Supplementary Information

Supplementary Methods

varianTalker execution and configuration

The workflow supports resumable execution through Nextflow's built-in caching mechanism. When a run is interrupted or repeated with unchanged inputs and parameters, completed processes are reused and execution resumes from the last non-completed step. This is particularly useful for large cohorts and computationally intensive annotation stages.

The workflow is invoked as follows:

```
nextflow run path_to/main.nf \
  -c nextflow.config \
  -profile singularity \
  --input samplesheet.csv \
  --outdir outdir
```

Runtime parameters are formally described in nextflow_schema.json in the repository, which provides a structured definition of accepted parameters and supports reproducible configuration of analyses.

Sample sheet tumour tissue vocabulary

Somatic entries in the sample sheet require a tumour tissue designation drawn from the following controlled vocabulary of 33 terms: Adrenal_Gland, Bile_Duct, Bladder, Blood, Bone, Bone_Marrow, Brain, Breast, Cancer_all, Cervix, Colorectal, Esophagus, Eye, Head_and_Neck, Inflammatory, Intrahepatic, Kidney, Liver, Lung, Lymph_Nodes, Nervous_System, Other, Ovary, Pancreas, Pleura, Prostate, Skin, Soft_Tissue, Stomach, Testis, Thymus, Thyroid, Uterus. Germline entries do not require tissue annotation.

Extended description of the integrated annotation resources

Funcotator can generate either MAF or VCF output containing annotations from all matching data sources. It supports both somatic and germline contexts and reports gene, transcript, coding consequence, protein consequence and source-specific annotation fields, keeping the database bundles separate. Reporting consequences on MANE Select transcripts reduces transcript-dependent discrepancies in variant consequence reporting and improves consistency across tools and sequencing platforms, since MANE provides matched RefSeq and Ensembl/GENCODE transcript pairs that are identical across the transcript sequence and can therefore be used consistently for clinical reporting.

Combining CancerVar and InterVar ensures that somatic and germline variants are evaluated under their appropriate guideline framework rather than being interpreted using a single generic pathogenicity model. CIViC evidence items are assigned a level reflecting the robustness of the underlying evidence, from A (a validated association) to E (an inferential association), providing structured, context-specific evidence linking a molecular alteration to a clinical or biological interpretation in cancer.

AlphaMissense output consists of a continuous pathogenicity score from 0 to 1, with higher values indicating a higher predicted probability of pathogenicity, together with categorical labels (likely benign, ambiguous, likely pathogenic). Predictions are restricted to missense variants.

ClinVar annotation includes categories ranging from pathogenic to benign together with review status information. The ClinVar review system reflects the level of supporting evidence behind an assertion rather than the severity or clinical impact of the variant itself.

RENOVo is trained on clinically classified variants and leverages publicly available annotation features to predict whether a variant is more likely to be pathogenic or benign, reporting a Pathogenicity Likelihood Score (PLS). Within varianTalker it contributes to the consensus prioritisation framework as an additional computational evidence layer.

ESCAT Tier I alterations are considered ready for routine clinical implementation, whereas lower tiers represent, in decreasing order, reduced levels of clinical evidence, evidence derived from other tumour types, preclinical support, and hypothetical actionability. The ESCAT layer can be compared with CancerVar, CIViC and other therapeutic annotations.

PostgreSQL database schema and the update_db module

The database is organised in two levels. First, a central registry keeps track of each analytical configuration used in a run, including the set of annotation tools, reference databases and their versions, plus the genome build, sequencing technology, somatic/germline mode and tumour tissue. Each configuration is identified by a hash and mapped to its own schema. A stored annotation is therefore reused only when the new sample is processed with exactly the same configuration; if a database is updated or the genome build changes, a new schema is created and outdated annotations are never reused by mistake.

Second, within each schema a single table stores the variants. Each record holds the variant coordinates, the full annotation as a JSON document, and a few key fields (gene symbol, CancerVar, InterVar, ClinVar and COSMIC) kept as separate columns for quick querying. Variants are uniquely identified by their chromosome, position and a hash of the reference and alternate alleles, which keeps lookups fast even for long insertions or deletions. New records are added in batches and updated only when the annotation has actually changed, so repeated runs do not create duplicates.

Functional annotation relies on external reference databases that are updated regularly, and these updates can change how a variant is interpreted over time; a variant first reported as a VUS may later be reclassified as pathogenic. The optional update_db module maintains the Funcotator and ANNOVAR databases on demand. It is invoked by the user and is never triggered by an analysis run, so the annotation resources used by a given run remain fixed unless the user updates them explicitly. When invoked, the module compares each local database with the latest available release. Resources with an active release cycle are downloaded and updated; a small number of resources have not been revised since their original release and are therefore only version-checked, with a warning raised if an update ever becomes available. Both the hg19 and hg38 genome builds are supported. Some ANNOVAR databases are also used by CancerVar and InterVar; when these change, the module regenerates the matching CancerVar and InterVar scripts and configuration files so that clinical interpretation stays consistent with the underlying data. The module is provided as a container image (Docker/Singularity) for reproducible use on HPC systems, and the source code is available at https://github.com/zhanyinx/variantalker.

MAFigate retention rules and interface

Independently of the technical, biological and clinical criteria, variants reaching a strong clinical call are retained by default: on the somatic arm a CancerVar tier I or II assignment, a ClinVar pathogenic or likely pathogenic classification, or CIViC evidence level A or B; on the germline arm an InterVar or ClinVar pathogenic or likely pathogenic classification. This rule can be switched off, and it overrides every other criterion except the population allele frequency cut-off, from which ClinVar pathogenic and likely pathogenic variants are in turn exempt.

Results are presented as an interactive, sortable table, a summary dashboard over the loaded file, and a per-variant detail panel. The dashboard reports variant classification, mutation type spectrum, chromosomal and allele fraction distributions, the most frequently mutated genes, and clinical significance for passing against non-passing variants. The detail panel shows clinical and therapeutic annotations, population frequencies, and direct links to external resources (gnomAD, ClinVar, dbSNP and the UCSC Genome Browser). For germline variants the panel adds a colour-coded rendering of the full InterVar ACMG/AMP classification, showing the significance call together with the pathogenic and benign criteria that were met, grouped and graded by evidence strength; for somatic variants it adds the corresponding rendering of CancerVar's twelve AMP/ASCO/CAP evidence categories and the tier they support.

Both the passing and non-passing variant sets, together with any manual notes and user-defined annotation columns, can be exported as CSV for downstream analysis or molecular tumour board discussion, and a text report of the parameters each run was made with can be downloaded alongside them. Detailed guidelines for operating the application and interpreting its output are provided as Supplementary File 1.

varianTalker output files

varianTalker generates a complete annotated MAF file containing all processed variants. The details of the content within the MAF file can be found at Supplementary Table S3. For germline samples, an HTML report is generated for the PharmGKB annotations, providing an accessible summary of pharmacogenomic findings. The standardized input VCF file is also retained in the output folder so that users can reuse it if needed. This output structure preserves both the full annotation record and the filtering decision applied to each variant, allowing users to recover excluded findings without re-running the annotation workflow.

SAREK-based WES preprocessing and variant calling

Raw WES FASTQ files from the internal cohort, comprising 45 tumour samples and 21 matched normal samples, were processed using nf-core/sarek v3.7.1 with Nextflow v25.10.4 and containerized software dependencies. All four available sequencing lanes were included for each sample. Analyses were performed in WES mode against the GATK GRCh38 reference bundle (GATK.GRCh38) provided through the nf-core/iGenomes configuration. Whole exome libraries were prepared with the Human Comprehensive Exome kit (Twist Bioscience). varianTalker reads the target interval file from a pipeline parameter, so any capture design can be substituted without modification of the workflow. Unless otherwise specified, pipeline and tool parameters were left at the nf-core/sarek v3.7.1 defaults.

For somatic variant calling, tumour and matched-normal samples were processed with GATK Mutect2 through nf-core/sarek. Because multiple tumour samples were available for some patients, tumour-normal comparisons were performed independently, resulting in 45 tumour-normal comparisons. The same normal sample was therefore used in one to three independent comparisons

depending on the number of tumour samples available for the corresponding patient. Samples were distributed across three Sarek input sample sheets to maintain independent tumour-normal comparisons. Germline variant calling was performed separately on the 21 normal samples using GATK HaplotypeCaller through nf-core/sarek.

DRAGEN-based WES processing and variant calling

As an independent variant-calling approach, WES data from the internal cohort were additionally processed using Illumina DRAGEN v4.4.4 against the GRCh38 human reference genome (GRCh38_full_analysis_set_plus_decoy_hla), with variant calling restricted to the target regions defined by Twist Bioscience exome capture technology. Reads were mapped, sorted, and duplicate-marked (--enable-duplicate-marking true), and alignments were written in CRAM format.

Normal samples were processed in normal-only (germline) mode against the DRAGEN pangenome hash table (hg38-alt_masked.cnv.graph.hla.methyl_cg.rna-11-r5.0-1), rather than the linear reference. A SNP panel of normals built from the 1000 Genomes hg38 call set was supplied for small-variant filtering (--panel-of-normals), and variant deduplication was disabled (--enable-variant-deduplication false). Somatic analyses were performed on matched tumour-normal pairs and comprised two DRAGEN invocations: a germline run on the normal sample alone, configured as described above, and a paired somatic run against the linear hash table (hg38-alt_masked.cnv.hla.methyl_cg.methylated_combined.rna-11-r5.0-1) using both tumour and normal FASTQ lists. Somatic small-variant calling applied systematic-noise filtering using the Illumina WES hg38 systematic-noise baseline v2.0.0 (--vc-systematic-noise) together with the 1000 Genomes SNP panel of normals (--panel-of-normals), and the variant-allele-frequency detection threshold was lowered to 0.01 (--vc-target-vaf 0.01). All parameters not stated above were left at DRAGEN defaults.

Ion Torrent sequencing and variant calling

A second internal cohort comprising 178 samples (89 matched tumour-normal pairs) was analysed using the Ion Torrent sequencing platform. Sequencing was performed on an S5 XL instrument using a targeted panel comprising approximately 650 genes. Sequencing reads were mapped to the hg19 human reference genome using TMAP (Torrent Mapping Alignment Program), and variants were called using Ion Reporter software (version 5.22.0.2). Variant calling was performed independently for normal and tumour samples using single mode custom AmpliSeq workflows derived from *Oncomine Comprehensive Assay*. No filters were applied at this level.

ClinVar VCF preprocessing

To generate a reference variant set compatible with the downstream annotation workflow, the ClinVar GRCh38 VCF release dated December 28, 2025 (`clinvar_20251228.vcf.gz`) was obtained from the NCBI ClinVar repository through the GATK Funcotator data sources and preprocessed to mimic the structure of a sample-derived VCF. Chromosome identifiers were standardized to the `chrN` convention, including conversion of mitochondrial chromosome identifiers to `chrM`, and records mapping to contigs absent from the GRCh38 reference used by the annotation workflow were excluded. The VCF header was reconstructed using contig definitions from the GRCh38 reference FASTA index, and all retained records were assigned a `PASS` filter status.

Because ClinVar records do not contain the sample-level sequencing fields expected by the downstream annotation workflow, synthetic genotype and read-support information was added solely to ensure compatibility with its input and filtering requirements. Variant records were assigned a heterozygous genotype (`GT=0/1`), a total depth of 200 reads (`DP=200`), balanced reference and

alternative allele depths (AD=100,100), and an allele fraction of 0.5 (AF=0.5). Records without an alternative allele were assigned GT=0/0, DP=200, AD=200,0, and AF=0. These values were arbitrary synthetic placeholders selected to avoid triggering downstream depth- or allele-fraction-based filters and do not represent experimental measurements. Duplicate variant records were removed.

During workflow development, 46 long-deletion records were identified as incompatible with the Funcotator-based annotation step and consistently caused annotation failure (Supplementary Table S4); these records were therefore excluded from the ClinVar test set prior to downstream analysis. In addition, complex variants outside the scope of the present analysis were removed. Complex variants were operationally defined after trimming identical leading bases from the REF and ALT alleles as records for which both residual alleles remained non-empty and at least one residual allele was longer than one nucleotide. This retained simple SNVs and insertions/deletions while excluding composite substitutions/indel representations. The resulting VCF was BGZF-compressed and indexed using bcftools v1.21 and used as the ClinVar input for subsequent annotation analyses.

VCF preprocessing for PCGR compatibility

Variant call format (VCF) files generated by the different variant-calling workflows, together with the preprocessed ClinVar variant set, underwent an additional preprocessing step external to Personal Cancer Genome Reporter (PCGR) to generate input files compatible with subsequent PCGR analysis. Because sequencing depth and variant allele fraction were encoded in sample-level FORMAT fields in the input VCFs, FORMAT/DP and FORMAT/AF values were extracted and copied to newly defined INFO fields (CUSTOMDP and CUSTOMAF, respectively), allowing these quantities to be explicitly provided to PCGR through its tumour depth and tumour allele fraction input parameters. For multi-sample tumour–normal VCFs, DP and AF were specifically extracted from the tumour sample, whereas the available sample column was used for single-sample VCFs. Existing CUSTOMDP and CUSTOMAF annotations, when present, were removed and regenerated from the selected sample to ensure consistent assignment across input files. Records lacking an alternative allele were excluded during preprocessing. For the ClinVar-derived dataset, CUSTOMDP and CUSTOMAF corresponded to the synthetic depth and allele-fraction values introduced during the preceding ClinVar preprocessing step rather than experimentally measured sequencing quantities. Finally, the resulting VCF files were BGZF-compressed and tabix-indexed using bcftools v1.21, generating the input files subsequently supplied to PCGR.

Generation of ClinVar subsets for computational benchmarking

The preprocessed ClinVar variant set was subsampled to generate synthetic VCF datasets for computational performance benchmarking. To assess pipeline scalability as a function of the number of input variants, VCFs containing 5,000, 10,000, 50,000, 100,000, 500,000, and 1,000,000 variants were generated by deterministic random sampling. For each target size, an initial pool corresponding to ten times the requested number of variants was randomly sampled, records lacking an alternative allele were excluded, and records sharing the same genomic position were deduplicated. The remaining records were randomly resampled to the requested dataset size and sorted by genomic coordinate. Five deterministic random seeds (12, 33, 55, 120, and 200) were initially evaluated. Despite the preceding ClinVar preprocessing, a subset of ClinVar records caused PCGR execution failures without allowing the problematic records to be unambiguously identified from the reported errors. Multiple seeds were therefore generated to obtain a complete series of technically processable benchmark VCFs. Seed 33 was retained because it was the first seed for which all six VCF sizes could be successfully processed by PCGR. Only the six VCFs generated using this seed were used for the subsequent single-VCF scalability benchmark.

A second benchmark dataset was generated to evaluate pipeline performance when processing a large number of independent input samples. A total of 1,000 synthetic VCFs, each containing 1,000 variants, were independently sampled from the same preprocessed ClinVar dataset. For each VCF, a deterministic sample-specific seed derived from a common seed of 15 and the sample index was used. An initial pool of 5,000 variants was sampled, followed by removal of records lacking alternative alleles and deduplication by genomic position; 1,000 records were then randomly selected and sorted by genomic coordinate. Each resulting VCF was assigned a unique synthetic sample identifier. Unlike the single-VCF scalability dataset, these VCFs were generated specifically to benchmark the annotation pipeline and were not required to undergo PCGR analysis. All benchmark VCFs were BGZF-compressed and tabix-indexed using bcftools v1.21.

Variant annotation with varianTalker

For the annotation concordance benchmark, the same datasets analysed with OpenCRAVAT, PCGR and VEP were processed with varianTalker, including all VCFs from the internal cohorts and the six ClinVar-derived subsets containing increasing numbers of variants. Analyses were performed using the native genome assembly of each dataset, with GRCh38 used for the ClinVar-, Sarek- and DRAGEN-derived VCFs and hg19/GRCh37 for the Ion Torrent-derived VCFs, without coordinate liftover. The internal cohort VCFs were processed according to their corresponding upstream sequencing and variant-calling configurations: Illumina WES with either Sarek or DRAGEN for the ALL cohort and Ion Torrent targeted sequencing for the second internal cohort. Dataset-specific parameters describing the upstream variant-calling pipeline, sequencing technology and exome status were supplied to varianTalker to configure the appropriate input normalization strategy. ClinVar-derived datasets were processed as non-exome GRCh38 inputs. Because these datasets did not originate from sequencing experiments, the sequencing technology and upstream variant-calling parameters assigned to them were used exclusively to configure input normalization and did not represent an actual experimental workflow. CIViC and PharmGKB annotations were disabled because these resources were not included in the comparative annotation assessment. The optional PostgreSQL-backed annotation retrieval mode was disabled for the concordance benchmark, ensuring that all variants underwent the standard varianTalker annotation workflow. The resulting annotated outputs were retained for subsequent cross-tool comparison with OpenCRAVAT, PCGR and VEP.

Computational performance with and without PostgreSQL-backed annotation retrieval was evaluated separately using the ClinVar-derived benchmark datasets. These experiments included the VCF series containing increasing numbers of variants and the dataset comprising 1,000 independently generated VCFs containing 1,000 variants each; multi-sample experiments were processed as single cohort-level varianTalker executions using sample sheets containing the corresponding number of input samples. For the standard annotation condition, varianTalker was executed without PostgreSQL-backed retrieval, requiring variants to undergo the complete annotation workflow. For the database-backed condition, a PostgreSQL database was pre-populated with the annotations required for the corresponding benchmark datasets using an additional database-enabled execution that was excluded from performance measurements. The same benchmark inputs were subsequently processed with PostgreSQL-backed retrieval enabled, allowing previously computed annotations to be retrieved from the database rather than recomputed. Thus, the performance comparison specifically evaluated standard de novo annotation against retrieval of pre-existing annotations and did not evaluate the time required to populate the database. Each measured benchmark condition was independently repeated five times to reduce the influence of transient fluctuations in computational performance. Execution metrics were collected from Nextflow trace, timeline and execution-report outputs for subsequent computational performance analysis.

Computational performance analysis

Computational performance was evaluated from the execution metadata generated by Nextflow for each varianTalker benchmark run, including the execution trace, timeline and execution report. Two scaling experiments were considered: increasing variant load using single ClinVar-derived VCFs containing 5,000, 10,000, 50,000, 100,000, 500,000, 1,000,000 variants and the complete 4,115,028 variants clinvar vcf and increasing cohort size using 10, 20, 30, 50, 80, 100, 200, 300, 500, 800 and 1,000 independently generated samples containing 1,000 variants each. Standard annotation and PostgreSQL-backed retrieval were evaluated separately, using the most stable five independent Nextflow executions available for each benchmark condition.

Runtime metrics were derived from the Nextflow timeline. To minimize variability caused by HPC scheduling and temporary resource unavailability, active workflow runtime was calculated as the union of all time intervals during which at least one Nextflow task was actively executing. Periods in which no task was running, including gaps caused by jobs remaining pending in the scheduler, were therefore excluded. The sum of individual task runtimes and the complete end-to-end execution span were additionally extracted for quality control. CPU consumption was obtained from the CPU-hours metric reported directly by the Nextflow execution report. Memory usage was summarized as the cumulative peak resident set size (RSS), calculated as the sum of the task-level peak RSS values reported in the Nextflow trace. Aggregate disk I/O was calculated by summing task-level read and write volumes separately. Memory and disk I/O quantities were converted from bytes to GiB.

For each dataset size and database condition, the available independent executions were summarized using the total range of values (ymin to ymax). Performance was evaluated as a function of either the number of input variants or the number of input samples, comparing standard de novo annotation with PostgreSQL-backed retrieval of previously computed annotations.

Annotation coverage and evidence integration

Annotation coverage across the resources integrated in varianTalker was evaluated using the internal cohort datasets, with germline and somatic variants analysed separately. Analyses were performed both on the complete variant sets and after restriction to non-silent variants. For each annotation resource, a variant was considered annotated when at least one of the corresponding output fields contained an informative value, defined as a value other than missing, an empty string, `"."`, or `"NA"`. For resources represented by multiple output fields, the presence of an informative value in any associated field was considered sufficient to define annotation coverage. ClinVar coverage was determined from the ClinVar-derived annotation fields, whereas InterVar or CancerVar, ReNOVo and ESCAT coverage was determined from their corresponding classification fields. Coverage of individual in silico pathogenicity predictors was similarly defined according to the availability of a prediction or score in the corresponding output field(s).

