



**Generade**  
Applied Genomics for Life



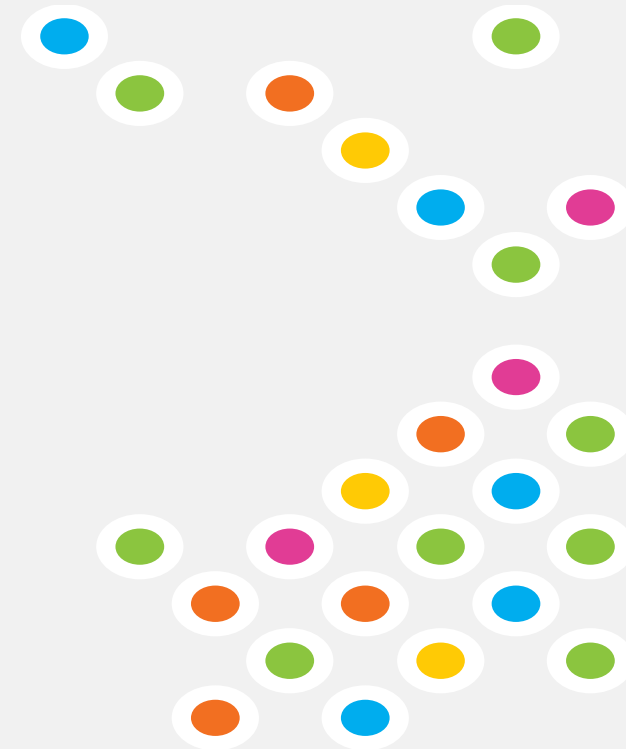
## Genetic profiles in relation to sports: a databased approach

NWO & FAPESP Sports & Healthy Living,

Sao Paulo, March 23 2016

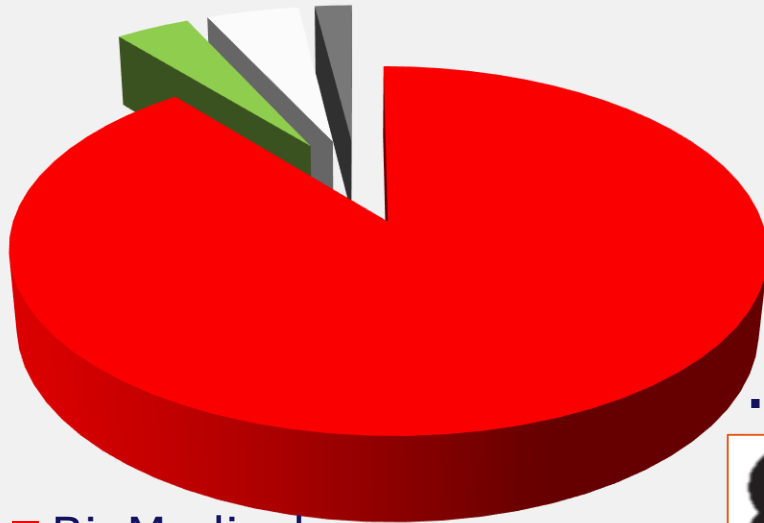
Peter Taschner, Professor Genome-based Health

taschner@generade.nl



# Generade CoE Genomics in perspective

## Leiden BioScience Park



[91] ■ BioMedical

[4] ■ Agro/Food/Plant

[5] ■ BioMarine

[2] ■ Industrieel

## ...and the surrounding region

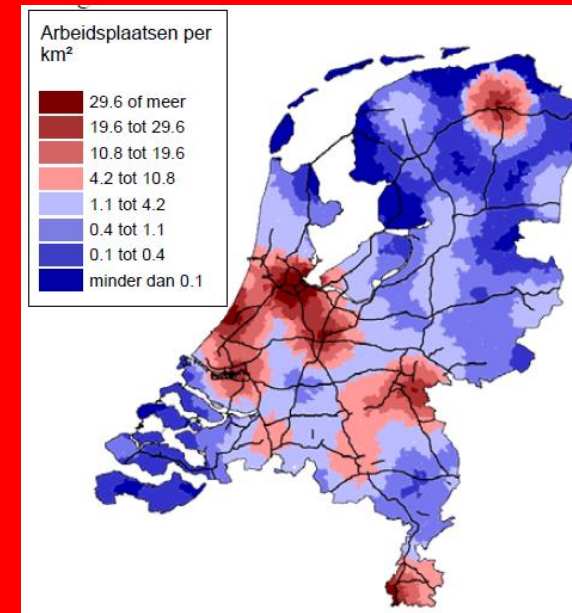


Leiden – Rotterdam – Delft

## Medical Delta

- 10 billion € revenue
- Over 300 life sciences companies
- 3 universities
- 4 universities of applied sciences
- 2 science parks
- Targeted molecular diagnostics, image & image guided therapies, interventions and care, vitality

## Red life sciences / Life Science and Health



# Dutch Sports Research Agenda



- Themes
- Value of Sports
- **Performance**
- **More people, more active, more often**



# Dutch National Research Agenda

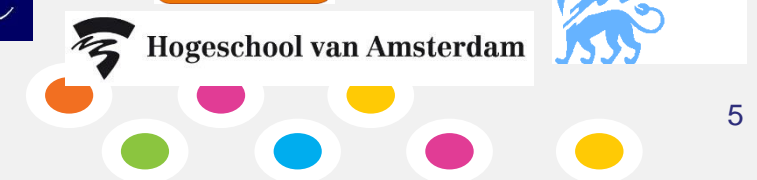


- **Sports and healthy living – research questions**
- 072: How do we improve health and prevent disease through healthy lifestyle and behavior?
- 075: How can we improve health through sports, exercise and nutrition and which effects will result from this?
- 102: How can we develop new drugs and treatments to remain as vital and healthy as possible?

