Clinical Chemistry
Trainee Council

PEARLS OF
LABORATORY MEDICINE

DNA Sequence Nomenclature and Variant Interpretation

DOI: 10.15428/CCTC.2013.214643

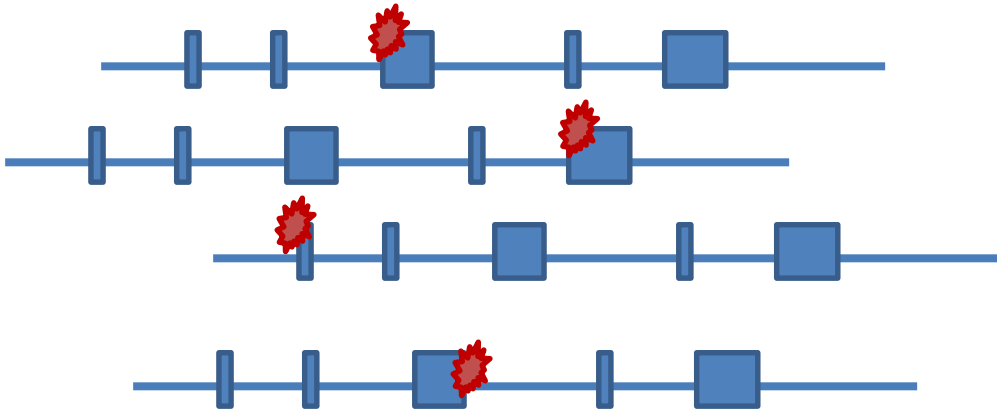
Lora J. H. Bean, PhD, FACMG
Emory University

www.traineecouncil.org

The Human Genome

- ~3 billion base pair locations
  - Gaps in sequence still exist
  - Exact position of a specific base changes with builds
  - Current build – Hg19
- Variation
  - Millions of single nucleotide polymorphisms (SNPs)
  - Copy number variants (CNVs)
- Nomenclature
  - Standard
  - Historic

Clinical Sequencing



One or a few pathogenic variants account for most cases in a population

Many private pathogenic variants or an unknown spectrum of mutations in the population

Targeted genotyping assays

Full gene sequencing

Standard Nomenclature

- Use of standard nomenclature allows for the precise, unambiguous identification of a genomic position
- Citing a specific reference sequence is critical for long-term understandability of results
- Necessary for targeted testing and proper interpretation of family studies

Standard Nomenclature

- Recommendations by the Human Genome Variation Society
- DNA / RNA / protein identity
  - Genomic denoted as “g.”
  - Coding denoted as “c.”
  - Mitochondrial denoted as “m.”
  - RNA denoted as “r.”
  - Protein denoted as “p.”
- Recommendations for
  - Single basepair changes
  - Small deletions, duplications, insertions
  - Large rearrangements
  - Intronic changes
  - Nearly any scenario that has ever been reported

Standard Nomenclature:
Recommendations by the Human Genome Variation Society

Nomenclature for Variation from Reference Sequence			
Type of Variation	**Genomic**	**cDNA**	**Protein**
Missense change	g.5248232T>A	c.20A>T	p.Glu7Val or p.E7V
Nonsense change	g.20763582C>A	c.139G>T	p.Glu47* or p.E47X
Deletion	g.117199646_ 117199648delCTT	c.1521_1523delCTT	p.Phe508del or p.F508del
Intronic change	g.103234177C>T	c.1315+1G>A	-

- Specifying the reference sequence is critical
  - These changes are meaningless without a reference
  - Different references result in different correct ways to describe a variant

http:// www.hgvs.org/mutnomen/
Ogino et al., 2007

Standard Nomenclature

The most common sickle cell disease mutation

Medical Records

Databases

Hemoglobin S

Reference

Historic

dbSNP:rs334

HbS

dbSNP:rs77121243 (retired)

