Genomic Crowdsourcing:

Allele Frequency Community Provides Expansive, Ethnically Diverse, Freely Available Community Resource for Allele Frequency Annotation

D Richards[1], D Bassett[1], Carlos Bustamante[2], D Cooper[3], G Eley[4], R Felciano[1], R Forsberg[1], L Furmanski[1], G Glusman[5], D Goldstein[6], M Hegde[7], P Hieter , A Joecker[1], T Kaminski[9], A Krämer[1], S Letovsky[9], T Love[1], B Macy[1], CE Mason[10], A Muthiah[1], N Nielsen[1], K Patel[1], N Pearson[11], H Rehm[12], H Rienhoff[13], F Schacherer[1], E Schadt[14], S Scott[1], S Shah[1], J Shendure[15], D Shiffman[1], B Solomon[3], A Subramanian[1], P van der Spek[16], JG Vockley[2], R Yip[1], X Zhu[1] , and hundreds of Allele Frequency Community members worldwide.

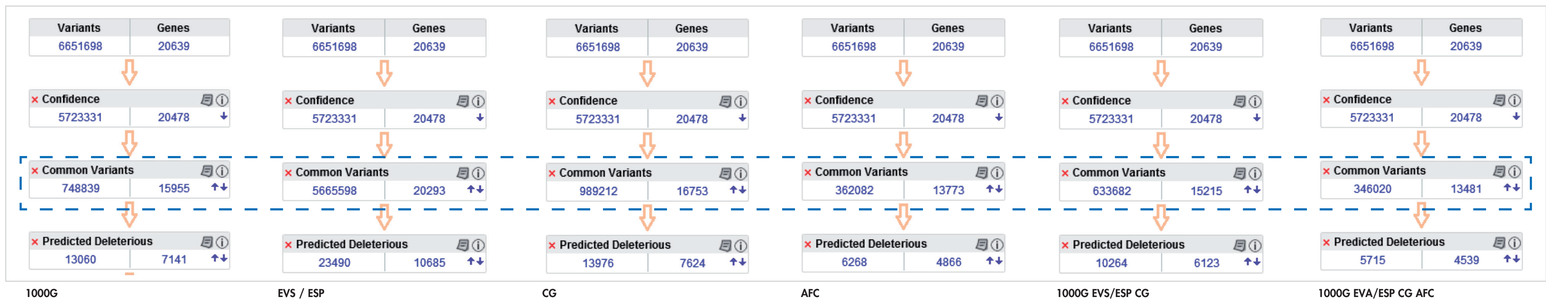
QIAGEN presents: Allele Frequency Community

Insights through Diversity

- Free to join - **www.allelefrequencycommunity.org**
- Over over 103,000 samples, including over 11,000 whole genomes
- Improves over time as it is used!
- Available today in QIAGEN's Ingenuity Variant Analysis - **www.ingenuity.com/variants**
- Benefits from foundational community resources, including 1000 Genomes Project, the Exome Aggregation Consortium (ExAC), Exome Variant Server (EVS), CGI public exomes, and Kaviar

How Well Does it Work?

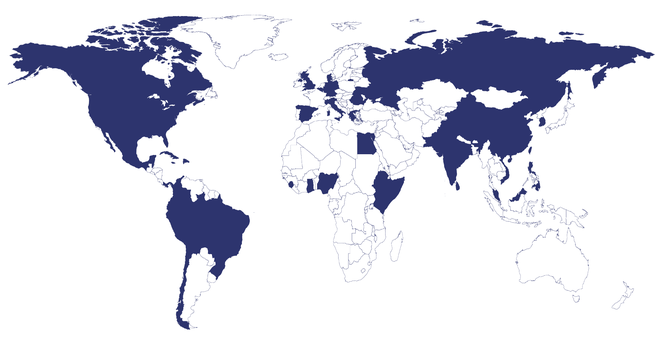
Whole Genome Data – 1% maximum allele frequency

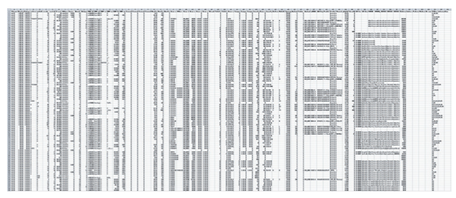


Ingenuity Variant Analysis

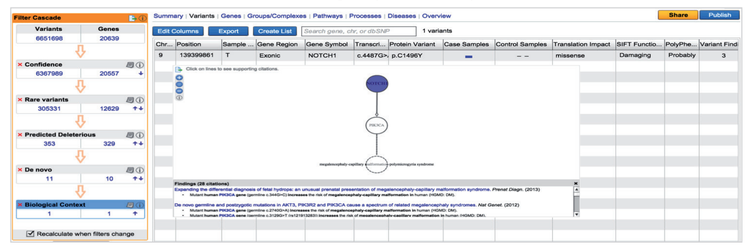
Identify causal variants from human sequencing data in just minutes.

