GeneGrid

finding disease causing variants in NGS data

Claudia Gugenmus

Genomatix Software GmbH
Bayerstrasse 85a. 80335 Munich, Germany
http://www.genomatix.com

Motivation

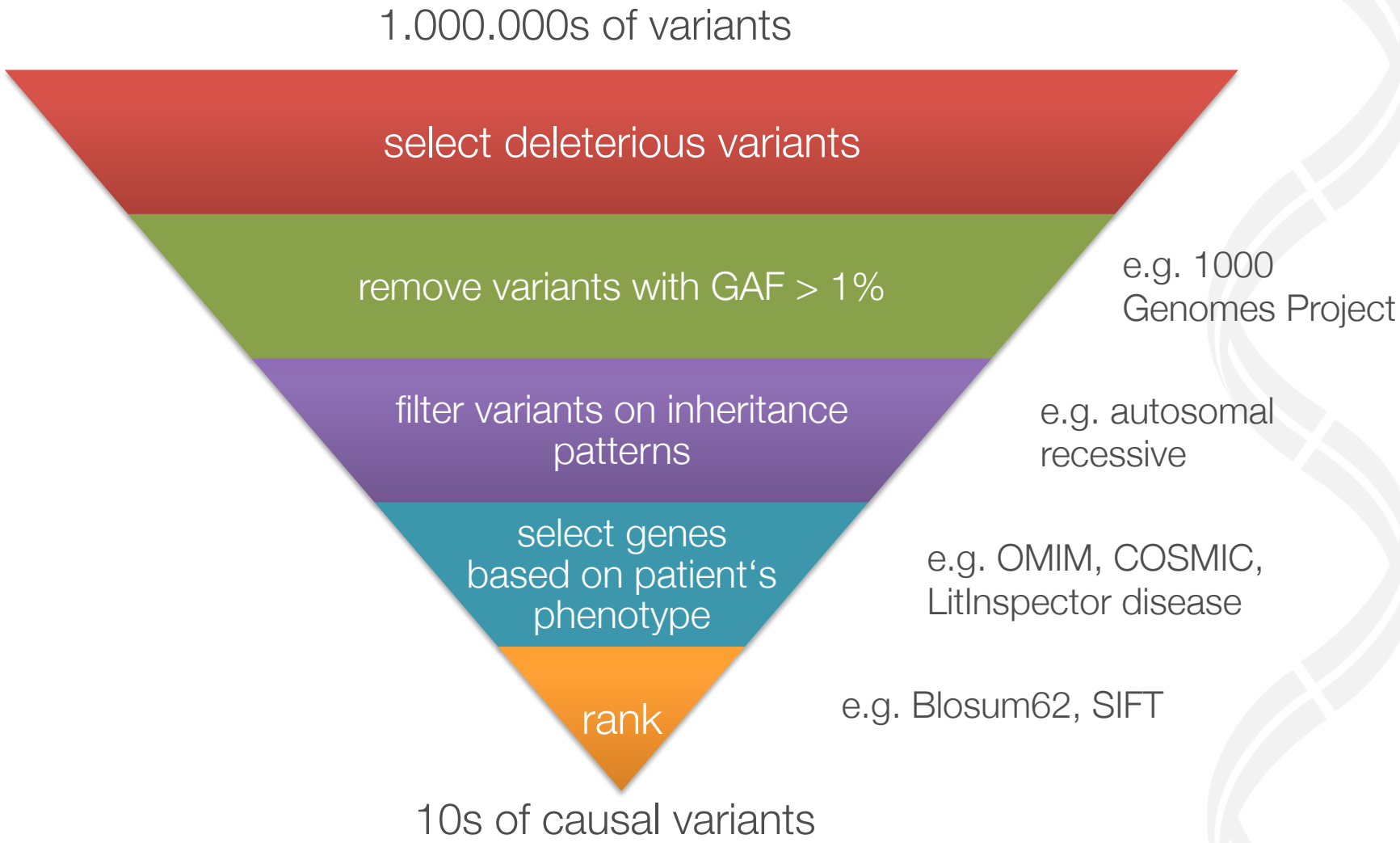
Genomic variants like SNPs or InDels are of major interest to biologists and clinicians.

