Automating the ACMG Guidelines with VSClinical

Gabe Rudy | VP of Product & Engineering

Thanks to NIH & Stakeholders

- NIH Grant Supported
  - Research reported in this publication was supported by the National Institute Of General Medical Sciences of the National Institutes of Health under Award Number R43GM128485. PI is Dr. Andreas Scherer, CEO Golden Helix. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

- ACMG Guidelines Author Collaborator:
  - Dr. Elaine Spector (Childrens Colorado, USA)
- Stakeholders:
  - Dr. Abdallah Elias (Shodair Children's Hospital, USA)
  - Dr. Ahmed Alfares, King Abdul Aziz Medical City, Saudi Arabia),
  - Dr. Bailey Glen (Medical University of South Carolina, USA)
  - Dr. Jim Weber (PreventionGenetics, USA)
  - Dr. Qin Hao and Dr. Line Larsen (Amplexa, Denmark)
  - Dr. Val Hyland (Pathology Queensland, Australia).

Q & A

Please enter your questions into your GoToWebinar Panel

Golden Helix – Who We Are

Golden Helix is a global bioinformatics company founded in 1998.

GlaxoSmithKline

varSEQ®

- Variant Calling
- Filtering and Annotation
- **Variant Interpretation**
- Clinical Reports
- CNV Analysis
- Pipeline: Run Workflows

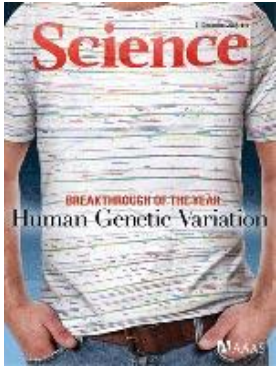
VARIANT WAREHOUSE

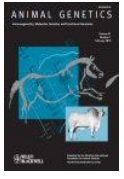
- Variant Warehouse
- Centralized Annotations
- Hosted Reports
- Sharing and Integration

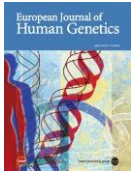
SNP & VARIATION SUITE

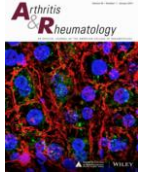
- **CNV Analysis**
- GWAS
- Genomic Prediction
- Large-N-Population Studies
- RNA-Seq
- Large-N CNV-Analysis

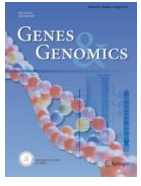
Cited in over 1,200 peer-reviewed publications

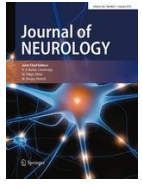












Over 350 customers globally

SickKids
THE HOSPITAL FOR SICK CHILDREN

Cincinnati Children's
Hospital Medical Center

The Children's Hospital of Philadelphia

Stanford University

UBC

Ucla

Yale

The Feinstein Institute for Medical Research
North Shore-Long Island Jewish Health System
North Shore LIJ