HBB: Glu6Val

NC_000011.9:g.5248232T>A

NP_000509.1:p.Glu7Val

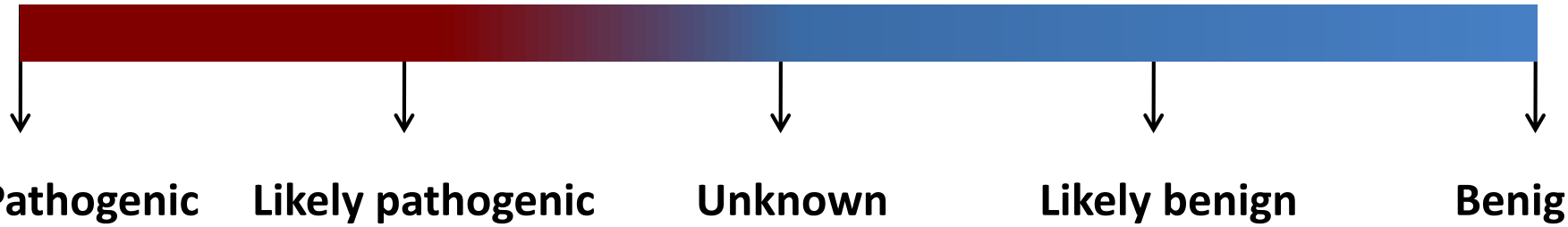
NM_000518.4:c.20A>T

Published Literature

Standard

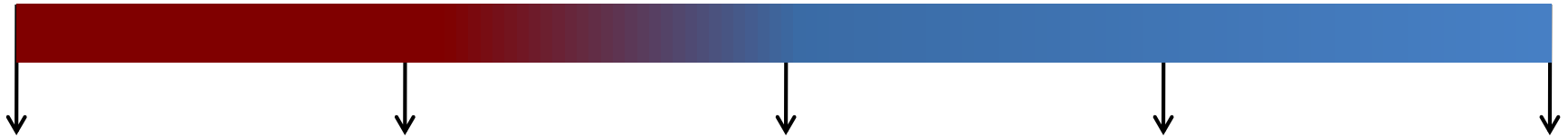
Genotyping Platforms

Sequence variant interpretation



- Richards et al., ACMG Recommendations for standards for interpretation and reporting of sequence variations: Revisions 2007. Genet Med 2008: 10(4): 294-300.
- Maddalena et al., Technical standards and guidelines: molecular genetic testing for ultra-rare disorders. Genet Med 2005: 10(8):571-583. Reviewed and Revised 2009.
- The ACMG Laboratory Practice Committee Working Group, ACMG Recommendations for Standards and Interpretation of Sequence Variants. Genet Med 2000: 2(5): 302-303.