Identifying causal variants is crucial for the diagnostics of rare and common diseases.

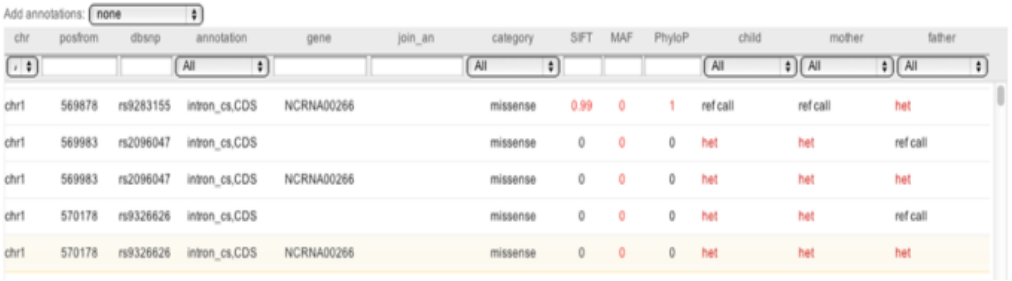
With today's Next Generation Sequencing (NGS) technology it is possible to detect millions of variants within an individual genome.

Which are the relevant ones?

Finding the needle in the haystack

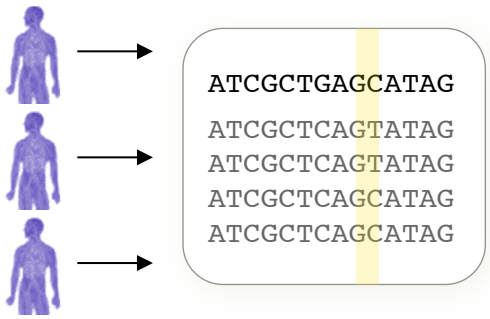


The Genomatix **GeneGrid** system



view, compare
filter, sort

Pre-knowledge database
annotations
disease & medical info



SNPs

import

annotate

GeneGrid database

dbSNP, 1000 Genomes, SIFT,
PhyloP, COSMIC, Genomatix
LitInspector

Proprietary user annotations:
e.g. cancer / disease panels

Results &
Reports

add annotations (e.g., COSMIC)

Filter by annotation: disease (Gx) Show only deleterious: Show only differential:

Compared 3 samples (case: 3) with 62315 variants

Coordinates			Annotation		Prediction						Samples		
chr	position	dbSNP	coding	gene	disease (Gx)	category	GAF	BLOSUM62	TFBS	cons	mother	father	daughter
			All		ALZHEIMER DISE	All		All			All	All	All
chr3	46,480,958	2073495	3utr,CDS	LTF	SIGN OR SYMPTC	missense	-1	-1		yes	het	ref call	het
chr3	46,487,937	6173931	3utr,CDS	LTF	SIGN OR SYMPTC	missense	-1	1		yes	het	ref call	het
chr3	46,501,287	8033738	5utr,CDS	LTF	SIGN OR SYMPTC	missense,synonyı	-1	-2		no	het	hom	het
chr3	51,395,571	1542574	CDS	DOCK3	MELANOMA,ALZI	missense	0.8	-3		no	hom	het	hom
chr3	113,187,648	7323915:	5utr,CDS	SPICE1	EXTRACTED LEAI	missense	-1	1		yes	ref call	het	het
chr3	123,987,496	7318817	CDS,promoter	KALRN	ISCHEMIC CEREBF	initiating	-1			no	het	ref call	het
chr3	124,351,316	1708303	5utr,CDS	KALRN	ISCHEMIC CEREBF	synonymous,nons	0.72	4		yes	het	hom	het
chr3	124,390,722	3565363!	CDS	KALRN	ISCHEMIC CEREBF	missense	-1	-2		yes	het	ref call	ref call
chr3	124,431,841	7906215	CDS	KALRN	ISCHEMIC CEREBF	missense	-1	-3		yes	ref call	ref call	het
chr3	183,960,252	706584	3utr,CDS	ALG3	ALZHEIMER DISEA	missense	-1	-2		no	het	ref call	het
chr3	183,976,241	6767237	CDS	ECE2	THYROID HYALIN	missense	-1	1		yes	het	ref call	het
chr3	183,995,210	1179550	3utr,5utr,CDS	ECE2	THYROID HYALIN	missense	-1	0		yes	ref call	ref call	het
chr3	196,865,242	1134986	5utr,CDS	DLG1	AMERICAN COLLE	missense	-1	1		yes	ref call	het	het
chr4	41,015,899	4861358	CDS	APBB2	PELVIC INFLAMMA	missense	0.69	1		yes	hom	het	het
chr4	109,841,743	1759670!	5utr,CDS,promoter	COL25A1	ALZHEIMER DISEA	missense,synonyı	-1	-2		yes	het	ref call	het