For comparative analyses, clinical annotations were harmonized into `pathogenic`, `benign`, `VUS`, `other`, and `unannotated` categories. ClinVar records containing multiple clinical significance terms were classified as pathogenic if at least one pathogenic term was present, as benign if at least one benign term was present in the absence of a pathogenic term, and as VUS when all informative terms corresponded to uncertain classifications; remaining annotated records were categorized as other. Single-category annotations from the other clinical interpretation tools were mapped to the same classification scheme. In silico predictors were evaluated for annotation availability rather than harmonized clinical classification. To assess the contribution of integrating complementary clinical evidence, classifications were subsequently combined cumulatively. When multiple resources

provided classifications for the same variant, the highest-priority available category was retained according to the hierarchy `pathogenic > VUS > benign > other > unannotated`.

Resolution of uncertain classifications and comparison with VarSome

The ability of complementary annotation resources to provide additional evidence for variants lacking a definitive classification was evaluated separately for germline and somatic variants. Clinical annotations were first harmonized into pathogenic, benign, VUS, other, or unannotated categories as described above, and duplicate records representing the same chromosome, genomic position, reference allele, and alternative allele were removed. For each pairwise comparison, variants classified as VUS or lacking an annotation in the source resource were identified and the corresponding classification assigned by the complementary resource was examined. In germline analyses, ClinVar classifications were compared with InterVar and the reciprocal comparison was also performed; in somatic analyses, analogous comparisons were performed between ClinVar and CancerVar. A variant was considered to have received a definitive complementary classification when a source VUS or unannotated variant was classified as either pathogenic or benign by the second resource. The proportions of source VUS and unannotated variants receiving each definitive classification were calculated separately.

A subset of these reclassifications was independently compared with VarSome classifications. All unique variants receiving a pathogenic classification from the complementary resource were selected, together with an equal-sized random sample of unique variants receiving a benign classification. Sampling was performed independently among variants originally classified as VUS and among variants lacking a source annotation, using a random seed of 123. The sampling procedure therefore retained all pathogenic reclassifications while providing a size-matched benign comparison set. Selected variants were manually queried in VarSome in July 2026. The structured **Germline Classification** reported by VarSome was recorded for germline variants, whereas the **Somatic Classification** was recorded for somatic variants. VarSome classifications were then compared with the corresponding InterVar, CancerVar, or ClinVar classifications using the appropriate germline or somatic classification framework.

Clinical filtering and expert review

The clinical prioritization strategy implemented in varianTalker was evaluated in 50 matched tumour–normal pairs of samples from the Ion Torrent internal cohort. Samples were processed using the standard varianTalker workflow and the germline- or somatic-specific filtering configurations described above. Variants satisfying the corresponding prioritization criteria were assigned to the PASS output, whereas excluded variants remained available in the NOPASS output. The effectiveness of the filtering strategy was evaluated by comparing the number and clinical composition of variants retained in PASS with the complete set of variants represented by PASS and NOPASS.

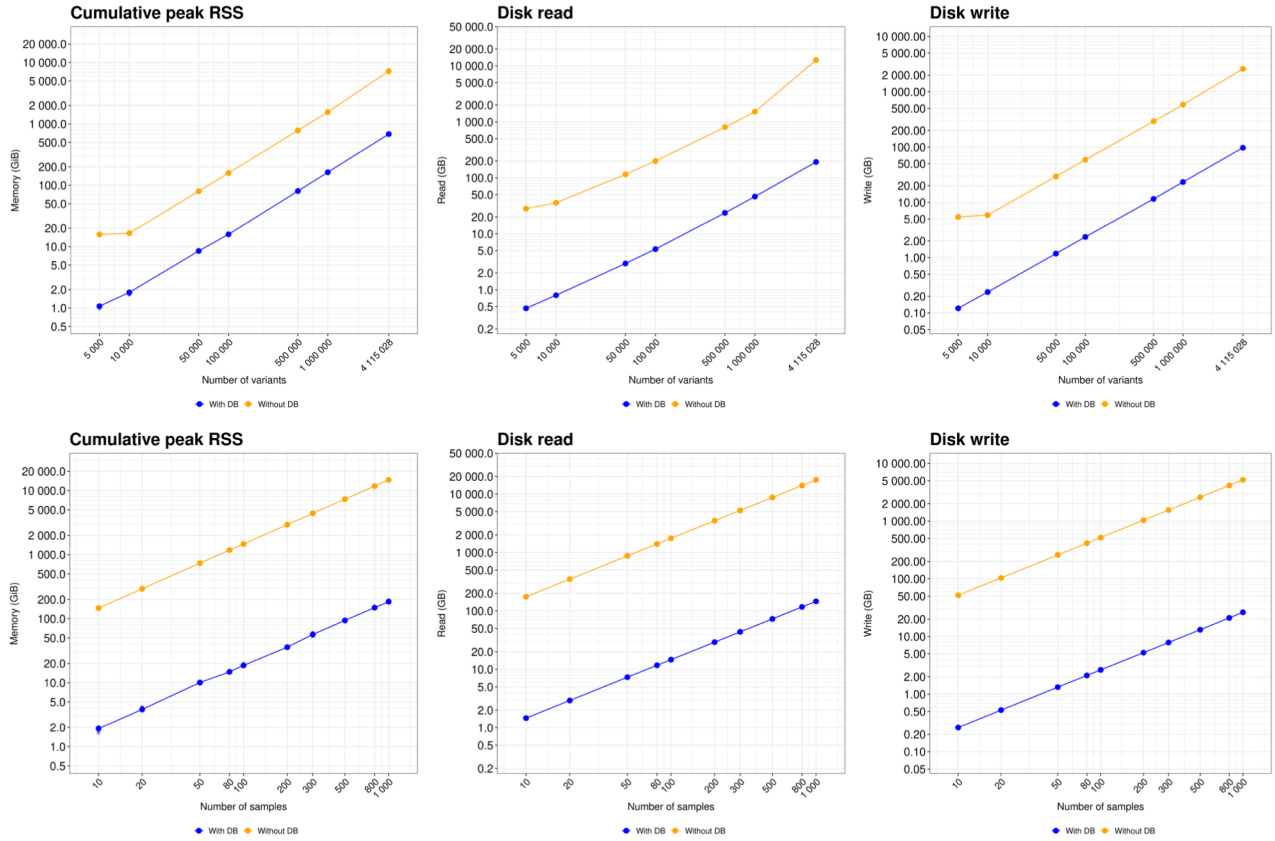
All PASS variants were individually reviewed by a clinical expert. Expert-assigned classifications were harmonized according to the priority hierarchy `P > LP > UV > LB > B`; when more than one assessment was recorded for a variant, the highest-priority classification was retained. Classifications were subsequently collapsed into three categories for summary analysis: pathogenic (`P`, including P and LP), uncertain (`UV`), and benign (`B`, including B and LB). NOPASS variants were not comprehensively reviewed individually. Instead, their classification distribution was derived computationally from the clinical annotations generated by the pipeline. ClinVar clinical significance terms were first simplified to pathogenic, likely pathogenic, uncertain, likely benign, or benign categories. For germline variants, the highest-priority available classification between ClinVar and

InterVar was retained, whereas for somatic variants the corresponding comparison was performed between ClinVar and CancerVar, using the hierarchy `P > LP > UV > LB > B`. The resulting classifications were subsequently collapsed into the same P, UV and B categories used for the expert-reviewed PASS variants. In addition, NOPASS variants carrying uncertain or conflicting clinical annotations were specifically inspected by the clinical expert to assess whether potentially pathogenic variants had been excluded by the prioritization procedure. The filtering assessment therefore evaluated both the reduction in the number of variants requiring primary manual review and the enrichment of the resulting PASS set for pathogenic and uncertain findings, while retaining all excluded variants in NOPASS for secondary or retrospective inspection.

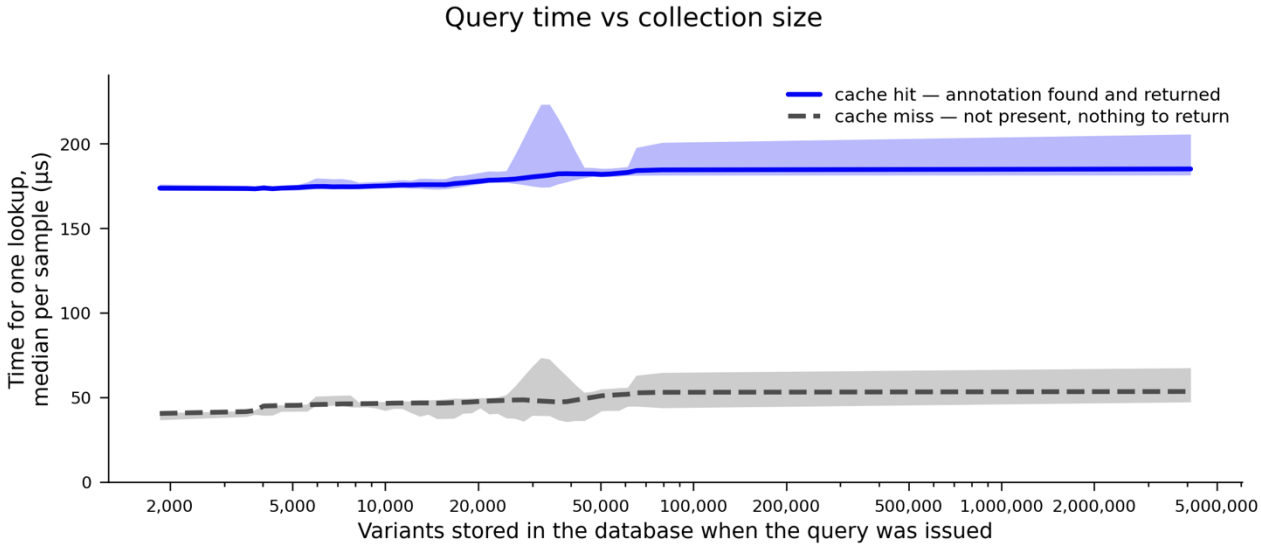
Declaration of generative AI use

During manuscript preparation, generative artificial intelligence tools were used solely for language editing and stylistic polishing. All scientific content, interpretations, and conclusions were developed by the authors, who reviewed and approved the final manuscript and take full responsibility for its content.

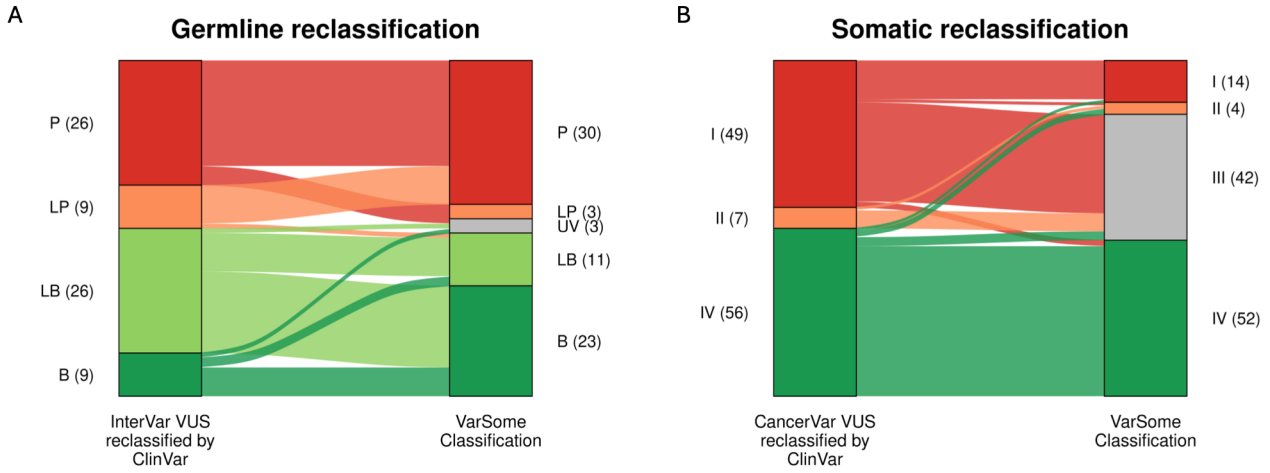
Supplementary Figures



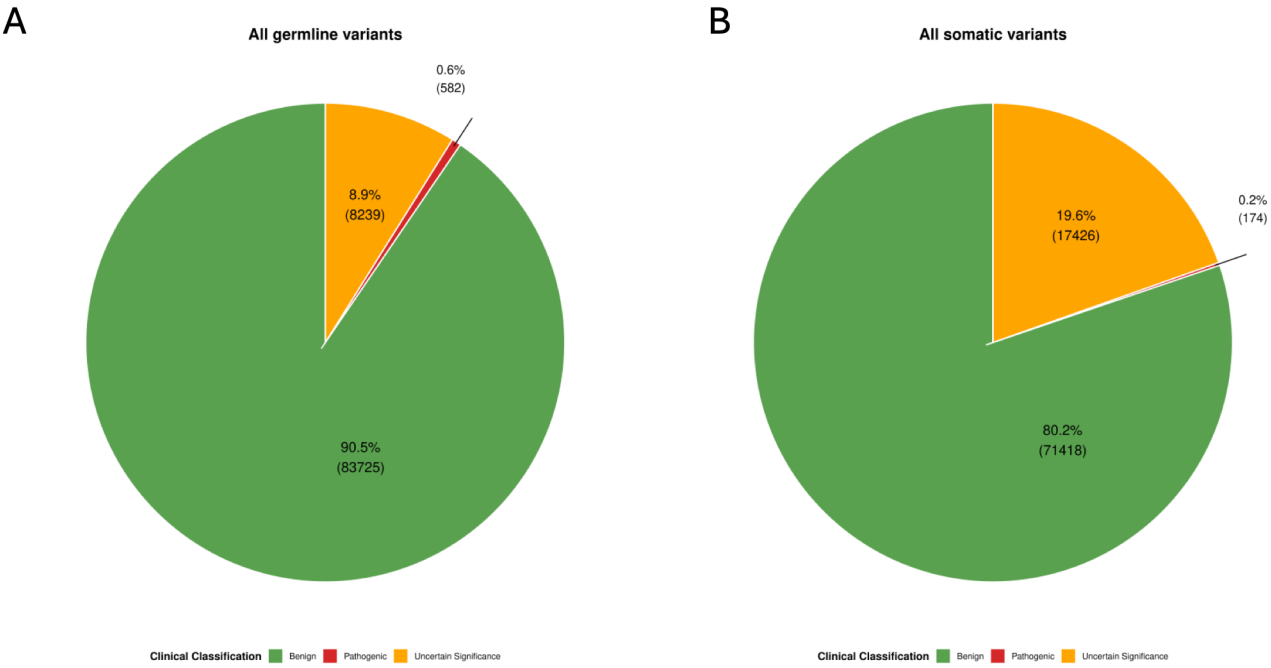
Supplementary Figure S1 – ***Resource utilization metrics of VarianTalker with and without PostgreSQL-backed retrieval.*** *Line plots show cumulative peak RSS, disk read and write for de novo annotation without PostgreSQL retrieval (orange) and PostgreSQL-backed retrieval of previously stored annotations (blue). (A) Performance across ClinVar-derived benchmark datasets of increasing variant number. (B) Performance across simulated multi-sample analyses of increasing sample number. Dots represent the mean of five independent runs. The error bar represents the range of the values, from minimum to maximum, across the 5 replicates for each run.*



Supplementary Figure S2 – ***PostgreSQL query performance according to the database size.*** *Line plots show the median time for a single-variant lookup against the number of variants stored in the database. For lookups that retrieved a stored annotation (blue) and lookpus that found none (grey, dashed). The shaded band represents the 5th to 95th percentile across 6 independent runs.*



Supplementary Figure S3 – ***Comparison of ClinVar reclassification of InterVar and CancerVar with VarSome.*** *Alluvial plots showing the correspondence between classifications assigned by the rule-based tools integrated into varianTalker and those reported by VarSome.* ***(A)*** *Germline InterVar VUS classified by ClinVar as pathogenic/likely pathogenic or benign/likely benign and their corresponding VarSome classifications.* ***(B)*** *Somatic CancerVar VUS classified by CancerVar as Tier I/II or Tier IV and their corresponding VarSome tier assignments. Flow widths represent the number of variants in each classification transition.*



Supplementary Figure S4 – ***Distribution of exprert classification categories in all variants.*** *Distribution of expert classification categories in all variants, shown for germline* ***(A)*** *and somatic* ***(B)*** *samples. Classifications in these panels were assigned by expert review and are independent of the varianTalker calls.*

Supplementary Tables

source	genome	category	database	version
annovar	hg19	-	AFR.sites.2015	08
annovar	hg19	-	ALL.sites.2015	08
annovar	hg19	-	AMR.sites.2015	08
annovar	hg19	-	avsnp	151
annovar	hg19	-	clinvar	20260102
annovar	hg19	-	cosmic	98
annovar	hg19	-	dbnsfp	47a
annovar	hg19	-	dbnsfp47a_interpro	47a
annovar	hg19	-	dbscsnv	v11
annovar	hg19	-	EAS.sites.2015	08
annovar	hg19	-	ensGene	
annovar	hg19	-	ensGeneMrna	
annovar	hg19	-	esp6500siv2_all	v2
annovar	hg19	-	EUR.sites.2015	08
annovar	hg19	-	exac	03
annovar	hg19	-	example_db_generic	
annovar	hg19	-	example_db_gff	3
annovar	hg19	-	gnomad211_exome	211
annovar	hg19	-	gnomad_genome	
annovar	hg19	-	icgc	28
annovar	hg19	-	intervar	20180118
annovar	hg19	-	kgXref	
annovar	hg19	-	knownGene	
annovar	hg19	-	knownGeneMrna	
annovar	hg19	-	MT_ensGene	
annovar	hg19	-	MT_ensGeneMrna	
annovar	hg19	-	refGene	
annovar	hg19	-	refGeneMrna	
annovar	hg19	-	refGeneVersion	
annovar	hg19	-	refGeneWithVer	
annovar	hg19	-	refGeneWithVerMrna	
annovar	hg19	-	rmsk	
annovar	hg19	-	SAS.sites.2015	08
annovar	hg38	-	AFR.sites.2015	08
annovar	hg38	-	ALL.sites.2015	08
annovar	hg38	-	AMR.sites.2015	08
annovar	hg38	-	avsnp	151
annovar	hg38	-	clinvar	20260102
annovar	hg38	-	cosmic	98
annovar	hg38	-	dbnsfp	47a

annovar	hg38	-	dbnsfp47a_interpro	47a
annovar	hg38	-	dbscsnv	v11
annovar	hg38	-	EAS.sites.2015	08
annovar	hg38	-	ensGene	
annovar	hg38	-	ensGeneMrna	
annovar	hg38	-	esp6500siv2_all	v2
annovar	hg38	-	EUR.sites.2015	08
annovar	hg38	-	exac	03
annovar	hg38	-	gnomad211_exome	211
annovar	hg38	-	gnomad_genome	
annovar	hg38	-	icgc	28
annovar	hg38	-	intervar	20180118
annovar	hg38	-	kgXref	
annovar	hg38	-	knownGene	
annovar	hg38	-	knownGeneMrna	
annovar	hg38	-	refGene	
annovar	hg38	-	refGeneMrna	
annovar	hg38	-	refGeneVersion	
annovar	hg38	-	rmsk	
annovar	hg38	-	SAS.sites.2015	08
funcotator	hg19	germline	Achilles	110303
funcotator	hg19	germline	ACMGLMMLof	1
funcotator	hg19	germline	ACMG_recommendation	v3.2
funcotator	hg19	germline	ClinVar_VCF	20251228
funcotator	hg19	germline	dbSNP	b157
funcotator	hg19	germline	Gencode	43
funcotator	hg19	germline	Gencode_XHGNC	75_37
funcotator	hg19	germline	Gencode_XRefSeq	75_37
funcotator	hg19	germline	gnomAD_exome	2.1
funcotator	hg19	germline	HGNC	Nov082023
funcotator	hg19	germline	LMMKnown	20180612
funcotator	hg19	somatic	Achilles	110303
funcotator	hg19	somatic	CGC	full_2012_03-15
funcotator	hg19	somatic	ClinVar	12.03.20
funcotator	hg19	somatic	ClinVar_VCF	20251228
funcotator	hg19	somatic	CosmicFusion	v98
funcotator	hg19	somatic	CosmicTissue	v98
funcotator	hg19	somatic	Cosmic	v98
funcotator	hg19	somatic	dbSNP	b157
funcotator	hg19	somatic	DNARepairGenes	20251229T120219
funcotator	hg19	somatic	Familial_Cancer_Genes	20110905
funcotator	hg19	somatic	Gencode	43
funcotator	hg19	somatic	Gencode_XHGNC	75_37

funcotator	hg19	somatic	Gencode_XRefSeq	75_37
funcotator	hg19	somatic	gnomAD_exome	2.1
funcotator	hg19	somatic	HGNC	Nov082023
funcotator	hg19	somatic	Oreganno	20160119
funcotator	hg19	somatic	Simple_Uniprot	2014_12
funcotator	hg38	germline	Achilles	110303
funcotator	hg38	germline	ACMGLMMLof	1
funcotator	hg38	germline	ACMG_recommendation	v3.2
funcotator	hg38	germline	ClinVar_VCF	20251228
funcotator	hg38	germline	dbSNP	b157
funcotator	hg38	germline	Gencode	43
funcotator	hg38	germline	Gencode_XHGNC	110_38
funcotator	hg38	germline	Gencode_XRefSeq	110_38
funcotator	hg38	germline	gnomAD_exome	4.1
funcotator	hg38	germline	HGNC	Nov082023
funcotator	hg38	germline	LMMKnown	20180618
funcotator	hg38	somatic	Achilles	110303
funcotator	hg38	somatic	CGC	full_2012_03-15
funcotator	hg38	somatic	ClinVar_VCF	20251228
funcotator	hg38	somatic	CosmicFusion	v98
funcotator	hg38	somatic	CosmicTissue	v98
funcotator	hg38	somatic	Cosmic	v98
funcotator	hg38	somatic	dbSNP	b157
funcotator	hg38	somatic	DNARepairGenes	20251229T120219
funcotator	hg38	somatic	Familial_Cancer_Genes	20110905
funcotator	hg38	somatic	Gencode	43
funcotator	hg38	somatic	Gencode_XHGNC	110_38
funcotator	hg38	somatic	Gencode_XRefSeq	110_38
funcotator	hg38	somatic	gnomAD_exome	4.1
funcotator	hg38	somatic	HGNC	Dec302025
funcotator	hg38	somatic	Oreganno	20160119
funcotator	hg38	somatic	Simple_Uniprot	2014_12