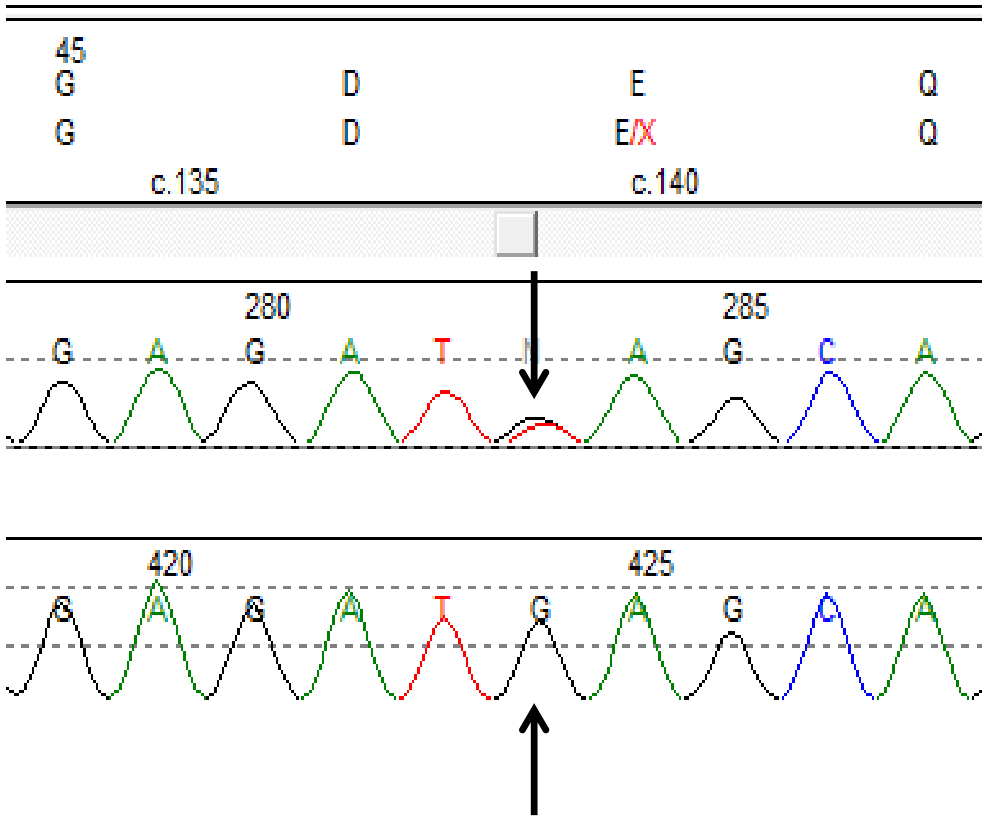
Pathogenic variants



Pathogenic Likely pathogenic Unknown Likely benign Benign

- May be interpreted as pathogenic
  - Nonsense
  - Frameshift (deletion, duplication, or insertion)
  - Splice site mutations (± 1 or ± 2 intronic position)
- May be interpreted as pathogenic
  - Missense or in-frame changes with sufficient evidence
  - Silent or intronic change proven to disrupt splicing

Pathogenic variants: *GJB2*: c.139G>T (p.E47X)



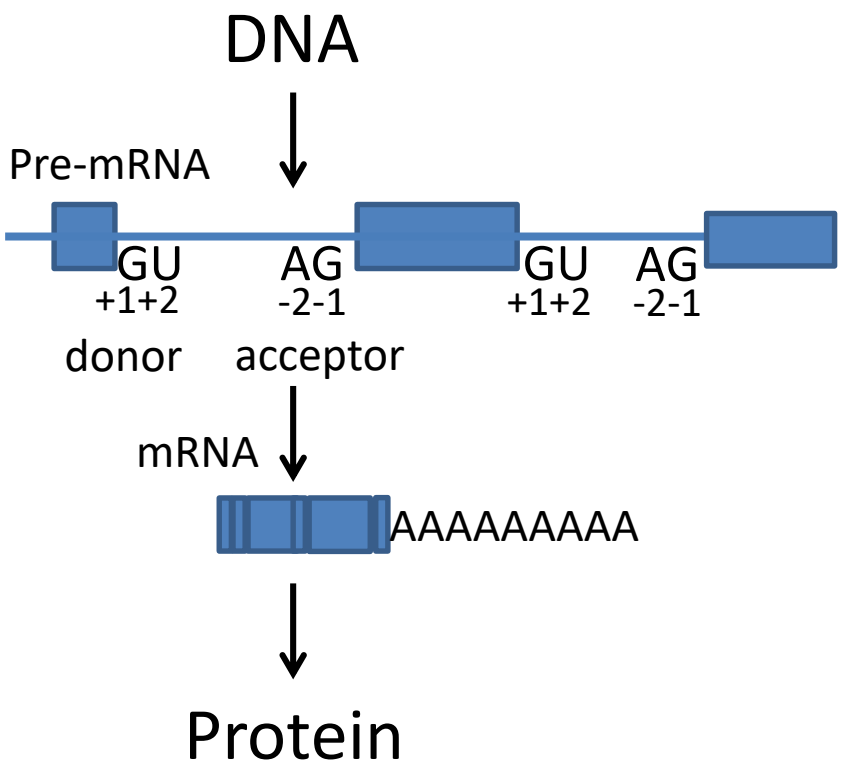
Patient:
GAG -> glutamate (GLU or E)
TAG -> STOP

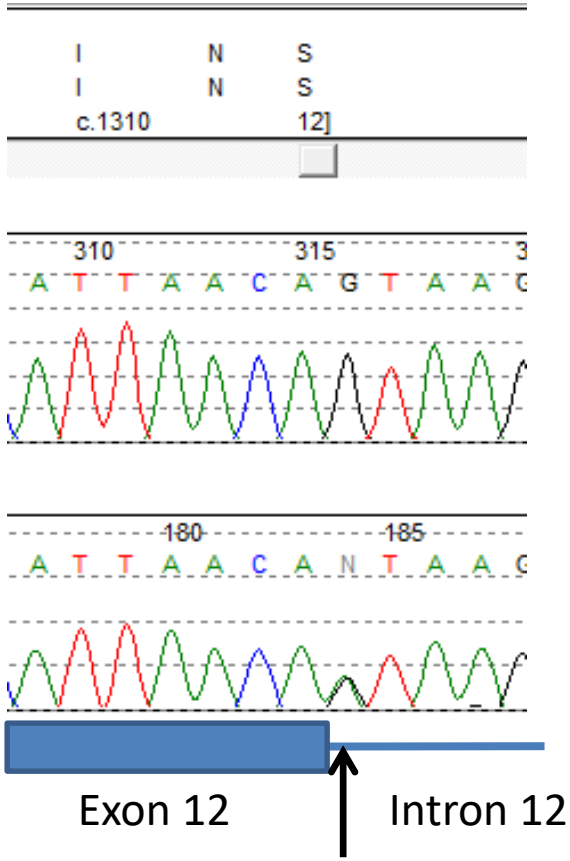
Reference:
GAG -> glutamate (GLU or E)

