The Human Phenotype Ontology (HPO)

Andreas Laner   (MGZ München; laner@mgz-muenchen.de)

Sebastian Köhler   (Charité – Universitätsmedizin Berlin; drseb.github.io.)

Peter Robinson   (Jackson Lab; peter.robinson@jax.org)

- History of HPO

- Current status

- Applications

- Workshop/ Demonstration

Why phenotypes matter ?

- Phenotypic abnormality = clinical feature

- Constellation / Pattern of phenotypes / clinical features defines a disease:

  … is a rare developmental disorder defined by the combination of **aplasia cutis congenita of the scalp vertex** and **terminal transverse limb defects**. In addition, **vascular anomalies** such as **cutis marmorata telangiectatica** … are recurrently seen.

OMIM

WILLIAMS-BEUREN SYNDROME; WBS

Alternative titles; symbols

CHROMOSOME 7q11.23 DELETION SYNDROME, 1.5- TO 1.8-MB
WILLIAMS SYNDROME; WMS; WS

Cytogenetic location: 7q11.23   Genomic coordinates (GRCh38): 7:72,700,000-77,900,000

Gene-Phenotype Relationships

Location	Phenotype	Phenotype MIM number	Inheritance	Phenotype mapping key
7q11.23	Williams-Beuren syndrome	194050	AD	4

Clinical Synopsis ▾

▾ **TEXT**

A number sign (#) is used with this entry because Williams-Beuren syndrome (WBS) is a contiguous gene deletion syndrome resulting from the hemizygous deletion of 1.5 to 1.8 Mb on chromosome 7q11.23.

For a discussion of the genes deleted in this syndrome and possible genotype/phenotype correlations, see below.

▾ **Description**

Williams-Beuren syndrome is a multisystem disorder caused by hemizygous deletion of 1.5 to 1.8 Mb on chromosome 7q11.23, which contains approximately 28 genes. Pober (2010) reviewed the clinical features of Williams-Beuren syndrome as well as the genomic and genetic basis and clinical management. ⊕

See also the distal chromosome 7q11.23 deletion syndrome (613729), which occurs between the WBS region and the MAGI2 gene (606382).

▾ **Clinical Features**

Williams et al. (1961) described a syndrome characterized by supravalvular aortic stenosis (SVAS), mental retardation, and distinctive facial features. Beuren et al. (1962) described a similar syndrome with the additional features of dental anomalies and peripheral pulmonary artery stenosis. Two features of the syndrome had been described as distinct entities: supravalvular aortic stenosis

- Free text phenotypic description
- Very expressive ....

OMIM

Clinical Synopsis (**CS**) section

WILLIAMS-BEUREN SYNDROME; WBS

Alternative titles; symbols

CHROMOSOME 7q11.23 DELETION SYNDROME, 1.5- TO 1.8-MB
WILLIAMS SYNDROME; WMS; WS

Cytogenetic location: 7q11.23   Genomic coordinates (GRCh38): 7:72,700,000-77,900,000

Gene-Phenotype Relationships

Location	Phenotype	Phenotype MIM number	Inheritance	Phenotype mapping key
7q11.23	Williams-Beuren syndrome	194050	AD	4

Clinical Synopsis

▾ TEXT

A number sign (#) is used with this entry because Williams-Beuren syndrome (WBS) is a contiguous gene deletion syndrome resulting from the hemizygous deletion of 1.5 to 1.8 Mb on chromosome 7q11.23.

For a discussion of the genes deleted in this syndrome and possible genotype/phenotype correlations, see below.

▾ Description

Williams-Beuren syndrome is a multisystem disorder caused by hemizygous deletion of 1.5 to 1.8 Mb on chromosome 7q11.23, which contains approximately 28 genes. Pober (2010) reviewed the clinical features of Williams-Beuren syndrome as well as the genomic and genetic basis and clinical management. ⊙

See also the distal chromosome 7q11.23 deletion syndrome (613729), which occurs between the WBS region and the MAGI2 gene (606382).

▾ Clinical Features

Williams et al. (1961) described a syndrome characterized by supravalvular aortic stenosis (SVAS), mental retardation, and distinctive facial features. Beuren et al. (1962) described a similar syndrome with the additional features of dental anomalies and peripheral pulmonary artery stenosis. Two features of the syndrome had been described as distinct entities: supravalvular aortic stenosis

INHERITANCE
- Autosomal dominant
GROWTH
Height
- Short stature
Weight
- Abnormal weight gain
Other
- Intrauterine growth retardation (IUGR)
HEAD & NECK
Face
- Medial eyebrow flare
- Flat midface
- Periorbital fullness (puffy eyes)
- Epicanthal folds
- Long philtrum
Ears
- Sensorineural hearing loss, mild to moderate
- Hyperacusis
- Phonophobia
- Abnormal brain auditory evoked responses (BAER)
- Decreased or absent ipsilateral acoustic reflex response to maximum stimulation
Eyes
- Stellate pattern of iris
- Strabismus
- Altered visual acuity
Nose
- Depressed nasal bridge
- Anteverted nares
Mouth
- Thick lips
Teeth
- Hypodontia
- Microdontia
CARDIOVASCULAR
Heart
- Supravalvular aortic stenosis

- Free text phenotypic description
- Very expressive ....

