



# Accurate and Rapid Genome Interpretation – in clinical care

---

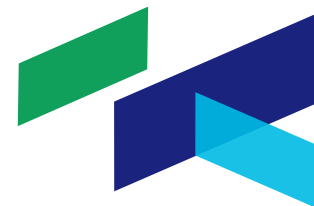
Martin Reese, Ph.D.  
Founder and CEO



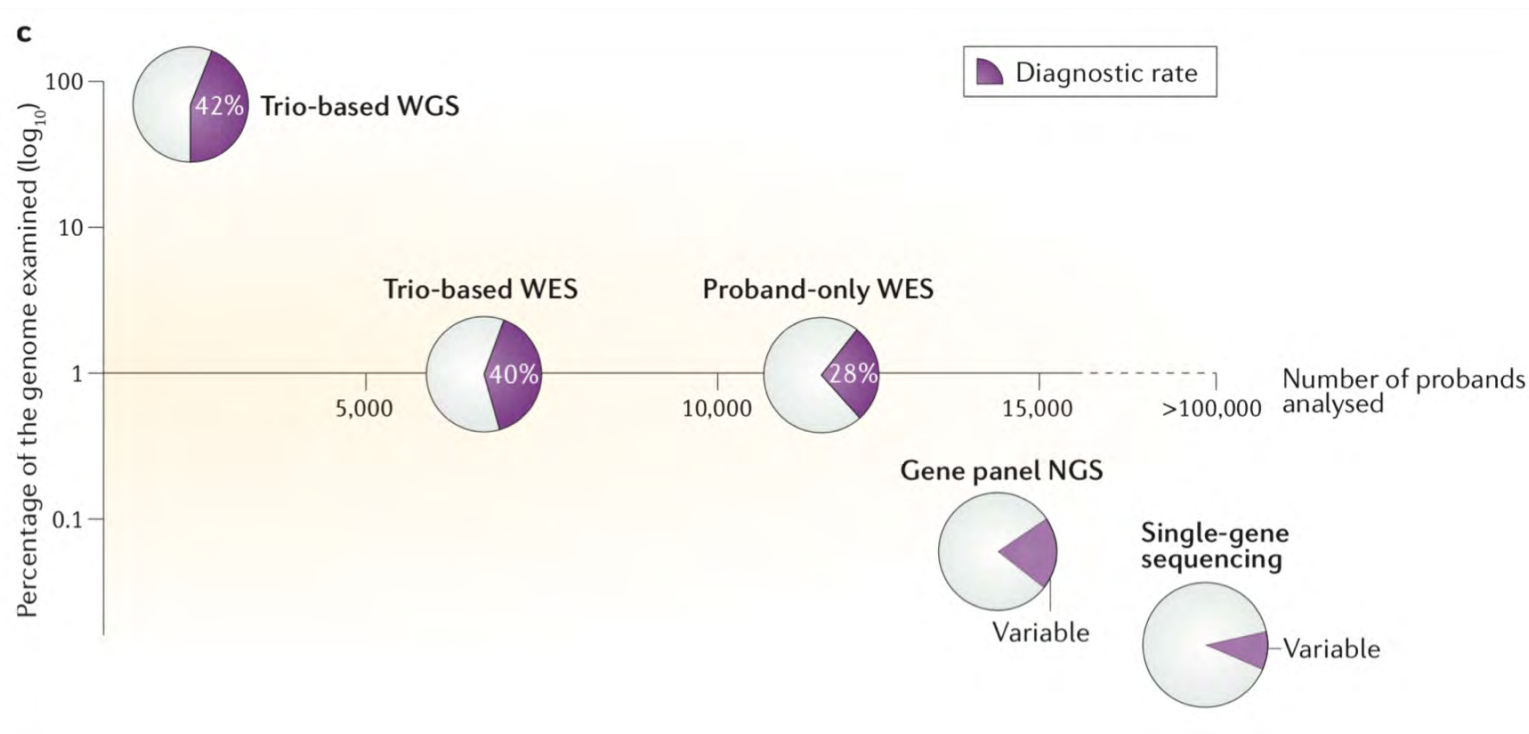
# Overview

---

- What is genome variant interpretation?
- Optimized variant ranking VAAST/Phevor - ready to scale
  - For Clinical Diagnostics and Gene Discovery on a population scale
- Collaborative clinical variant sharing for improved diagnosis
- Two clinical genomics examples
  - Undiagnosed genetic disease patients – Genomics England 100,000 Genome Project
  - First line diagnosis for NICU/ICU babies –Rady’s Children Hospital (Dr. Kingsmore)



# Diagnostic rate of genetic diseases



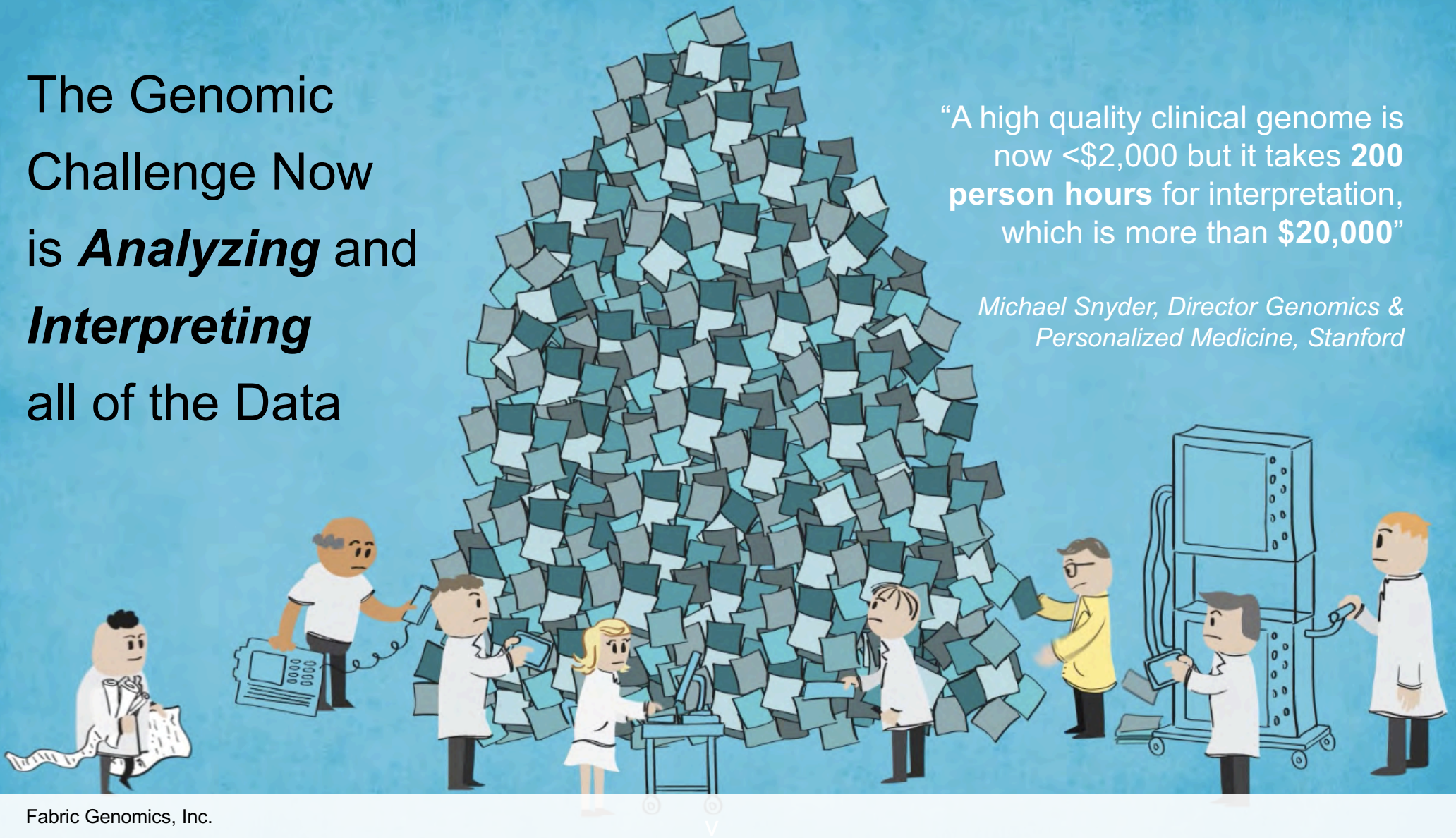
Wright, C. F., FitzPatrick, D. R. & Firth, H. V. Paediatric genomics: diagnosing rare disease in children. *Nature Publishing Group* **10**, 1–16 (2018).



The Genomic  
Challenge Now  
is *Analyzing* and  
*Interpreting*  
all of the Data

“A high quality clinical genome is  
now <\$2,000 but it takes **200  
person hours** for interpretation,  
which is more than **\$20,000**”