Sequencing traces from a patient (top) with a heterozygous premature stop codon compared to a reference sequence (bottom). The stop codon is interpreted as a pathogenic change.

Source: Emory Genetics Laboratory

Pathogenic variants: *PAH:* IVS12+1G>A

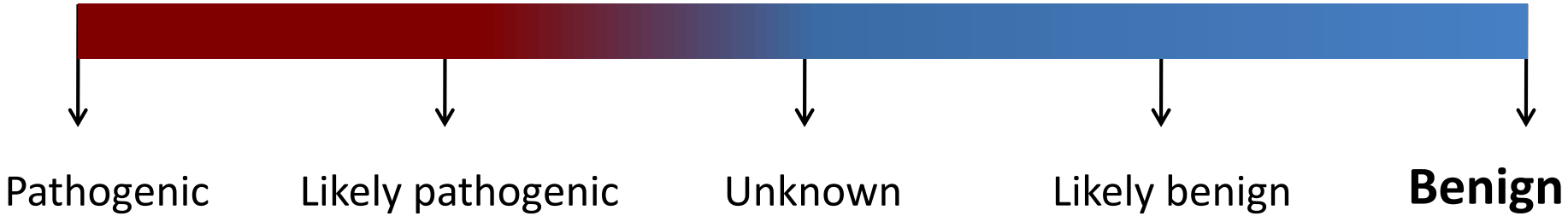




Sequencing traces from a patient (bottom) with a pathogenic variant in the consensus donor site of exon 12 / intron 12 of the *PAH* gene compared to a reference sequence (top).

Source: Emory Genetics Laboratory

Benign variants



- Population frequency inconsistent with disease
  - Frequency in the general population
  - Frequency in a control population
- Evidence in the published literature that variant has no effect on function