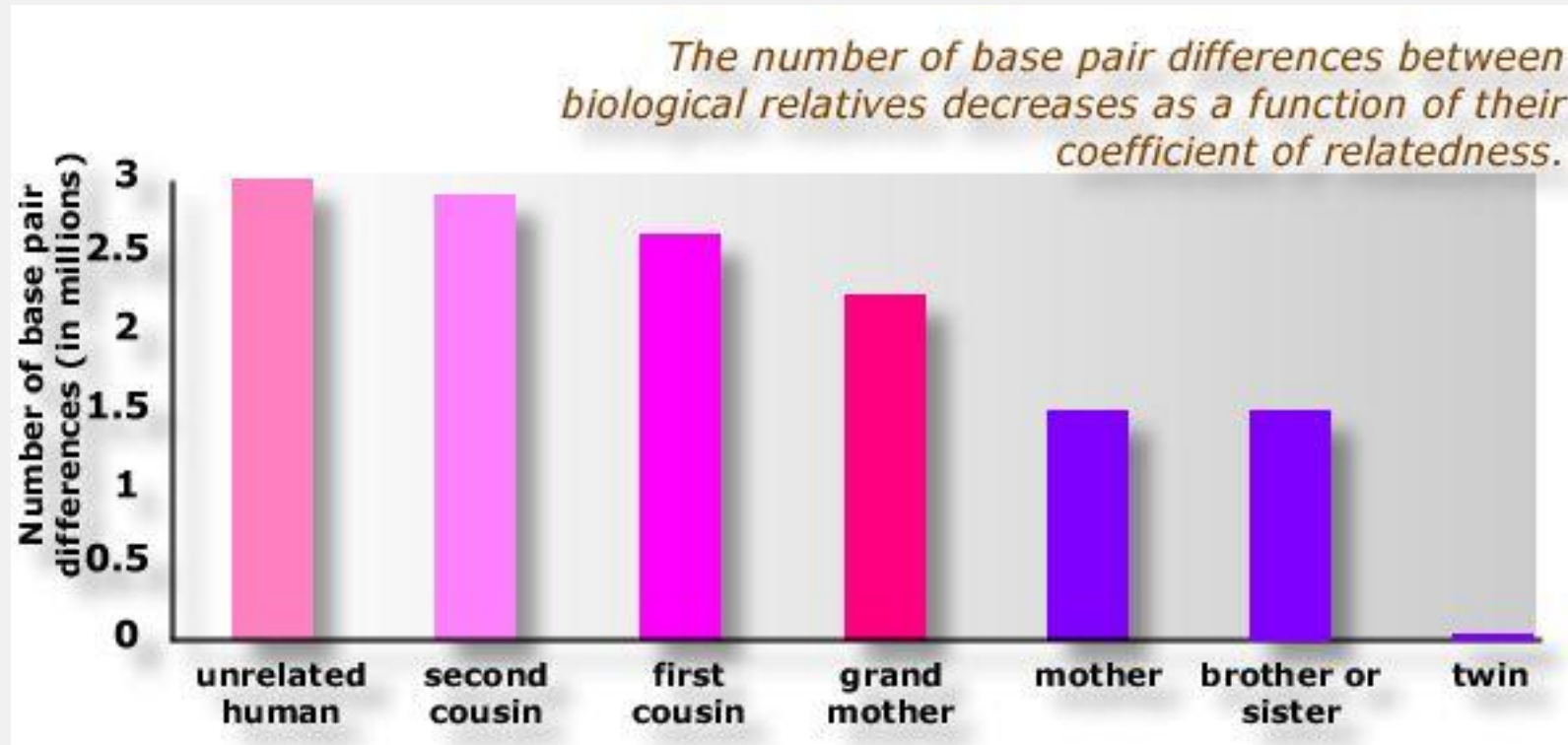


# Current research questions and projects

- How can we identify talent at an early age?
  - Genetic (DNA) profiles?
- How can we select the optimal personal training intensity?
- How to monitor recovery of top athletes, normal and revalidating individuals?
  - From better understanding to improving the processes of training and recovery
- Gene expression (RNA) profiles?
- Other profiles?



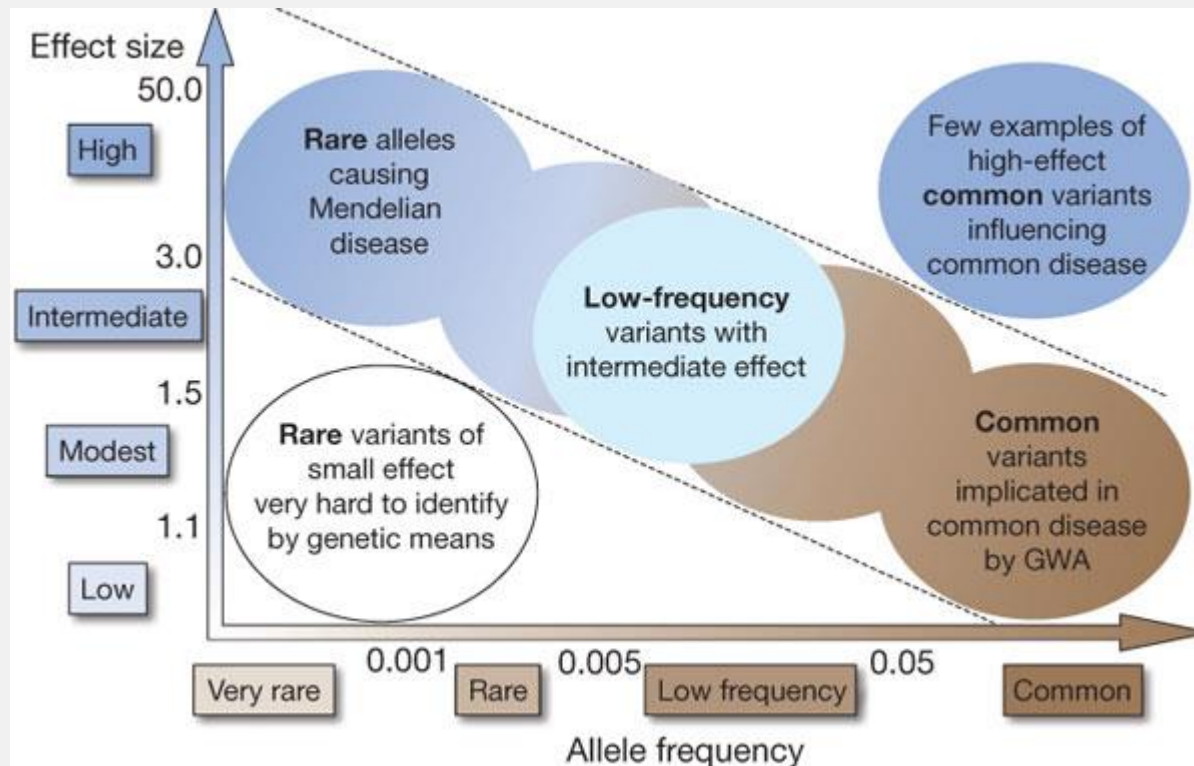
# Much natural variation between individuals



Most variants shared between family members



# Effect size/frequency also applies to sports



TA Manolio *et al. Nature* **461**, 747-753 (2009) doi:10.1038/nature08494

Most variants: small effects

Rare variants:  
specific for family or individual

Variants causing  
sub-clinical musculoskeletal disease?



# DNA profile – Sports genes



**ELITE ATHLETE GENETIC PROFILE**

## JENNY MEADOWS



**AGE:** 32

**WEIGHT:** 50KG

**FASTEST TIME 800M:** 1.57.93

**RECOVERY PROFILE:** FAST

**INJURY RISK:** HIGH

**VO<sub>2</sub> MAX RESPONSE:** MEDIUM

**POWER POTENTIAL:** 48.8%

**51.2% ENDURANCE POTENTIAL**




TEAM GB - 800M

- ★ Jenny Meadows, born 17.04.81. Wigan, UK. Event: 800m
- ★ 2009 World Championships - Bronze
- ★ 2010 European Championships - Bronze
- ★ 2010 World Indoor Championships - Silver
- ★ 2011 European Indoor Championships - Gold
- ★ Third fastest British woman ever over 800m

#DNAthlete dnafit.com

## JENNY MEADOWS



**ENDURANCE / POWER PROFILE & VO<sub>2</sub> MAX POTENTIAL**

GENE	ALLELE	RESULT EFFECT
ACE	ID	Endurance / Power mix
ADRB2	GG	Lower VO <sub>2</sub> max capacity
AGT	CT	No measured impact
ACTN3	CT	Advantage for sprint and power profile, OK for endurance
BDKRB2	TT	Associated with endurance
CRP	GA	Exercise positive for VO <sub>2</sub> max / Endurance profile
IL6	GG	Associated with power performance
NRF	AA	No measured impact on fitness
PPARA	GG	Associated with endurance
PPARGC1A	GG	Power/ Endurance mix
TRHR	TT	No measured impact on fitness
VEGF	CG	Intermediate VEGF production
VDR	CC	Better strength gain, muscle growth

**POST EXERCISE RECOVERY & INJURY RISK**

CRP	GA	Regular exercise has positive impact on recovery
IL6	GG	No measured impact on fitness
IL6R	AC	Associated with intermediate fatigue and longer recovery times
SOD2	TC	Nutritional support for antioxidant function
TNF	GG	Regular exercise has positive impact on recovery
COL1A1	GG	May be more prone to ligament injury
GDF	CT	Intermediate tendinopathy risk

