ePTRI

European Paediatric Translational Research Infrastructure

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 777554

Omics application in drug target discovery and preclinical research

Pier Luigi Martelli – University of Bologna – ELIXIR-IT

Biotechnology to bring innovation in the paediatric drug development

October 2, 2020

The event is part of the European Biotech Week 2020

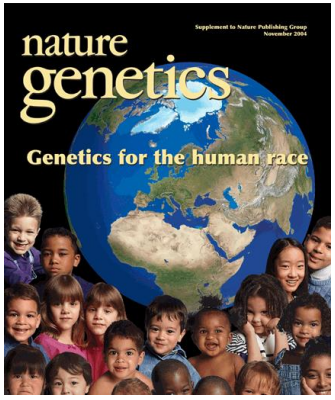
PART OF EUROPEAN BIOTECH WEEK

INNOVATION IS IN OUR GENES

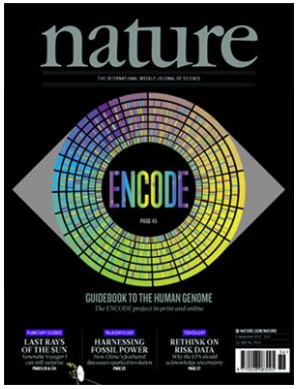
The "omic" revolution

The analysis of the components of a living organism in its entirety

15/02/2001



Human Variations 11/2004



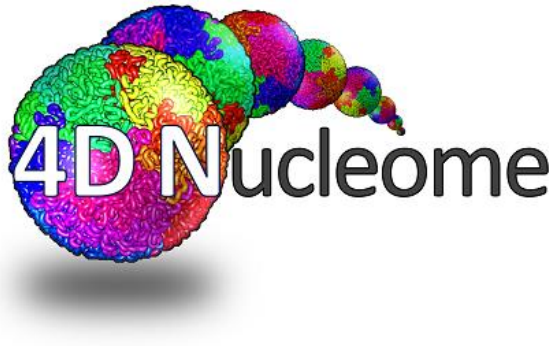
ENCODE 09/2012

RoadMap Epigenomics 02/2015

16/02/2001

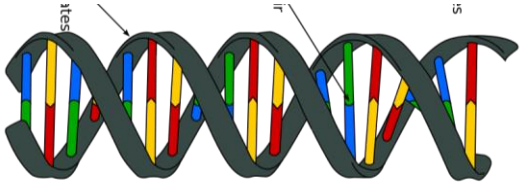
Human Proteome 05/2014

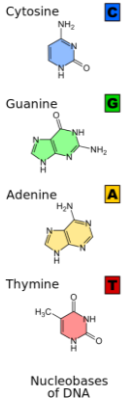
Next…



EPTRI
EUROPEAN PAEDIATRIC TRANSLATIONAL RESEARCH INFRASTRUCTURE

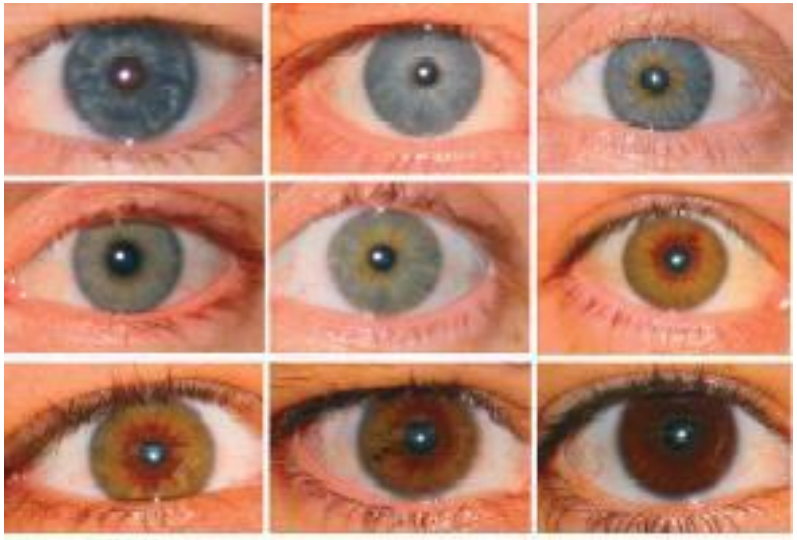
Genotype & Phenotype



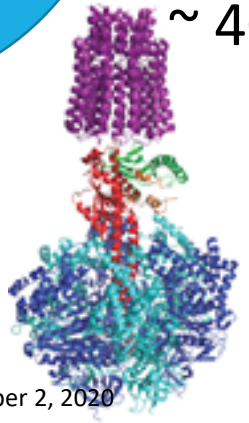


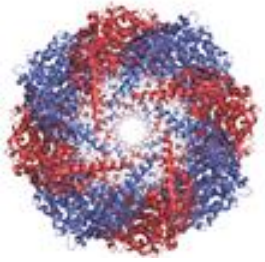
~ 3 billion bases
~ 20,000 protein coding genes

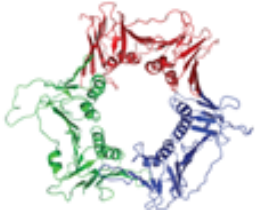
Coding for
~ 230,000 RNA transcripts
~ 40,000 proteins



