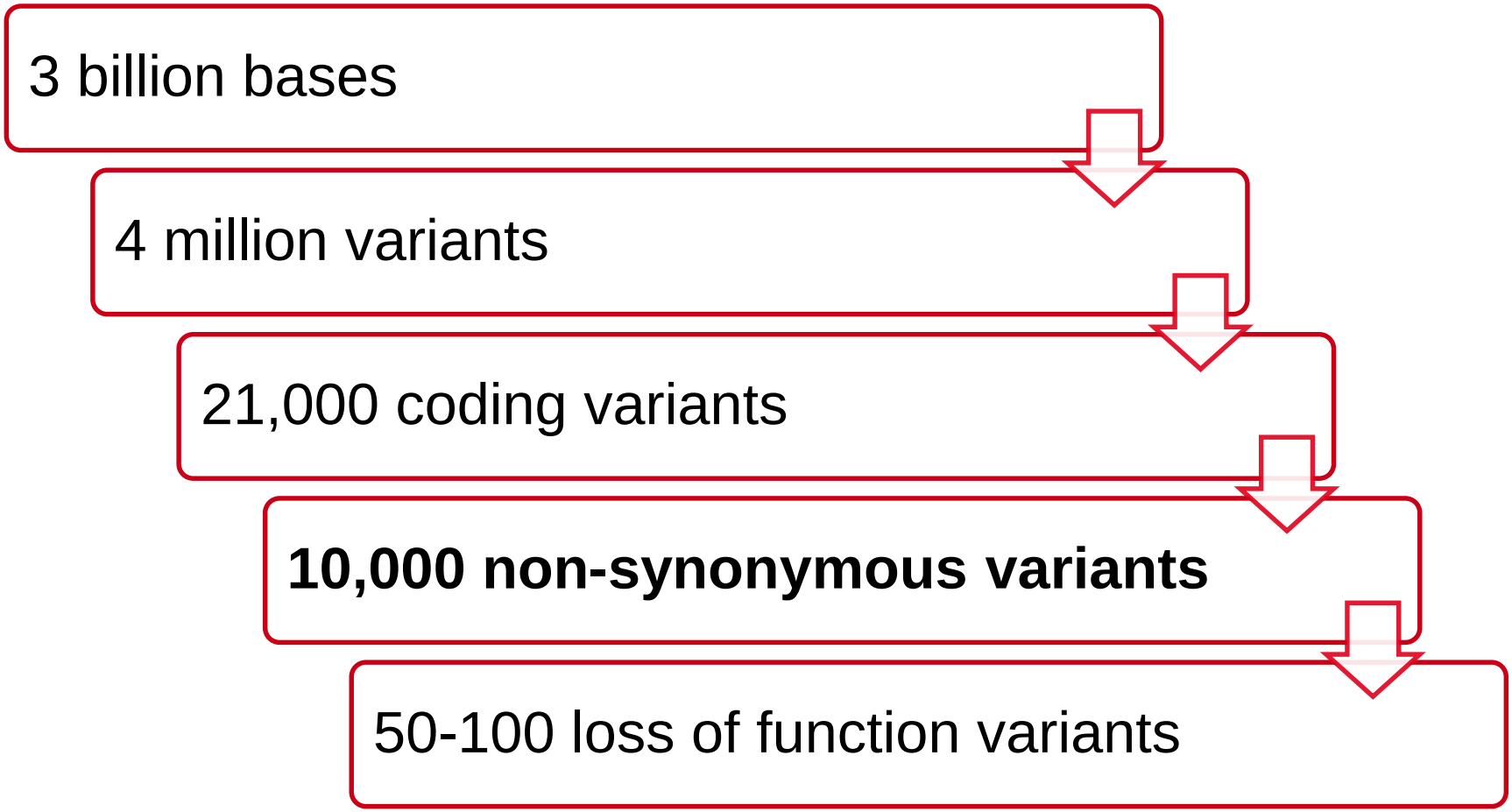
Interpreting Genome Data for Personalised Medicine

