Deep Phenotyping for Translational Research and Precision Medicine

NIH Symposium: Linking Disease Model Phenotypes to Human Conditions

Peter Robinson

Charité Universitätsmedizin Berlin

September 10–11, 2015

Thanks!


Matthew Brush


Nathan Dunn


Melissa Haendel


Harry Hochheiser


Sebastian Köhler


Suzanna Lewis


Julie McMurry


Christopher Mungall

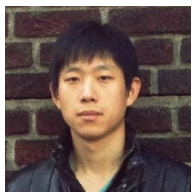

Peter Robinson


Kent Shefchek


Damian Smedley


Nicole Vasilevsky


Nicole Washington

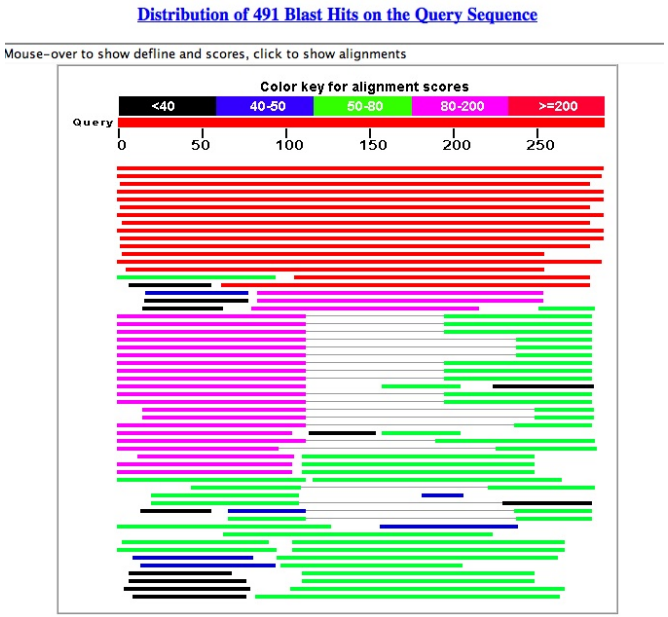

Zhou Yuan

http://monarchinitiative.org

Plan

Bioinformatics

- Since the beginnings of the field of Bioinformatics in the 1960s, a central theme has been the development of algorithms that calculate similarity scores between biological entities and use them to rank lists

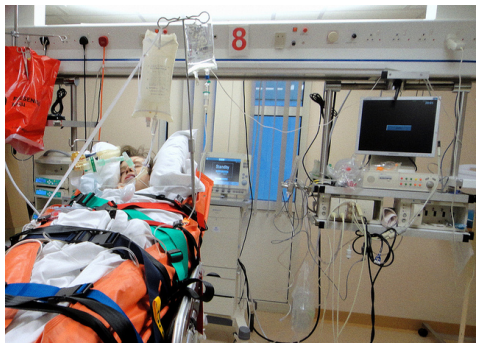


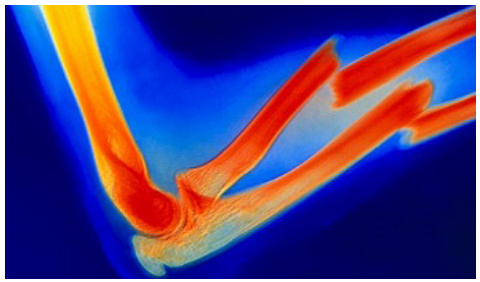
BLAST: Find and rank homologous sequences

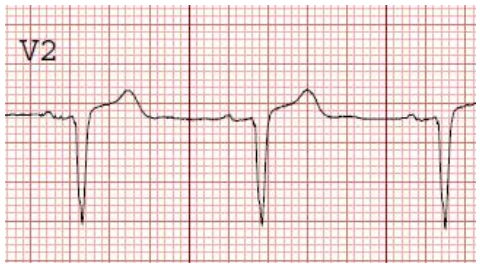
Margaret Dayhoff, originator of PAM matrices

Bioinformatics for medicine?

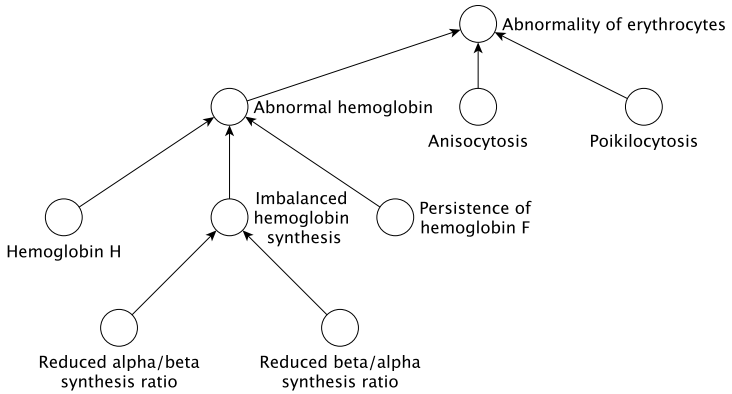
But how exactly do we calculate the similarity between diseases, symptoms, patients,. . .?







The Human Phenotype Ontology



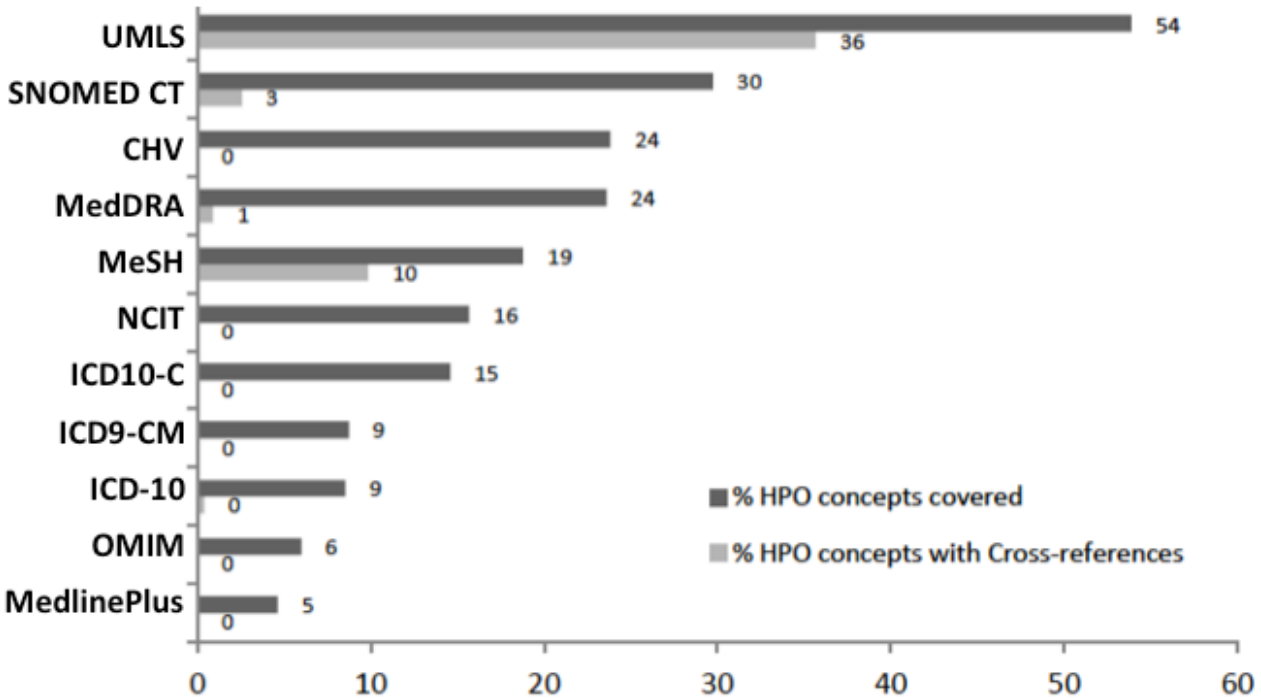
- 11,030 terms
- 117,348 annotations for $\sim$ 7000 mainly monogenic diseases
- http://www.human-phenotype-ontology.org

- Widely used in rare disease community: UK 100,000 genomes; NIH Undiagnosed Diseases Network; DDD/DECIPHER, GA4GH, etc.
- Applications:
  - linking human diseases to animal models
  - inferring novel drug interactions
  - prioritizing gene-disease targets
  - describing rare clinical disorders

- Interoperable with *model organism data* and *basic research standards*
- A *computable* representation of human disease

Why HPO?

- Substantially better coverage of phenotype concepts than any other terminology



Winnenburg and Bodenreider,

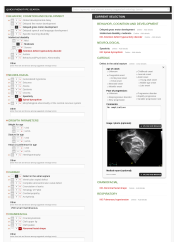
ISMB PhenoDay, 2014

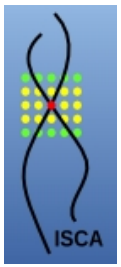
Widely used in the community

Databases & Bioinformatics Resources Using HPO

DECIPHER (Sanger Institute)
DDD (Sanger Institute)
ECARUCA
FORGE (Genome Canada)
GWAS Central
IRDiRC
ISCA
NCBI Genetic Testing Registry
NIH Undiagnosed diseases program
UK 100,000 Genomes Program
UMLS
Phenotips (Brudno Group, U Toronto)
...

Major credits go to OMIM and Orphanet





OMIM

orphanet

D966–D974 Nucleic Acids Research, 2014, Vol. 42, Database issue Published online 11 November 2013
doi:10.1093/nar/gkt1026

The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data

Sebastian Köhler[1,2,*], Sandra C. Doelken[1], Christopher J. Mungall[3], Sebastian Bauer[1], Helen V. Firth[4,5], Isabelle Bailleul-Forestier[6], Graeme C. M. Black[7,8], Danielle L. Brown[9], Michael Brudno[10,11], Jennifer Campbell[9,12], David R. FitzPatrick[13], Janan T. Eppig[14], Andrew P. Jackson[13], Kathleen Freson[15], Marta Girdea[10,11], Ingo Helbig[16], Jane A. Hurst[17], Johanna Jähn[16], Laird G. Jackson[18], Anne M. Kelly[19], David H. Ledbetter[20], Sahar Mansour[21], Christa L. Martin[20], Celia Moss[22], Andrew Mumford[23], Willem H. Ouwehand[4,19], Soo-Mi Park[5], Erin Rooney Riggs[20], Richard H. Scott[24], Sanjay Sisodiya[25], Steven Van Vooren[26], Ronald J. Wapner[27], Andrew O. M. Wilkie[28], Caroline F. Wright[4], Anneke T. Vulto-van Silfhout[29], Nicole de Leeuw[29], Bert B. A. de Vries[29], Nicole L. Washingthon[3], Cynthia L. Smith[14], Monte Westerfield[30], Paul Schofield[14,31], Barbara J. Ruef[30], Georgios V. Gkoutos[32], Melissa Haendel[33], Damian Smedley[4], Suzanna E. Lewis[3] and Peter N. Robinson[1,2,34,*]