#DNAthlete dnafit.com



Myostatin-related muscle hypertrophy  
*MSTN*: LRG\_200t1:c.373+5G>A

Genome-wide association studies (GWAS)  
 vs  
 Monogenic studies

**OMIM**  
 Online Mendelian Inheritance in Man

**GWAS CENTRAL**



# DNA profile – The new heel prick?

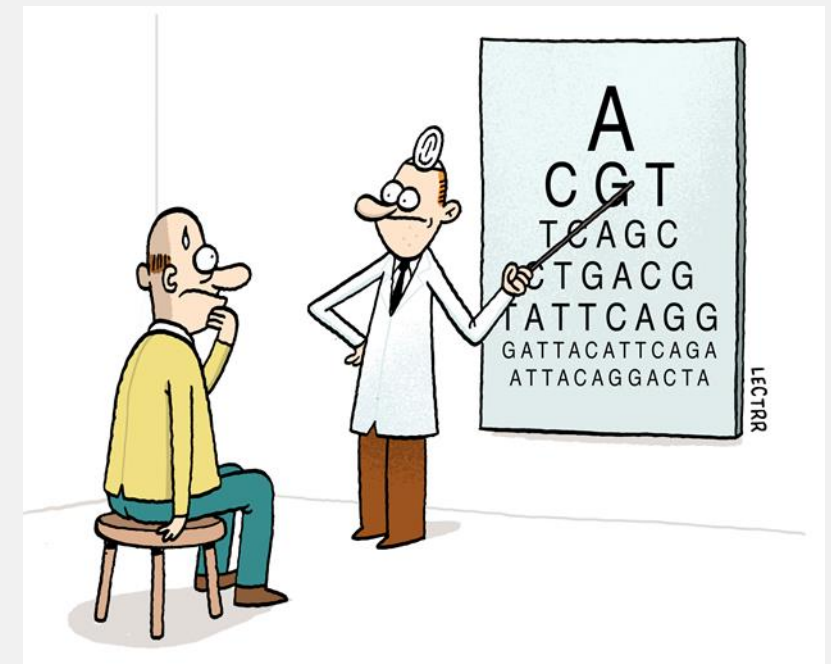
1000 US\$

EDITORIAL

nature  
biotechnology

Knocking on the clinic door

High-throughput sequencing for clinical purposes faces technical and quality challenges, but it's worth it.



LABORATORY FOR MOLECULAR MEDICINE  
65 Landsdowne St, Cambridge, MA 02139  
Phone: (617) 768-8500 / Fax: (617) 768-8513  
http://pcpgm.partners.org/lmm



CENTER FOR PERSONALIZED  
GENETIC MEDICINE



\*\*\*\*\*EXAMPLE REPORT\*\*\*\*\*

**Name:** DOE, JOHN      **Accession ID:** PMXX-12345  
**DOB:** 01/23/1900      **MRN:** 0123456789      **Family #:** F1234657  
**Sex:** Female      **Specimen:** Blood, Peripheral      **Referring physician:** Dr. Med Seq  
**Race:** Caucasian      **Received:** 05/03/2013      **Referring facility:** Brigham and Women's  
**Indication for testing:** MedSeq, Primary Care      **Test:** WGS-pnIA, SeqConV2, WGS-GGR

## GENOME REPORT

### RESULT SUMMARY

Sequencing of this individual's genome was performed and covered 95.3% of all positions at 8X coverage or higher, resulting in over 5.2 million variants compared to a reference genome. These data were analyzed to identify previously reported variants of potential clinical relevance as well as novel variants that could reasonably be assumed to cause disease (see methodology below). All results are summarized on page 1 with further details on subsequent pages.

#### A. MONOGENIC DISEASE RISK: 0 VARIANTS IDENTIFIED

This test did NOT identify genetic variants that may be responsible for existing disease or the development of disease in this individual's lifetime.

#### B. CARRIER STATUS: 3 VARIANTS IDENTIFIED

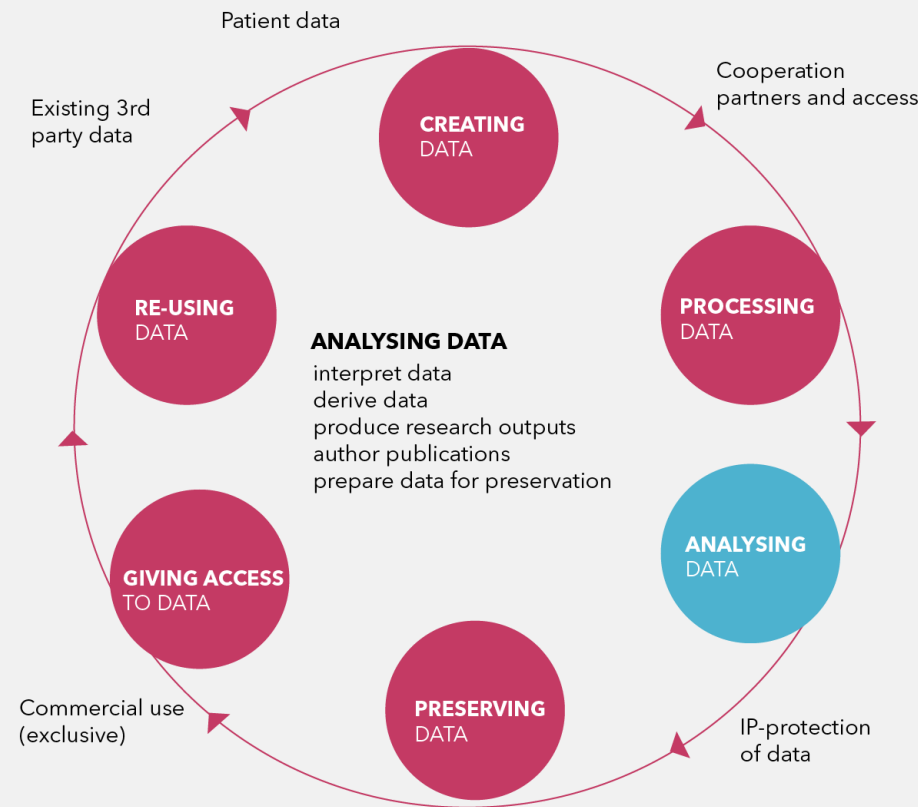
This test identified carrier status for 3 autosomal recessive disorders.

Disease Inheritance	Gene Transcript	Zygoty Variant	Classification	Carrier Phenotype*
B1. Congenital myasthenic syndrome Autosomal recessive	RAPSN NM_005055.4	Heterozygous c.264C>A p.Asn88Lys	Pathogenic	None reported
B2. Cutis laxa, type IC Autosomal recessive	LTBP4 NM_003573.2	Heterozygous c.254delT p.Leu85ArgfsX15	Pathogenic	None reported
B3. Usher syndrome type II Autosomal recessive	USH2A NM_206933	Heterozygous c.609_610insC p.Gly204ArgfsX12	Pathogenic	None reported

As a carrier for recessive genetic variants, this individual is at higher risk for having a child with one or more of these highly penetrant disorders. To determine the risk for this individual's future children to be affected, the partner of this individual would also need to be tested for variants in these genes. Other biologically related family members may also be carriers of these variants. \*Carriers for some recessive disorders may be at risk for certain phenotypes. Please see variant descriptions for more information.



# Research data life cycle



<http://data4lifesciences.nl/hands/handbook-for-adequate-natural-data-stewardship/>



NETHERLANDS FEDERATION OF UMCs  
DATA4LIFESCIENCES

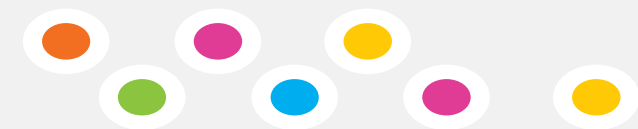
## FAIR data:

**Findable:** data uniquely and persistently identifiable. Others should be able to find your data.

**Accessible:** conditions for use clear to machines and humans.

**Interoperable:** machine-readable data, terminologies, vocabularies, or ontologies commonly used in the field;

**Reusable:** compliant with the above, sufficiently well described with metadata and provenance information supporting linking or integration with other data sources, enabling proper citation