*Supplementary Table S1 – **Annotation databases used by ANNOVAR and Funcotator.** Summary of the databases used for variant annotation with ANNOVAR and Funcotator. For each resource, the table reports the annotation tool (source), reference genome assembly (genome), variant category (germline or somatic, where applicable), database name, and database version or release used in the analysis.*

Capability	VarianTalker	PCGR	OpenCRAVAT	Ensembl VEP
Primary scope	Cancer-focused annotation, prioritisation and biomarker extraction for somatic and germline variants; research use only	Tumour-focused somatic and multi-omic interpretation and reporting for precision oncology	General-purpose modular genomic variant interpretation, annotation and scoring; cancer modules are optional	General-purpose variant consequence annotation and prioritisation support
Supported molecular input types / variant classes	Somatic and germline VCF variants, including SNVs and indels, plus copy-number inputs from CNVKit .cnr or Dragen VCF files	Somatic SNVs/indels, allele-specific copy-number segments, RNA fusions and bulk RNA-expression data	Substitutions, insertions and deletions; coverage depends on the installed mapper and annotator modules	SNVs, indels, copy-number variants and other structural variants
Dedicated somatic and germline interpretation in the same workflow/run	Yes; a single annotation-run samplesheet may contain somatic, germline and copy-number inputs	PCGR does not classify raw germline variants; an optional CPSR-classified germline result can be integrated into the PCGR report	No dedicated somatic or germline clinical modes are documented; multiple files or samples can be processed together	No dedicated clinical interpretation modes; somatic and germline records can be annotated together and are treated uniformly
Somatic classification framework	CancerVar using the AMP/ASCO/CAP/CGC framework; ESCAT and CIViC provide actionability and clinical evidence	ClinGen/CGC/VICC oncogenicity classification and AMP/ASCO/CAP actionability tiers	No native somatic clinical-classification framework documented; cancer-specific annotations are module-dependent	No native somatic clinical-classification framework; external classifications can be added through plugins or custom files
Germline classification framework	InterVar ACMG/AMP classification, with ReNOVo as an additional ML pathogenicity predictor	Companion CPSR implements a key subset of ACMG/AMP criteria with gene-specific adaptations	No native ACMG rule engine documented; precomputed classifications can be added through modules such as ClinVar ACMG	No core ACMG rule engine; optional plugins can retrieve externally generated ACMG classifications
Therapeutic actionability	CIViC and an ESCAT-linked in-house actionability resource	CIViC and the Cancer Genome Interpreter Cancer Biomarkers Database, with optional OncoKB integration requiring an API token	Module-dependent evidence, including CIViC, Cancer Genome Interpreter and DGIdb; no integrated native tiering engine documented	No native therapeutic matcher; external evidence can be added through plugins or custom files
Pharmacogenomic annotation	PharmCAT-based pharmacogenomic reporting (PharmGKB)	No native PCGR pharmacogenomic component found in the official v2.3.2 documentation	PharmGKB annotation through an optional module	No native or officially listed PGx component found in the documentation; custom integration remains possible
Transcript selection and standardisation	Funcotator CANONICAL transcript-selection mode by default, with optional user-supplied transcript list (MANE select list by default)	Configurable VEP pick order; the default prioritises MANE Select, MANE Plus Clinical and canonical transcripts	Mapper-based transcript annotation; the result viewer uses a representative MANE transcript by default	User-configurable transcript flags and selection, including MANE Select, MANE Plus Clinical and custom pick order
Annotation-resource update model	Separate updater for selected Funcotator and ANNOVAR resources; not all resources are automatically updated	Date-versioned reference bundle distributed separately but tied to compatible PCGR/VEP versions	Annotator modules are independently installed, versioned and updated through the module store	Core cache is release-matched; plugin and custom-annotation data may be managed separately

Institutional or custom annotations	Custom Funcotator data-source directories and user evidence files are configurable; local XSV annotations are technically supported	No plugin-style custom annotator framework is documented but selected pre-existing VCF INFO annotations can be retained	User-developed custom annotator modules	User-supplied custom annotation files and plugins
Reuse of previously computed annotations	PostgreSQL-backed variant-annotation cache	No persistent cross-sample result cache found in the official documentation reviewed	No persistent cross-job cache documented; duplicate variants are consolidated within multi-file jobs	No persistent result cache found; the VEP cache stores reference annotation data
Execution model and containerisation	Nextflow DSL2; Docker and Singularity/Apptainer	Python-led workflow with an R reporting environment; Conda, Docker and Singularity/Apptainer	Python package distributed through pip, installers and Bioconda; Docker deployment files are available	Perl command-line application; official Docker, documented Singularity and Nextflow execution, and a community Bioconda package
Interactive results/review interface	MAFigate local Streamlit application with filtering, reviewer notes, custom columns and export	Interactive HTML report for exploration and filtering; reviewer-note curation is not documented	Local/web GUI with job submission and interactive result viewer; formal reviewer-note curation is not documented	Web submission, result browsing, filtering and download; reviewer-note curation is not documented

Supplementary Table S2 – ***Comparison of varianTalker with widely used variant annotation and interpretation tools.*** *Qualitative comparison of the documented capabilities of varianTalker, PCGR, OpenCRAVAT, and Ensembl VEP. The table summarises key features of each tool, including their primary scope, supported variant classes, annotation and interpretation capabilities, and other functionalities relevant to cancer variant analysis and prioritisation.*

column_name	example_value	description	primary_evidence_url	source_artifact_url_or_path
avsnp151	rs535965616; rs140269415; rs779620679	dbSNP rs identifier(s) from the ANNOVAR avsnp151 database.	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/
ExAC_AFR	0.0; 0.0002; 0.0001	Adjusted alternate-allele frequency in the African/African American group of the ExAC dataset.	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/
ExAC_AMR	0.0; 0.0002; 0.0003	Adjusted alternate-allele frequency in the Latino/Admixed American group of the ExAC dataset.	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/
ExAC_EAS	0.0; 0.0001; 0.0002	Adjusted alternate-allele frequency in the East Asian group of the ExAC dataset.	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/
ExAC_FIN	0.0; 0.0002; 0.0003	Adjusted alternate-allele frequency in the Finnish group of the ExAC dataset.	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/
ExAC_NFE	0.0; 1.499e-05; 0.0001	Adjusted alternate-allele frequency in the non-Finnish European group of the ExAC dataset.	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/
ExAC_OTH	0.0; 0.0011; 0.0022	Adjusted alternate-allele frequency in the other/uncategorized group of the ExAC dataset.	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/
ExAC_SAS	0.0; 0.0001; 0.0002	Adjusted alternate-allele frequency in the South Asian group of the ExAC dataset.	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/
Freq_1000g2015aug_all	0.000199681; 0.000399361; 0.000599042	Alternate-allele frequency in the ANNOVAR 1000 Genomes August 2015 all-population dataset.	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/
Freq_esp6500siv2_all	7.7e-05; 0.0002; 0.0003	Alternate-allele frequency in the ANNOVAR ESP6500si-v2 all-samples dataset.	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/
Freq_ExAC_ALL	0.0001; 0.0002; 0.0003	Alternate-allele frequency in the ANNOVAR ExAC all-samples dataset.	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/
AAChange.knownGene	UNKNOWN; MT-ND4L:ENST00000361335.1:exon1:c.1_297del:p.M1_*99del; ENST00000256474.3:ENST00000256474.3:exon2:c.341_463del:p.G114_P154del	ANNOVAR transcript-level coding/protein change annotation using UCSC Known Gene models. Entries can include gene, transcript, exon, cDNA change and protein change; multiple transcript annotations may be comma-separated.	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/
AAChange.refGene	UNKNOWN; MSH2:NM_000251:exon1:c.1_211del:p.M1?,MSH2:NM_001258281:exon2:c.1_13del:p.M1?; BRCA1:NM_007297:exon18:c.5053_5136del:p.H1685_K1712del,BRCA1:NM_007298:exon18...	ANNOVAR transcript-level coding/protein change annotation using NCBI RefSeq gene models. Entries can include gene, transcript, exon, cDNA change and protein change; multiple transcript annotations may be comma-separated.	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/
clinvar: Clinvar	clinvar: Uncertain_significance; clinvar: Likely_benign; clinvar: UNK	InterVar simplified ClinVar annotation string (for example “clinvar: Benign” or “clinvar: UNK”), derived from the ClinVar/ANNOVAR resource configured for InterVar. It is not the standard ClinVar VCF CLNSIG field.	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/	https://annovar.openbioinformatics.org/en/latest/user-guide/filter/
AAChange.ensGene	UNKNOWN; MT-ND4L:ENST00000361335.1:exon1:c.1_297del:p.M1_*99del; VHL:ENST00000256474.3:exon2:c.341_463del:p.G114_P154del	ANNOVAR transcript-level coding/protein change annotation using Ensembl gene models. Entries can include gene, transcript, exon, cDNA change and protein change; multiple transcript annotations may be comma-separated.	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/
ExonicFunc.ensGene	nonsynonymous SNV; synonymous SNV; stopgain	ANNOVAR exonic consequence class using Ensembl gene models (for example nonsynonymous SNV, synonymous SNV, stopgain, frameshift insertion/deletion).	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/
ExonicFunc.knownGene	nonsynonymous SNV; synonymous SNV; stopgain	ANNOVAR exonic consequence class using UCSC Known Gene models (for example nonsynonymous SNV, synonymous SNV, stopgain, frameshift insertion/deletion).	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/
ExonicFunc.refGene	nonsynonymous SNV; synonymous SNV; stopgain	ANNOVAR exonic consequence class using NCBI RefSeq gene models (for example nonsynonymous SNV, synonymous SNV, stopgain, frameshift insertion/deletion).	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/
Func.ensGene	exonic; intron; intronic	ANNOVAR functional region/category relative to Ensembl gene models (for example exonic, intronic, splicing, UTR, upstream/downstream or intergenic).	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/
Func.knownGene	exonic; intron; intronic	ANNOVAR functional region/category relative to UCSC Known Gene models (for example exonic, intronic, splicing, UTR, upstream/downstream or intergenic).	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/

Gene.knownGene	TTN; BRCA2,ENST00000380152.8; ENST00000358273.9,NF1	Gene symbol(s) assigned by ANNOVAR using UCSC Known Gene models; for intergenic records, neighboring genes may be reported.	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/
Gene.refGene	TTN; BRCA2; ATM	Gene symbol(s) assigned by ANNOVAR using NCBI RefSeq gene models; for intergenic records, neighboring genes may be reported.	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/
GeneDetail.ensGene	dist=1; ENST00000543555.6:exon2:UTR5; dist=4	Additional ANNOVAR detail for the Ensembl gene models annotation, such as transcript/exon information for splice-related records or distances/neighboring-gene details for intergenic/upstream/downstream records. Exact syntax is consequence-dependent.	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/
GeneDetail.knownGene	dist=1; ENST00000543555.6:exon2:UTR5; dist=8	Additional ANNOVAR detail for the UCSC Known Gene models annotation, such as transcript/exon information for splice-related records or distances/neighboring-gene details for intergenic/upstream/downstream records. Exact syntax is consequence-dependent.	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/
GeneDetail.refGene	dist=1; dist=4; dist=11	Additional ANNOVAR detail for the NCBI RefSeq gene models annotation, such as transcript/exon information for splice-related records or distances/neighboring-gene details for intergenic/upstream/downstream records. Exact syntax is consequence-dependent.	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/
cosmic	ID=COSV99049576;OCCURENCE=2(lung); ID=COSV59705710;OCCURENCE=1(breast); ID=COSV58986988;OCCURENCE=1(upper_aerodigestive_tract),1(large_intestine)	ANNOVAR COSMIC annotation string for the exact variant, typically containing a COSV identifier and tissue-specific occurrence counts from the configured COSMIC release.	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/
cosmic98	ID=COSV99049576;OCCURENCE=2(lung); ID=COSV59705710;OCCURENCE=1(breast); ID=COSV58986988;OCCURENCE=1(upper_aerodigestive_tract),1(large_intestine)	ANNOVAR COSMIC v98 annotation string for the exact variant, typically containing a COSV identifier and occurrence counts by tissue.	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/	https://annovar.openbioinformatics.org/en/latest/user-guide/gene/
CGC_Cancer_Germline_Mut	yes; -'	Indicator that the gene has documented germline cancer-predisposition mutations in the Cancer Gene Census source; gene-level, not a statement about the current variant.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_Cancer_Molecular_Genetics	Rec; Dom; Dom, Rec	Cancer-predisposition inheritance/molecular-genetics descriptor for the gene (for example dominant, recessive or X-linked); controlled values are source-version-dependent.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_Cancer_Somatic_Mut	yes	Indicator that the gene has documented somatic mutations in cancer in the Cancer Gene Census source; gene-level, not a statement about the current variant.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_CANCER_SYNDROME	Hereditary breast/ovarian cancer; Ataxia-telangiectasia; Neurofibromatosis type 1	Cancer-predisposition syndrome(s) associated with the gene in the Cancer Gene Census.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_CHR_BAND	13q12; 17q12; 11q22.3	Cytogenetic chromosome band of the gene.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_CHROMOSOME		Chromosome containing the gene.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_COSMIC_GENE_ID		COSMIC gene identifier for the Cancer Gene Census record.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_GeneID	675.0; 472.0; 4763.0	NCBI Gene identifier for the Cancer Gene Census gene.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_GENOME_START		Genomic start coordinate of the Cancer Gene Census gene record on the datasource assembly; it is not the start coordinate of the current variant.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_GENOME_STOP		Genomic end coordinate of the Cancer Gene Census gene record on the datasource assembly; it is not the end coordinate of the current variant.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_GERMLINE		Flag/descriptor indicating documented germline involvement of the gene in cancer predisposition.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_MOLECULAR_GENETICS		Curated molecular-genetics/inheritance descriptor for the gene; controlled codes and delimiters depend on the Cancer Gene Census release.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms

CGC_Mutation_Type	D, Mis, N, F, S; T; Mis	Cancer-associated alteration types curated for the gene (for example missense, nonsense, frameshift, splice-site, deletion, amplification or translocation); code definitions are release-dependent.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_MUTATION_TYPES		Cancer-associated alteration types curated for the gene; code definitions are release-dependent and should not be expanded using unsupported guesses.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_NAME	familial breast/ovarian cancer gene 2; ataxia telangiectasia mutated; neurofibromatosis type 1 gene	Full name of the Cancer Gene Census gene.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_Other_Diseases	Hirschsprung disease; Osteogenesis imperfecta; Piebald trait	Other disease associations recorded for the gene in the Cancer Gene Census source.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_OTHER_GERMLINE_MUT	yes; Yes	Additional germline-mutation/disease information recorded for the gene in the Cancer Gene Census source.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_OTHER_SYNDROME		Additional syndrome associations recorded for the gene.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_ROLE_IN_CANCER		Cancer Gene Census role assigned to the gene, such as oncogene, tumour-suppressor gene or fusion partner; this is gene-level, not a classification of the current variant.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_SOMATIC		Flag/descriptor indicating documented somatic involvement of the gene in cancer.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_SYNONYMS		Alternative symbols/identifiers for the Cancer Gene Census gene.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_TIER		Cancer Gene Census evidence tier for the gene. Tier 1 genes have strong documented evidence of cancer-driving activity; Tier 2 genes have strong but less extensive/emerging evidence.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_TISSUE_TYPE	E; L; L, E	Broad tumour/tissue class associated with the Cancer Gene Census gene record (for example epithelial, lymphoid, mesenchymal or other); it does not indicate tissue-specific expression of the current variant.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_Translocation_Partner	BRAF; NPM1, TPM3, TFG, TPM4, ATIC, CLTC, MSN, ALO17, CARS, EML4, KIF5B, C2orf22; TRB@	Gene(s) reported as recurrent translocation/fusion partners of the Cancer Gene Census gene.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_TRANSLOCATION_PARTNER	BRAF; NPM1, TPM3, TFG, TPM4, ATIC, CLTC, MSN, ALO17, CARS, EML4, KIF5B, C2orf22; TRB@	Gene(s) reported as translocation/fusion partners of the Cancer Gene Census gene.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_Tumor_Types_Germline	breast, ovarian, pancreatic, leukemia (FANCB, FANCD1); leukemia, lymphoma, medulloblastoma, glioma; hamartoma, renal cell	Tumor types linked to germline alteration of the gene in the Cancer Gene Census.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_Tumor_Types_Somatic	breast, ovarian, pancreatic; T-PLL; neurofibroma, glioma	Tumor types linked to somatic alteration of the gene in the Cancer Gene Census.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_TUMOUR_TYPES_GERMLINE		Tumor types linked to germline alteration of the gene in the Cancer Gene Census.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
CGC_TUMOUR_TYPES_SOMATIC		Tumor types linked to somatic alteration of the gene in the Cancer Gene Census.	https://cancer.sanger.ac.uk/census	CGC CSV/TSV downloaded under the applicable COSMIC terms
COSMIC_fusion_genes	ENST00000358273.8(NF1):r.1_4091::ENST00000225823.6(ASIC2):r.1582_3443(2); ENST00000318522.9(EML4):r.1_1751::ENST00000389048.7(ALK):r.4080_6220(367)|ENS...; ENST00000277541.7(NOTCH1):r.1_5638::ENST00000259455.3(GABBR2):r.2354_5788(2)|...	COSMIC fusion transcripts/genes involving the annotated gene, serialized using HGVS-like RNA-level breakpoint notation; each record may include a count in parentheses. This is primarily a gene/fusion catalogue annotation, not evidence that the current small variant causes the fusion.	https://cancer.sanger.ac.uk/census	licensed/downloaded COSMIC mutation or fusion file plus local datasource config