Eye Colors
Strum and Frudakis (2004) Trends in Genetics 20:327



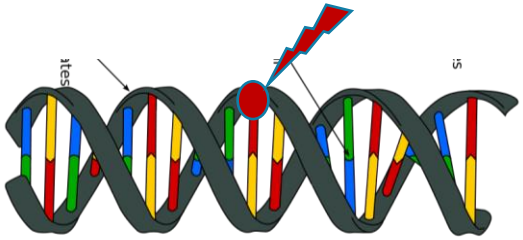


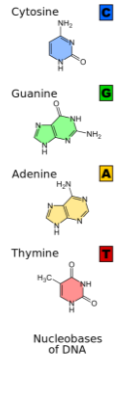


ePTRI
European Paediatric Translational Research Infrastructure

PART OF EUROPEAN BIOTECH WEEK
INNOVATION IS IN OUR GENES

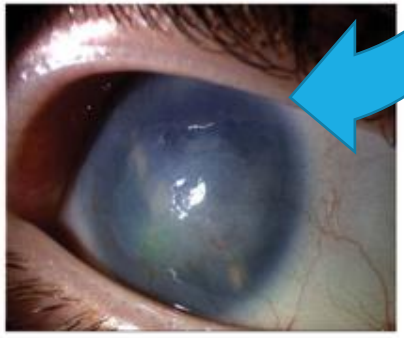
Genes & Diseases





WT

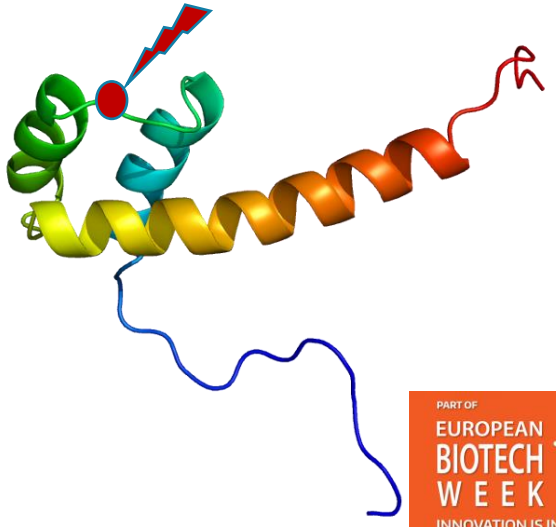
mut



Aniridia

PAX6$^{+/-}$

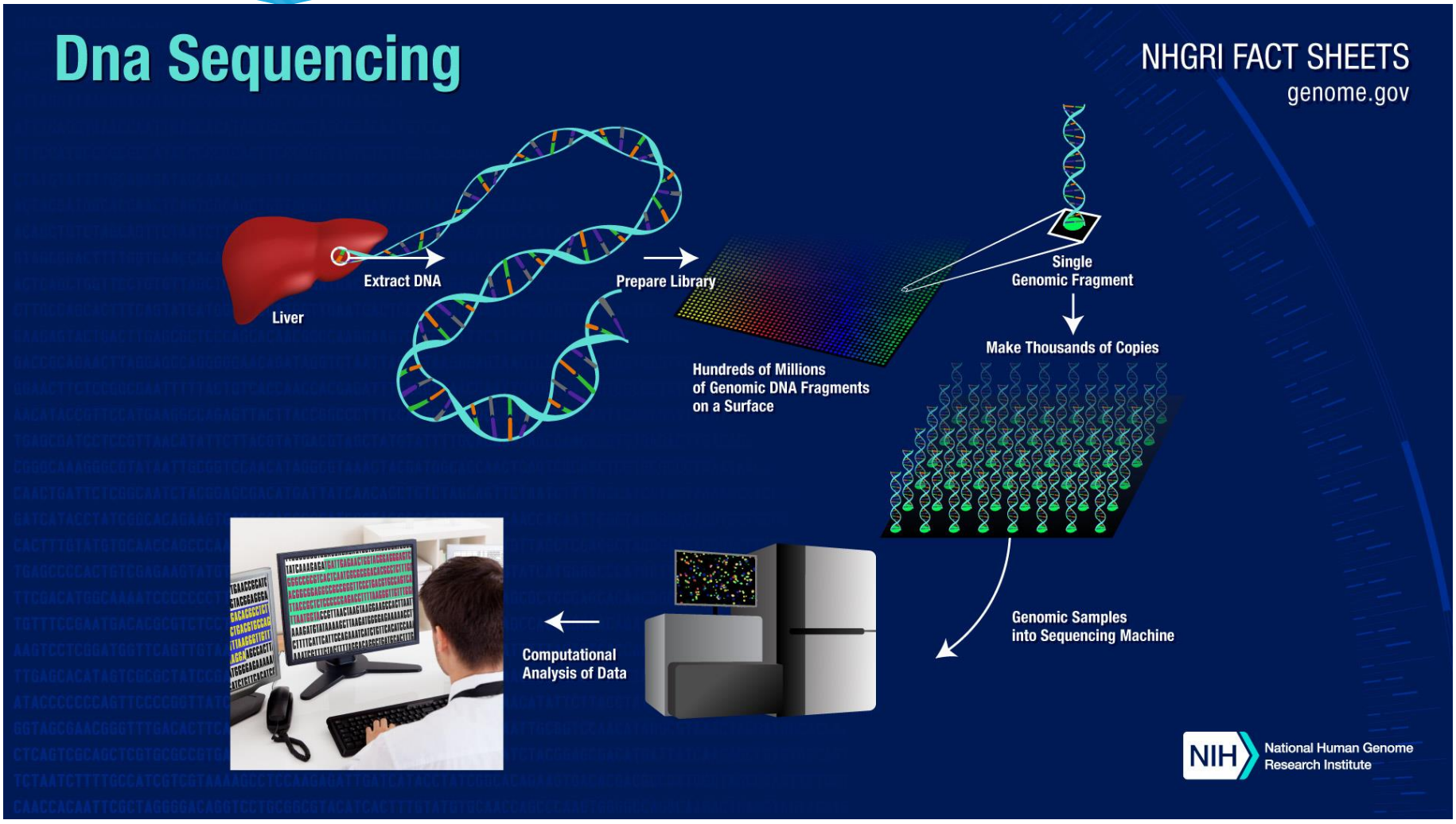
A single variation can lead to disease insurgence