View 1 - 175 of 175

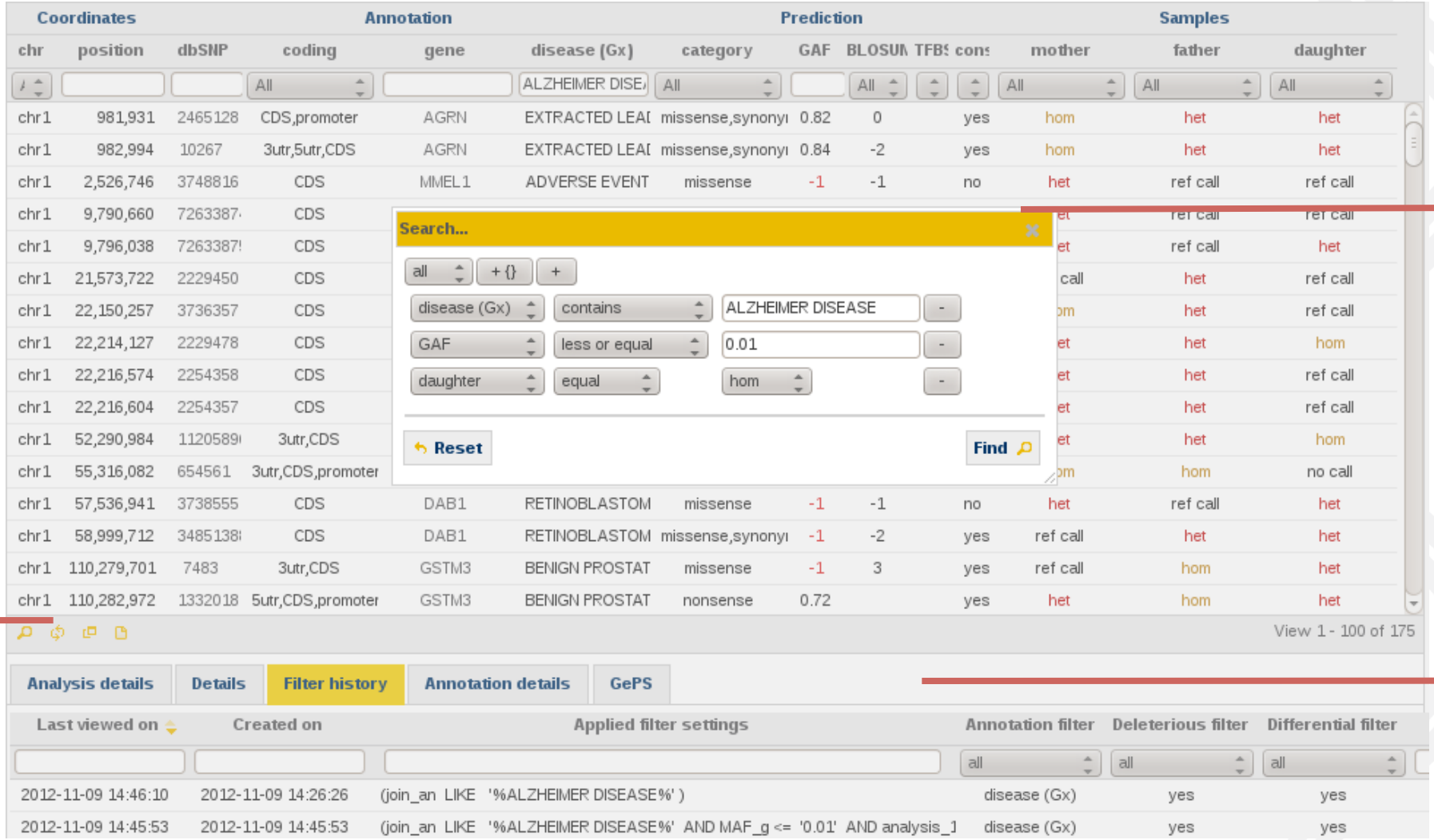
add gene based annotations

additional info

Analysis details | Details | Filter history | Annotation details | GePS

All details details for chr3 113187648

sample	Genotype		Allele			Coverage				Quality		Prediction					
	type	GT	ref	alt1	alt2	total	ref	alt1	alt2	qual	GQ	category	del	Blosum	SIFT	PhyloP	TFBS
daughter	SNP_het	0/1	C	T		87				999	99	missense	yes	1	0.96	0.97745	
father	SNP_het	0/1	C	T		29				999	99	missense	yes	1	0.96	0.97745	
mother	REF_hom	0/0	C			73				999	99		no				



filter, sort and export capabilities

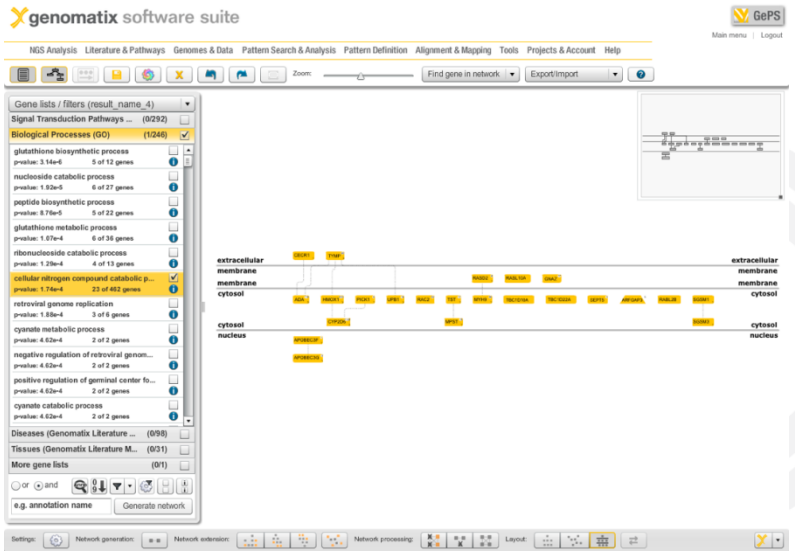
custom filter dialog

filter history

Coordinates			Annotation		Prediction						Samples		
chr	position	dbSNP	coding	gene	disease (Gx)	category	GAF	BLOSUM	TFBS	cons	mother	father	daughter
			All			nonsense	0.01	All			All	All	All
chr1	36,563,467	1492512	3utr,CDS,promoter	COL8A2	BULLOUS KERAT	nonsense	-1			yes	ref call	ref call	het
chr1	47,882,497	3408235	CDS	FOXE3	ABNORMALITY,A	nonsense	-1			yes	no call	ref call	het
chr1	115,236,057	1760272	CDS	AMPD1	CORONARY DISE	nonsense	-1			yes	het	ref call	het

chromosomal position

affected genes



genome browser

pathway system

GeneGrid was used in the CLARITY challenge

Genomatix, CeGaT and the Department of Prostate Cancer Research at the University Hospital Bonn participated in the Boston Children's Hospital's CLARITY challenge.

