Role of Clinical Exome Sequencing in Diagnostic Odyssey

Pinar Bayrak-Toydemir, MD, PhD
Professor, Department of Pathology, University of Utah
Medical Director, Molecular Genetics and Genomics, ARUP Laboratories

Outline

- Description of exome sequencing

- Results of our clinical exome cases

  Detection rate based on clinical findings and trio vs proband

- Exome Sequencing interesting case discussions

- Guidelines and Recommendations

Next Generation Sequencing in Molecular Diagnosis

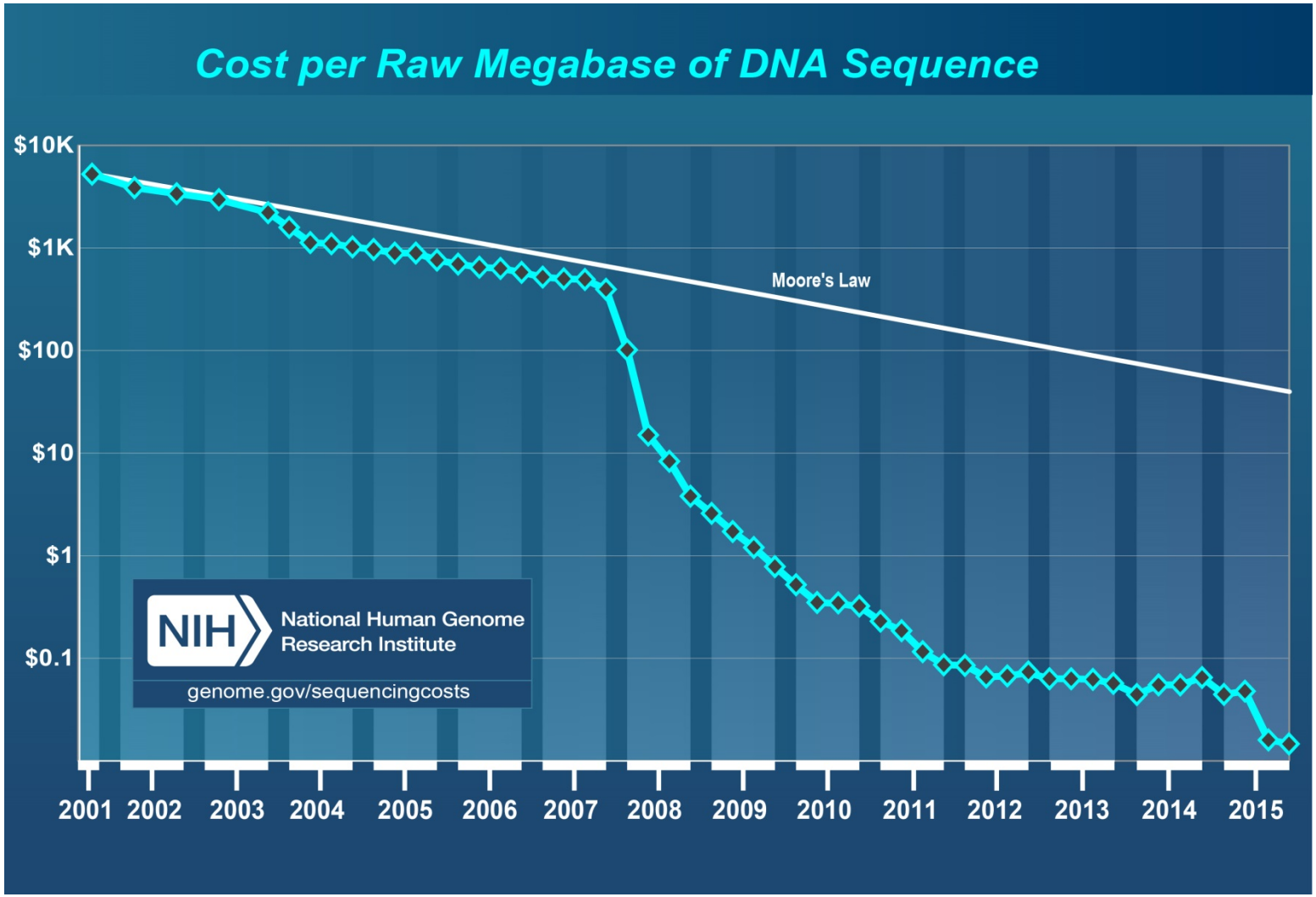
A powerful tool for gene discovery

200 genes are discovered every year

Now a powerful diagnostic tool !

Changed the way we think about
scientific approaches in basic, applied and clinical research and diagnostics

Next Generation Sequencing Cost Dropping



http://www.genome.gov/sequencingcosts/

Of approximately ~**19,000 protein-coding genes** predicted to exist in the human genome, variants that cause **Mendelian phenotypes have been identified in ~3,303** genes

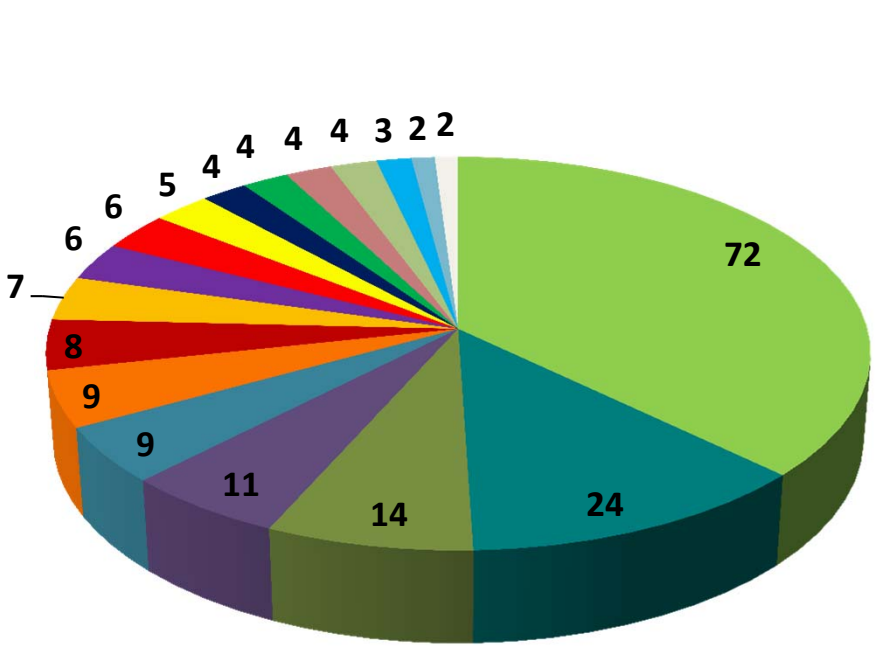
Dissected OMIM Morbid Map Scorecard (Updated January 9th, 2017) :

Class of phenotype	Phenotype	Gene *
Single gene disorders and traits	4,887	3,303
Susceptibility to complex disease or infection	699	498
"Nondiseases"	143	113
Somatic cell genetic disease	205	117

https://www.omim.org/statistics/geneMap

TT

New 2016 OMIM Disease-Associated Genes



N=179

- Neurological
- Syndromes
- Immunological
- Eye
- Skeletal
- Muscle
- Cardiac
- Mitochondrial
- Metabolic
- Blood
- Reproductive
- Ciliopathies
- Gastrointestinal
- Kidney
- Hearing
- Ectodermal
- Heterotaxy

Exome Sequencing

Sequencing of coding regions of all known genes

- Balanced to cover and obtain full coverage across the medically relevant genes in the human exome

- 100% coverage of all exons in 3,000 of the 4,600 disease associated genes making it the most comprehensive exome sequencing test available

Exome sequencing

- **Allows for identification of pathologic variants in newly identified disease genes**
- **Useful for conditions with locus heterogeneity (long molecular differentials)**
- **Unexpected/expanded phenotypic variation**

Exome Diagnostic Yield in Known Disease Genes in Children

Sawyer et al.

Table 1. Broad phenotypes and associated diagnostic rates in known disease genes of the families studied using WES in FORGE

Broad phenotype	Total families ($N = 362$)	Families with known genes ($N = 105$)	Diagnostic rate (%)
Neurodevelopmental	98	31	31.6
Dysmorphic syndromes	80	18	22.5
Ocular	40	11	27.5
Metabolic	31	12	38.7
Neuromuscular	30	7	23.3
Ciliopathy	27	12	44.4
Congenital malformation syndromes	19	4	21.1
Immunological	17	2	11.8
Other (isolated cardiac, endocrinology, skeletal dysplasia, connective tissue disorders, mental illness, lung disorder)	20	8	40.0

FORGE, Finding Of Rare Disease GEnes; WES, whole-exome sequencing.

Clin Genet 2016; 89: 275–228.

TT

Clinical Sensitivity

Clinical sensitivity may change based on the test ordered and also based on clinical presentation.

- **Neurodevelopmental disorders- yield around 73%**
- **Autism- yield around 28%**
- **Epilepsy-30%**

(Soden et al, 2014)
(Lee et al., 2014)
(Juusola et al., 2015)

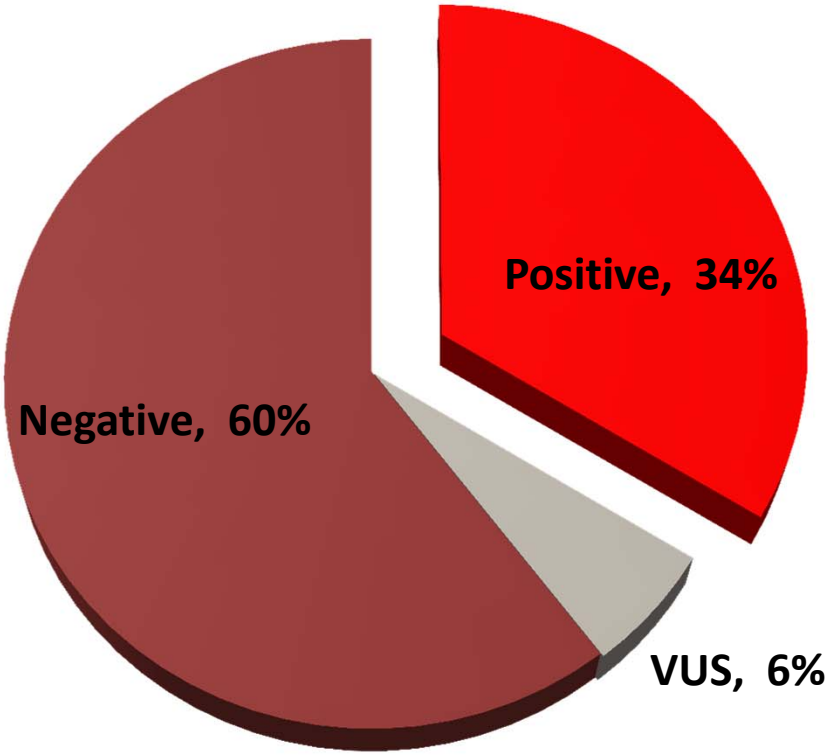
Clinical Sensitivity

De novo variants are reported when both parent's samples are available for exome sequencing; 35-50% of diagnoses were achieved by identification of de novo variants.