(Un)Controlled Vocabularies

- Non-standardized method for describing phenotypes

- Not designed to be machine interpretable

- Spelling problems

(Un)Controlled Vocabularies

- Non-standardized method for describing phenotypes

- Not designed to be machine interpretable

- Spelling problems

Incomplete:
Fulltext contains phenotype information; absent in **CS**

Inconsistent:
No handling of synonyms

```
'Height: short stature'
'Reduced adult height'
'Final adult height, 84-128cm'
```

(Un)Controlled Vocabularies

- Non-standardized method for describing phenotypes

- Not designed to be machine interpretable

- Spelling problems

CS contains symptoms such as:
'Heart: Prolonged QTc interval'
  or
'T-wave abnormalities'

Imagine query for

'ECG Abnormalities', how to ensure the examples above are found?

(Un)Controlled Vocabularies

- Non-standardized method for describing phenotypes

- Not designed to be machine interpretable

- Spelling problems

E.g.:

hypereflexia	– hyperreflexia
congential	– congenital
defeciency	– deficiency

Homonyms
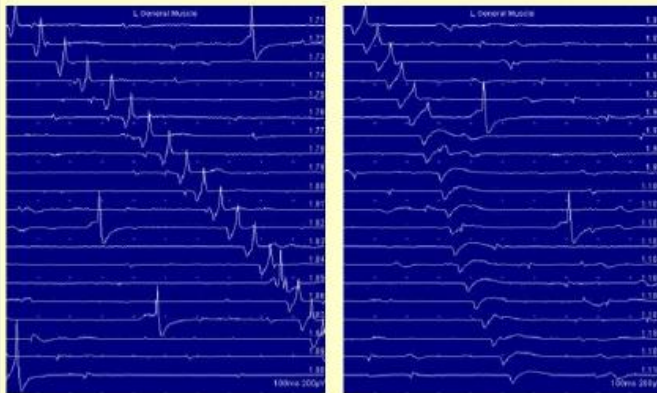
Motivation for HPO Development

OMIM Query	Number of Results
large bones	264
large bone	785
enlarged bones	87
enlarged bone	156
big bones	16
huge bones	4
massive bones	28
hyperplastic bones	12
hyperplastic bone	40
bone hyperplasia	134
increased bone growth	612

Washington et al. *PLoS Biology* (2009)
Linking human diseases to animal models using ontology-based phenotype annotation

Goal of HPO

- Computer-interpretable clinical features

  - Compare diseases based on clinical features

  - Compare patients based on clinical features

  - Compare patients with diseases based on clinical features

  - .....

  - Prioritization of variants in high troughput sequencing assays

- Easy to use and freely available

The Human Phenotype Ontology (HPO)