Professor Dame Janet Thornton

EMBL-EBI

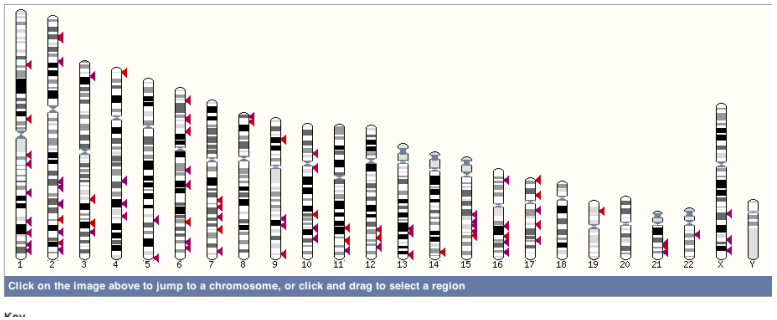
EMBL-EBI

Deciphering a genome



Deciphering a variant

e.g. Variants associated with Coronary heart Disease



- **Many different types of variants**
  - SNPs – single nucleotide polymorphisms
  - Copy Number Variants
  - Splice site variants
  - Variants in Promoters/enhancers
  - Variants which affect post-translational modifications (signalling) and epigenetic signals

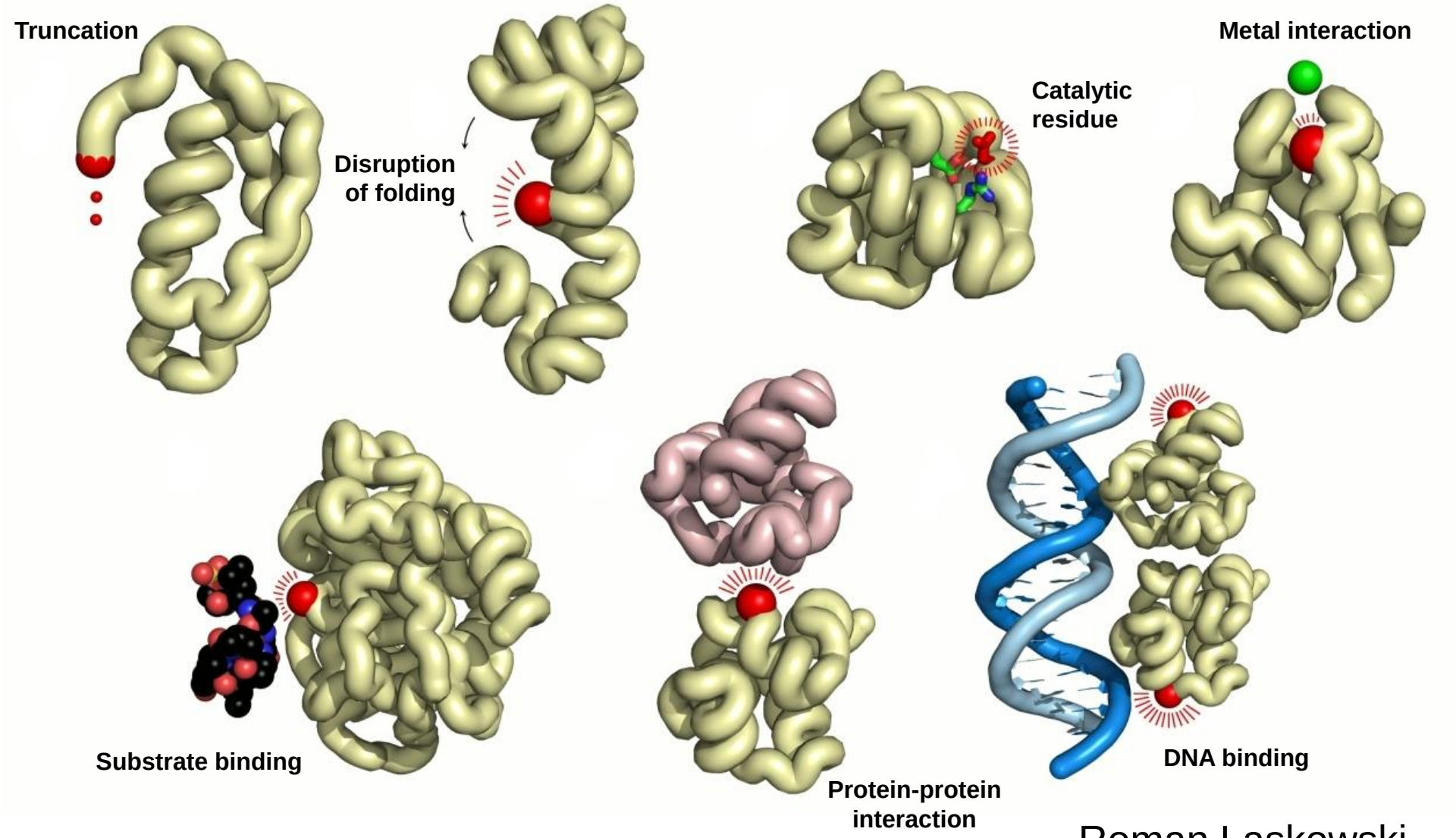
All have potential to cause disease