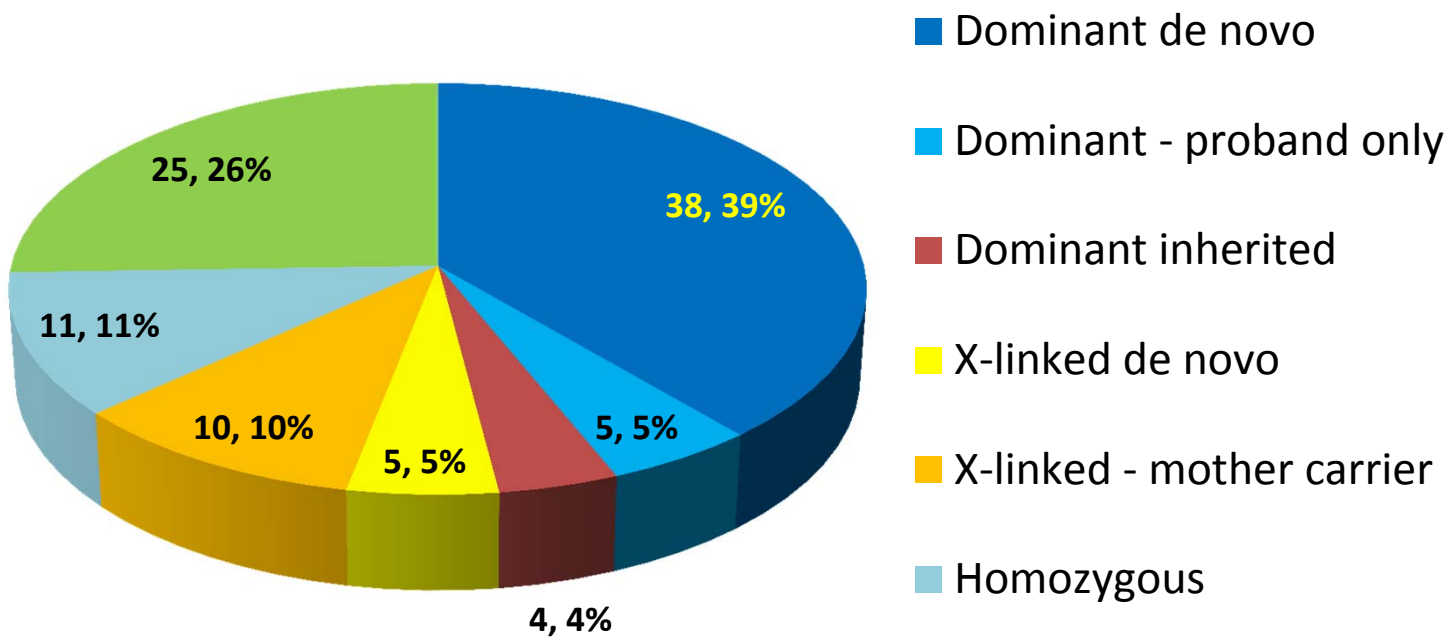
Compound heterozygous/homozygous variants (30%) are reported for autosomal recessive conditions related to the patient's symptoms.