[1]Institute for Medical Genetics and Human Genetics, Charité-Universitätsmedizin Berlin, Augustenburger Platz 1, 13353 Berlin, Germany, [2]Berlin-Brandenburg Center for Regenerative Therapies, Charité-Universitätsmedizin Berlin, Augustenburger Platz 1, 13353 Berlin, Germany, [3]Lawrence Berkeley National Laboratory, Mail Stop 84R0171, Berkeley, CA 94720, USA, [4]The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire CB10 1SA, UK, [5]Department of Medical Genetics, Cambridge University Addenbrooke's Hospital, Cambridge CB2 2QQ, UK, [6]Université Paul Sabatier, Faculté de Chirurgie Dentaire, CHU Toulouse, France, [7]Centre for Genomic Medicine, Central Manchester University Hospitals NHS Foundation Trust, Manchester Academic Health Sciences Centre (MAHSC), Manchester, UK, [8]Centre for Genomic Medicine, Institute of Human Development, Faculty of Medical and Human Sciences, University of Manchester, MAHSC, Manchester M13 9WL, UK, [9]Institute of Genetic Medicine. Newcastle University, Central Parkway, Newcastle upon Tyne, NE1 3BZ, UK, [10]Department of Computer Science, University of Toronto, Ontario, Canada, [11]Centre for Computational Medicine, Hospital for Sick Children, Toronto, Ontario, Canada, [12]Department of Clinical Genetics, Leeds Teaching Hospitals NHS Trust, Leeds LS2 9NS, UK, [13]MRC Human Genetics Unit, MRC Institute of Genetic and Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, UK, [14]The Jackson Laboratory, Bar Harbor, ME 04609, USA, [15]Center for Molecular and Vascular Biology, University of Leuven, Belgium, [16]Department of Neuropediatrics, University Medical Center Schleswig-Holstein, Kiel Campus, 24105 Kiel, Germany, [17]NE Thames Genetics Service, Great Ormond Street Hospital, London WC1N 3JH, UK, [18]Drexel University College of Medicine, Philadelphia, PA 19102, USA, [19]Department of Haematology, University of Cambridge and NHS Blood and Transplant Cambridge, CB2 0PT Cambridge, UK, [20]Autism and Developmental Medicine Institute, Geisinger Health System, Danville, PA 17822, USA, [21]SW Thames Regional Genetics Service, St George's Healthcare NHS Trust, Tooting, London SW17 0RE, UK, [22]Department of Dermatology, Birmingham Children's Hospital, Birmingham, UK, [23]Bristol Heart Institute, University of Bristol, Bristol, UK, [24]Department of Clinical Genetics, Great Ormond Street Hospital, London and Clinical and Molecular Genetics Unit, UCL Institute of Child Health, London, UK, [25]Department of Clinical and Experimental Epilepsy, UCL Institute of Neurology, London, UK, [26]Cartagenia, Leuven, Belgium, [27]Department of Obstetrics and Gynecology, Columbia University Medical Center, New York, NY 10032, USA, [28]Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DS, UK, [29]Department of Human Genetics, Radboud University Medical Centre, 6500 HB Nijmegen, The Netherlands, [30]ZFIN, University of Oregon, Eugene, OR 97403-5291, USA, [31]Department of Physiology, Development and Neuroscience, Downing Street, Cambridge CB2 3EG, UK, [32]Department of Computer Science, Aberystwyth

Plan

What's The Problem?

- Phenotypic descriptions that are very evocative for humans but meaningless for computers:

  myopathic electromyography
  still walking 25 years after onset

- The following descriptions mean the same thing to you: "`generalized amyotrophy`", "`generalized muscle atrophy`", "`muscular atrophy, generalized`" (etc)[1]
- Many publications have little[2] information about the actual phenotypic features seen in patients with particular mutations
- Databases cannot talk to one another about phenotypes

A tale of two footballs

A football . . .

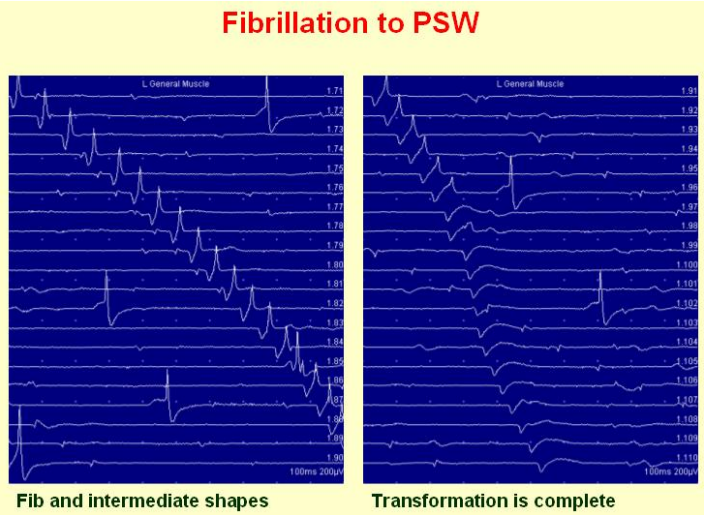
A football . . .

American Football = Football ≠ Football = European Football = Soccer

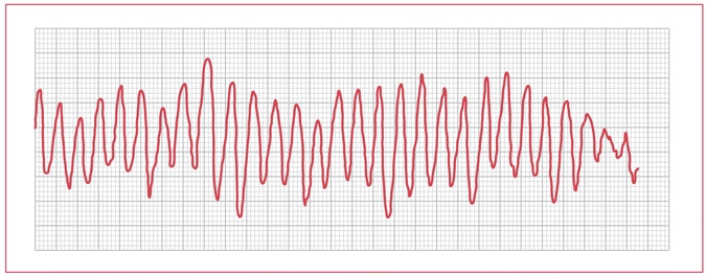
When you see "football", your computer sees:

01000110011011110110111101110100011000100110000101101100011011 00

A tale of two fibrillations



fibrillation . . .



fibrillation . . .

muscle fibrillation = fibrillation = fibrillation = ventricular fibrillation

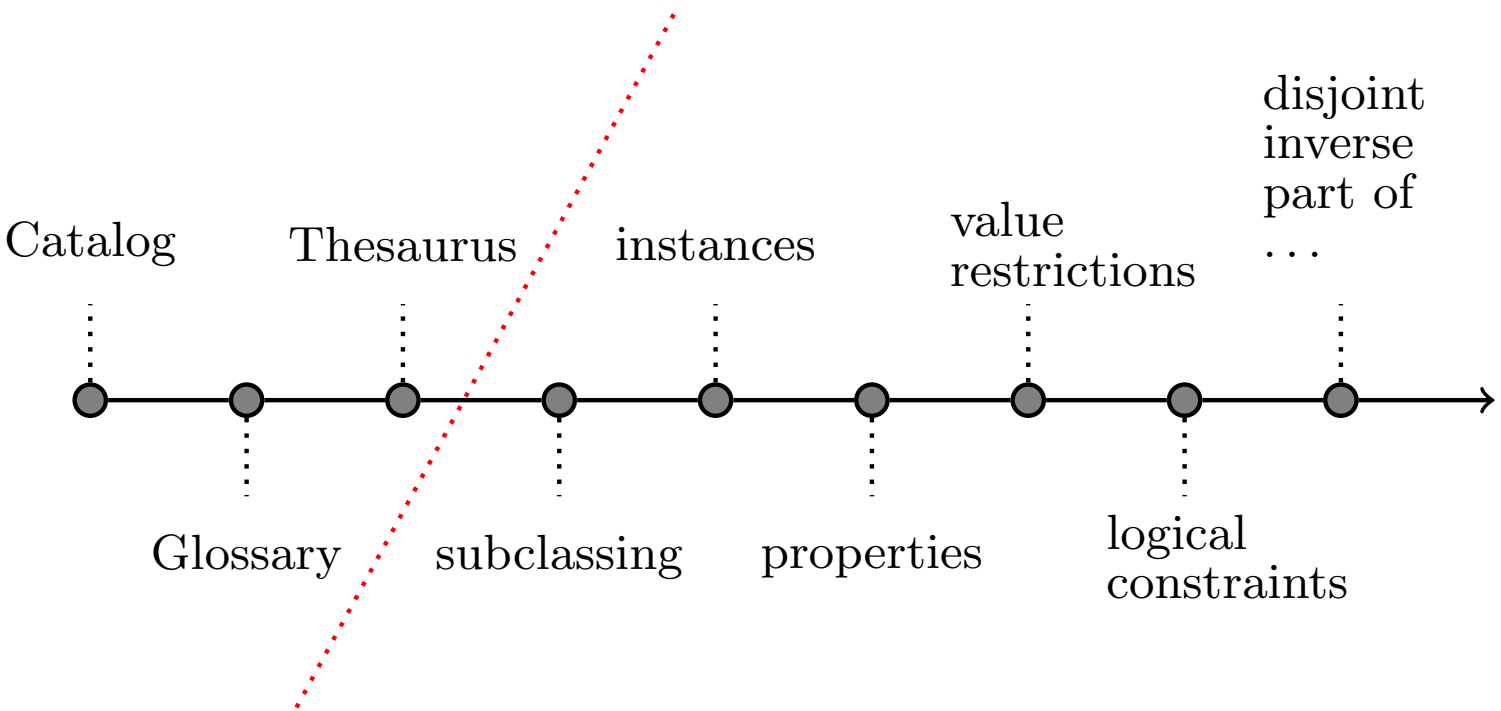
When you see “fibrillation”, your computer sees:

01100110011010010110001001110010011010010110110001101100011000010111010001101001

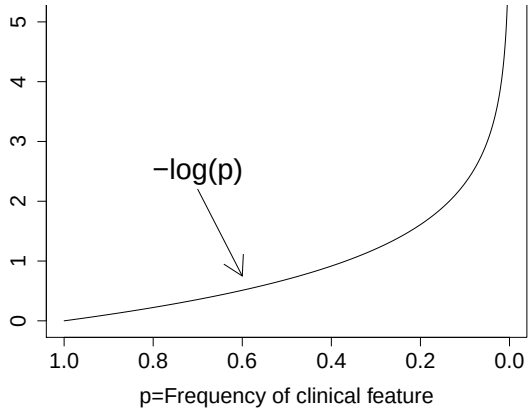
0110111101101110

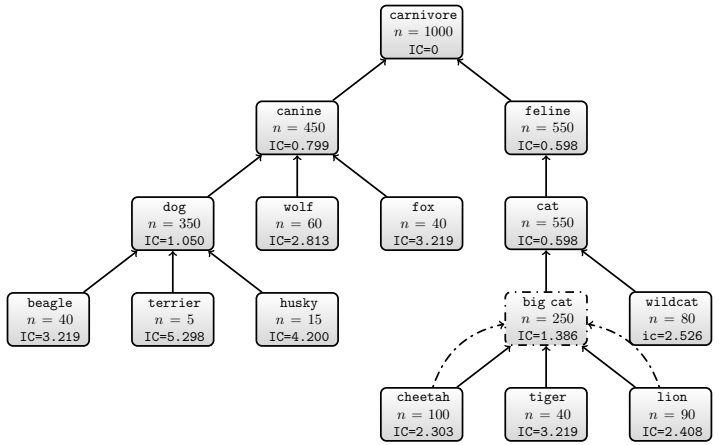
What is an Ontology?