PREVENTION GENETICS
DISEASE PREVENTION THROUGH GENETIC TESTING

MAYO CLINIC

LINEAGEN

MOFFITT CANCER CENTER

Dartmouth-Hitchcock
NORRIS COTTON CANCER CENTER

BAYER

Lilly

abbvie

TEVA

NATIONAL CANCER INSTITUTE

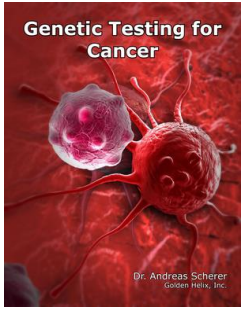
USDA
USDA Agricultural Research Service

FDA

Golden Helix – Who We Are

When you choose a Golden Helix solution, you get more than just software

- REPUTATION
- TRUST
- EXPERIENCE



- INDUSTRY FOCUS
- THOUGHT LEADERSHIP
- COMMUNITY

- TRAINING
- SUPPORT
- RESPONSIVENESS

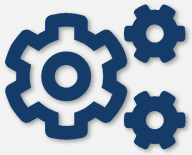
- INNOVATION and SPEED
- CUSTOMIZATIONS

Genetic Testing Process

Golden Helix Clinical Suite

Sample Prep

Sequencing



Align & Call

Annotate & Filter

Variant Interpretation

Report

Sentieon & VS-CNV

VarSeq

VSClinical

VSReports

VSWarehouse

Aggregate Variants, Reports, Knowledgebase

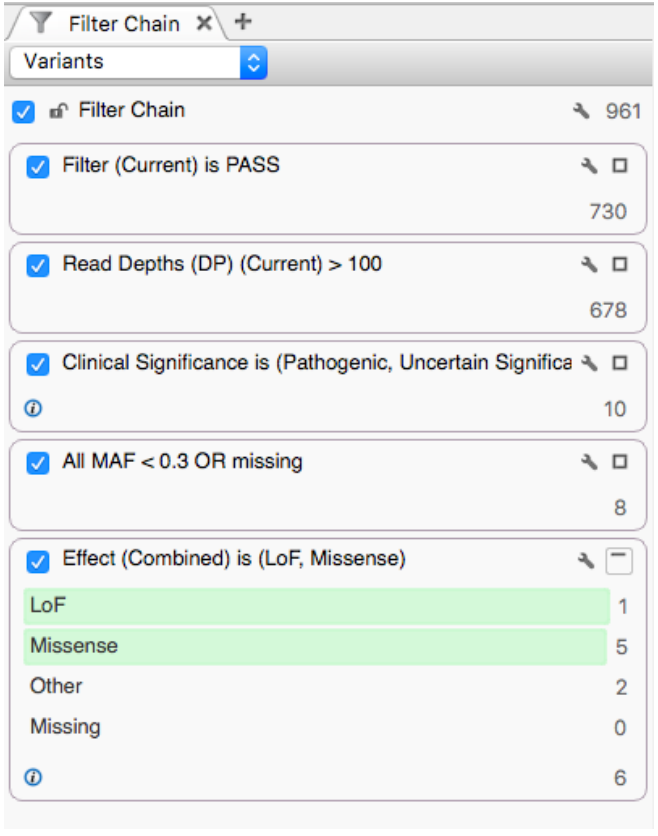
VSClinical

- **Complete Support for ACMG Guideline Workflow:**
  - Implements a guided workflow for following the ACMG guideline scoring and classifying
  - Place criteria into conceptually related groups, paired with their opposites, and formatted as answerable question.
- **Aggregate and Automate:**
  - ACMG Classifier algorithm in VarSeq runs automatable rules
  - Recommendation engine in VSClinical provides list of criteria and reasons for recommendations
- **Expert and Beginner Friendly:**
  - Descriptive summaries and recommendations for a variant
  - Deep dive into Population Catalogs, Gene Impact, Published Studies and Clinical tabs
  - Integrated documentation, readings on scoring criteria and citations