# Where to find genetic variant information?

Efficient information retrieval



**PC Spreadsheet**

Exon	coding DNA description	Protein description	Remarks	VariantID
1	c.120A>A		France, Reported as Promoter Polymorphism	SDHC_00005
1	c.1-7-20>Tdel	p.?	France	SDHC_00043
1	c.1A>G	p.(Met17)	Spain	SDHC_00006
1	c.2T>A	p.(Met17)	Germany	SDHC_00022
1	c.3G>A	p.(Met17)	Germany	SDHC_00001
1	c.20>2A>C	p.?	France	SDHC_00016
1	c.20>12nsTG	p.(?)	France, Allelic freq controls 11/200, patients 63/420	SDHC_00032
1	c.20>13nsTG	p.(?)	Portugal	SDHC_00010
1	c.20>62T>C	p.(?)	France, rs11265689, Allelic freq. controls 10/156, patients 9/420	SDHC_00033
1	c.21>96C>T	p.(?)	France, rs4265463, Allelic freq. controls 17/200, patients 69/426	SDHC_00036
2	c.21>2A>C	p.?	France	SDHC_00035
2	c.21-77>T>Gat	p.?	France, LOH at SDHC locus; SDHC protein expression=0	SDHC_00034
2	c.23A>A	p.(His93Arg*12)	Germany	SDHC_00023
2	c.39C>A	p.(Cys137)	Germany	SDHC_00007
2	c.43C>T	p.(Arg157)	France, Germany, GIST	SDHC_00013
2	c.77-30T>A	p.(?)	Netherlands, No predicted splice site change, not in dbSNP	SDHC_00012
3	c.78A>G	p.?	Poland, Germany, Proven exon 3 skipping, internal gtpo & EA-PGL	SDHC_00017
3	c.126G>A	p.(Trp42*)	Italy	SDHC_00011
3	c.128A>G	p.(His43Gln)	China	SDHC_00046
3	c.166A>T	p.(His59Phe)	Germany	SDHC_00025
3	c.173T>C	p.(His59Thr)	Germany	SDHC_00026
4	c.148C>T	p.(Arg93Cys)	Germany	SDHC_00024
4	c.180T-2628>T>del	p.?	Germany	SDHC_00030
4	c.183G>A	p.(Trp61*)	France	SDHC_00019
4	c.210C>G	p.(Cys70Trp)	Germany	SDHC_00027
4	c.214C>T	p.(Arg72Cys)	Turkey, Poland, functional domain, conserved residue, Q164 controls	SDHC_00008



**Drawer**

**Intranet**

**Internet**

- National Databases

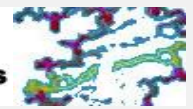


- General Databases



- Ethnic Databases

**dbSNP**  
Short Genetic Variations



**Journals**

**nature genetics**

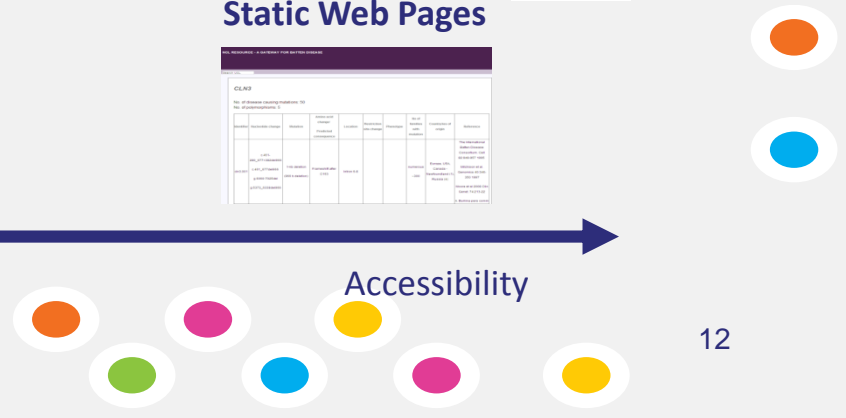
A systematic, large-scale resequencing screen of X-chromosome coding exons in mental retardation

Patrick S Tarczy<sup>1</sup>, Raffaella Smith<sup>1</sup>, Erin Pleasance<sup>1</sup>, Annabel Whibley<sup>2</sup>, Sarah Edkins<sup>3</sup>, Claire Hardy<sup>4</sup>, Sarah O'Meara<sup>5</sup>, Calli Latimer<sup>6</sup>, Ed Dicka<sup>6</sup>, Andrew Monzon<sup>6</sup>, Phil Stephens<sup>6</sup>, Matt Blau<sup>6</sup>, Chris Greenman<sup>6</sup>, Yuh Xue<sup>6</sup>, Chris Tike-Smith<sup>6</sup>, Deborah Thompson<sup>6</sup>, Kristian Gray<sup>6</sup>, Jesse Andrews<sup>6</sup>, Syd Bartholomew<sup>6</sup>, Gemma Buck<sup>6</sup>, Jennifer Cole<sup>6</sup>, Rebecca Danmore<sup>6</sup>, David Jones<sup>6</sup>, Mark Maddison<sup>6</sup>, Tatiana Mironenko<sup>6</sup>, Rachel Turner<sup>6</sup>, Kelly Turrell<sup>6</sup>, Jennifer Varian<sup>6</sup>, Sofie Wood<sup>6</sup>, Sara Wicks<sup>6</sup>, Paul Wray<sup>6</sup>, Jon Tague<sup>6</sup>, Adam Butler<sup>6</sup>, Andrew Jankinson<sup>6</sup>, Mingming Jia<sup>6</sup>, David Richardson<sup>6</sup>, Rebecca Shepherd<sup>6</sup>, Richard Wooster<sup>6</sup>, M Isabel Trisal<sup>6</sup>, Francisco Martinez<sup>6</sup>, Gemma Carvill<sup>6</sup>, Rene Gahrke<sup>6</sup>, Arjan P M de Bruijn<sup>6</sup>, Hans van Bokhoven<sup>6</sup>, Hilde Van Eeck<sup>6</sup>, Jamel Chelly<sup>6</sup>, Martine Raymond<sup>6</sup>, Hans-Wilger Reppert<sup>6</sup>, Fatima E Abadi<sup>6</sup>, Anand K Srivastava<sup>6</sup>, James Cox<sup>6</sup>, Ting Luo<sup>6</sup>, Uma Mahalingam<sup>6</sup>, Jesse Moon<sup>6</sup>, Josef Parrao<sup>6</sup>, Shuhli Mohammed<sup>6</sup>, John E Tolmie<sup>6</sup>, Cheryl Shoshitaishvili<sup>6</sup>, Mark Corbett<sup>6</sup>, Alban Gardier<sup>6</sup>, Eric Haan<sup>6</sup>, Sinduborn Rajarathand<sup>6</sup>, Marie Shaw<sup>6</sup>, Lucienne Vandeleur<sup>6</sup>, Ted Fulstone<sup>6</sup>, Douglas F Easton<sup>6</sup>, Jackie Royle<sup>6</sup>, Michael Partridge<sup>6</sup>, Anne Hacken<sup>6</sup>, Michael Field<sup>6</sup>, Cindy Skinner<sup>6</sup>, Roger E Stevenson<sup>6</sup>, Martin Bokros<sup>6</sup>, Gillian Turner<sup>6</sup>, Charles E Schwartz<sup>6</sup>, Josef Guce<sup>6</sup>, P Lucy Raymond<sup>6</sup>, P Andrew Futreal<sup>6</sup> & Michael R Stratton<sup>6</sup>



**Static Web Pages**

**Accessibility**



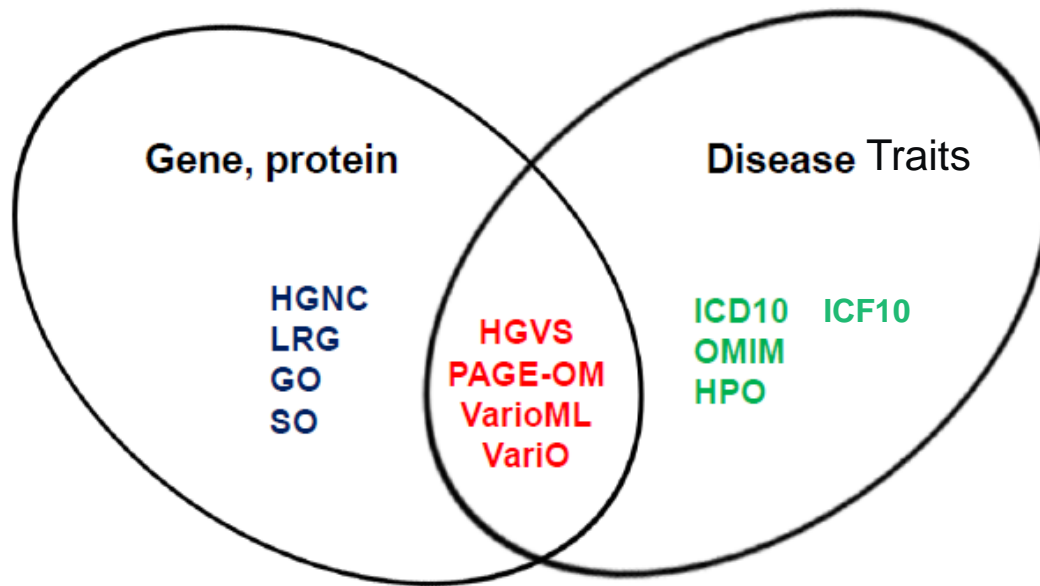
# Genetic variant information sources

- Highly heterogeneous
- Vast amounts of data
- Limited data integration or standardization
- Limited search possibilities across databases
- Difficult to maintain or submit new data