"An ontology is a specification of a conceptualization."
Tom Gruber, 1993



Information content

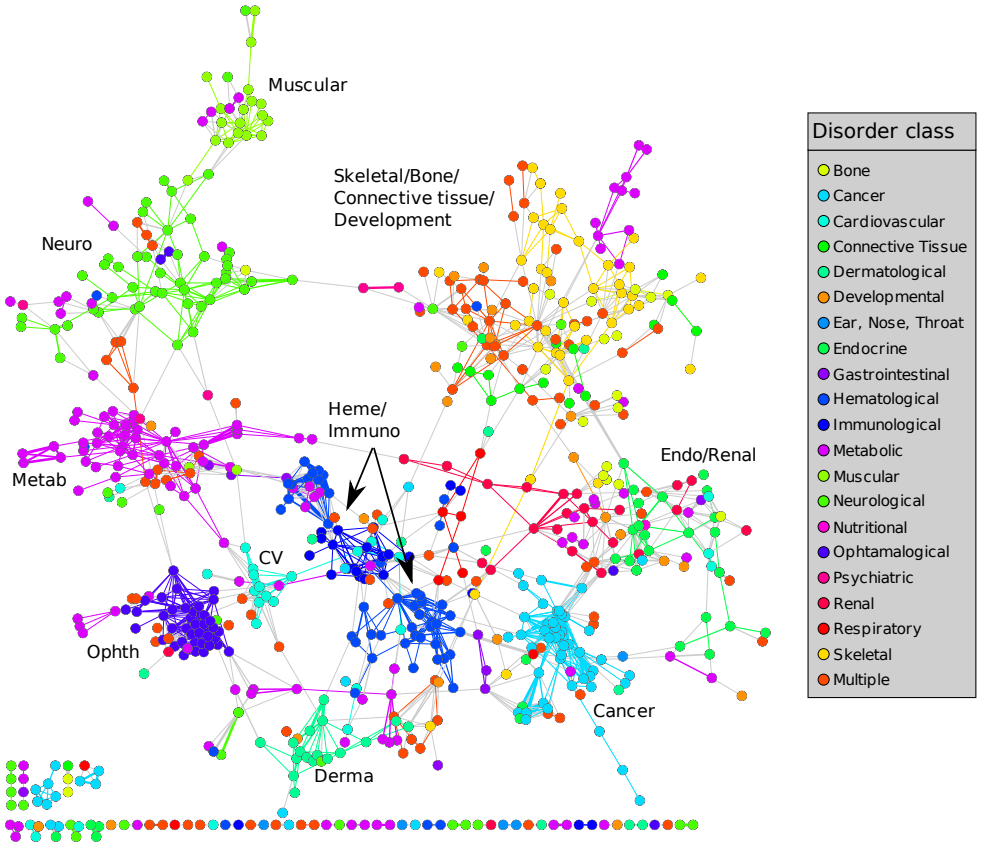




$$\mathrm{IC}(t) = -\log p(t),$$

- Information content of common ancestor: Similarity between ontology terms
- Average similarity between terms can be used to compare two diseases

The Human Phenome: Network of Human Diseases and Disease Genes



$$\text{sim}(d_1, d_2) = 0.5 \cdot \text{avg}\left[\sum_{s \in d_1} \max_{t \in d_2} \text{sim}(s,t)\right] + 0.5 \cdot \text{avg}\left[\sum_{s \in d_2} \max_{t \in d_1} \text{sim}(s,t)\right]$$

FOL: Klingon Opera

$$(\forall x)(\texttt{Klingon}(x) \Rightarrow \texttt{OperaLover}(x))$$

$$\texttt{Klingon}(Worf)$$

- If $\texttt{Klingon}(Worf)$ is true, we can infer that Worf is an opera lover.

$$\texttt{OperaLover}(Worf)$$

- Analogous algorithms are the basis for human $\Leftrightarrow$ model organism comparisons

What is a phenotype ontology?

Precise language (and thinking), interoperability, improved database models to reliably capture and interpret phenotype information.

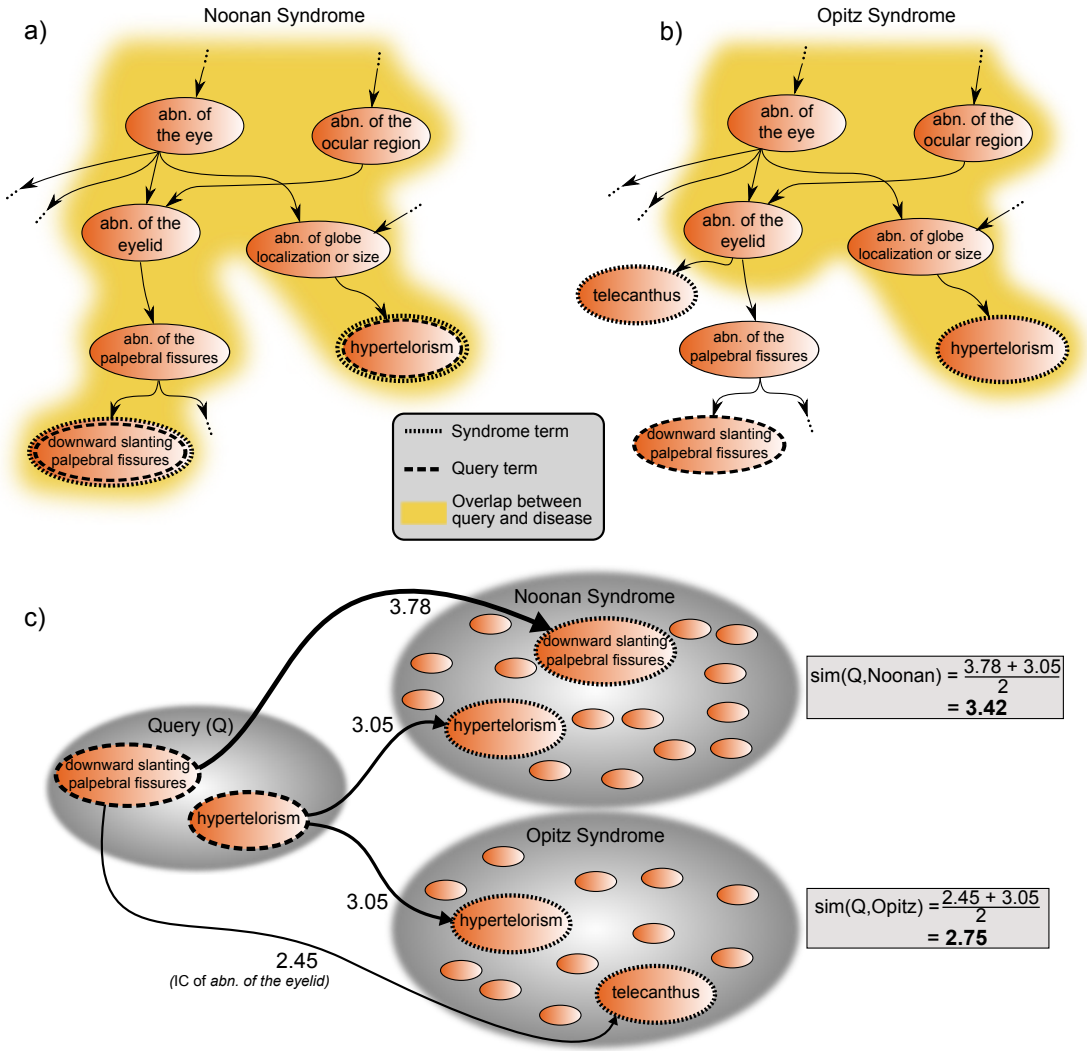
- A *medical* phenotype ontology describes the individual manifestations of diseases:
  1. signs
  2. symptoms
  3. laboratory findings
  4. imaging studies
  5. etc.
- Deep phenotype: The precise and comprehensive analysis of phenotypic abnormalities
- Individual components of disease rather than "gestalt"

Robinson PN, Webber C (2014) Phenotype ontologies and cross-species analysis for translational research. *PLoS Genet* **10**:e1004268.

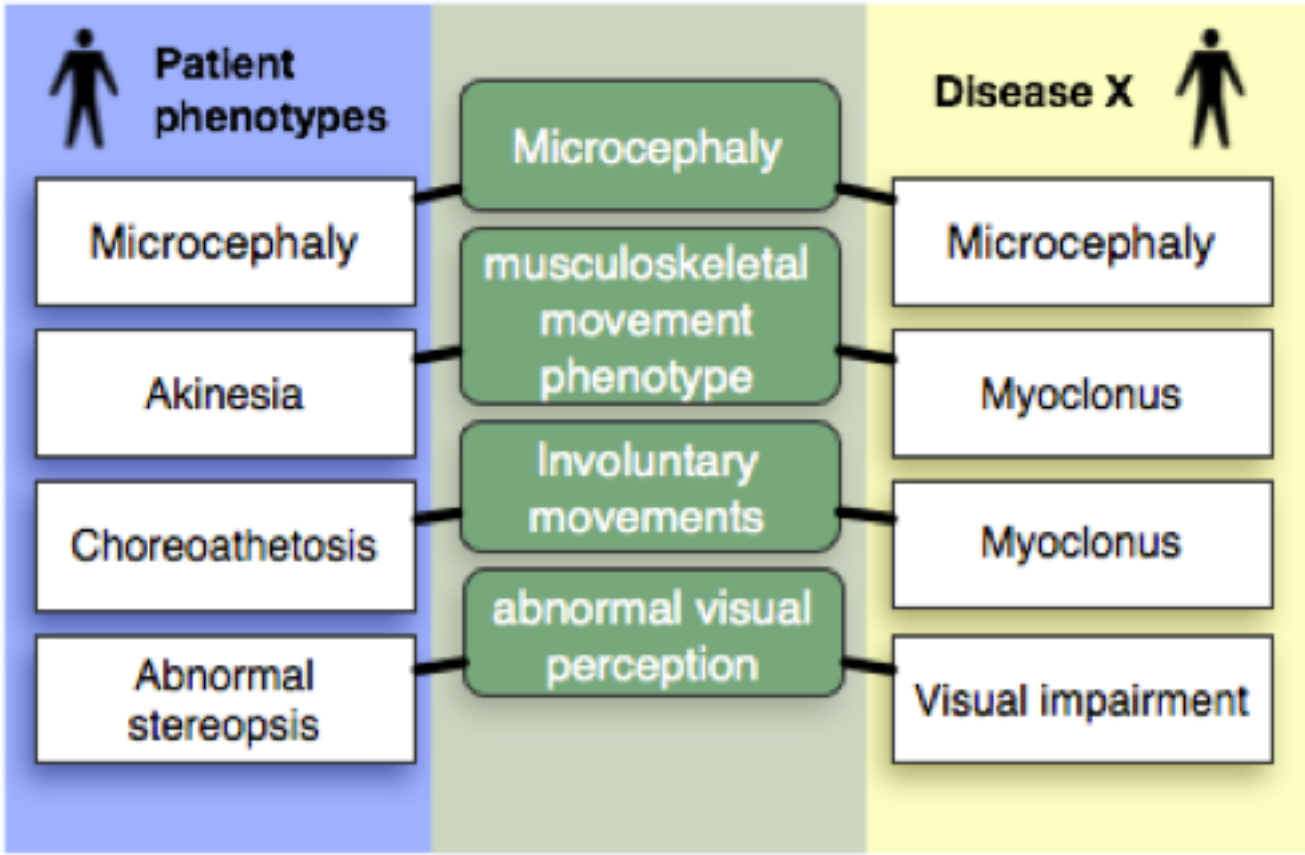
PN Robinson (2012) Deep phenotyping for precision medicine. *Hum Mutat* **33**: 777–780 (Special Issue of *Human Mutation* on Deep Phenotyping)

