GENEYX

Geneyx Analysis Technical Datasheet: A clinical genetics data management platform

GENEYX ANALYSIS TECHNICAL DATASHEET: A CLINICAL GENETICS DATA MANAGEMENT PLATFORM

Overview:

Geneyx Analysis is a comprehensive bioinformatic solution for analyzing and interpreting DNA/RNA sequencing data for both NGS and microarrays. This solution offers scalability for short-read and long-read workflows and provides novel biomedical insights for variant-gene and gene-disease associations. With these capabilities, *Geneyx Analysis* has been adopted globally for research, academia, and clinical diagnostic workflows.

Introduction:

Geneyx Analysis empowers hospitals and genetic laboratories with an innovative solution for clinical applications and services spanning germline and somatic workflows. Ingenuity and useability are at the forefront of the application with alignment to both secondary and tertiary workflows. This enables FASTQ and VCF data uploads, among other file formats, for both individual and batch processes. For optimization, Geneyx Analysis leverages comprehensive and up-to-date annotation sources for all variant types. This includes single nucleotide variants (SNVs), insertions and deletions (Indels), copy number variations (CNVs), repeat expansions, and structural variations (SVs). Following variant analysis and interpretation, the results can be easily rendered into a customizable clinical report and integrated into a LIMS or EHR system.

Features:

Variant Analysis

For variant analysis, the platform offers an intuitive and robust variant browsing interface. Variants are prioritized using customizable filter logic and can be easily sorted using phenotypic prioritization. Annotations are condensed to display the most useful resources but can also be expanded to view the entire genomic information. Hyperlinks are interactive and data can be easily visualized in genome browse for validation purposes. Once a variant has been interpreted, the case can be transferred to other users in your organization and the status can be changed to reflect the current stage in the pipeline.

Rapid & Comprehensive Annotations

Geneyx Analysis optimizes annotation updates for over 50 data sources. These updates occur bi-monthly and include standard ACMG-recommended databases together with all publicly available annotation sources. Licensed annotations can also be integrated on request and all annotation source information can be pulled into a customizable clinical report.

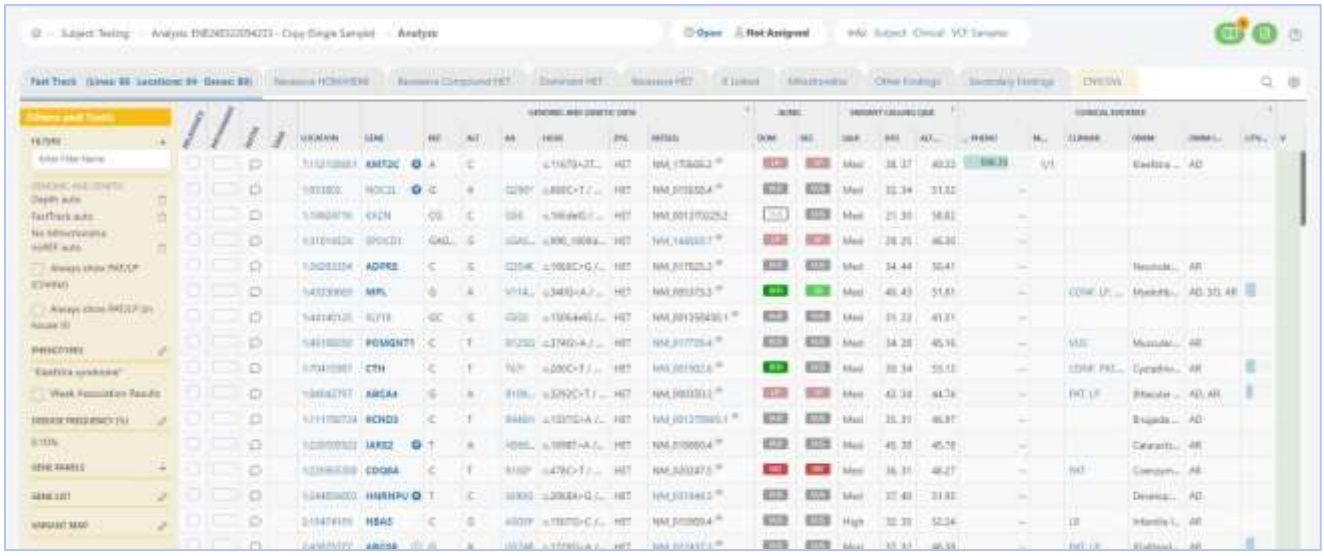


Figure 1: Variant interpretation interface

Annotating a VCF file includes:

1. Variant frequency, clinical significance, and damage effect predictions.
2. SFARI and Orphanet gene data are integrated.
3. OMIM Clinical Synopsis enables users to view gene-specific summaries of clinical features.
4. The AlphaMissense provides pathogenicity scores for missense variants, improving variant interpretation.
5. ClinGen Dosage Sensitivity is available: ClinGen Dosage Sensitivity collects evidence supporting or refuting the haploinsufficiency and triplosensitivity of genes and genomic regions.
6. OMIM for CNV/SV**:** OMIM phenotype information is now available for CNV/SV analysis. This new field will display OMIM information for all overlapping events observed in the database.
7. Incorporated CoLoRS data for detecting complex variants from long-read sequencing.
8. ClinVar Origin Annotation: The allele origin, being either germline or somatic, is now displayed when viewing the ClinVar hyperlink for SNVs of interest.
9. Recessive Het Tab: To accommodate for carrier screening workflows, or heterozygous variants with autosomal recessive mode of inheritance, a Recessive Het tab has been developed that includes the ACMG classification using the autosomal recessive model.
10. Splitting multi-allelic variants and merging identical allelic variants on different transcripts as per ACMG interpretation guidelines.
11. Clinical interpretation of pathogenicity based on ACMG guidelines.
12. Annotations based on the user account internal databases, including allele frequency and previous variant and gene annotations