X-linked mutations are 10%

Diagnostic Yield

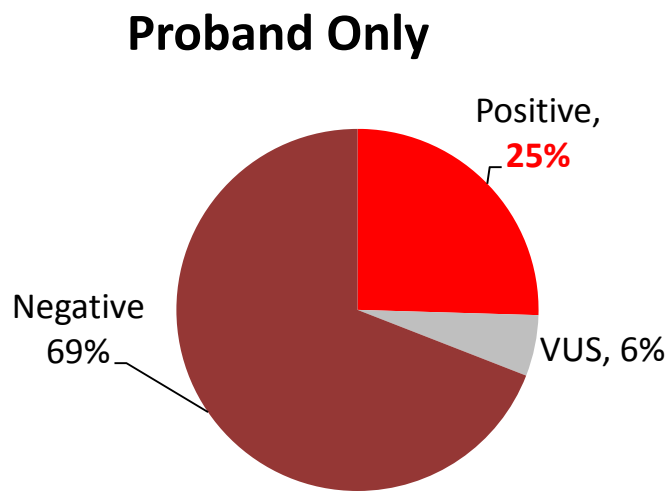


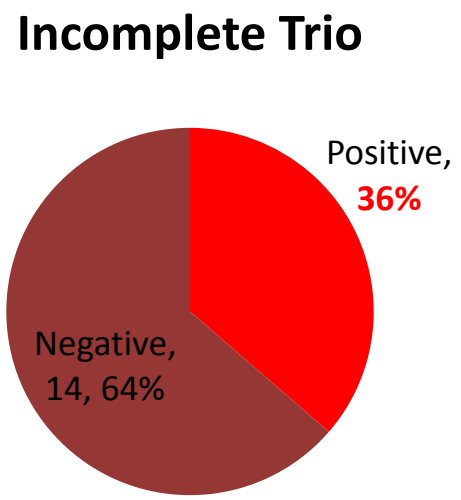
Inheritance Pattern Positive Cases

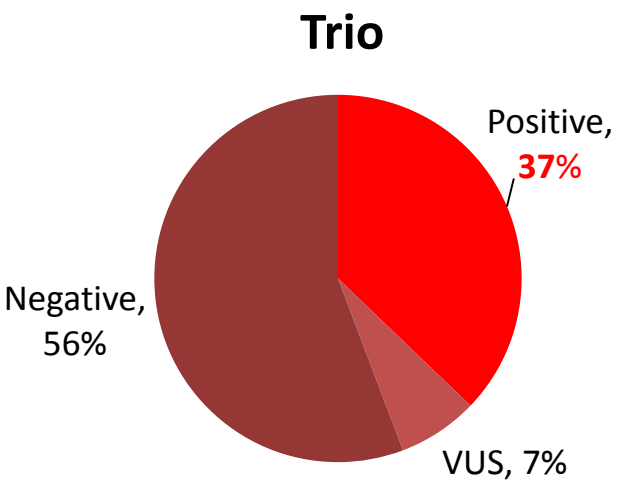


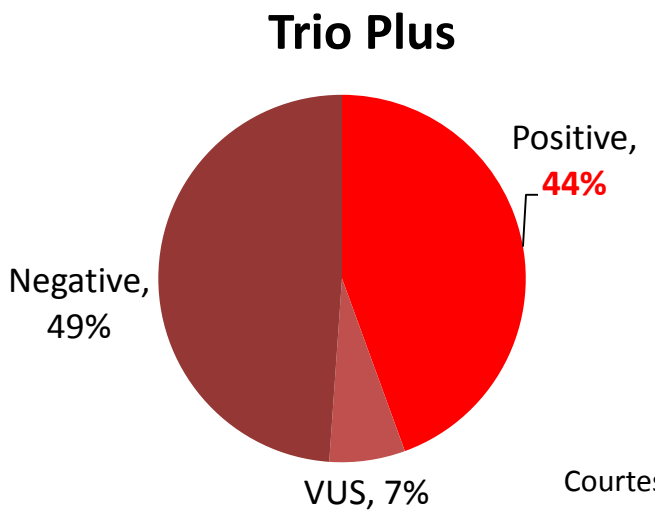
Courtesy of Tatiana Tvrdik

Proband Only vs Trios Diagnostic Yield







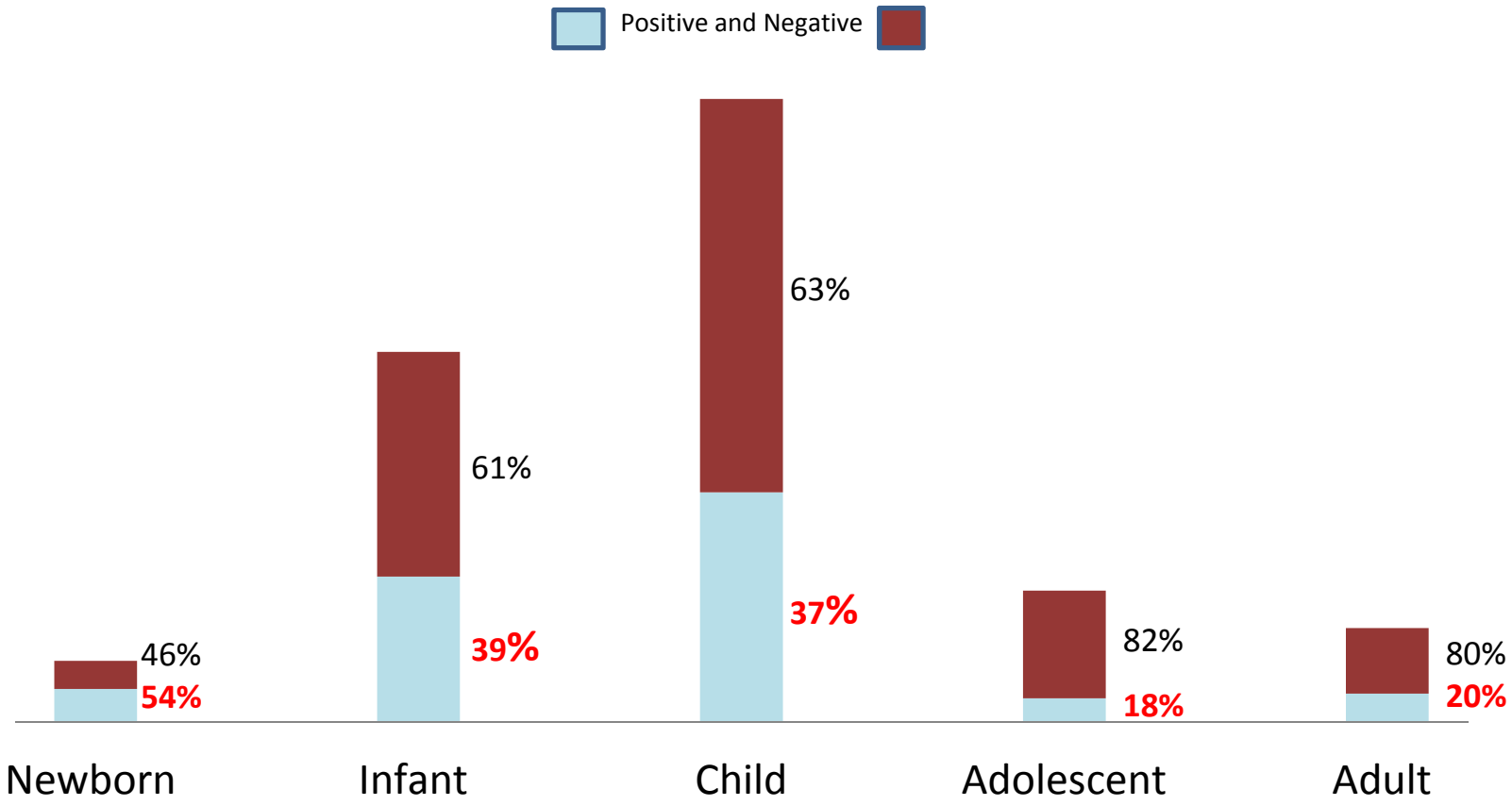


Courtesy of Tatiana Tvrdik

Power of Trio in Exome Testing

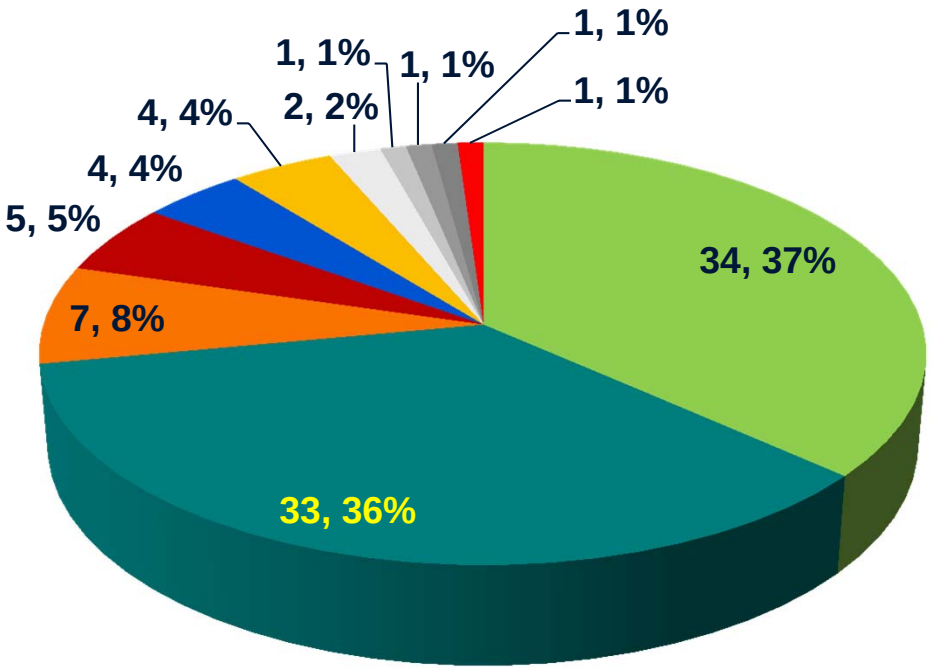
- De novo variants
- Potential to identify parent-of-origin of de novo variants
- Compound heterozygotes and complex variants
- Homozygous vs apparent homozygous variants
- Reduced number of variants to be considered as causative

Diagnostic Yield by Age



Courtesy of Tatiana Tvrdik

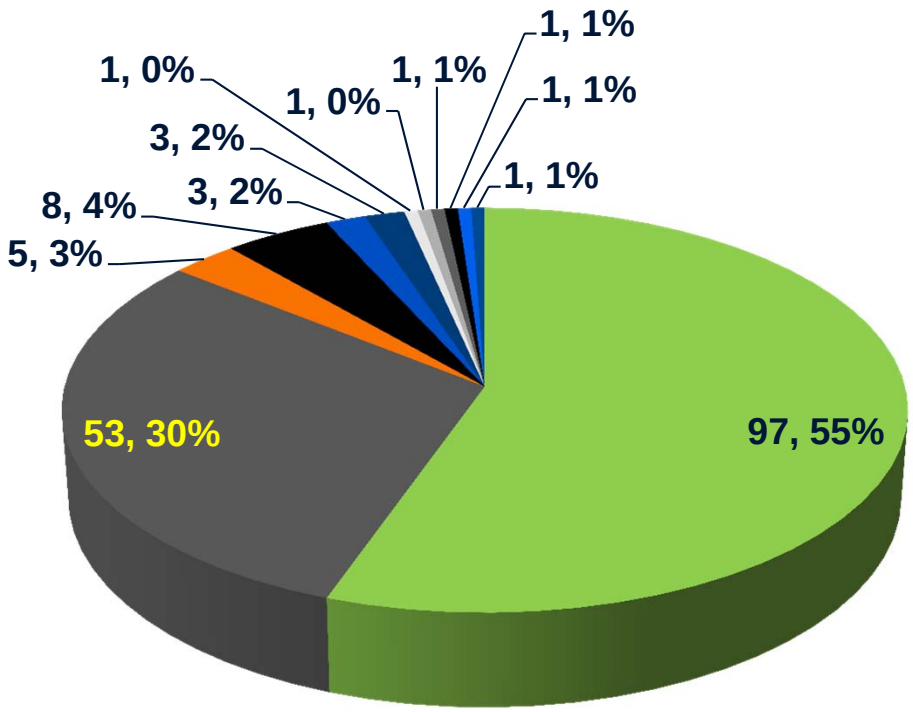
Causative Disorders



- Syndromes
- Neurological
- Muscular
- Vascular
- Metabolic
- Mitochondrial
- Skeletal
- Cilliopathy
- Hearing
- Gastrointestinal
- Hematological

Courtesy of Tatiana Tvrdik

Cases with No Molecular Diagnosis



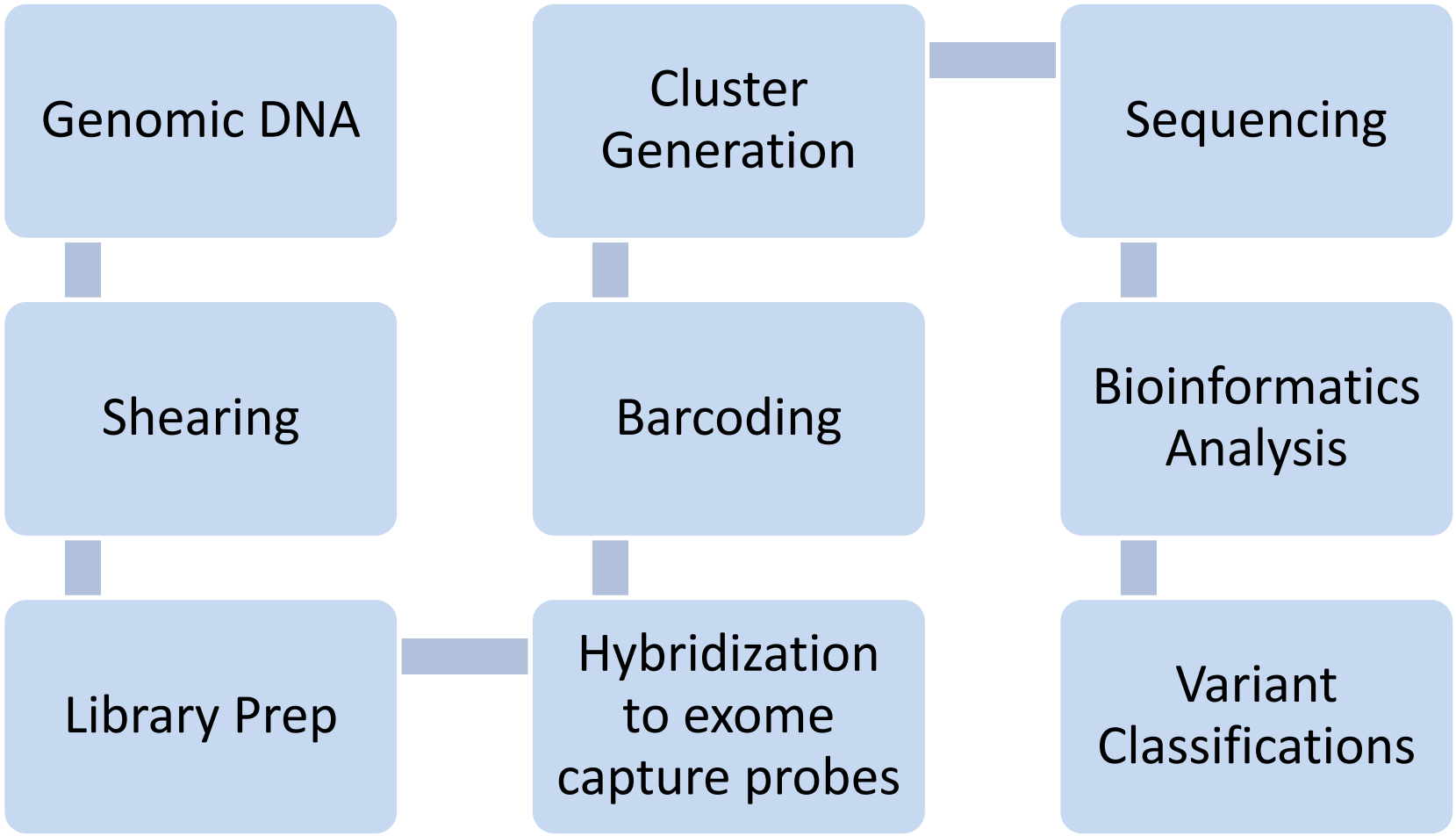
Courtesy of Tatiana Tvrdik

Limitations of Our Exome Sequencing

The following will not be identified:

- Some coding regions, amenable to capture
- Any genetic changes residing outside of the targeted regions
- Repeat expansions
- Low level of mosaicism
- Structural DNA variation: translocations, inversions, insertions/deletions (indels) and copy number variations
- Mitochondrial genome variants

Exome Sequencing Laboratory Workflow

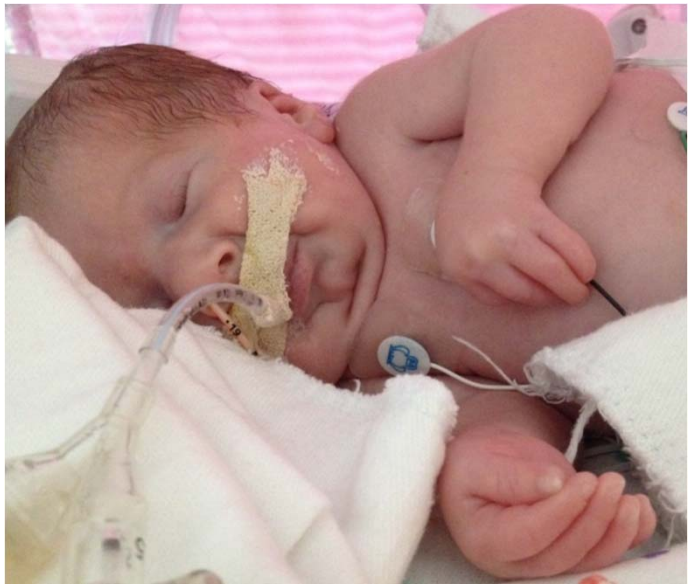


Courtesy of Tatiana Tvrdik

Case Discussion

CASE 1

Dysmorphic features
Narrow palpebral fissures,
blepharophimosis,
prominent nasolabial folds,
small mouth, dimpling on chin,
retrognathia and low-set ears



Distal Arthrogryposis: finger elbow and knee contractures, ulnar deviation, and fixed thumb adduction, difficulty in opening jaw