Plan

Ontological diagnostics

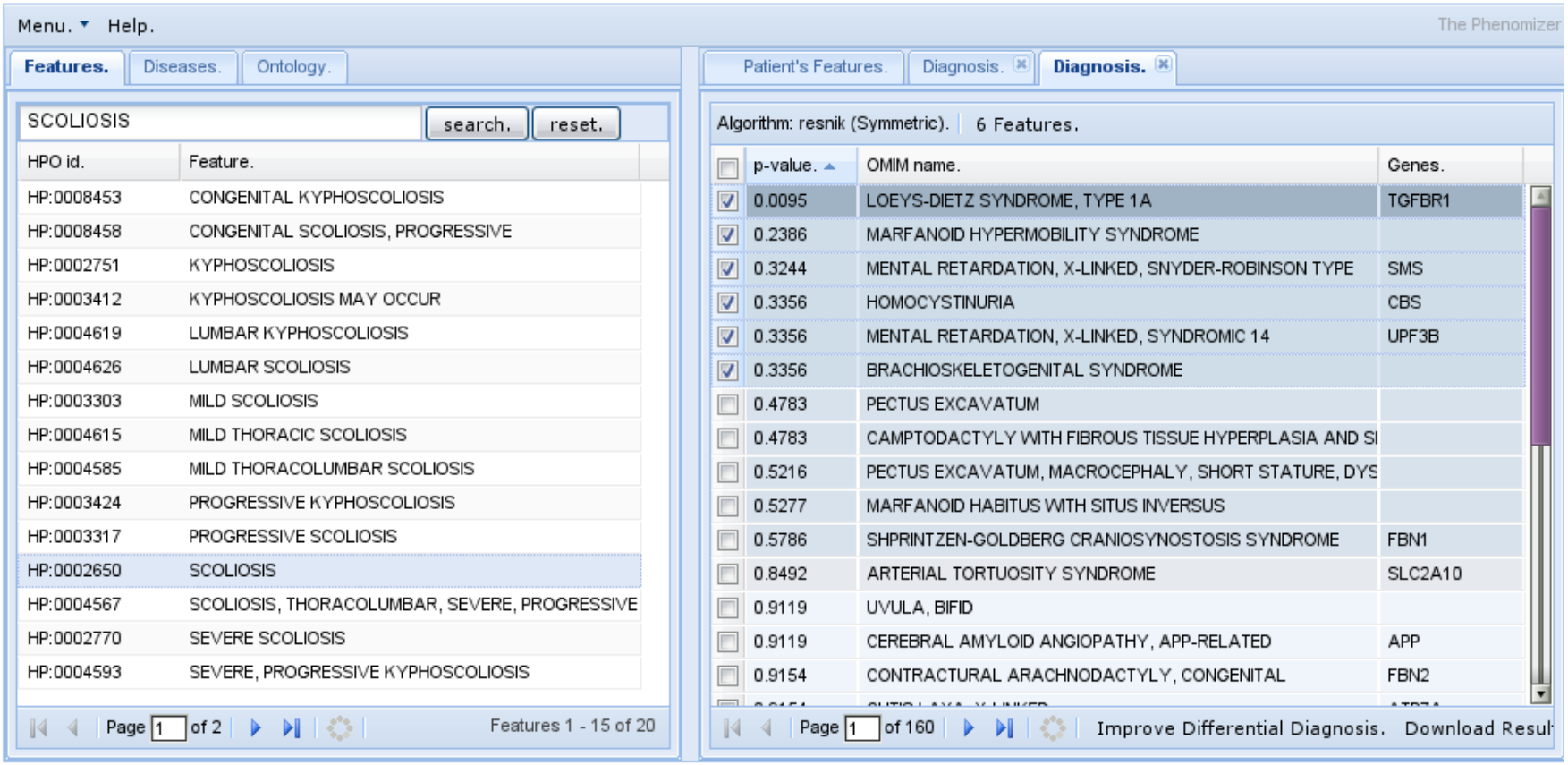


$$\text{sim}(Q \rightarrow d) \quad = \quad \text{avg}\left[\sum_{s \in Q} \max_{t \in d} \text{sim}(s,t)\right]$$



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

The Phenomizer

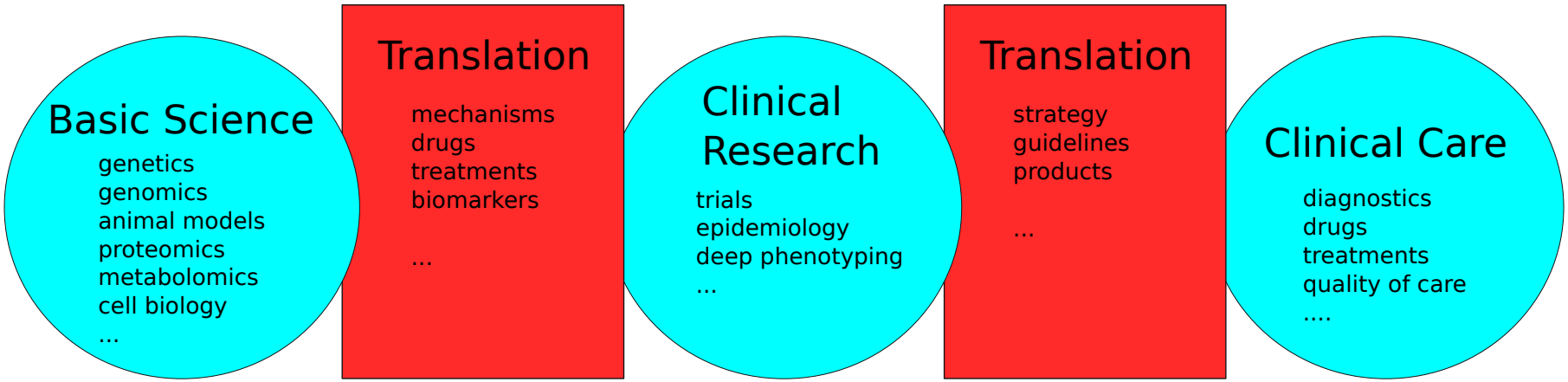


- Sebastian Köhler et al. (2009) Clinical Diagnostics with Semantic Similarity Searches in Ontologies. *Am J Hum Genet*, **85**:457–64.

http://compbio.charite.de/Phenomizer

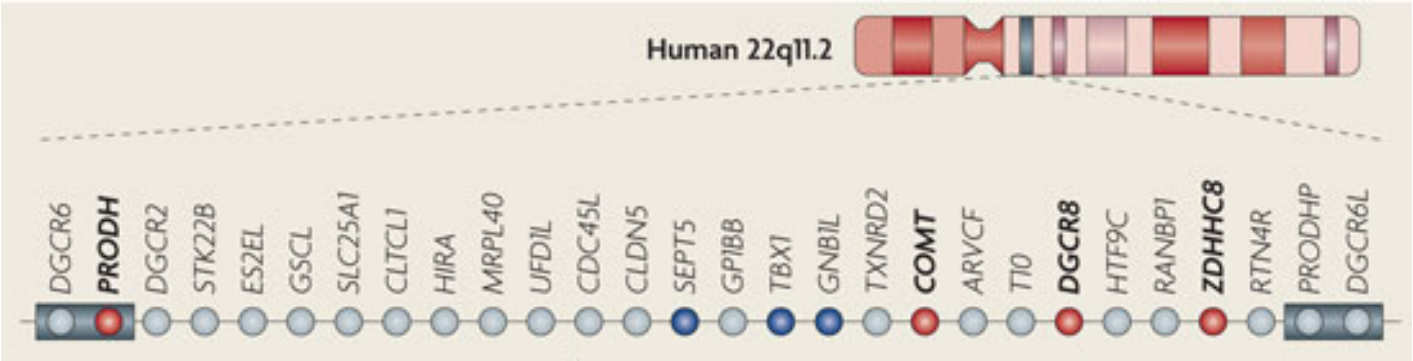
Plan

HPO for translational research



- The HPO is a sophisticated computational resource that can be used to link data from anatomy, histology, pathology, gene function, model organisms, etc., in order to perform integrative computationally driven translational research

Traditional view of CNV pathogenesis

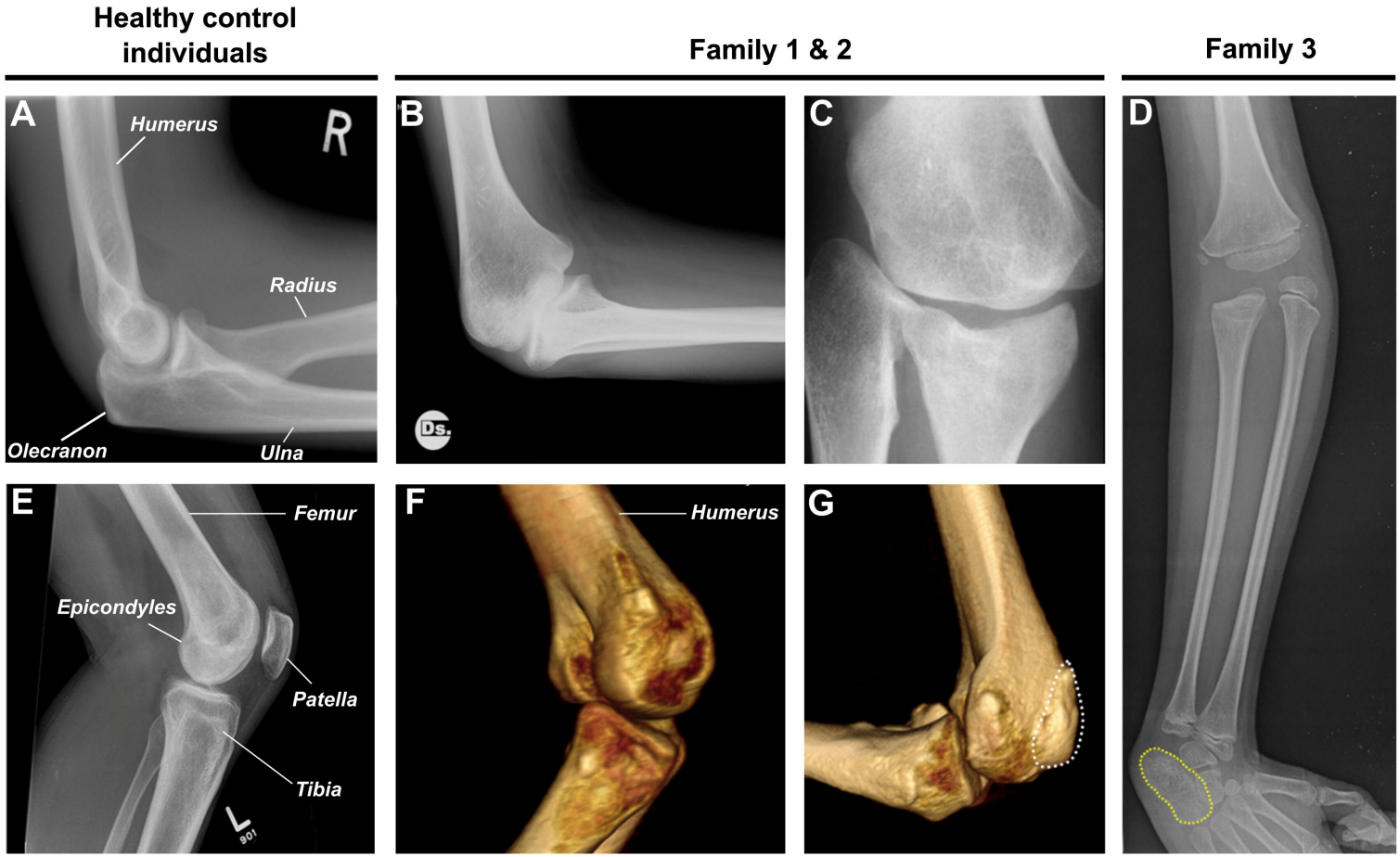


- Phenotype results from dosage effects of one or more affected genes
- Diagnostic problem: Distinguish pathogenic from neutral CNVs
- Scientific and medical problem: Decide which genes are responsible for the phenotypic features?