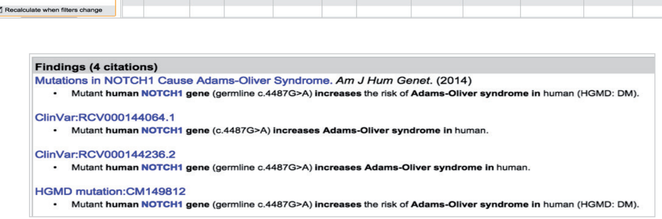
- **Intuitive, user-friendly interface** - No bioinformatics skills needed.
  - **Interactive Filter Cascade** - Rapidly eliminate common and non-deleterious variants with a basic set of filters.
  - **Iterative analysis** - Apply a hypothesis, visualize and evaluate results in real-time.
- **Knowledge driven algorithms and analytics** - Genetics, Statistics, Functional Prediction.
  - **Curated up to date content at your fingertips** - Utilize The Ingenuity Knowledge Base™ containing millions of biomedical findings and mutations, accurate, up-to-date and curated by experts, from the literature and public databases.
  - **Causal Network Analytics** - Identify variant in genes within 1- or 2- network "hops" of upstream or downstream mutated genes.
- **Sharing and Publication Tools** - Easily collaborate with colleagues and peers. Export data and graphics to aid in manuscript preparation and publications.
- **Scalable processing capacity** - Analyze thousands of samples simultaneously.



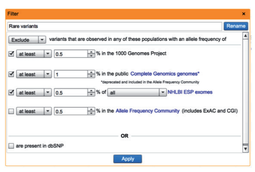


QIAGEN Ingenuity Variant Analysis





Analysis without Allele Frequency Community filtering



Chr...	Position	Sample Allele	Gene Region	Gene Symbol	Transcript Variant	Protein Variant	Case Samples	Control Samples	Translation Impact	SIFT Functi...	PolyPhen-2 Funct...	Classification
2	131350534	T	Exonic	CFC1/CFC1E	c.363C>A, c.473C>	p.P121P, p.P1		–	missense, synonyn			Uncertain Signific
8	61757808	A	Splice Site	CHD7	c.5051-1G>A			– –				Pathogenic
9	2039777		Exonic	SMARCA2	c.703_705delCAG	p.Q238del		– –	in-frame			Likely Benign
9	107556793	AAA	Splice Site	ABCA1	c.5383-3_5383-2ins			– –				Likely Pathogenic

Demo case

- Male
- Tetralogy of Fallot
- Choanal atresia
- Hypoplasia of the semicircular canal
- Mode of inheritance - likely de novo

Analysis with Allele Frequency Community filtering

Chr...	Position	Sample Allele	Gene Region	Gene Symbol	Transcript Variant	Protein Variant	Case Samples	Control Samples	Translation Impact	SIFT Functi...	PolyPhen-2 Funct...	Classification
8	61757808	A	Splice Site	CHD7	c.5051-1G>A			– –				Pathogenic

> 75 % reduction

Allele Frequency Community Founders

- David Goldstein, Columbia University Institute for Genomic Medicine
- Madhuri Hegde, Emory Genetics Laboratory
- Tom Kaminski & Stan Letovsky, Enlighten Health Genomics, a business of Laboratory Corporation of America® Holdings (LabCorp®)
- Peter van der Spek, Erasmus University Medical Center
- Eric Schadt, Icahn Institute for Genomics and Multiscale Biology at Mount Sinai
- Hugh Rienhoff, Imago Biosciences
- Joe Vockley & Greg Eley, Inova Translational Medicine Institute
- Gustavo Glusman, The Institute for Systems Biology
- Nathan Pearson, New York Genome Center
- Heidi Rehm, Partners Healthcare
- Carlos Bustamante, Stanford University
- Phil Hieter, University of British Columbia
- Jay Shendure, University of Washington
- Chris Mason, Weill Cornell Medical College
- QIAGEN

Co-author affiliantions

1. QIAGEN Bioinformatics, 1700 Seaport Blvd, Third Floor, Redwood City, CA, 94063, USA
2. Stanford School of Medicine, 291 Campus Drive, Stanford, CA 94305-5101, USA
3. Cardiff University School of Medicine, Cardiff, CF14 4XN
4. Inova Translational Medicine Institute, 8110 Gatehouse Road, Falls Church, VA, 22042, USA
5. Institute for Systems Biology, 401 Terry Avenue North, Seattle, WA, 98109, USA
6. Columbia University Institute for Genomic Medicine, 630 West 168th St, New York, NY 10032, USA
7. Emory Genetics Laboratory, 2165 N Decatur Rd, Decatur, GA, 30033, USA
8. University of British Columbia, 2329 W Mall, Vancouver, BC, V6T 1Z4, Canada
9. Enlighten Health Genomics, a business of Laboratory Corporation of America® Holdings (LabCorp®), T W Alexander Drive, Research Triangle Park, North Carolina 27709, USA
10. Weill Cornell Medical College, 1305 York Ave, New York, NY 10021, USA
11. New York Genome Center, 101 Avenue of the Americas, New York, NY 10013, USA
12. Harvard Medical School and Brigham & Women's Hospital, 65 Landsdowne Street, Cambridge, MA, 02139, USA
13. Imago Biosciences, San Francisco, California, USA
14. Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, New York, NY, 10029, USA
15. University of Washington, 3720 15th Ave NE, Seattle, WA, 98195, USA
16. Erasmus University Medical Center, 3000 CA, Rotterdam, The Netherlands

Sample to Insight