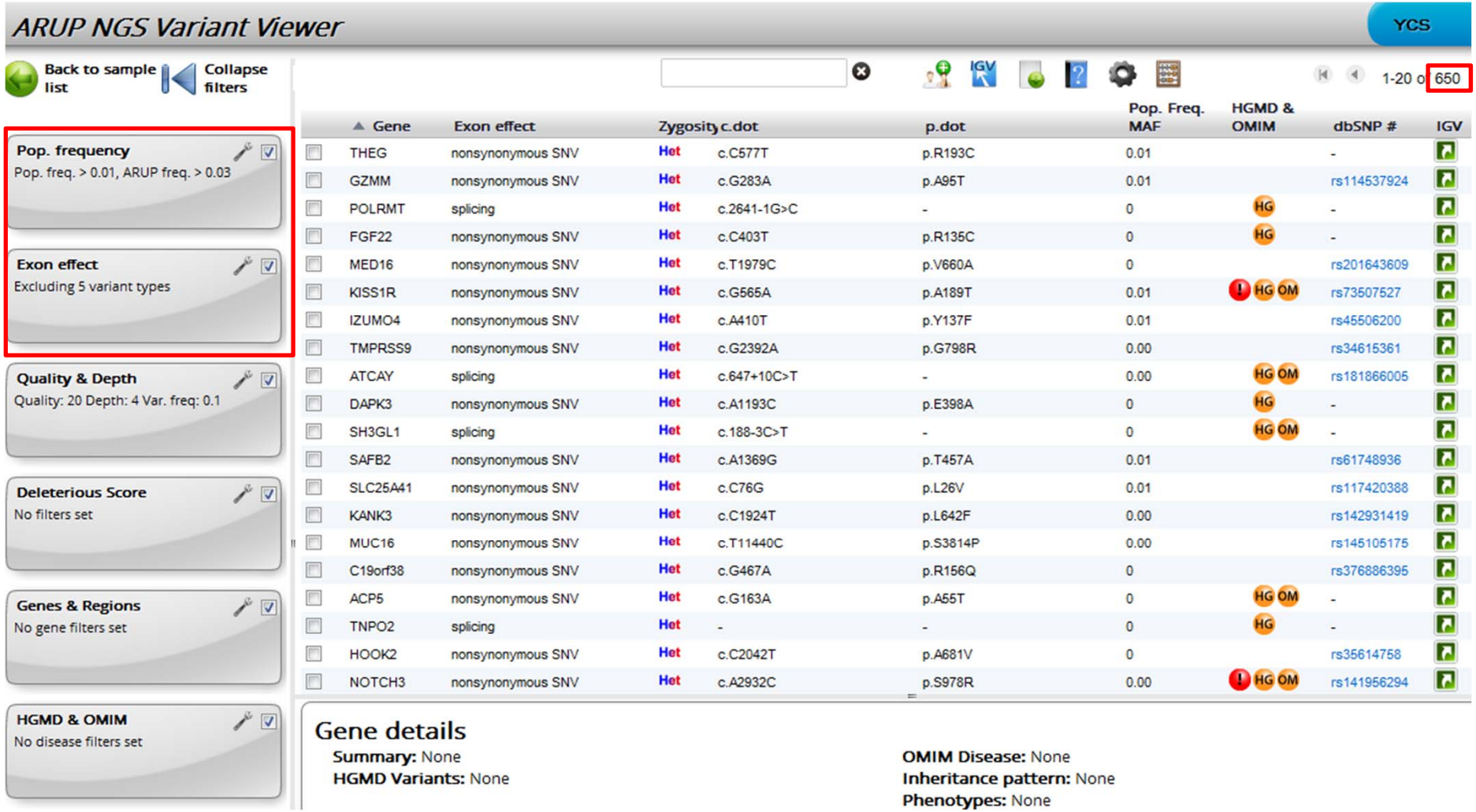
MicroArray : 409kb gain at 4q32.2

Dave Stevenson, MD, Kathryn Swoboda, MD

CLINICAL DIFFERENTIAL DIAGNOSIS

DISORDER	CLINICAL MANIFESTATION	GENETIC BASIS	TEST RESULT
Stuve Wiedmann syndrome (SWS)	Argthrogryposis Long bone bowing Autonomic dysregulation Early death	LIFR Autosomal Recessive	Negative_1 Variant VUS
Freeman Sheldon Syndrome	Face, hands, and feet "whistling face"; chin dimple shaped like an "H" or "V"; malignant hyperthermia	MYH3 Autosomal Dominant	Negative_No disease causing mt noted

FILTERING METHODS

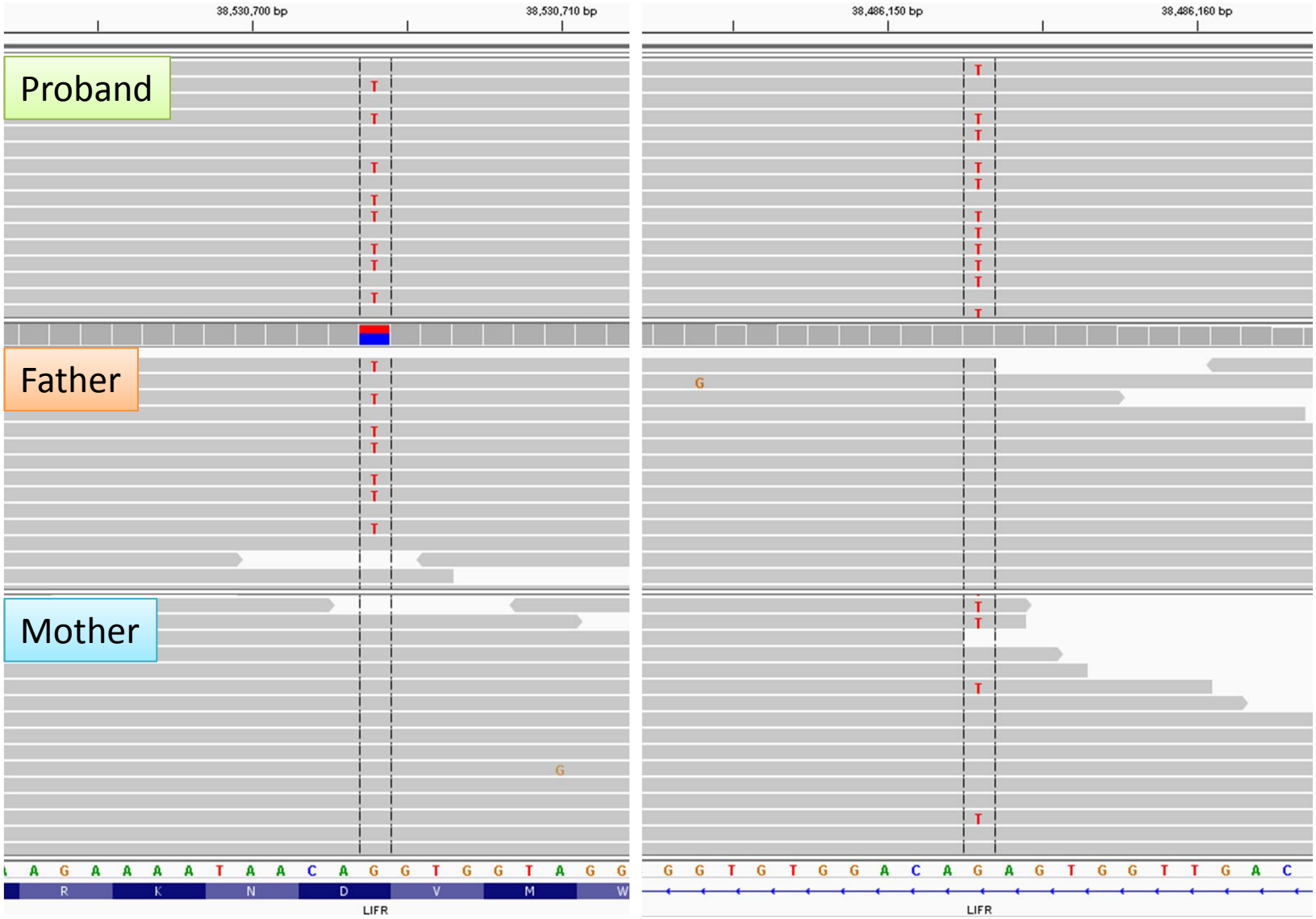


Gene	Exon effect	Zygosity	c.dot	p.dot	Pop. Freq. MAF	HGMD & OMIM	dbSNP #	IGV
THEG	nonsynonymous SNV	Het	c.C577T	p.R193C	0.01		-	
GZMM	nonsynonymous SNV	Het	c.G283A	p.A95T	0.01		rs114537924	
POLRMT	splicing	Het	c.2641-1G>C	-	0	HG	-	
FGF22	nonsynonymous SNV	Het	c.C403T	p.R135C	0	HG	-	
MED16	nonsynonymous SNV	Het	c.T1979C	p.V660A	0		rs201643609	
KISS1R	nonsynonymous SNV	Het	c.G565A	p.A189T	0.01	! HG OM	rs73507527	
IZUMO4	nonsynonymous SNV	Het	c.A410T	p.Y137F	0.01		rs45506200	
TMPRSS9	nonsynonymous SNV	Het	c.G2392A	p.G798R	0.00		rs34615361	
ATCAY	splicing	Het	c.647+10C>T	-	0.00	HG OM	rs181866005	
DAPK3	nonsynonymous SNV	Het	c.A1193C	p.E398A	0	HG	-	
SH3GL1	splicing	Het	c.188-3C>T	-	0	HG OM	-	
SAFB2	nonsynonymous SNV	Het	c.A1369G	p.T457A	0.01		rs61748936	
SLC25A41	nonsynonymous SNV	Het	c.C76G	p.L26V	0.01		rs117420388	
KANK3	nonsynonymous SNV	Het	c.C1924T	p.L642F	0.00		rs142931419	
MUC16	nonsynonymous SNV	Het	c.T11440C	p.S3814P	0.00		rs145105175	
C19orf38	nonsynonymous SNV	Het	c.G467A	p.R156Q	0		rs376886395	
ACP5	nonsynonymous SNV	Het	c.G163A	p.A55T	0	HG OM	-	
TNPO2	splicing	Het	-	-	0	HG	-	
HOOK2	nonsynonymous SNV	Het	c.C2042T	p.A681V	0		rs35614758	
NOTCH3	nonsynonymous SNV	Het	c.A2932C	p.S978R	0.00	! HG OM	rs141956294	