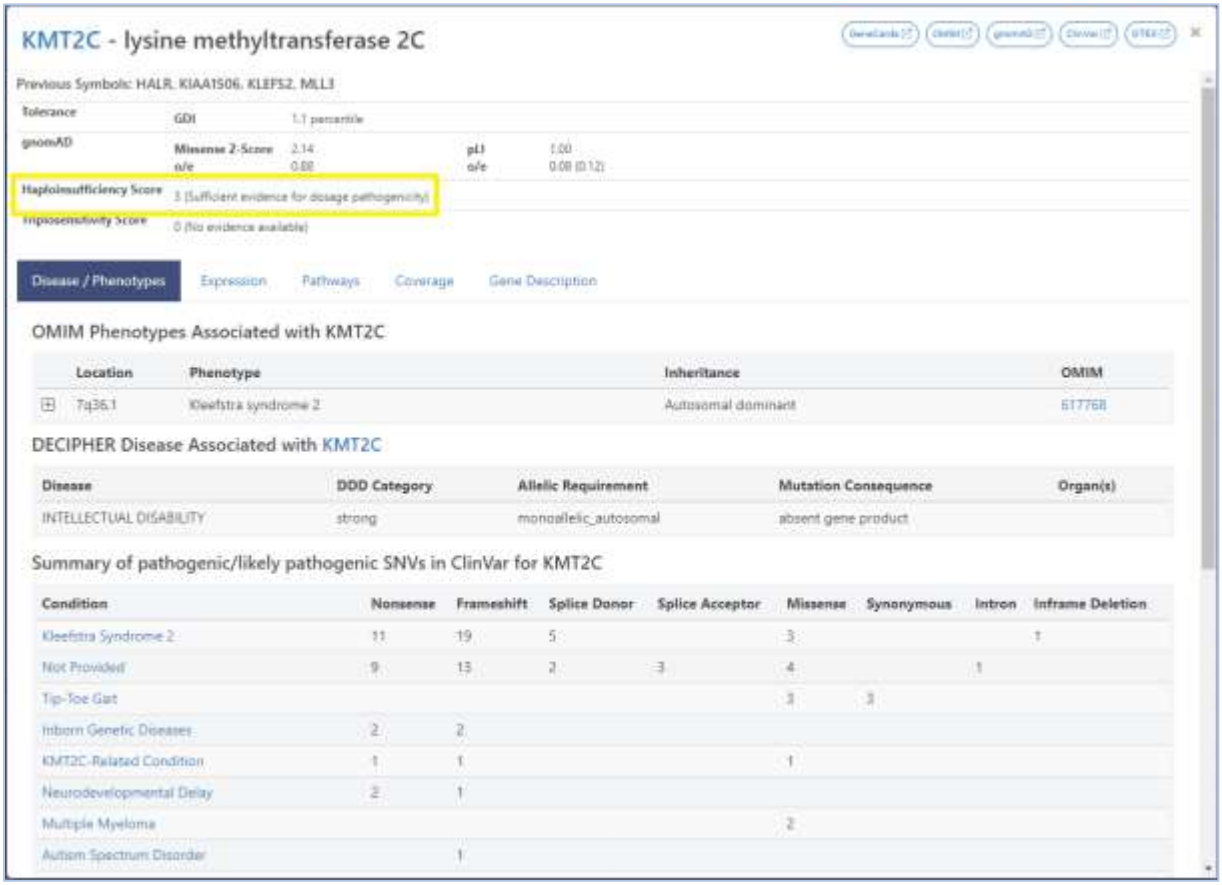


Figure 2: HI/TP scores are available via ClinGen

CNVs and SVs can be investigated at the gene level and are annotated with clinical data and can be prioritized according to associated phenotype. Compound mutations can also be assessed to view SNVs within the CNV/SV with clinical significance.

Phenotype-Driven Variant Prioritization

- *Free Text Phenotypes* – Variants can be prioritized based on their association with any free-text biological term or HPO terms, and phenotype specificity is displayed to show gene heterogeneity of the phenotype entered. Multiple phenotypes can be entered using the advanced options and phenotypes can be modified once an analysis has been created.
- *Phenotypic Evidence Score* – Variant scores are created from investigating over 50+ biomedical data sources with the ability to deep-dive into relevant literature to find novel disease associations.
- Geneyx's phenotyper will highlight association with regulatory regions, as well as enhancers.

CACTGTT

- If a phenotype has been observed in a different sample and occurs in the same gene, the phenotypic score will increase based on integration of the users' Clinical Knowledgebase. This is performed on the backend and is proprietary to the users account.

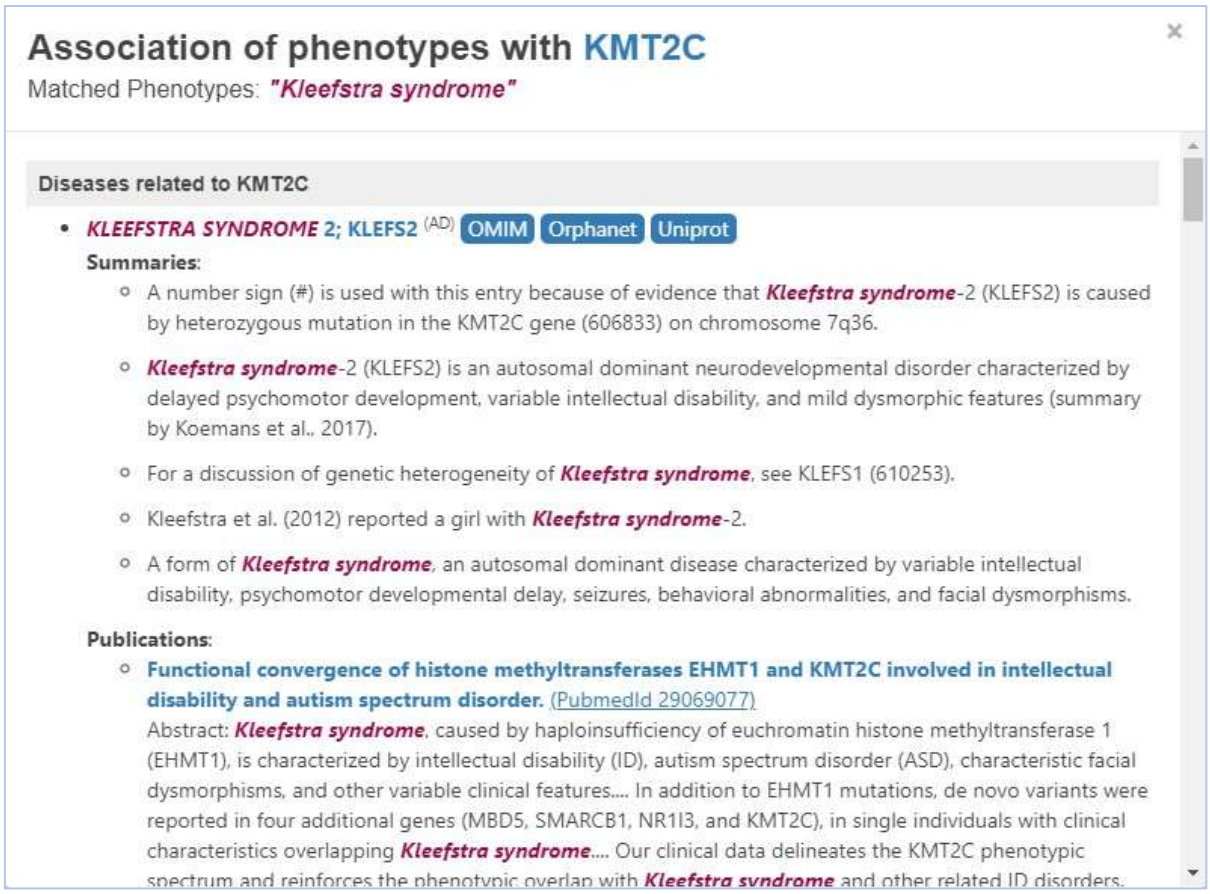


Figure 3: Phenotypic prioritization using key search terms

ACMG CNV/SV Guidelines

The CNV/SV tab in Geneyx Analysis adheres to the joint ACMG and ClinGen technical standards for reporting constitutional copy number variants (CNVs). ACMG guidelines are automatically applied to all deletion and duplication events, and the interface allows modification of criteria using internal evidence. Users gain full transparency with a detailed view of activated, inactivated, and unmet criteria, as well as comprehensive audit trails for all actions performed.

Key Features:

Interactive Criteria Interface: Criteria are categorized by ACMG-defined standards, with tooltips offering evidence summaries for each criterion. Users can activate/inactivate or adjust scores to tailor assessments to specific cases.

Legend for Automated Criteria: Visual indicators identify scored, unmet, or inactive criteria, ensuring clarity in evaluation workflows.

Stored Interpretations: Adjustments and notes are securely saved, enabling consistent reuse for future encounters of the same variant.

This streamlined approach supports standardized, precise, and reusable interpretation workflows for CNVs and structural variants (SVs) such as deletions and duplications. It aligns with ACMG/ClinGen best practices, enhancing diagnostic accuracy while optimizing efficiency.