COSMIC_overlapping_mutations	p.D145_E14445dup(25); p.D45_S1245del(1)|p.D45_K1373del(1)|p.S7_N1269del(10)|p.G124_K1373del(1); p.D145_E1407dup(17)|p.D145_K1362dup(38)|p.D145_E14445dup(25)|p.D145_Q1494dup(...	COSMIC coding mutation(s) overlapping the current protein change; a count in parentheses is the datasource-reported number of affected samples/records after its filters.	https://cancer.sanger.ac.uk/census	licensed/downloaded COSMIC mutation or fusion file plus local datasource config
COSMIC_tissue_types_affected	biliary_tract(2)|breast(11)|central_nervous_system(44)|large_intestine(11)|pa...; biliary_tract(2)|breast(11)|central_nervous_system(44)|large_intestine(11)|li...; biliary_tract(2)|breast(11)|central_nervous_system(44)|large_intestine(11)|lu...	COSMIC tissue categories in which alterations of the annotated gene are represented; counts in parentheses are source-dataset alteration/sample counts, not the number of samples tested for the current variant.	https://cancer.sanger.ac.uk/census	licensed/downloaded COSMIC mutation or fusion file plus local datasource config
COSMIC_total_alterations_in_gene	91.0; 90.0; 92.0	Total alteration count associated with the annotated gene in the configured COSMIC datasource; this is gene-level and is not the number of samples tested for the current variant.	https://cancer.sanger.ac.uk/census	licensed/downloaded COSMIC mutation or fusion file plus local datasource config
CancerVar	Tier_IV_benign; Tier_III_Uncertain; Tier_II_potential	CancerVar somatic clinical tier: Tier_I_strong, Tier_II_potential, Tier_III_Uncertain or Tier_IV_benign, based on the bundled CancerVar implementation and evidence database.	https://cancervar.wglab.org/	https://cancervar.wglab.org/about.php
CancerVar: CancerVar and Evidence	CancerVar: -1#Tier_IV_benign EVS=[0, 0, 0, 0, 0, 0, -1, 0, 0, 0]; CancerVar: -1#Tier_IV_benign EVS=[0, 0, 0, 0, 0, 0, 0, 0, -1, 0, 0]; CancerVar: 0#Tier_IV_benign EVS=[0, 0, 0, 0, 0, 0, 0, 0, 0, 0]	CancerVar aggregate output containing the numerical evidence score, tier and a 12-element evidence vector (EVS) corresponding to the CancerVar evidence criteria implemented in the bundled version.	https://cancervar.wglab.org/	https://cancervar.wglab.org/about.php
ICGC_Occurrence	UCEC-US|1|531|0.00188324; SKCM-US|1|466|0.00214592; COAD-US|1|402|0.00248756	ICGC occurrence summary for the exact variant, serialized by project with mutated-patient count, project total and/or frequency according to the configured annotation release.	https://dcc.icgc.org/	https://www.icgc-argo.org/
icgc28	MU82293966; MU129190874; MU129162988	ANNOVAR ICGC release 28 annotation for the variant; "." indicates no matching annotation in that source.	https://dcc.icgc.org/	https://www.icgc-argo.org/
AF		Alternate-allele frequency emitted by the upstream caller. In Torrent Variant Caller it is based on flow-evaluator observation counts; in other callers the exact definition may differ.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
FR		Caller-provided reason(s) that the allele was filtered.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
FSAF		Flow Evaluator alternate-allele observations on the forward strand.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
FSAR		Flow Evaluator alternate-allele observations on the reverse strand.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
FSRF		Flow Evaluator reference-allele observations on the forward strand.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
FSRR		Flow Evaluator reference-allele observations on the reverse strand.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
FWDB		Forward-strand bias metric reported by Torrent Variant Caller.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header

FXX		Flow Evaluator failed-read ratio/metric reported by Torrent Variant Caller.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
GCM		Torrent Variant Caller indicator that the locus/allele is associated with a GC-motif context.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
HRUN		Homopolymer-run length in the reference context at or adjacent to the variant, as reported by Torrent Variant Caller.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
LEN		Caller-reported allele/event length.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
MISC		Miscellaneous caller-provided information about the variant call.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
MLLD		Mean log-likelihood delta per read between allele hypotheses.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
OALT		Original alternate allele(s) before normalization/decomposition.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
OID		Original hotspot/allele identifier(s) before normalization/decomposition.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
OMAPALT		Mapping between original OID/OPOS/OREF/OALT entries and the normalized ALT allele(s).	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
OPOS		Original allele position(s) before normalization/decomposition.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
OREF		Original reference allele(s) before normalization/decomposition.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
PB		Relative read-position bias statistic comparing variant-supporting and reference-supporting reads (Mann–Whitney U rho form).	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
PBP		P-value for relative read-position bias between variant-supporting and reference-supporting reads.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
PPD		Number of prefix-padding bases added to the normalized alternate allele representation.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
QD		Quality-by-depth metric reported by the caller; in Torrent Variant Caller it is defined from QUAL and flow-evaluator depth using the caller's formula.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
RBI		Distance of the fitted bias parameters from zero, as reported by Torrent Variant Caller.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header

REFB		Reference-hypothesis bias metric reported by Torrent Variant Caller.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
REVB		Reverse-strand bias metric reported by Torrent Variant Caller.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
RO		Reference-allele observation count reported by the caller.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
SAF		Alternate-allele observations on the forward strand.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
SAR		Alternate-allele observations on the reverse strand.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
SPD		Number of suffix-padding bases added to the normalized alternate allele representation.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
SRF		Reference-allele observations on the forward strand.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
SSEN		Strand-specific error metric/prediction for the negative strand.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
SSSB		Strand-specific strand-bias metric for the allele.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
STB		Strand-bias statistic comparing variant and reference observations.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
STBP		P-value for strand bias comparing variant and reference observations.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
TYPE		Caller-reported allele type, commonly snp, mnp, ins, del or complex.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
VARB		Variant-hypothesis bias metric reported by Torrent Variant Caller.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
AO		Alternate-allele observation count reported by the caller.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
FDP		Flow Evaluator read depth at the locus.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
FDVR		Flow-disruption metric comparing the alternate and reference hypotheses, as reported by Torrent Variant Caller.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header

SRR		Reference-allele observations on the reverse strand.	https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf	actual Torrent Variant Caller VCF header
Annotation_Transcript	ENST00000589042.5; ENST00000380152.8; ENST00000675843.1	Identifier of the transcript selected by Funcotator for detailed annotation (typically an Ensembl/GENCODE transcript accession, including version when available).	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
cDNA_Change	c.e3+1G>A; c.e2+1G>A; c.e3-1G>A	HGVS-like coding-DNA or transcript change for the selected transcript, for example c.1427T>C. The number is a transcript/CDS coordinate, not a codon number.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
Chromosome	chr1; chr2; chr17	Reference contig/chromosome on which the annotated allele occurs, as emitted by Funcotator and consistent with the input variant position.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
Codon_Change	c.e2+1; c.e3+1; c.e4+1	Codon-aligned nucleotide change for the selected coding transcript. The reported sequence is aligned to complete codons (except frameshifts); changed bases are capitalized and contextual reference bases are lower-case.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
Description	titin; BRCA2 DNA repair associated; ATM serine/threonine kinase	Full gene description/name supplied by the corresponding annotation data source.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
End_Position	45770205; 83134243; 147912050	1-based inclusive genomic end coordinate used by Funcotator for the annotated allele. It equals Start_Position for an SNV; for indels and multi-base events, interpret it together with Start_Position and the MAF-style Reference_Allele/Tumor_Seq_Allele2 representation.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
Entrez_Gene_Id	7273.0; 675.0; 472.0	NCBI Gene (Entrez) numeric identifier for the gene associated with the annotated transcript or locus.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
gc_content	0.6109725685785536; 0.6034912718204489; 0.600997506234414	GC fraction in the reference-sequence window used by the Funcotator GENCODE annotation around the allele.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
Gencode_43_secondaryVariantClassification	INTRON; MISSENSE; SILENT	Additional Funcotator classification for alleles whose primary Variant_Classification is SPLICE_SITE; otherwise the field is empty. The “Gencode_43” prefix identifies the data-source version/name.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
Genome_Change	g.chr1:66927delG; g.chr1:69134A>G; g.chr1:69308A>G	Funcotator genomic-change string summarizing the event on the reference sequence. The exact HGVS-like syntax varies for SNPs, MNPs, insertions and deletions (for example g.chr2:29320870A>G for an SNP).	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
Hugo_Symbol	TTN; BRCA2; ATM	Gene symbol assigned to the annotated allele by the selected GENCODE transcript annotation; “Unknown” may be used when no transcript/gene boundary is matched.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
Matched_Norm_Sample_Barcode	Germline test: clinvar_subset_1000_913; somatic test: empty	Identifier stored in the matched-normal barcode field of the MAF. In the current VariantAlker germline output it carries the analysed sample ID while Tumor_Sample_Barcode is __UNKNOWN__; in somatic mode it is empty unless a matched normal is supplied.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
NCBI_Build	hg38	Reference assembly used for the Funcotator GENCODE annotation: hg19 or hg38.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
Other_Transcripts	NEB_ENST00000427231.7_Intron|NEB_ENST00000409198.5_Intron; RYR1_ENST00000355481.8_Intron; CHD7_ENST00000524602.5_Intron	Funcotator summary of non-selected overlapping transcripts. A standard entry has the form HUGO_SYMBOL_TRANSCRIPT_ID_VARIANT_CLASSIFICATION[_PROTEIN_CHANGE], and standard Funcotator documentation separates multiple entries with “/” or "|".	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications

Protein_Change	p.M1V; p.M1I; p.M1T	Predicted amino-acid change on the selected transcript/protein in Funcotator notation (for example p.V476A); empty when no protein consequence can be assigned.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
ref_context	GTGATCCCCCCAGCAGCAGC; ACTCCGTCTCAAAAAAAAAA; ACTCCATCTCAAAAAAAAAA	Reference-sequence context surrounding the allele, as emitted by the Funcotator GENCODE annotation.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
Reference_Allele	C; G; A	Reference allele represented by Funcotator for the annotated event; insertions are represented as "-".	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
Start_Position	45770204; 83134242; 147912049	1-based inclusive genomic start coordinate of the annotated allele, corresponding to the start position used by Funcotator after the input allele representation/normalization.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
Strand	+	Genomic strand used for the MAF allele representation, normally "+"; do not confuse it with Transcript_Strand.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
Transcript_Exon	2.0; 1.0; 3.0	1-based exon number on the selected transcript when the allele lies in an expressed exon; empty for annotations such as intronic or intergenic events.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
Transcript_Position	629; 628; 575	1-based inclusive position of the allele on the selected transcript; multi-base events may be reported as start_end, and non-transcribed/intergenic events may be empty.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
Transcript_Strand	-; +	Strand of the selected annotation transcript: + or -.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
Tumor_Sample_Barcode	ATND0404NX-1TAD104	Identifier of the tumor or analyzed sample in MAF-compatible output. A matched-normal identifier may be present separately; this field does not imply that a "normal subtraction" operation was performed.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
Tumor_Seq_Allele1	C; G; A	Reference allele in the Funcotator MAF representation. Despite the historical field name, Funcotator writes the reference allele here.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
Tumor_Seq_Allele2	T; A; G	Alternate allele represented by this MAF record; for a biallelic variant this is the annotated variant allele.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
Variant_Classification	Missense_Mutation; Silent; Intron	Funcotator/MAF functional consequence category for the selected transcript, using MAF terms such as Missense_Mutation, Silent, Nonsense_Mutation, Frame_Shift_Del, Splice_Site or Intron.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
Variant_Type	SNP; DEL; INS	DNA-level MAF variant type. Common values include SNP, INS and DEL; Funcotator can also emit multi-nucleotide categories such as DNP, TNP, ONP or MNP.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
hotspot		Caller/database flag indicating that the locus is present in the configured hotspot list; hotspot status may alter calling/filtering behavior but does not by itself establish pathogenicity.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	local VCF, target BED and caller logs
n_alt_count		Number of reads/observations supporting the alternate allele in the normal or matched-normal sample, when available.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	actual input VCF/MAF and output table
n_ref_count		Number of reads/observations supporting the reference allele in the normal or matched-normal sample, when available.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	actual input VCF/MAF and output table

t_alt_count	23	Number of reads/observations supporting the alternate allele in the tumor/analyzed sample.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	actual input VCF/MAF and output table
t_ref_count	25	Number of reads/observations supporting the reference allele in the tumor/analyzed sample.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	actual input VCF/MAF and output table
tumor_f	0,479166667	Tumor/analyzed-sample variant allele fraction. VariantAlker recalculates it as t_alt_count/(t_alt_count+t_ref_count) when those counts are available.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	actual input VCF/MAF and output table
Achilles_Top_Genes	Esophageal Squamous(138;838 1285 7957 30353 30468 36915 49332); NSCLC(107;854 1218 9699 17025 28335 47076 52975)|Esophageal Squamous(32;282 5...; Colon(161;2205 2542 7338 31318)	Gene-level Achilles/DepMap annotation. In the current somatic output, values are cancer-type labels followed by parenthesized numeric lists rather than a yes/no flag. The exact meaning of the numeric list and the selection threshold must be read from the local Funcotator datasource schema.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
ACMG_recommendation_Disease_Name	Dilated cardiomyopathy (truncating variants only) (MIM 604145); Breast-ovarian cancer, familial 2 (MIM 612555); Adenomatous polyposis coli (MIM 175100)	Disease/condition name linked to the gene in the configured ACMG-recommendation Funcotator data source (typically an ACMG secondary-findings gene/condition list). The field is gene-level and version-dependent.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
ACMGLMMLof_LOF_Mechanism	YES; NO	Gene–disease-level loss-of-function mechanism flag/descriptor from the configured ACMG LoF-mechanism Funcotator data source; it does not assert that the current variant is loss-of-function.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
ACMGLMMLof_Mode_of_Inheritance	AD; AR; XLD/XLR	Mode-of-inheritance code for the gene–disease relationship in the configured ACMG LoF-mechanism data source (for example AD, AR, XLD or XLR).	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
ACMGLMMLof_Notes	YES FOR AR; ?; YES for Hirschsprung disease (not returnable?)	Source notes for the gene–disease loss-of-function mechanism record in the configured ACMG LoF data source.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
DNARepairGenes_Accession_number_linked_to_NCBI_Entrez	NM_000051; NM_007294; NM_000135	RefSeq nucleotide accession associated with the DNA-repair gene in the custom datasource, for example an NM_ accession. This is gene-level, not variant-specific.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
DNARepairGenes_Chromosome_location_linked_to_Genome_Data_Viewer	11q22.3; 17q21.31; 15q26.1	Cytogenetic chromosome location associated with the DNA-repair gene in the custom datasource, such as 11q22.3. This is gene-level, not variant-specific.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
Familial_Cancer_Genes_Reference	Familial Cancer Database	Reference/source label for the familial-cancer gene/disease annotation in the paired fields.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
Familial_Cancer_Genes_Synonym	Pancytopenia Dysmelia, FA (several complementation groups); Ciliary Dyskinesia, Primary (CILD1 - CILD13), Immotile Cilia Syndrome; Brain tumor-colorectal polyp(osis) syndrome; ;Hereditary Non-Polyposis Colore...	Familial cancer disorder/syndrome names or aliases associated with the gene in the configured familial-cancer data source.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
FamilialCancerDatabase_Syndromes	Fanconi Anemia; Kartagener syndrome; Turcot syndrome;Muir-Torre syndrome;Lynch syndrome;Constitutional Mismatch Re...	Syndromes associated with the gene in the configured Familial Cancer Database source.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications

LMMKnown_LMM_FLAGGED	False; True	Flag from the configured loss-of-function mechanism/known-variant data source indicating that the record was flagged by that source. Interpret only with the exact data-source documentation and version.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
ClinVar_VCF_CLNVCSO	SO:0001483; SO:0000159; SO:1000035	Sequence Ontology identifier corresponding to the ClinVar variant type reported in CLNVC; it is not restricted to single-nucleotide variants.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
ClinVar_VCF_CLNVI	Genetic_Testing_Registry_(GTR):GTR000569522|Genetic_Testing_Registry_(GTR):GT...; OMIM:154270.0001; OMIM:131230.0001	External variant identifiers submitted to ClinVar, represented as source:identifier pairs; these are identifiers, not evidence-source publications.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
ClinVar_VCF_GENEINFO	BRCA2:675; TTN:7273|TTN-AS1:100506866; APC:324	Gene symbol and NCBI Gene ID pair(s) associated with the variant, encoded as SYMBOL:GeneID and separated by "|" when multiple.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
ClinVar_VCF_ONCDISDB	Human_Phenotype_Ontology:HP:0002664,Human_Phenotype_Ontology:HP:0003008,Human...; Human_Phenotype_Ontology:HP:0002858,Human_Phenotype_Ontology:HP:0006754,MONDO...; Human_Phenotype_Ontology:HP:0100723,MONDO:MONDO:0011719,MeSH:D046152,MedGen:C...	Condition database identifier(s) submitted for ClinVar oncogenicity classifications.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
ClinVar_VCF_ORIGIN	1.0; 0.0; 32.0	Allele-origin bitmask used by ClinVar. Values can be combined: 0 unknown, 1 germline, 2 somatic, 4 inherited, 8 paternal, 16 maternal, 32 de novo, 64 biparental, 128 uniparental, 256 not tested, 512 tested inconclusive, and 1073741824 other.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
ClinVar_VCF_SCIREVSTAT	criteria_provided,_single_submitter; no_classification_for_the_single_variant; no_assertion_criteria_provided	ClinVar review status for the aggregate somatic clinical-impact classification.	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications	https://gatk.broadinstitute.org/hc/en-us/articles/360035531732-Funcotator-Annotation-Specifications
GO_Biological_Process	regulation of transcription, DNA-templated (GO:0006355)|transcription, DNA-te...; adult heart development (GO:0007512)|blood coagulation (GO:0007596)|cardiacm...; brain development (GO:0007420)|cell aging (GO:0007569)|centrosome duplication...	Gene Ontology Biological Process annotations for the gene product, reported as GO term names and GO identifiers. These are curated/evidence-backed annotations, not necessarily every process in which the gene can participate.	https://geneontology.org/docs/ontology-documentation/	https://geneontology.org/docs/ontology-documentation/
GO_Cellular_Component	nucleus (GO:0005634); integral component of membrane (GO:0016021)|plasma membrane (GO:0005886); integral component of membrane (GO:0016021)	Gene Ontology Cellular Component annotations describing cellular locations or complexes associated with the gene product, reported with GO identifiers.	https://geneontology.org/docs/ontology-documentation/	https://geneontology.org/docs/ontology-documentation/
GO_Molecular_Function	calcium ion binding (GO:0005509); DNA binding (GO:0003677)|metal ion binding (GO:0046872); actin filament binding (GO:0051015)|actinin binding (GO:0042805)|ATP binding ...	Gene Ontology Molecular Function annotations describing molecular activities of the gene product, reported with GO identifiers.	https://geneontology.org/docs/ontology-documentation/	https://geneontology.org/docs/ontology-documentation/
AlphaMissense_pred	B; P; A	dbNSFP AlphaMissense class: B (likely benign), A (ambiguous) or P (likely pathogenic), using the thresholds encoded by the dbNSFP release.	https://github.com/google-deepmind/alphamissense	https://github.com/google-deepmind/alphamissense