TWO *LIFR* VARIANTS ON DIFFERENT CHROMOSOMES

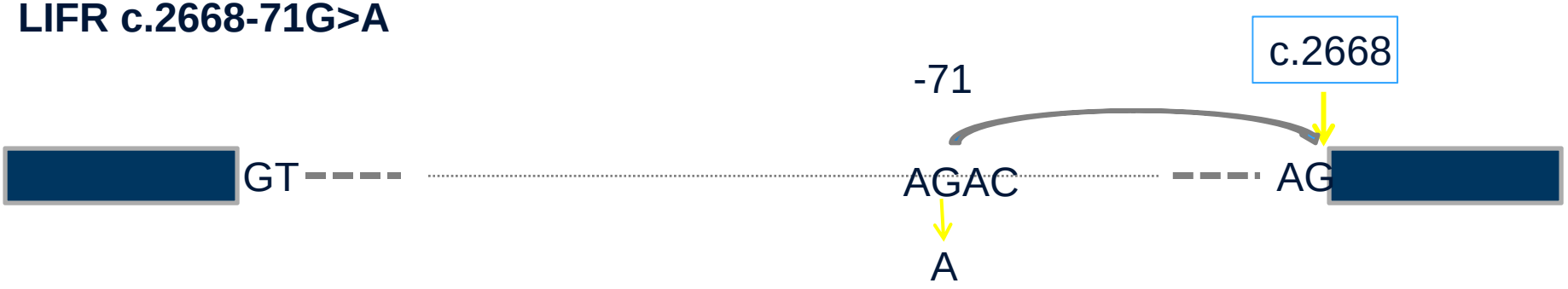
c.46G>A; p.Asp16Asn

2688-71G>A



DONOR SPLICE SITE

LIFR c.2668-71G>A



WT

Splice site predictions for 1 sequence with donor score cutoff 0.10, acc

Donor site predictions for 67.148.8.170.25727.0 :

Start	End	Score	Exon	Intron
92	106	0.42	tctgcag	gtttgctt
236	250	0.16	acagaat	gtgtgtga
354	368	0.14	cgtgggg	gtagagct
370	384	0.78	gaacaag	gtagtggc
435	449	0.29	gcagttg	gtgagaca
444	458	0.59	gagacag	gtgtgggt
471	485	0.44	gaaacgg	gttagaga

c.2668-71G>A

Splice site predictions for 1 sequence with donor score cutoff 0.10, acce

Donor site predictions for 67.148.8.170.25783.0 :

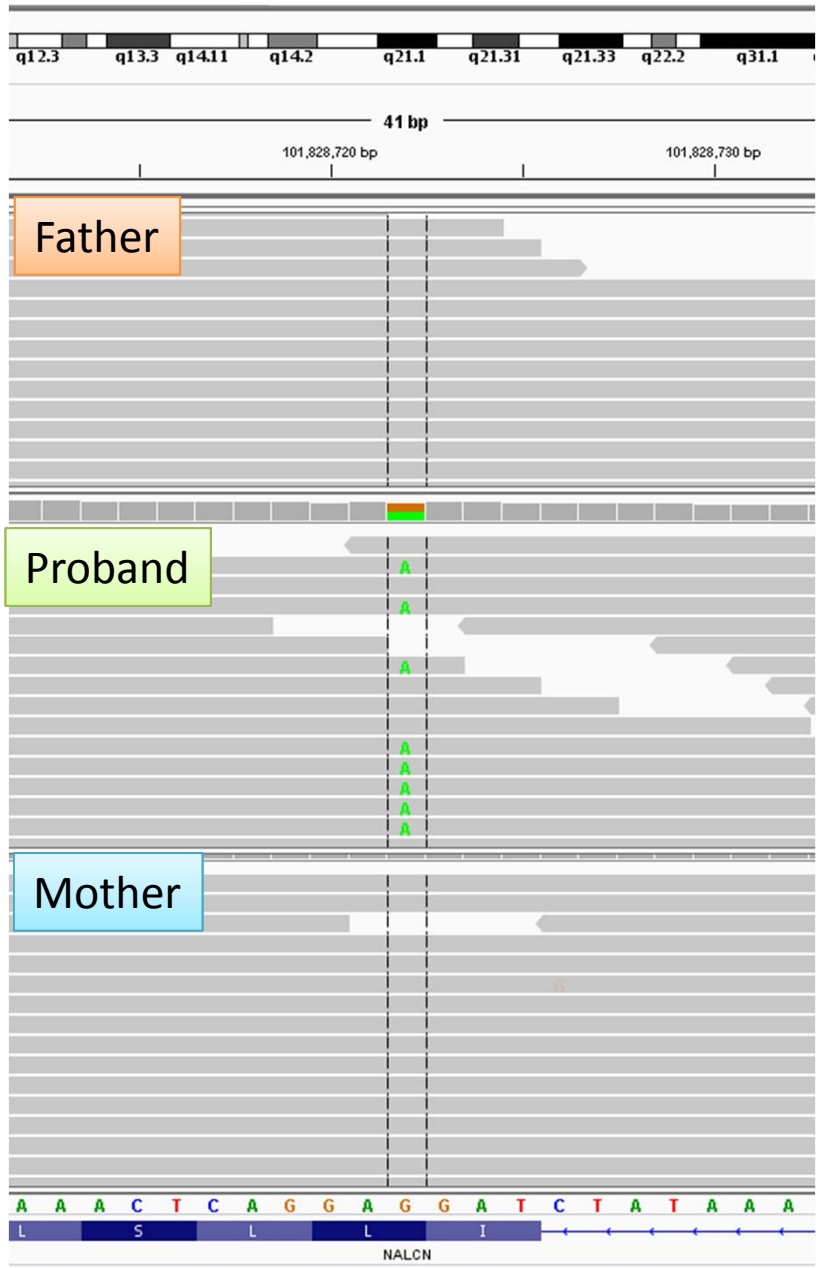
Start	End	Score	Exon	Intron
92	106	0.42	tctgcag	gtttgctt
236	250	0.16	acagaat	gtgtgtga
354	368	0.14	cgtgggg	gtagagct
370	384	0.78	gaacaag	gtagtggc
444	458	0.80	gaaacag	gtgtgggt
471	485	0.44	gaaacgg	gttagaga

REPORTING OF LIFR

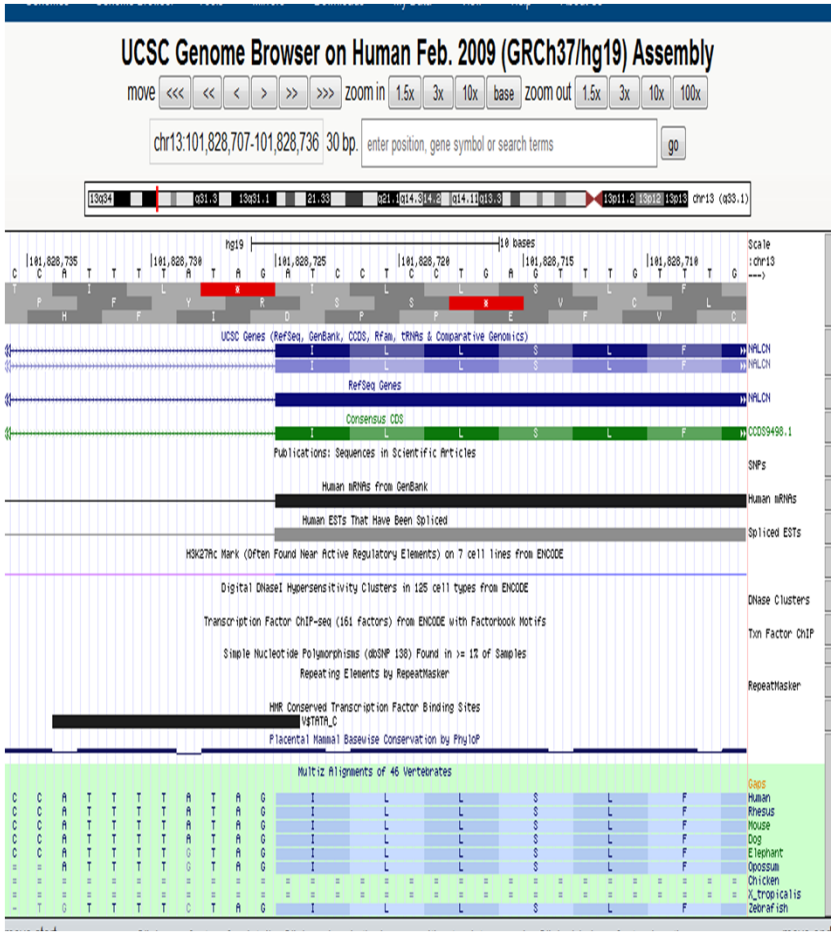
- **Gene: LIFR (NM_002310)**
- **Variant:**
  - **c.46G>A; p.Asp16Asn (one copy) - Variant of Uncertain Significance**
  - **c.2336-71G>A (one copy)- Variant of Uncertain Significance**
- **Inheritance pattern: Autosomal recessive**

DE NOVO MUTATION ?

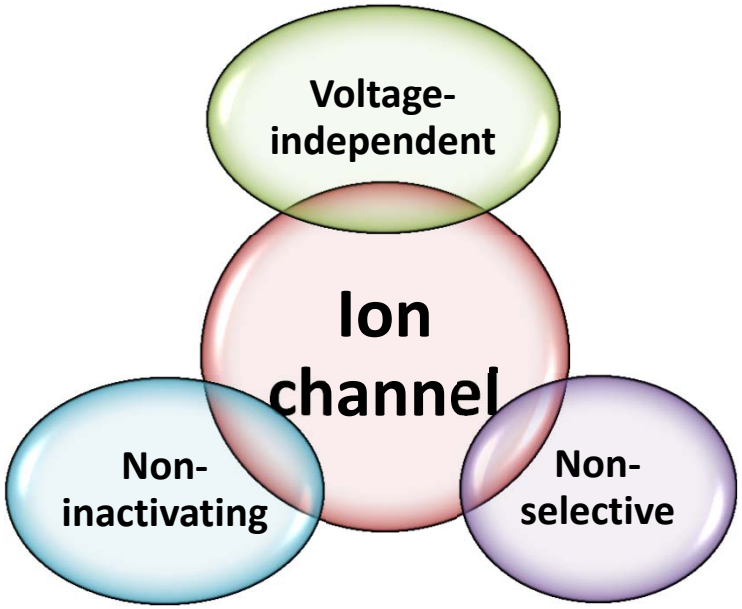
DE NOVO MUTATION: *NALCN* GENE

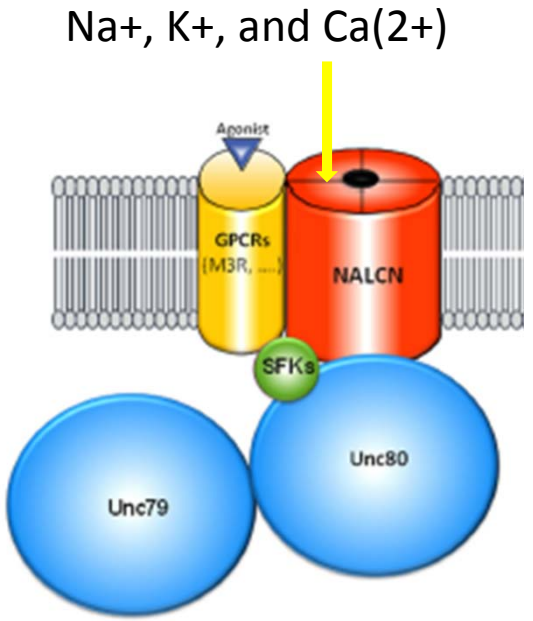


c.1768C>T; p.Leu590Phe



SODIUM LEAK CHANNEL (NALCN)