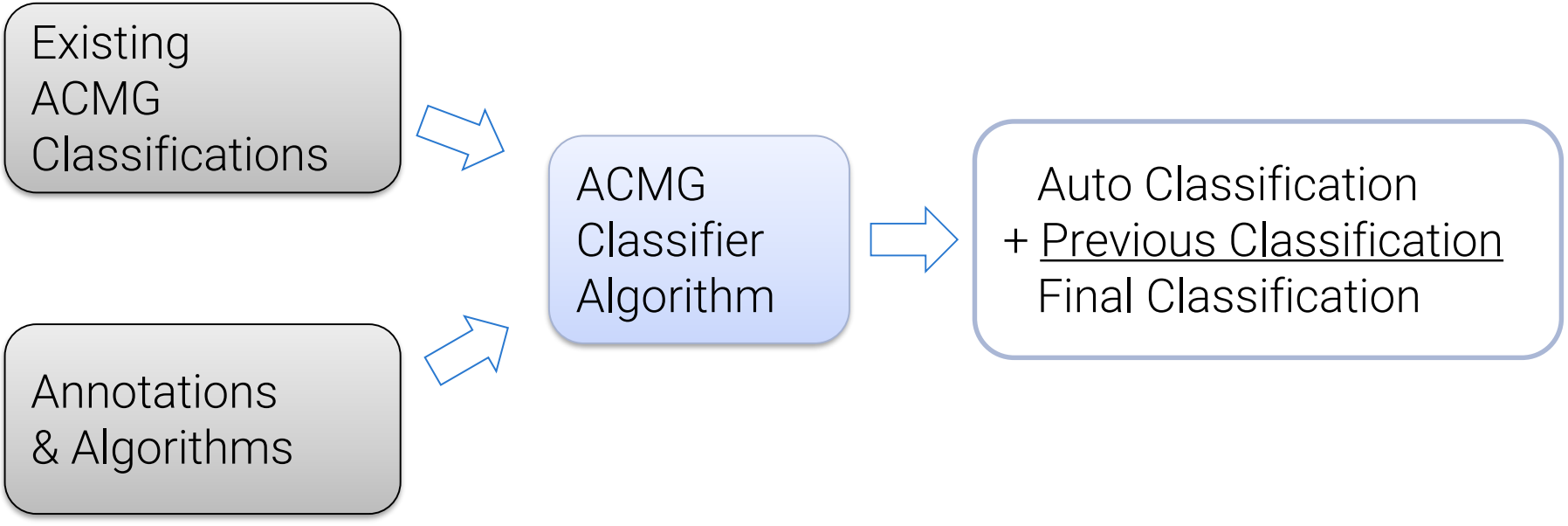
VSClinical®

Analysis Workflow with VSClinical

1. **Follow your existing VarSeq annotation and filtering workflow**
2. **Add new ACMG Auto Classifier algorithm:**
   - Looks up if variant annotated in previous sample
   - Scores 18 criteria based on available evidence from 7 sources
3. **Select variants to evaluate using the ACMG Guidelines**
4. **Score and Finalize each variant, selecting which to report**
5. **Finalize the sample, review and report**



The ACMG Classifier Algorithm



- 1000 Genomes Frequencies w/ Genotype Counts
- gnomAD Exomes Frequencies w Genotype Coun
- **GERP++ / PhyloP Conservation Scores**
- **SIFT & PolyPhen2 MSA Missense Predictions**

- Transcript Annotations with Default Transcript
- **Splice Site Prediction Algorithms**
- Gene Preferences (Recessive/Dominant Model)
- ClinVar Variants & Transcript Counts

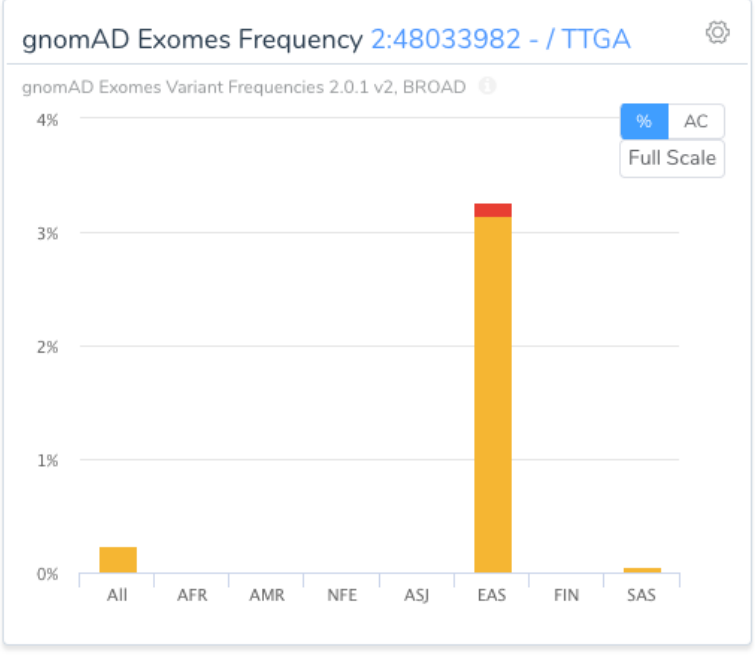
[Demo in VarSeq]

Population Criteria

For all frequency thresholds, the max sub-population frequency is used.