For instance, haploinsufficiency of *TBX1* contributes to the heart defects seen in 22q11.2 deletion syndrome

Liebenberg syndrome

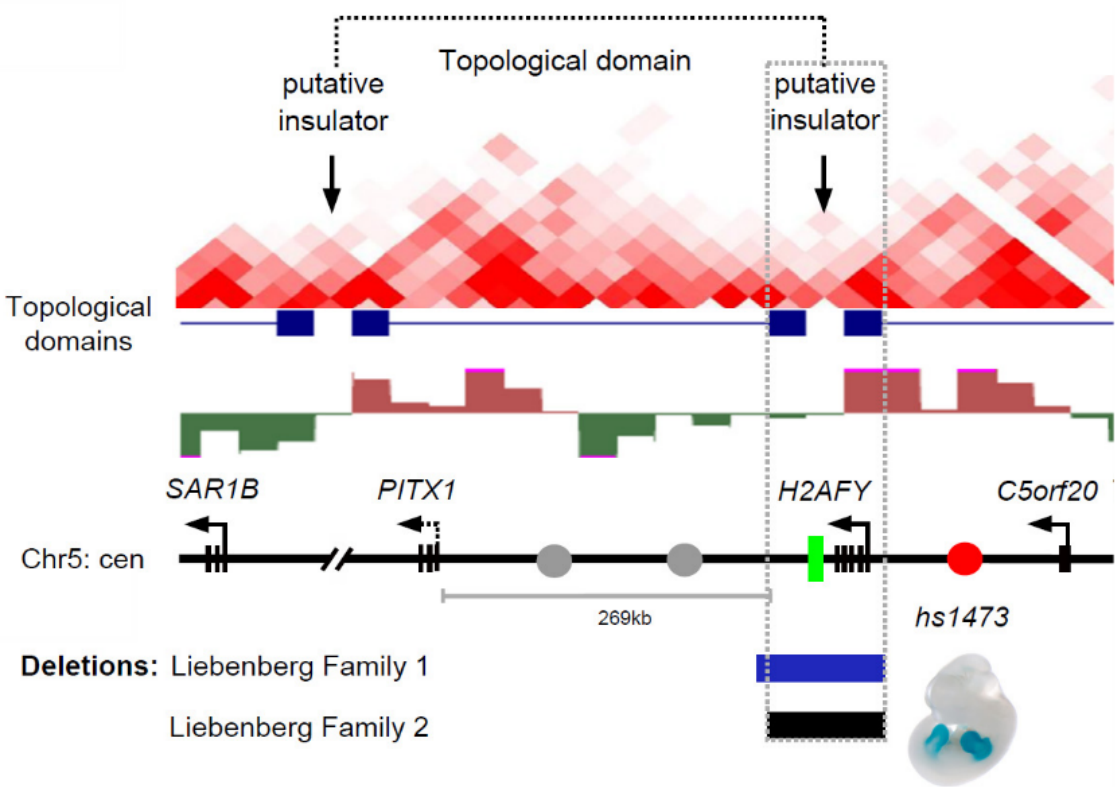
- Recent observations challenge the gene-dosage model



Spielmann et al., (2012) Homeotic Arm-to-Leg Transformation Associated with Genomic Rearrangements at the PITX1 Locus. *Am J Hum Genet* **91**:629–635

Long Range Control of Gene Expression

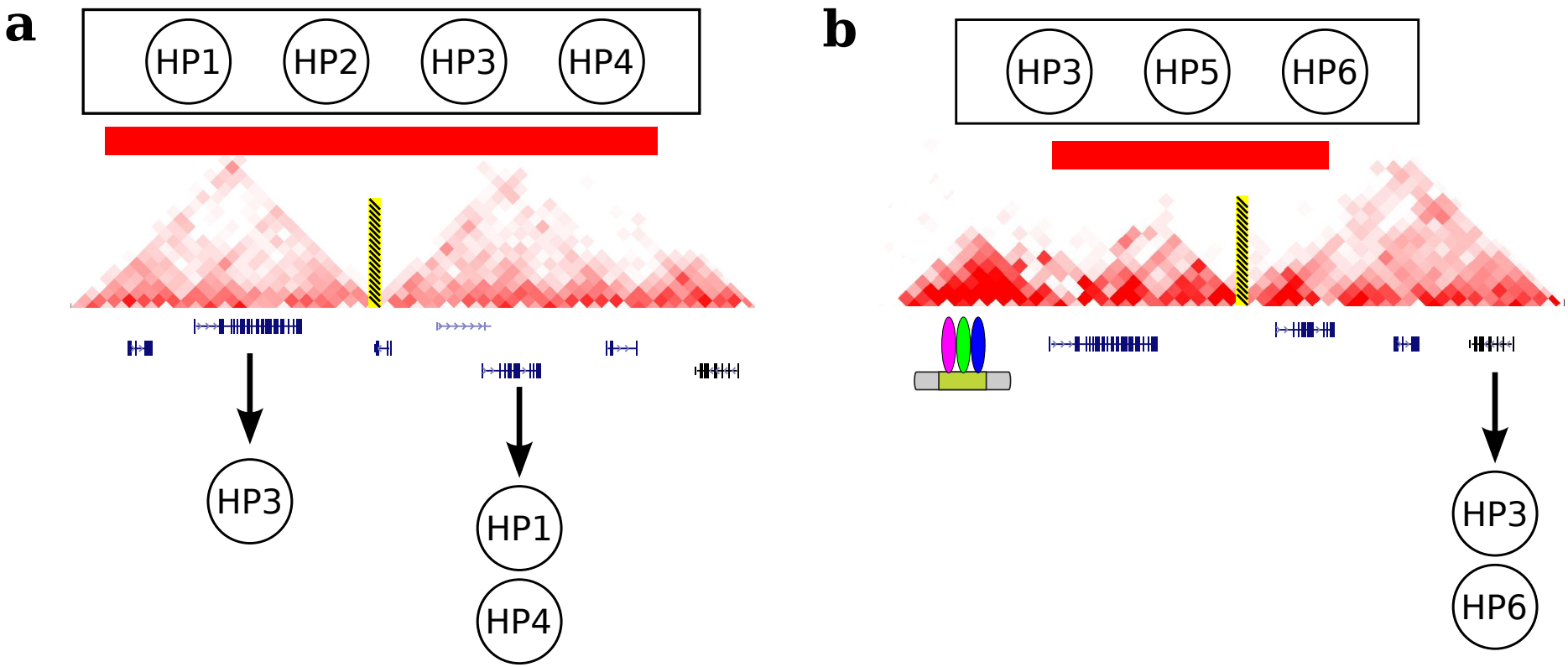
- Re-examination of the Liebenberg deletion indicated deletion of a topological domain barrier between a forelimb enhancer and a PITX1
- . . . leading to ectopic gene expression ("enhancer adoption")



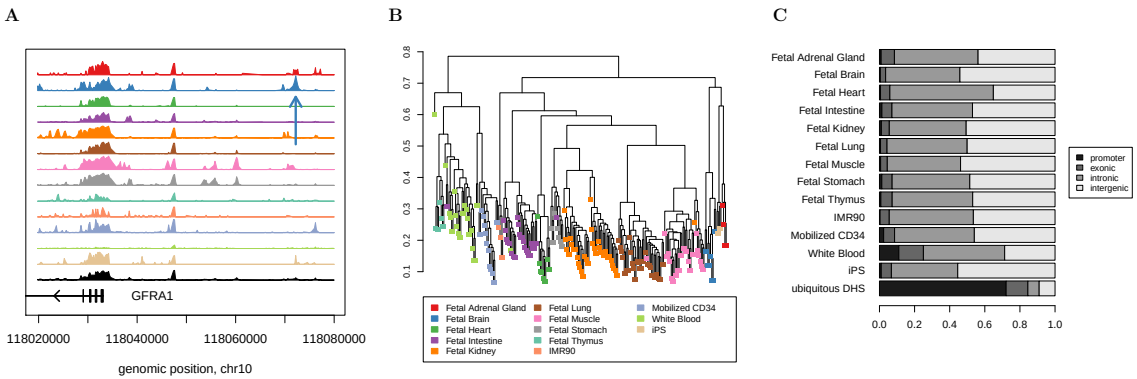
Spielmann M, Mundlos S (2013) *Bioessays* **35**:533-43.

Gene-Dosage vs. Enhancer adoption

- Therefore, we decided to address the question of how common haploinsufficiency (gene-dosage; **a**) and TDB disruption (**b**) are amongst CNVs associated with congenital disease



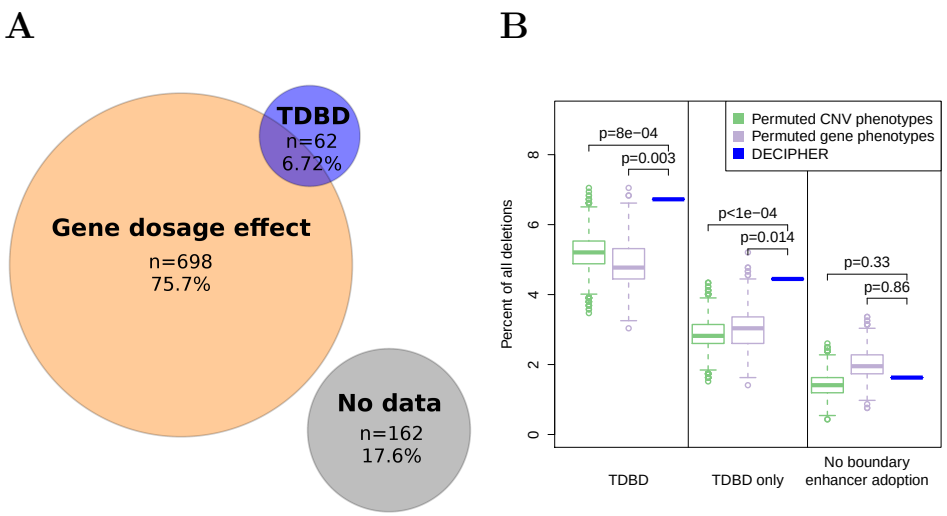
Tissue-Specific Enhancers and Phenotypic categories



- Our genome: $>$ 400,000 enhancers
- We identified cell type specific enhancers as DNase I hypersensitive sites (DHS)
- Assigned tissue-specific enhancers to HPO terms

Tissue	HPO term name	Term ID	Descendant HPO terms	Genes	Cases
Adrenal	Abnormality of the adrenal glands	HP:0000834	65	75	2 (0.217%)
Brain	Abnormality of the forebrain	HP:0100547	213	640	276 (29.9%)
Heart	Abnormality of the heart	HP:0001627	273	491	236 (25.6%)
Intestine	Abnormality of the intestine	HP:0002242	121	260	17 (1.84%)
Kidney	Abnormality of the kidney	HP:0000077	184	383	77 (8.35%)
Lung	Abnormality of the lung	HP:0002088	149	529	9 (0.976%)
Muscle	Abnormality of the musculature	HP:0003011	667	1079	291 (31.6%)
Stomach	Abnormality of the stomach	HP:0002577	24	116	10 (1.08%)
Thymus	Abnormality of the thymus	HP:0000777	9	26	0 (0.0%)
WBC	Abnormality of leukocytes	HP:0001881	195	256	4 (0.434%)

How common is enhancer adoption?