- Description of **phenotypic abnormalities** (=clinical features) in humans

- Synonyms merged into one term

- Creation of textual and logical definitions for each term
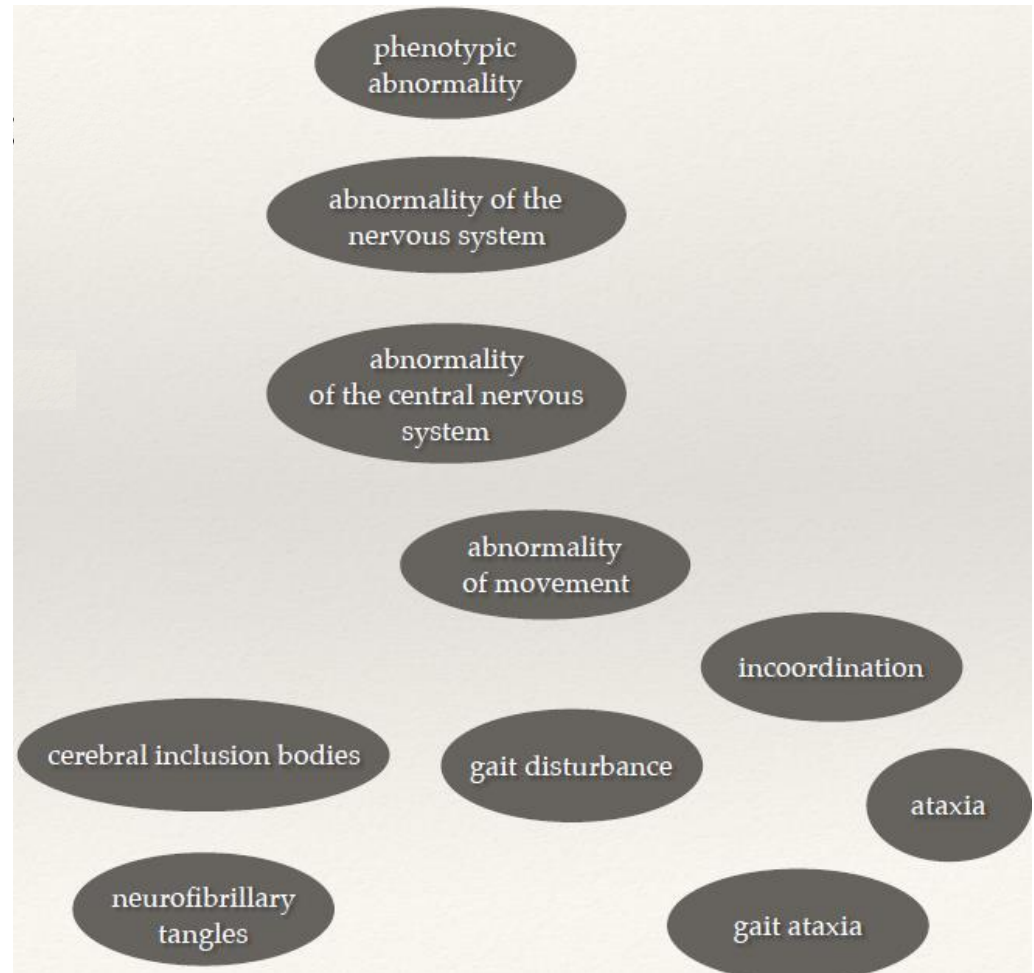
The Human Phenotype Ontology (HPO)
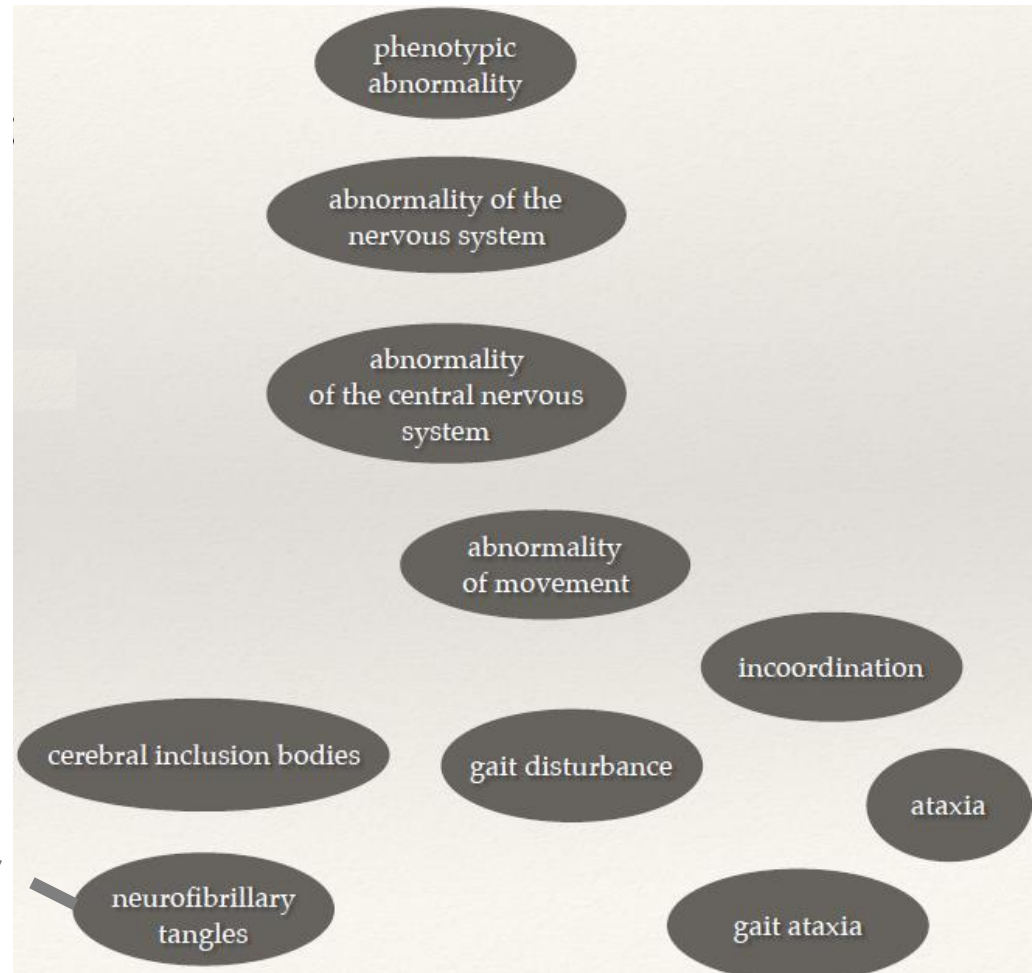
The Human Phenotype Ontology (HPO)

➡ Semantic relations
(is a sublass of)

➡ From top to bottom terms get
more specific
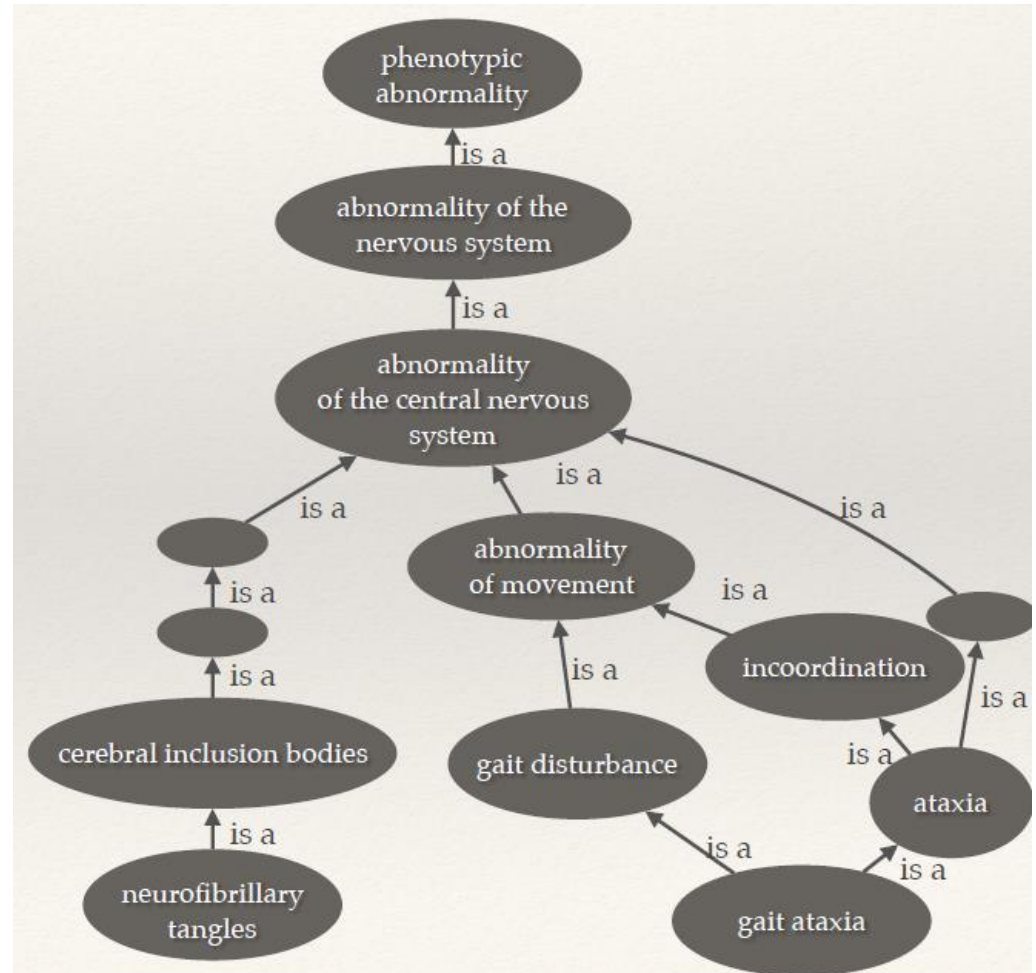
Annotation of diseases