- **BA1**: Variant is common (assumed benign) in one or more population catalogs
- **BS1**: Allele frequency greater than expected for disorder
- **BS2**: Observed in a healthy adult individual with causal genotype for early onset, highly penetrant disorder.
- **PM2**: Absent from controls in population catalogs

	Recessive	Dominant
BA1	1.00%	0.50%
BS1	0.30% or 2+ Hom	0.10% or 1+ Hom
BS2	Hom + Hemi in 1kG > 0	Het in 1KG > 0
PM2	0.15%	Absent in all



Variant Type Specific Criteria

Analysis Based on Current Transcript Variant Effect

Loss of Function Variants and Protein Length Changes

- **PVS1**: Null variant in a gene where LOF is a known mechanism of disease
  - Will not score if within last 50bp of pen-ultimate exon
  - 1+ LoF Pathogenic Variant in ClinVar with 1+ star rating
- **PM4**: Protein length changes as a result of a stop-loss variant
- **PM4**/**BP3**: Protein length changes as a result of in-frame deletions/insertions in a non-repeat region
  - For in-frame insertions/deletions, if changed amino acid sequence is repeated less than two times PM4, else BP3

NM_001009944.2(Reverse Strand) Coding Change

Coding DNA Sequence:

cDNA	GCC TGG TGT GCC TCC CTG GCC CAC GGG CTC AGC CTG CTC	
Pos	10,654	10,692
cDNA	CTG GTG GCT GTG GCT GT~~G~~ ~~GCT~~ ~~GT~~C TCA GGG TGG GTG GGT	
Pos	10,693	10,731
cDNA	GCG AGC TTC CCC CCG GGC GTG AGT GTT GCG TGG CTC CTG	
Pos	10,732	10,770

Amino Acid Sequence:

AA	Ala Trp Cys Ala Ser Leu Ala His Gly Leu Ser Leu Leu	
Pos	3,552	3,564
AA	Leu Val Ala Val Ala Val ~~Ala~~ ~~Val~~ Ser Gly Trp Val Gly	
Pos	3,565	3,577
AA	Ala Ser Phe Pro Pro Gly Val Ser Val Ala Trp Leu Leu	
Pos	3,578	3,590

The p.A3571_V3572del variant is a in-frame deletion of an amino acid sequence that is repeated 2 times in the surrounding region.

Variant Type Specific Criteria

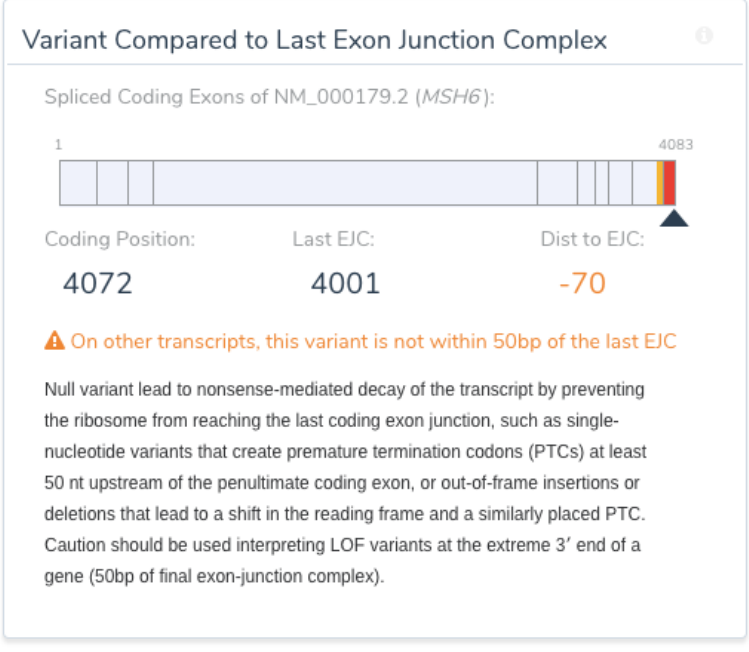
Analysis Based on Current Transcript Variant Effect

Missense & In-Frame Variants

- **PM1**: Mutational hot spot without benign variation
  - Within 6 amino acids, at last 2 pathogenic variants, and more than there are benign
  - No benign variants within 3 amino acids

Missense

- **PP2**/**BP1**: Missense in gene with low rate of benign missense variants and pathogenic missenses common
  - Missense Z Score is > 1
  - One or more Pathogenic/Likely Pathogenic missense in gene



The p.K1358Dfs variant occurs in the last exon of MSH6.
There are no other pathogenic loss of function variants downstream of the variant p.K1358Dfs.