ePTRI
European Paediatric Translational Research Infrastructure

PART OF EUROPEAN BIOTECH WEEK
INNOVATION IS IN OUR GENES

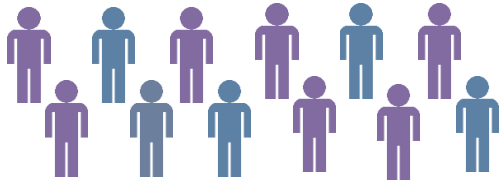
Deep sequencing



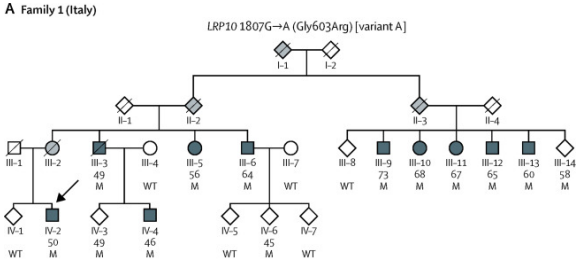
Genomic Techniques

Drug target discovery

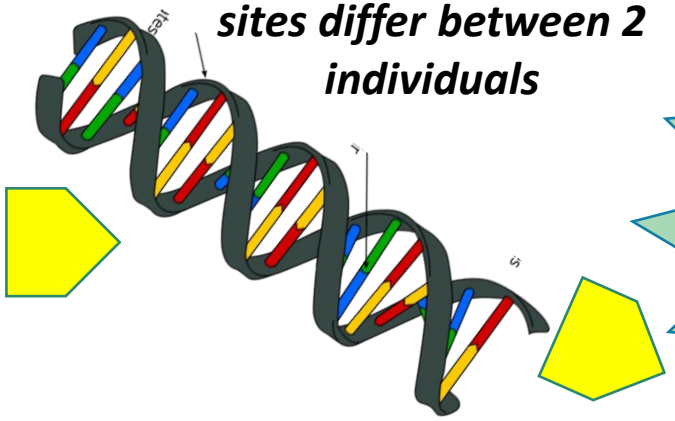
Case/control studies



Pedigree studies

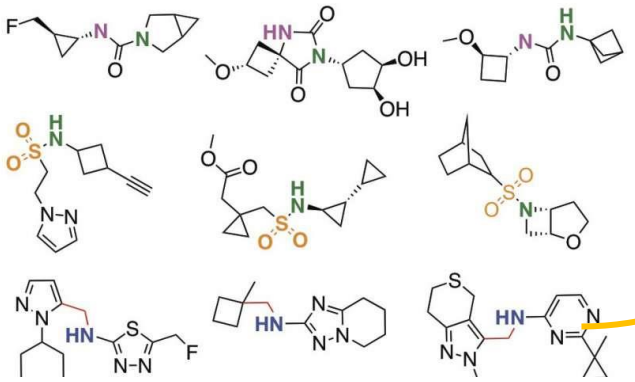


Around 4 millions sites differ between 2 individuals

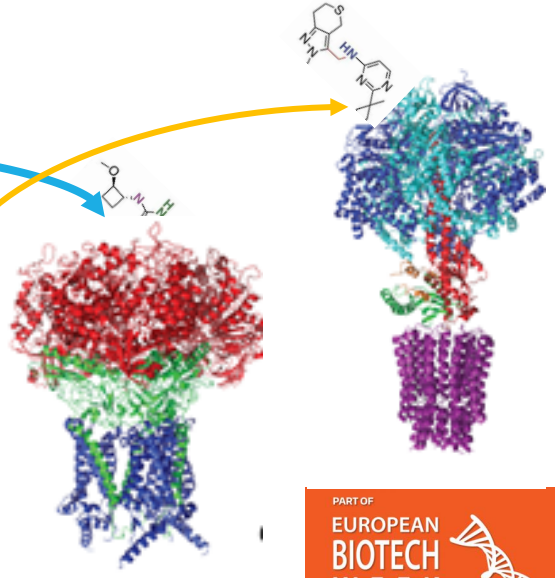


Significant variations must be prioritised

Mutated genes/proteins



Drug libraries



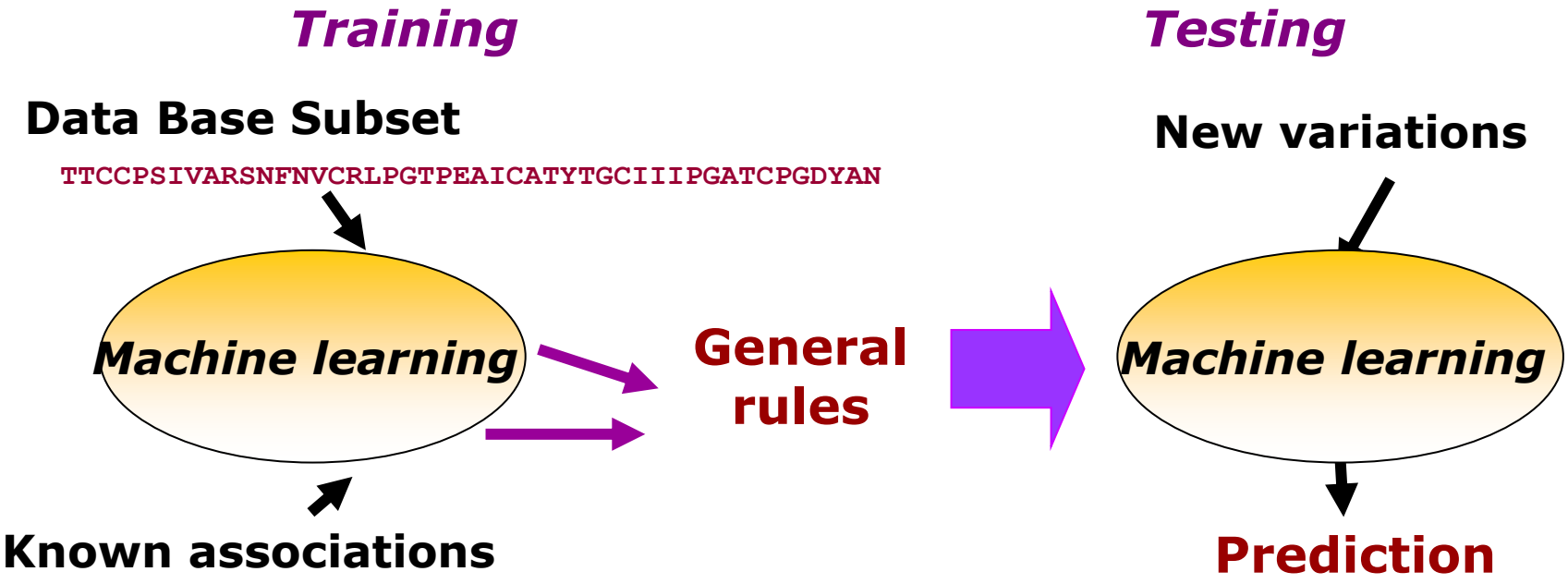
EPTRI
EUROPEAN PAEDIATRIC TRANSLATIONAL RESEARCH INFRASTRUCTURE

PART OF EUROPEAN BIOTECH WEEK
INNOVATION IS IN OUR GENES

Computational prioritization

- To date, some 30.000 variations in 13.000 human proteins are associated to disease

