illumina®

TruSight™ Software Suite

Bringing efficiency and high confidence to case management, variant analysis, and interpretation in rare disease.

Highlights

- **Comprehensive genomic evaluation**
  Analyze, visualize, and interpret small variants, structural variants, mitochondrial variants, repeat expansions, runs of homozygosity, and *SMN1/SMN2* variants
- **Ready-made, integrated workflow**
  Keep pace with evolving technology using a ready-made infrastructure and simplified integration of diverse analytical tools to optimize the benefits of next-generation sequencing

Introduction

Whole-genome sequencing (WGS) and whole-exome sequencing (WES) using next-generation sequencing (NGS) technologies are powerful methods for investigating variants linked to genetic disease. In a meta-analysis of literature from January 2011 to August 2017, 37 studies comprising 20,068 children were reviewed for the diagnostic utility of three testing approaches: chromosomal microarray (CMA), WES, and WGS. Results showed 8.3× greater odds of diagnosis with NGS methods, compared to CMA (Figure 1).[1]

WGS and WES provide a high-resolution, unbiased view across the entire genome to discover causative variants associated with rare diseases. However, the vast amounts of data produced by these methods represent a significant bottleneck and require comprehensive data analysis tools that can efficiently translate the raw sequencing data into meaningful, interpretable results. To address this challenge, Illumina offers TruSight Software Suite. This software as a service (SaaS) integrates with BaseSpace™ Sequence Hub and Illumina sequencing systems to access run monitoring, run metrics, and automated sequencing data upload. It includes cloud-based access to the DRAGEN™ Bio-IT Platform, enabling comprehensive, streamlined secondary and tertiary analysis workflows for NGS (Figure 2).

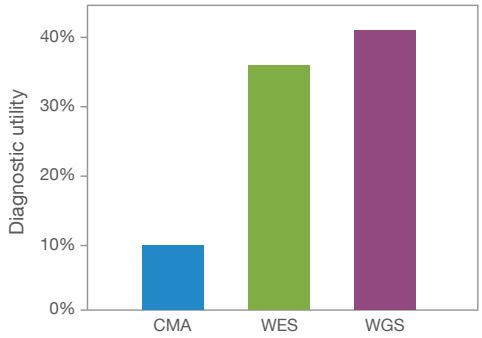

Figure 1: WGS and WES have higher diagnostic utility than CMA—Quantitative analyses of 37 studies comprising 20,068 children for diagnostic utility of first-line genomic tests showed 36% and 41% utility for WES and WGS, respectively, compared to 10% for CMA. 95% CI: 4.7-14.9, P < 0.0001.

Variant analysis in TruSight Software Suite

Secondary analysis includes:

- Alignment and variant calling using the DRAGEN platform

Tertiary analysis includes:

- Variant annotation
- Variant filtering and triage
- Variant visualization
- Variant curation
- Variant interpretation and customized reporting

Powered by the DRAGEN platform

TruSight Software Suite is powered by the DRAGEN (Dynamic Read Analysis for GENomics) Bio-IT Platform, providing secondary analysis of genomic data. Fundamental features of the DRAGEN platform address common challenges in genomic analysis, such as lengthy

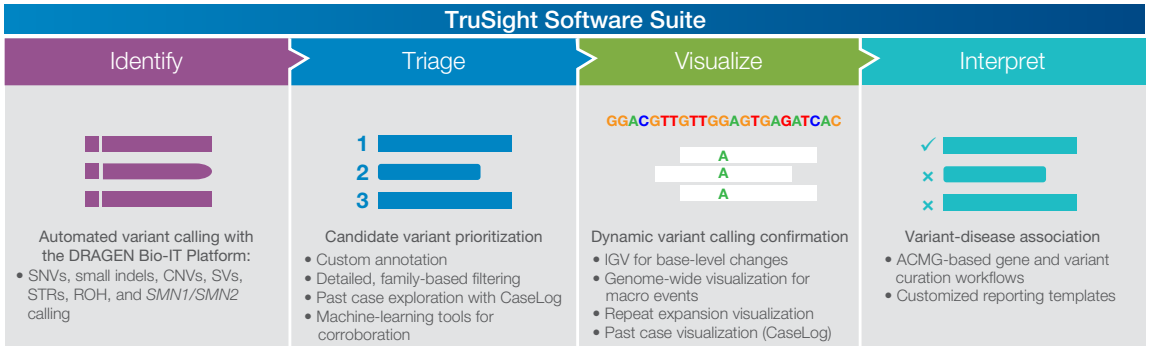

Figure 2: Variant analysis in TruSight Software Suite—Variant analysis in TruSight Software Suite begins with automatic alignment and variant calling using the DRAGEN platform, requiring approximately three hours for a 30× trio, followed by triaging, visualizing, and interpreting variants.

