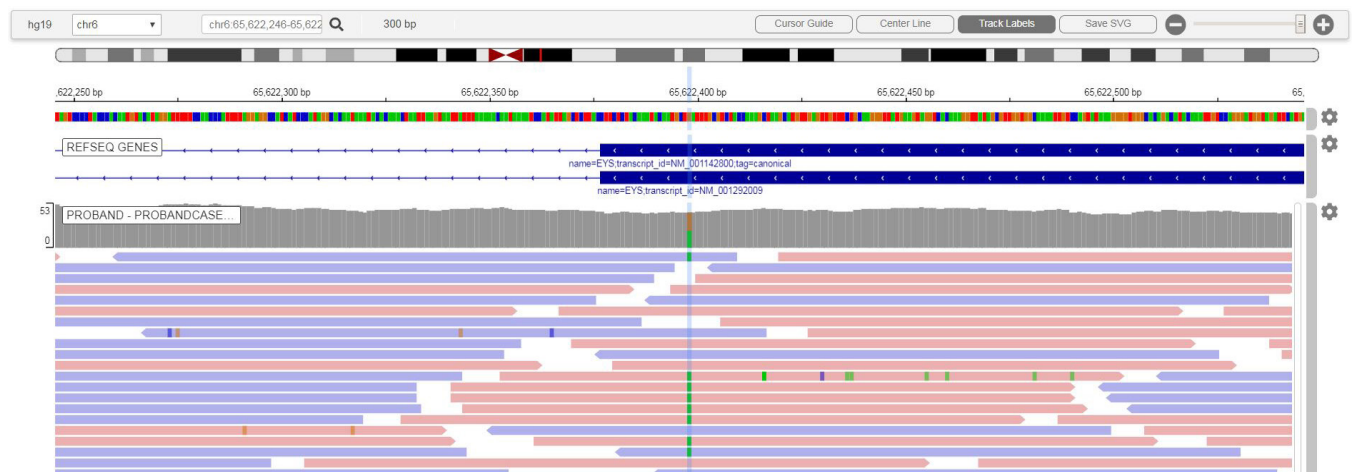


Figure 3: Variant calling confirmation—The lab analyst can access an embedded IGV for detailed variant inspection and comparison to other cases and external databases.