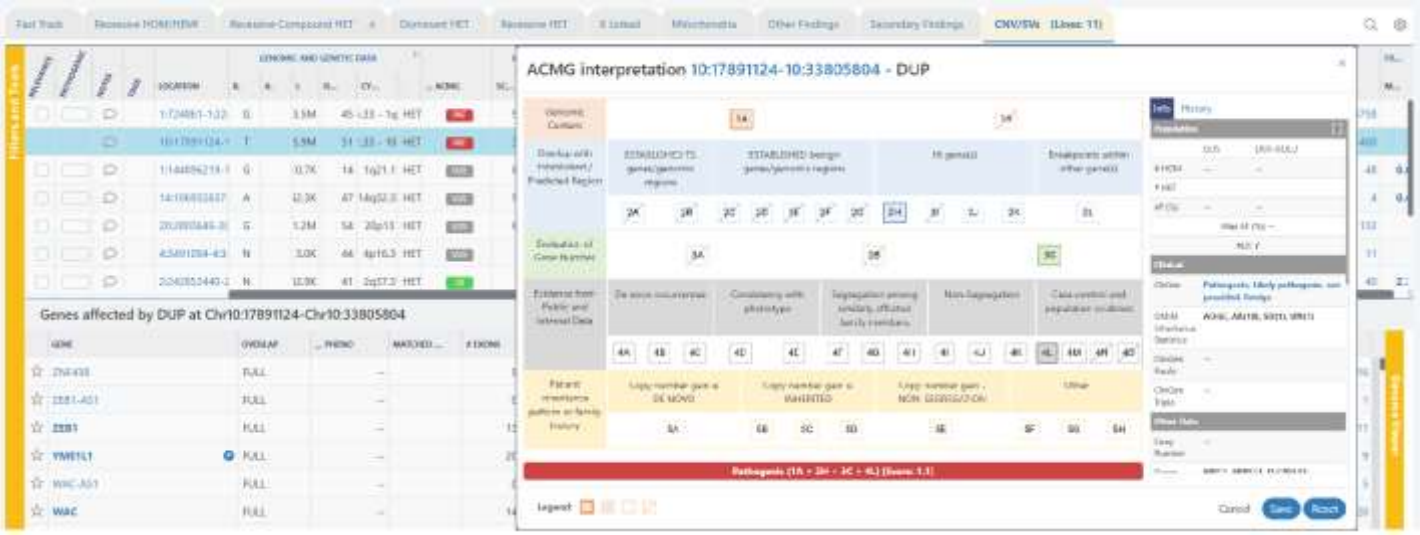


Figure 4. ACMG Guidelines for CNV/SV

ACMG SNV Guidelines

The American College of Medical Genetics and Genomics (ACMG) represents the interests of clinical geneticists, clinical laboratory geneticists, and genetic counselors by providing best practice guidelines. Geneyx Analysis adheres to these guidelines by providing:

- *ACMG Automation* – Geneyx Analysis automates 18/28 ACMG criteria, reducing workflow complexity and ensuring best practices are followed. Users can define thresholds for application of a subset of criteria.
- *Sample & Case-context classification*- ACMG classifications automatically include case-level information such as inheritance models and relevant findings in associated samples.
- *Ability to modify* – If additional information or expertise is available, users can incorporate their data, modify criteria, make notes, and store changes for all downstream cases.
- Easy access to information relevant to manual curation of variant pathogenicity
- Support for reanalysis with updated evidence and annotations

Geneyx Analysis adherence to the ACMG Guidelines has been validated by labs in the United States, Europe, Israel, Hong Kong, and China. Further, it has been approved for use in CLIA-CAP genetic laboratories.

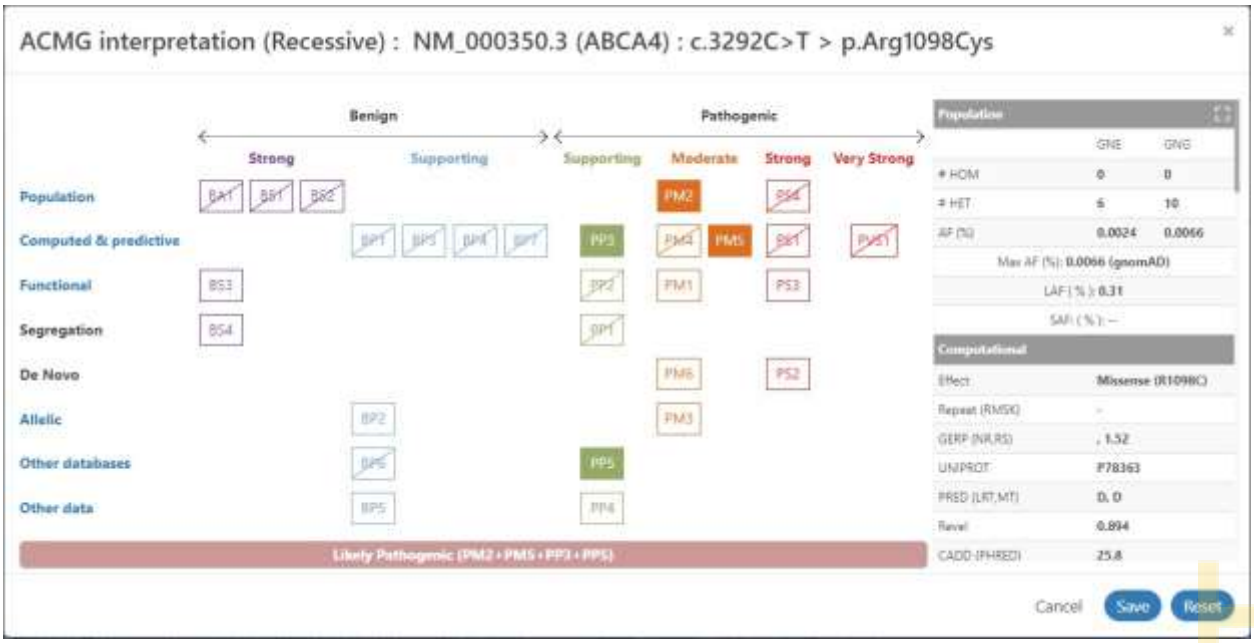


Figure 4: ACMG dialog with ability to modify criteria

Customer Local Database & Inter-Lab Sharing

Geneyx Analysis automatically calculates an 'in-house' allele frequency annotation for all variants in an account that is derived from all imported samples. This is useful to eliminate common variants or enable variant selection among different ethnic groups. The in-house database also includes annotations and interpretations previously entered by the analysts, which assists in applying accumulated in-house knowledge to new cases. Laboratories may also participate in

Figure 5: In-house database in green shows allele frequency and previous classifications from other samples

Whole Genome Sequencing Interpretation

Geneyx Analysis incorporates a comprehensive genome-wide map of regulatory elements including promoters and enhancers together with their gene associations from various sources. This enables the association of variants in non-coding regions to gene-phenotype relationships and subsequent prioritization.

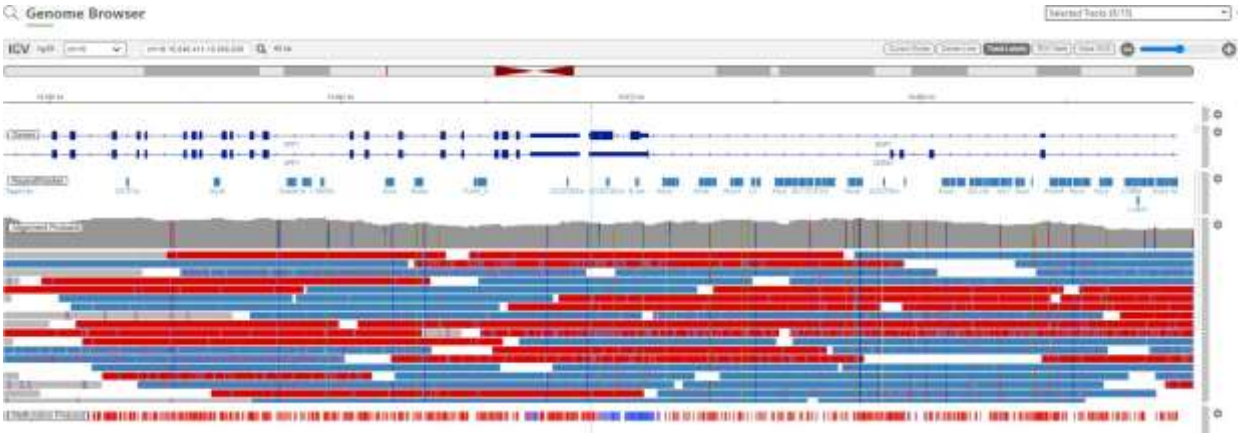


Figure 6: BAM file visualization using integrated IGV

Evidence Collection and Reporting

- *Automated evidence collection* – Geneyx Analysis supports automatic generation of comprehensive clinical reports leveraging available annotations & phenotypic evidence.
- *Customizable report templates* – Reports are fully customizable to include the organization logo and information that the user deems necessary for a distinctive design.
- Selected Variants' feature: Once the relevance and associated information is set, user can review the selections across all the genetic models using this feature.
- *Exporting flexibility* – Reports can be rendered in PDF and JSON format and integrated with LIMs/EHR systems where necessary.