NALCN
(Cochet-Bissuel 2014)

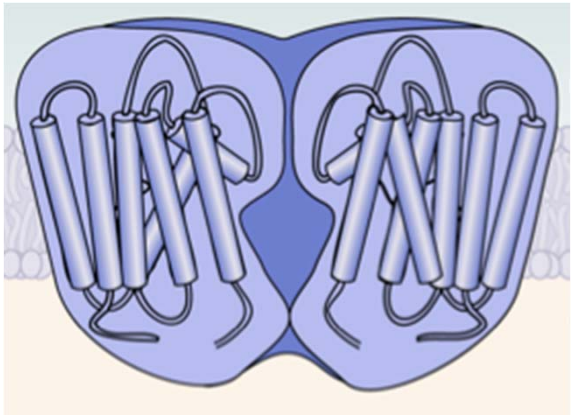
- Mainly expressed in CNS
- Synapse development and synaptic density (Lu et al., 2007)
- KO mice: die of respiratory rhythm

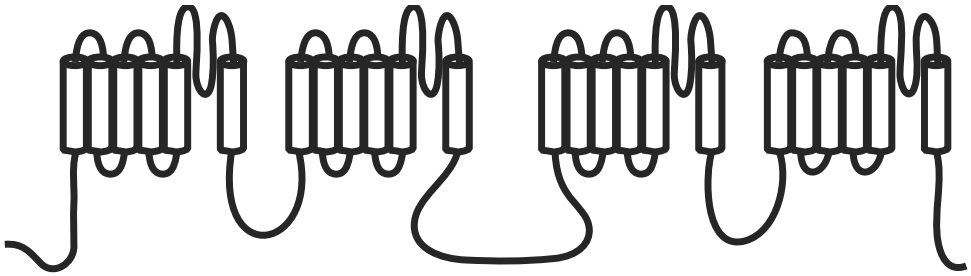
Courtesy of Eric Bend and Erik Jorgensen

NALCN mutations result in:

Infantile hypotonia with psychomotor retardation and facial dysmorphism (IHPRF; MIM #615419)

LOSS-OF-FUNCTION MUTATION	GAIN-OF-FUNCTION MUTATION
Autosomal recessive inheritance	Putative dominant inheritance ?
Mild to severe hypotonia	Hypertonia - distal contractures ?
Viable	Infant mortality ?



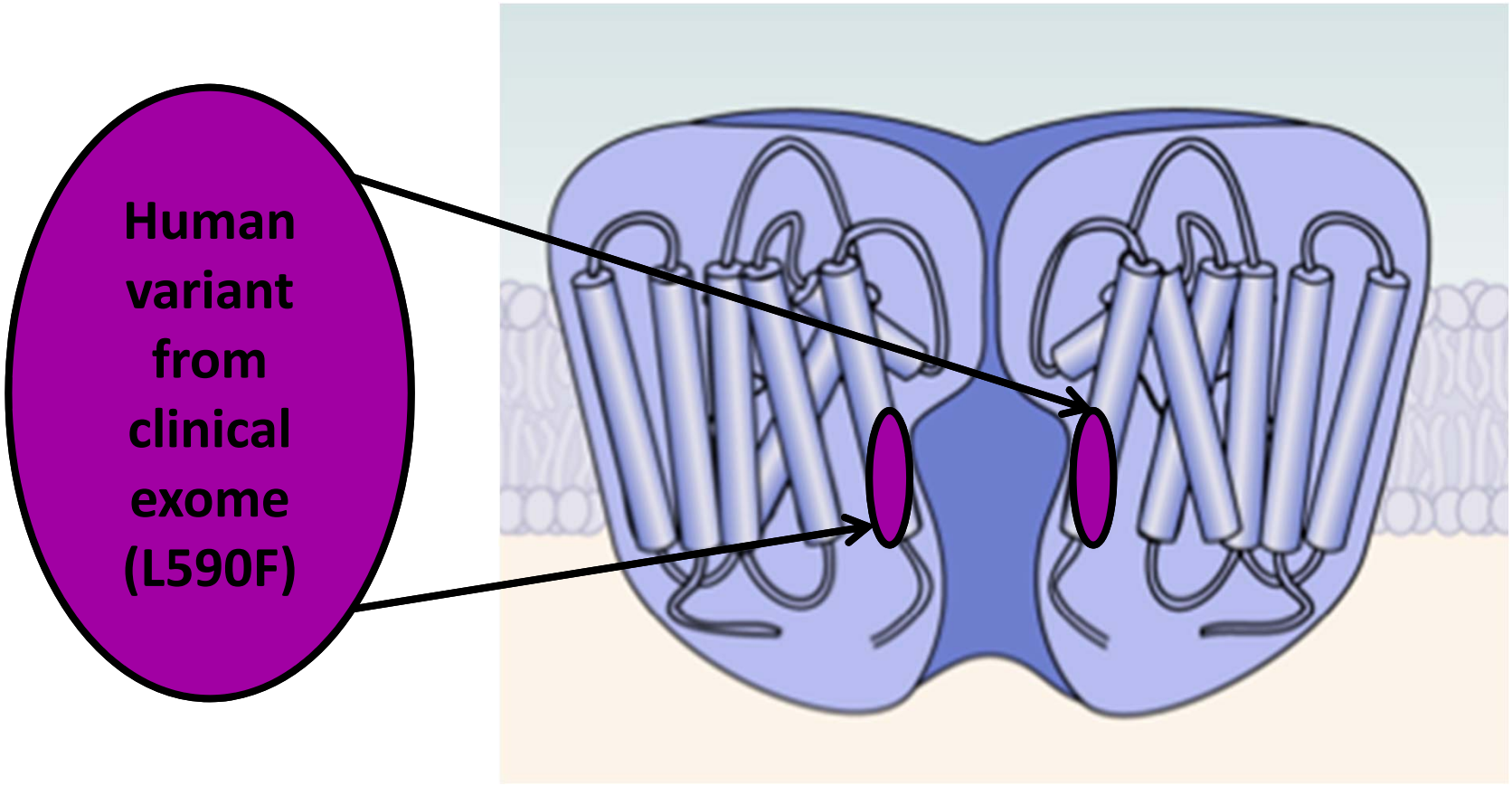


Al-Sayed et al. J Hum Genet 2013

Koroglu et al. J Med Genet 2013

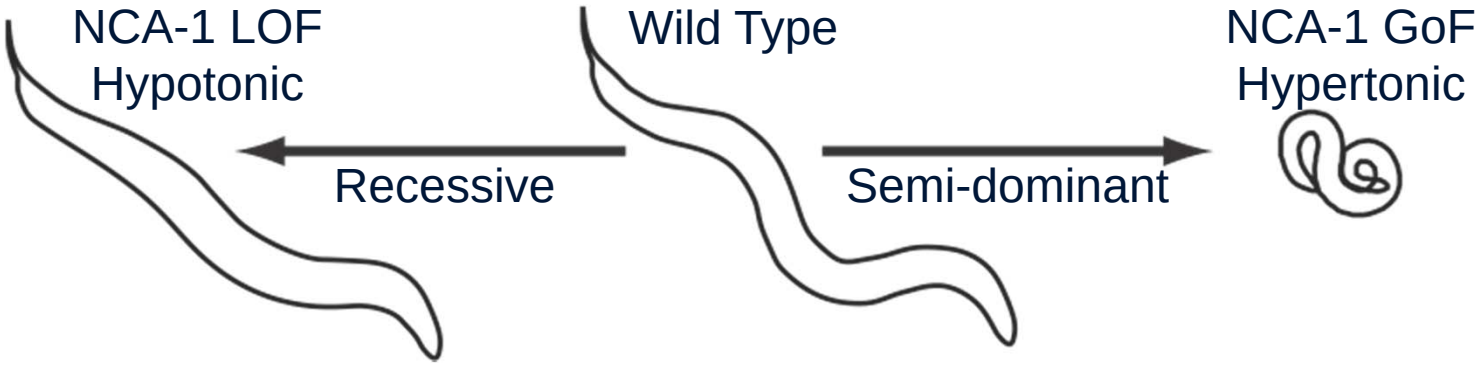
Courtesy of Eric Bend and Erik Jorgensen

THE C-TERMINUS OF THE S6 IS CRITICAL FOR CHANNEL GATING



Courtesy of Eric Bend and Erik Jorgensen

DOES THE NALCN SNP CAUSE A GAIN-OF-FUNCTION CHANNELOPATHY?