**NOT FAIR: Standardization necessary**



# Standardization



HUGO Gene Nomenclature Committee  
Locus Reference Genomic  
Gene Ontology  
Sequence Ontology

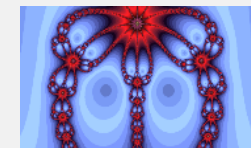
International Classification of Diseases  
Online Mendelian Inheritance in Man  
Human Phenotype Ontology

Human Genome Variation Society Nomenclature  
Phenotype and Genotype Experiment Object Model  
VariOML  
Variation Ontology

## Human Variome Project Recommended systems



<http://www.hgvs.org/mutnomen>



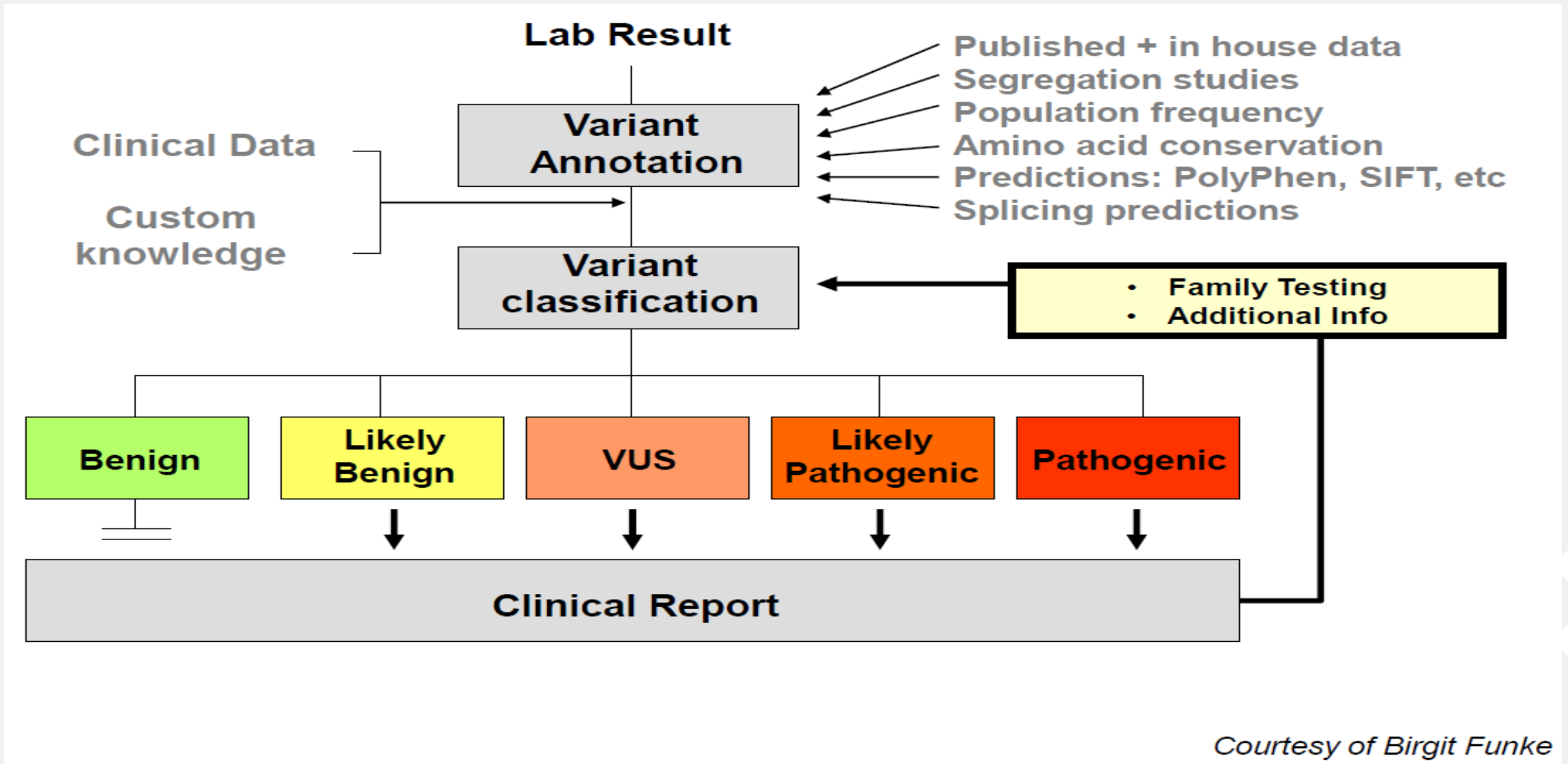
<https://mutalyzer.nl/>



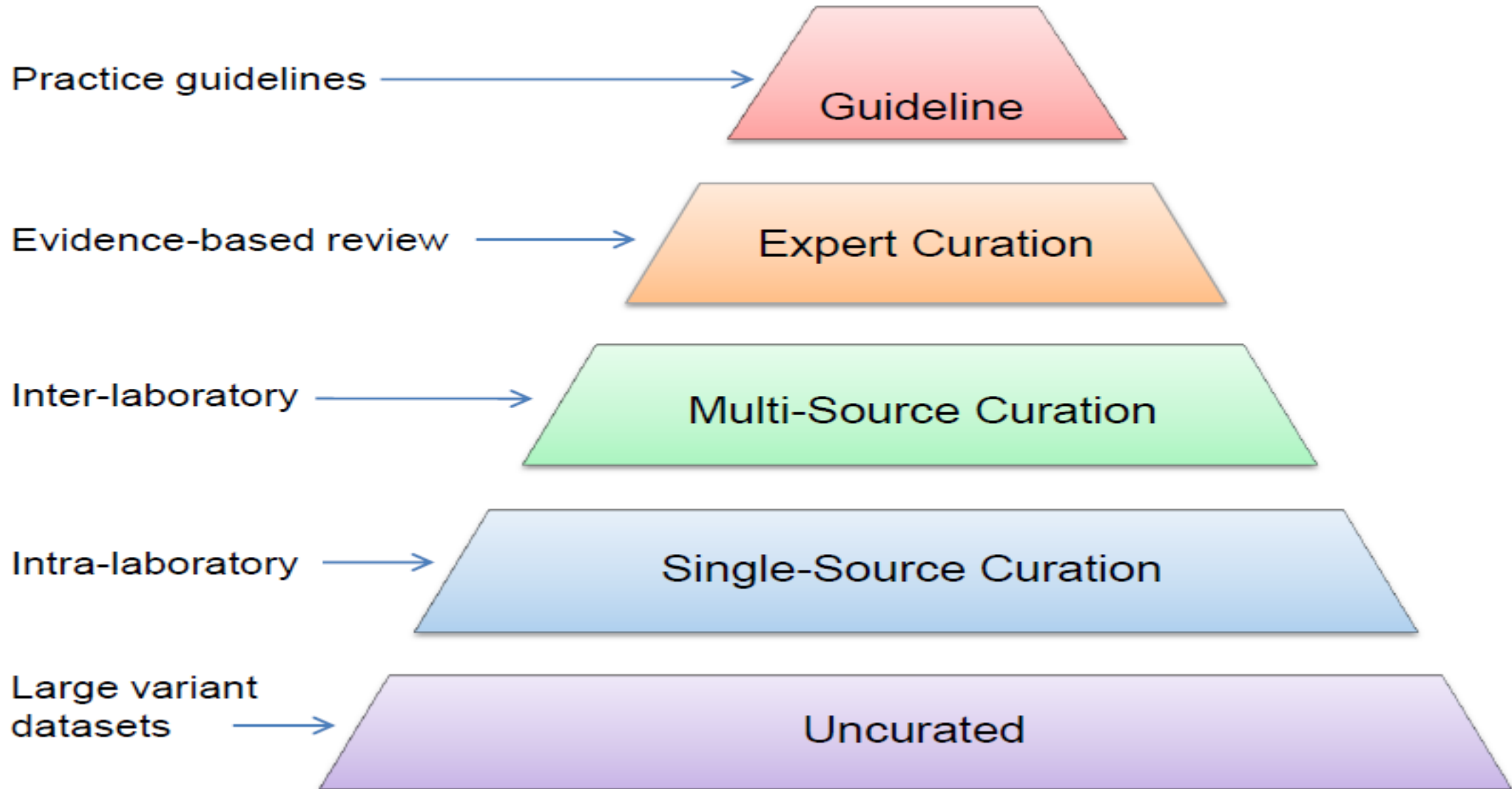
<http://www.lovd.nl/>



# Variant assessment and classification



# Mark variants by level of curation





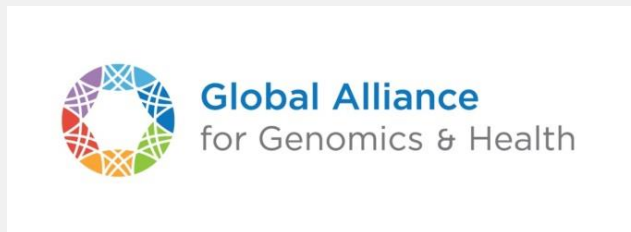
# Matchmaker Exchange



Data sharing:

Genome/exome/variants:  
Who has seen this variant  
or phenotype?

Causative variant identification  
Variant effect determination



[www.ga4gh.org](http://www.ga4gh.org)

Beacons:

Advertise variant information

# BioSHaRE



IT sub-program established to address informatic challenges



Working with BioSHaRE-ELSI to address ethico-legal aspects

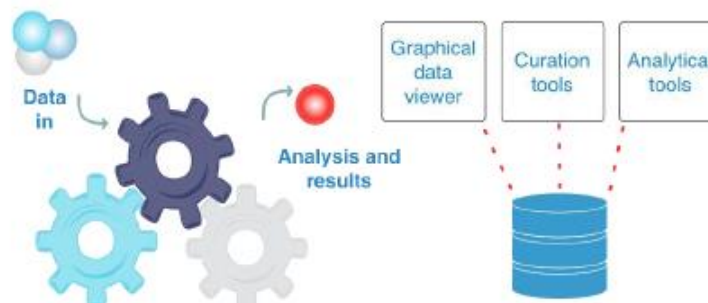
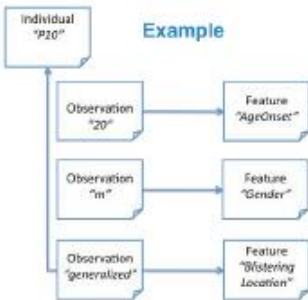
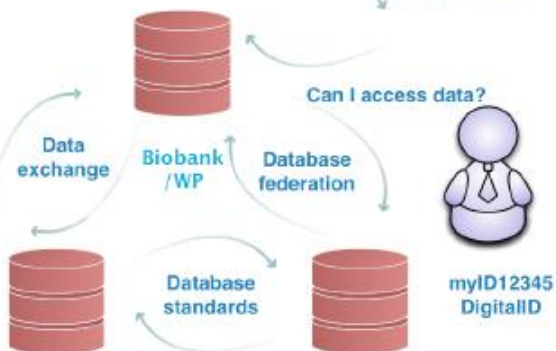
BioSHaRE-ELSI

BioSHaRE-IT

Partner biobanks, BioSHaRE core projects and work packages

Who are you?

Model



Promote standards, formats and digital identities for use in and between biobanks

Data model development for core molecular, phenotype, and environmental data

Implement distributed grid technologies to deal with computational problems and remote analysis

Develop user interface tools

# The promise of genomics for sports, health

- Better prediction of personal risk factors
- You in charge: optimise your lifestyle, nutrition, training to increase your performance and health



Gattaca, 1997



Star Trek Into Darkness, 2009



# Your own genome in your hand



Genome sequence in an Illumina iPad app

Foto: Adam Jeffery CNBC

Apps and databases:

Efficient links to other databases?

Annotation by patients without privacy issues?

Suitable for DNA bank account?

Transfer/provide access to (anonymized) personal data

Different levels: health care and research



# Genotype known: deep phenotyping needed

**Generade - Personal Genomes Database**

Genes | Transcripts | Variants | Individuals | Diseases | Screenings | Submit | Documentation

[View all genes](#)

19810 entries on 199 pages. Showing entries 1 - 100.