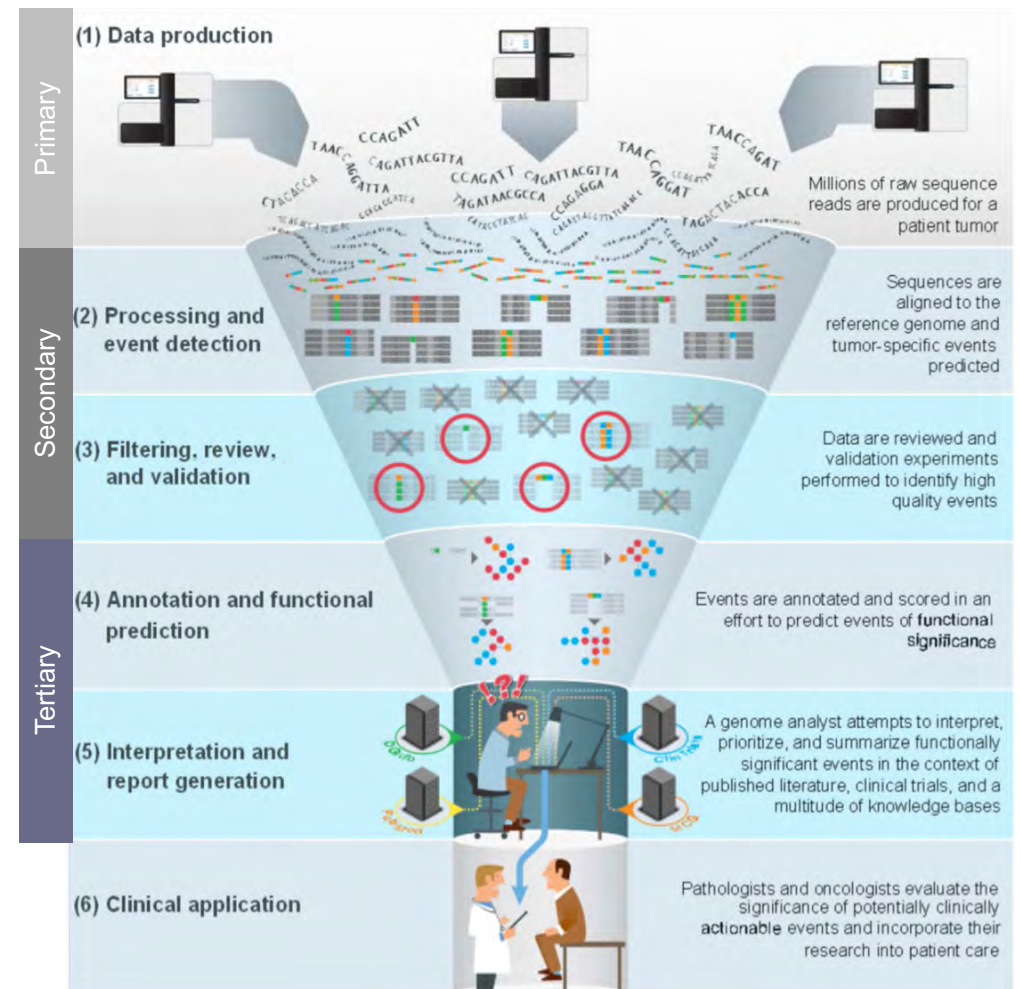
*Michael Snyder, Director Genomics &  
Personalized Medicine, Stanford*





# Variant Interpretation

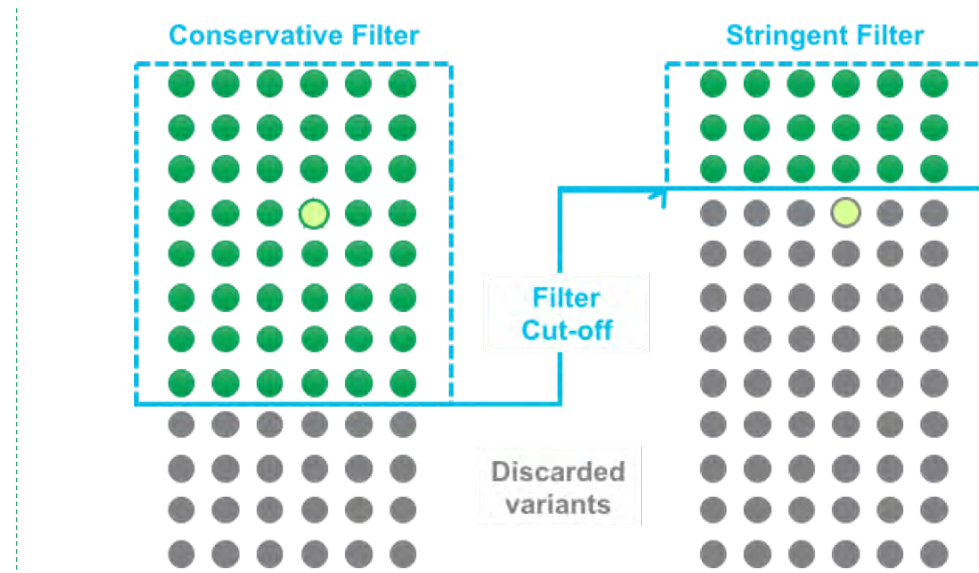
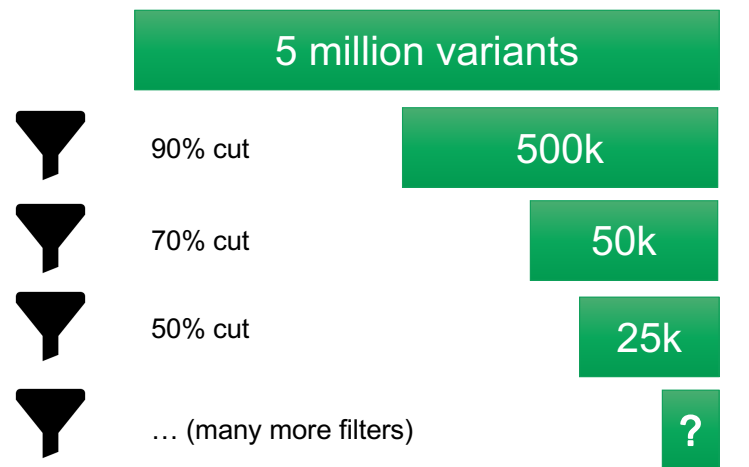
- How do we find the variants important for the disease or phenotype of the patient(s)?
- How do we find the important variants quickly, reliably, and reproducibly?
- How do we find the important variants at scale?





WGS/WES uncovers a huge amount of information  
Which findings actually matter?

## Conventional Analysis - Filtering



**100s of unranked variants to review**

**Limit Throughput**

**Drive Costs**

***Can easily spend days interpreting un-prioritized candidates for a single case***



# Alternative Variant Prioritization: Automatization without filtering caveats



## Algorithm-Driven Prioritization

5 million variants



*Actual disease-causing variant*

**Result: Specific ranked list of variants to review**  
**Diagnostic candidate consistently near top**

- Combine **phenotype, disease** and **genotype** information for **phenotype driven prioritization**
- harmonize information across hundreds of sources
- Algorithmic methods provide complementary data points



# VAAST-Phevor Algorithms for Variant Prioritization

## Peer-Reviewed and Clinically Validated

- **VAAST:** Evaluate and rank **every variant** based on **effect, conservation and frequency**
- **Phevor:** Re-prioritize variants based on patient's **phenotype**
- **Validated, Published and Clinically Proven** for **phenotype driven genome-wide variant prioritization**

*VAAST-Phevor has been shown to consistently elevate the causative variant to the top*

Resource

A preprint for peer review

Mark Yan, Jinchuan Jin, Marc V. Sills, Rebecca L. Chad D. H.

VAAST and Phevor (AAS) frame use variant conservation and frequency to prioritize variants commonly associated with disease, and present a novel approach to whole-genome variant prioritization. The past technologies for panels of polymorphisms of human disease (2008). With the advent of high-throughput sequencing, the host of clinical genetic disorders, complex families, and complex diseases. Recent technologies should be used to present significant genomic data. Broadly, the form of sequence data is not the context of a matter of predicting the function, a clinical association study. Much work is still needed in several areas.

**These authors contributed equally and significantly to the work and equally designed the study. All authors read and approved the final manuscript.**

21:1529-1542

© 2018 Flygare et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

ARTICLE

### Phevor Combines Multiple Biomedical Ontologies for Accurate Identification of Disease-Causing Alleles in Single Patients

Flygare et al. *BMC Bioinformatics* (2018) 19:37  
<https://doi.org/10.1186/s12859-018-2056-y>

BMC Bioinformatics

Open Access

CrossMark

#### SOFTWARE

## The VAAST Variant Prioritizer (VWP): ultrafast, easy to use whole genome variant prioritization tool

Steven Flygare<sup>1,2</sup>, Edgar Javier Hernandez<sup>1,2</sup>, Lon Phan<sup>3</sup>, Barry Moore<sup>1,2</sup>, Man U<sup>1</sup>, Anthony Fejes<sup>3</sup>, Hao Hu<sup>4</sup>, Karen Ellbeck<sup>5,6</sup>, Chad Huff<sup>1,2</sup>, Lynn Jorde<sup>1,2</sup>, Martin G. Reese<sup>5</sup> and Mark Yandell<sup>1,2\*</sup>

**Abstract**

**Background:** Prioritization of sequence variants for diagnosis and discovery of Mendelian diseases is challenging, especially in large collections of whole genome sequences (WGS). Fast, scalable solutions are needed for discovery research, for clinical applications, and for curation of massive public variant repositories such as dbSNP and gnomAD. In response, we have developed VWP, the VAAST Variant Prioritizer. VWP is ultrafast, scales to even the largest variant repositories and genome collections, and its outputs are designed to simplify clinical interpretation of variants of uncertain significance.

**Results:** We show that scoring the entire contents of dbSNP (>155 million variants) requires only 95 min using a machine with 4 cpus and 16 GB of RAM, and that a 60X WGS can be processed in less than 5 min. We also demonstrate that VWP can score variants anywhere in the genome, regardless of type, effect, or location. It does so by integrating sequence conservation, the type of sequence change, allele frequencies, variant burden, and zygosity. Finally, we also show that VWP scores are consistently accurate, and easily interpreted, traits not shared by many commonly used tools such as SIFT and CADD.

**Conclusions:** VWP provides rapid and scalable means to prioritize any sequence variant, anywhere in the genome, and its scores are designed to facilitate variant interpretation using ACMG and NHG guidelines. These traits make it well suited for operation on very large collections of WGS sequences.

**Keywords:** Variant prioritization, Genomics, Human genome, Variants of uncertain significance

**Background**

Variant prioritization is the process of determining which variants identified in the course of genetic testing, exome, or whole-genome sequencing are likely to damage gene function (for review [1–3]). Variant prioritization is central to discovery efforts, and prioritization scores are increasingly used for disease diagnosis as well. Both the American College of Medical Genetics and National Health Service of the United Kingdom have published guidelines for employing prioritization scores during clinical review of variants of unknown significance, or VUS [4–6].

The advent of whole genome sequencing (WGS), along with ever-growing clinical applications, has produced a host of new bioinformatics challenges for variant prioritization. Ideally, a tool should compute upon any type of variant, scale to large discovery efforts, and integrate the diverse data types that inform the prioritization process. Its scores also need to be intelligible to clinical genetics professionals. Meeting all of these requirements with a single tool is no easy matter.