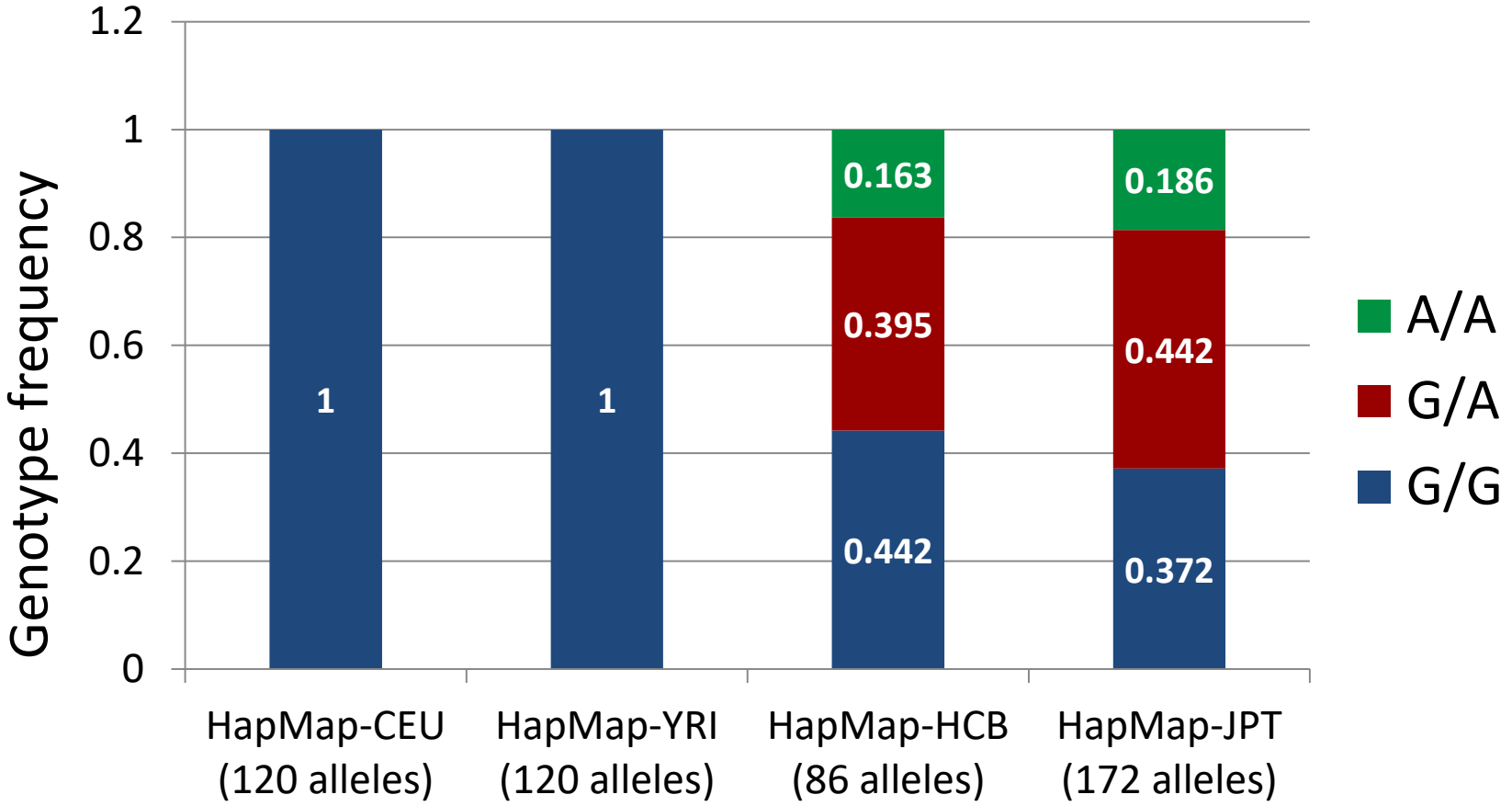
Benign variation *GJB2*

- The NHLBI GO Exome Sequencing Project
  - Large scale next-generation sequencing project
  - Focus on heart, lung, and blood disorders
  - European and African American populations
- *GJB2*: c.-34C>T

  European allele frequency: 0.08% (A=7 / G=8587)
  African allele frequency: 23.38% (A=1030 / G=3376)
- *GJB2*: c.79G>A (p.V27I)

  European allele frequency: 0.21% (A=18 / G=8582)
  African allele frequency: 0.34% (A=15 / G=4391)

Benign variation *GJB2*: c.79G>A (p.V27I)a

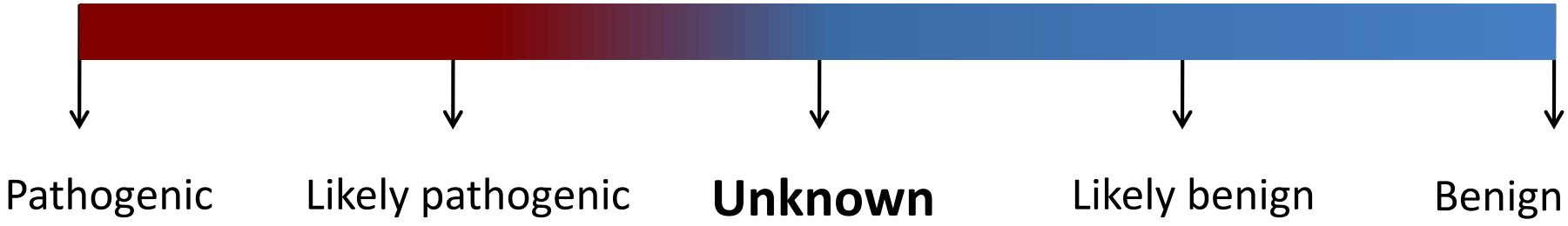


Genotype frequencies of the c.79G>A (p.V27I) variant in HapMap populations.