Figure 7 : Annotating variants to pull into report and in-house databases

Storage and Pipeline Management

- *Unlimited Cloud-based Genetic Data Storage* – Easily manage and stores genetic data in a secure private Geneyx Azure cloud, (HIPAA and GDPR-compliant).
- *Local or Hybrid Configurations* – On-premise configurations and Hybrid

architecture are also available if an organization needs to remain behind local firewalls.*Account Management* – Restrictions can be set at a user or group level to ensure assigned, fixed workflows are utilized to streamline processing.

- Storage Allocation: Display the storage space allocated or consumed by the data files within the specified timeline. This information allows users to track their storage usage and manage their resources effectively.
- *Audit Trails* – All information from FASTQ to report is documented and can be reviewed at any time.

Figure 8 : Complete visibility of data workflows and management features

Advanced Callers and Low-Pass Whole Genome Sequencing (WGS)

Geneyx Analysis supports advanced sequencing pipelines with options tailored for different reference genome assemblies and analysis requirements.

Pipeline Options:

DRAGEN Pipelines:

- Exome: Available for both hg19 and hg38 genome assemblies.
- Whole Genome: Optimized for hg19 and hg38 assemblies with support for DRAGEN v4.0.5 and optional upgrades to v4.2.4.

Advanced Callers (DRAGEN v4.2.4):
Enhanced sequence-graph realignment improves variant calling accuracy for challenging genes with high-identity paralogs, such as GBA, SMN1, HBA, LPA, and CYP2D6. Results are accessible in the Advanced Analysis section for WGS workflows with ≥30X coverage.

Low-Pass WGS:
Low-pass whole genome sequencing (<5x coverage) is now supported for efficient and cost-effective genomic analysis. Users can select the “Whole Genome Low Pass” target for these workflows.

Sentieon SNV Pipeline:
Sentieon offers accuracy, speed, and scalability for diverse applications, including cancer genomics, rare disease studies, and population genetics.

By selecting the appropriate pipeline based on sequencing data and reference genome assembly, users ensure accurate and efficient variant calling and analysis.

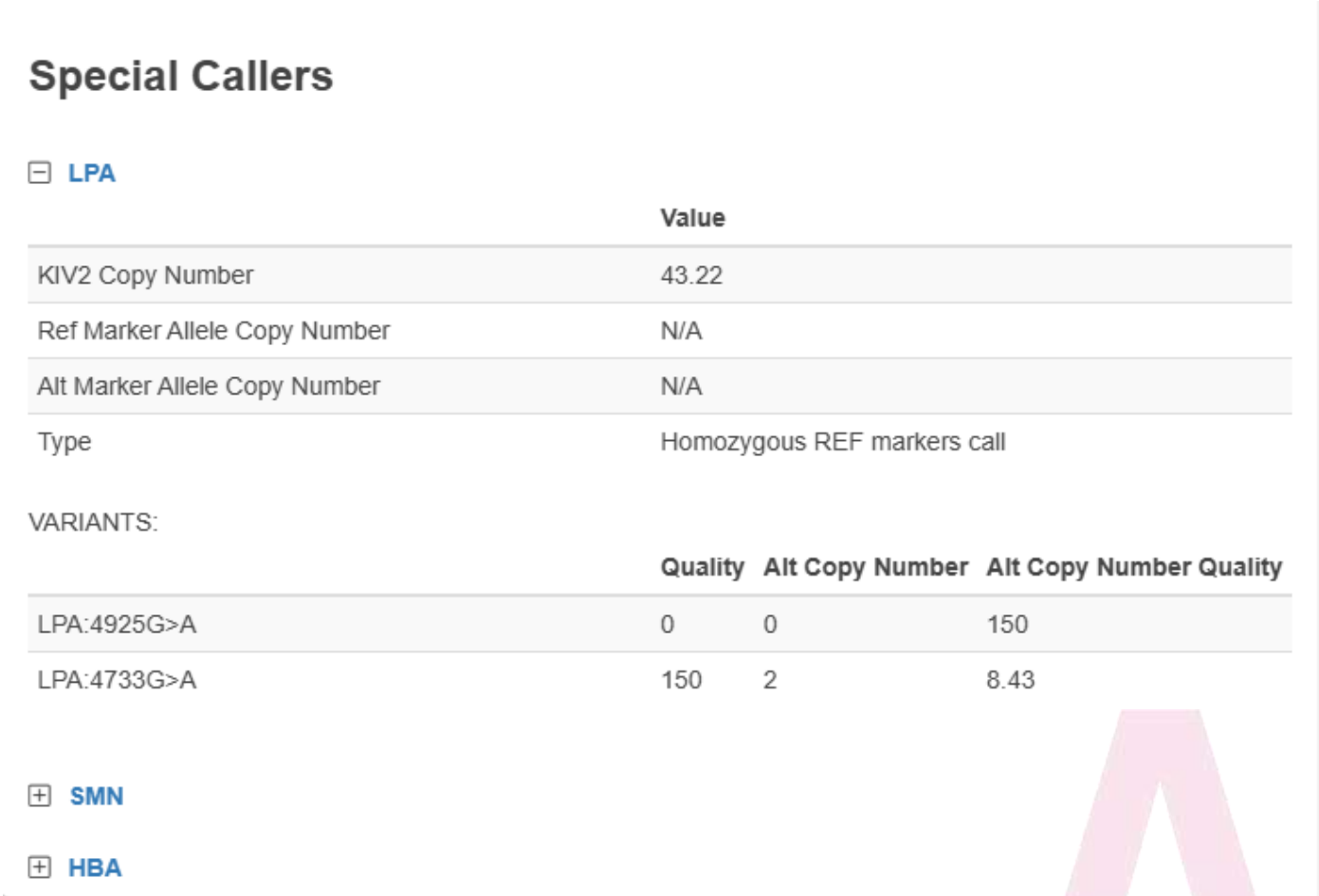


Figure 9: Special callers overview

Customizable Workflows and Protocols

Geneyx Analysis provides customizable protocols which enable laboratories to standardize genetic test and analysis workflows. Protocols can include combinations of simple or complex filters on any set of annotation entities for any gene lists or genomic regions and can be locked

as a protocol for defined users. Protocols can also be optimized to automatically render a report on the remaining variants and include any user interpretations for given variants.