Another challenge is how best to incorporate population and gene-specific variation rates into prioritization scores. The density of variation is not constant within a gene; for example, intronic variation is more frequently observed than exonic [7–9]. Moreover, the amount of potentially damaging variation varies between genes, a phenomenon referred to as “burden” [2, 10]. Zygosity is

\* Correspondence: [mflygare@slu.edu](mailto:mflygare@slu.edu)

<sup>1</sup>Department of Human Genetics, University of Utah, Salt Lake City, UT, USA

<sup>2</sup>Utah Center for Genetic Diagnosis, Salt Lake City, UT, USA

Full list of author information is available at the end of the article

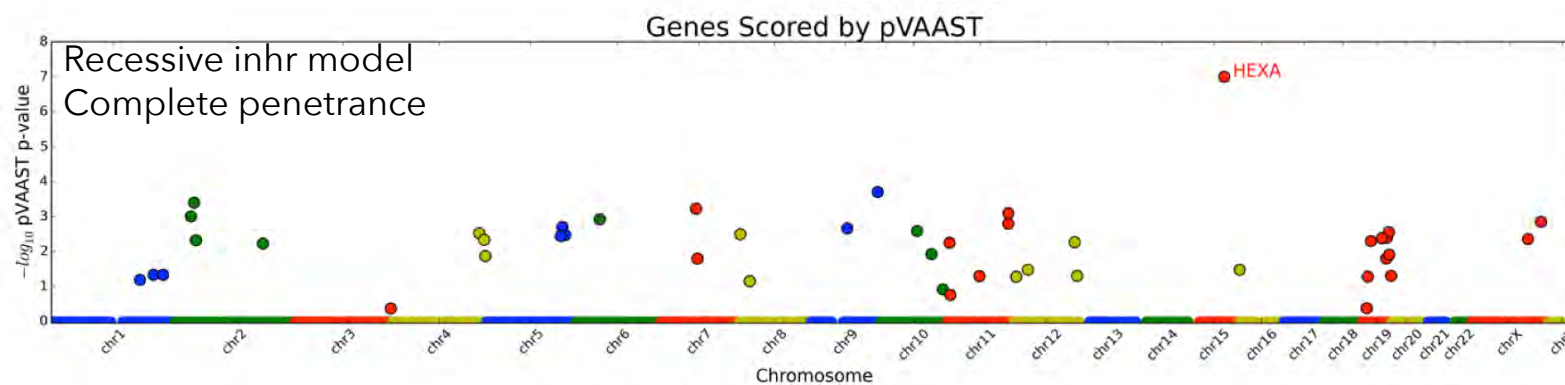
© The Author(s). 2018 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

BioMed Central





# VAAST, Families and Disease






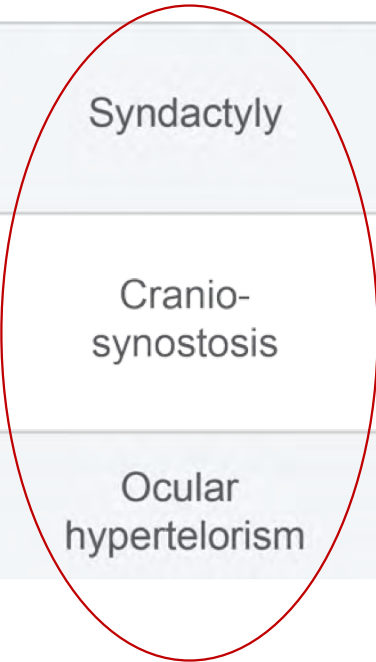
A family with a Tay-Sachs affected child





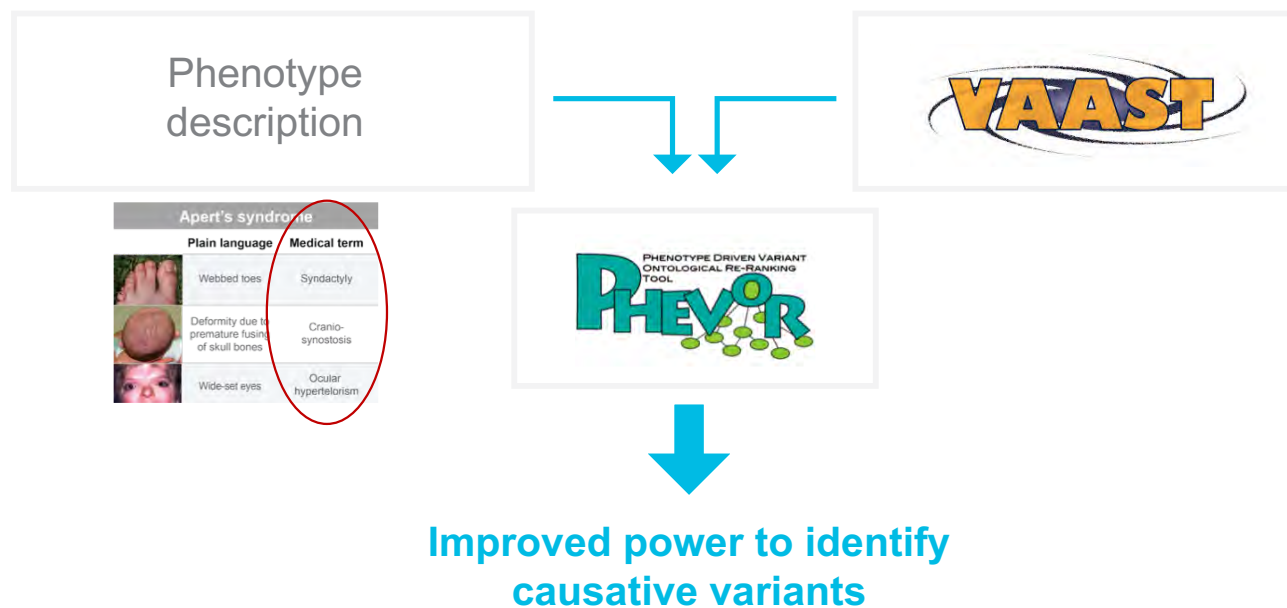
# Patient phenotype (symptoms) are big clues

	<b>Plain language</b>	<b>Medical term</b>
	Webbed toes	Syndactyly
	Deformity due to premature fusing of skull bones	Cranio-synostosis
	Wide-set eyes	Ocular hypertelorism





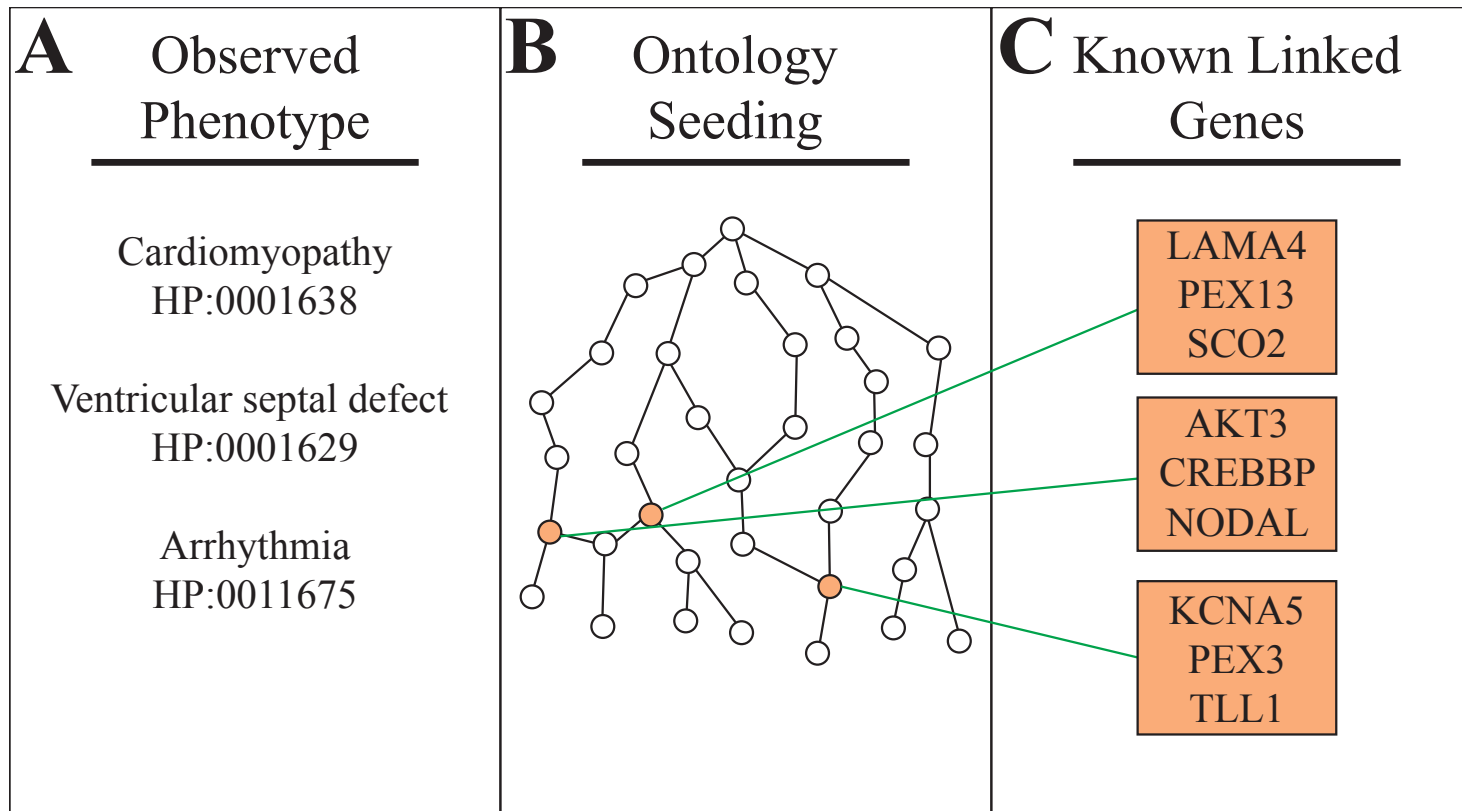
# Phevor Combines Phenotype and Genotype