Interpreting variants is not just about the DNA!

- Impact on
  - Protein, Protein Structure, Pathways
- Using data from Evolutionary biology
  - Per-base conservation, mouse orthologs, mouse phenotypes
- Considering Association with small molecule biology
  - Metabolism, Pre-clinical drugs, Marketed drugs
- Sample tracking and ontologies
  - Place where cell type, tissue and phenotypes integrate

Coding variants that cause diseases

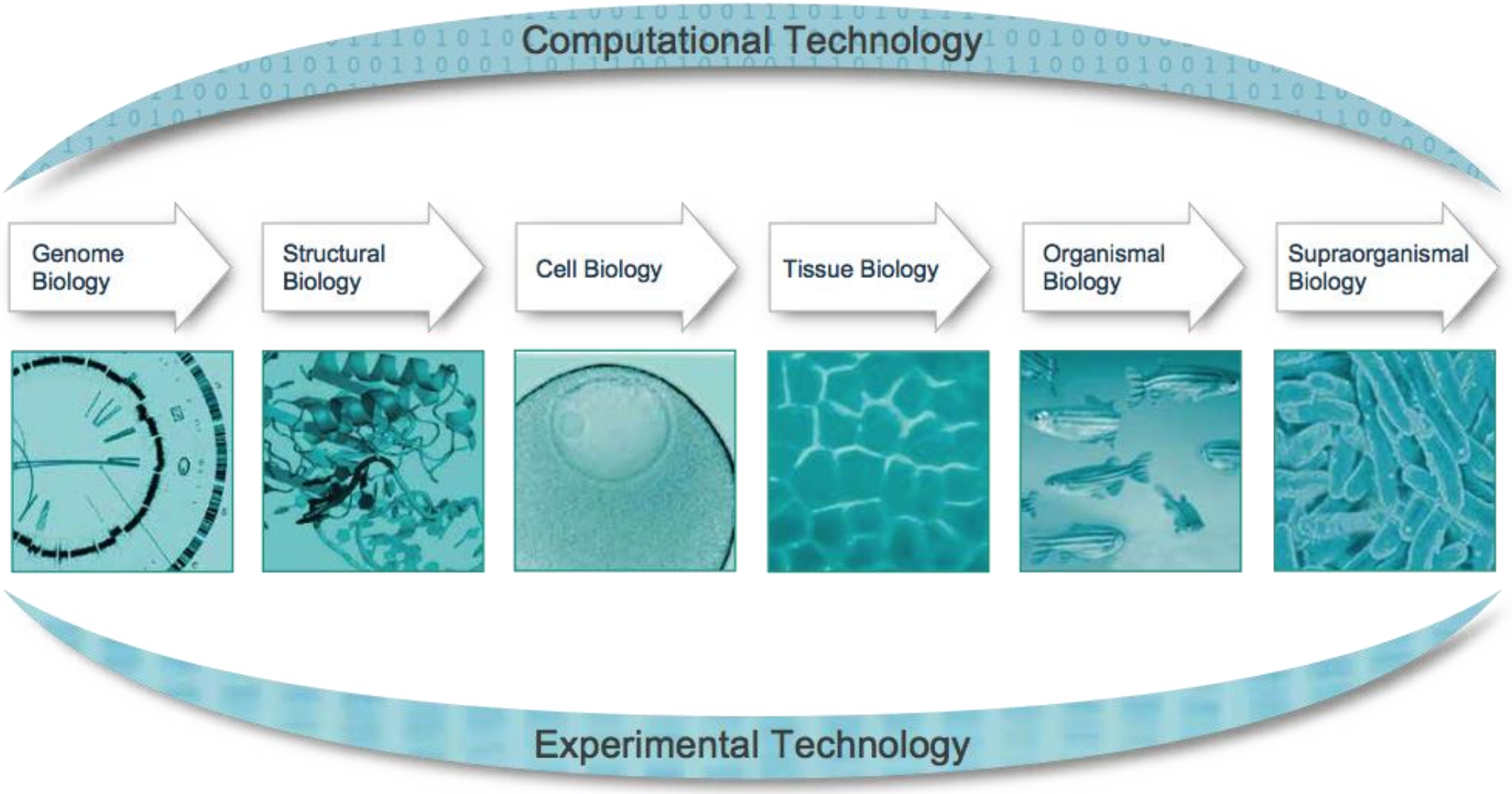
Structure can help explain a variant's disruption of protein function