Workflow highlights:

Case assignment

Gene panel configuration

Completely customizable filters

Protocols for germline, somatic, and carrier screening

- Role-based-permissions

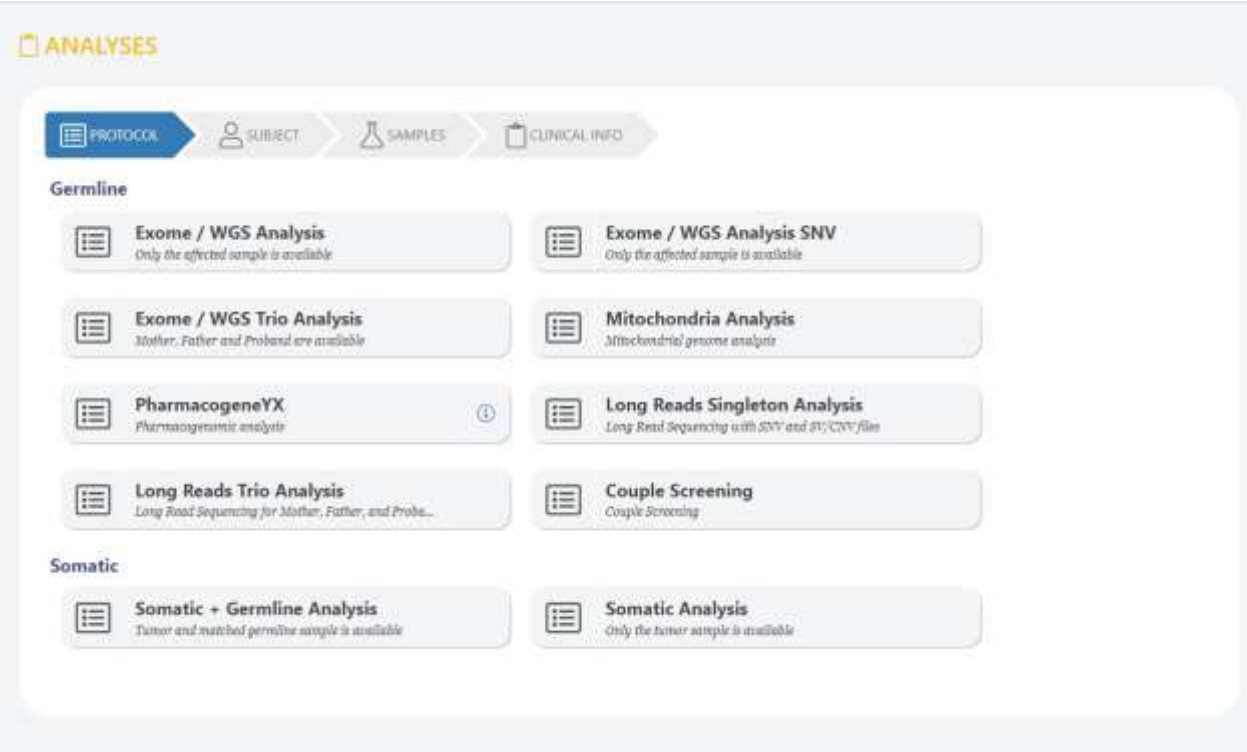


Figure 10: Default and customizable protocols are available

Pharmacogenomics (PGx) Workflow

Pharmacogenomics (PGx) is an integral part of precision medicine, focusing on the relationship between genetic variations and drug metabolism. The Geneyx PGx workflow offers comprehensive interpretations for 13 CPIC level A/B genes influencing the metabolism of 68 commonly prescribed drugs. This data is curated from reliable sources, including CPIC, FDA, and PharmGKB, with hyperlinks to these databases integrated into the report.

Key Features:

Customizable PGx Reports:

- Includes gene name, genotype, and the impact of genetic variations on drug metabolism.
- Users can exclude genes with low-quality or missing SNPs and make final adjustments using the report editor.
- Electronic signature sign-off ensures compliance and accountability.

Enrichment Kit for PGx Analysis:

- Utilizes PGx BED files for hg19 and hg38 genome builds.
- Supports both FASTQ and VCF input, with enrichment processes ensuring accurate star allele detection.

Star Allele Visualization:

- Easily access genes included in the report via a star icon feature.
- Genes with low coverage, poor quality, or missing SNPs can be excluded before report generation.

Licensing and Support:

A PGx license is required to access this workflow. For licensing or assistance with implementation, contact support@geneyx.com.

By leveraging advanced enrichment protocols and comprehensive data integration, the Geneyx PGx workflow provides powerful insights to optimize drug efficacy and minimize adverse reactions.

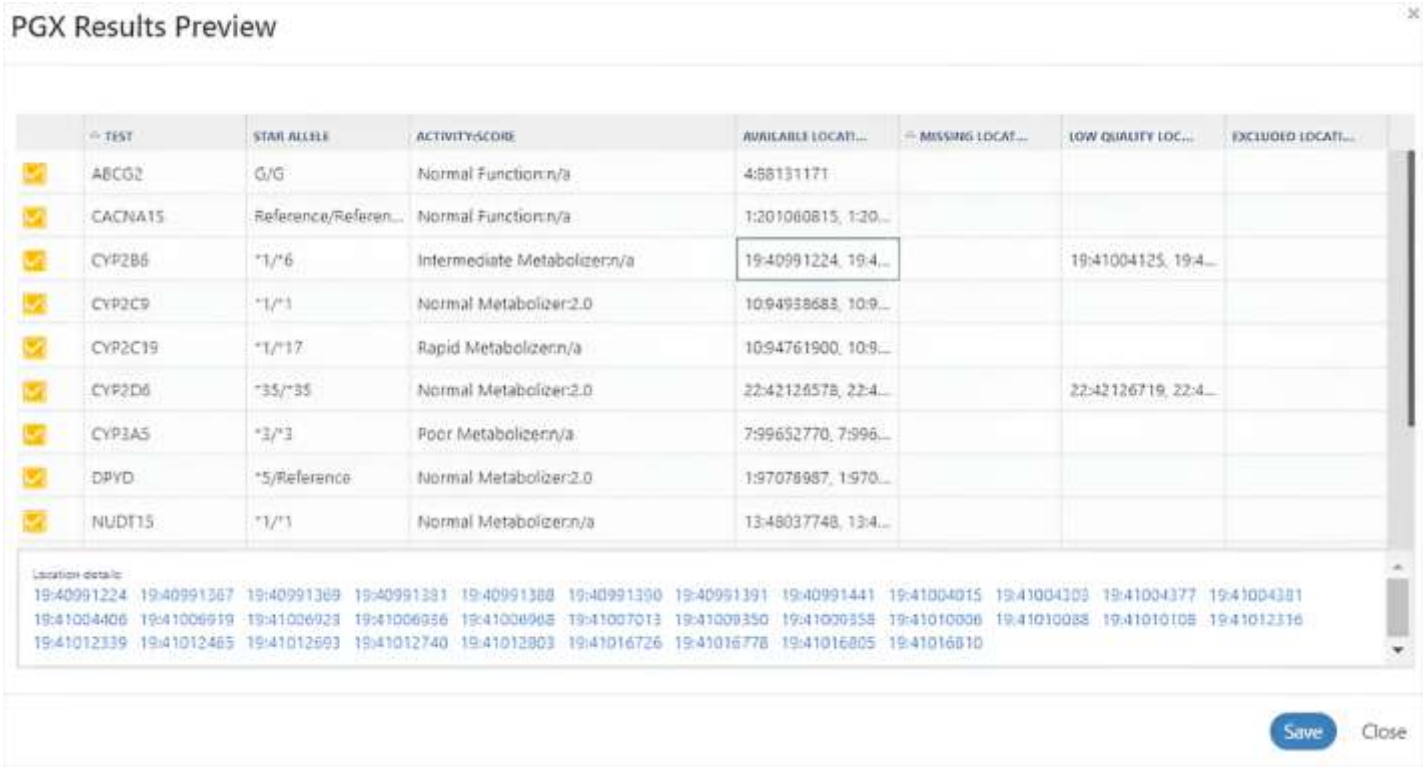


Figure 11: PGx Result Preview

Repeat Expansion Analysis

Repeat expansions are genetic mutations where DNA sequences are repeated beyond normal thresholds, leading to disorders such as Huntington's Disease, Fragile X Syndrome, Myotonic Dystrophy, Spinocerebellar Ataxias (SCAs), and Frontotemporal Dementia (FTD). The severity of these conditions often correlates with the length of the repeat expansion.

Key Features of Repeat Expansion Analysis:

Integrated Analysis:

- Displays repeat expansions in a two-column format: observed copies (allele 1 and allele 2) and repeat unit.