compute times and massive volumes of data. Without compromising accuracy, the DRAGEN platform delivers quickness, flexibility, and cost efficiency, enabling labs of all sizes and disciplines to do more with their genomic data. Comprehensive variant calling includes single nucleotide variants (SNVs), insertions/deletions (indels), copy number variants (CNVs), structural variants (SVs), short tandem repeats (STRs), repeat expansions, runs of homozygosity (ROH), *SMN1/SMN2* calling, and more (Figure 3). While TruSight Software Suite is compatible with both WGS and WES outputs, the DRAGEN platform only supports repeat expansion calling and *SMN1/SMN2* calling with whole-genome samples. See the TruSight Software Suite user guide for additional details.

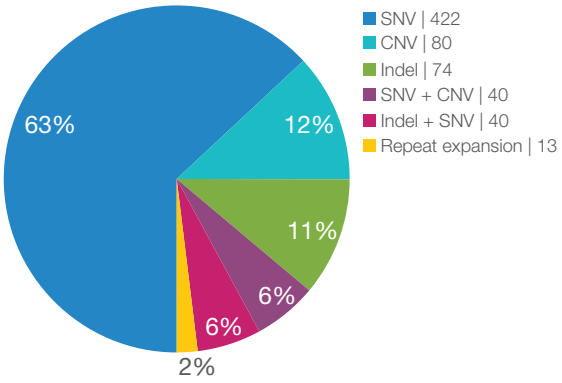

Figure 3: Variant types identified with TruSight Software Suite—Analysis of WGS data resulted in detection of different variant types. Percentages indicate number of cases out of 669 total cases.

Integration with other platforms and systems

Many labs struggle to keep pace with integrating new genomics technology, instruments, and methodologies. TruSight Software Suite simplifies the process, integrating seamlessly with the NovaSeq™ 6000, NextSeq™ 2000, or other systems via BaseSpace Sequence Hub for automating WGS and WES analysis. Furthermore, TruSight Software Suite represents the final piece in the rare disease workflow of an integrated, DNA-to-report WGS solution, including Illumina DNA PCR-Free Prep, Tagmentation, and the NovaSeq 6000 System. Compatibility with Application Programming Interfaces (APIs) enables integration with other institutional laboratory information management systems (LIMS). The software provides a complete data storage architecture to manage short- and long-term storage of FASTQ, VCF, BAM files, etc, in a cost-effective and secure manner.

Simplified, customizable case management

TruSight Software Suite features a Case Management Portal that allows users to create new cases, import data files, and associate sequencing data files with each case. Alternatively, this information can be easily imported via an API. Users enter familial relationship information about each case (up to five individuals), including family structure, proband gender, proband phenotypic features (optional), and affected status of family members to improve variant filtering and prioritization.

Cases can be assigned to specific roles or functions within a laboratory to improve efficiency. Real-time updates of case status are displayed in the TruSight Software Suite dashboard, a single view to monitor a laboratory's entire caseload. This allows managers and other personnel to monitor progress through the analysis workflow.

Intuitive, high-powered interpretation

TruSight Software Suite uses critical data aggregation, variant visualization, variant curation, and machine-learning tools to promote efficient and informed interpretation.

Variant triage

Using the Interpretation tab in TruSight Software Suite, variants can be filtered following a custom plan or a prebuilt filter plan. Family-based filtering in TruSight Software Suite enables comparison of the proband with other family members. Additional options include filtering on population frequencies from sources such as the Genome Aggregation Database (gnomAD), variant consequences, modes of inheritance, ClinVar pathogenicity, and more.