CEU - Utah Residents, YRI - Yoruba in Ibadan, Nigeria, HCB - Han Chinese in Beijing, China, JPT - Japanese in Tokyo, Japan

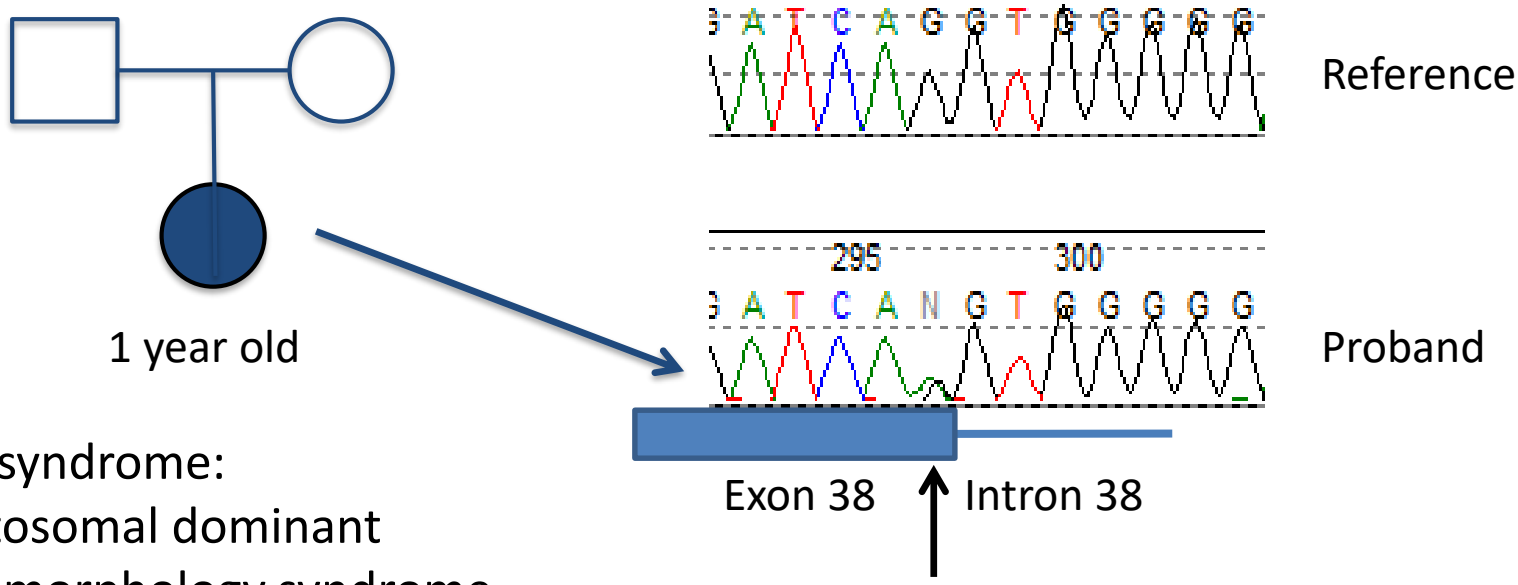
http://www.ncbi.nlm.nih.gov/projects/SNP (rs2274084)
http://hapmap.ncbi.nlm.nih.gov/

Variants of Unknown Significance



- Previously unreported missense changes
- Previously unreported in-frame deletions and duplications
- Genes mutated in conditions with no definite clinical or biochemical diagnosis

Variant of Unknown Significance: *MLL2:* c.10740G>A (p.Q3580Q)



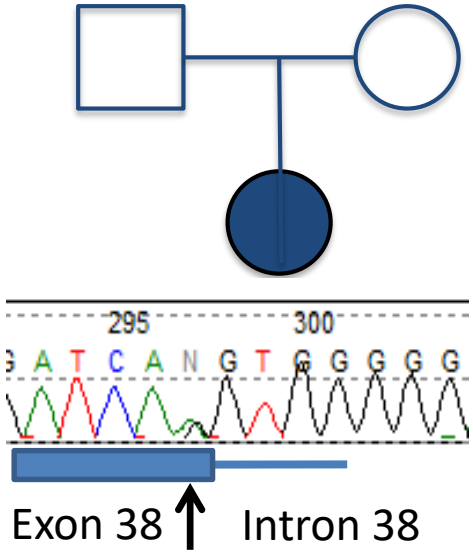
Kabuki syndrome:

- Autosomal dominant
- Dysmorphology syndrome
- Caused by mutations in *MLL2*

Sequencing traces from a patient (bottom) with a heterozygous variant (marked with an arrow) compared to a reference sequence (top).