Color-Coded Results for Clinical Implications:

- Red: Pathogenic range (requires medical attention).
- Orange: Premutation range (clinically significant).
- Green: Outside mutation/premutation range (normal).
- No Color: Insufficient clinical data.

Hovering over repeats provides wildtype allele copy details.

Enhanced Filtering Options:

Users can filter repeats by color or clinical impact, focusing analyses efficiently.

In-House Data Presentation:

- Displays counts for homozygotes, heterozygotes, and hemizygotes within an in-house database.
- Hovering over sample counts reveals repeat allele distributions.
- Clickable sample count opens a detailed table and histogram for repeat frequency visualization.

This analysis enables users to identify, interpret, and prioritize clinically significant repeat expansions, ensuring streamlined workflows for genetic disorder evaluations.

Figure 12: Repeat Expansion analysis showing full mutation in HOXA13 gene

Automation and APIs

Geneyx Analysis API feature allows users to connect with existing LIMS/EHR systems and upload data using command line workflows.

1. Integration with primary and secondary analysis pipelines to allow automated import and annotation of VCF files. This works with single sample and batch workflows.
2. Integration with laboratory information management systems (LIMS) or electronic health record (EHR) systems enables the creation of automated analyses of patient clinical information and the streamlining of reports from Geneyx Analysis to the LIMS/EHR.
3. Auto-Reporting for SNVs/CNV/SVs to generate TSV files based on your customized filtering.

Other API features provided by Geneyx Analysis can be found here: https://github.com/geneyx/geneyx.analysis.api/.

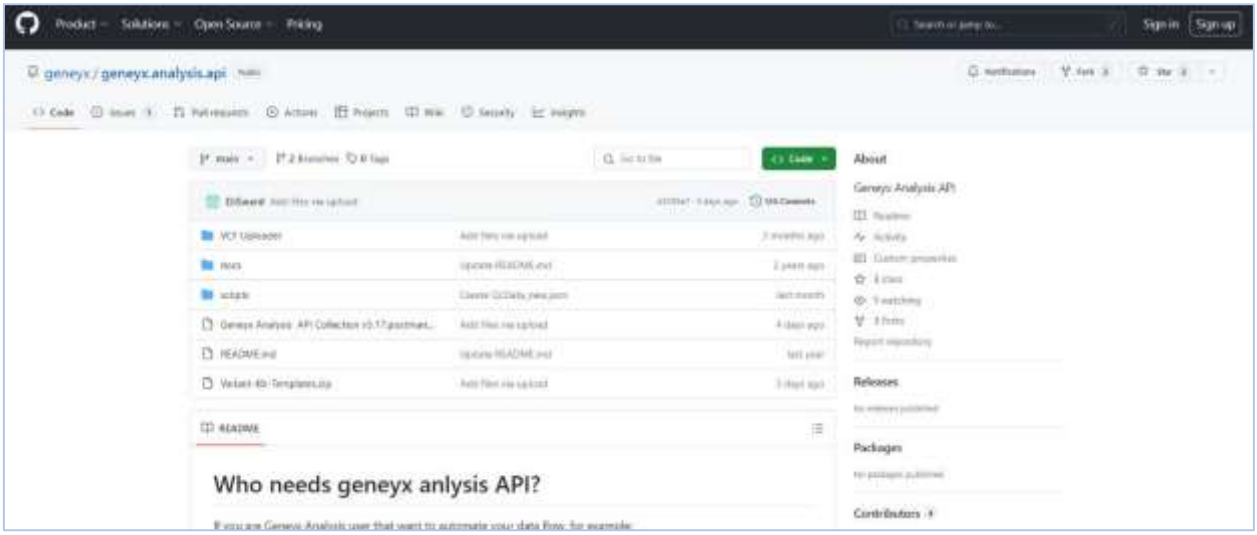


Figure 13: APIs available on our GitHub

Variant Knowledgebase

Geneyx's classified variants are organized within a knowledgebase, uniquely tied to each account. Users can now harness the capability to export and import this valuable data. An exemplary use case involves lifting over annotated variants to an alternate genome assembly, leveraging the transformed data as an allele frequency backlog for augmented evidence across different assemblies

Figure 14: Variant Knowledgebase of interpreted variants

Variant Browser

The Variant Browser in Geneyx Analysis provides a robust platform for querying and exploring Single Nucleotide Variants (SNVs) and Copy Number/Structural Variants (CNVs/SVs) across large-scale datasets. Designed to support up to 200,000 samples simultaneously, the Variant Browser ensures efficient and scalable genomic data analysis.

Key Features:

Flexible Filtering:

- Users can define filters based on criteria such as genomic coordinates, variant type, allele frequency, and functional impact.
- Enables targeted exploration of specific variants or genomic regions.

Streamlined Workflow:

- A Query Builder allows users to configure filters, select genome builds, and define analysis logic.
- Results are displayed in an interactive table for further exploration or export.

Comprehensive Insights:

- Facilitates sophisticated analyses to identify disease-causing variants and understand the genetic basis of complex disorders.

The Variant Browser empowers researchers and clinicians to navigate extensive datasets, enabling precise identification of clinically significant variants and facilitating impactful genomic research.