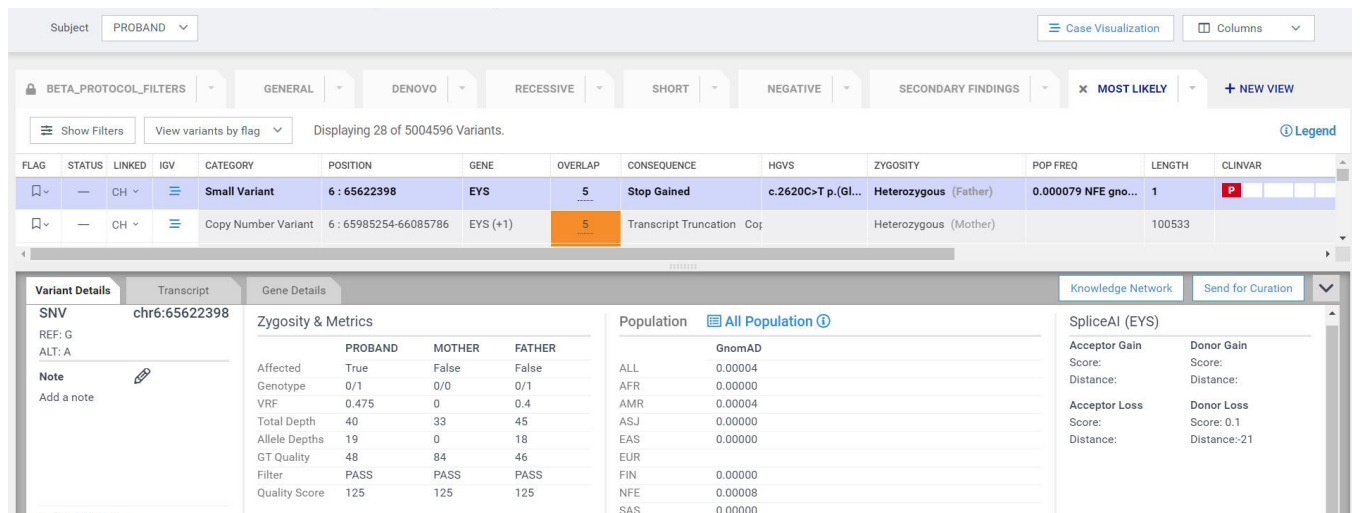


Figure 4: Variant interpretation and curation—Users can search multiple external databases within the software to aid in variant interpretation.

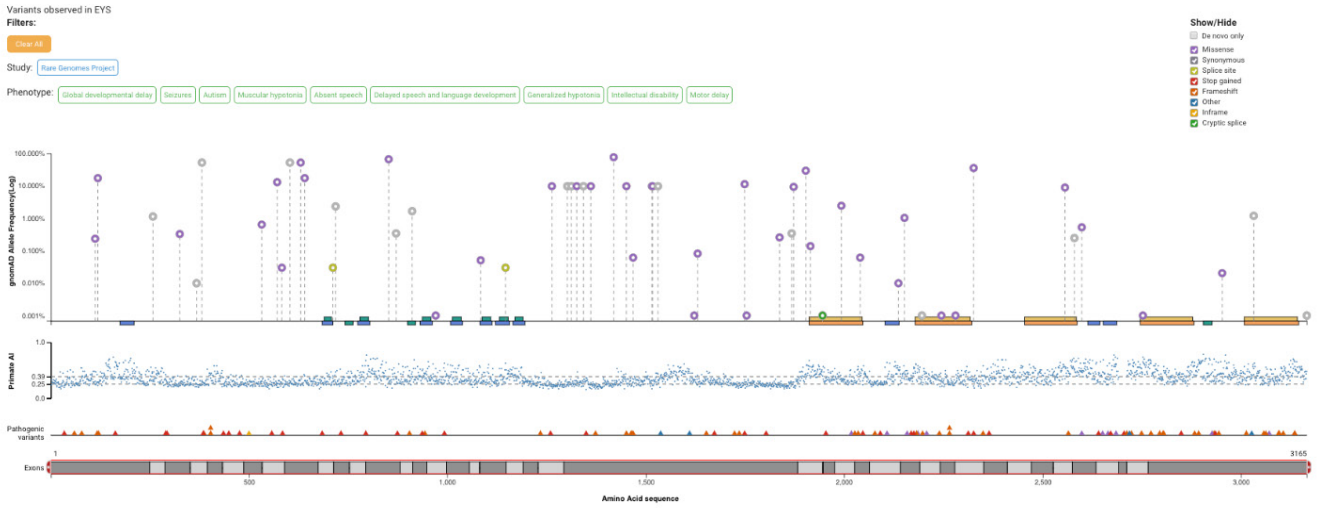


Figure 5: CaseLog—The lab analyst can visualize aggregate data for both variants and genes of interest.