Roman Laskowski

Interpretation the clinical effect of a mutation depends primarily on current knowledge

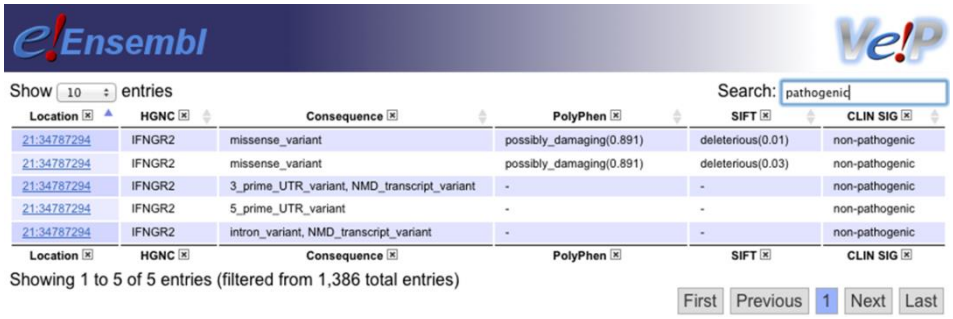
From Genotype to Disease: many layers of complexity

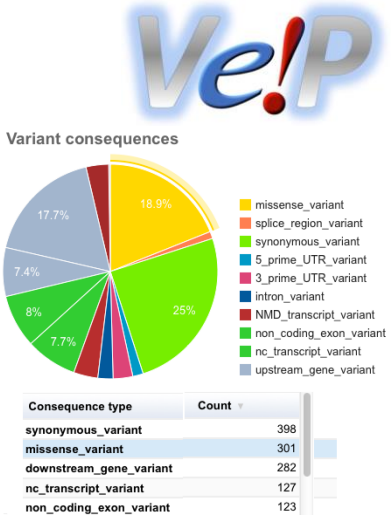


How can we improve variant interpretation for Personalised Medicine?

Through Research

- Research to understand basis of individual diseases
- Including development of Tools to interpret data
  - Eg Variant Effect Predictor VeP





McLaren et al. 2016 (Genome Biology, *in press*), www.ensembl.org/vep

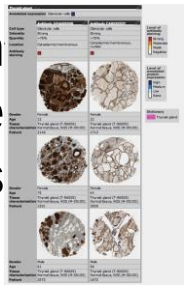
How can we improve variant interpretation for Personalised Medicine?

Develop Shared (global) public data resources

- To ensure knowledge is captured
- To ensure up to date
- To identify actionable variants
- Data accessible to all

orphanet

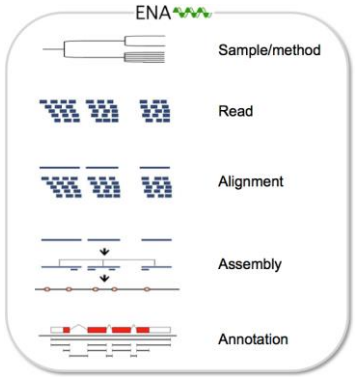
Human Proteome Atlas



Ensembl

GWAS Catalog

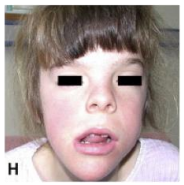
The NHGRI-EBI Catalog of published genome-wide association studies