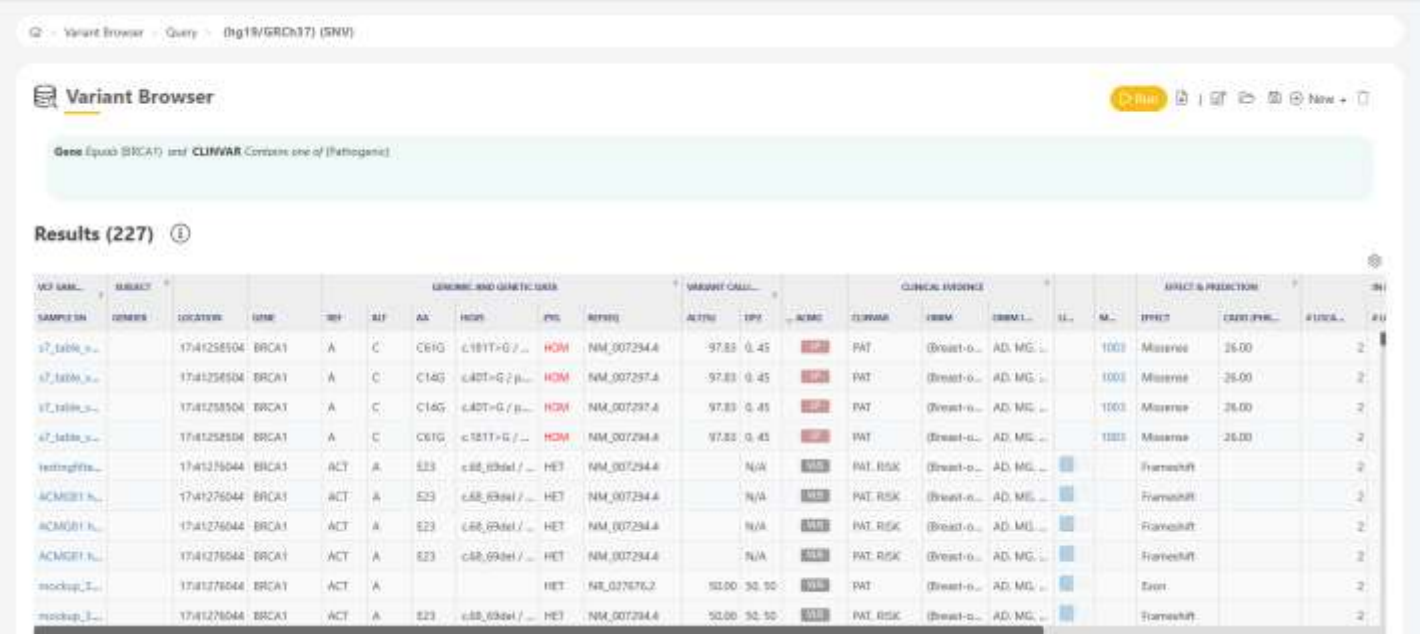


Figure 15: Variant Browser

VUS Monitor

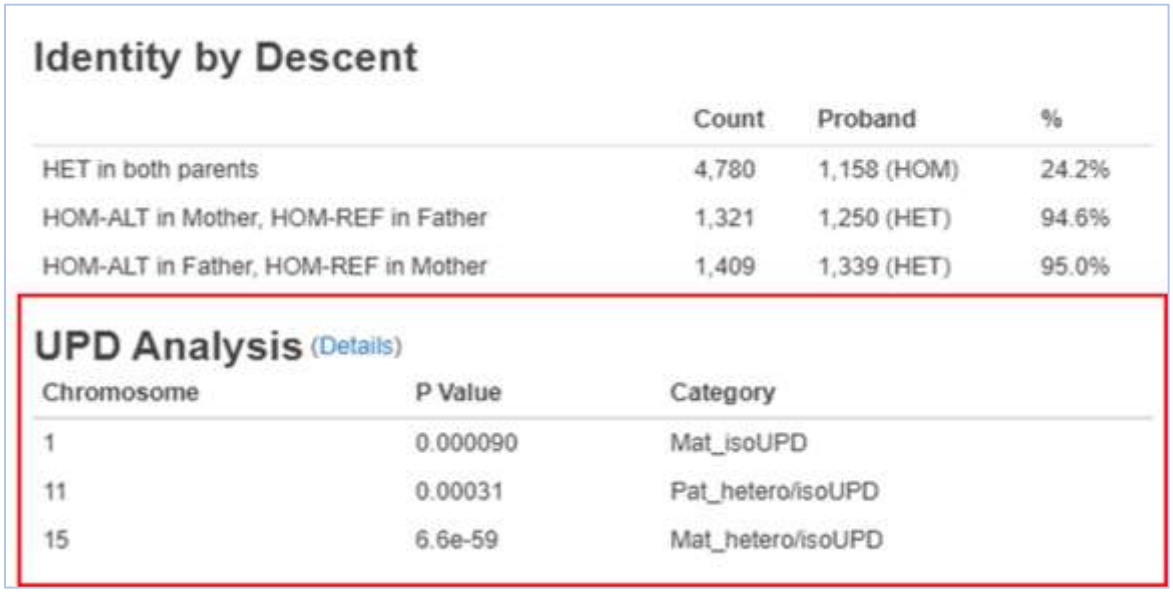
Keeping track of variant classification changes can be a daunting task. To simplify this process, we have implemented an auto-notification feature that alerts you whenever variants undergo

classification changes according to ClinVar. This proactive notification provides comprehensive details, including a direct link to the analysis, enabling you to perform retrospective analyses with ease and precision.

Figure 16: VUS Monitor to track variants of unknown significance through annotation updates

Uniparental Disomy

UPD is automatically calculated in *Geneyx Analysis* when running trio exomes and genomes by comparing each variant of the proband to the parents. Notifications are provided if significant results are identified.



Identity by Descent

	Count	Proband	%
HET in both parents	4,780	1,158 (HOM)	24.2%
HOM-ALT in Mother, HOM-REF in Father	1,321	1,250 (HET)	94.6%
HOM-ALT in Father, HOM-REF in Mother	1,409	1,339 (HET)	95.0%

UPD Analysis (Details)

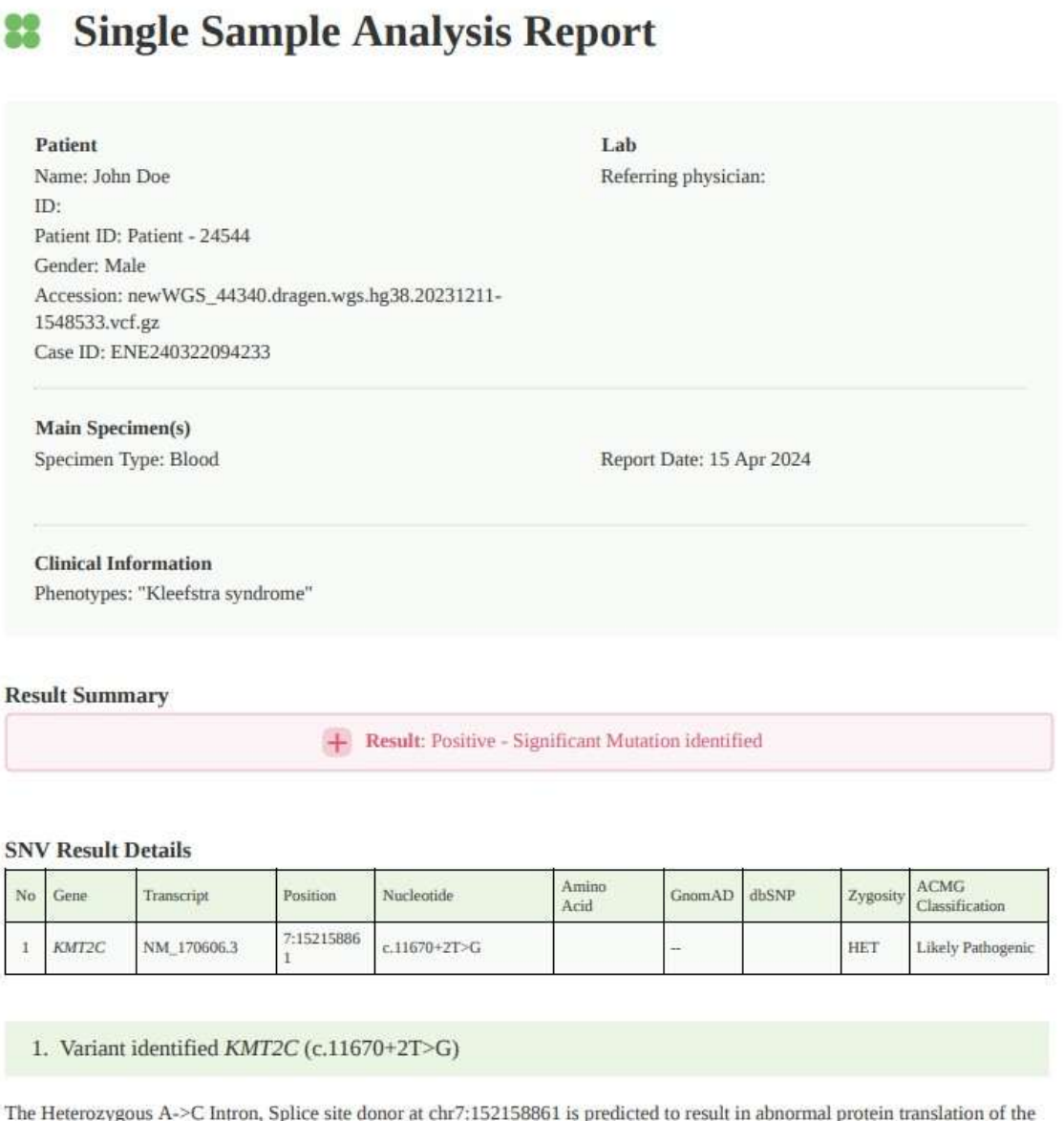
Chromosome	P Value	Category
1	0.000090	Mat_isoUPD
11	0.00031	Pat_hetero/isoUPD
15	6.6e-59	Mat_hetero/isoUPD

Figure 17: UPD Analysis is calculated for all trio workflows

Comprehensive and Customizable Reports

A comprehensive and flexible reporting system of modular tools and features that can be combined and customized to create personalized analysis pipelines. These tools cover a wide range of applications, including variant calling, annotation, filtration, pathway analysis, and interpretation.