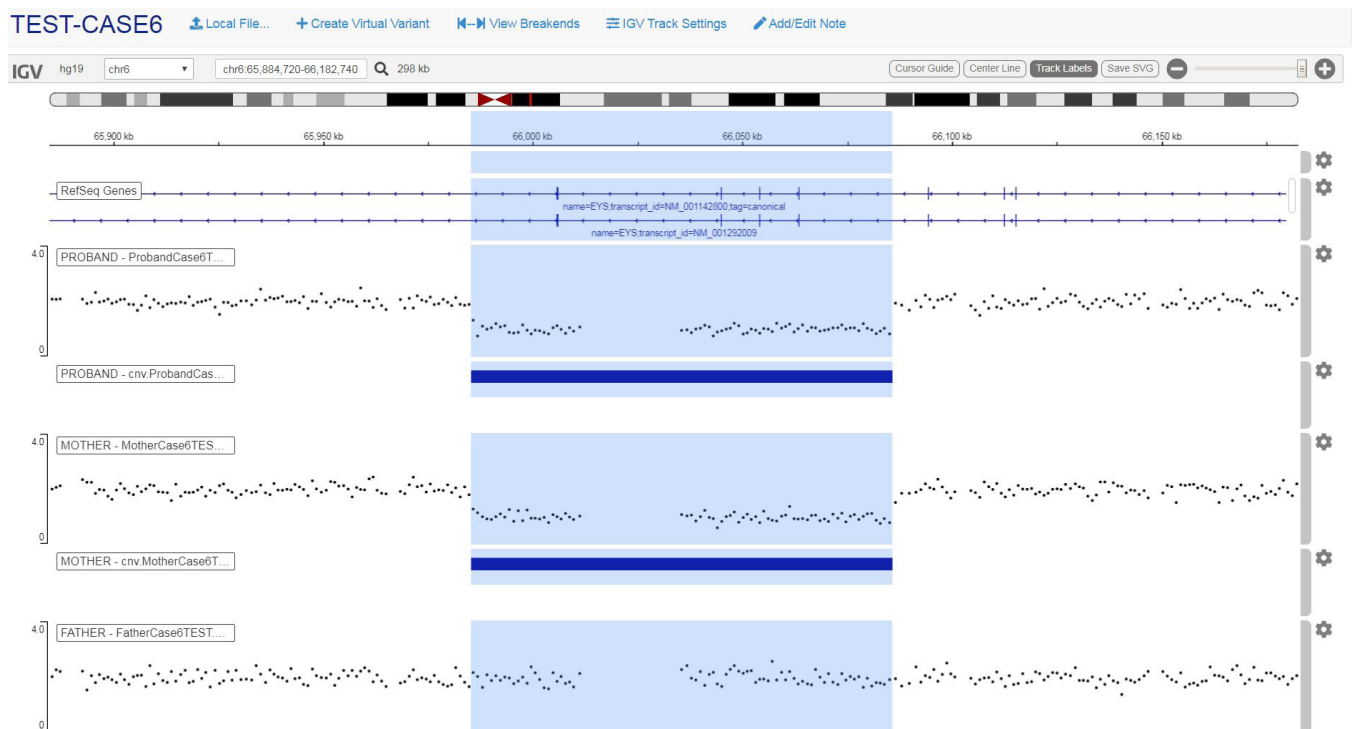


Figure 6: Variants of interest in Family 1—Using the tools in the TruSight Software Suite, the lab analyst has identified variants of interest in the proband: a compound heterozygous variant in the *EYS* gene consisting of a small stop-gain variant and a CNV. This image depicts the normalized coverage plot with a clear copy number loss in both the proband and the mother, highlighted in the blue box.

(Figure 2). Also, the analyst can access a growing list of tools powered by artificial intelligence (AI)-based machine learning, such as SpliceAI and PrimateAI to further prioritize variants of interest. These deep learning networks provide unbiased, highly accurate classification of mRNA splice sites and missense variants, respectively.^{1,2} After triage, the lab analyst has filtered out millions of irrelevant variants and is able to focus on the top variants of interest for the case for visualization and interpretation.

Variant visualizations

The analyst uses embedded visualization tools such as the Integrative Genomics Viewer (IGV) to further inspect genomic data such as read-level alignments and pair-end reads, b-allele frequency and normalized coverage tracks for all subjects within a case (Figure 3). In addition to variant level visualizations, the analyst can zoom out to view either an entire chromosome or whole genome to look for large anomalies. After variant calls have been confirmed, the lab analyst proceeds with interpretation for the case.