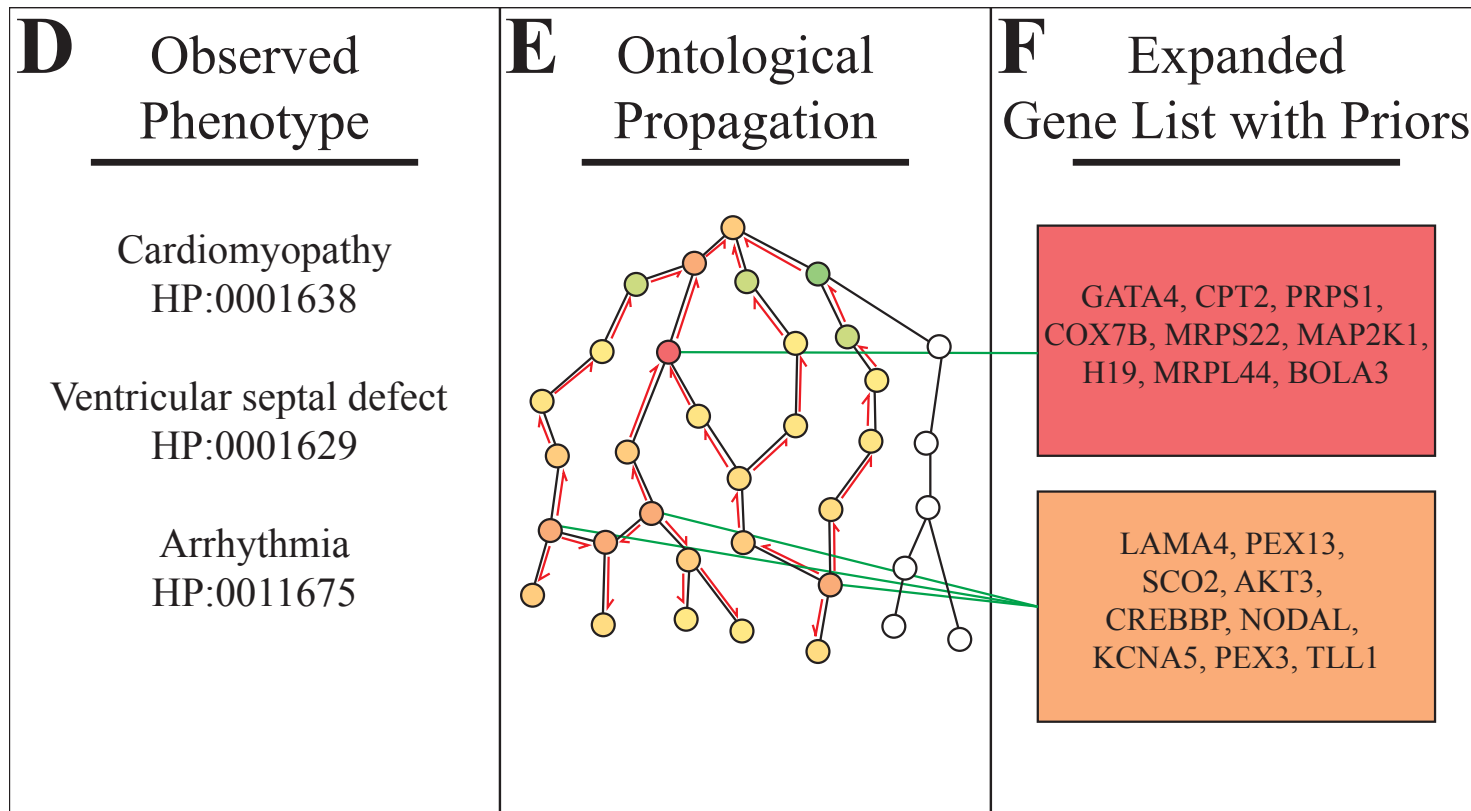
*Phevor Combines Multiple Biomedical Ontologies for Accurate Identification of Disease-Causing Alleles in Single Individuals and Small Nuclear Families.* Singleton M., Guthery SL., Voelkerding KV., et al., Reese MG., Jorde LB., Huff CD., Yandell M. ( 2014) . Am J Hum Genet. 2014 Apr 3; 94(4):599-610.



# How Phevor Works (seeding)

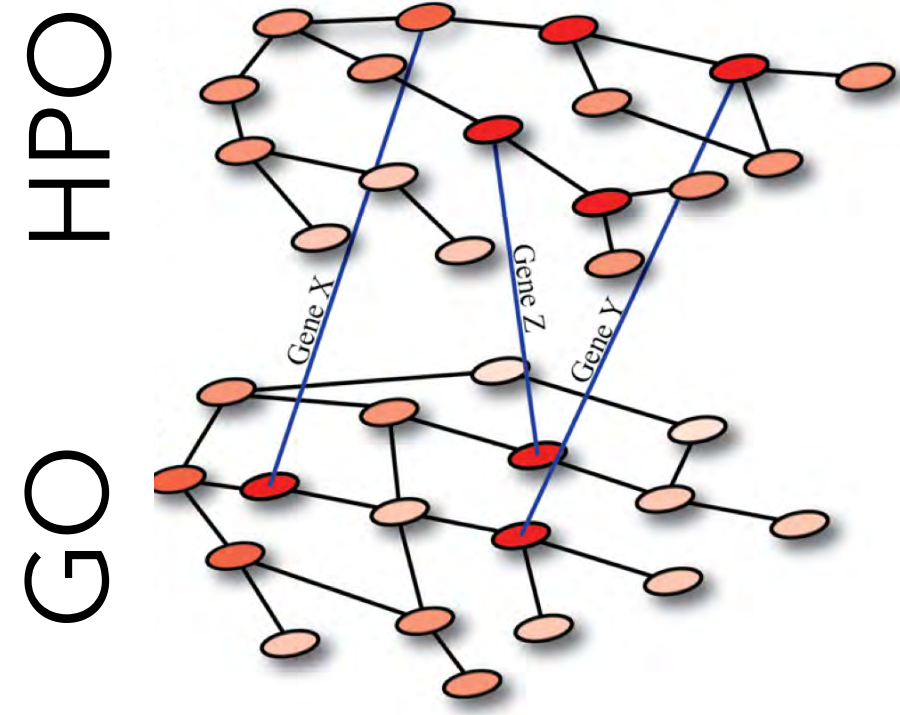


# How Phevor Works (seeding)

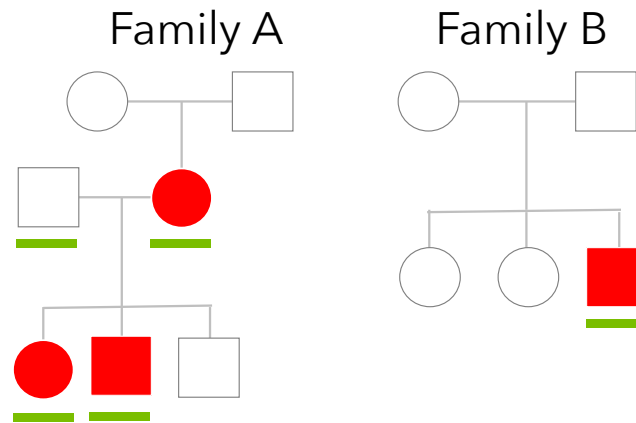




# How Phevor Works (combining ontologies)






# Common Variable Immune disease (CVID)



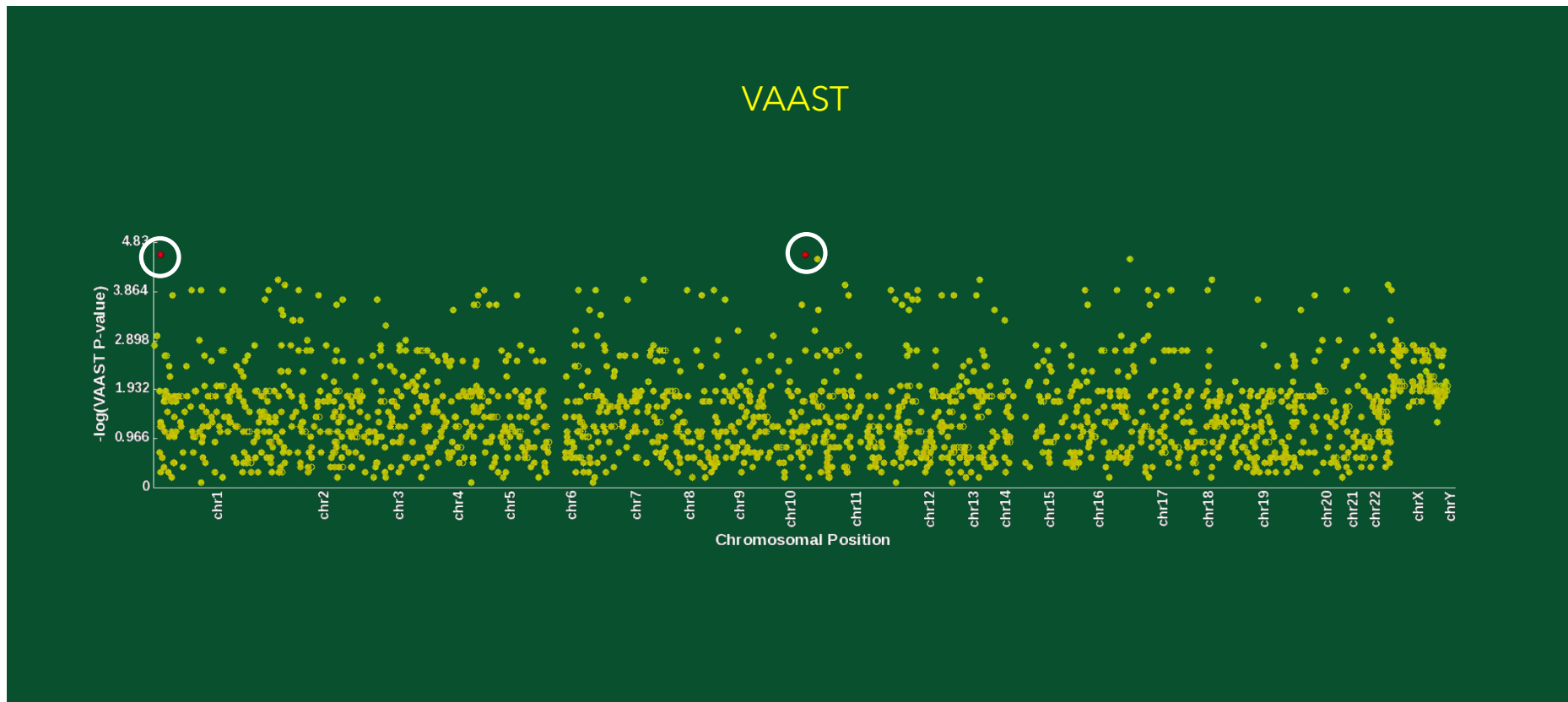
## Phenotype

Recurrent infections (HPO:0002719)

Abnormality of Humoral immunity (HPO:0005368)