➡ HPO terms are used to **annotate** (describe) diseases

- E.g. *neurofibrillary tangles* is used to annotate Alzheimer Disease

➡ Note: Annotation with *neurofibrillary tangles* induces annotation to all ancestor terms

Köhler et al.
The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data ; NAR (2014)
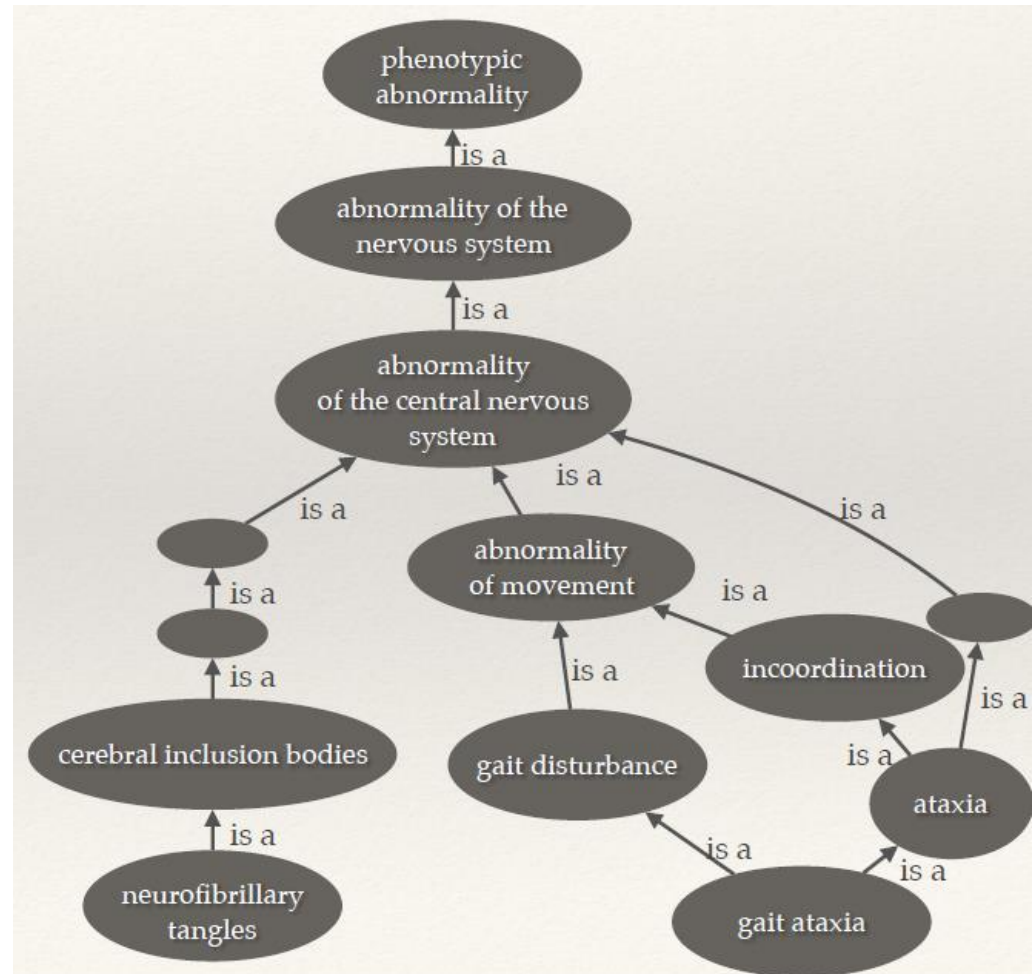
Current Status

➡ 4 root classes

- **Phenotypic abnormality**
- Mode of Inheritance
- Clinical modifier
- Mortality/ Aging

➡ Over 13.000 terms in HPO

➡ Over 156.000 annotations of > 7.700 rare diseases (OMIM, Orphanet, DECIPHER)

➡ Over 133.000 annotations of > 3.100 common diseases

Recent projects

➡ Translations of labels, synonyms and textual definitions (crowdsourcing)

Recent projects

➡ Translations of labels, synonyms and textual definitions (crowdsourcing)

Needs Translation:

Arabic	Chinese Simplified	French	German	Hindi
translated: 1%	translated: 0%	translated: 1%	translated: 10%	translated: 0%

Italian	Japanese	Portuguese	Russian	Spanish
translated: 19%	translated: 0%	translated: 1%	translated: 9%	translated: 19%
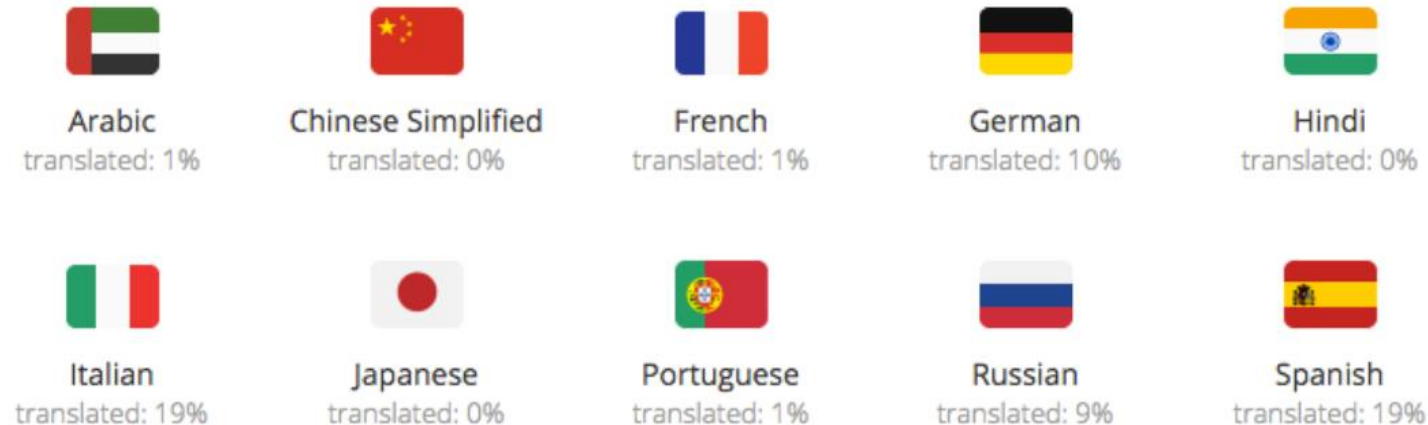
➡ Numbers from 2017

➡ The HPO team needs help to get HPO available in our language as well

Sebastian Köhler   (Charité – Universitätsmedizin Berlin; **drseb.github.io**)

Sources of HPO Terms

Sources of HPO Terms

Sources of HPO Terms

Sources of HPO Terms

Adoption of HPO

Table 1. A selection of public-facing clinical databases using HPO to annotate patient data for disease-gene discovery projects

Name	URL
PhenomeCentral	phenomecentral.org
DDD (Deciphering Developmental Disorders)	www.ddduk.org
DECIPHER (DatabasE of genomiC varIation and Phenotype in Humans using Ensembl Resources)	decipher.sanger.ac.uk
ECARUCA (European Cytogeneticists Association Register of Unbalanced Chromosome Aberrations)	http://umcecaruca01.extern.umcn.nl:8080/ecaruca/ecaruca.jsp
The 100 000 Genomes Project	https://www.genomicsengland.co.uk/
Geno2MP (Exome sequencing data linked to phenotypic information from a wide variety of Mendelian gene discovery projects)	http://geno2mp.gs.washington.edu
NIH UDP (Undiagnosed Diseases Program)	available via phenomecentral.org
NIH UDN (Undiagnosed Diseases Network)	available via phenomecentral.org
HDG (Human Disease Gene Website series)	www.humandiseasegenes.com
Phenopolis(anopen platform for harmonization and analysis of sequencing andphenotype data)	https://phenopolis.github.io
GenomeConnect (Patient portal developed by ClinGen (67)	www.genomeconnect.org
FORGE Canada & Care4Rare Consortium	available via phenomecentral.org
RD-Connect	platform.rd-connect.eu
Genesis	thegenesisprojectfoundation.org

Köhler et al.; The human phenotype ontology in 2017. NAR (2016)

Adoption of HPO

Table 2. Tools and applications using HPO

Tool

Phenotype-driven differential diagnosis
Phenomizer
BOQA
FACE2GENE
Phenolyzer

Phenotype-driven exome/genome analysis
Exomiser
PhenIX
Phevor
PhenoVar
eXtasy
OMIMExplorer
Phen-Gen
Geno2MP
Genomiser
SimReg
ontologySimilarity

Functional and network analysis
TopGene/ToppFunn
WebGestalt
SUPERFAMILY
GREAT
Random walk on heterogeneous network
PANDA
PREDICT

Clinical data management and analysis
Phenotips
Patient Archive
GENESIS (GEM.app)

Cross-species phenotype analysis
PhenoDigm
MouseFinder
Monarch
PhenomeNet
UberPheno
MORPHIN
PhenogramViz

Phenotype knowledge resources and databases
Orphanet
MalaCards
NIH genetic testing registry
OMIM
dcGO
ClinVar
GeneSetDB
MSeqDR
DIDA (digenic diseases database)
Genetic and Rare Diseases (GARD) Information Center

Visualization
PhenoStacks
PhenoBlocks
DECIPHER (phenogram)
phenogrid
ontologyPlot

Köhler et al.; The human phenotype ontology in 2017. NAR (2016)