AlphaMissense_rankscore	0.99795; 0.99585; 0.99395	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://github.com/google-deepmind/alphamissense	https://github.com/google-deepmind/alphamissense
AlphaMissense_score	0.077; 0.07; 0.076	AlphaMissense pathogenicity score from 0 to 1; larger values indicate greater predicted pathogenicity for a missense variant.	https://github.com/google-deepmind/alphamissense	https://github.com/google-deepmind/alphamissense
am_class	likely_benign; likely_pathogenic; ambiguous	AlphaMissense class derived from am_pathogenicity: likely_benign for scores <0.34, likely_pathogenic for scores >0.564, and ambiguous otherwise.	https://github.com/google-deepmind/alphamissense	https://github.com/google-deepmind/alphamissense
am_pathogenicity	0.9999; 0.9998; 0.9997	Calibrated AlphaMissense pathogenicity score from 0 to 1; larger values indicate a higher predicted probability of clinical pathogenicity for the missense substitution.	https://github.com/google-deepmind/alphamissense	https://github.com/google-deepmind/alphamissense
RENOVO_Class	IP Benign; HP Benign; HP Pathogenic	ReNOVo output class combining predicted label (Pathogenic or Benign) with a precision/confidence band: HP, IP or LP (high, intermediate or low precision/confidence, according to the ReNOVo implementation).	https://github.com/mazzalab-ieo/renovo	local ReNOVo model and output schema
RENOVO_pls	0.0143116588143269; 0.0; 0.0020201064852864	ReNOVo Pathogenicity Likelihood Score (PLS) produced by the random-forest model; larger values provide stronger model support for the pathogenic prediction and are used with the predicted label/confidence band to derive RENOVO_Class.	https://github.com/mazzalab-ieo/renovo	local ReNOVo model and output schema
project_id	breast_project; code02234; internal_att	Project identifier propagated from the VariantAlker input metadata.	https://github.com/zhanyinx/variantalker	local samplesheet.csv, nextflow.config and execution report
tumor_tissue	Breast; Uterus	Tumor tissue/site label supplied in the VariantAlker sample metadata and used by tissue-aware annotations such as ESCAT/CIViC.	https://github.com/zhanyinx/variantalker	local samplesheet.csv, nextflow.config and execution report
filter	PASS; NOPASS	VariantAlker prioritization result. PASS means the variant satisfied the configured annotation/depth/VAF/gene filters or was retained by the pathogenic-retention rule; NOPASS means it did not. This is not the original VCF FILTER field, because VariantAlker first retains only input records with VCF FILTER=PASS.	https://github.com/zhanyinx/variantalker	https://github.com/zhanyinx/variantalker
variantalker_naive	Uncertain_Significance; Likely_Benign; Pathogenic	VariantAlker naive clinical summary. The pipeline maps available values from CancerVar, InterVar, ClinVar_VCF_CLNSIG, RENOVO_Class and CIViC_Evidence_Level to ordered numeric tiers, then returns the strongest (minimum-numbered) class: Pathogenic, Likely_Pathogenic, Uncertain_Significance, Likely_Benign, Benign, Not_Provided or Unknown. If none of the five fields contains usable data, the output is "🔍 No Clinical Data".	https://github.com/zhanyinx/variantalker	https://github.com/zhanyinx/variantalker
CLINSIG	Uncertain_significance; Likely_benign; Benign	Aggregate ClinVar germline clinical significance/classification copied by the configured legacy/ANNOVAR data source; multiple values and delimiters are release-dependent.	https://github.com/zhanyinx/variantalker/blob/main/resources/configs/config.init.intervar	https://www.ncbi.nlm.nih.gov/clinvar/docs/
gnomAD_exome_AF		Alternate-allele frequency in the gnomAD exome dataset for the represented allele and release.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_afr	0.0; 2.98686e-05; 5.97372e-05	Adjusted alternate-allele frequency in gnomAD exomes for African/African American genetic-ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_fin	0.0; 1.87196e-05; 1.87203e-05	Adjusted alternate-allele frequency in gnomAD exomes for Finnish genetic-ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_fin_female		Adjusted alternate-allele frequency in gnomAD exomes for Finnish genetic-ancestry group among female samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_fin_male		Adjusted alternate-allele frequency in gnomAD exomes for Finnish genetic-ancestry group among male samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_nfe_swe		Adjusted alternate-allele frequency in gnomAD exomes for nfe swe group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_female		Adjusted alternate-allele frequency in gnomAD exomes for female group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_afr_female		Adjusted alternate-allele frequency in gnomAD exomes for African/African American genetic-ancestry group among female samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_afr_male		Adjusted alternate-allele frequency in gnomAD exomes for African/African American genetic-ancestry group among male samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_amr	0.0; 2.23594e-05; 4.47187e-05	Adjusted alternate-allele frequency in gnomAD exomes for Admixed American genetic-ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_amr_female		Adjusted alternate-allele frequency in gnomAD exomes for Admixed American genetic-ancestry group among female samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help

gnomAD_exome_AF_amr_male		Adjusted alternate-allele frequency in gnomAD exomes for Admixed American genetic-ancestry group among male samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_asj	0.0; 3.82614e-05; 3.82643e-05	Adjusted alternate-allele frequency in gnomAD exomes for Ashkenazi Jewish genetic-ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_asj_female		Adjusted alternate-allele frequency in gnomAD exomes for Ashkenazi Jewish genetic-ancestry group among female samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_asj_male		Adjusted alternate-allele frequency in gnomAD exomes for Ashkenazi Jewish genetic-ancestry group among male samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_eas	0.0; 2.51889e-05; 5.03778e-05	Adjusted alternate-allele frequency in gnomAD exomes for East Asian genetic-ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_eas_female		Adjusted alternate-allele frequency in gnomAD exomes for East Asian genetic-ancestry group among female samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_eas_jpn		Adjusted alternate-allele frequency in gnomAD exomes for eas jpn group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_eas_kor		Adjusted alternate-allele frequency in gnomAD exomes for eas kor group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_eas_male		Adjusted alternate-allele frequency in gnomAD exomes for East Asian genetic-ancestry group among male samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_eas_oea		Adjusted alternate-allele frequency in gnomAD exomes for eas oea group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_grpmax	2.98686e-05; 0.00017337; 2.23594e-05	Adjusted alternate-allele frequency in gnomAD exomes for the genetic-ancestry group with the highest adjusted allele frequency.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_grpmax_non_ukb	5.65163e-05; 2.28624e-05; 0.00024108	Adjusted alternate-allele frequency in gnomAD exomes for the genetic-ancestry group with the highest adjusted allele frequency after excluding UK Biobank samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_male		Adjusted alternate-allele frequency in gnomAD exomes for male group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_mid	0.0; 0.00017337; 0.0003467409999999	Adjusted alternate-allele frequency in gnomAD exomes for Middle Eastern genetic-ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_nfe	0.0; 8.99271e-07; 1.79854e-06	Adjusted alternate-allele frequency in gnomAD exomes for non-Finnish European genetic-ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_nfe_bgr		Adjusted alternate-allele frequency in gnomAD exomes for nfe bgr group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_nfe_est		Adjusted alternate-allele frequency in gnomAD exomes for nfe est group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_nfe_female		Adjusted alternate-allele frequency in gnomAD exomes for non-Finnish European genetic-ancestry group among female samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_nfe_male		Adjusted alternate-allele frequency in gnomAD exomes for non-Finnish European genetic-ancestry group among male samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_nfe_nwe		Adjusted alternate-allele frequency in gnomAD exomes for nfe nwe group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_nfe_onf		Adjusted alternate-allele frequency in gnomAD exomes for nfe onf group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_nfe_seu		Adjusted alternate-allele frequency in gnomAD exomes for nfe seu group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_non_ukb	0.0; 1.59038e-06; 1.5903899999999996e-06	Adjusted alternate-allele frequency in gnomAD exomes after excluding UK Biobank samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_non_ukb_afr	0.0; 5.65163e-05; 0.0001130329999999	Adjusted alternate-allele frequency in gnomAD exomes for African/African American genetic-ancestry group after excluding UK Biobank samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_non_ukb_amr	0.0; 2.28624e-05; 4.57247e-05	Adjusted alternate-allele frequency in gnomAD exomes for Admixed American genetic-ancestry group after excluding UK Biobank samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_non_ukb_asj	0.0; 4.76554e-05; 4.76599e-05	Adjusted alternate-allele frequency in gnomAD exomes for Ashkenazi Jewish genetic-ancestry group after excluding UK Biobank samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_non_ukb_eas	0.0; 2.77239e-05; 5.54477e-05	Adjusted alternate-allele frequency in gnomAD exomes for East Asian genetic-ancestry group after excluding UK Biobank samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_non_ukb_fin	0.0; 1.88168e-05; 1.88175e-05	Adjusted alternate-allele frequency in gnomAD exomes for Finnish genetic-ancestry group after excluding UK Biobank samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_non_ukb_mid	0.0; 0.00024108; 0.00048216	Adjusted alternate-allele frequency in gnomAD exomes for Middle Eastern genetic-ancestry group after excluding UK Biobank samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_non_ukb_nfe	0.0; 2.85626e-06; 2.85628e-06	Adjusted alternate-allele frequency in gnomAD exomes for non-Finnish European genetic-ancestry group after excluding UK Biobank samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_non_ukb_raw	1.59037e-06; 3.18074e-06; 0.0	Raw alternate-allele frequency in gnomAD exomes for all samples after excluding UK Biobank samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help

gnomAD_exome_AF_non_ukb_remaining	0.0; 3.02133e-05; 3.02151e-05	Adjusted alternate-allele frequency in gnomAD exomes for remaining individuals not assigned to one of the listed genetic-ancestry groups after excluding UK Biobank samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_non_ukb_sas	0.0; 1.43271e-05; 1.43275e-05	Adjusted alternate-allele frequency in gnomAD exomes for South Asian genetic-ancestry group after excluding UK Biobank samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_oth		Adjusted alternate-allele frequency in gnomAD exomes for legacy “other” ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_oth_female		Adjusted alternate-allele frequency in gnomAD exomes for legacy “other” ancestry group among female samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_oth_male		Adjusted alternate-allele frequency in gnomAD exomes for legacy “other” ancestry group among male samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_popmax		Adjusted alternate-allele frequency in gnomAD exomes for the legacy population group with the highest adjusted allele frequency.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_raw	6.84044e-07; 1.36809e-06; 2.73618e-06	Raw alternate-allele frequency in gnomAD exomes for all samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_remaining	0.0; 1.6557400000000002e-05; 1.65579e-05	Adjusted alternate-allele frequency in gnomAD exomes for remaining individuals not assigned to one of the listed genetic-ancestry groups.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_sas	0.0; 1.15931e-05; 1.15934e-05	Adjusted alternate-allele frequency in gnomAD exomes for South Asian genetic-ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_sas_female		Adjusted alternate-allele frequency in gnomAD exomes for South Asian genetic-ancestry group among female samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_AF_sas_male		Adjusted alternate-allele frequency in gnomAD exomes for South Asian genetic-ancestry group among male samples.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_genome_AFR		Alternate-allele frequency in the gnomAD genome dataset for the African/African American genetic-ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_genome_AMR		Alternate-allele frequency in the gnomAD genome dataset for the Admixed American genetic-ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_genome_ASJ		Alternate-allele frequency in the gnomAD genome dataset for the Ashkenazi Jewish genetic-ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_genome_EAS		Alternate-allele frequency in the gnomAD genome dataset for the East Asian genetic-ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_genome_FIN		Alternate-allele frequency in the gnomAD genome dataset for the Finnish genetic-ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_genome_NFE		Alternate-allele frequency in the gnomAD genome dataset for the non-Finnish European genetic-ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_genome_OTH		Alternate-allele frequency in the gnomAD genome dataset for the legacy “other” ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_afr_XX	0.0; 1.8100000000000001e-05; 1.811e-05	gnomAD exome filtering allele-frequency threshold for African/African American genetic-ancestry group in the XX karyotype stratum, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_afr_XY	0.0; 4.946e-05; 4.948e-05	gnomAD exome filtering allele-frequency threshold for African/African American genetic-ancestry group in the XY karyotype stratum, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_amr_XX	0.0; 1.288e-05; 3.092e-05	gnomAD exome filtering allele-frequency threshold for Admixed American genetic-ancestry group in the XX karyotype stratum, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_amr_XY	0.0; 3.662e-05; 3.661e-05	gnomAD exome filtering allele-frequency threshold for Admixed American genetic-ancestry group in the XY karyotype stratum, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_eas_XX	0.0; 1.696e-05; 1.697e-05	gnomAD exome filtering allele-frequency threshold for East Asian genetic-ancestry group in the XX karyotype stratum, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help

gnomAD_exome_faf95_eas_XY	0.0; 3.67e-05; 3.671e-05	gnomAD exome filtering allele-frequency threshold for East Asian genetic-ancestry group in the XY karyotype stratum, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_mid_XX	0.0; 0.00014098; 0.00032597	gnomAD exome filtering allele-frequency threshold for Middle Eastern genetic-ancestry group in the XX karyotype stratum, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_mid_XY	0.0; 0.00021765; 0.00021777	gnomAD exome filtering allele-frequency threshold for Middle Eastern genetic-ancestry group in the XY karyotype stratum, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_nfe_XX	0.0; 5.8e-07; 1.4e-06	gnomAD exome filtering allele-frequency threshold for non-Finnish European genetic-ancestry group in the XX karyotype stratum, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_nfe_XY	0.0; 1.23e-06; 2.96e-06	gnomAD exome filtering allele-frequency threshold for non-Finnish European genetic-ancestry group in the XY karyotype stratum, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_non_ukb_afr_XX	0.0; 3.404e-05; 7.889e-05	gnomAD exome filtering allele-frequency threshold for African/African American genetic-ancestry group in the XX karyotype stratum after excluding UK Biobank samples, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_non_ukb_afr_XY	0.0; 9.53e-05; 0.00022045	gnomAD exome filtering allele-frequency threshold for African/African American genetic-ancestry group in the XY karyotype stratum after excluding UK Biobank samples, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_non_ukb_amr_XX	0.0; 1.322e-05; 3.172e-05	gnomAD exome filtering allele-frequency threshold for Admixed American genetic-ancestry group in the XX karyotype stratum after excluding UK Biobank samples, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_non_ukb_amr_XY	0.0; 3.723e-05; 3.724e-05	gnomAD exome filtering allele-frequency threshold for Admixed American genetic-ancestry group in the XY karyotype stratum after excluding UK Biobank samples, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_non_ukb_eas_XX	0.0; 1.906e-05; 4.435e-05	gnomAD exome filtering allele-frequency threshold for East Asian genetic-ancestry group in the XX karyotype stratum after excluding UK Biobank samples, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_non_ukb_eas_XY	0.0; 3.905e-05; 3.906e-05	gnomAD exome filtering allele-frequency threshold for East Asian genetic-ancestry group in the XY karyotype stratum after excluding UK Biobank samples, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_non_ukb_mid_XX	0.0; 0.00018546; 0.00042619	gnomAD exome filtering allele-frequency threshold for Middle Eastern genetic-ancestry group in the XX karyotype stratum after excluding UK Biobank samples, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help

gnomAD_exome_faf95_non_ukb_mid_XY	0.0; 0.0003176; 0.00031785	gnomAD exome filtering allele-frequency threshold for Middle Eastern genetic-ancestry group in the XY karyotype stratum after excluding UK Biobank samples, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_non_ukb_nfe_XX	0.0; 2.05e-06; 4.92e-06	gnomAD exome filtering allele-frequency threshold for non-Finnish European genetic-ancestry group in the XX karyotype stratum after excluding UK Biobank samples, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_non_ukb_nfe_XY	0.0; 3.54e-06; 8.49e-06	gnomAD exome filtering allele-frequency threshold for non-Finnish European genetic-ancestry group in the XY karyotype stratum after excluding UK Biobank samples, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_non_ukb_sas_XX	0.0; 2.377e-05; 5.564e-05	gnomAD exome filtering allele-frequency threshold for South Asian genetic-ancestry group in the XX karyotype stratum after excluding UK Biobank samples, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_non_ukb_sas_XY	0.0; 1.197e-05; 1.198e-05	gnomAD exome filtering allele-frequency threshold for South Asian genetic-ancestry group in the XY karyotype stratum after excluding UK Biobank samples, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_non_ukb_XX	0.0; 1.16e-06; 2.79e-06	gnomAD exome filtering allele-frequency threshold for XX group after excluding UK Biobank samples, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_non_ukb_XY	0.0; 1.94e-06; 4.66e-06	gnomAD exome filtering allele-frequency threshold for XY group after excluding UK Biobank samples, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_sas_XX	0.0; 1.601e-05; 3.702e-05	gnomAD exome filtering allele-frequency threshold for South Asian genetic-ancestry group in the XX karyotype stratum, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_sas_XY	0.0; 1.032e-05; 1.033e-05	gnomAD exome filtering allele-frequency threshold for South Asian genetic-ancestry group in the XY karyotype stratum, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_XX	0.0; 4.5e-07; 1.09e-06	gnomAD exome filtering allele-frequency threshold for XX group, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf95_XY	0.0; 9.2e-07; 2.2e-06	gnomAD exome filtering allele-frequency threshold for XY group, computed using the 95% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_afr_XX	0.0; 7.39e-06; 2.236e-05	gnomAD exome filtering allele-frequency threshold for African/African American genetic-ancestry group in the XX karyotype stratum, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_afr_XY	0.0; 2.033e-05; 2.034e-05	gnomAD exome filtering allele-frequency threshold for African/African American genetic-ancestry group in the XY karyotype stratum, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help

gnomAD_exome_faf99_amr_XX	0.0; 4.82e-06; 1.656e-05	gnomAD exome filtering allele-frequency threshold for Admixed American genetic-ancestry group in the XX karyotype stratum, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_amr_XY	0.0; 1.492e-05; 1.493e-05	gnomAD exome filtering allele-frequency threshold for Admixed American genetic-ancestry group in the XY karyotype stratum, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_eas_XX	0.0; 6.97e-06; 2.061e-05	gnomAD exome filtering allele-frequency threshold for East Asian genetic-ancestry group in the XX karyotype stratum, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_eas_XY	0.0; 1.495e-05; 4.502e-05	gnomAD exome filtering allele-frequency threshold for East Asian genetic-ancestry group in the XY karyotype stratum, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_mid_XX	0.0; 5.829e-05; 0.0001737	gnomAD exome filtering allele-frequency threshold for Middle Eastern genetic-ancestry group in the XX karyotype stratum, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_mid_XY	0.0; 9.062e-05; 9.066e-05	gnomAD exome filtering allele-frequency threshold for Middle Eastern genetic-ancestry group in the XY karyotype stratum, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_nfe_XX	0.0; 2.2e-07; 3.9e-07	gnomAD exome filtering allele-frequency threshold for non-Finnish European genetic-ancestry group in the XX karyotype stratum, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_nfe_XY	0.0; 4.6e-07; 8.2e-07	gnomAD exome filtering allele-frequency threshold for non-Finnish European genetic-ancestry group in the XY karyotype stratum, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_non_ukb_afr_XX	0.0; 1.396e-05; 4.226e-05	gnomAD exome filtering allele-frequency threshold for African/African American genetic-ancestry group in the XX karyotype stratum after excluding UK Biobank samples, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_non_ukb_afr_XY	0.0; 3.986e-05; 0.00011682	gnomAD exome filtering allele-frequency threshold for African/African American genetic-ancestry group in the XY karyotype stratum after excluding UK Biobank samples, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_non_ukb_amr_XX	0.0; 4.94e-06; 1.678e-05	gnomAD exome filtering allele-frequency threshold for Admixed American genetic-ancestry group in the XX karyotype stratum after excluding UK Biobank samples, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_non_ukb_amr_XY	0.0; 1.515e-05; 1.514e-05	gnomAD exome filtering allele-frequency threshold for Admixed American genetic-ancestry group in the XY karyotype stratum after excluding UK Biobank samples, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_non_ukb_eas_XX	0.0; 7.75e-06; 2.3e-05	gnomAD exome filtering allele-frequency threshold for East Asian genetic-ancestry group in the XX karyotype stratum after excluding UK Biobank samples, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help

gnomAD_exome_faf99_non_ukb_eas_XY	0.0; 1.583e-05; 4.858e-05	gnomAD exome filtering allele-frequency threshold for East Asian genetic-ancestry group in the XY karyotype stratum after excluding UK Biobank samples, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_non_ukb_mid_XX	0.0; 7.686e-05; 0.00022751	gnomAD exome filtering allele-frequency threshold for Middle Eastern genetic-ancestry group in the XX karyotype stratum after excluding UK Biobank samples, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_non_ukb_mid_XY	0.0; 0.0001327; 0.00013279	gnomAD exome filtering allele-frequency threshold for Middle Eastern genetic-ancestry group in the XY karyotype stratum after excluding UK Biobank samples, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_non_ukb_nfe_XX	0.0; 7.7e-07; 2.36e-06	gnomAD exome filtering allele-frequency threshold for non-Finnish European genetic-ancestry group in the XX karyotype stratum after excluding UK Biobank samples, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_non_ukb_nfe_XY	0.0; 1.32e-06; 1.33e-06	gnomAD exome filtering allele-frequency threshold for non-Finnish European genetic-ancestry group in the XY karyotype stratum after excluding UK Biobank samples, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_non_ukb_sas_XX	0.0; 9.5e-06; 2.912e-05	gnomAD exome filtering allele-frequency threshold for South Asian genetic-ancestry group in the XX karyotype stratum after excluding UK Biobank samples, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_non_ukb_sas_XY	0.0; 4.48e-06; 1.496e-05	gnomAD exome filtering allele-frequency threshold for South Asian genetic-ancestry group in the XY karyotype stratum after excluding UK Biobank samples, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_non_ukb_XX	0.0; 4.4e-07; 7.7e-07	gnomAD exome filtering allele-frequency threshold for XX group after excluding UK Biobank samples, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_non_ukb_XY	0.0; 7.3e-07; 2.29e-06	gnomAD exome filtering allele-frequency threshold for XY group after excluding UK Biobank samples, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_sas_XX	0.0; 6.61e-06; 1.897e-05	gnomAD exome filtering allele-frequency threshold for South Asian genetic-ancestry group in the XX karyotype stratum, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_sas_XY	0.0; 3.86e-06; 1.286e-05	gnomAD exome filtering allele-frequency threshold for South Asian genetic-ancestry group in the XY karyotype stratum, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_XX	0.0; 1.7e-07; 3e-07	gnomAD exome filtering allele-frequency threshold for XX group, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_faf99_XY	0.0; 3.4e-07; 6.1e-07	gnomAD exome filtering allele-frequency threshold for XY group, computed using the 99% confidence setting. It is a disease-filtering threshold estimate, not an ordinary allele frequency or a generic “upper confidence bound”.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help