- 6.7% of 922 DECIPHER deletion cases potentially related to TDBD
- Even higher rate of TDBD predicted by analysis that includes ontologically mapped model organism phenotype data (mouse, fish): 11.9%
- **No difference** if we simulate deletions so as not to disrupt TDBs!

Ibn-Salem J, Köhler S, (13 coauthors), Spielmann M, Robinson PN (2014) Deletions of chromosomal regulatory boundaries are associated with congenital disease. *Genome Biology* **15**:423.

Plan

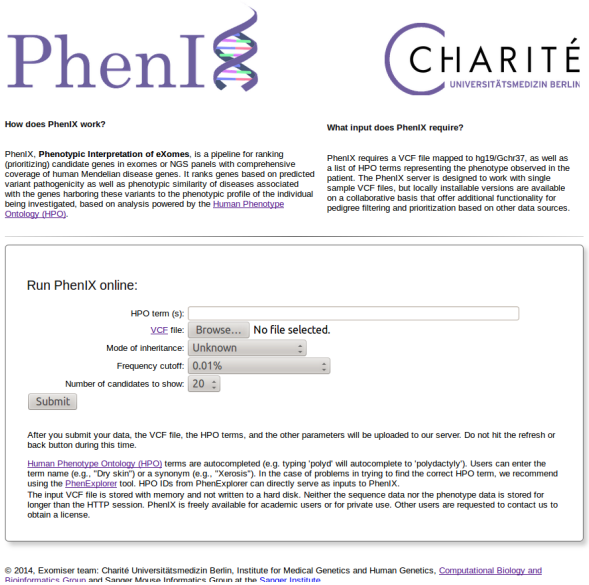
Clinical Diagnostics in Medical Genetics

Getting a precise diagnosis for individuals with rare disease can be difficult

- Roughly 7,000 Mendelian diseases. Although these diseases are individually rare, up to 8% of the population is affected by a specific genetic disorder
- Diagnosis useful for avoiding unnecessary investigations, exact prognosis, personalised clinical management, recurrence risk, "closure" and reduction of feelings of guilt
- Diagnostic rate

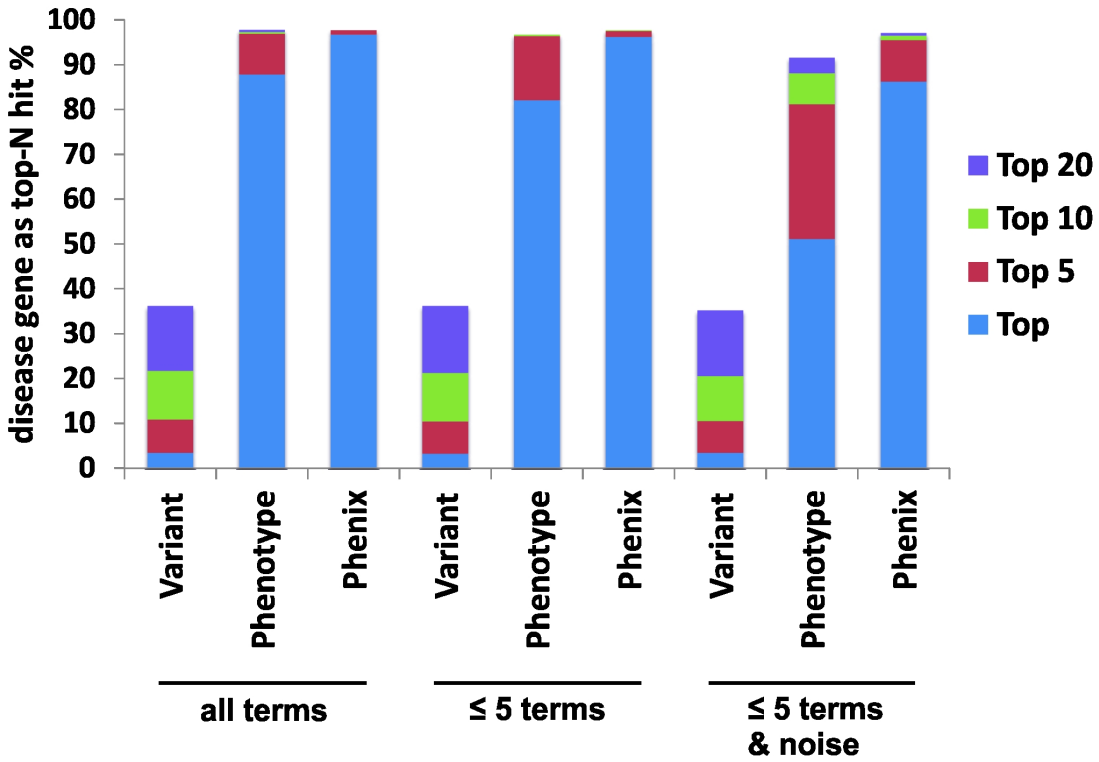
  Cytogenetics: $\sim$ 4%
  Array-CGH: $\sim$ 10–15 %
  Targeted Sanger sequencing: $\sim$ 5–95% (depending on indication)
  Total with traditional workup: **Less than 50% overall**[3]

PhenIX



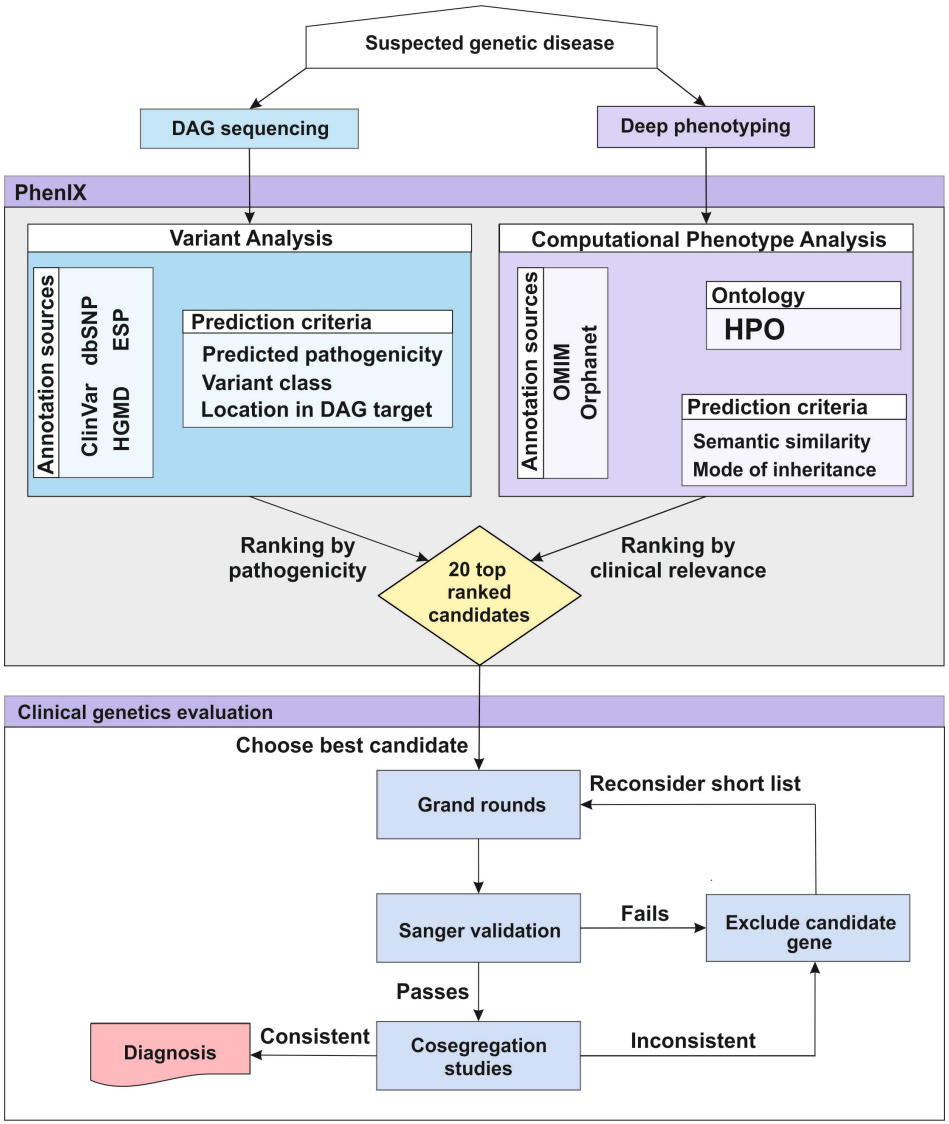
- Basic algorithm: (i) Identify predicted pathogenic mutations in (clinical) exome (typically up to 100); (ii) rank the corresponding genes according to phenotypic relevance with the Phenomizer
- PhenIX Phenotypic Interpretation of eXomes
- Designed for **clinical diagnostics** with exome or DAG panel

Phenotypic Interpretation of eXomes: PhenIX



- In 10,000 simulations with mutations from HGMD and corresponding HPO terms, the correct gene was ranked in first place in over 86% of cases

PhenIX: Workflow



- Clinical/Bioinformatic workflow in ca. 60 min

PhenIX: Prospective Validation

ID	Age, Sex	Presentation	Gene	Rank	Diagnosis
P1	3y (f)	Intellectual disability + complex phenotype	MLL	2	Wiedemann Steiner syndrome
P2	5y (f)	Intellectual disability + complex phenotype	SYNGAP1	4	Mental retardation, MRD5 (41)
P3	6y (f)	Skeletal phenotype	FGFR2	1	Pfeiffer syndrome
P4	d. 5.5m (f)	Complex phenotype without intellectual disability	SH3PXD2B	6	Frank-ter Haar syndrome
P5	6m (f)	Intellectual disability + neurological abnormalities	SLC6A3	1	Parkinsonism-dystonia
P6	Fetus	Skeletal phenotype	ALPL	2	Infantile hypophosphatasia
P7	7y (m)	Eye phenotype	NHS	2	Nance-Horan Syndrome
P8	14y (m)	Intellectual disability + complex phenotype	MLL	1	Wiedemann-Steiner syndrome
P9	6y (f)	Intellectual disability + complex phenotype	DYRK1A	4	Mental retardation, MRD7
P10	1.5–7y	Intellectual disability + complex phenotype	MCOLN1	1	Type IV mucolipidosis
P11	3y (m)	Intellectual disability + complex phenotype	RBM10	3	TARP syndrome

- Diagnosis rate was 100% in 52 retrospective “solved” cases and 28% in the 40 “unknown” cases