Adoption of HPO

Table 2. Tools and applications using HPO

Tool

Phenotype-driven differential diagnosis
➡ Phenomizer
BOQA
FACE2GENE
Phenolyzer
Phenotype-driven exome/genome analysis
➡ Exomiser
➡ PhenIX
Phevor
PhenoVar
eXtasy
OMIMExplorer
Phen-Gen
Geno2MP
Genomiser
SimReg
ontologySimilarity
Functional and network analysis
TopGene/ToppFunn
WebGestalt
SUPERFAMILY
GREAT
Random walk on heterogeneous network
PANDA
PREDICT
Clinical data management and analysis
➡ Phenotips
Patient Archive
GENESIS (GEM.app)

Cross-species phenotype analysis
PhenoDigm
MouseFinder
Monarch
PhenomeNet
UberPheno
MORPHIN
PhenogramViz
Phenotype knowledge resources and databases
Orphanet
MalaCards
NIH genetic testing registry
OMIM
dcGO
ClinVar
GeneSetDB
MSeqDR
DIDA (digenic diseases database)
Genetic and Rare Diseases (GARD) Information Center
Visualization
PhenoStacks
PhenoBlocks
DECIPHER (phenogram)
phenogrid
ontologyPlot

Köhler et al.; The human phenotype ontology in 2017. NAR (2016)

Phenotyp-driven Differential Diagnosis

Phenomizer: Search for diseases or differential diagnosis with HPO terms

Phenotyp-driven Differential Diagnosis

Phenomizer: Search for diseases or differential diagnosis with HPO terms

Phenotyp-driven Differential Diagnosis

Phenomizer: Search for diseases or differential diagnosis with HPO terms
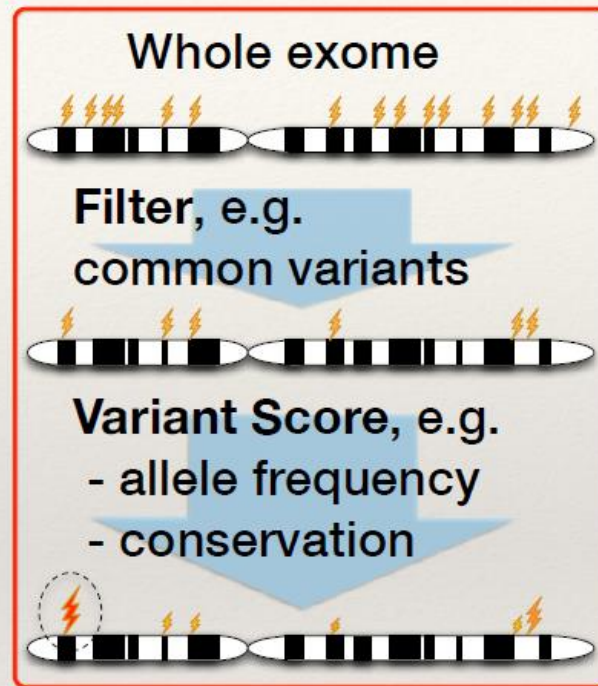
Clinical Data Management and Analysis

Adoption of HPO

* Basic idea of ontological search: Do not need exact match! But semantically similar diseases score well.

* Image a BLAST-search for sets of clinical features. (Phenomizer)

Clinical Genomics

* "Standard" clinical exome pipeline

* Predicts causative variant based on information in genome of patient and background genomic data

* Each human genome harbors about 100 genuine loss-of-function SNVs with ~20 genes completely inactivated (3) and around 50-100 CNVs. (DG MacArthur et al., *Science* 2012)

Clinical Genomics
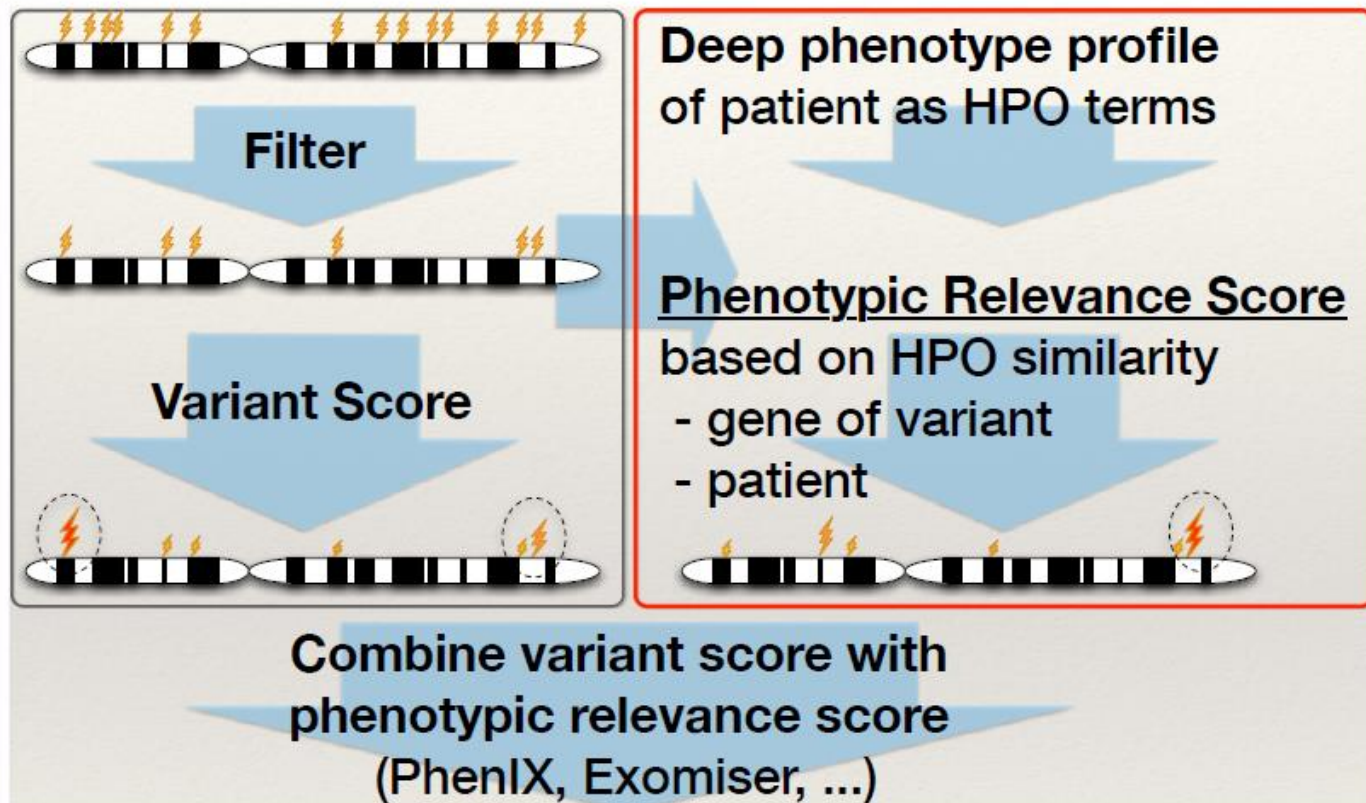
Robinson, Köhler, et al.
Improved exome prioritization of disease genes through cross-species phenotype comparison
Genome Research (2013)