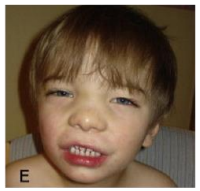


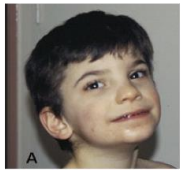
European Nucleotide Archive

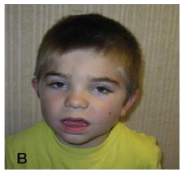
DECIPHER
GRCh37

Deciphering Developmental Disorders



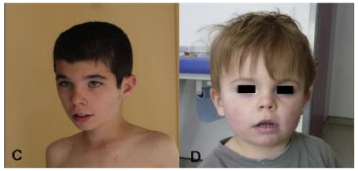


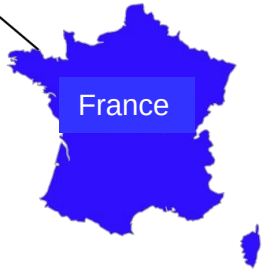


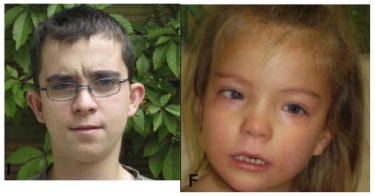


A novel microdeletion syndrome at 9q21.13 characterised by mental retardation, speech delay, epilepsy and characteristic facial features

Eur J Med Genet
March 2013







How can we improve variant interpretation for Personalised Medicine?

Develop Global Standards for Data

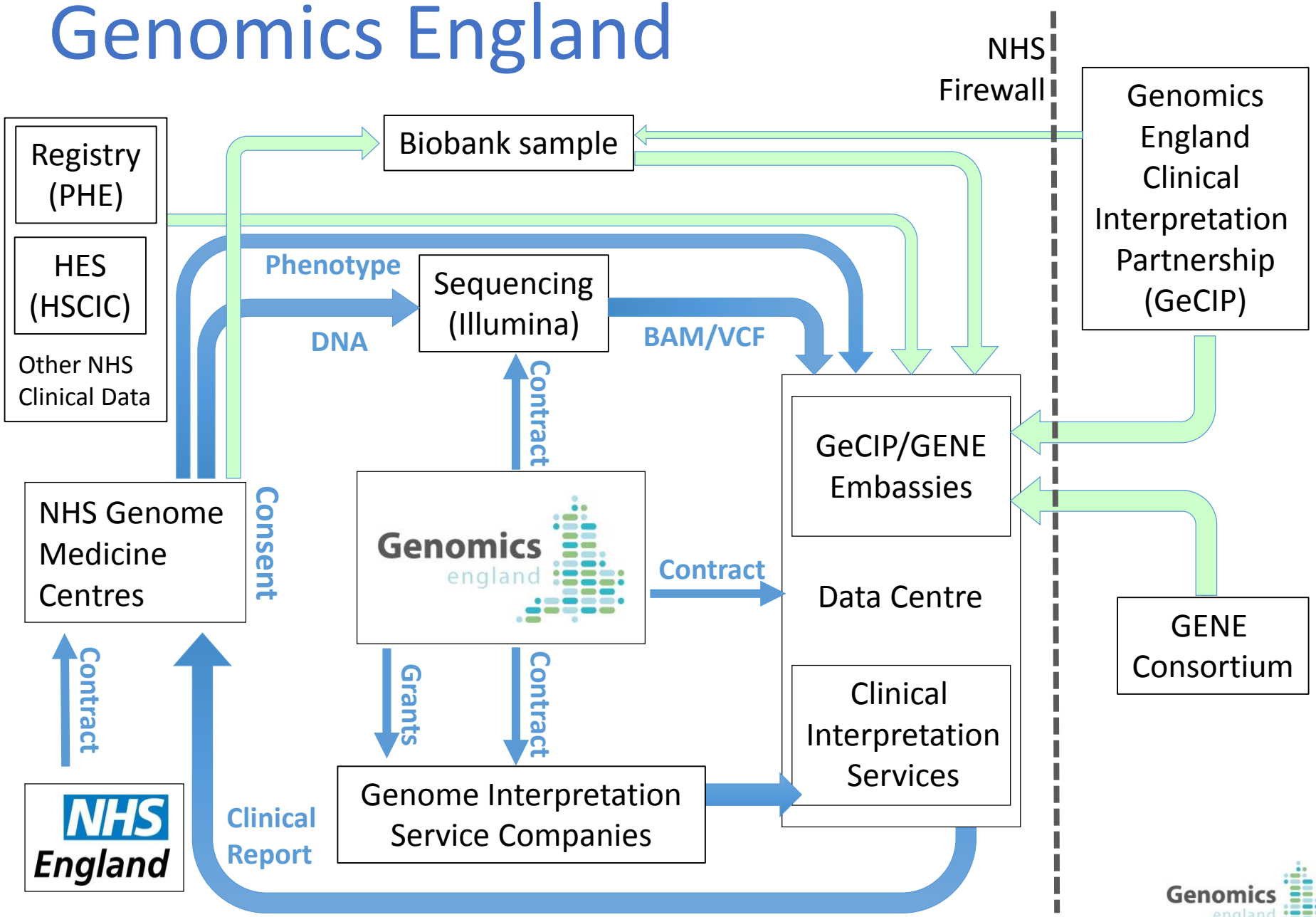
- ELIXIR
- Global Alliance for Genomes and Health
  http://genomicsandhealth.org/