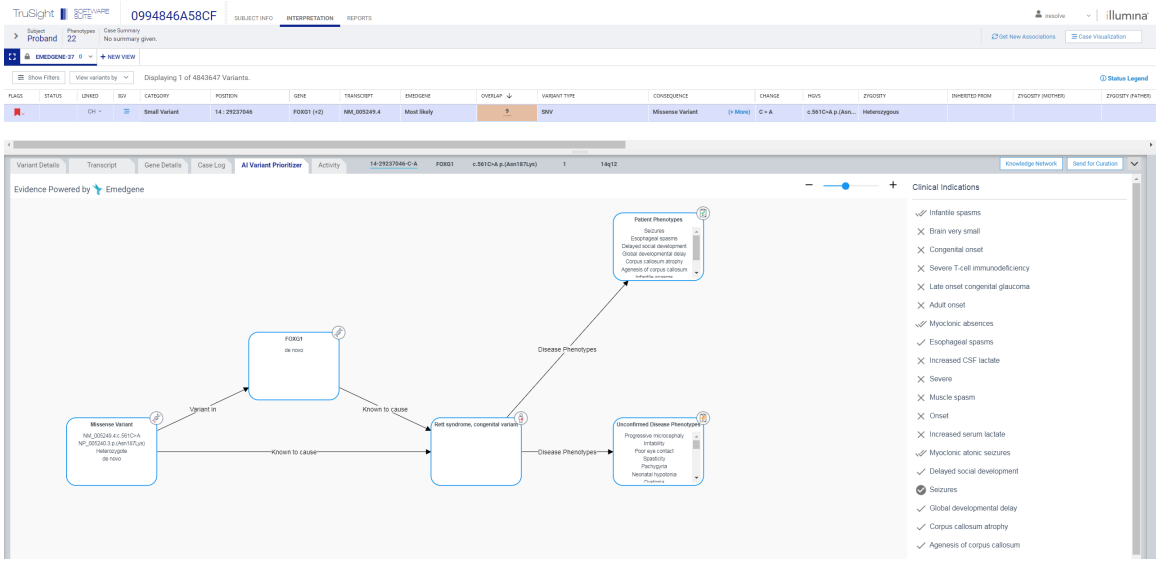

Figure 4: Variant filtering and prioritization—Family-based variant filtering enables identification of inherited and *de novo* variants. The variant grid, which is customizable by each user, shows information on category of variant, chromosomal position, gene affected, overlap (number of overlapped phenotypes for the variant), consequence of the variant, population frequency of the variant (if known), and more.

Variant analysis with machine learning

TruSight Software Suite incorporates a genomics artificial intelligence (AI) engine, powered by Emedgene, to rank variants and highlight the most-likely candidates. The engine generates a knowledge-graph showing supporting evidence for the variant prioritization, like disease–gene relationships, generated by the application of natural language processing (NLP) to various data sources (Figure 4).

Illumina worked with collaborators at Stanford University, University of California, San Francisco, University of Florida, University of Chicago, and the Broad Institute to develop the SpliceAI and PrimateAI analysis tools. These state-of-the-art deep neural networks are powered by machine learning to find disease-causing mutations. SpliceAI and PrimateAI provide unbiased, highly accurate classification of mRNA splice sites and missense variants, respectively.[1-3] Using these automated prioritization tools, users can add depth to their analysis and quickly filter out millions of variants to focus on the top, candidate variants of interest for visualization and interpretation.

Variant visualizations

TruSight Software Suite features embedded visualization tools, such as the Integrative Genomics Viewer (IGV), for further inspection of genomic data, including read alignments, variants, B-allele frequency, and coverage tracks for all subjects within a case. In addition to variant-level visualizations, the IGV offers views of an entire chromosome or whole genome to look for large anomalies.

Variant interpretation and curation

TruSight Software Suite offers various features to help determine which prioritized variants are relevant to the current case, enabling interrogation of gene– and variant–disease associations with overlapping phenotypic features similar to those in the proband.

TruSight Software Suite aggregates and integrates data from preferred external databases, such as the Online Mendelian Inheritance in Man (OMIM) catalog, ClinVar, and others, into the Variant Details tab. The aggregated data can be viewed to gain valuable insights into a variant, eliminating the need for repeated online searches in separate databases. Additionally, 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 access and storage of gene-level information (eg, tolerance to loss-of-function variation, etc.) and characteristics of gene–disease relationships (via preferred external databases). Transcript-level information is also displayed for each variant, and both canonical and noncanonical transcripts can be selected for interpretation. Features such as the Note field can be used to add case-specific notes pertaining to a specific variant. The Comments field can be used to record case-independent information about the variant or gene, which may be valuable if observed in future cases.

CaseLog: a customer-specific database

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 interactive database stores public, rare disease data sets and cases previously seen by a lab to inform curation, interpretation, and reporting of genes or variants of interest based on new discoveries in the scientific community.

Results and custom report generation

Interpretation is complete when variants have been identified and curated with known disease associations. Customers can use templates in TruSight Software Suite to customize reports of gene and variant associations relevant to cases (Figure 6). The report can be

Figure 5: CaseLog—The CaseLog feature in TruSight Software Suite enables visualization of aggregate data for both variants and genes of interest.