- Each complete sequence reports millions of variations.
- Case-control or pedigree analyses report tens to hundreds variations



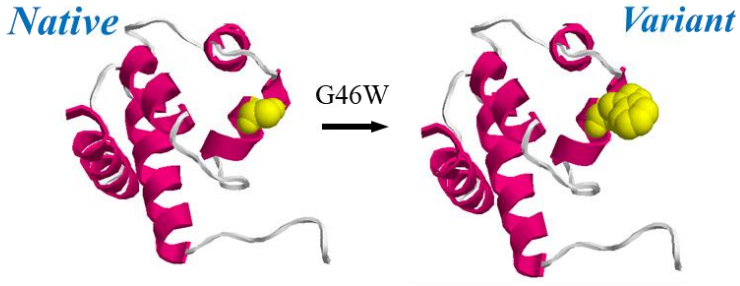
Computational methods can leverage available information and generalize to new uncharacterized variations

EPTRI
European Paediatric Translational Research Infrastructure

Part of European Biotech Week
Innovation is in our genes

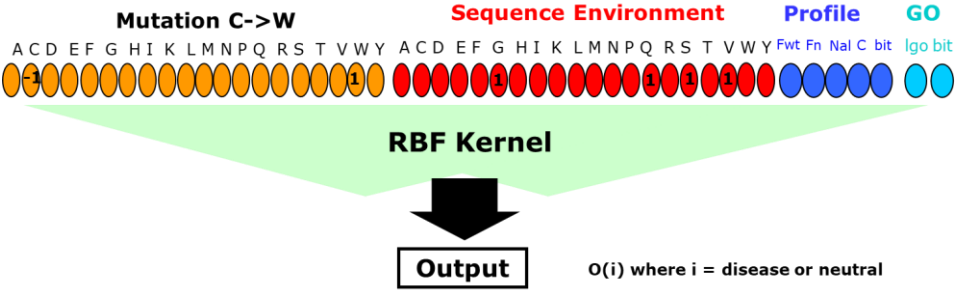
Predicting the effect of variations

www.biocomp.unibo.it/predictors



SNPs&GO

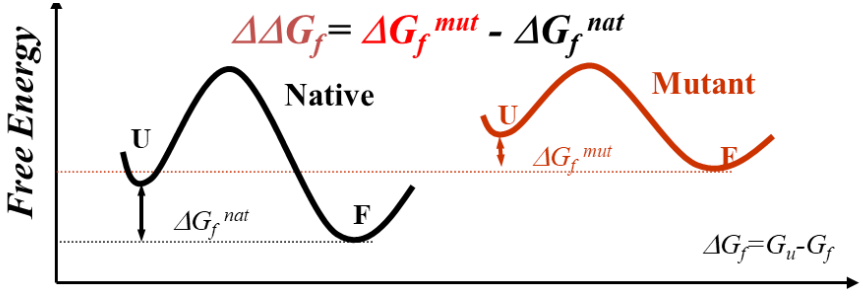
Relation to disease



Method	Q2	P[D]	Q[D]	P[N]	Q[N]	C	PM (%)
PolyPhen[a]	0.71	0.76	0.75	0.63	0.64	0.39	58
SIFT[b]	0.76	0.75	0.76	0.77	0.75	0.52	93
PANTHER[c]	0.74	0.77	0.73	0.71	0.76	0.48	76
Eremorph[d]	0.74	0.83	0.64	0.68	0.85	0.50	82
HybridMeth[e]	0.74	0.74	0.70	0.74	0.77	0.47	100
SNPs&GO	0.82	0.83	0.78	0.80	0.85	0.63	100

INPS-MD

Effects on protein stability



Method	Blind test Corr / SE	P53 blind set Corr / SE	
INPS	0.68 / 1.26	0.69 / 1.45	*sequence*
EASE-MM	0.69 / 1.34	0.56 / 1.13	
INPS3D	0.72 / 1.15	0.76 / 1.35	
MAESTRO°	0.70 / 1.13°	0.44 / 1.75	*structure*
mCSM^	0.73 / 1.08^	0.68 / 1.40	

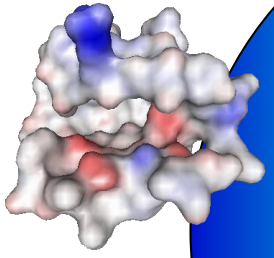
EPTRI
European Paediatric Translational Research Infrastructure

BIOTECH WEEK
INNOVATION IS IN OUR GENES

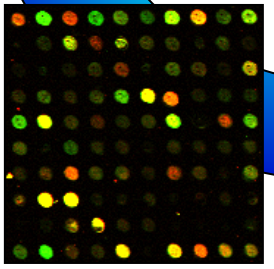
Genotype to Phenotype relation is complex

Genes in DNA...

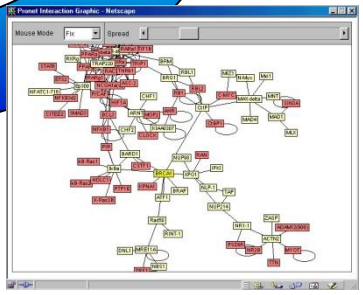
…code for proteins...



…proteins correspond to functions...