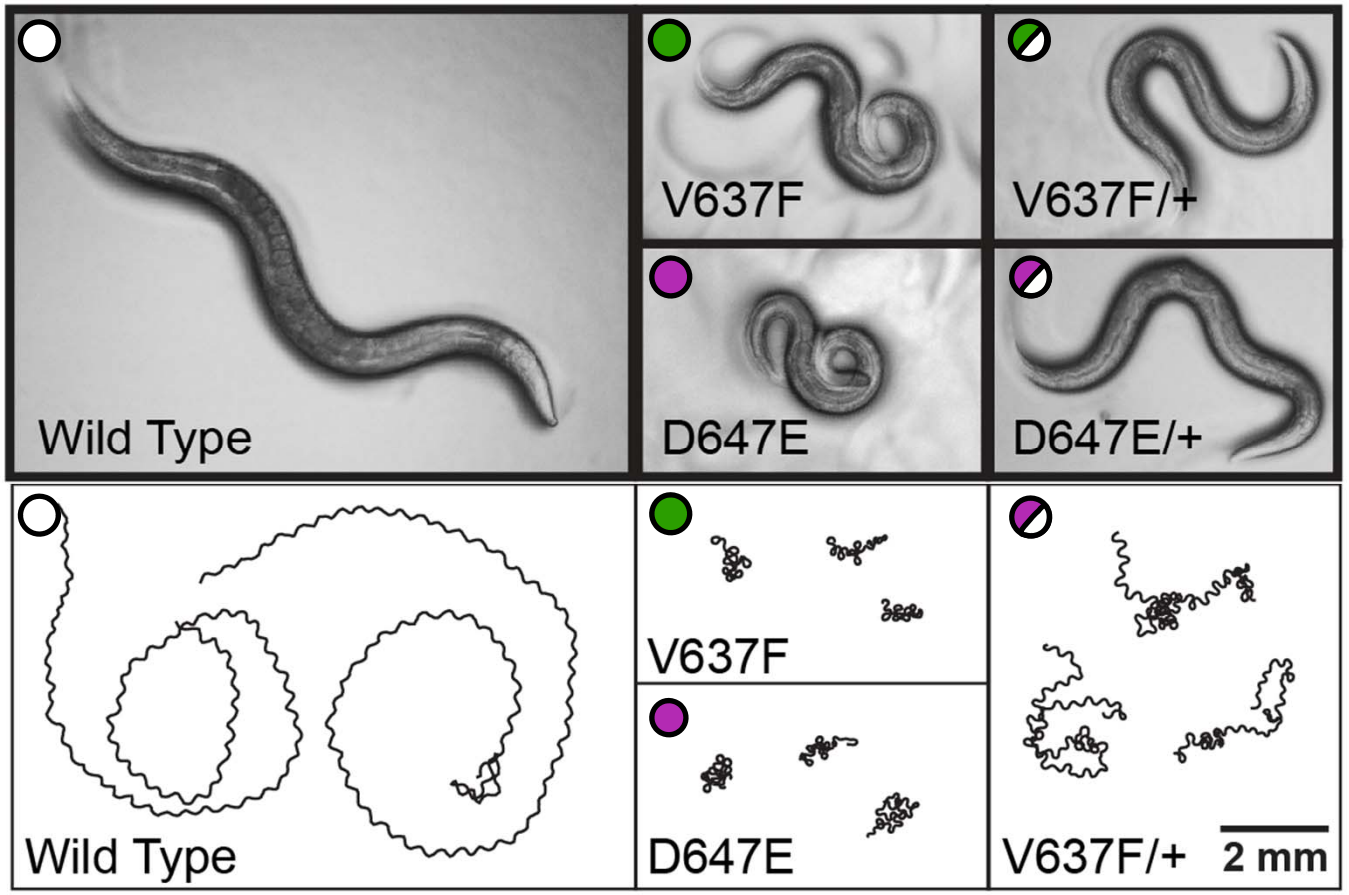


Predictions

- ☑ Dominant inheritance
- ☑ Hypertonia
- ☐ Increased neurotransmission

Courtesy of Eric Bend and Erik Jorgens

THE HUMAN SNP PHENOCOPIES A GAIN-OF-FUNCTION NCA-1 ALLELE



- Wild Type
- Gain-of-Function
- Human SNP

Courtesy of Eric Bend and Erik Jorgensen

Am J Hum Genet. 2015 Mar 5;96(3):462-73. doi: 10.1016/j.ajhg.2015.01.003. Epub 2015 Feb 12.

De novo mutations in NALCN cause a syndrome characterized by congenital contractures of the limbs and face, hypotonia, and developmental delay.

Chong JX[1], McMillin MJ[1], Shively KM[1], Beck AE[1], Marvin CT[1], Armenteros JR[1], Buckingham KJ[1], Nkinsi NT[1], Boyle EA[2], Berry MN[3], Bocian M[4], Foulds N[5], Uzielli ML[6], Haldeman-Englert C[3], Hennekam RC[7], Kaplan P[8], Kline AD[9], Mercer CL[5], Nowaczyk MJ[10], Klein Wassink-Ruiter JS[11], McPherson EW[12], Moreno RA[13], Scheuerle AE[14], Shashi V[15], Stevens CA[16], Carey JC[17], Monteil A[18], Lory P[18], Tabor HK[19], Smith JD[2], Shendure J[2], Nickerson DA[2]; University of Washington Center for Mendelian Genomics, Bamshad MJ[20].

Neurology. 2016 Sep 13;87(11):1131-9. doi: 10.1212/WNL.0000000000003095. Epub 2016 Aug 24.

NALCN channelopathies: Distinguishing gain-of-function and loss-of-function mutations.

Bend EG[1], Si Y[1], Stevenson DA[1], Bayrak-Toydemir P[1], Newcomb TM[1], Jorgensen EM[2], Swoboda KJ[2].

CASE 2

Public database filtering

CASE 2:

a 7-year-old boy of hispanic/native american/caucasian ancestry

Clinical Findings:

Pre and postnatal overgrowth,
Moderate ID,
Not typical Sotos face,
Advanced bone age,
History of laryngomalacia,
Hypotonia,
No history of seizure,
Mild optic nerve hypoplasia

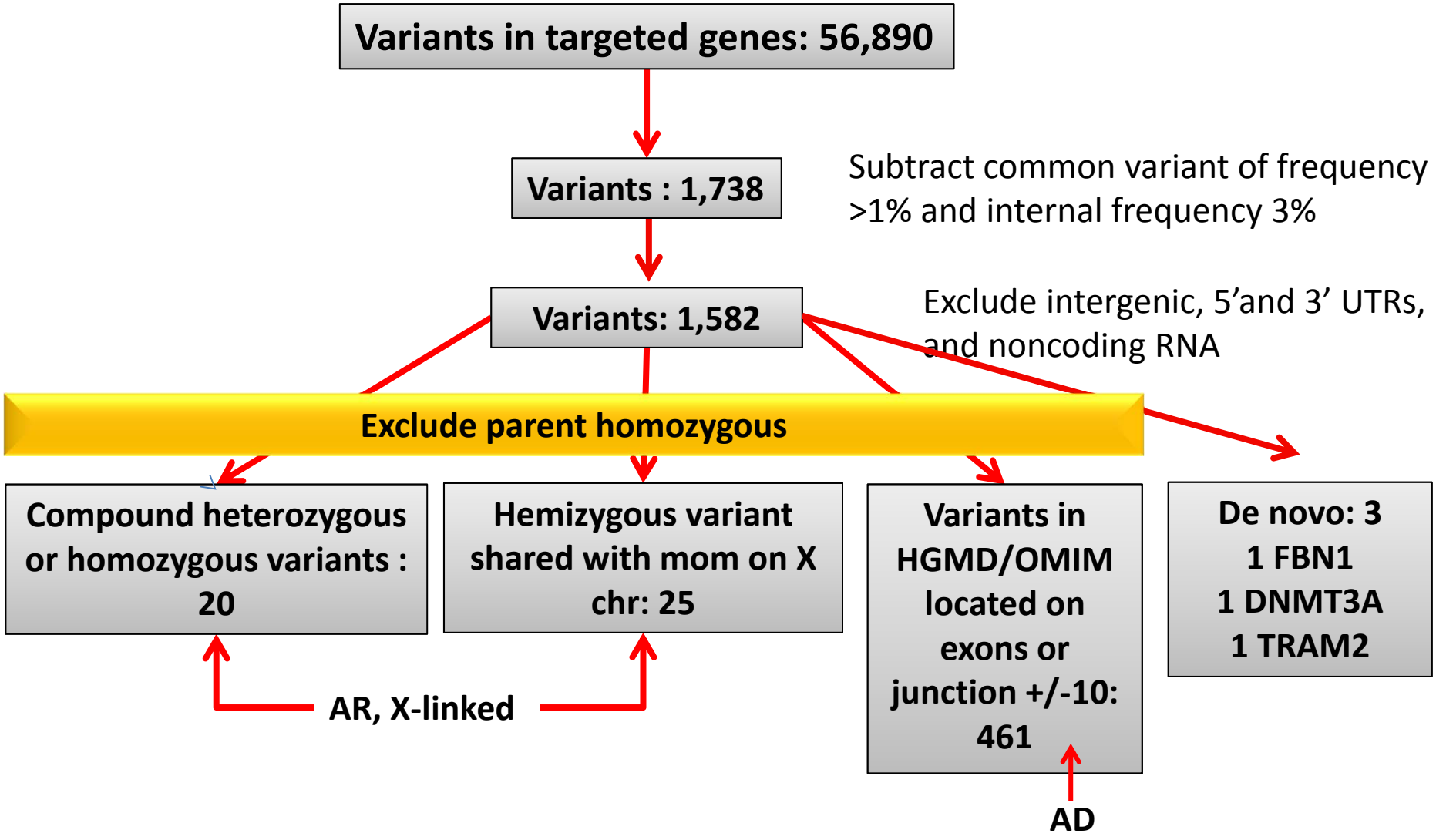
Other Testing Results:

MRI showed a small optic chiasm, focal encephalomalacia or dilated perivascular spaces.

The patient had a normal genomic microarray.

John Carey, MD

EXOME DATA


Variants in targeted genes: 56,890
Variants : 1,738
Subtract common variant of frequency >1% and internal frequency 3%
Variants: 1,582
Exclude intergenic, 5'and 3' UTRs, and noncoding RNA
Exclude parent homozygous
Compound heterozygous or homozygous variants : 20
Hemizygous variant shared with mom on X chr: 25
Variants in HGMD/OMIM located on exons or junction +/-10: 461
De novo: 3
1 FBN1
1 DNMT3A
1 TRAM2
AR, X-linked
AD

De Novo Variant: *FBN1* c.4894C>T,p.Arg1632Cys

This FBN1 variant (p.Arg1632Cys) alters a moderately conserved amino acid and creates an extra cysteine residue between cysteine residues 4 and 5 (Cys1631 and Cys1633) in the EGF-like calcium-binding domain 27.

FBN1 protein contains 47 epidermal growth factor (EGF)-like domains which are characterized by six conserved cysteine residues. These six cysteine residues form three disulfide bonds that are critical for the normal protein structure of FBN1.

Cysteine substitutions that disrupt one of the three disulfide bonds are frequent causes of Marfan syndrome.

De Novo Variant: *FBN1* c.4894C>T,p.Arg1632Cys

Computational analyses predict that this *FBN1* variant (p.Arg1632Cys) will affect protein function (SIFT: deleterious, MutationTaster: disease causing, PolyPhen-2: probably damaging).

In addition, it is only reported in one individual in the Exome Aggregation Consortium database (1 out of 121378 alleles).

Although this particular *FBN1* variant (p.Arg1632Cys) has not been reported in the literature, a different amino acid alteration at the same codon (p.Arg1632His) has been reported in a patient that met Ghent criteria for Marfan syndrome with ocular findings and no skeletal or cardiovascular findings .