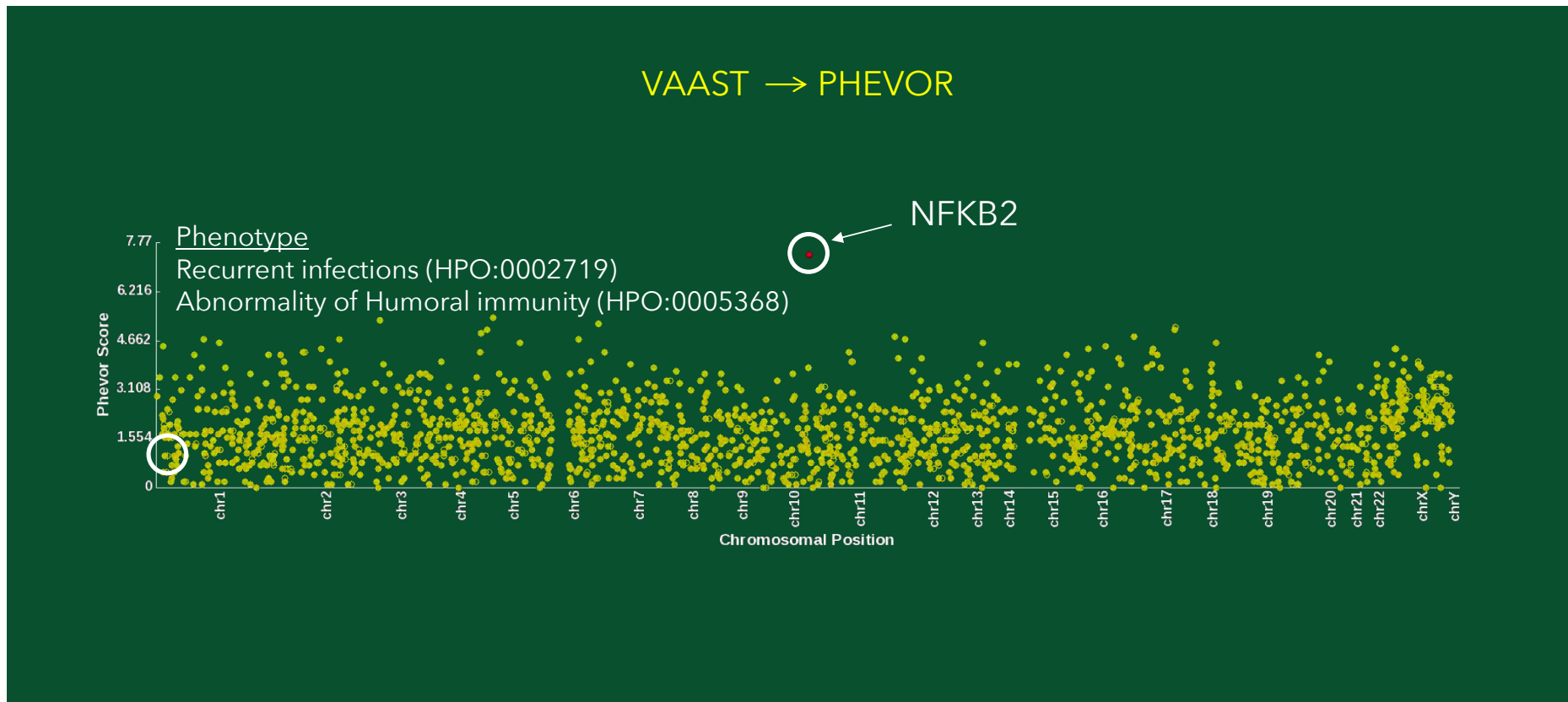
-  affected
-  unaffected
-  exome sequenced

# A Family with Common Variable Immune disease (CVID)



K. Chen MD, EM. Coonrod PhD, A. Kumanovics, K. Voelkerding MD, et al.

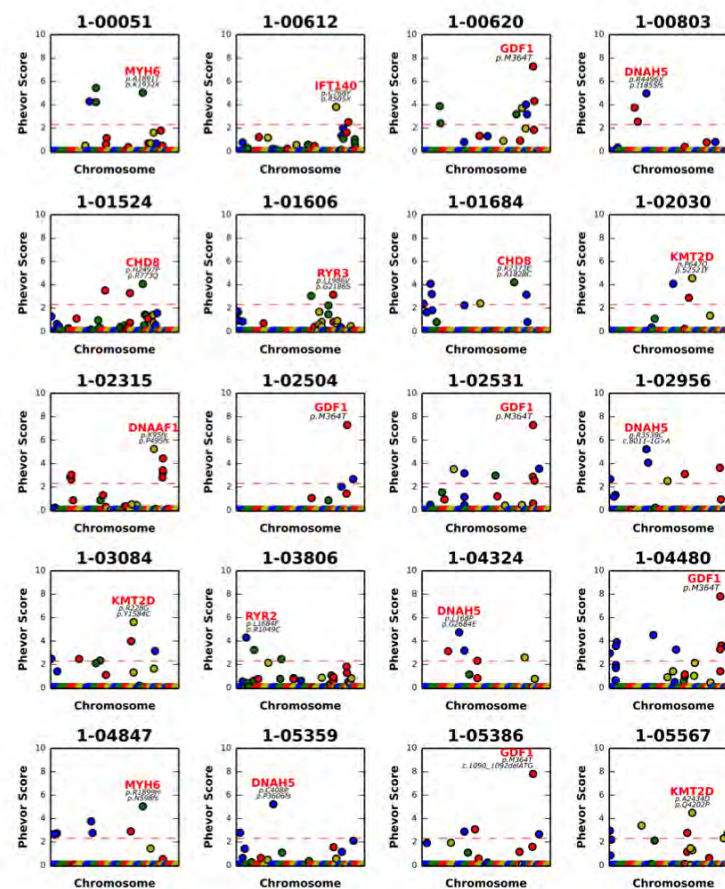
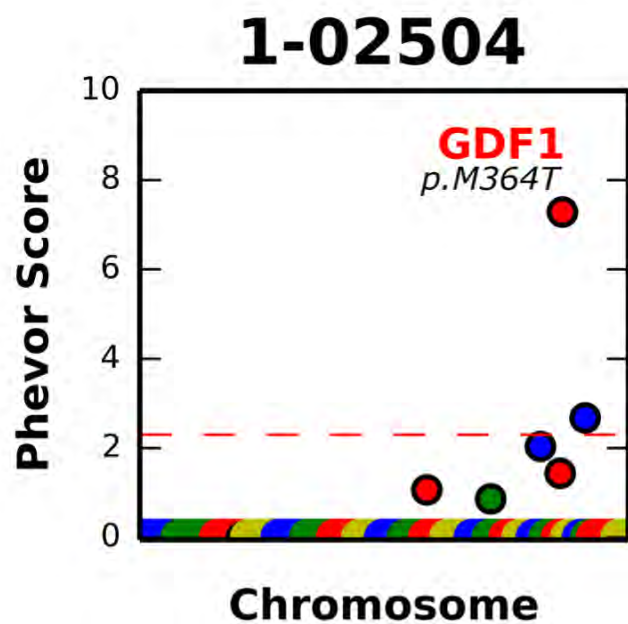
# A Family with Common Variable Immune disease (CVID)



K. Chen MD, EM. Coonrod PhD, A. Kumanovics, K. Voelkerding MD, et al.



Selected *Recessive* Congenital Heart disease genes identified with VAAST & PHEVOR in a cohort of 2,871 congenital heart disease probands \*



\*Jin SC., Homsy J. et al. Contribution of rare transmitted and de novo variants among 2,871 congenital heart disease probands. Submitted.



# Country Sequencing Program

## *Genomics England & 100,000 Genomes Project*

---



### **Goal:**

Sequence 100,000 whole genomes to discover and diagnose the genetic basis for rare disease and cancer



## The UK's 100,000 Genomes Project

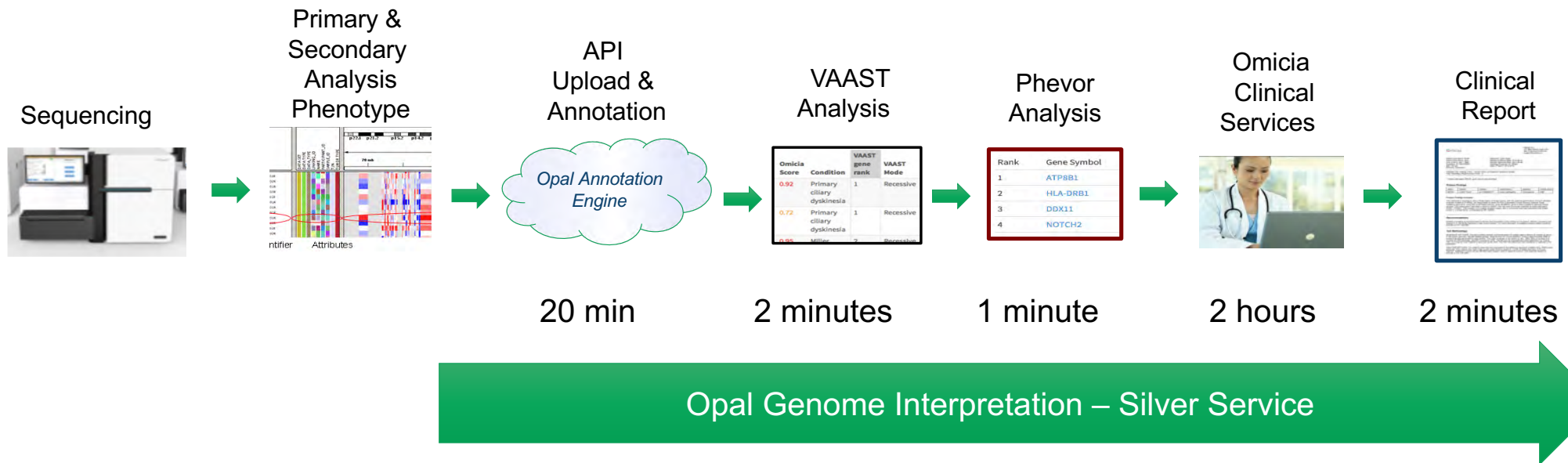
is the largest initiative of its kind in the world. It is ground-breaking ....and is establishing protocols and standards that will be applied across the entire healthcare system.