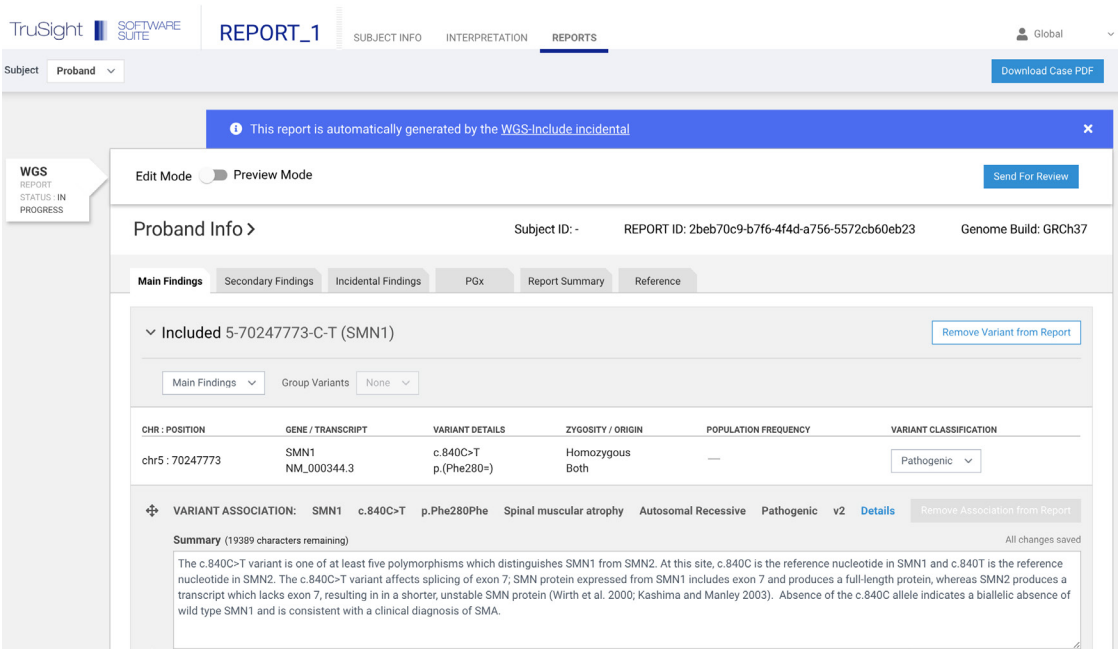

Figure 6: Customizable report generation—TruSight Software Suite offers a template for customization of reports of gene and variant associations relevant to the case.

sent for additional review and approval within the software. For ease of data sharing, reports can be downloaded in a PDF or JSON format.

Secure, compliant environment

TruSight Software Suite is ISO-27001 and ISO-13485 certified and complies with Health Insurance Portability and Accountability Act (HIPAA) (third-party audited) and the principles of the General Data Protection Regulation (GDPR). TruSight Software Suite also offers options to integrate with a lab's single sign-on policy and other security settings.

Summary

TruSight Software Suite offers an intuitive and comprehensive rare disease analysis and interpretation solution. It integrates with Illumina sequencing systems and includes the DRAGEN Bio-IT Platform for ultra-rapid variant calling and features tools to visualize, triage, and interpret variants associated with genetic disease. Results can be output using customizable templates for customer-specific reports.

Learn more

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

References

1. Clark MM, Stark Z, Farnaes L, et al. Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected diseases. *NPJ Genom Med*. 2018;3:16.
2. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, et al. Predicting splicing from primary sequence with deep learning. *Cell*. 2019;176(3):535–548.
3. 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.

Ordering information

Illumina offers a 30-day free trial, allowing customers to work with example cases available in TruSight Software Suite or upload and evaluate their own cases within the software.

Product	No. of samples	Catalog no.
TruSight Software Suite	48 WGS/96 WES	20041943
TruSight Software Suite	96 WGS/192 WES	20041944
TruSight Software Suite	288 WGS/576 WES	20041945
TruSight Software Suite	480 WGS/960 WES	20041946
TruSight Software Suite	960 WGS/1920 WES	20041947
TruSight Software Suite	2400 WGS/4800 WES	20041948
TruSight Software Suite	4800 WGS/9600 WES	20041949
TruSight Software Suite	9600 WGS/19,200 WES	20042010
TruSight Software Suite 30 day, 15 sample free trial		20042019
TruSight Software Suite Training at customer site (1 day)		20042020
TruSight Software Suite Training at Illumina Solutions Center (1 day)		20042021

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