gnomAD_exome_FILTER	AC0; AS_VQSR; AS_VQSR;AC0	Quality-control FILTER status/labels from the gnomAD exome variant record.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_ID	rs56986587; rs11309117; rs3078066	Variant identifier carried by the gnomAD exome record, commonly a dbSNP rs identifier when available.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_monoallelic	False; True; FALSE	gnomAD monoallelic flag from the source VCF. Interpret it according to the exact gnomAD release/header rather than as a generic dominant/recessive property.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomAD_exome_non_par	False; True; FALSE	Flag indicating that a variant on chromosome X or Y lies outside the pseudoautosomal regions.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomad41_genome_AF	1.314e-05; 6.57e-06; 0.0002	Adjusted alternate-allele frequency in the full gnomAD v4.1 genome dataset.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomad41_genome_AF_afr	0.0; 2.413e-05; 2.412e-05	Adjusted alternate-allele frequency in the gnomAD v4.1 genome dataset for the African/African American genetic-ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomad41_genome_AF_ami	0.0; 0.0011; 0.0022	Adjusted alternate-allele frequency in the gnomAD v4.1 genome dataset for the Amish genetic-ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomad41_genome_AF_amr	0.0; 0.0001; 0.0003	Adjusted alternate-allele frequency in the gnomAD v4.1 genome dataset for the Admixed American genetic-ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomad41_genome_AF_asj	0.0; 0.0003; 0.0006	Adjusted alternate-allele frequency in the gnomAD v4.1 genome dataset for the Ashkenazi Jewish genetic-ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomad41_genome_AF_eas	0.0; 0.0002; 0.0004	Adjusted alternate-allele frequency in the gnomAD v4.1 genome dataset for the East Asian genetic-ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomad41_genome_AF_fin	0.0; 0.0002; 0.0003	Adjusted alternate-allele frequency in the gnomAD v4.1 genome dataset for the Finnish genetic-ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomad41_genome_AF_grpmax	0.0002; 1.47e-05; 0.0001	Adjusted alternate-allele frequency in the gnomAD v4.1 genetic-ancestry group with the highest AF for the variant.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomad41_genome_AF_mid	0.0; 0.0034; 0.0068	Adjusted alternate-allele frequency in the gnomAD v4.1 genome dataset for the Middle Eastern genetic-ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomad41_genome_AF_nfe	0.0; 1.47e-05; 2.94e-05	Adjusted alternate-allele frequency in the gnomAD v4.1 genome dataset for the non-Finnish European genetic-ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomad41_genome_AF_raw	6.567e-06; 1.313e-05; 6.566e-06	Raw (unadjusted) alternate-allele frequency in the gnomAD v4.1 genome dataset.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomad41_genome_AF_remaining	0.0; 0.0005; 0.0009	Adjusted alternate-allele frequency among gnomAD v4.1 “remaining individuals” not assigned to one of the listed genetic-ancestry groups.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomad41_genome_AF_sas	0.0; 0.0002; 0.0004	Adjusted alternate-allele frequency in the gnomAD v4.1 genome dataset for the South Asian genetic-ancestry group.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomad41_genome_AF_XX	0.0; 1.285e-05; 1.284e-05	Adjusted alternate-allele frequency in the gnomAD v4.1 genome XX-karyotype stratum; this should not be relabelled simply as female/male.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomad41_genome_AF_XY	0.0; 1.345e-05; 1.346e-05	Adjusted alternate-allele frequency in the gnomAD v4.1 genome XY-karyotype stratum; this should not be relabelled simply as female/male.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomad41_genome_faf95	0.0; 2.18e-06; 5.24e-06	gnomAD v4.1 genome filtering allele-frequency threshold using the 95% confidence setting for the overall eligible dataset. It is a conservative frequency threshold intended for variant filtering, derived with the stated confidence setting.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomad41_genome_faf99	0.0; 8.2e-07; 2.45e-06	gnomAD v4.1 genome filtering allele-frequency threshold using the 99% confidence setting for the overall eligible dataset. It is a conservative frequency threshold intended for variant filtering, derived with the stated confidence setting.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomad41_genome_fafmax_faf95_max	4.88e-06; 0.0002; 1.171e-05	Maximum gnomAD v4.1 genome filtering allele-frequency (FAF) value across the eligible genetic-ancestry groups at the 95% confidence setting; it is not the FAF of the total population.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
gnomad41_genome_fafmax_faf99_max	1.83e-06; 0.0002; 6.25e-06	Maximum gnomAD v4.1 genome filtering allele-frequency (FAF) value across the eligible genetic-ancestry groups at the 99% confidence setting; it is not the FAF of the total population.	https://gnomad.broadinstitute.org/help	https://gnomad.broadinstitute.org/help
DrugBank	Caffeine(DB00201); Collagenase(DB00048); Caffeine(DB00201)|Dantrolene(DB01219)|Suramin(DB04786)	DrugBank drugs associated with the annotated gene or protein target. This gene-level association does not by itself establish that the specific variant predicts response.	https://go.drugbank.com/	local DrugBank XML/CSV export
eQTLGen_snp_id	rs7417106; rs1921; rs113288277	Variant/SNP identifier used for the eQTLGen association record.	https://gtexportal.org/home/documentationPage	https://gtexportal.org/home/documentationPage

GTEx_V8_eQTL_gene	TOLLIP-AS1; AC007966.1|FSIP2-AS1|AC007966.1|FSIP2-AS1|AC007966.1|FSIP2-AS1|AC007966.1|FSI...; B3GALT1	Target gene for a significant GTEx v8 expression quantitative trait locus (eQTL) association involving the variant.	https://gtexportal.org/home/documentationPage	https://gtexportal.org/home/documentationPage
GTEx_V8_eQTL_tissue	Cells_Cultured_fibroblasts; Testis; Thyroid	GTEx v8 tissue in which the significant variant–gene eQTL association was detected.	https://gtexportal.org/home/documentationPage	https://gtexportal.org/home/documentationPage
GTEx_V8_gene		Gene identifier(s) linked to the variant by significant GTEx v8 QTL associations. This does not mean that the variant itself is expressed.	https://gtexportal.org/home/documentationPage	https://gtexportal.org/home/documentationPage
GTEx_V8_sQTL_gene	TOLLIP|TOLLIP; AC007966.1|FSIP2-AS1|FSIP2|AC007966.1; FKBP7	Target gene for a significant GTEx v8 splicing quantitative trait locus (sQTL) association involving the variant.	https://gtexportal.org/home/documentationPage	https://gtexportal.org/home/documentationPage
GTEx_V8_sQTL_tissue	Testis; Cells_Cultured_fibroblasts; Whole_Blood	GTEx v8 tissue in which the significant variant–gene sQTL association was detected.	https://gtexportal.org/home/documentationPage	https://gtexportal.org/home/documentationPage
GTEx_V8_tissue		GTEx v8 tissue(s) in which the corresponding variant–gene QTL association was detected; this is not a list of tissues expressing the variant.	https://gtexportal.org/home/documentationPage	https://gtexportal.org/home/documentationPage
AQ		Systematic-noise score reported by the DRAGEN caller for the alternate allele.	https://help.dragen.illumina.com/dragen-v4.4/product-guide/dragen-v4.4/dragen-dna-pipeline/small-variant-calling	actual DRAGEN VCF header
FractionInformativeReads	0,96	Fraction of reads considered informative by the DRAGEN caller for this variant. It is a caller metric and should not be described as a universal measure of overall sequencing quality.	https://help.dragen.illumina.com/dragen-v4.4/product-guide/dragen-v4.4/dragen-dna-pipeline/small-variant-calling	actual DRAGEN VCF header
MNVTAG		Identifier of the multi-nucleotide variant (MNV) call to which this record belongs, in a caller-defined CHROM:POS_REF->ALT representation.	https://help.dragen.illumina.com/dragen-v4.4/product-guide/dragen-v4.4/dragen-dna-pipeline/small-variant-calling	actual DRAGEN VCF header
SoftClipRatio	0	Ratio of soft-clipped bases to total bases among reads supporting the variant, as reported by the caller.	https://help.dragen.illumina.com/dragen-v4.4/product-guide/dragen-v4.4/dragen-dna-pipeline/small-variant-calling	actual DRAGEN VCF header
SUBSET		1-based ALT-allele index or indices identifying genotyped allele(s) that are strict supersets of this allele, as defined by the upstream caller.	https://help.dragen.illumina.com/dragen-v4.4/product-guide/dragen-v4.4/dragen-dna-pipeline/small-variant-calling	actual DRAGEN VCF header
GermlineStatus	Germline_DB	DRAGEN/TSO500 status used for TMB-oriented origin filtering, such as Germline_DB, Germline_Proxi, Somatic, Somatic_Putative_CH and (in applicable workflows) Somatic_Hotspot. These are algorithmic status predictions intended for the TMB workflow, not definitive constitutional/somatic diagnoses.	https://help.dragen.illumina.com/dragen-v4.4/product-guide/dragen-v4.4/dragen-dna-pipeline/small-variant-calling/somatic-mode#germline-tagging-in-the-tumor-only-pipeline	actual TSO500/DRAGEN output header
OMIM	188840.0; 600185.0; 607585.0	OMIM gene identifier associated with the annotated gene (equivalent in intent to the corresponding HGNC OMIM cross-reference field).	https://hgnc.genenames.org/download/archive/	https://storage.googleapis.com/public-download-files/hgnc/tsv/tsv/hgnc_complete_set.txt
ESCAT	IA; IIA; IB	ESCAT actionability tier assigned by VariantAlker after matching the gene and relevant alteration context (for example protein change, exon, variant type and tissue) against the configured ESCAT table. Empty means no qualifying match.	https://raw.githubusercontent.com/zhanyinx/variantalker/main/bin/add_guidelines_escat_to_funcotator.py	https://www.esmo.org/frameworks-and-tools/esmo-scale-for-clinical-actionability-of-molecular-targets-escat
ESCAT_CANCER	MBC; PC; AML	Cancer indication associated with the ESCAT record selected by VariantAlker.	https://raw.githubusercontent.com/zhanyinx/variantalker/main/bin/add_guidelines_escat_to_funcotator.py	https://www.esmo.org/frameworks-and-tools/esmo-scale-for-clinical-actionability-of-molecular-targets-escat
ESCAT_TISSUE	BREAST; PROSTATE; BONE_MARROW	Tissue/site associated with the ESCAT record selected by VariantAlker.	https://raw.githubusercontent.com/zhanyinx/variantalker/main/bin/add_guidelines_escat_to_funcotator.py	https://www.esmo.org/frameworks-and-tools/esmo-scale-for-clinical-actionability-of-molecular-targets-escat
Prog_list	1821,,1822; ,897,896; ,741	Comma-separated internal CancerVar marker-record indices for prognostic evidence matched to the variant. They are implementation/database row references, not stable external identifiers, and are meaningful only with the exact bundled CancerVar evidence database version.	https://raw.githubusercontent.com/zhanyinx/variantalker/main/resources/CancerVar/CancerVar.py	https://cancervar.wglab.org/
Ref.Gene	TTN; BRCA2; ATM	Gene symbol used by the CancerVar/ANNOVAR branch for the variant.	https://raw.githubusercontent.com/zhanyinx/variantalker/main/resources/CancerVar/CancerVar.py	https://cancervar.wglab.org/

Therap_list	,1843,1824,1819,1844,1823,1826,1825,1831,1830,1837,1833,1839,1829,1827,1834,1...; ,895,912,900,911,898,908,909,910,904,903,907,894,893,901,877,902,899,906; ,7718,7720,7729,7714,7727,7722,7737,7704,7736,7707,7710,7734,7731,7733,7712,7...	Comma-separated internal CancerVar marker-record indices for therapeutic evidence matched to the variant. They are implementation/database row references, not stable external identifiers, and are meaningful only with the exact bundled CancerVar evidence database version.	https://raw.githubusercontent.com/zhanyinx/variantalker/main/resources/CancerVar/CancerVar.py	https://cancervar.wglab.org/
DP	50	Read depth emitted or derived by the upstream caller for the sample/locus. Exact filtering and whether this is INFO-level or FORMAT-level depth depend on the caller; VariantAlker supports DRAGEN, nf-sarek and Ion Torrent inputs.	https://samtools.github.io/hts-specs/VCFv4.3.pdf	actual input VCF header
MQ	244,54	Root-mean-square mapping quality reported for reads at the variant locus.	https://samtools.github.io/hts-specs/VCFv4.3.pdf	actual input VCF header
RPA		Number of tandem-repeat units for each allele, including the reference allele, as a list aligned to the REF/ALT alleles.	https://samtools.github.io/hts-specs/VCFv4.3.pdf	actual input VCF header
RU		Tandem-repeat unit sequence.	https://samtools.github.io/hts-specs/VCFv4.3.pdf	actual input VCF header
STR		Flag indicating that the variant lies in a short tandem repeat.	https://samtools.github.io/hts-specs/VCFv4.3.pdf	actual input VCF header
InterVar: InterVar and Evidence	InterVar: Benign PVS1=0 PS=[0, 0, 0, 0, 0] PM=[1, 0, 0, 0, 0, 0] PP=[0, 0...; InterVar: Benign PVS1=0 PS=[0, 0, 0, 0, 0] PM=[0, 0, 0, 0, 0, 0] PP=[0, 0...; InterVar: Benign PVS1=1 PS=[0, 0, 0, 0, 0] PM=[0, 0, 0, 0, 0, 0] PP=[0, 0...	InterVar aggregate classification and ACMG evidence-code vector.	https://wintervar.wglab.org/	https://github.com/WGLab/InterVar
InterVar	Benign; Uncertain significance; Likely benign	InterVar germline classification under the ACMG/AMP five-class framework: Pathogenic, Likely pathogenic, Uncertain significance, Likely benign or Benign.	https://wintervar.wglab.org/	https://github.com/WGLab/InterVar
DEOGEN2_score	0.004; 0.005; 0.002	DEOGEN2 score from 0 to 1; larger values indicate greater predicted deleteriousness.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
MetaLR_pred	T; D	MetaLR prediction: D (damaging) or T (tolerated), based on the release threshold.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
MetaLR_rankscore	0.00011; 0.04558; 0.0451	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
PROVEAN_converted_rankscore	0.22727; 0.20576; 0.2186	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
PROVEAN_pred	N; D	PROVEAN prediction: D (damaging) or N (neutral), using the release threshold.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
FATHMM_converted_rankscore	0.44461; 0.44065; 0.43672	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
FATHMM_pred	T; D	FATHMM prediction: D (damaging) or T (tolerated), based on the original score/release threshold.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
CADD_raw_rankscore	0.76383; 0.81072; 0.75465	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the CADD raw-score distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
MutationTaster_score	1.0; 0.999999; 0.999998	Original MutationTaster probability-like score. Its interpretation must be combined with MutationTaster_pred (A, D, N or P); direction is not safely inferred from the score alone across all classes.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
SIFT_score	0.0; 0.001; 0.002	Original SIFT score from 0 to 1; smaller values indicate greater predicted deleteriousness (commonly <0.05 for deleterious).	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP

Interpro_domain.1	.\x3b.; .\x3b.\x3b.; .\x3b.\x3b.\x3b.	Functional protein domain(s) associated with the affected protein/position in the secondary InterPro-derived data source; multiple entries may be separated by semicolons.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
Aloft_Confidence	High; High;.; High;.;.	ALoFT confidence associated with the loss-of-function impact/inheritance prediction.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
Aloft_pred	Recessive; Recessive;.; .;Recessive	ALoFT prediction for a putative loss-of-function variant/transcript: Recessive, Dominant or Tolerant (source-version vocabulary).	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
BayesDel_addAF_pred	T; D	BayesDel prediction including allele-frequency features: D (deleterious) or T (tolerated), using the dbNSFP release threshold.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
BayesDel_addAF_rankscore	0.99412; 0.98066; 0.98769	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
BayesDel_addAF_score	0.625005; 0.618043; 0.622132	BayesDel ensemble deleteriousness score including allele-frequency features; larger values indicate greater predicted deleteriousness.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
BayesDel_noAF_pred	T; D	BayesDel prediction excluding allele-frequency features: D (deleterious) or T (tolerated), using the dbNSFP release threshold.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
BayesDel_noAF_rankscore	0.99401; 0.98035; 0.98747	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
BayesDel_noAF_score	0.66; 0.65; 0.655873	BayesDel ensemble deleteriousness score excluding allele-frequency features; larger values indicate greater predicted deleteriousness.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
bStatistic	970.0; 940.0; 976.0	Background-selection (B) statistic from CADD/dbNSFP, scaled approximately 0–1000. Lower values indicate stronger expected loss of neutral diversity from background selection; values near 1000 indicate little/no background selection.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
bStatistic_converted_rankscore	0.06235; 0.13648; 0.04745	dbNSFP rankscore after converting bStatistic to -bStatistic; larger values therefore correspond to stronger background-selection signal (lower original B statistic).	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
ClinPred_pred	D; T	ClinPred binary prediction: D (deleterious) or T (tolerated), using the dbNSFP release threshold.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
ClinPred_rankscore	0.99146; 0.9909; 0.99363	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
ClinPred_score	1.0; 0.0002953502280559; 0.999809563159943	ClinPred score from 0 to 1; larger values indicate greater predicted pathogenicity/deleteriousness.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
DANN_rankscore	0.98852; 0.98721; 0.98917	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
DANN_score	0.9268219051870434; 0.9988428841597063; 0.9344018693407072	DANN functional-impact score from 0 to 1; larger values indicate a higher predicted probability of a damaging effect.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
dbscSNV_ADA_SCORE	1.0; 0.0; 0.0001	dbscSNV AdaBoost score for predicted splice-altering impact of an SNV near a splice junction; larger values indicate stronger splice-impact prediction.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
dbscSNV_RF_SCORE	0.0; 0.002; 0.938	dbscSNV random-forest score for predicted splice-altering impact of an SNV near a splice junction; larger values indicate stronger splice-impact prediction.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
DEOGEN2_pred	T; D	DEOGEN2 prediction: D (deleterious) or T (tolerated).	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
DEOGEN2_rankscore	0.02402; 0.0306; 0.02667	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
Eigen-PC-phred_coding	10.771; 10.669; 10.018	Phred-scaled Eigen-PC functional score for coding SNVs; larger values indicate greater predicted functional/deleterious impact.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP

Eigen-PC-raw_coding	1.12357875994551; 1.00967386719567; 1.14230968258289	Eigen-PC raw score for coding SNVs, a principal-component-adjusted unsupervised functional score; larger values indicate greater predicted functional/deleterious impact.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
Eigen-PC-raw_coding_rankscore	0.82322; 0.78636; 0.82694	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
Eigen-phred_coding	12.619; 1.998; 1.179	Phred-scaled Eigen functional score for coding SNVs; larger values indicate greater predicted functional/deleterious impact.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
Eigen-raw_coding	1.21748086594471; 1.07572918353069; 1.26225632023751	Eigen raw unsupervised functional score for coding SNVs; larger values indicate greater predicted functional/deleterious impact.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
Eigen-raw_coding_rankscore	0.74848; 0.77282; 0.82018	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
ESM1b_pred	T; D	dbNSFP binary ESM-1b prediction derived using the release threshold (D=deleterious, T=tolerated); the original ESM-1b authors did not prescribe a universal clinical threshold.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
ESM1b_rankscore	0.24549; 0.27112; 0.20432	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
ESM1b_score	-3.775; -3.918; -3.731	ESM-1b protein-language-model log-likelihood-ratio score; more negative values indicate a greater predicted pathogenic effect. Interpret transcript-specific multiple values with their matching transcript IDs.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
EVE_rankscore	0.12352; 0.12714; 0.1256	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
EVE_score	0.1249725419417923; 0.1251981576316652; 0.1262599386802879	EVE evolutionary-model score from 0 to 1; larger values indicate greater predicted pathogenicity for a missense variant.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
FATHMM_score	0.97; 0.93; 0.95	Original FATHMM score; lower/more negative values indicate a greater predicted damaging effect (commonly ≤−1.5 for damaging in this dbNSFP release).	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
fathmm-MKL_coding_group	AEFBI; AEFDBI; AEFDBCI	Feature-group labels (A–J) used by the fathmm-MKL coding model; consult the method documentation for the group definitions.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
fathmm-MKL_coding_pred	D; N	fathmm-MKL coding prediction: D (damaging) or N (neutral), typically using a score threshold of 0.5 in this release.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
fathmm-MKL_coding_rankscore	0.00017; 0.91343; 0.90087	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
fathmm-MKL_coding_score	0.0; 0.99108; 0.99022	fathmm-MKL coding deleteriousness score, ranging from 0 to 1; values above the documented threshold are predicted deleterious. It is not a p-value.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
fathmm-XF_coding_pred	N; D	fathmm-XF coding prediction: D for damaging or N for neutral.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
fathmm-XF_coding_rankscore	0.03337; 0.99996; 0.01776	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
fathmm-XF_coding_score	0.031355; 0.925375; 0.024797	fathmm-XF coding score from 0 to 1; larger values indicate greater predicted deleteriousness.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
GenoCanyon_rankscore	0.74766; 0.98316; 0.51787	Rank-normalised GenoCanyon score in dbNSFP, computed relative to other GenoCanyon scores; larger rankscore values correspond to larger source scores.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP

GenoCanyon_score	1.0; 0.999999999999999; 0.999999999999998	GenoCanyon functional-potential score for a genomic position, derived from conservation and biochemical annotations; larger values indicate greater predicted functionality.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
GERP++_NR	6.17; 5.65; 5.55	GERP++ neutral evolutionary rate at the genomic position.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
GERP++_RS	6.17; 5.65; 5.63	GERP++ rejected-substitutions score; larger positive values indicate stronger evolutionary constraint/conservation.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
GERP++_RS_rankscore	0.99707; 0.86881; 0.86108	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more evolutionarily constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
GM12878_confidence_value	0.0; 1.0; 2.0	fitCons confidence category for the epigenomic fingerprint at this position: 0 highly significant, 1 significant, 2 informative, 3 other (approximate p-value ranges are source-defined).	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
GM12878_fitCons_rankscore	0.89359; 0.26702; 0.74545	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the GM12878 fitCons distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
GM12878_fitCons_score	0.724815; 0.573888; 0.702456	fitCons score for the GM12878 cell type, estimating the fraction of positions in the same epigenomic functional class that are under selective pressure; larger values indicate greater inferred functional importance.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
gMVP_rankscore	0.99989; 0.99999; 0.99998	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
gMVP_score	0.068169099379579; 0.0571118262807137; 0.187271898306529	gMVP missense pathogenicity score from 0 to 1; larger values indicate greater predicted pathogenicity.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
H1-hESC_confidence_value	0.0; 1.0; 2.0	fitCons confidence category for the epigenomic fingerprint at this position: 0 highly significant, 1 significant, 2 informative, 3 other (approximate p-value ranges are source-defined).	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
H1-hESC_fitCons_rankscore	0.87919; 0.75317; 0.68683	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the H1-hESC fitCons distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
H1-hESC_fitCons_score	0.724815; 0.709663; 0.702456	fitCons score for the H1-hESC cell type, estimating the fraction of positions in the same epigenomic functional class that are under selective pressure; larger values indicate greater inferred functional importance.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
HUVEC_confidence_value	0.0; 1.0; 2.0	fitCons confidence category for the epigenomic fingerprint at this position: 0 highly significant, 1 significant, 2 informative, 3 other (approximate p-value ranges are source-defined).	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
HUVEC_fitCons_rankscore	0.83352; 0.26826; 0.41899	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the HUVEC fitCons distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
HUVEC_fitCons_score	0.714379; 0.564101; 0.613276	fitCons score for the HUVEC cell type, estimating the fraction of positions in the same epigenomic functional class that are under selective pressure; larger values indicate greater inferred functional importance.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
integrated_confidence_value	0.0; 1.0; 2.0	Confidence category for the integrated fitCons annotation: 0 highly significant, 1 significant, 2 informative, 3 other (approximate p-value ranges are source-defined).	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
integrated_fitCons_rankscore	0.73137; 0.92422; 0.61202	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the integrated fitCons distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
integrated_fitCons_score	0.706548; 0.732398; 0.706298	Integrated fitCons score across GM12878, H1-hESC and HUVEC epigenomic fingerprints; larger values indicate a greater estimated fraction of functionally constrained positions in the assigned class.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP

Interpro_domain	GPCR, rhodopsin-like, 7TM; GPCR, rhodopsin-like, 7TM;GPCR, rhodopsin-like, 7TM; .;.;Fibronectin type III|Fibronectin type III|Fibronectin type III|Fibronecti...	InterPro domain or conserved-site annotation overlapping the affected protein position. Multiple transcript/protein-specific values may be separated by semicolons; any counts/parentheses follow the dbNSFP release encoding.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
LINSIGHT	0.046875; 0.0489367; 0.0414318	LINSIGHT probability-like score for negative selection at a non-coding genomic site; larger values indicate stronger predicted selection/functional importance.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
LINSIGHT_rankscore	0.08445; 0.06002; 0.01019	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more negatively selected/functional end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
LIST-S2_pred	T; D	LIST-S2 prediction: D (deleterious) or T (tolerated), using the dbNSFP release threshold.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
LIST-S2_rankscore	0.65058; 0.61998; 0.6628	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
LIST-S2_score	0.9; 0.889011; 0.915608	LIST-S2 missense deleteriousness score; larger values indicate greater predicted deleteriousness.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
LRT_converted_rankscore	0.8433; 0.62929; 0.55875	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
LRT_Omega	0.0; 0.061993; 0.064498	Estimated nonsynonymous-to-synonymous substitution-rate ratio (omega) reported by the LRT method.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
LRT_pred	D; N; U	LRT prediction: D (deleterious), N (neutral) or U (unknown); it is not determined solely by a simple score cutoff.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
LRT_score	0.0; 1e-06; 2e-06	Original LRT two-sided p-value in dbNSFP. Lower values provide stronger evidence against neutrality and therefore greater evidence of a deleterious effect.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
M-CAP_pred	D; T	M-CAP prediction: D (deleterious) or T (tolerated), using the dbNSFP release threshold.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
M-CAP_rankscore	0.15101; 0.10121; 0.13001	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
M-CAP_score	0.004; 0.005; 0.006	M-CAP score for missense variants; larger values indicate greater predicted pathogenicity (the original method commonly uses 0.025 as the high-sensitivity threshold).	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
MetaLR_score	0.015; 0.011; 0.013	MetaLR ensemble score; larger values indicate greater predicted deleteriousness.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
MetaRNN_pred	T; D	MetaRNN prediction: D (deleterious) or T (tolerated), using the dbNSFP release threshold.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
MetaRNN_rankscore	3e-05; 0.00198; 0.00187	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
MetaRNN_score	0.0826554; 0.98437643; 0.122805506	MetaRNN pathogenicity score; larger values indicate greater predicted pathogenicity/deleteriousness.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
MetaSVM_pred	T; D	MetaSVM prediction: D (damaging) or T (tolerated), based on the release threshold.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
MetaSVM_rankscore	0.21005; 0.23747; 0.21168	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
MetaSVM_score	1.097; 1.095; 1.099	MetaSVM ensemble score; larger/more positive values indicate greater predicted deleteriousness.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP

MPC_rankscore	0.97175; 0.97893; 0.93298	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
MPC_score	2.40020819914; 2.55136621799; 1.94467123622	MPC (Missense badness, PolyPhen-2 and Constraint) score; larger values indicate greater predicted missense deleteriousness. The scale is not limited to 0–1.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
MutationAssessor_pred	M; L; N	MutationAssessor impact class: H (high), M (medium), L (low) or N (neutral).	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
MutationAssessor_rankscore	0.06538; 0.11182; 0.14455	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
MutationAssessor_score	0.0; 0.345; 0.55	MutationAssessor functional-impact score; larger values indicate greater predicted functional impact.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
MutationTaster_converted_rankscore	0.81001; 0.08975; 0.58761	dbNSFP rankscore derived after converting the MutationTaster score according to its prediction class; larger values indicate the more damaging end of the converted distribution.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
MutationTaster_pred	D; N; A	MutationTaster prediction: A (disease_causing_automatic), D (disease_causing), N (polymorphism) or P (polymorphism_automatic).	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
MutPred_rankscore	0.45709; 0.43573; 0.43738	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
MutPred_score	0.417; 0.404; 0.405	MutPred score; larger values indicate greater predicted probability of a deleterious/disease-associated effect.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
MVP_rankscore	0.03247; 0.061; 0.0136	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
MVP_score	0.043077524339; 0.0675242888579; 0.0297737177859	MVP missense pathogenicity score; larger values indicate greater predicted pathogenicity.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
phastCons100way_vertebrate	1.0; 0.0; 0.001	phastCons conservation probability based on the 100-vertebrate alignment; values are generally 0–1, with larger values indicating greater conservation.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
phastCons100way_vertebrate_rankscore	0.71638; 0.06391; 0.13787	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more conserved end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
phastCons17way_primate	0.998; 0.999; 1.0	phastCons conservation probability based on the 17-primate alignment; values are generally 0–1, with larger values indicating greater conservation.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
phastCons17way_primate_rankscore	0.85391; 0.91618; 0.97212	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more conserved end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
phastCons30way_mammalian	0,395	phastCons conservation probability based on the 30-mammal alignment; values are generally 0–1, with larger values indicating greater conservation.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
phastCons30way_mammalian_rankscore	0,256	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more conserved end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
phastCons470way_mammalian	1.0; 0.0; 0.999	phastCons conservation probability based on the 470-mammal alignment; values are generally 0–1, with larger values indicating greater conservation.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
phastCons470way_mammalian_rankscore	0.68203; 0.08366; 0.35428	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more conserved end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
phyloP100way_vertebrate	7.905; 10.003; 9.325	phyloP conservation/acceleration score based on the 100-vertebrate alignment; positive values indicate conservation and negative values indicate acceleration relative to neutrality, with larger values more conserved.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP

phyloP100way_vertebrate_rankscore	0.86479; 0.99689; 0.96006	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more conserved end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
phyloP17way_primate	0.599; 0.676; 0.665	phyloP conservation/acceleration score based on the 17-primate alignment; positive values indicate conservation and negative values indicate acceleration relative to neutrality, with larger values more conserved.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
phyloP17way_primate_rankscore	0.4025; 0.7674; 0.62972	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more conserved end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
phyloP30way_mammalian	-0,147	phyloP conservation/acceleration score based on the 30-mammal alignment; positive values indicate conservation and negative values indicate acceleration relative to neutrality, with larger values more conserved.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
phyloP30way_mammalian_rankscore	0,111	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more conserved end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
phyloP470way_mammalian	-20.0; 7.662; 7.675	phyloP conservation/acceleration score based on the 470-mammal alignment; positive values indicate conservation and negative values indicate acceleration relative to neutrality, with larger values more conserved.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
phyloP470way_mammalian_rankscore	0.00162; 0.6429; 0.64017	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more conserved end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
Polyphen2_HDIV_pred	B; D; P	PolyPhen-2 HumDiv class: D (probably damaging), P (possibly damaging) or B (benign), using the release thresholds.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
Polyphen2_HDIV_rankscore	0.90584; 0.02946; 0.77913	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
Polyphen2_HDIV_score	1.0; 0.0; 0.999	PolyPhen-2 HumDiv probability score from 0 to 1; larger values indicate greater predicted damage.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
Polyphen2_HVAR_pred	B; D; P	PolyPhen-2 HumVar class: D (probably damaging), P (possibly damaging) or B (benign), using the release thresholds.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
Polyphen2_HVAR_rankscore	0.97372; 0.92359; 0.04355	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
Polyphen2_HVAR_score	0.001; 1.0; 0.0	PolyPhen-2 HumVar probability score from 0 to 1; larger values indicate greater predicted damage.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
PrimateAI_pred	T; D	PrimateAI prediction: D (deleterious) or T (tolerated), using the dbNSFP release threshold.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
PrimateAI_rankscore	0.03626; 0.02946; 0.02491	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
PrimateAI_score	0.507606744766; 0.297095894814; 0.589073061943	PrimateAI missense score; larger values indicate greater predicted pathogenicity/deleteriousness.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
PROVEAN_score	-0.83; -0.79; -0.73	Original PROVEAN score; smaller/more negative values indicate a greater predicted damaging effect (commonly ≤−2.5 for damaging).	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
Reliability_index	10.0; 9.0; 8.0	Number of observed component predictor scores used by MetaSVM/MetaLR (excluding the maximum 1000 Genomes frequency component), typically 1–10; larger values indicate fewer imputed/missing components and therefore greater model reliability.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
REVEL_rankscore	0.11227; 0.08419; 0.09122	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP

REVEL_score	0.042; 0.034; 0.036	REVEL ensemble score from 0 to 1 for missense variants; larger values indicate greater predicted pathogenicity.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
SIFT_converted_rankscore	0.91255; 0.7849; 0.72154	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
SIFT_pred	D; T	SIFT prediction: D (deleterious) or T (tolerated).	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
SIFT4G_converted_rankscore	0.92824; 0.83351; 0.79402	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
SIFT4G_pred	T; D	SIFT4G prediction: D (deleterious) or T (tolerated).	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
SIFT4G_score	0.0; 0.001; 1.0	SIFT4G score for a missense substitution; lower scores indicate a greater likelihood of a deleterious effect.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
SiPhy_29way_logOdds	20.879; 16.822; 20.86	SiPhy 29-mammal log-odds score comparing a constrained (purifying-selection) model with a neutral model; larger values indicate stronger conservation evidence.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
SiPhy_29way_logOdds_rankscore	0.9987; 0.85699; 0.99763	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more conserved end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
SiPhy_29way_pi	0.0:1.0:0.0:0.0; 0.0:0.0:1.0:0.0; 1.0:0.0:0.0:0.0	Estimated stationary nucleotide distribution (A, C, G and T) at the site from the SiPhy 29-mammal model.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
VARITY_ER_LOO_rankscore	0.07967; 0.04959; 0.06897	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
VARITY_ER_LOO_score	0.114046186; 0.08271042; 0.045069125	VARITY_ER leave-one-variant-out pathogenicity score for extremely rare missense variants; larger values indicate greater predicted pathogenicity.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
VARITY_ER_rankscore	0.06898; 0.07968; 0.07048	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
VARITY_ER_score	0.114046186; 0.08271042; 0.06319954	VARITY_ER pathogenicity score for extremely rare missense variants, from 0 to 1; larger values indicate greater predicted pathogenicity.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
VARITY_R_LOO_rankscore	0.0; 2e-05; 3e-05	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
VARITY_R_LOO_score	0.97404546; 0.90579534; 0.7982764	VARITY_R leave-one-variant-out pathogenicity score for rare missense variants; larger values indicate greater predicted pathogenicity.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
VARITY_R_rankscore	0.0; 3e-05; 2e-05	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
VARITY_R_score	0.97404546; 0.8279641; 0.96526915	VARITY_R pathogenicity score for rare missense variants, from 0 to 1; larger values indicate greater predicted pathogenicity.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
VEST4_rankscore	0.1134; 0.08088; 0.09349	dbNSFP rank-normalized score (approximately 0–1). Larger values place the observation toward the more damaging/pathogenic or more constrained end of the method-specific distribution; interpret with the corresponding raw score and dbNSFP version.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
VEST4_score	0.093; 0.098; 0.099	VEST4 score from 0 to 1; larger values indicate a greater predicted probability of functional impact/disease association. Multiple transcript-specific values may be separated by semicolons.	https://www.dbnsfp.org/	https://sites.google.com/site/jpopgen/dbNSFP
HGNC_Accession_Numbers	X90568; U43746; AB209133	INSDC (GenBank/ENA/DDBJ) nucleotide-sequence accession number(s) associated with the HGNC gene record.	https://www.genenames.org/download/statistics-and-files/	https://hgnc.genenames.org/download/archive/

HGNC_CCDS_IDs	CCDS54424, CCDS54422, CCDS59435, CCDS54421, CCDS54423, CCDS33337; CCDS9344; CCDS86245, CCDS31669	Consensus Coding Sequence (CCDS) identifier(s) associated with the HGNC gene record.	https://www.genenames.org/download/statistics-and-files/	https://hgnc.genenames.org/download/archive/
HGNC_Chromosome	16p13.3; 2q31.2; 19p13.2	Cytogenetic location of the gene (for example 2p23.2-p23.1).	https://www.genenames.org/download/statistics-and-files/	https://hgnc.genenames.org/download/archive/
HGNC_Date_Modified	2023-01-20; 2023-03-15; 2023-07-17	Date on which the HGNC record was last modified.	https://www.genenames.org/download/statistics-and-files/	https://hgnc.genenames.org/download/archive/
HGNC_Date_Name_Changed	2015-11-18; 2016-02-04; 2015-11-13	Date on which the HGNC-approved gene name was last changed.	https://www.genenames.org/download/statistics-and-files/	https://hgnc.genenames.org/download/archive/
HGNC_Date_Symbol_Changed	2015-03-03; 2013-05-09; 2005-11-24	Date on which the HGNC-approved gene symbol was last changed.	https://www.genenames.org/download/statistics-and-files/	https://hgnc.genenames.org/download/archive/
HGNC_Ensembl_Gene_ID	ENSG00000155657; ENSG00000139618; ENSG00000149311	Ensembl gene identifier associated with the HGNC record.	https://www.genenames.org/download/statistics-and-files/	https://hgnc.genenames.org/download/archive/
HGNC_Enzyme_IDs	2.7.10.1; 2.7.11.1; 2.1.1.364	Enzyme Commission (EC) accession number(s), when applicable.	https://www.genenames.org/download/statistics-and-files/	https://hgnc.genenames.org/download/archive/
HGNC_Gene_Family_ID	28; 1492; 1668|593|555	HGNC gene-group/family identifier(s) assigned to the gene.	https://www.genenames.org/download/statistics-and-files/	https://hgnc.genenames.org/download/archive/
HGNC_Gene_Family_Name	Zinc fingers C2H2-type; Armadillo like helical domain containing; Myosin light chain kinase family|I-set domain containing|Fibronectin type III...	HGNC gene-group/family name(s) assigned to the gene.	https://www.genenames.org/download/statistics-and-files/	https://hgnc.genenames.org/download/archive/
HGNC_HGNC_ID	HGNC:12403; HGNC:1101; HGNC:795	Stable HGNC identifier in the form HGNC:<number>.	https://www.genenames.org/download/statistics-and-files/	https://hgnc.genenames.org/download/archive/
HGNC_Locus_Group	protein-coding gene; other; non-coding RNA	HGNC locus-group category grouping related locus types (for example protein-coding gene, non-coding RNA, pseudogene or other).	https://www.genenames.org/download/statistics-and-files/	https://hgnc.genenames.org/download/archive/
HGNC_Locus_Type	gene with protein product; complex locus constituent; RNA, long non-coding	Specific HGNC locus type (for example gene with protein product, RNA long non-coding, RNA micro, RNA transfer or a pseudogene subtype).	https://www.genenames.org/download/statistics-and-files/	https://hgnc.genenames.org/download/archive/
HGNC_Name_Synonyms	BRCA1/BRCA2-containing complex, subunit 2; TEL1, telomere maintenance 1, homolog (S. cerevisiae); neurofibromatosis, "von Recklinghausen disease", "Watson disease"	Alias/alternative gene names recorded by HGNC.	https://www.genenames.org/download/statistics-and-files/	https://hgnc.genenames.org/download/archive/
HGNC_OMIM_ID(supplied_by_OMIM)	188840.0; 600185.0; 611731.0	OMIM identifier(s) supplied by OMIM for the HGNC gene record.	https://www.genenames.org/download/statistics-and-files/	https://hgnc.genenames.org/download/archive/
HGNC_Previous_Name	cardiomyopathy, dilated 1G (autosomal dominant); Fanconi anemia, complementation group D1, "breast cancer 2, early onset", "br...; ataxia telangiectasia mutated (includes complementation groups A, C and D), "...	Gene name(s) previously approved by HGNC.	https://www.genenames.org/download/statistics-and-files/	https://hgnc.genenames.org/download/archive/
HGNC_Previous_Symbols	CMD1G; FANCD1, FACD, FANCD; ATA, ATDC, ATC, ATD	Gene symbol(s) previously approved by HGNC.	https://www.genenames.org/download/statistics-and-files/	https://hgnc.genenames.org/download/archive/
HGNC_Pubmed_IDs	2129545, 10051295; 8091231, 7581463, 15057823; 7792600, 7545545, 8589678	PubMed identifiers linked to the HGNC gene record. They are gene-record references and are not necessarily publications about the specific variant.	https://www.genenames.org/download/statistics-and-files/	https://hgnc.genenames.org/download/archive/
HGNC_RefSeq(supplied_by_NCBI)	NM_001256850; NM_000059; NM_000051	RefSeq nucleotide accession(s) supplied by NCBI for the HGNC gene record.	https://www.genenames.org/download/statistics-and-files/	https://hgnc.genenames.org/download/archive/
HGNC_RefSeq_IDs	NM_133378; NM_000059; NM_000051	RefSeq nucleotide accession(s) curated/recorded by HGNC for the gene; this field does not define a universal “main transcript”.	https://www.genenames.org/download/statistics-and-files/	https://hgnc.genenames.org/download/archive/
HGNC_Status	Approved	Status of the HGNC symbol report, principally Approved or Entry Withdrawn in the complete-set schema.	https://www.genenames.org/download/statistics-and-files/	https://hgnc.genenames.org/download/archive/
HGNC_Synonyms	CMPD4, FLJ32040, TMD, CMH9, LGMD2J, MYLK5; FAD, FAD1, BRCC2, XRCC11; TEL1, TELO1	Alias/alternative gene symbols recorded by HGNC.	https://www.genenames.org/download/statistics-and-files/	https://hgnc.genenames.org/download/archive/