---

SIR JOHN CHISOLM,  
EXECUTIVE CHAIR  
OF GENOMICS ENGLAND



# Fabric Delivers Automated, Intelligent Interpretation Workflows



## Performance highlights:

**1st**  
Company to deliver  
clinical reports  
(April, 2016)

**1,900+**  
Cases returned as  
of February 2018

**40.8%**

**Disease causing  
candidate yield<sup>1</sup>**  
(industry average: 29%)<sup>2</sup>

**Initial WGS  
interpretation  
< 2 hours**

**53.3%**

**Candidates identified  
by Phevor-VAAST  
as Top hit<sup>1</sup>**

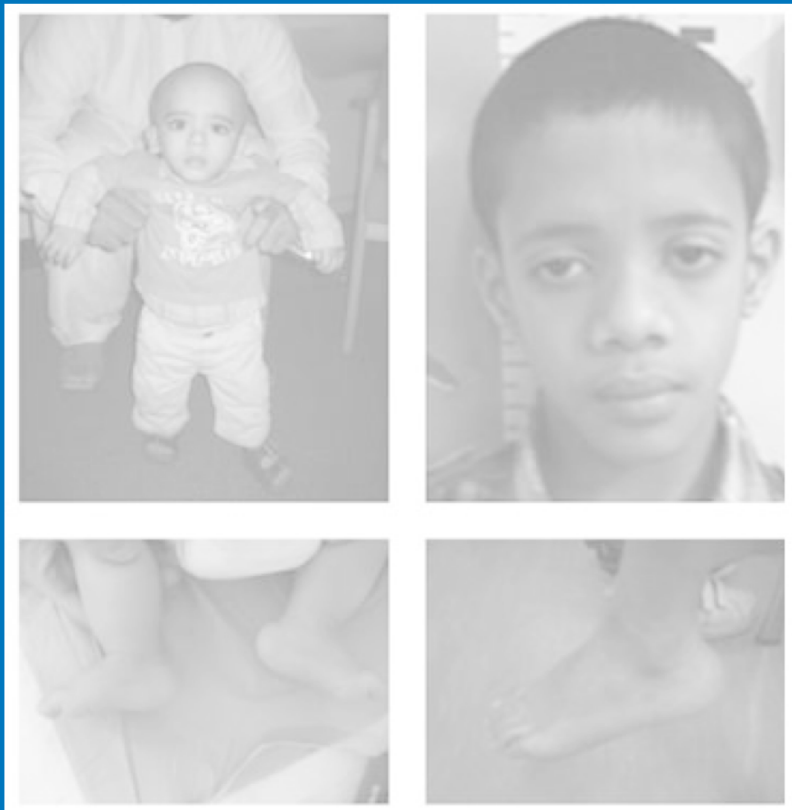


**Delivering reports  
to multiple  
main program  
sites**

<sup>1</sup>Identified disease-causing candidates in first 609 cases. Babcock et al, *Increased Yield of Clinically Relevant Candidates in the UK 100,000 Genomes Project Using Opal™ Clinical for Hereditary Disease*. Poster session presented at: ACMG; 2017 Mar 21-25; Phoenix, AZ.

<sup>2</sup>"The success rate of 29 percent, which is about twofold higher of conventional genetic evaluations for such patients, makes WES a reasonable diagnostic approach for patients on a diagnostic odyssey" says Dr. Lazaridis, Mayo Clinic, Individualized Medicine Clinic.

# Recent case – solved by Fabric Genomics



## Patient diagnosed with:

- Intellectual Disability
- Multiple Epiphyseal Dysplasia

## Findings:

- Misdiagnosis
- Fabric Genomics discovered a change in OBSL1 gene, associated with Three M syndrome\*

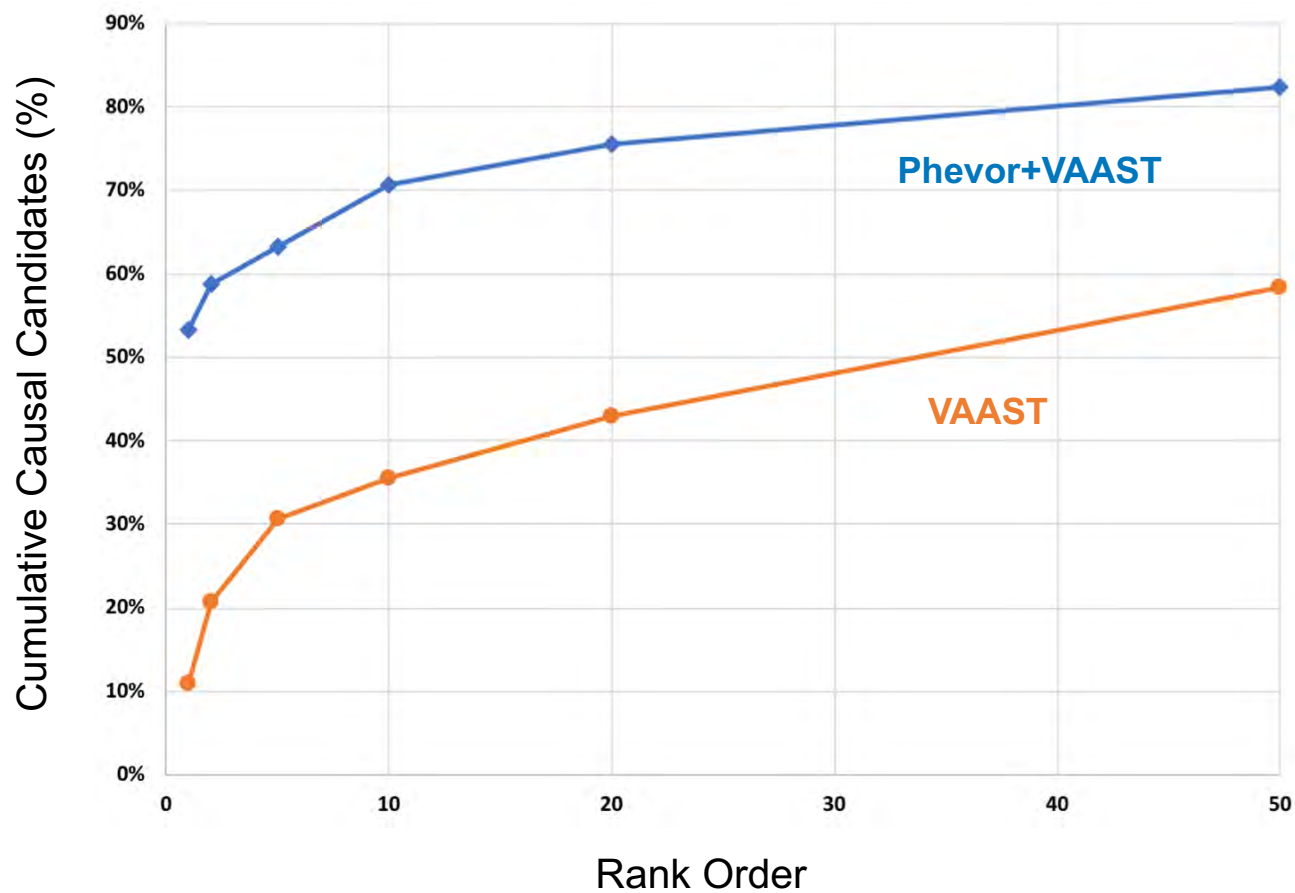
## Result:

- Patient was accurately diagnosed with Three M syndrome

\*Three M syndrome is an extremely rare genetic disorder characterized by low birth weight, short stature (dwarfism), characteristic abnormalities of the head and facial (craniofacial) area, distinctive skeletal malformations, and/or other physical abnormalities.



# Phevor significantly improves ranking of variants



Genome England cases

- Causative candidate variants n=644
- 450 Proband with positive findings





# Rady Children's Hospital Launches Rapid Genome Service

- **Goal:** Set up rapid genome testing for newborns and infants in Intensive Care Unit
- **Need:** 1 hour turnaround time for interpretation
- **With Fabric Genomics' solution:** Delivering 1 hour interpretation made it possible to now routinely turnaround a sample from collection, sequencing to clinical report in < 36 hours



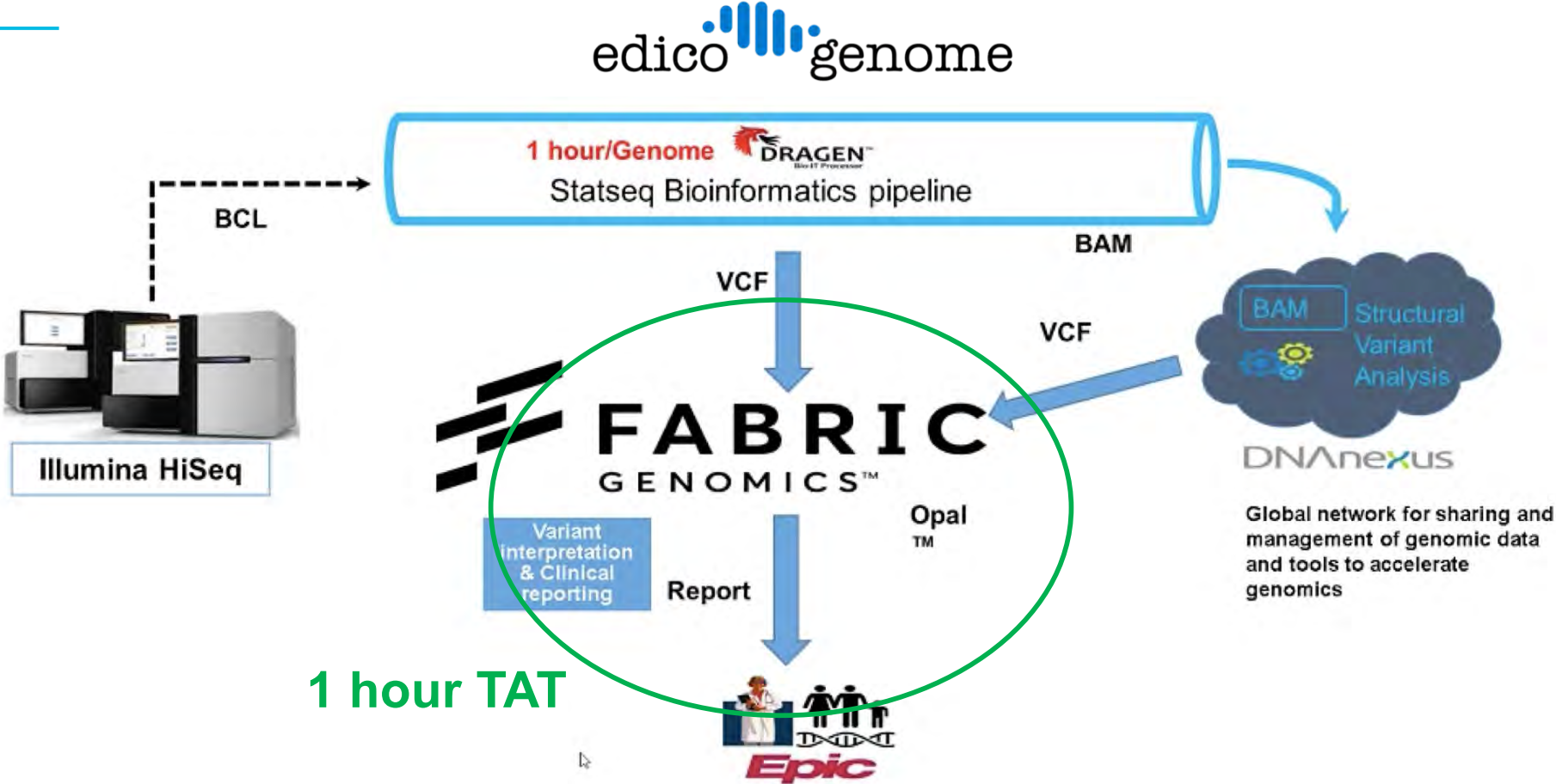
“ Diagnosing acutely ill babies is a race against the clock, which is why it's so essential for physicians to have access to technology that will provide answers faster and help set the course of treatment. ”