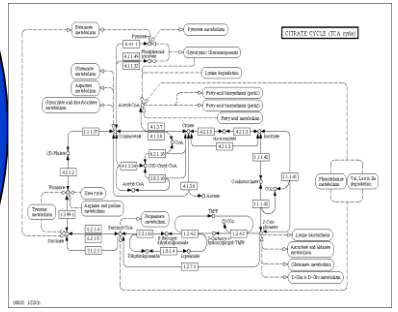


…when they are expressed



Proteins interact

….in metabolic pathways



…with different effects depending on variability

EPTRI
EUROPEAN PAEDIATRIC TRANSLATIONAL RESEARCH INFRASTRUCTURE

PART OF EUROPEAN BIOTECH WEEK
INNOVATION IS IN OUR GENES

Sequencing enables more Omics

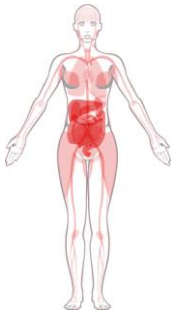
Cell/Tissue specific genome	→	Somatic variations (e.g. in cancer)
Cell/Tissue/Process specific transcriptome *(RNA-seq)*	→	When and where genes are expressed
Protein-DNA/RNA interactome *(Chip-seq)*	→	Mechanisms of gene expression regulation
DNA epigenome *(Chip-seq, Methyl-seq, HI-seq,…)*	→	Structure and state of chromatin (important for gene regulation)
Microbiome	→	Genome/transcriptome of host organisms important in health and disease

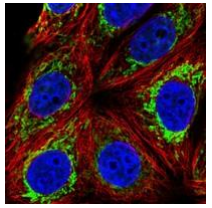
More omic levels can be investigated

Cell/Tissue/Process specific proteome

When and where proteins are expressed

Mass-spec, immunofluorescence, …

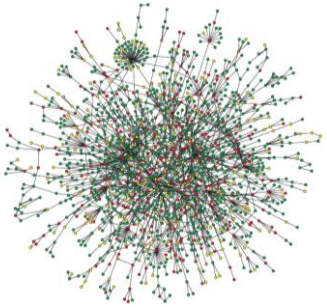




Protein interactome

Protein-protein interaction networks

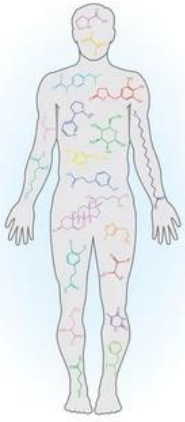
Yeast 2-Hybrid, affinity purification + Mass-spec, …

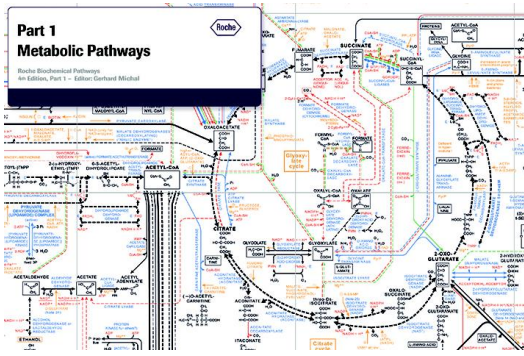


Cell/Tissue/Process specific metabolome

Concentration of small molecules

NMR, Mass-spec, HT chromatography..



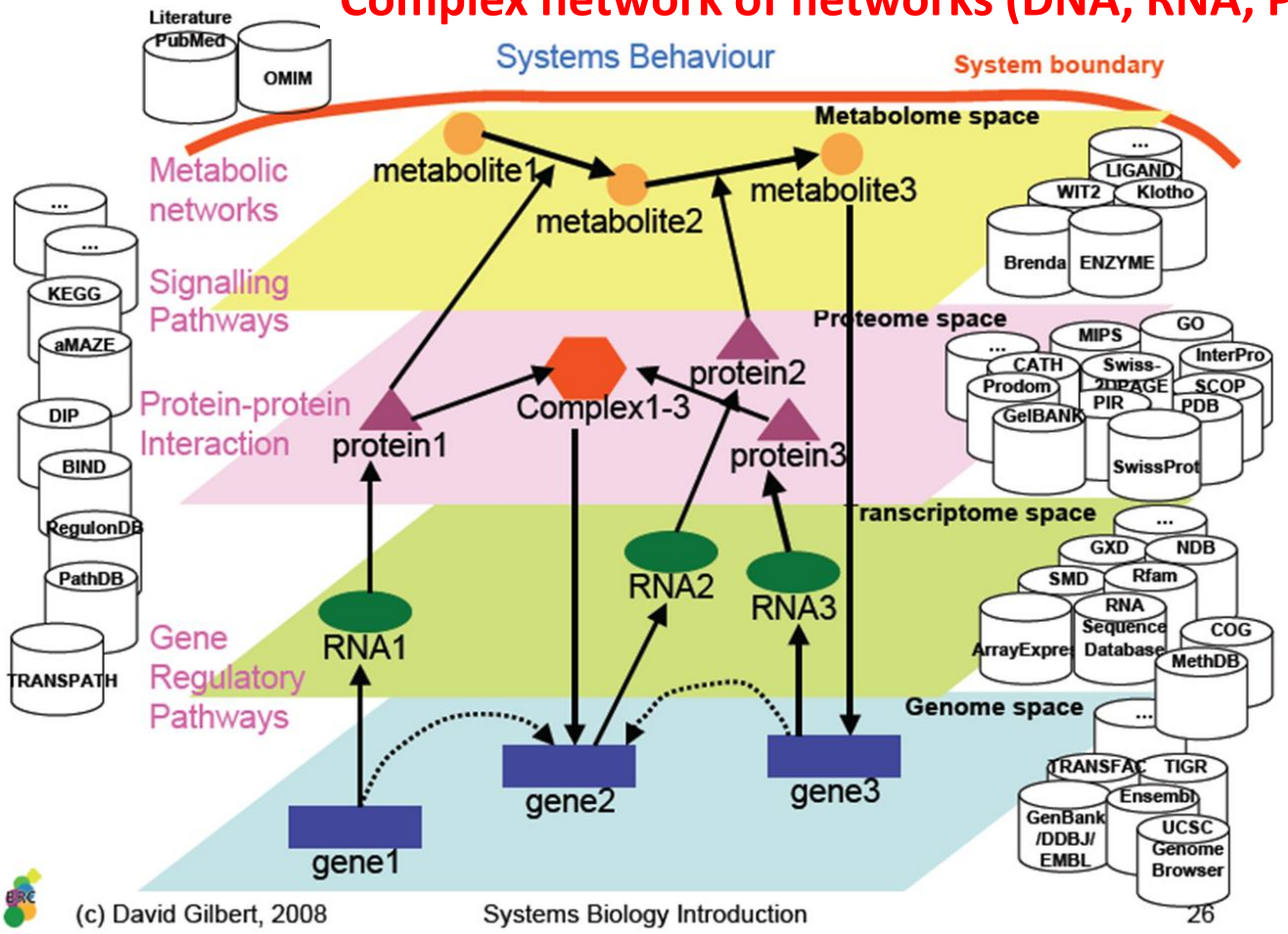


ePTRI
European Paediatric Translational Research Infrastructure

Part of European Biotech Week
Innovation is in our genes

Drug targets must be searched in the context of biological complexity

Complex network of networks (DNA, RNA, Proteins, Metabolites)