How can we improve variant interpretation for Personalised Medicine?

Build a suitable Computer Infrastructure(s)

- Appropriate IT systems/sample management in each hospital
  - EHRs for all individuals
  - Accessible to doctor and patient but also **secure**
  - Accessible for researchers, but **secure**
- National Infrastructure to coordinate hospital data/analysis - would probably be a hub & spokes model

Genomics England



An ‘Institute for Biomedical Data and Informatics’ in each country

- Each country could establish an “Institute for Biomedical data and informatics” - BMI
  - This institute would ‘look after’ the personal genomics data
  - In bigger nations, this is likely to be a network, but with a centre of gravity
- Europe (EU) could then establish a coordination project/infrastructure to ‘join up’ these institutes
- Such an effort would build on and interact closely with current biological EU infrastructures eg ELIXIR; BBMRI

How can we work together to facilitate Personalised Medicine & Biomedical Data Interpretation across Europe?

Coordination across Euro

Technical platforms

- *Data*
- *Interoperability*
- *Tools*
- *Compute*
- *Training*
- *Industry*

An international distributed infrastructure for biological data

ELIXIR's support to Personalised Medicine

excelerate

Data Archiving

EGA @ EBI (European Genome-phenome Archive) ⟷ EGA @ CRG (Centre de Regulació Genòmica)

Secure data access remote API (GA4GH)

High speed encrypted data transfer

GridFTP/Globus/Aspera

Data

Sequencing centers

Managing Samples

BBMRI Biobanking and Biomolecular Resources Research Infrastructure

EMBL-EBI BioSamples

Managing Access

Data Owner
Data Access Agreement
Data Access Committee
Data Request

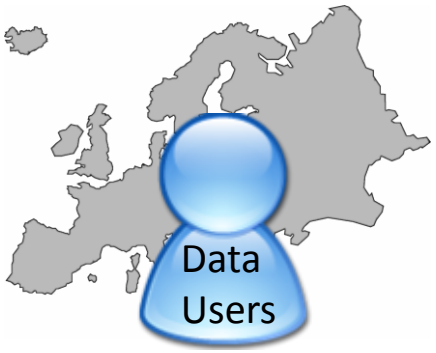
Authorization Management Tools
(EGA and CSC REMS)

elixir

Services and Coordination

- Federated Authentication
- Authorization
- Dataset registry
- Data transfer hub
- Policy and Legal Framework

Moving data to users



Bringing users to data

Secure Compute Clouds

BSC

CSC

EMBL-EBI

How can we improve variant interpretation for Personalised Medicine?

Training for Biomedical bioinformaticians/IT staff

- Develop a cadre of Clinical Scientists who are experts in Genomic Medicine
- Training In Data Handling & Interpretation in Clinical Context
  - E.g. Masters in Genomic Medicine

Training for hospital staff

- CPD training for doctors (in hospitals and for GPs)
- Nurses & caring staff