Users can select the relevant tools for each step of the analysis, configuring them to meet their specific requirements. The comprehensive clinical reports can be customized according to the organization's logo and formatting requirements to incorporate static and dynamic information. The reports are also offered in multiple languages, preventing the need to translate generic information.



Single Sample Analysis Report

Patient
Name: John Doe
ID:
Patient ID: Patient - 24544
Gender: Male
Accession: newWGS_44340.dragen.wgs.hg38.20231211-1548533.vcf.gz
Case ID: ENE240322094233

Lab
Referring physician:

Main Specimen(s)
Specimen Type: Blood

Report Date: 15 Apr 2024

Clinical Information
Phenotypes: "Kleefstra syndrome"

Result Summary

Result: Positive - Significant Mutation identified

SNV Result Details

No	Gene	Transcript	Position	Nucleotide	Amino Acid	GnomAD	dbSNP	Zygosity	ACMG Classification
1	*KMT2C*	NM_170606.3	7:152158861	c.11670+2T>G		–		HET	Likely Pathogenic

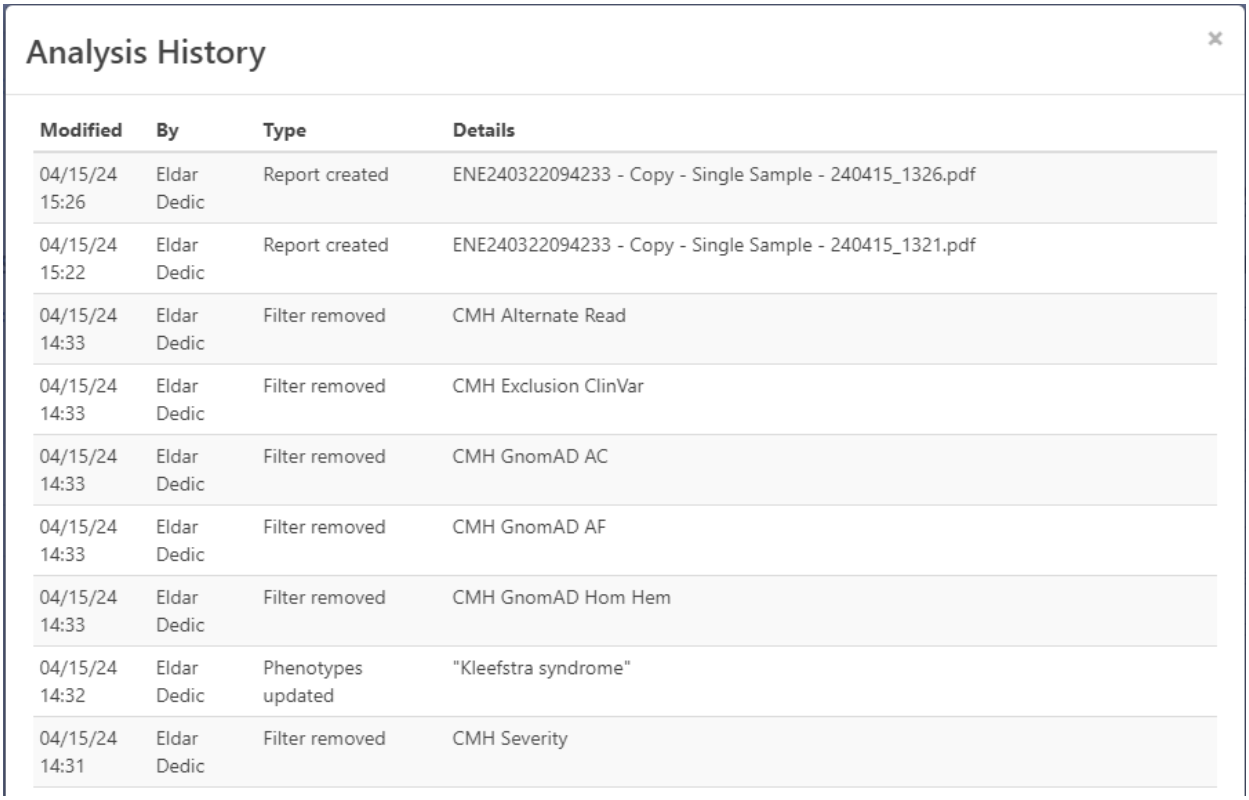
1. Variant identified *KMT2C* (c.11670+2T>G)

The Heterozygous A->C Intron, Splice site donor at chr7:152158861 is predicted to result in abnormal protein translation of the

Figure 18: Reporting outputs in Geneyx are fully customizable

Account Management and Audit Trails

Account management capabilities enable the group leader to set permissions for individual users to ensure complete control of all processes and steps performed by individual users to streamline workflow and manage data applications. All processes performed in the application are stored and can be reviewed, offering a comprehensive audit trail for laboratories.



Modified	By	Type	Details
04/15/24 15:26	Eldar Dedic	Report created	ENE240322094233 - Copy - Single Sample - 240415_1326.pdf
04/15/24 15:22	Eldar Dedic	Report created	ENE240322094233 - Copy - Single Sample - 240415_1321.pdf
04/15/24 14:33	Eldar Dedic	Filter removed	CMH Alternate Read
04/15/24 14:33	Eldar Dedic	Filter removed	CMH Exclusion ClinVar
04/15/24 14:33	Eldar Dedic	Filter removed	CMH GnomAD AC
04/15/24 14:33	Eldar Dedic	Filter removed	CMH GnomAD AF
04/15/24 14:33	Eldar Dedic	Filter removed	CMH GnomAD Hom Hem
04/15/24 14:32	Eldar Dedic	Phenotypes updated	"Kleefstra syndrome"
04/15/24 14:31	Eldar Dedic	Filter removed	CMH Severity

Figure 19: Audit trails output for all actions performed in an analysis

Long-Read sequencing

Geneyx Analysis provides a comprehensive tertiary analysis platform for the efficient processing of Oxford Nanopore Technologies and PacBio Long-Read sequencing data. We've introduced innovative features that facilitate rapid data ingestion and annotation of long-read data, streamlining the entire process of variant interpretation and reporting.

- Phasing analysis: Our specialized tools, including phasing and methylation analysis simplify the bioinformatics process, ensuring accurate variant annotation and classification.

- Comprehensive view of overlapping CNVs and SNVs for comphound heterozygous variant detection
- SMART Filtering to quickly reduce background noise associated with long read sequencing data
- Repeat Expansion analysis aimed to identify deleterious repeats and potential gene silencing events
- Ability to import all variant file outputs

Figure 20: Allele phasing integration where variant is observed in trans

Data Protection Compliance

- HIPAA (Health Insurance Portability and Accountability Act)
- GDPR (General Data Protection Regulation) compliant.
- ISO 13485 is designed for medical equipment and software manufacturers to ensure an internationally recognized standard for quality management.)
- ISO 27001 (**standard for information security management systems (ISMS))**
- All private patient information is encrypted.

Overview Summary

Geneyx Analysis is a powerful secondary and tertiary analysis platform for annotation, analysis, and prioritization of coding and non-coding genomic variants using short and long-read sequencing analysis. This platform provides access to an extensive knowledge base of genomic annotations and offers intuitive and flexible configurations. The ease of use and transparency allows quick adaptation and implementation into clinical workflows and innovative features that simplify and accelerate variant interpretation. *Geneyx Analysis* is an end-to-end solution demonstrating a unique paradigm that can increase diagnostic yield and reduce turnaround time for all users.

Geneyx Analysis is available for trial at www.geneyx.com

Publications:

Genome analysis and knowledge-driven variant interpretation with TGex. PMID: 31888639

Clinical evaluation of long-read sequencing-based episignature detection in developmental disorders PMID: 39789644

Somatic and germline mutations in endometrial cancer PMID: 39296440

Complex rearrangement in TBC1D4 in an individual with diabetes due to severe insulin resistance syndrome PMID: 38086948.