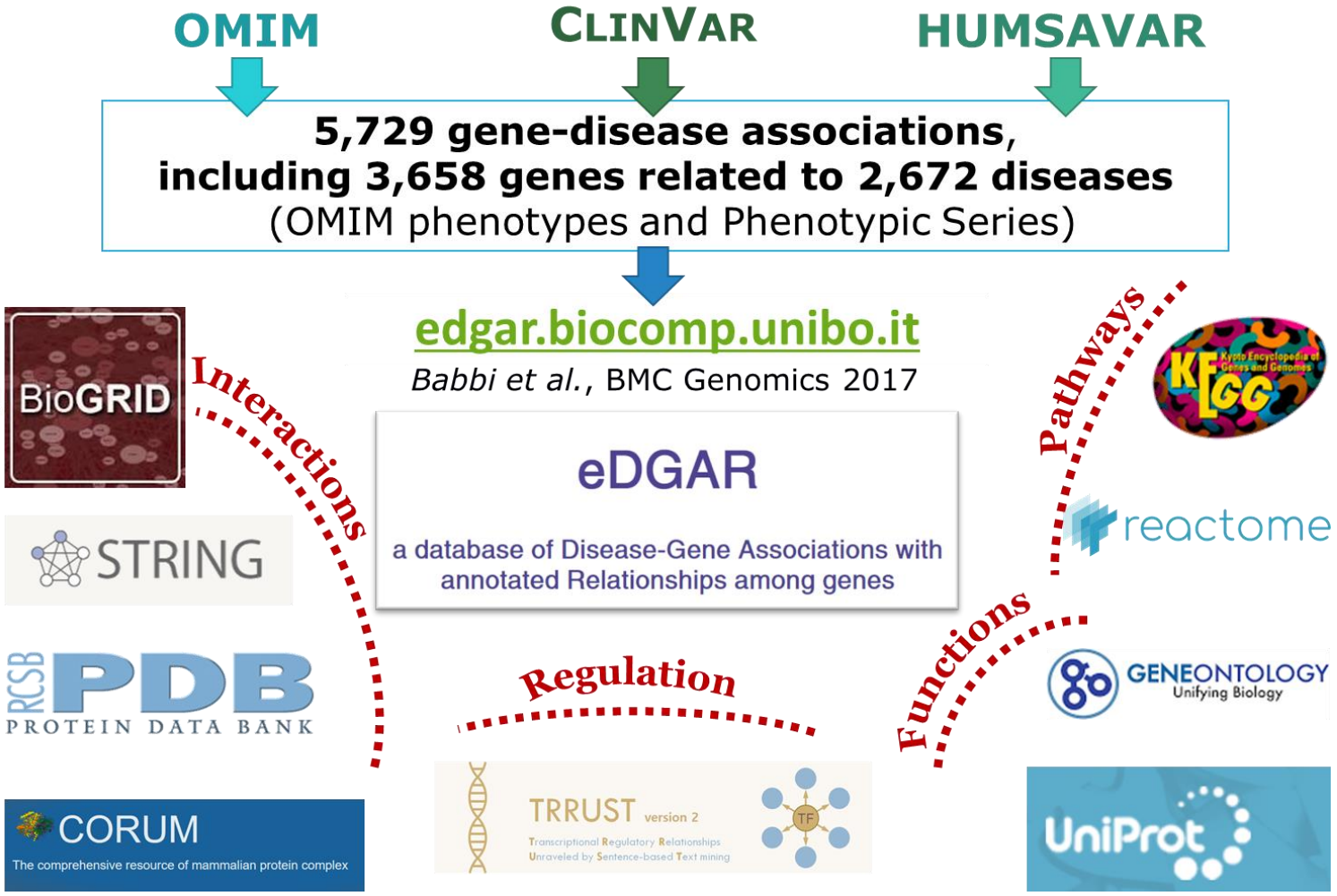


New targets can be predicted

New (unexpected) effects of a drug can be explained or predicted

Integrative computational methods are needed

Integrating the information available in different databases

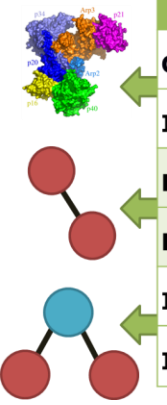


EPTRI
EUROPEAN PAEDIATRIC TRANSLATIONAL RESEARCH INFRASTRUCTURE

PART OF EUROPEAN BIOTECH WEEK
INNOVATION IS IN OUR GENES

eDGAR shows disease genes in their context

Interactions
among genes involved in polygenic diseases

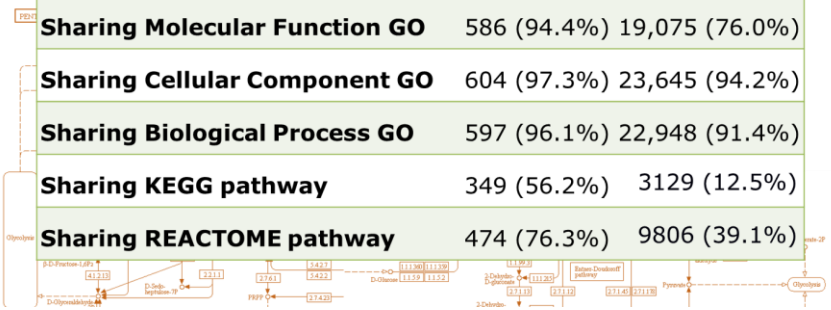


Pairs of genes:	# diseases	# genes
Co-resolved in PDB	96 (15.5%)	257 (0.8%)
In the same CORUM complex	86 (13.8%)	469 (1.9%)
Directly linked in STRING	291 (46.9%)	1,535 (6.1%)
Directly linked in BIOGRID	250 (40.3%)	4,355 (17.4%)
Indirectly linked in STRING	115 (18.5%)	944 (3.8%)
Indirectly linked in BIOGRID	160 (25.8%)	5,228 (20.8%)