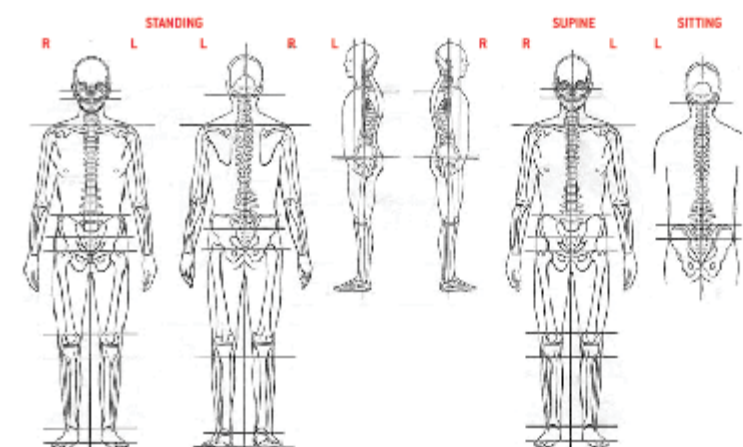
100 per page | << First | < Prev | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | ... | Next | Last >>

Symbol	Gene	Chr	Band	Transcripts	Variants	Unique variants
A1BG	alpha-1-B glycoprotein	19	q13.43	1	13	13
A1CF	APOBEC1 complementation factor	10	q21.1	6	49	49
A2M	alpha-2-macroglobulin	12	p13.31	1	13	13
A2ML1	alpha-2-macroglobulin-like 1	12	p13	2	31	31
A3GALT2	alpha 1,3-galactosyltransferase 2	1	p35.1	1	9	9
A4GALT	alpha 1,4-galactosyltransferase	22	p13.2	6	96	96
A4GNT	alpha-1,4-N-acetylglucosaminyltransferase	3	p14.3	1	13	13
AAAS	achalasia, adrenocortical insufficiency, alacrima	12	q13	2	15	15
AACS	acetoacetyl-CoA synthetase	12	q24.31	4	12	12
AADAC	arylacetamide deacetylase	3	q25.1	3	31	31
AADACL2	arylacetamide deacetylase-like 2	3	q25.1	1	17	17
AADACL3	arylacetamide deacetylase-like 3	1	p36.21	2	39	39
AADACL4	arylacetamide deacetylase-like 4	1	p36.21	1	1	1
AADAT	aminoadipate aminotransferase	4	q33	2	13	13
AAED1	AhpC/TSA antioxidant enzyme domain containing 1	9	q22.32	3	1	1
AAGAB	alpha- and gamma-adaptin binding protein	15	q22.33-q23	4	39	39
AAK1	AP2 associated kinase 1	2	p13.3	9	44	44
AAMDC	adipogenesis associated, Mth938 domain containing	11	q14.1	2	15	15
AAMP	angio-associated, migratory cell protein	2	q	1	14	14
AANAT	aralkylamine N-acetyltransferase	17	q25.1	2	27	27
AAR2	AAR2 splicing factor homolog (S. cerevisiae)	20	q11.23	3	13	13
AARD	alanine and arginine rich domain containing protein	8	q24.11	1	16	16
AARS	alanyl-tRNA synthetase	16	q22.1	1	5	5
AARS2	alanyl-tRNA synthetase 2, mitochondrial	6	p21.1	2	47	47
AARS1	alanyl-tRNA synthetase domain containing 1	17	q21.31	1	19	19
AASDH	aminoadipate-semialdehyde dehydrogenase	4	q12	5	40	40
AASDHPTT	aminoadipate-semialdehyde dehydrogenase-phosphopantetheinyl transferase	11	q22	1	12	12
AASS	aminoadipate-semialdehyde synthase	7	q31.3	1	9	9
AATF	apoptosis antagonizing transcription factor	17	q12	1	9	9
AATK	apoptosis-associated tyrosine kinase	17	q25.3	13	31	31
ABAT	4-aminobutyrate aminotransferase	16	p13.2	4	199	199
ABCA1	ATP-binding cassette, sub-family A (ABC1), member 1	9	q31	1	92	92
ABCA10	ATP-binding cassette, sub-family A (ABC1), member 10	17	q24	2	44	44
ABCA12	ATP-binding cassette, sub-family A (ABC1), member 12	2	q34	2	27	27
ABCA13	ATP-binding cassette, sub-family A (ABC1), member 13	7	p12.3	1	37	37
ABCA2	ATP-binding cassette, sub-family A (ABC1), member 2	9	q34	2	26	26
ABCA3	ATP-binding cassette, sub-family A (ABC1), member 3	16	p13.3	1	21	21
ABCA4	ATP-binding cassette, sub-family A (ABC1), member 4	1	p22	1	16	16