HGNC_UniProt_ID(supplied_by_UniProt)	Q8WZ42; P51587; Q13315	UniProt protein accession(s) supplied by UniProt for the HGNC gene record.	https://www.genenames.org/download/statistics-and-files/	https://hgnc.genenames.org/download/archive/
HGNC_Vega_ID	OTTHUMG00000154448; OTTHUMG00000017411; OTTHUMG00000166480	Vega gene identifier associated with the HGNC record.	https://www.genenames.org/download/statistics-and-files/	https://hgnc.genenames.org/download/archive/
Pathway	~ko04151 PI3K-Akt signaling pathway~ko04510 Focal adhesion~ko04512 ECM-recept...; ~ko01524 Platinum drug resistance~ko03430 Mismatch repair~ko05200 Pathways in...; ~ko05202 Transcriptional misregulation in cancer ~	KEGG pathway identifier(s) and name(s) associated with the annotated gene in the CancerVar data source.	https://www.genome.jp/kegg/pathway.html	https://www.genome.jp/kegg/pathway.html
ClinVar_ASSEMBLY		Reference genome assembly associated with the legacy ClinVar/ANNOVAR record.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_HGMD_ID		HGMD cross-reference identifier carried by the legacy ClinVar datasource, when present.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_rs		dbSNP rs identifier associated with the ClinVar record.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_SYM		Gene symbol associated with the ClinVar record.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_TYPE		Variant type reported by the legacy ClinVar/ANNOVAR annotation source.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_AF_ESP	8e-05; 0.00015; 0.0002299999999999	Alternate-allele frequency from the NHLBI Exome Sequencing Project as carried in the ClinVar VCF/datasource.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_AF_EXAC	1e-05; 2e-05; 3e-05	Alternate-allele frequency from ExAC as carried in the ClinVar VCF/datasource.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_AF_TGP	0.0002; 0.0004; 0.0006	Alternate-allele frequency from the 1000 Genomes Project (TGP) as carried in the ClinVar VCF/datasource.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_ALLELEID	3544463.0; 2193183.0; 4039319.0	ClinVar Allele ID for the allele represented by the record.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_CLNDISDB	MedGen:CN169374; MedGen:C3661900; MeSH:D030342,MedGen:C0950123	Condition database identifiers associated with the aggregate germline classification, serialized as database:identifier values; multiple conditions are separated according to the ClinVar VCF encoding.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_CLNDISDBINCL	MedGen:CN169374; MedGen:CN297638; MONDO:MONDO:0018998,MeSH:D057130,MedGen:C0339527,OMIM:PS204000,Orphanet:65	Condition database identifiers for a haplotype/genotype record that includes this variant rather than for the single variant alone.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_CLNDN	not_specified; not_provided; Inborn_genetic_diseases	ClinVar preferred condition name(s) associated with the germline classification; ordering corresponds to the condition identifier groups in CLNDISDB.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_CLNDNINCL	not_specified; Vancomycin_response; Leber_congenital_amaurosis	ClinVar preferred condition name(s) for a haplotype/genotype record that includes this variant.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_CLNHGVS	NC_000001.11:g.66927del; NC_000001.11:g.69134A>G; NC_000001.11:g.69308A>G	Representative ClinVar HGVS expression for the allele/Variation ID, commonly a genomic expression on the stated assembly.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_CLNREVSTAT	criteria_provided,_single_submitter; criteria_provided,_multiple_submitters,_no_conflicts; criteria_provided,_conflicting_classifications	ClinVar review status for the aggregate classification, using ClinVar controlled vocabulary. It reports the review/consensus process and should not be interpreted as a direct quantitative measure of biological evidence.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_CLNSIG	Uncertain_significance; Likely_benign; Benign	Aggregate germline clinical significance/classification for the single variant in ClinVar; multiple condition-specific values may be separated by "|".	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/

ClinVar_VCF_CLNSIGCONF	Uncertain_significance(1)|Likely_benign(1); Uncertain_significance(2)|Likely_benign(1); Uncertain_significance(1)|Likely_benign(2)	Details of conflicting germline clinical-significance submissions for the single variant, when present.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_CLNSIGINCL	1723892:drug_response; 440728:risk_factor; 5345:Benign	Clinical significance for a haplotype or genotype that includes this variant, rather than for the single variant alone; encoded using ClinVar's included-record syntax.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_CLNSIGSCV	SCV005419006; SCV003526545; SCV006120601	SCV accession(s) for submissions contributing to the aggregate germline classification; multiple values may be separated by "|".	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_CLNVC	single_nucleotide_variant; Deletion; Duplication	ClinVar variant class/type for the allele (for example single_nucleotide_variant).	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_ID	3385321.0; 2205837.0; 3925305.0	ClinVar Variation ID.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_MC	SO:0001583|missense_variant; SO:0001627|intron_variant; SO:0001819|synonymous_variant	Molecular consequence annotation(s), encoded as Sequence Ontology ID|term pairs; multiple consequences use the ClinVar VCF delimiter for that release.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_ONC	Likely_oncogenic; no_classification_for_the_single_variant; Uncertain_significance	Aggregate ClinVar oncogenicity classification for the single variant; multiple values may be separated by "|".	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_ONCDN	Neoplasm; Meningioma; Gastrointestinal_stromal_tumor	ClinVar preferred condition name(s) corresponding to ONCDISDB for oncogenicity classifications.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_ONCREVSTAT	criteria_provided,_single_submitter; no_classification_for_the_single_variant; no_assertion_criteria_provided	ClinVar review status for the aggregate oncogenicity classification.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_ONCSCV	SCV005094042; SCV005871698; SCV005094153	SCV accession(s) contributing to the aggregate oncogenicity classification.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_RS	56986587.0; 11309117.0; 193922936.0	dbSNP rs identifier number associated with the ClinVar allele (typically without the "rs" prefix in the VCF value).	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_SCI	Tier_I_-_Strong; no_classification_for_the_single_variant; Tier_II_-_Potential	Aggregate ClinVar somatic clinical-impact classification for the single variant; multiple values may be separated by "|".	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_SCIDISDB	Human_Phenotype_Ontology:HP:0006743,MONDO:MONDO:0009993,MedGen:C0206656,Orpha...; MONDO:MONDO:0957196,MedGen:C4289688; MONDO:MONDO:0850196,MedGen:C4331965	Condition database identifier(s) submitted for somatic clinical-impact classifications.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_SCIDN	Embryonal_rhabdomyosarcoma; Diffuse_midline_glioma,_H3_K27M-mutant; Medulloblastoma_WNT_activated	ClinVar preferred condition name(s) corresponding to SCIDISDB for somatic clinical-impact classifications.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
ClinVar_VCF_SCISCV	SCV007104356; SCV007105254; SCV007105256	SCV accession(s) contributing to the aggregate somatic clinical-impact classification.	https://www.ncbi.nlm.nih.gov/clinvar	https://www.ncbi.nlm.nih.gov/clinvar/docs/maintenance_use/
Refseq_mRNA_Id	NM_001267550.2; NM_000059.4; NM_000051.4	Versioned NCBI RefSeq transcript accession for the annotated RNA (for example NM_ for curated protein-coding transcripts, NR_ for curated non-coding transcripts, XM_/XR_ for predicted models).	https://www.ncbi.nlm.nih.gov/refseq/about/	https://www.ncbi.nlm.nih.gov/refseq/about/
Refseq_prot_Id	NP_001254479.2; NP_000050.3; NP_000042.3	Versioned NCBI RefSeq protein accession corresponding to a coding transcript (typically NP_ for curated proteins or XP_ for predicted proteins).	https://www.ncbi.nlm.nih.gov/refseq/about/	https://www.ncbi.nlm.nih.gov/refseq/about/
dbSNP_ASS	false; False; true	Flag indicating an acceptor splice-site annotation.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/

dbSNP_CLNACC	.,.,RCV002663130.2,.,RCV002184323.12,RCV001393531.5,RCV002111771.5,RCV0021079...; .,RCV004013714.2,RCV003321087.1,.,RCV001262895.1|RCV002256726.1,.,RCV00058180...; .,.,.,RCV001518047.3,RCV001510843.3,RCV001515042.3,RCV000979124.4,RCV00094456...	For each allele, a pipe-delimited list of ClinVar RCV phenotype accession.version values associated with that allele; allele groups are comma-delimited.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_CLNDISDB	.,MedGen:C3661900; .,MedGen:CN169374; .,MedGen:C3661900,.	ClinVar condition database identifiers carried in the dbSNP VCF for each allele; serialization is allele- and release-dependent.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_CLNDN	.,not_provided; .,not_specified; .,not_provided,.	ClinVar preferred condition names carried in the dbSNP VCF for each allele.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_CLNHGVS	chr3:g.25638315_25638352=,chr3:g.25638325_25638352del,chr3:g.25638326_2563835...; chr2:g.47414421_47414447=,chr2:g.47414429_47414447del,chr2:g.47414430_4741444...; chrX:g.147912050_147912080=,chrX:g.147912051CGG[3],chrX:g.147912051CGG[4],chr...	ClinVar HGVS expression(s) carried in the dbSNP VCF for each allele; ordering corresponds to the other allele-specific clinical INFO fields.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_CLNORIGIN	.,1; .,1,.; .,.,1	ClinVar allele-origin code(s) carried in the dbSNP VCF for each allele.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_CLNREVSTAT	.,single; .,single,.; .,.,single	ClinVar review-status code(s) carried in the dbSNP VCF; legacy abbreviations such as single, mult, conf, exp or guideline are release-dependent.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_CLNSIG	.,0; .,3; .,0,.	ClinVar clinical-significance code(s) carried in the dbSNP VCF for each allele. Numeric code meanings are release-dependent and should be interpreted against the matching VCF header.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_CLNVI	.,; .,,.; .,.,	Variant identifiers maintained by external organizations, encoded as source:identifier pairs and separated by "|".	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_COMMON	false; False; true	Flag indicating that the RefSNP is common under dbSNP's stated frequency criterion for that release.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_dbSNPBuildID	144; 157; 151	First dbSNP build in which the RefSNP identifier was created; may be absent for records created solely from frequency data.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_DSS	false; False; true	Flag indicating a donor splice-site annotation.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_FREQ	GnomAD_exomes:1,1.679e-06; TOPMED:1,3.778e-06|dbGaP_PopFreq:1,0; GnomAD_exomes:1,1.241e-06	Ordered allele frequencies from contributing studies: reference allele first, followed by ALT alleles in VCF order; study blocks are separated by "|".	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_GENEINFO	BRCA2:675; TTN:7273|TTN-AS1:100506866; APC:324	Gene symbol:NCBI Gene ID pair(s) associated with the RefSNP, separated by "|"; pseudogenes are reported separately.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_GNO	true; false; True	Flag indicating that genotype data are available for the RefSNP.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_ID	rs56986587; rs11309117; rs193922936	dbSNP RefSNP identifier, including the "rs" prefix.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_INT	false; true; False	Flag indicating an intronic annotation.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_NSF	false; False; true	Flag indicating a non-synonymous frameshift coding annotation.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_NSM	true; false; True	Flag indicating a non-synonymous missense coding annotation.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_NSN	false; False; true	Flag indicating a non-synonymous nonsense/stop-gained coding annotation.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_PM	false; False; false|false	Flag indicating that the variant has an associated publication.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_PSEUDOGENEINFO	TPTE2P5:100616668; MROH5:389690; SSPOP:23145	Pseudogene symbol:NCBI Gene ID pair(s), separated by "|".	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_PUB	false; False; true	Flag indicating that the RefSNP or an associated submitted SNP is cited in a publication.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_R3	false; False; true	Flag indicating annotation in a 3-prime gene region.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_R5	false; False; true	Flag indicating annotation in a 5-prime gene region.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/

dbSNP_SAO		Variant allele-origin code: 0 unspecified, 1 germline, 2 somatic, 3 both.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_SSR	0; 0.0; 1	Bitmask of dbSNP suspect-reason codes; multiple reasons can be combined. Interpret values against the matching dbSNP VCF header.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_SYN	false; true; False	Flag indicating a synonymous coding annotation.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_U3	false; False; true	Flag indicating a 3-prime untranslated-region annotation.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_U5	false; False; true	Flag indicating a 5-prime untranslated-region annotation.	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
dbSNP_VC	SNV; INDEL; DEL	dbSNP variation class (for example SNV).	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/	https://www.ncbi.nlm.nih.gov/snp/docs/RefSNP_about/
OREGANNO_ID	OREG1277197; OREG1499998; OREG1500837	ORegAnno record identifier for a regulatory element overlapping the variant/locus. Included in Funcotator data sources bundle.	https://plone.bcgsc.ca/platform/bioinfo/software/ORegAnno	https://gatk.broadinstitute.org/hc/en-us/articles/360035889931-Funcotator-Information-and-Tutorial
OREGANNO_Values	Outcome=POSITIVE OUTCOME|Type=REGULATORY REGION|Regulatory_Element_Symbol=CTC...; Type=TRANSCRIPTION FACTOR BINDING SITE|Gene_Symbol=ZNF469|Gene_ID=ENSG00000022...; Type=TRANSCRIPTION FACTOR BINDING SITE|Gene_Symbol=FLNC|Gene_ID=ENSG000001285...	ORegAnno regulatory annotation value(s) for overlapping records; multiple records may be delimited in the source-specific serialization. Included in Funcotator data sources bundle.	https://plone.bcgsc.ca/platform/bioinfo/software/ORegAnno	https://gatk.broadinstitute.org/hc/en-us/articles/360035889931-Funcotator-Information-and-Tutorial
Orpha	289377|-|Unknown|Autosomal recessive|Childhood|611705 ~140922|LGMD2J|<1 / 10...; 791|-|1-5 / 10 000|Autosomal dominant or Autosomal recessive or&nb...; 244|PCD|-|Autosomal dominant or Autosomal recessive or X-link...	Serialized Orphanet gene–disorder annotation from the configured data source, containing fields such as Orphanet number, disorder/syndrome label, prevalence, inheritance, age of onset and OMIM cross-reference. Exact position and missing-value conventions are source-version-dependent.	https://www.orphadata.com/	downloaded Orphadata XML/JSON/CSV and product documentation
OrphaNumber	289377;140922;154;178464;609;; 791;; 244;	Orphanet disorder identifier(s) associated with the annotated gene.	https://www.orphadata.com/	downloaded Orphadata XML/JSON/CSV and product documentation
Phenotype_MIM	600334;603689;604145;608807;611705;613765;; 114480;155255;176807;194070;605724;612555;613029;613347;; 114480;208900;	OMIM phenotype identifier(s) associated with the gene/disease annotation.	https://www.orphadata.com/	downloaded Orphadata XML/JSON/CSV and product documentation
rmsk	Name=G-rich; Name=GA-rich; Name=(T)n	RepeatMasker repeat annotation for the genomic region containing the variant.	https://www.repeatmasker.org/	local rmsk table/BED and track schema
Simple_Uniprot_alt_uniprot_accessions	A6NKB1|E7EQE6|E7ET18|K7ENY1|Q10465|Q10466|Q15598|Q2XUS3|Q32Q60|Q4U1Z6|Q4ZG20|...; O00183|O15008|Q13879|Q5TBJ7; F8WCL5|F8WCP0|Q15346|Q53QQ2|Q53TG8	Secondary/alternative UniProt accession(s) associated with the protein entry; these are redirects/aliases, not separate gene identifiers.	https://www.uniprot.org/help/entry_name	https://www.uniprot.org/help/entry_name
SwissProt_acc_Id	Q8WZ42; P51587; P20929	Stable UniProtKB/Swiss-Prot accession for a reviewed protein entry; isoforms may use a suffix such as "-2".	https://www.uniprot.org/help/entry_name	https://www.uniprot.org/help/entry_name
SwissProt_entry_Id	TITIN_HUMAN; BRCA2_HUMAN; NEBU_HUMAN	UniProtKB/Swiss-Prot mnemonic entry name (for example ALK_HUMAN). It is human-readable but less stable than the accession.	https://www.uniprot.org/help/entry_name	https://www.uniprot.org/help/entry_name

Supplementary Table S3 – Description of the annotation fields included in the varianTalker MAF output. Detailed description of the annotation fields reported in the complete MAF file generated by varianTalker. For each field, the table provides the column name, an example value, a description of its content and interpretation, the primary evidence or documentation source, and the corresponding source resource or artifact used to generate the annotation.

#CHROM	POS	ID
chr11	110279332	983524
chr11	22214829	4070939
chr1	146135233	4287598
chr15	24952616	1697252
chr15	72714243	983424
chr17	19656568	659058
chr17	35117694	3336907
chr17	3644575	870585
chr17	48109265	4285397
chr17	61860483	989417
chr20	50899224	1879783
chr2	178444500	4287285
chr22	26470846	4073704
chr22	28725113	1713222
chr22	28725114	665014
chr2	47783231	1048851
chr3	36993509	635950
chr5	110742637	1068766
chr6	166951306	413
chr6	33308994	646471
chr6	87555066	4528930
chr7	117498400	3337030
chr9	95313839	929797
chr17	48112739	4285398
chr2	178447584	4287286
chr3	36995227	3396542
chr7	117498442	1705873
chr3	36995927	89666
chr3	36999241	90028
chr3	37002980	90153
chr3	37003580	90159
chr3	37003588	90158
chr3	37004155	90155
chr3	37006559	90196
chr3	37008148	90237
chr3	37008308	90235
chr3	37008347	90234
chr3	37008348	90233
chr3	37011458	90268
chr3	37019712	90423
chr3	37019815	90422
chr3	37006051	1735212
chr3	37006916	989413
chr3	37007688	242651
chr3	37010176	242650
chr3	37018276	3337903

Supplementary Table S4 – ***Long-deletion ClinVar records incompatible with Funcotator annotation.*** *Genomic coordinates and ClinVar identifiers of the 46 long-deletion records that consistently caused failure during the Funcotator-based annotation step of varianTalker. These variants were identified during workflow development and excluded from the ClinVar test set prior to downstream annotation and benchmarking analyses.*