Regulatory relations
among genes involved in polygenic diseases

Pairs of genes:	# diseases	# genes
TF/target pairs (both involved in the disease)	39 (6.3%)	81 (0.3%)
Co-regulated by the same TF (not involved in the disease)	273 (44.0%)	2,308 (9.2%)

Functional relations
among genes involved in polygenic diseases

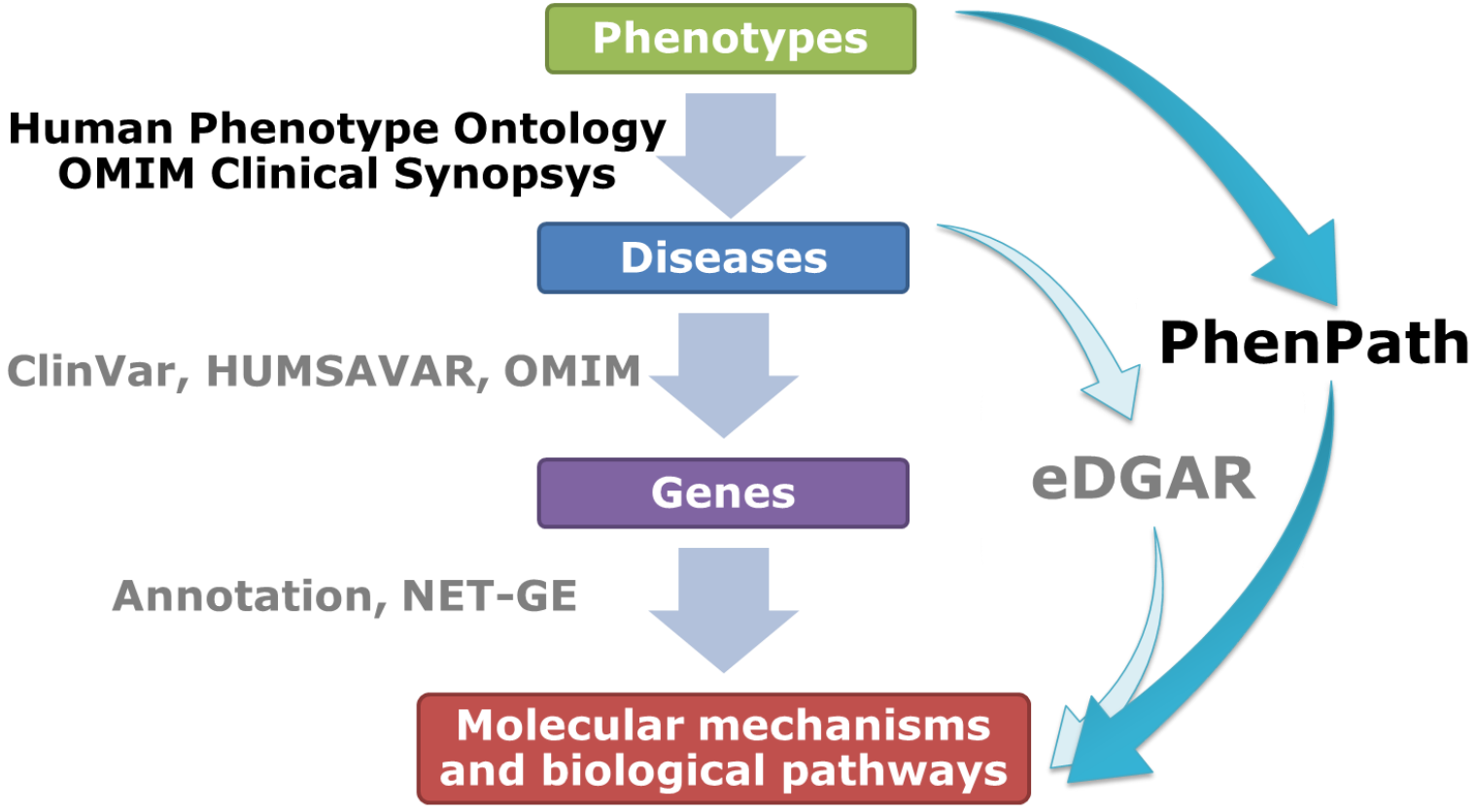


Pairs of genes:	# diseases	# genes
Sharing Molecular Function GO	586 (94.4%)	19,075 (76.0%)
Sharing Cellular Component GO	604 (97.3%)	23,645 (94.2%)
Sharing Biological Process GO	597 (96.1%)	22,948 (91.4%)
Sharing KEGG pathway	349 (56.2%)	3129 (12.5%)
Sharing REACTOME pathway	474 (76.3%)	9806 (39.1%)