Zemojtel, Köhler et al.
Effective diagnosis of genetic disease by computational phenotype analysis of the disease-associated genome
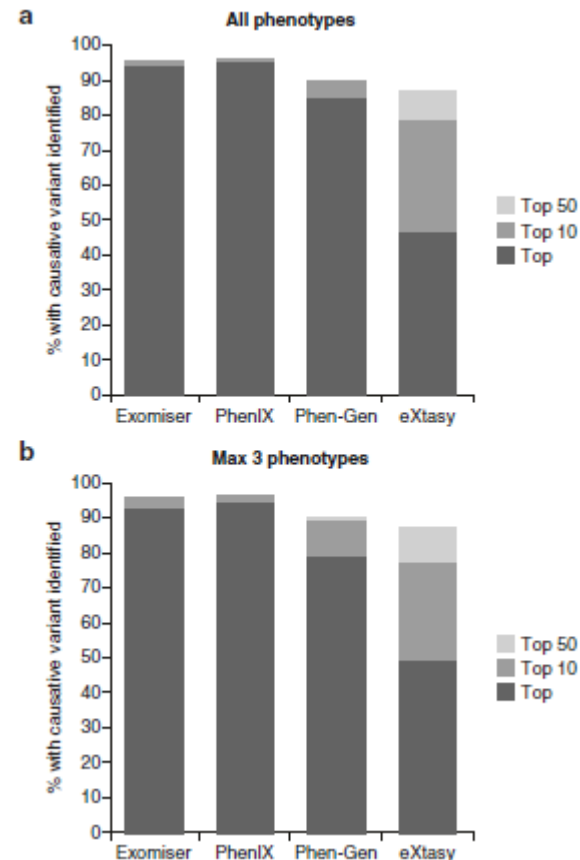Science Translational Medicine (2014)
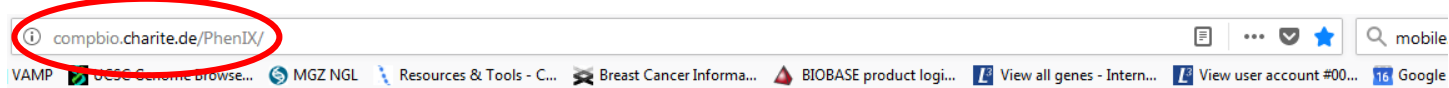
Clinical Genomics

Performance
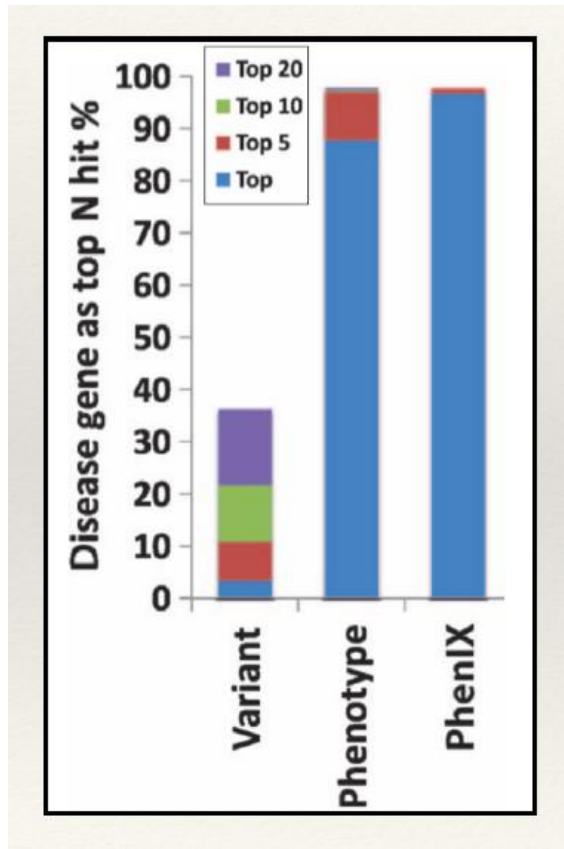
Combination of variant score
and phenotype score is key

Other tools

Phenotyp-driven Exome/Genome Analysis

Phenotyp-driven Exome/Genome Analysis

Exomiser – filtered genes assigned a phenotypic relevance score based on comparison with mouse and zebrafish models with mutations in orthologous genes

Genes with variants that survive the initial filtering steps are screened for mouse models with phenotypic to the human disease
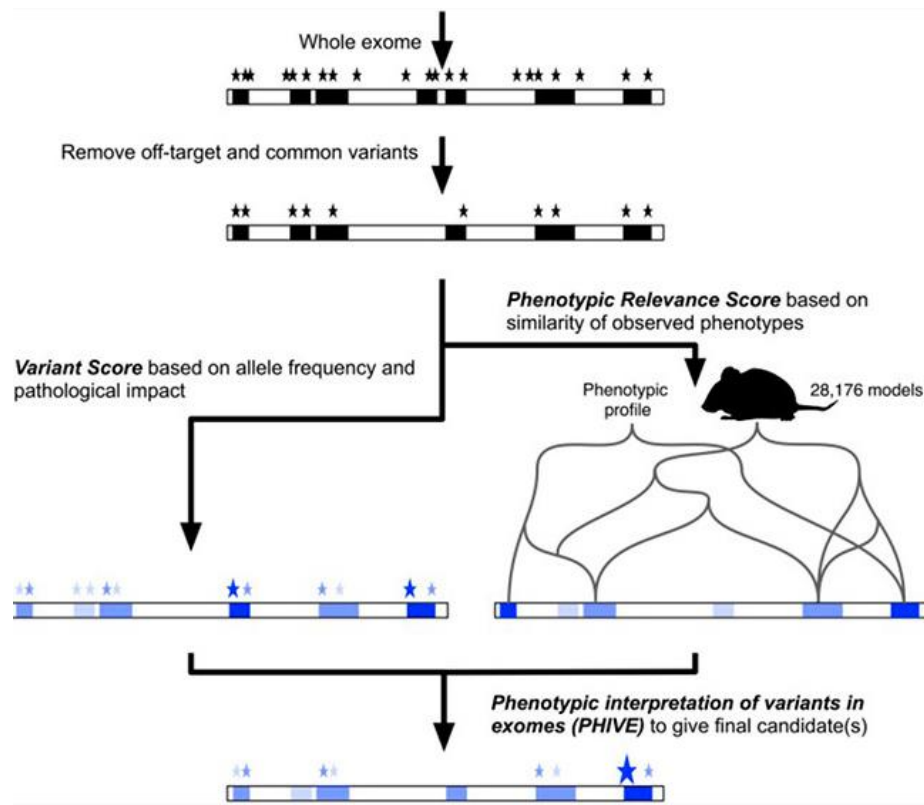
Phenotyp-driven Exome/Genome Analysis

Exomiser – filtered genes assigned a phenotypic relevance score based on comparison with mouse and zebrafish models with mutations in orthologous genes
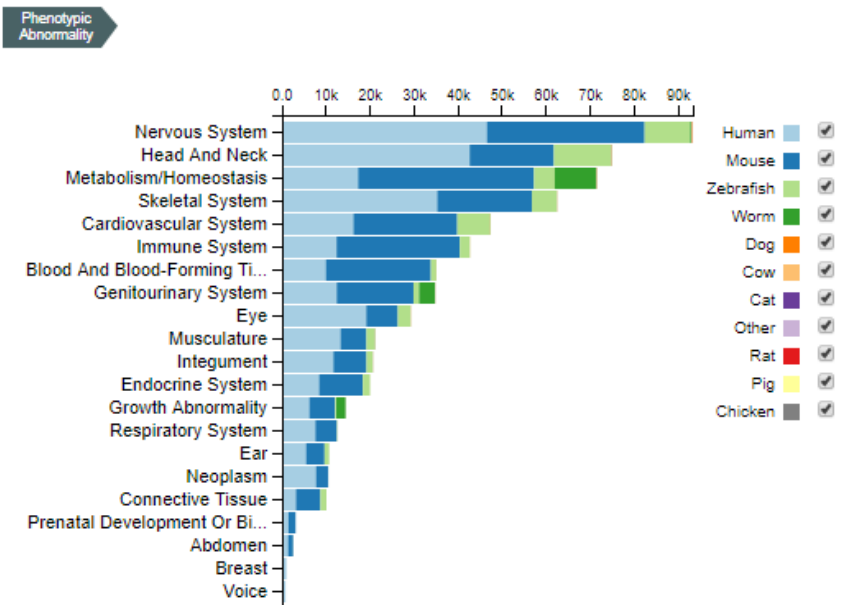

genes with variants that survive the initial filtering steps are then screened for mouse models with phenotypic to the human disease

Summary

HPO – a controlled vocabulary of phenotypic abnormalities for human genetics

- Freely available
- Open source

Novel approaches towards:

- Differential diagnosis tools (e.g. Phenomizer)
- Variant prioritization tools (e.g. Exomiser)
- Standard patient description in projects world-wide

Caveat: Phenotyp-driven variant prioritization

- Phenotypes in your patient (may) change over time
- Phenotypes are not always expressed / observed or described properly (two diseases)
- Phenotypes are not determined for (very) rare diseases (OMIM updates)
- Phenotypes are not very selective for e.g. mental retardation (Phenomizer; e.g. „intellectual disability" + „seizures" + „developmental delay")

... Challenges to (H)WPO

The famous Bavarian "Wolpertinger"