Variant interpretation and curation

Finally, the lab analyst works to determine which prioritized variants are relevant to the current case. To do this, the lab analyst interrogates gene and variant disease associations with overlapping phenotypic features similar to those in the proband. TruSight Software helps by aggregating and integrating preferred external databases such as the Online Mendelian Inheritance in Man (OMIM) catalog, ClinVar, and others into the Variant Details tab (Figure 4). The lab analyst can view the aggregated data to gain valuable insights into a variant, eliminating the need for repeated online searches in separate databases. Additionally, the TruSight Software Suite integrates standard terminology for variant classification developed by the American College of Medical Genetics (ACMG), which helps with recording variant details and associations. This enables customers to access and store gene-level information (eg, tolerance to loss-of-function variation, etc) and characteristics of gene-disease relationships (via preferred external databases).

The lab analyst can also use features such as the Note field to add case-specific notes pertaining to a specific variant, or the Comments field to record case-independent information about the variant or gene, which may be valuable if observed in future cases. CaseLog is used to view and aggregate gene, variant, and phenotypic information for each case across both private and public data sets (Figure 5). This customer-specific database of variant information enables reanalysis of historical data based on new discoveries in the scientific community.

Case results and reporting

After interpretation, the lab analyst has identified and curated variants of interest in the proband with known disease associations. A compound heterozygous variant was found in the *EYS* gene that includes a small stop-gain variant and a CNV that results in deletion of multiple exons (Table 2, Figure 6). Mutations in the *EYS* gene have known associations with retinal pigmentosa (RP), a genetic disease that results in an incomplete loss of vision that includes night and peripheral vision.

The lab analyst creates a new case report and adds gene and variant associations relevant to the case. The lab analyst can send the report for additional review and approval within the software. For ease of data sharing, the report can be downloaded as a PDF or JSON format. The Case status is now updated to Reported in the Case Management Portal.

Table 2: Family 1 proband variant interpretation

Variant of interest	Interpretation
<i>EYS</i> Small Variant 6:65622398 (c.2620C>T; p.Gln874Ter) Heterozygous	The c.2620C>T; p.Gln874Ter variant introduces a premature stop codon that results in early truncation of the <i>EYS</i> protein.
<i>EYS</i> CNV 6:65985254-66085786 Heterozygous	The CNV results in a loss of ~100 Kb of sequence that includes multiple exons of the <i>EYS</i> gene.

Summary

As demonstrated in this use case, a laboratory analyst can use the intuitive and comprehensive features and tools of the TruSight Software Suite to visualize, triage, and interpret variants associated with genetic disease.

Learn more

Learn more about the TruSight Software Suite at www.illumina.com/trusight-software-suite.html

References

- Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, et al. [Predicting splicing from primary sequence with deep learning.](#) *Cell*. 2019;176(3):535–548.
- Sundaram L, Gao H, Padigepati SR, et al. [Predicting the clinical impact of human mutation with deep neural networks.](#) *Nat Genet*. 2018;50(8):1161–1170.

About TruSight Software Suite

The TruSight Software Suite is a software as a service (SaaS) analytics solution that integrates with the NovaSeq 6000 Sequencing System and the DRAGEN Bio-IT Platform to enable comprehensive and streamlined secondary and tertiary analysis workflows for NGS. The TruSight Software Suite offers:

- **Simplified, customizable case management**—Manage cases from sequencing data to report, assign cases to users, configure pipeline settings, and set quality control (QC) thresholds
- **Comprehensive, ultra-rapid variant calling**—Use DRAGEN secondary analysis to call small variants, CNVs, and structural variants, mitochondrial variants, repeat expansions, and more
- **Intuitive, high-powered interpretation**—Filter variants via gene lists, inheritance modes, custom annotations, and complex logic; flag, sort, and prioritize important variants

For a full list of current and future functionality, please contact an Illumina sales associate for additional information.