The competition challenged 30 teams of researchers worldwide to interpret the genomes and exomes of three families with children who suffer from undiagnosed diseases.

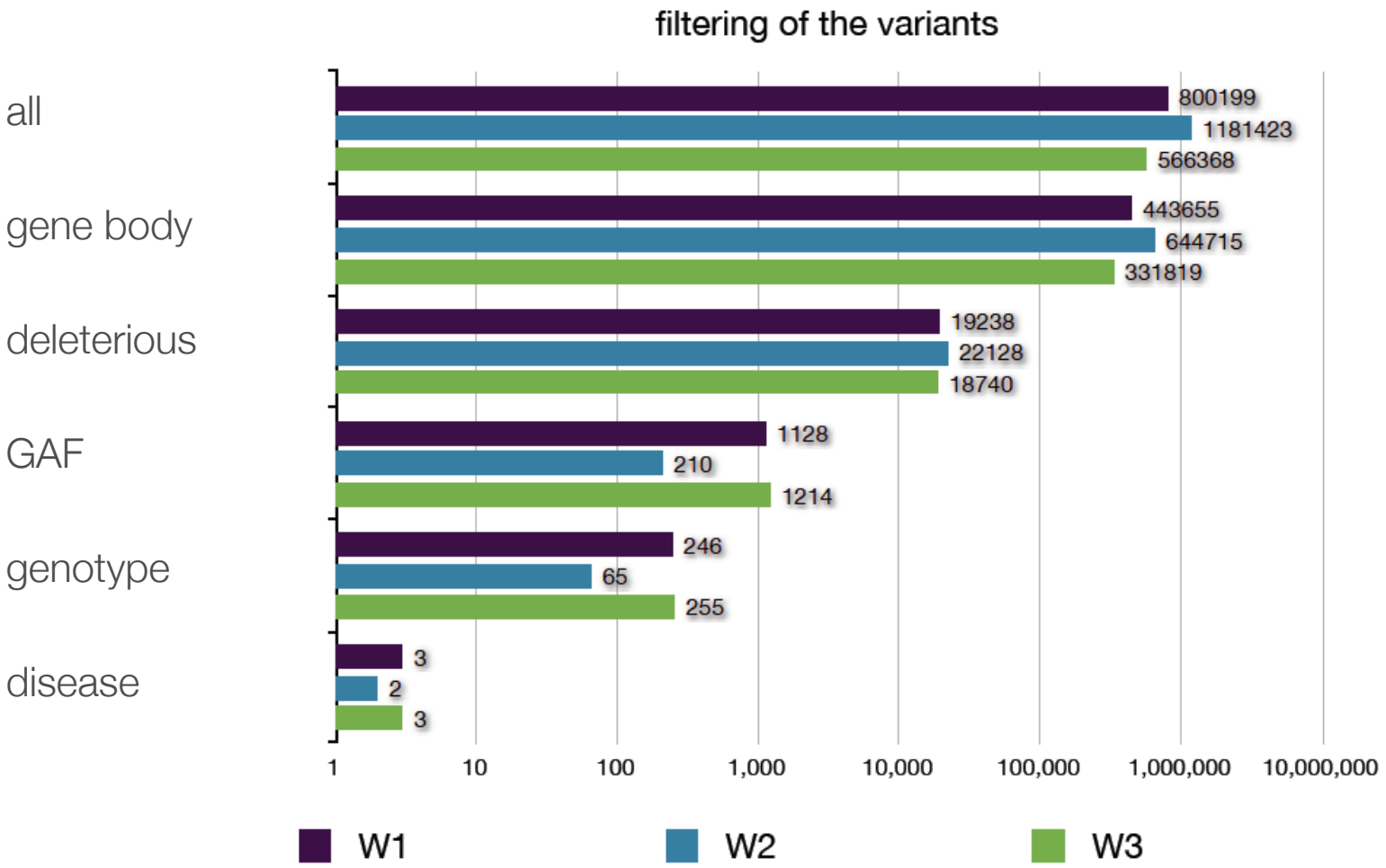
Two of the patients had undiagnosed neuromuscular diseases, the other a cardiovascular disorder.