Stephen Kingsmore, M.D.,  
D.Sc., president and CEO  
of Rady Children's Institute  
for Genomic Medicine



# Rady's Rapid Whole Genome Interpretation Pipeline



# Case Study: Fabric Genomics / Rady Children's Hospital STAT Pediatric Clinical Genome – 19.5 hours from sample to report



## New GUINNESS WORLD RECORDS™ Title Set for Fastest genetic diagnosis

Scientists at the [Rady Children's Institute for Genomic Medicine \(RCIGM\)](#) have compressed the time needed to decode rare genetic disorders in newborns through DNA sequencing to less than a day. RCIGM set a new GUINNESS WORLD RECORDS™ title for Fastest genetic diagnosis in 19.5 hours. Dr. Kingsmore has pioneered the rapid turnaround and delivery of genetic test results to neonatal and pediatric intensive care (NICU/PICU) physicians.

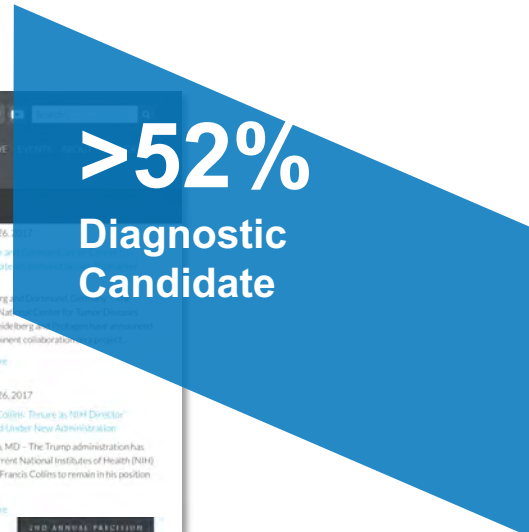
Fabric Genomics' platform was a critical technology enabler. Fabric's clinical decision support software enables rapid diagnosis by helping to pinpoint the source of genetic disease out of hundreds of possibilities. Fabric's platform includes the Fabric STAT pipeline that provides guaranteed rapid turnaround time for urgent pediatric genomic tests.

Genetic diseases are the leading cause of death in infants in North America, affecting an estimated four percent of newborns. Rare genetic diseases also account for approximately 15 percent of admissions to children's hospitals.

"Our evolving ability to find the answers to medical mysteries through rapid Whole Genome Sequencing is providing hope for babies and children with rare, genetic diseases," said Dr. Kingsmore. "By speeding delivery of genomic insights, we are equipping physicians with the information they need to provide precision care for the youngest and most fragile patients."

Our mission is honored to support this great milestone in Stephen Kingsmore said

Date: February 12, 2018





## Find the right answer faster using VAAST/Phevor

	Candidate Yield	Avg Rank Causative Variant	% Cases Where Causative Variant Ranked in	
			Top 1	Top 20
Rady Institute for Genomic Medicine	52%	4.5	<b>60%</b>	93%
Genomics England	45%	12.5	<b>45%</b>	74%
Labcorp	54%	12	<b>52%</b>	89%

***Immediately identify causative variant in ~50% of cases***



# Financing Modeling

★ Matched control

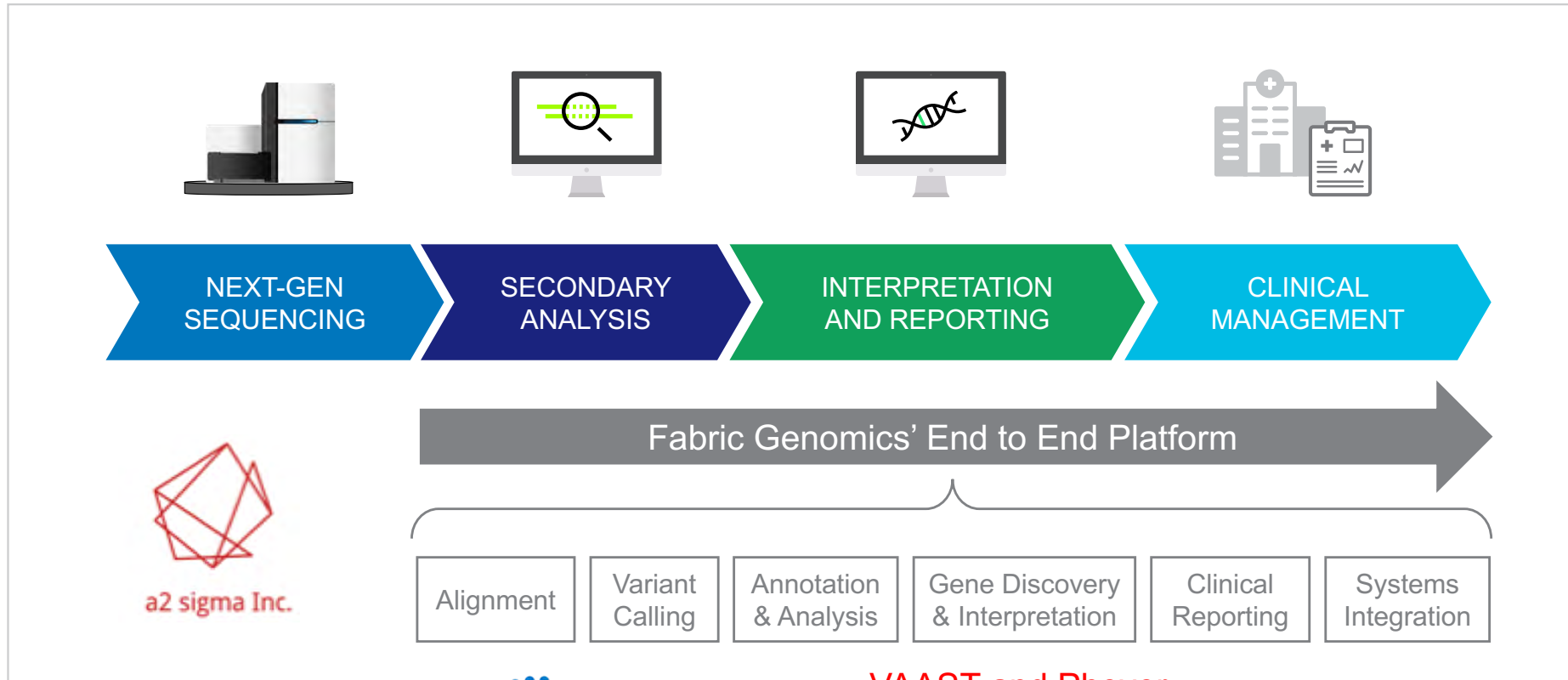


Subject ID	Site	Presentation	Gene	Admission to Diagnosis, days/(Test Method)	Type of cost	Decrease in		Medicaid		Commercial			
						Hospital stay (days)	Hospital Stay, days (%)	WRVU	WRVU Avoided	Payment, \$1000	Payment Avoided, \$1000 (%)	Payment, \$1000	Payment Avoided, \$1000 (%)
★ 6041 Ctrl 1	NICU	Seizures	KCNQ2	4 (rWGS)	Case: Rapid diagnosis and precision treatment	18		120		85		157	
				42 (std)	Model: Case + 6 week time to diagnosis	59	41 (69%)	280	156 (57%)	234	149 (64%)	429	272 (63%)
					Control: With standard time to diagnosis			326	205 (63%)	210	125 (60%)	383	226 (59%)
★ 6014 Ctrl 2	NICU	Hypotonia	NEB1	7 (rWGS)	Case: no muscle biopsy	n.a.		n.a.					
				n.a.	Control: whole hospitalization	35	2 (6%)	246	24 (8%)	165	32 (20%)	302	59 (20%)
					Control: muscle biopsy, pathology, postop. care x 1 day	2		24		33		59	
★ 6026 Ctrl 3	PICU	Cholestasis & CHD	JAG1	3 (rWGS)	Case: Kasai cancelled	11		n.d.		40		72	
				n.a.	Control: Kasai surgery whole hospitalization			104	90 (87%)	80	54 (68%)	147	100 (68%)
					Control: Kasai surgery & postop. care only	14	3 (21%)	90		54		100	
6053	NICU	Hypoglycemia	ABCC8	7 (rWGS)	Case: rapid diagnosis and transport for surgery	10		105	175 (62%)	60	108 (64%)	109	197 (64%)
				28 (std)	Model: average published time to diagnosis	31	21 (68%)	280		168		306	
6012	NICU	Complex	ARID1B	26 (rWGS)	Case: Palliative care started 250 days after admission	250		1,714		1,941		3,538	
					Model: Case + 6 weeks antibiotics	292	262 (90%)	2,001	1,624 (81%)	2,587	2,362 (91%)	4,707	4,299 (91%)
					Model: Palliative care started 30 days after admitted	30		377		224		408	
6011	GI	Cholestasis	NPC1	7 (rWGS)	Case: First hospitalization (rWGS unavailable)	8		26	33 (56%)	26	27 (51%)	47	50 (51%)
					Case: Second hospitalization (rWGS available)	15	15 (35%)	33		27		50	
Average						76	57 (76%)	499	354 (71%)	546	477 (87%)	994	826 (83%)