## ASSESSMENT

Please find your results of your completed assessment below:

Name & Surname: Shane Goodwill  
 Age: 37 ID: 7604220080084  
 Contact Number: 082 143 4310 Email: shane.goodwill@gmail.com  
 Medical Aid: Discovery Medical Aid No: 02875210  
 Date: 28 May 2013



- Postural Variation
- Knock - Knee
- Medial Rotation Of The Hip
- Antepulsion
- Pelvic Antversion
- Knee Hyperextension
- Lumbar Hyperlordosis
- Valgus Ankle
- Shoulder Imbalance
- Lateral Pelvic Inclination
- Scoliosis
- Trunk Rotation
- Thoracic Hyperkyphosis
- Winged Scapula
- Shoulder Protraction
- Abducted Scapula
- Medial Rotation Of Shoulders
- Head Tilt



# How to increase performance?

## Designing Babies



**Congratulations, it's a Ronaldo!**

Poll: Let parents choose a child's traits?



Guido de Wert, medical ethicist Maastricht UMC:

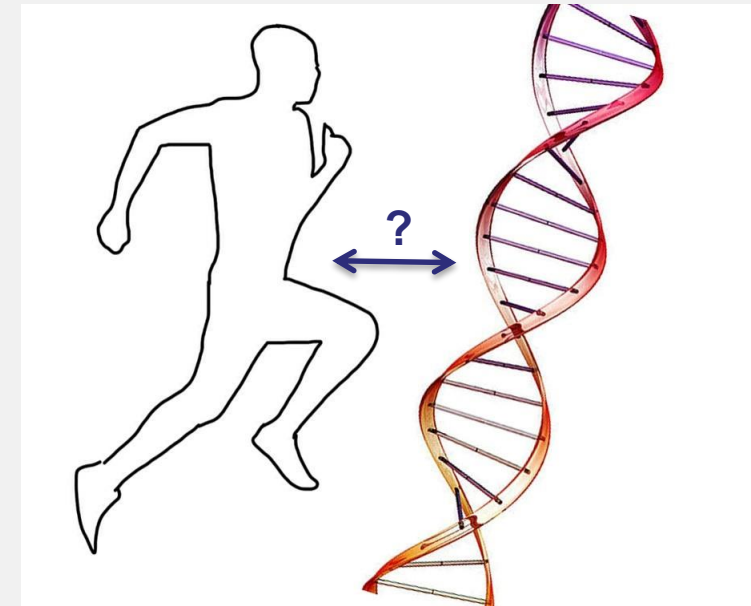
'Research shows that people are not at all obsessed getting perfect children.

Medical science enables us to prevent severe suffering.'



# Challenges

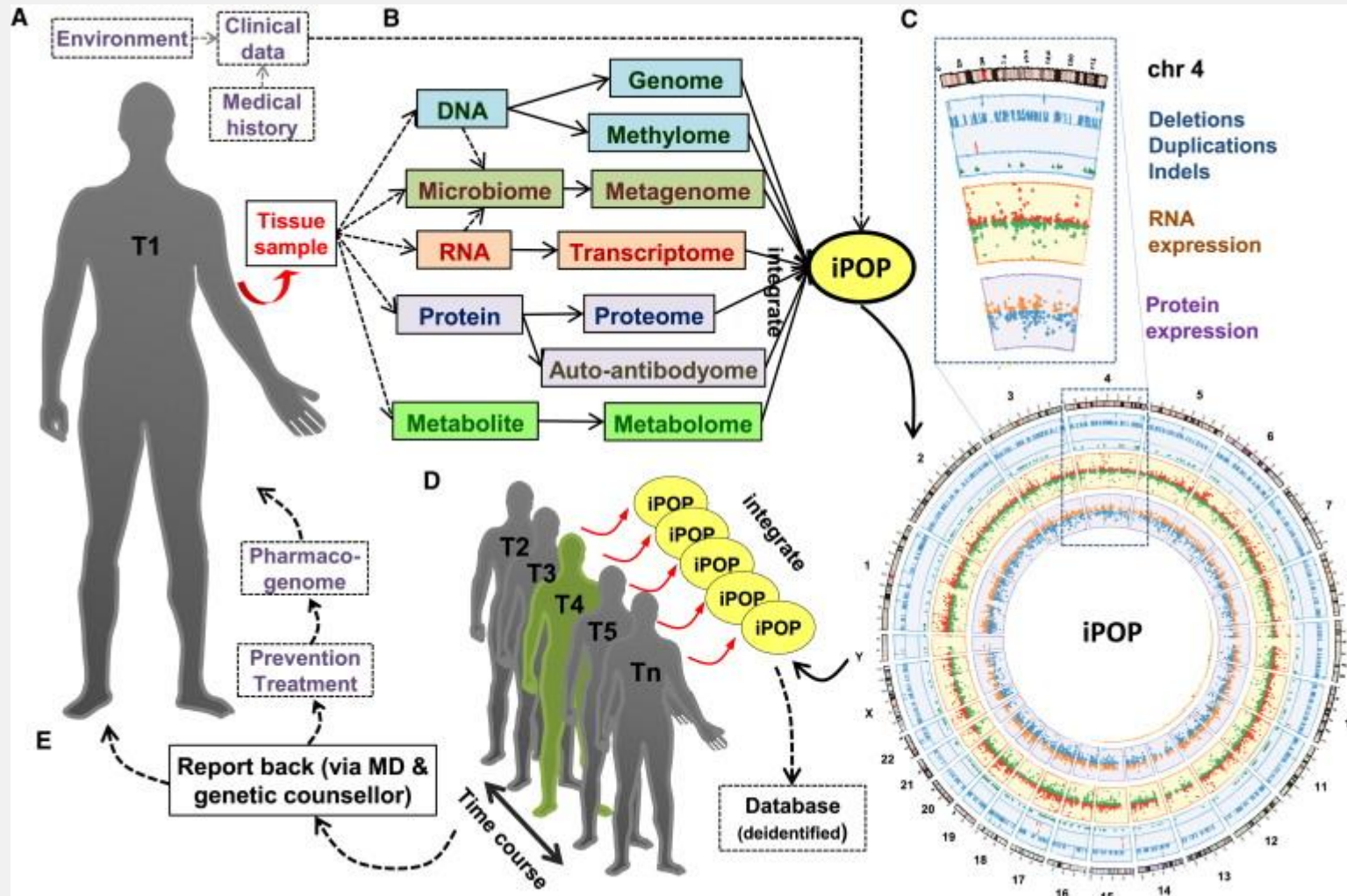
- Data management
- Implementation of innovation by the sports community
- Multidisciplinary research
  