ePTRI
European Paediatric Translational Research Infrastructure

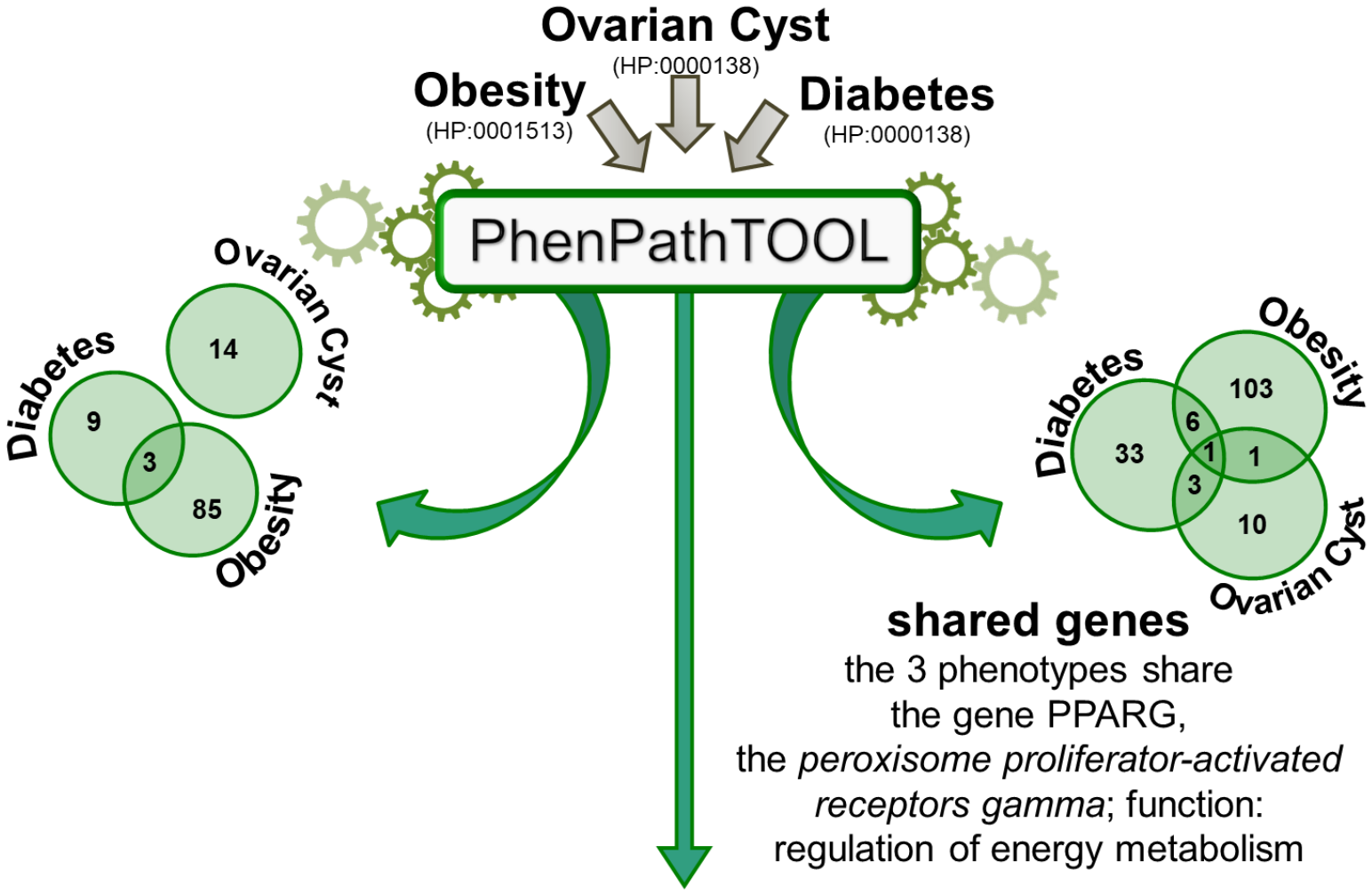
PART OF EUROPEAN BIOTECH WEEK
INNOVATION IS IN OUR GENES

From phenotypes to pathways and genes



http://phenpath.biocomp.unibo.it/phenpath/

Target proteins/pathways can be identified starting from phenotypes



Conclusions

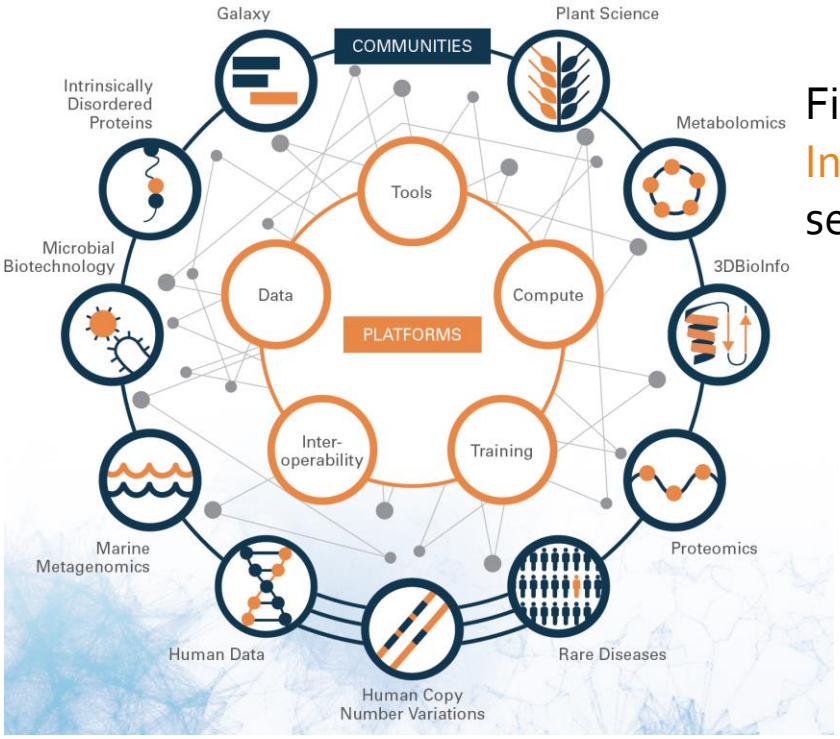
Deep sequencing, Mass spectrometry, Y2H and other techniques provide tools to dissect the complexity of biological systems

Drug targets and drug (side) effects must be discovered in the context of this complexity

Data needs to be collected, standardized, integrated and analysed

ELIXIR: a Research Infrastructure to face the Big Data challenge in Biology in Biology

ELIXIR is an **intergovernmental organisation** (Landmark European Research Infrastructure) that brings together "**bioinformatic resources" for life sciences** from across Europe. These resources include **databases**, **software tools**, **training materials**, **best practices**, **cloud storage** and **supercomputers**.



Five technical **platforms** for Compute, Data, Tools , Interoperability and Training complemented by several **user communities**

1 hub 23 national nodes including over 160 Research Organizations.

EPTRI
European Paediatric Translational Research Infrastructure

BIOTECH WEEK
INNOVATION IS IN OUR GENES

www.biocomp.unibo.it