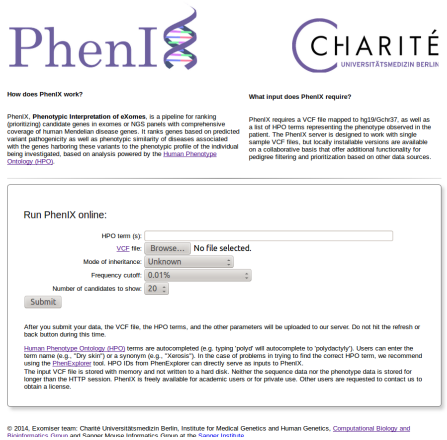
PhenIX

RESEARCH ARTICLE

GENETIC DIAGNOSIS

Effective diagnosis of genetic disease by computational phenotype analysis of the disease-associated genome

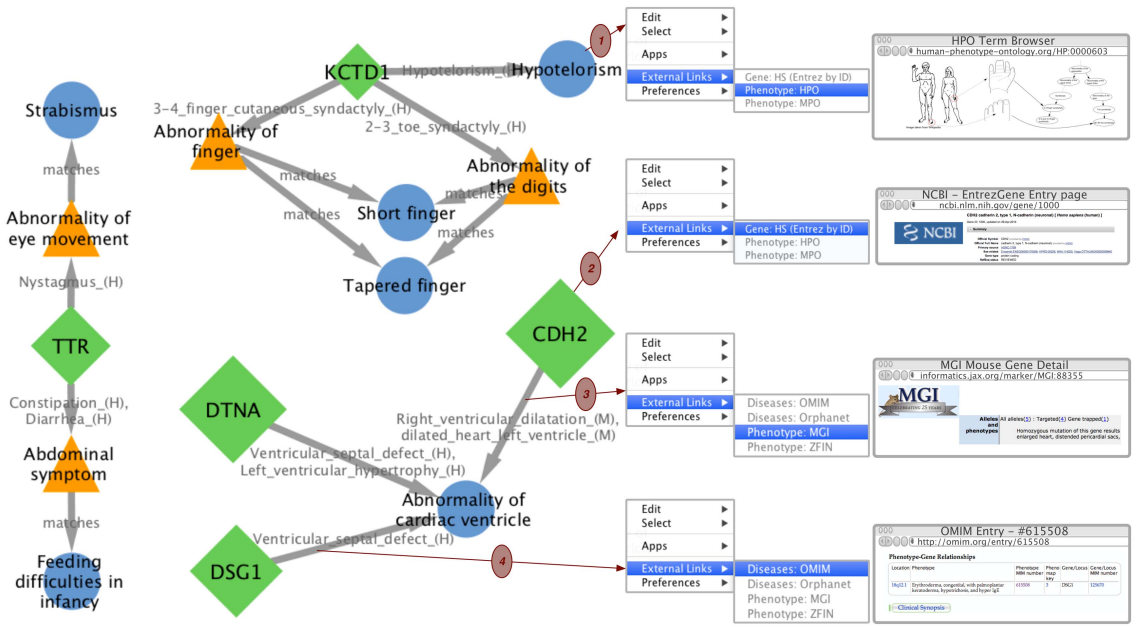
Tomasz Zemojtel,[1,2,3]* Sebastian Köhler,[1]* Luisa Mackenroth,[1]* Marten Jäger,[1] Jochen Hecht,[4,5] Peter Krawitz,[1,4] Luitgard Graul-Neumann,[1] Sandra Doelken,[1] Nadja Ehmke,[1] Malte Spielmann,[1,4] Nancy Christine Øien,[1,6] Michal R. Schweiger,[1,4,7] Ulrike Krüger,[1] Götz Frommer,[8] Björn Fischer,[1,4] Uwe Kornak,[1,4] Ricarda Flöttmann,[1] Amin Ardeshirdavani,[9] Yves Moreau,[9] Suzanna E. Lewis,[10] Melissa Haendel,[11] Damian Smedley,[12] Denise Horn,[1] Stefan Mundlos,[1,4,5] Peter N. Robinson[1,4,5,13†]



- http://compbio.charite.de/phenix
- Zemojtel T et al (2014) Effective diagnosis of genetic disease by computational phenotype analysis of the disease-associated genome. *Science Translational Medicine* **6**:252ra123

The value of models

- We remain largely ignorant of the genetic basis of human disease
- Large scale mouse and zebrafish programs are providing phenotype data for many thousands of genes for which no human disease is currently known
- A number of talks in this Symposium!

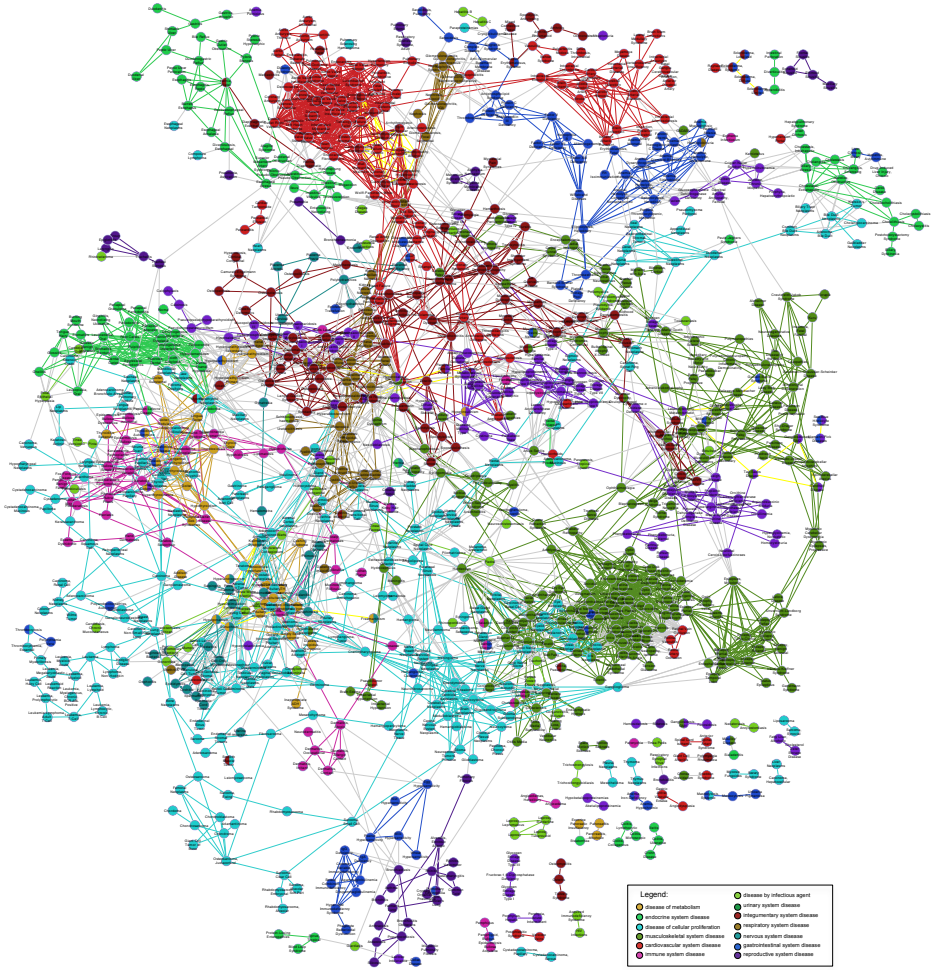


Köhler S, Schoeneberg U, et al (2014) Clinical interpretation of CNVs with cross-species phenotype data.

J Med Genet **51**:766-72.

Plan

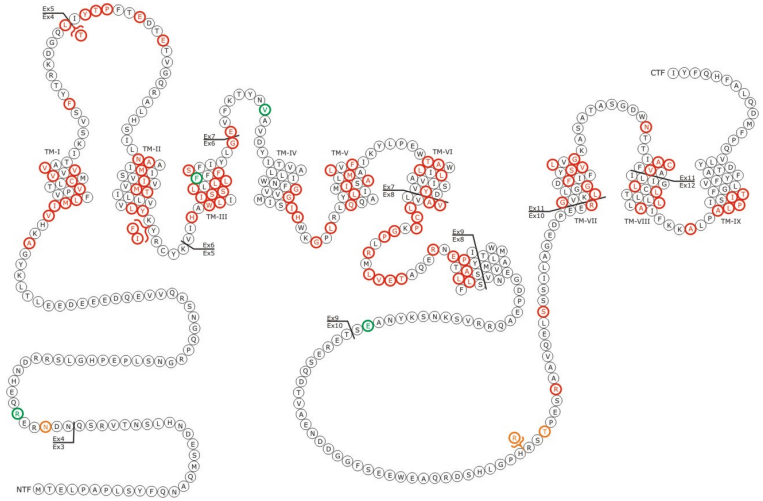
HPO and Common Disease



- Groza T, et al. (2015) The Human Phenotype Ontology: Semantic unification of common and rare disease (*Am J Hum Genet*, **97**:111-24)
- 132,006 annotations to terms of the HPO for 3,145 common human diseases

Phenotypic overlap: Rare and Common

- Mutations in the coding sequence of PSEN1 are associated with early-onset familial Alzheimer's disease



- Mutations in the promoter of *PSEN1* are associated with an increased risk of late-onset Alzheimer's disease
- **Question**: How common is this across the spectrum of all Mendelian disease?

Phenotypic overlap: Rare and Common

We compared phenotypes of rare diseases and common diseases with non-coding GWAS hits in the region of the gene

Disease	Common HPO terms
rs840016: Rheumatoid arthritis[43] *CD247*: Immunodeficiency due to defect in CD3-ζ	Edema (HP:0000969) Arthralgia (HP:0002829) Arthritis (HP:0001369) Autoimmunity (HP:0002960)
rs2268361: Polycystic ovary syndrome[44] *FSHR*: Ovarian hyperstimulation syndrome & Ovarian dysgenesis 1	Polycystic ovaries (HP:0000147) Decreased fertility (HP:0000144) Amenorrhea (HP:0000141)
rs13081389: Type 2 diabetes mellitus[45] *PPARG*: Lipodystrophy, familial partial, type 3	Hyperglycemia (HP:0003074) Hyperinsulinemia (HP:0000842) Hypertension (HP:0000822)
rs9644568: Hypertriglyceridemia[46] *LPL*: Type I hyperlipoproteinemia	Hypercholesterolemia (HP:0003124) Combined hyperlipidemia (HP:0008356) Atherosclerosis (HP:0002621) Pancreatitis (HP:0001733)
rs34778348: Parkinson's disease[47] *LRRK2*: Parkinson disease-8	Rigidity (HP:0002063) Bradykinesia (HP:0002067) Dementia (HP:0000726) Tremor (HP:0001337)
rs12726330: Parkinson's disease[48] *GBA*: Gaucher disease, various types	Akinesia (HP:0002304) Dementia (HP:0000726) Dementia (HP:0000726) Rigidity (HP:0002063)
rs12149070: COPD[49] *HYDIN*: Ciliary dyskinesia, primary, 5	Respiratory tract infection (HP:0011947) Respiratory insufficiency (HP:0002093)