- More knowledge required:
- Genetic basis of sports talent (nature):
  - Variants with small effect size
  - Structural variation
  - Variants in regulatory sequences
  
- Influence of and adaptation to environmental factors: lifestyle, training, etc (nurture)



Phenotype vs DNA-profile (genotype)



# The future: Integrated Personal Omics Profiling



Monitoring:

- at different levels
- longitudinal studies
- preferably on location



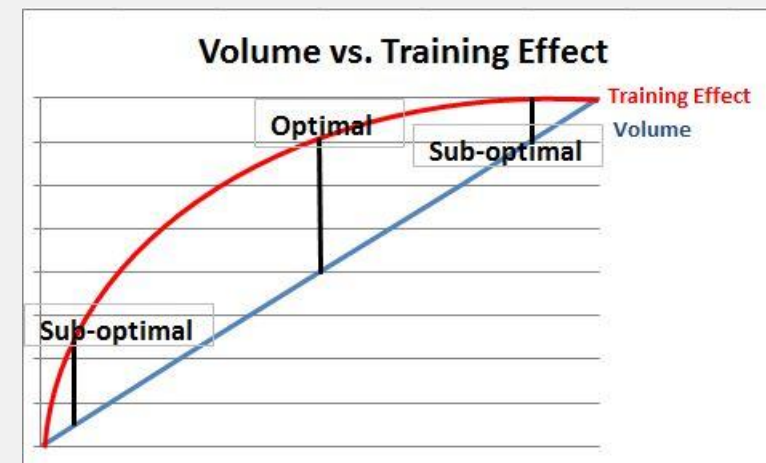
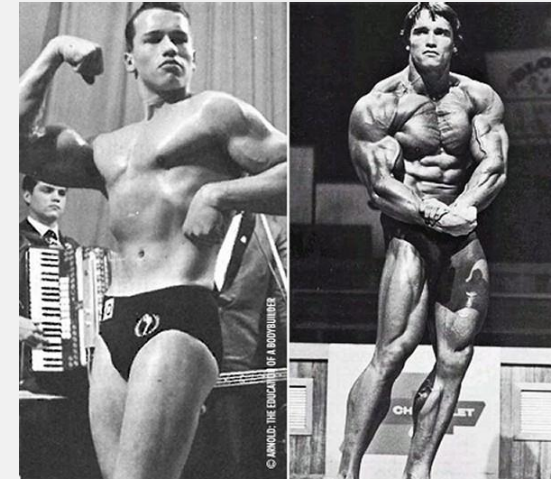
RNA-profiling device?

In combination with traditional sports performance data, apps, wearables

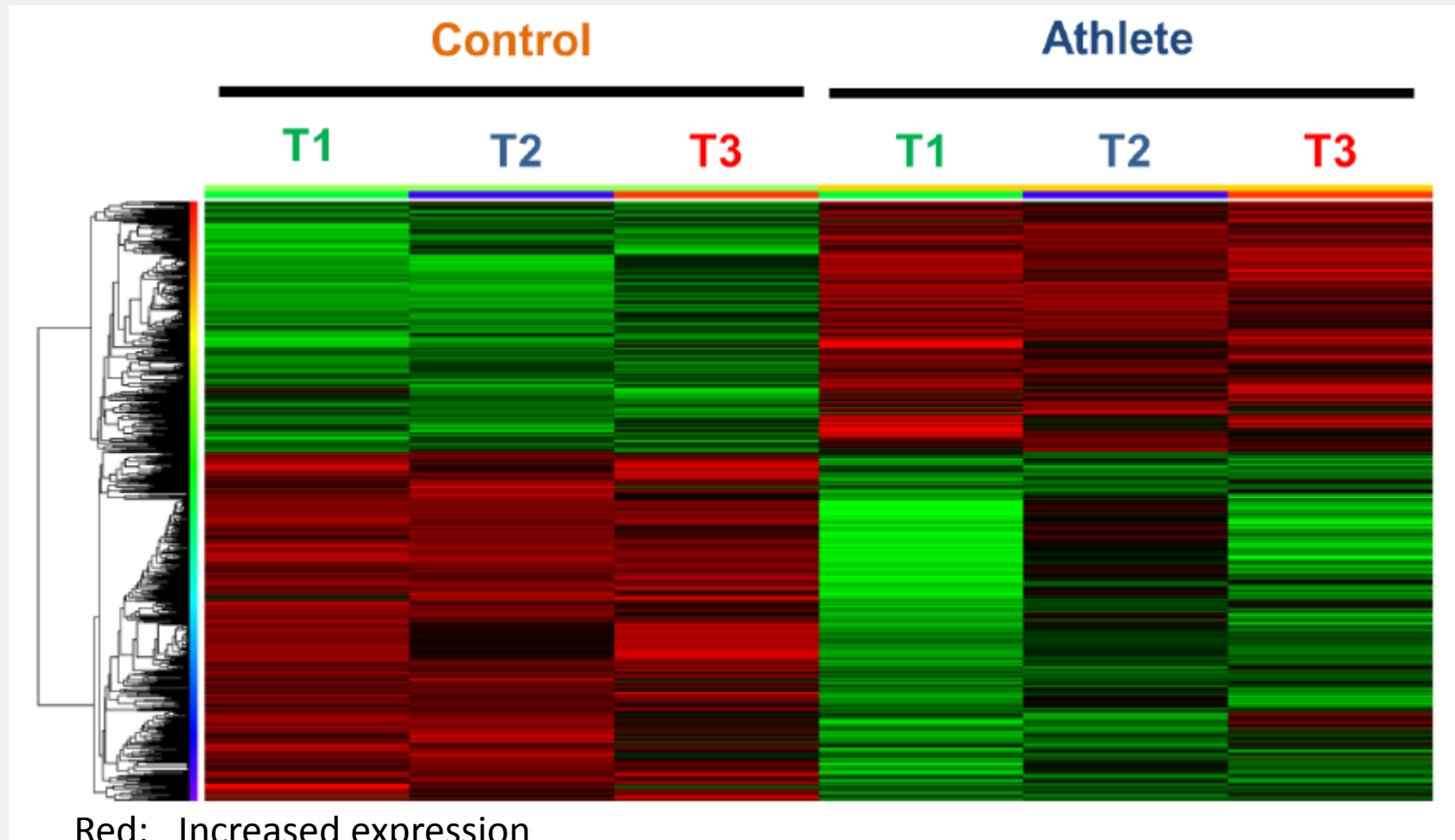


# Monitoring of training effects using RNA-profiles

- How to determine the optimal training intensity?
- Concept:
  - Individuals respond differently to training
  - Training induces stress, changes gene expression
  - Personal RNA-profiles in time (RNA-seq)
  - RNA-profile changes as indicators
- Execution: Blood samples taken at different time points
  - In rest (No intense activity in last 48 hours)
  - Directly after exercise
  - After 48-72 hour recovery



# Significant gene expression differences



T1: Before exercise  
T2: After exercise  
T3: After recovery

Repeat measurement after  
3-monthly training periods

Red: Increased expression

Green: Decreased expression



# Collaboration with interesting partners?

Human sarcopenia studies  
UAS HAN & University Maastricht, NL

Mouse models for sarcopenia +  
Accelerator mass spectrometry  
TNO Metabolic Health Research, Leiden, NL



**Risk factors:** →

- malnutrition
- aging
- immobilization
- COPD
- chemotherapy
- others



Human intervention + ergometer studies  
UAS Utrecht & Wageningen University, NL

food

supplements

medical nutrition

pharma

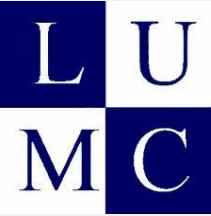
# International data sharing projects

- BRCA challenge ([www.variome.org/brca-challenge.html](http://www.variome.org/brca-challenge.html))
- Globin 2020 ([www.variome.org/gg2020/index.html](http://www.variome.org/gg2020/index.html))
- Aims:
  - Collecting all variants involved in breast cancer and hemoglobinopathies
  - Co-developing infrastructure and capacity in participating countries





[taschner@generade.nl](mailto:taschner@generade.nl)



Thanks!

Questions?