Variant Type Specific Criteria

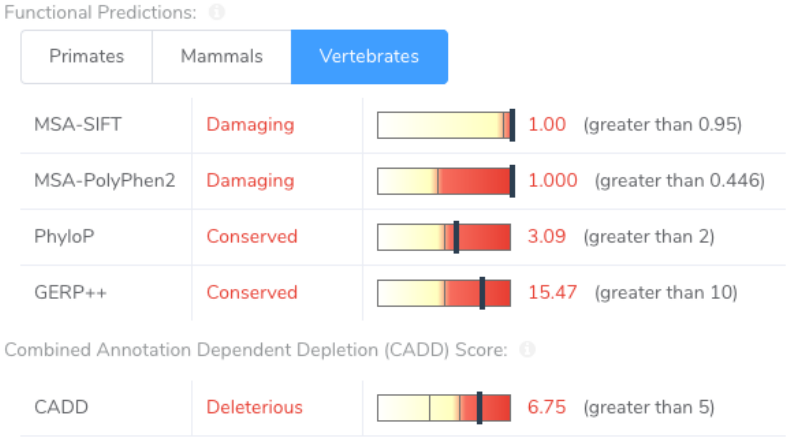
Analysis Based on Current Transcript Variant Effect

In-Silico Evidence (for Non-LoF Variants)

- **PP3**: 3 or 4 out of 4 splice site predictions of damaging
- **PP3**: In-silico predictions in agreement variant is damaging & conserved
- **BP4**: If variant amino acid present in mammalian species
- **BP4**: In-silico predictions in agreement that variant is tolerated & not conserved

Synonymous / Intronic Variants

- **BP7**: Not predicted to disrupt a canonical splice site (0 or 1 of 4 predicted disrupted)



- The p.R1076C missense variant is predicted to be damaging by both SIFT and PolyPhen2.
- The arginine residue at codon 1076 of MSH6 is conserved in all mammalian species.
- The nucleotide c.3226 in MSH6 is predicted conserved by GERP++ and PhyloP across 100 vertebrates.

Published Clinical / Functional Evidence

Required Manual Confirmation and Validation of Cited Evidence

Clinical Studies

- **PS1**: Missense variant with same amino acid change in ClinVar as a 1+ star rating as Pathogenic or Likely
- **PM5**: In-frame or Missense where ClinVar has 1+ star rated variant that is Pathogenic or Likely Pathogenic at same residue position.

Functional Studies

- Must check by hand, but ClinVar assertions provided.

Reputable Source

- **PP5**: Not applied PS1/PM5 and exact variant in ClinVar as Pathogenic or Likely
- **BP6**: In ClinVar as Benign

	Variant	HGVS	AA Change	Clinical Significance
>	428337	c.3226C>G p.Arg1076Gly	Arg → Gly CGC → GGC	Uncertain Significance ★★☆☆
>	This Variant	c.3226C>T p.Arg1076Cys	Arg → Cys CGC → TGC	Likely pathogenic ★★★☆
>	186361	c.3227G>A p.Arg1076His	Arg → His CGC → CAC	Uncertain Significance ★★☆☆

- The p.R1076C variant is a missense mutation resulting in an amino acid change which is shared by the previously classified pathogenic variant p.Arg1076Cys.

[Another Variant if Time]

Next Steps & Questions

- Special Pricing Offer!
  - 15-month license for all VSClinical annual subscriptions made in the month of June
- Contact us to evaluate!
  - VSClinical is separately licensed product
  - Includes OMIM, CADD, VSReports
  - Splice site predictions
  - Functional predictions
- Part of VarSeq 2.0 release now available
- Contact us if you are interested in other workflows or customizations

VSClinical®