The aim of the challenge was not only to solve the three cases but also to standardize methods for using genomic information in a clinical setting.

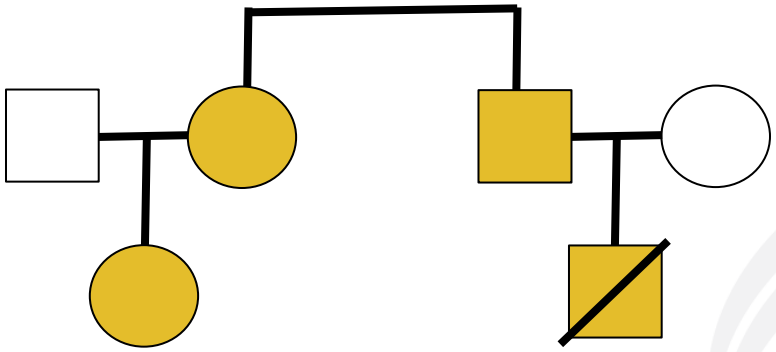
The GeneGrid technology was used to search for genetic disorders.

Children's Hospital Boston

More than 99% of all called variants could be removed

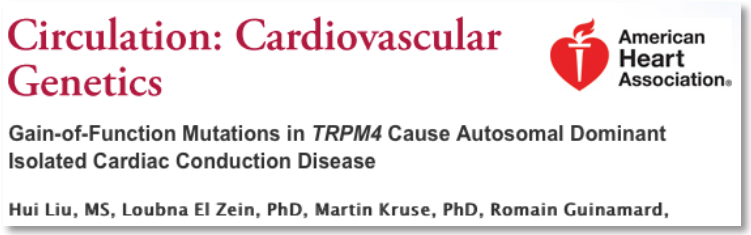


Family 2 – Right bundle branch block and AV block



W2 - RBBB	DB	diagnosis	MESH (1 step)	MESH (2 steps)	panel
terms		Bundle-Branch Block or AV Block	heart block	cardiac arrhythmias	HED
autosomal dominant (47 genes)	OMIM	none	**TRPM4**	none	**TRPM4**
	LitInspector	none	none	**TRPM4**,PHB	**TRPM4**

Family 2 – TRPM4 is associated with familial heart block



Circulation: Cardiovascular Genetics

American Heart Association.

Gain-of-Function Mutations in *TRPM4* Cause Autosomal Dominant Isolated Cardiac Conduction Disease

Hui Liu, MS, Loubna El Zein, PhD, Martin Kruse, PhD, Romain Guinamard,

JCI The Journal of Clinical Investigation

2011 IMPACT FACTOR: 13.069 FREE ACCESS TO ALL RESEARCH

Impaired endocytosis of the ion channel TRPM4 is associated with human progressive familial heart block type I

Martin Kruse[1], Eric Schulze-Bahr[2,3], Valerie Corfield[4],

Published in Volume 119, Issue 9 (September 1, 2009)
J Clin Invest. 2009;119(9):2737–2744. doi:10.1172/JCI38292.
Copyright © 2009, American Society for Clinical Investigation

Challenge Results

Bio·IT World

Next-Gen Technology • Big Data • Personalized Medicine

10th ANNIVERSARY 2002-2012

The Genomatix team where amongst the Top 3 of the CLARITY challenge

„The Genomatix consortium was the only team that "got the right answers" for all three patients" *(BIO-IT World)*

GeneGrid pre-release

GeneGrid is now available for:

- SNPs

GeneGrid will also integrate the following aspects in the next release cycle:

- Small InDels
- large genomic rearrangements and gene fusions
- Epigenetic modifications (DNA Methylation and Histone Modifications
- RNA Expression

Thank you!

Part of this work has been funded by the European Union's BLUEPRINT project.