 micia is now

 **FABRIC**  
GENOMICS™

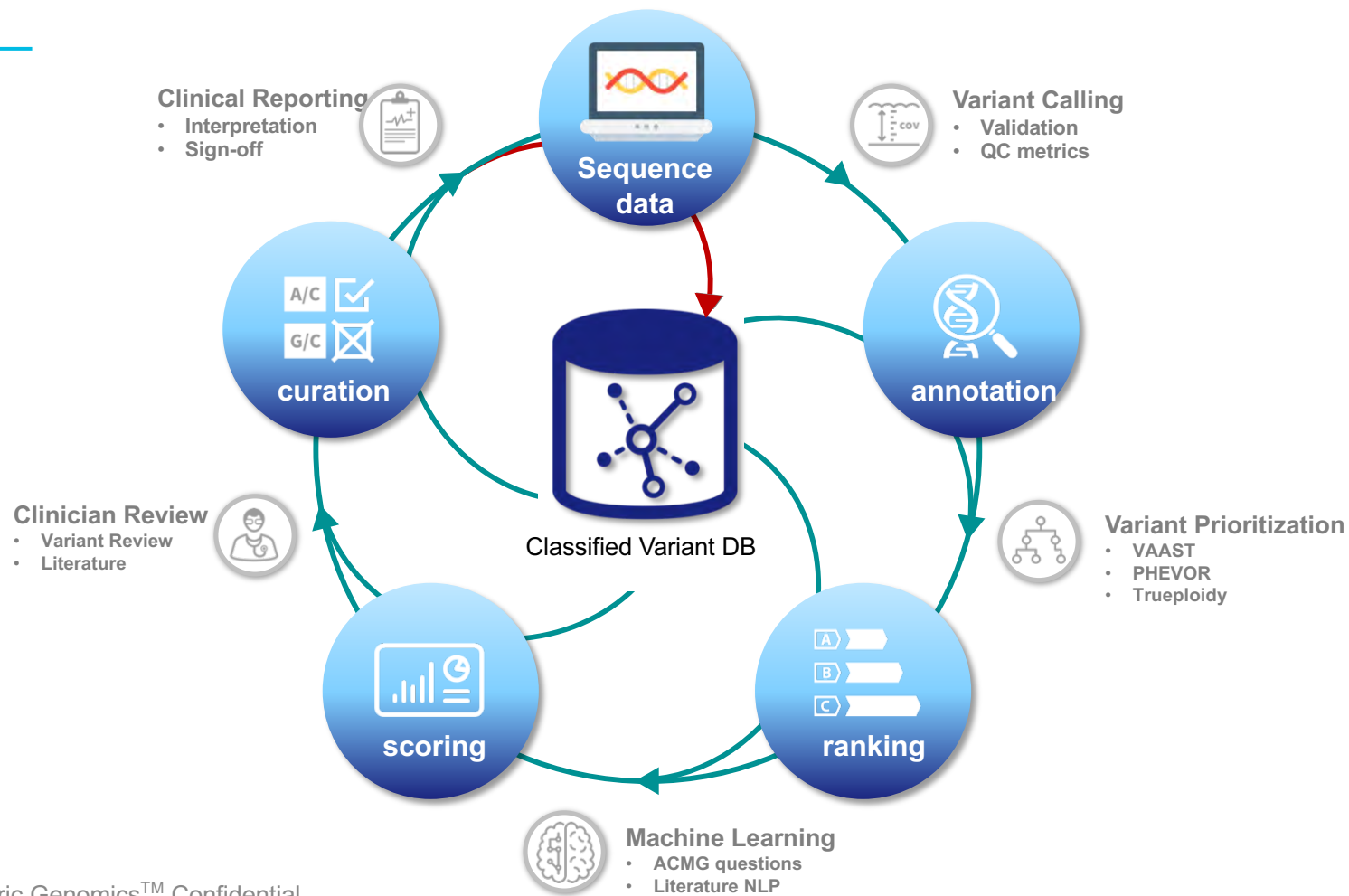


Fabric Genomics™

 edico genome

VAAST and Phevor

# Classified Variant Database to Provide Largest, Highest Quality Database of Human Genetic Variation via Network Effect



- Variants classification are shared within organization
- Option to share with other labs
- Share at different levels to protect patient privacy
- Leverage massive data and sharing to enable more accurate diagnoses

# Opal™ Clinical – Tertiary Analysis Decision Support System for Hereditary Disease



- 90+ public and proprietary databases and analysis algorithms per variant
- All testing types supported: solo, trio, and flexible family
- Proven health systems integrations with electronic medical records (Epic and Cerner) and LIS

Opal 4.13.1 Home Projects Admin Case Studies - csonrigby@

Case Studies | Clinical Reports | Interpret Variants

Clinical Report ID: 6521 Test: Non-syndromic Epilepsies Panel  
Genome ID: 214317 Filter: Consequence, No introns  
Genome: EESolo\_Single&Panel\_STXBP1frameshift\_NA12882\_IL1\_E1\_R0H COSMIC included: False  
Assay Type: TruSight One

Interpret Variants

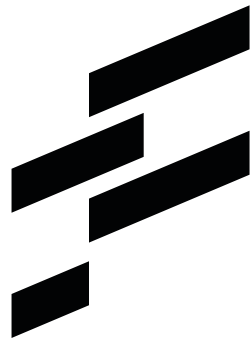
Report Filters	Gene	Position	Change	IGV	Effect	Zygoty	Quality GQ Coverage	1KG AF EVS AF ExAC AF	Omicia Score
Reporting Status: All	GPR98	chr5 89988504 rs2366826	A → G c.7034A>G p.Asn2345Ser		missense	●○	27436 99 65 : 33 : 32	0.32228 0.29139 0.37290	0.945
Gene Symbol: CPAG	CPAG	chr8 68334782 rs72654981	G → A c.1271C>T p.Ala424Val		missense	●○	3809 99 50 : 26 : 24	0.00020 0.00046 0.00032	0.945
Filter By: Require	GPR98	chr5 89979698 rs4916685	C → T c.5960C>T p.Pro1987Leu		missense	●○	50105 99 68 : 36 : 32	0.33726 0.30535 0.35417	0.808
Exclude	STXBP1	chr9 130428534	CTA → C c.754_755delAAT p.Met252Glu		frameshift	●○	34872 98 86 : 37 : 49		0.8
Zygoty: any	GPR98	chr5 89985882 rs10037067	A → G c.6695A>G p.Tyr2232Cys		missense	●○	17344 99 55 : 32 : 23	0.34984 0.31810 0.38742	0.751
Exclude: No-call, Fail VCF Filter	GPR98	chr5 90151630 rs56382582	C → G c.17667C>G p.His5889Gln		missense	●○	3206 99 49 : 31 : 18	0.00020 0.00008 0.00002	0.718
	KCNT1	chr9 136683984 rs74533482	A → G c.3685A>G p.Thr1229Ala		missense	●○	1469 99 40 : 17 : 23	0.00040 0.00046 0.00052	0.645

# Thank you!



**Martin Reese, Ph.D.**  
martin@fabricgenomics.com





# FABRIC

GENOMICS™



A global healthcare platform for genomic data analysis

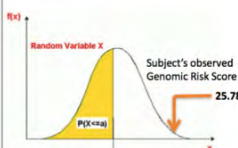




# Comprehensive Computational Genomics Platform

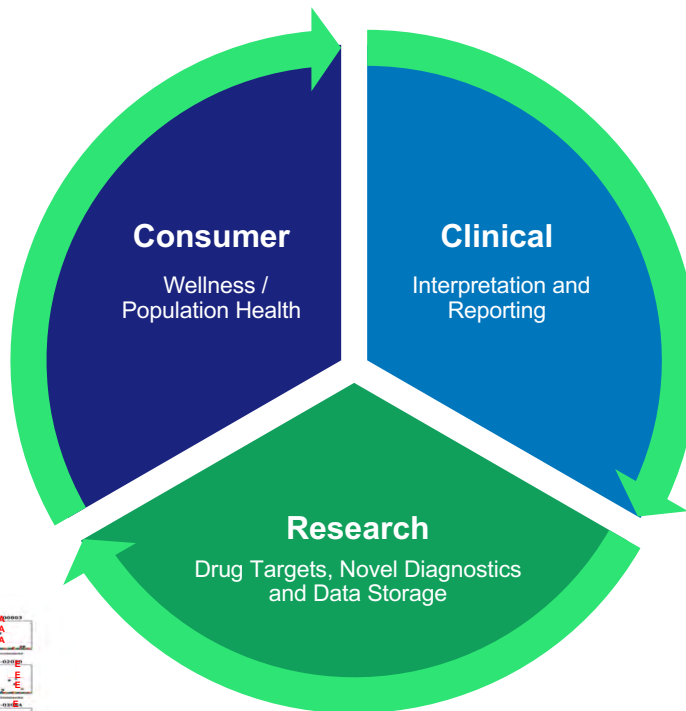
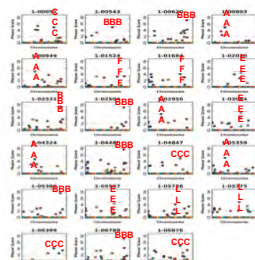
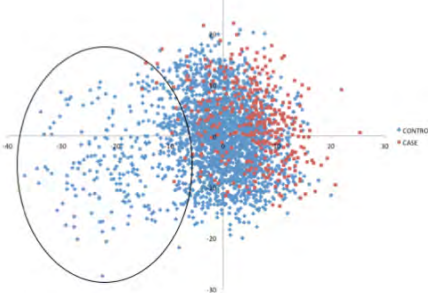
## Grouper

Genetic disease burden scores for wellness applications and population health management



## VAAST3 for Target Discovery

Case / Control studies



## Opal™ Clinical (VAAST + Phevor)

End-to-end clinical software that turns raw genome sequence data and phenotype into high-quality clinically meaningful insights



## BioGraph™

Advanced graph technology for structural variant detection, population-specific, reference-free diagnostics, genome data storage