SECOND DE NOVO VARIANT

*602769

DNA METHYLTRANSFERASE 3A; DNMT3A

HGNC Approved Gene Symbol: DNMT3A

Cytogenetic location: 2p23.3 *Genomic coordinates (GRCh38): 2:25,232,960-25,342,589* (from NCBI)

Gene-Phenotype Relationships

Location	Phenotype	Phenotype MIM number	Inheritance (in progress)	Phenotype mapping key
2p23.3	Tatton-Brown-Rahman syndrome	615879	AD	3

#615879

TATTON-BROWN-RAHMAN SYNDROME; TBRS

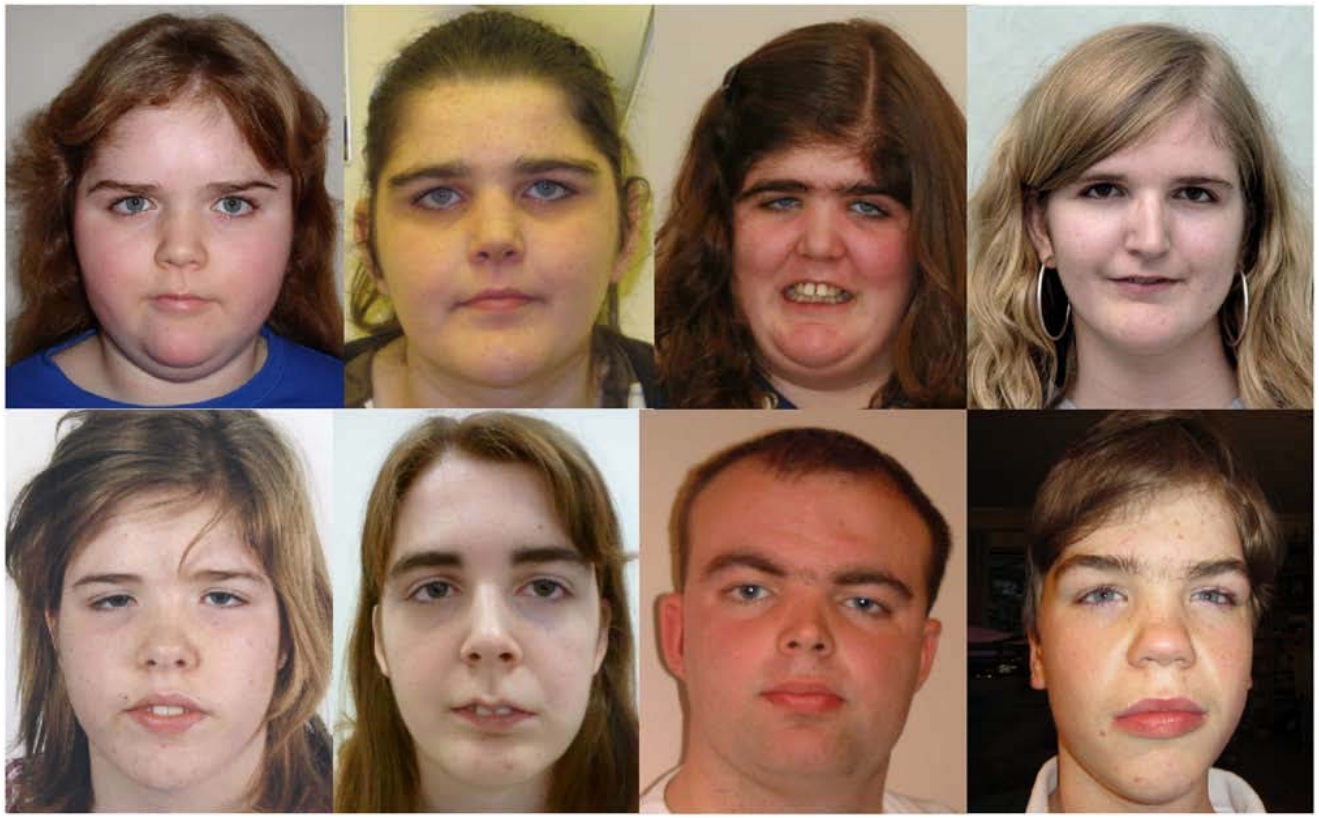
CATEGORY	SUBCATEGORY	FEATURES
Inheritance	-	Autosomal dominant
Growth	Height	Tall stature (+3 S.D)
Head and Neck	Head	Large head circumference (+2.5 SD)
	Face	Round face [EoM image]
	Eyes	Heavy horizontal eyebrows Narrow palpebral fissures
Cardiovascular	Heart	Atrial septal defect (less common)
Abdomen	External Features	Umbilical hernia (less common)
Skeletal	Spine	Scoliosis (less common)
Neurologic	Central Nervous System	Intellectual disability, mild to moderate Seizures (less common)
Miscellaneous	-	All reported cases result from de novo mutation (last curated July 2014)
Molecular Basis	-	Caused by mutation in the DNA methyltransferase 3A gene (DNMT3A, 602769.0001)

#615879

TATTON-BROWN-RAHMAN SYNDROME; TBRS

CATEGORY	SUBCATEGORY	FEATURES
Inheritance	-	Autosomal dominant
Growth	Height	Tall stature (+3 S.D)
Head and Neck	Head	Large head circumference (+2.5 SD)
	Face	Round face [EoM image]
	Eyes	Heavy horizontal eyebrows Narrow palpebral fissures
Cardiovascular	Heart	Atrial septal defect (less common)
Abdomen	External Features	Umbilical hernia (less common)
Skeletal	Spine	Scoliosis (less common)
Neurologic	Central Nervous System	Intellectual disability, mild to moderate Seizures (less common)
Miscellaneous	-	All reported cases result from de novo mutation (last curated July 2014)
Molecular Basis	-	Caused by mutation in the DNA methyltransferase 3A gene (DNMT3A, 602769.0001)

Characteristic facial appearance in DNMT3A overgrowth syndrome



Tatton-Brown et al. Nat Genet 2014

Pathogenic DNMT3A variant is present in a "normal" population database

Annotations

This variant falls on 10 transcripts in 1 genes:

missense

- DNMT3A Transcripts

3' UTR

- DNMT3A - ENST00000380756

Note: This list may not include additional transcripts in the same gene that the variant does not overlap.

Population Frequencies

Population	Allele Count	Allele Number	Number of Homozygotes	Allele Frequency
African	9	10352	0	0.0008694
European (Finnish)	5	6600	0	0.0007576
East Asian	6	8648	0	0.0006938
European (Non-Finnish)	38	66594	0	0.0005706
South Asian	8	16480	0	0.0004854
Latino	0	11544	0	0
Other	0	904	0	0
Total	**66**	**121122**	**0**	**0.0005449**

Information about 66 individuals from ExAC Database

- **Median age 60**
- **Age range 55-80**
- **Allelic ratio: Ranges 10-48%**
- **11 individuals–Somatic data from cancer tissue**
- **55 individuals- Germline data**

DNMT3A: c. 2645G>A, p.Arg882His

It affects the highly conserved methyltransferase domain and reduces methyltransferase activity by approximately 80% compared to the wild type protein, which results in focal hypomethylation at specific CpG sites throughout genome

Somatic *DNMT3A* variants are commonly found in patients with hematologic malignancies and in patients with age-related clonal hematopoiesis without overt disease but with increased risk for subsequent development of a hematologic malignancy.

The p.Arg882His variant is the most common somatic variant of *DNMT3A* observed in patients with age-related clonal hematopoiesis or hematologic malignancies

GUIDELINES/REGULATIONS
CLIA/CAP/ACMG

Next Generation Sequencing

Next Generation Sequencing (NGS) incorporates two processes: (1) the analytical wet bench process of sample and library preparation and sequence generation and (2) the bioinformatics process or pipeline of sequence alignment, annotation and variant calling. These two processes are inextricably linked as the output from each process supports the optimization of the other. The large volumes of data produced by NGS platforms put substantial demands on laboratories in terms of the requirements for documentation, validation, quality control and assurance, monitoring, data storage, as well as assessment and implementation of new technology and software releases.

Inspector Instructions:

- Sampling of next generation sequencing policies and procedures
- Records of wet bench processing and bioinformatics process validation
- QM program records with corrective action for component failure
- Sampling of exception log records

Guide validation of samples, analysis and reporting

Clinical Laboratory Standards for Next Generation Sequencing

Heidi L. Rehm, PhD[1,2], Sherri J Bale, PhD[3], Pinar Bayrak-Toydemir, PhD[4], Jonathan S Berg, MD[5], Kerry K Brown, PhD[6], Joshua L Deignan, PhD[7], Michael J Friez, PhD[8], Birgit H Funke, PhD[1,2], Madhuri R Hegde, PhD[9], Elaine Lyon, PhD[5], A Working Group of the American College of Medical Genetics and Genomics Laboratory Quality Assurance Committee

Clinical Laboratory Improvement Amendments

Assuring the Quality of Next-Generation Sequencing in Clinical Laboratory Practice

Next-generation Sequencing: Standardization of Clinical Testing (Nex-StoCT) Workgroup Principles and Guidelines

Supplementary Guidelines

American College of Medical Genetics and Genomics

ACMG Recommendations for Reporting of Incidental Findings in Clinical Exome and Genome Sequencing

Robert C. Green, MD, MPH[1,2], Jonathan S. Berg, MD, PhD[3], Wayne W. Grody, MD, PhD[4-6], Sarah S. Kalia, ScM, CGC[1], Bruce R. Korf, MD, PhD[7], Christa L. Martin, PhD, FACMG[8], Amy McGuire, JD, PhD[9], Robert L. Nussbaum, MD[10], Julianne M. O'Daniel, MS, CGC[11], Kelly E. Ormond, MS, CGC[12], Heidi L. Rehm, PhD, FACMG[2,13], Michael S. Watson, MS, PhD, FACMG[14], Marc S. Williams, MD, FACMG[15], Leslie G. Biesecker, MD[16]

"Direct laboratories to return with each genomic sequencing order results from 57 genes in which mutations greatly increase risk of 24 serious, but treatable diseases, even if clinicians do not suspect patients have them."

American College of Medical Genetics and Genomics

ACMG Recommendations for Reporting of Incidental Findings in Clinical Exome and Genome Sequencing

Robert C. Green, MD, MPH[1,2], Jonathan S. Berg, MD, PhD[3], Wayne W. Grody, MD, PhD[4-6], Sarah S. Kalia, ScM, CGC[1], Bruce R. Korf, MD, PhD[7], Christa L. Martin, PhD, FACMG[8], Amy McGuire, JD, PhD[9], Robert L. Nussbaum, MD[10], Julianne M. O'Daniel, MS, CGC[11], Kelly E. Ormond, MS, CGC[12], Heidi L. Rehm, PhD, FACMG[2,13], Michael S. Watson, MS, PhD, FACMG[14], Marc S. Williams, MD, FACMG[15], Leslie G. Biesecker, MD[16]

59 genes

"Direct laboratories to return with each genomic sequencing order results from ~~57~~ genes in which mutations greatly increase risk of 24 serious, but treatable diseases, even if clinicians do not suspect patients have them."

What are incidental (or secondary) findings?

Variants found by exome/genome sequencing , which are unrelated to the disease of interest

- majority of them are benign
- a small number of them (between 1-5) might be well-described, disease-associated mutations

Incidental Findings

The ACMG Working Group recommended that the laboratory **actively search** for the specified types of mutations in the specified genes listed in these recommendations.

Mandatory reporting known mutations for the disorders:
- Hereditary cancers,
- Marfan syndrome,
- Long QT syndrome,
- Brugada syndrome,
- Certain cardiomyopathies

Returning incidental findings in children

"Recommendations for seeking and reporting incidental findings not be limited by the age of the person being sequenced.

The ethical concerns about providing children with genetic risk information about adult-onset diseases were outweighed by the potential benefit to the future health of the child and the child's parent of discovering an incidental finding where intervention might be possible."

Patient Consent and Opt-in/out option

- Proband and family members need to consent for exome sequencing and incidental finding

- the ACMG Working Group revised document offers the patient a preference as to whether or not to receive the minimum list of incidental findings described in these recommendations.

Around 90% of cases would like to receive secondary findings

ARUP secondary finding frequency is 1-2%

Conclusion

- Clinical exome sequencing is effective to diagnosis heterogeneous disorders, non-specific or atypical presentation, especially for neurological and neuromuscular disorders

- Sensitivity depends on
  - Medical Exome enrichment
  - Including intronic regions and promoter regions to our bed file
  - Collaboration with clinicians
  - Follow up functional studies

- Quality control measures, data analysis and reporting of incidental findings will continue to evolve and improve