Source: Emory Genetics Laboratory

Variant of Unknown Significance: *MLL2:* c.10740G>A (p.Q3580Q)

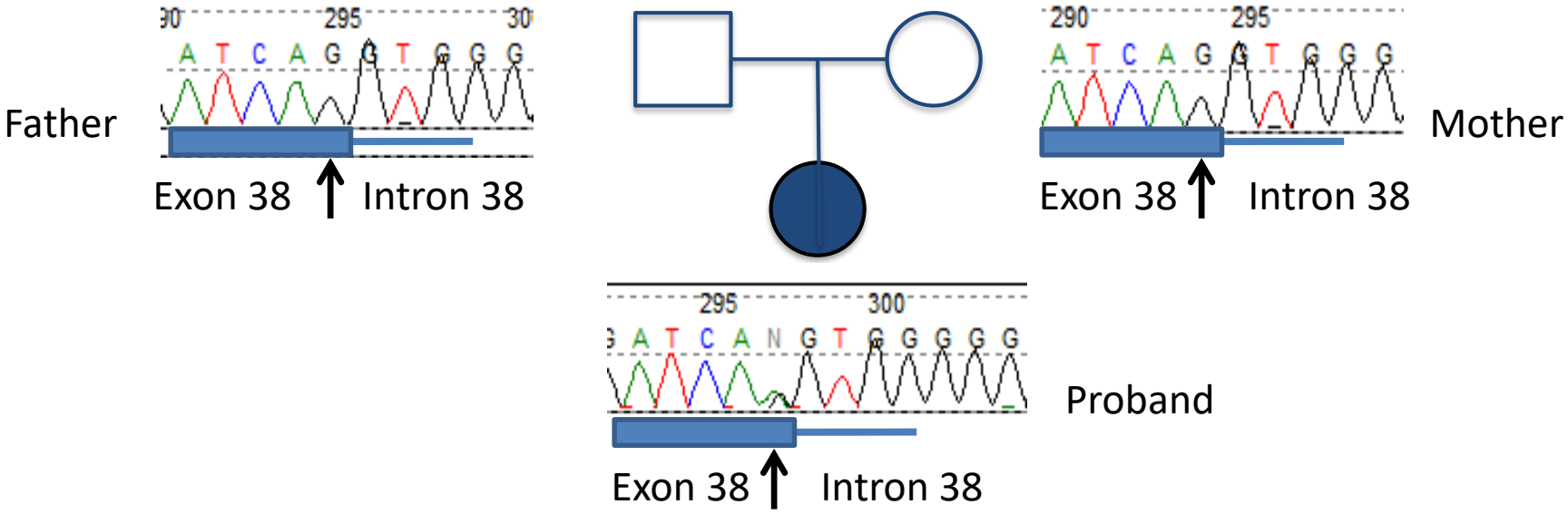


- Not reported in individuals with Kabuki syndrome
- Not reported in population studies
- Not predicted to change the amino acid at p.Q3580

→ Test the parents

Source: Emory Genetics Laboratory

Variant of Unknown Significance: *MLL2:* c.10740G>A (p.Q3580Q)

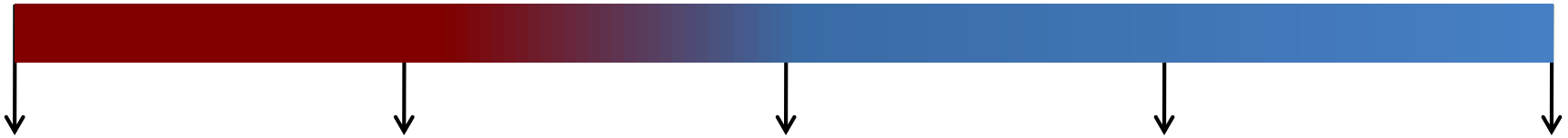


- Not reported in individuals with Kabuki syndrome
- Not reported in population studies
- Not predicted to change the amino acid at p.3580
- Parental studies demonstrate this change occurred *de novo* (parental identity was confirmed)