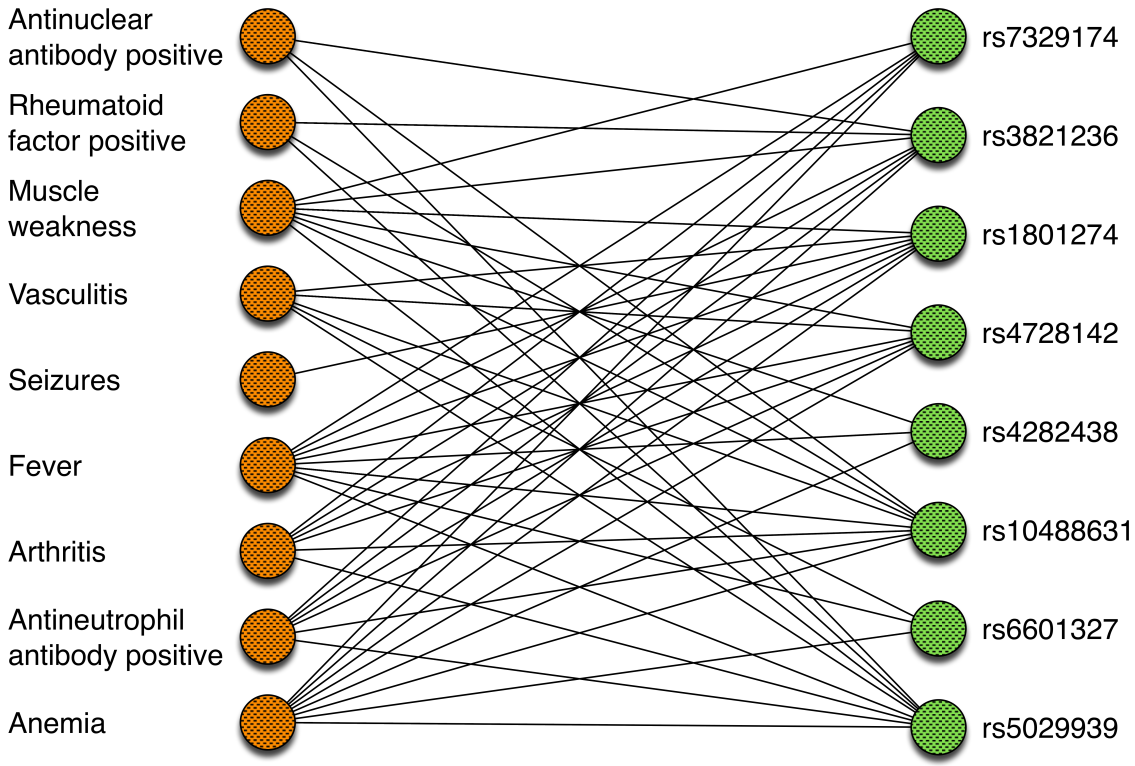
- Hundreds of candidates
- Overlap was significantly higher than randomized networks ($p = 1.6 \times 10^{-7}$)

Phenotypic Networks of Common Disease

GWAS studies have identified over 6,000 strong associations ($p < 10^{-8}$) to common complex diseases

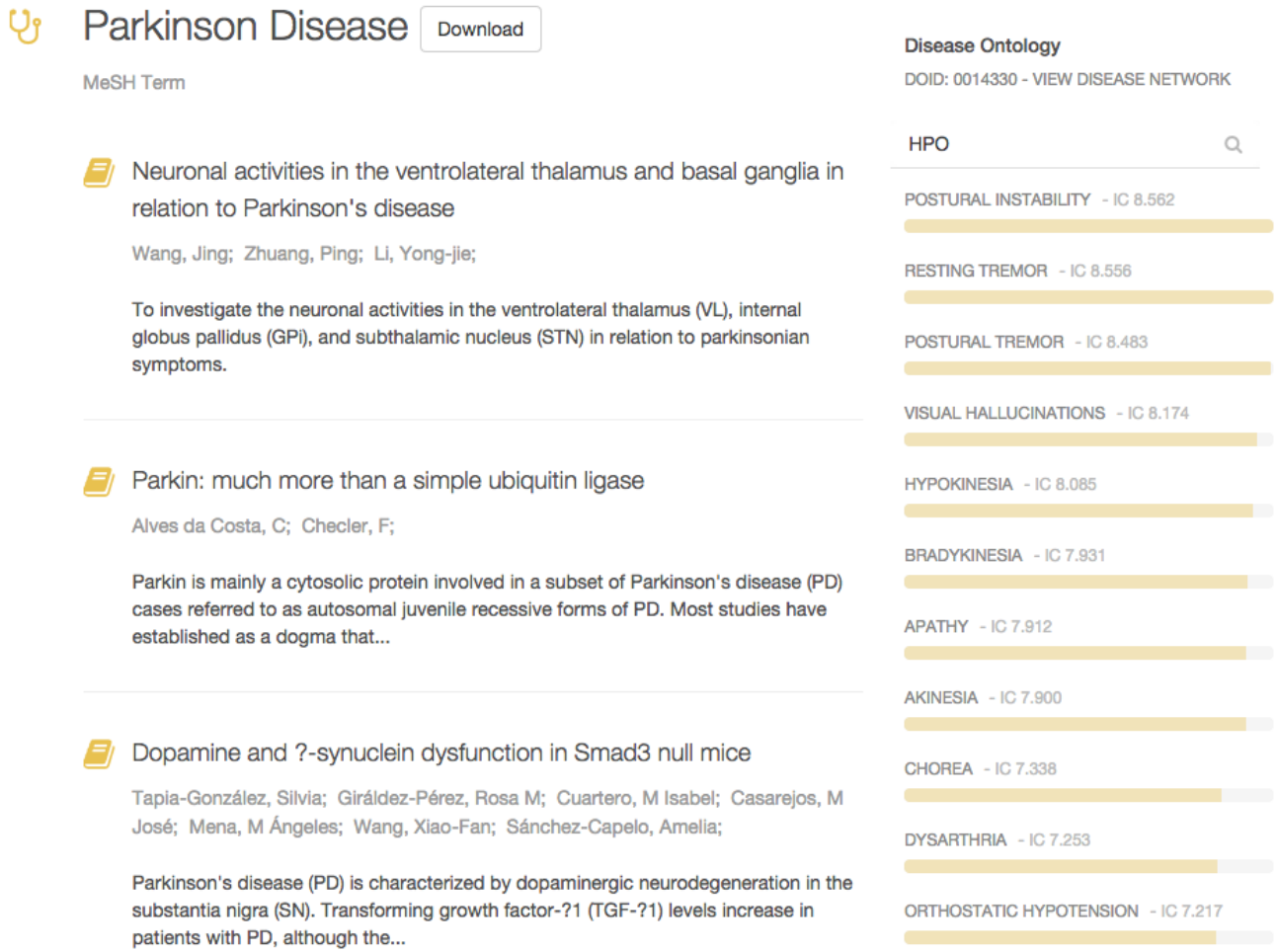
- Some GWAS hits have been associated with multiple diseases
- For instance, rs1344706 is associated with both schizophrenia and bipolar disease
- Of 16,152 GWAS hits analysed (GWAS Central), 863 were associated with 2 or more diseases
- For these 863 GWAS hits, we compared the HPO annotations of the diseases and looked for overlap

Phenotypic Networks of Common Disease



- We identified a substantial amount of phenotypic overlap including many dense subnetworks
- Highly statistically significant ($p = 2.3 \times 10^{-57}$)

Common Disease Annotations: Browsing



- The common disease annotations are available for browsing and download at `http://pubmed-browser.human-phenotype-ontology.org`

Plan

Cross species phenotype analysis

- The full value of projects such as UDP/UDN, 100K Genomes, IMPC, and many others will not be gotten without comprehensive computational phenotype resources
- Monarch and HPO are growing into a comprehensive interlinked database of the human phenome and diseasome with relevant model organism data, but are still underfunded
- Some areas in HP and MP will require extension

  Behaviour
  Metabolism (e.g., metabolomics)
  Craniofacial

Precision Medicine

- While current algorithms are working well for rare disease, more sophisticated representations of the phenotype will be required for common (complex) disease including cancer
  - Time course
  - Multimorbidity
  - Medications and treatments
  - Side effects
- Integrated algorithms for matching phenotype to molecular pathophysiology
  - Enable & shorten time to diagnosis
  - Identify actionable subtypes
  - Understand natural history and gender differences
- Connect to molecular taxonomy of disease
- Animal models of common disease

Genome sequencing & Non-coding Variation

- We are just beginning to explore the role of the entire genome in human disease
- Regulatory variation is probably more common than we think
- By understanding regulatory mutations in rare disease, we will have a path towards precision medicine – the great majority of GWAS hits are non-coding.
- How does phenotype differ from that of coding mutations in rare disease (e.g., *SHH* or *PTF1A*)
- How do the myriad variants found in all of us contribute to the phenotypic spectrum of common disease?
- Animal models of gene regulation are many but not currently annotated in a way that could be integrated in medical analysis

Thank you for your attention ....

Computational Biology and Bioinformatics Group, Charité

Sebastian Bauer (alumnus)
Peter Hansen
Verena Heinrich
Jonas Ibn-Salem
Marten Jäger
Sebastian Köhler
Peter Krawitz
Peter N. Robinson
Oliver Stolpe
Na Tzu
Tomasz Zemojtel
http://compbio.charite.de

Institute of Medical Genetics, Charité

Johannes Grünhagen (alumnus)
Gao Guo (alumnus)
Claus-Eric Ott
Sandra Dölken (alumna)
Denise Horn
Ulrike Krüger
Stefan Mundlos

ZFIN: The Zebrafish Model Organism Database, Eugene, OR, USA

Monte Westerfield
Barbara Ruef

Max Planck Institute for Molecular Genetics, Berlin

Marcel H. Schulz

University of Cambridge UK

George V. Gkoutous
Michael Ashburner
Paul Schofield

MRC Harwell UK

John Hancock

Berkeley Bioinformatics and Ontology Project (BBOP) California

Suzi Lewis
Chris Mungall
Nicole Washington

Mouse Informatics Group, Sanger Institute

Damian Smedley
Jules Jacobsen

Oregon Health & Sciences University

Melissa Haendel
Nicole Vasilevsky

HPO: Find out more

Overview articles	
Robinson PN et al 2008	The Human Phenotype Ontology: a tool for annotating and analyzing human hereditary disease. *Am J Hum Genet* **83**:610-5.
Robinson PN, Mundlos S 2010	The human phenotype ontology. *Clin Genet* **77**:525-34.
Köhler et al 2014	The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data. *Nucleic Acids Res.* **42**:D966-74.
Computational semantic reasoning over phenotypes	
Köhler et al (2013)	Construction and accessibility of a cross-species phenotype ontology along with gene annotations for biomedical research. *F1000Research* **2**:30
Köhler S et al. (2011)	Improving ontologies by automatic reasoning and evaluation of logical definitions. *BMC Bioinformatics* **12**:418.
Robinson PN, Webber C (2014)	Phenotype ontologies and cross-species analysis for translational research. *PLoS Genet* **10**:e1004268.
Decision support/exome analysis	
Köhler S et al 2009	Clinical diagnostics in human genetics with semantic similarity searches in ontologies. Am J Hum Genet. 2009 Oct;85(4):457-64.
Algorithms	
Schulz MH et al 2011	Exact score distribution computation for ontological similarity searches. BMC Bioinformatics. 12;12:441.
Bauer S et al 2012	Bayesian ontology querying for accurate and noise-tolerant semantic searches. Bioinformatics;28(19):2502-8.
Robinson PN, Bauer S 2011	Introduction to Bio-Ontologies. Chapman & Hall/CRC Mathematical & Computational Biology