→ Likely pathogenic

Source: Emory Genetics Laboratory

Variants of Unknown Significance

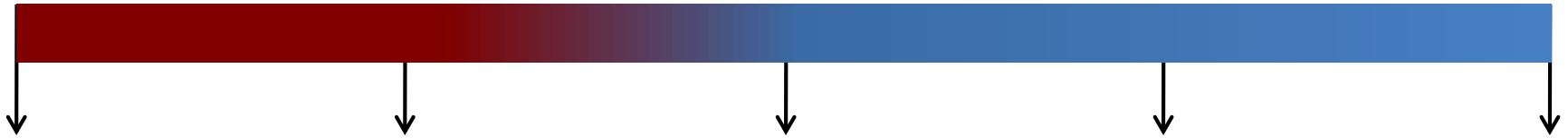


Pathogenic **Likely pathogenic** Unknown **Likely benign** Benign

- Segregates with a known disease causing pathogenic variant in an affected person (recessive) or segregates with disease in multiple affected people or occurred *de novo* (dominant)
- Nucleotide and/or amino acid at that position is evolutionarily conserved
- Not present in publically available population databases at a frequency consistent with being a benign variant

- Found in heterozygous state (dominant) or homozygous state (recessive) in multiple unaffected individuals.
- Found in *cis* with a pathogenic variant in multiple unrelated individuals
- For dominant disorders, the change is found in an unaffected family member

Interpretation of Sequence Variants

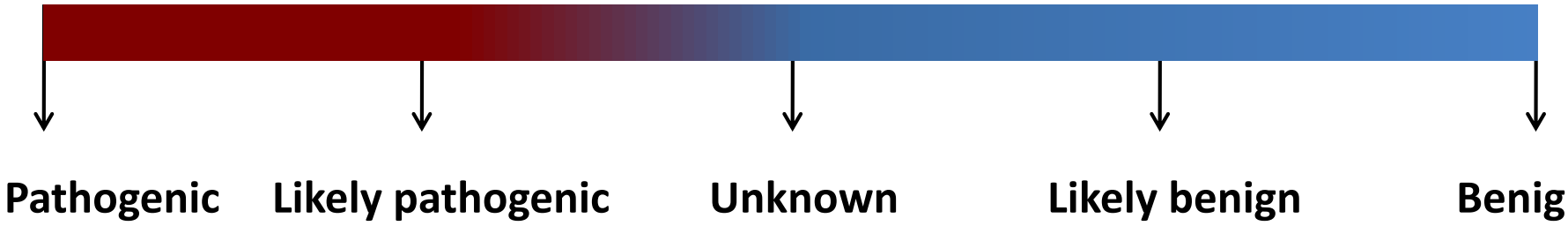


Pathogenic | Likely pathogenic | Unknown | Likely benign | Benign

- Population databases
- Disease-specific databases
- Literature
- Predicted change in protein

- Patient information
  - age
  - physical exam
- Family information
- Testing of parents or other family members
- Supporting lab results (e.g. biochemical testing)

Sequence variant interpretation



- ClinVar:
  www.ncbi.nlm.nih.gov/clinvar/
- The Human Variome Project:
  www.humanvariomeproject.org/

References

1. Maddalena A, Bale S, Das S, Grody W, Richards S. Technical standards and guidelines: molecular genetic testing for ultra-rare disorders. Genet Med 2005;7:571-83.

2. Ogino S, Gulley ML, den Dunnen JT, Wilson RB. Standard mutation nomenclature in molecular diagnostics: practical and educational challenges. J Mol Diagn 2007;9:1-6.

3. Richards CS, Bale S, Bellissimo DB, Das S, Grody WW, Hegde MR, et al. ACMG recommendations for standards for interpretation and reporting of sequence variations: Revisions 2007. Genet Med 2008;10:294-300.

4. The ACMG Laboratory Practice Committee Working Group. ACMG Recommendations for Standards and Interpretation of Sequence Variants. Genet Med 2000;2:302-3.

Disclosures/Potential Conflicts of Interest

Upon Pearl submission, the presenter completed the Clinical Chemistry disclosure form. Disclosures and/or potential conflicts of interest:

- **Employment or Leadership**: Senior Director, Molecular Genetics Laboratory, Emory Genetics Laboratory
- **Consultant or Advisory Role:** None declared
- **Stock Ownership:** None declared
- **Honoraria:** None declared
- **Research Funding:** None declared
- **Expert Testimony:** None declared
- **Patents:** None declared

Thank you for participating in this
Clinical Chemistry Trainee Council
Pearl of Laboratory Medicine.

Find our upcoming Pearls and other
Trainee Council information at
www.traineecouncil.org

Download the free *Clinical Chemistry* app
on iTunes today for additional content!

Follow us: