

Global Data Resources for Human Genomics & Health

Mallory Freeberg, PhD
Human Genomics Team Lead, EMBL-EBI

Australian BioCommons Webinar

20 September 2024

What is EMBL-EBI?

- World leading source of public biomolecular data
- Our vision is to benefit humankind by advancing scientific discovery and impact through bioinformatics.
- Part of the European Molecular Biology Laboratory (EMBL), Europe's flagship laboratory for the life sciences.



The European Molecular Biology Laboratory



EMBL-EBI

Bioinformatics



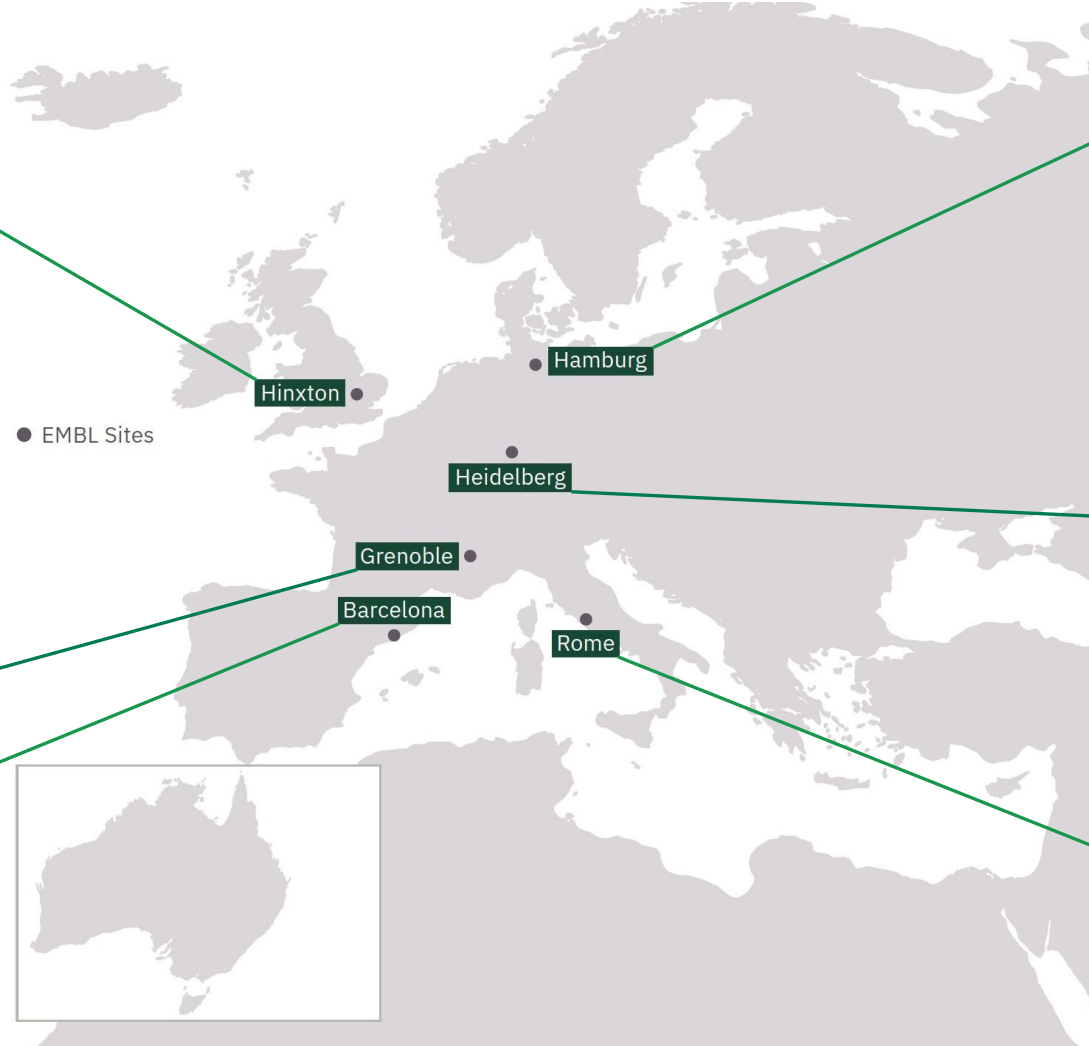
Grenoble

Structural biology



Barcelona

Tissue biology
and disease
modelling



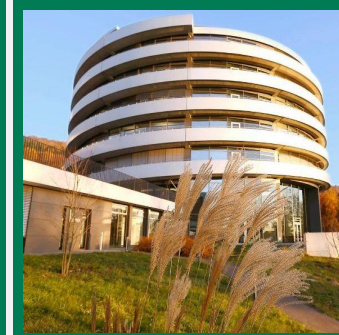
Hamburg

Structural biology



Heidelberg

Life sciences



Rome

Epigenetics
and neurobiology



EMBL member states

Member states (29)

| | |
|---------------------|---------------------|
| Austria 1974 | Portugal 1998 |
| Denmark 1974 | Ireland 2003 |
| France 1974 | Iceland 2005 |
| Germany 1974 | Croatia 2006 |
| Israel 1974 | Luxembourg 2007 |
| Italy 1974 | Czech Republic 2014 |
| Netherlands 1974 | Malta 2016 |
| Sweden 1974 | Hungary 2017 |
| Switzerland 1974 | Slovakia 2018 |
| United Kingdom 1974 | Montenegro 2018 |
| Finland 1984 | Lithuania 2019 |
| Greece 1984 | Poland 2019 |
| Norway 1985 | Estonia 2023 |
| Spain 1986 | Latvia 2024 |
| Belgium 1990 | |



Associate member state

Australia 2008



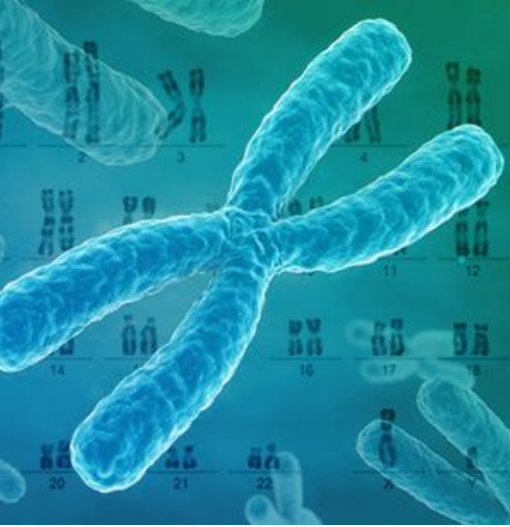
Prospect member states

Serbia

Bulgaria



EMBL-EBI's mission



Deliver data
resources

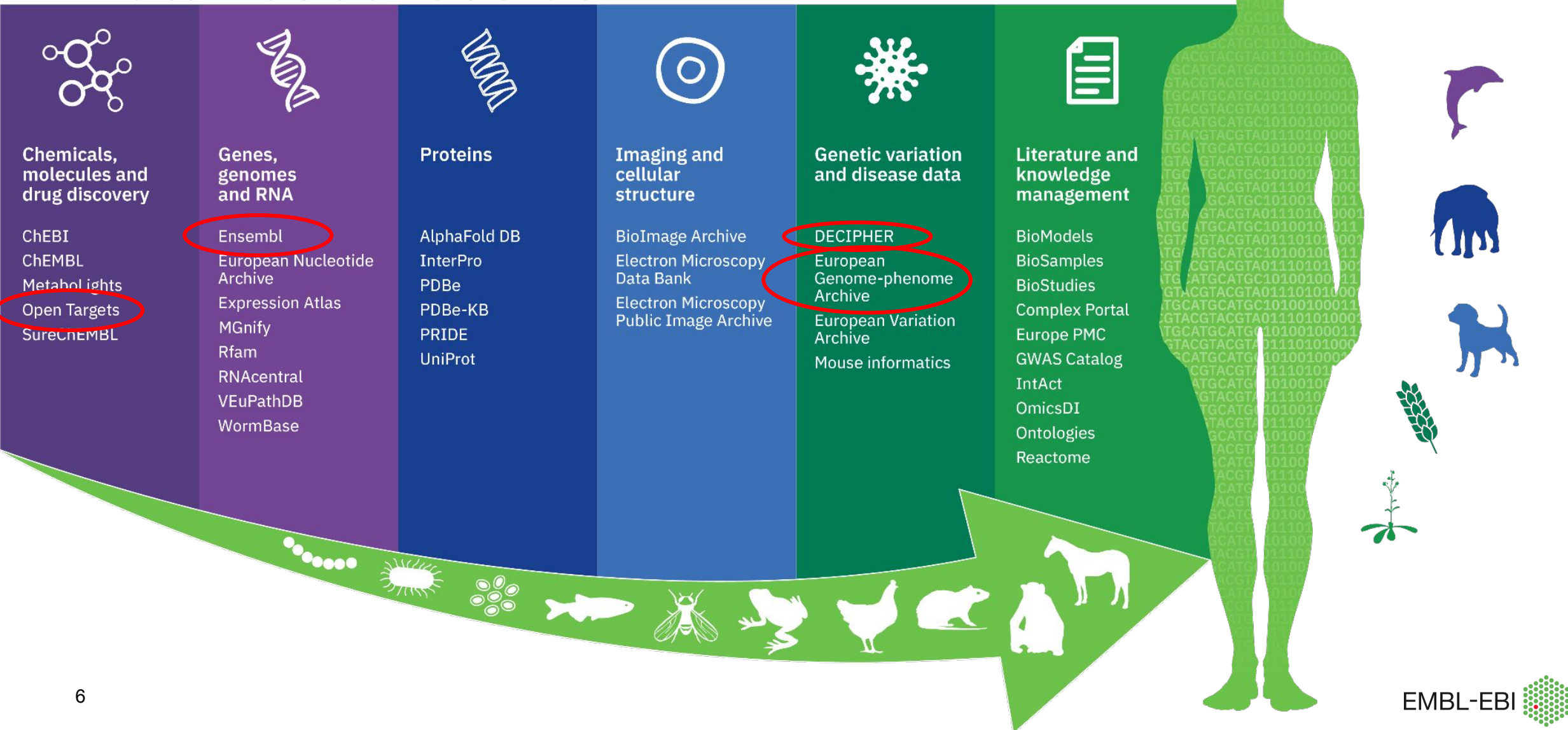
Perform excellent
research

Train the next
generation of
scientists

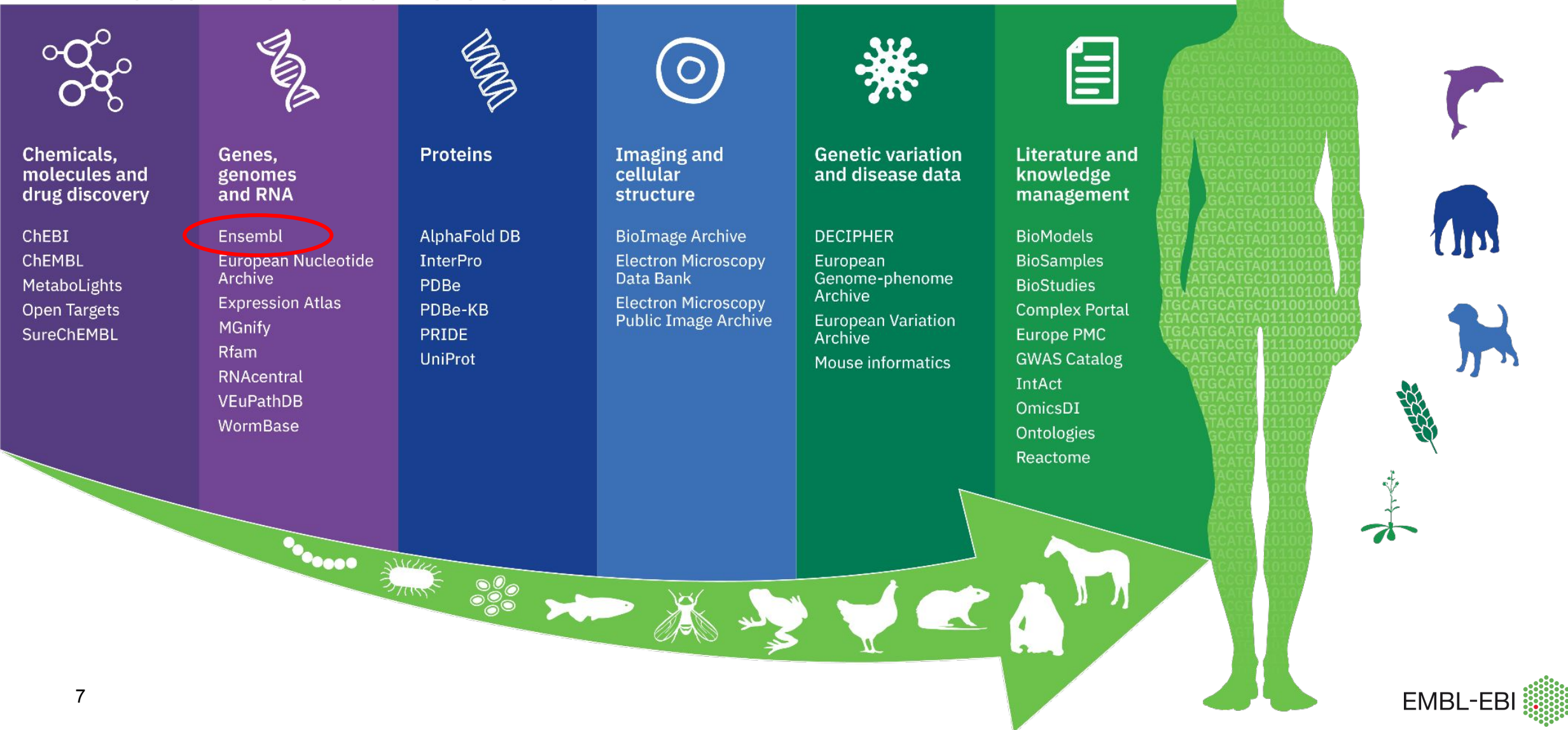
Engage with
industry

Coordinate
bioinformatics
in Europe

Data resources at EMBL-EBI



Data resources at EMBL-EBI



Tools

[All tools](#)

BioMart >

Export custom datasets from Ensembl with this data-mining tool

BLAST/BLAT >

Search our genomes for your DNA or protein sequence

Variant Effect Predictor >

Analyse your own variants and predict the functional consequences of known and unknown variants

Search

All species for

e.g. BRCA2 or rat 5:62797383-63627669 or rs699 or coronary heart disease

All genomes

-- Select a species --



Pig breeds
Pig reference genome and 12 additional breeds

[View full list of all species](#)

Favourite genomes



Human
GRCh38.p14

[Still using GRCh37?](#)



Mouse
GRCm39



Zebrafish
GRCz11

Ensembl is a genome browser for vertebrate genomes that supports research in comparative genomics, evolution, sequence variation and transcriptional regulation. Ensembl annotate genes, computes multiple alignments, predicts regulatory function and collects disease data. Ensembl tools include BLAST, BLAT, BioMart and the Variant Effect Predictor (VEP) for all supported species.

Ensembl Release 112 (May 2024)

- Many new fish genomes have been added to Ensembl
- Population frequency data is available for chicken, dog, goat and sheep through VEP
- Update to our current regulation annotation. The promoters now align with the 5' ends of known transcripts
- VEP will be updated to use the dbNSFP commercial data release

[More release news](#) on our blog

Ensembl Rapid Release

New assemblies with gene and protein annotation every two weeks.

Note: species that already exist on this site will continue to be updated with the full range of annotations.

The Ensembl Rapid Release website provides annotation for recently produced, publicly available vertebrate and non-vertebrate genomes from biodiversity initiatives such as Darwin Tree of Life, the Vertebrate Genomes Project and the Earth BioGenome Project.

[Rapid Release news](#) on our blog

Compare genes across species



Find SNPs and other variants for my gene



Gene expression in different tissues



Retrieve gene sequence

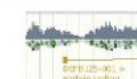
```

GCCTGACITTCGGGTTGG
GGCTTGTGGGGGGAGC
GGGCTCTGCTGGGCTT
AGGGACAGATTTGTGA
CACCTCTGGAGCGGTTI
CCCAGTCCAGCGTGGC
    
```

Find a Data Display



Use my own data in Ensembl



EMBL-EBI Ensembl creates, integrates and distributes reference datasets and analysis tools that enable genomics. We are based at [EMBL-EBI](#) and our software and data are freely available. Our [acknowledgements page](#) includes a list of current and previous funding bodies. [How to cite Ensembl](#) in your own publications.





About the ENSEMBL project

<https://beta.ensembl.org/>


ENSEMBL

Genome data & annotation

About using Ensembl  

Species selector 

Create & manage your own species list

Genome browser 

Look at genes & transcripts in their genomic context

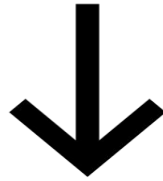
Entity viewer 

Get gene & transcript information



What is Ensembl?

```
AGTGCTTAATGTGGCTAGTGGCACCGGTTTGGACAGCACAGCTGTAAAATGTTCCCATCCTCAC
AGTAAGCTGTTACCGTTCAGGAGATGGGACTGAATTAGAATTCAAACAAATTTTCCAGCGCTT
CTGAGTTTTACCTCAGTCACATAATAAGGAATGCATCCCTGTGTAAGTGCATTTTGGTCTTCTG
TTTTGCAGACTTATTTACCAAGCATTGGAGGAATATCGTAGGTAAAAATGCCTATTGGATCCAA
AGAGAGGCCAACATTTTTTGAATTTTTAAGACACGCTGCAACAAAGCAGGTATTGACAAATTT
TATATAACTTTATAAATTACACCGAGAAAGTGTCTTCTAAAAAATGCTTGCTAAAAACCCAGTA
CGTCACAGTGTGCTTAGAACCATAAACTGTTCCCTTATGTGTGTATAAATCCAGTTAACAACAT
AATCATCGTTTGCAGGTTAACCACATGATAAATATAGAACGTCTAGTGGATAAAGAGGAAACTG
GCCCCTTGACTAGCAGTAGGAACAATTACTAACAAATCAGAAGCATTAATGT
```



Ensembl annotates and maps genomic features from genome sequences

What is Ensembl?

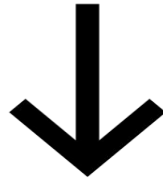
```
AGTGCTTAATGTGGCTAGTGGCACCGGTTTGGACAGCACAGCTGTAAAATGTTCCCATCCTCAC
AGTAAGCTGTTACCGTTCAGGAGATGGGACTGAATTAGAATTCAAACAAATTTTCCAGCGCTT
CTGAGTTTTACCTCAGTCACATAATAAGGAATGCATCCCTGTGTAAGTGCATTTTGGTCTTCTG
TTTTGCAGACTTATTTACCAAGCATTGGAGGAATATCGTAGGTAAAAATGCCTATTGGATCCAA
AGAGAGGCCAACATTTTTTGAATTTTTAAGACACGCTGCAACAAAGCAGGTATTGACAAATTT
TATATAACTTTATAAATTACACCGAGAAAGTGTCTTCTAAAAAATGCTTGCTAAAAACCCAGTA
CGTCACAGTGTGCTTAGAACCATAAACTGTTCCCTTATGTGTGTATAAATCCAGTTAACAACAT
AATCATCGTTTGCAGGTTAACCACATGATAAATATAGAACGTCTAGTGGATAAAGAGGAAACTG
GCCCCTTGACTAGCAGTAGGAACAATTACTAACAAATCAGAAGCATTAATGT
```



Ensembl is an ‘added value resource’ bringing together information from a wide range of other databases in a single site

What is Ensembl?

```
AGTGCTTAATGTGGCTAGTGGCACCGGTTTGGACAGCACAGCTGTAAAATGTTCCCATCCTCAC
AGTAAGCTGTTACCGTTCCAGGAGATGGGACTGAATTAGAATTCAAACAAATTTTCCAGCGCTT
CTGAGTTTTACCTCAGTCACATAATAAGGAATGCATCCCTGTGTAAGTGCATTTTGGTCTTCTG
TTTTGCAGACTTATTTACCAAGCATTGGAGGAATATCGTAGGTAAAAATGCCTATTGGATCCAA
AGAGAGGCCAACATTTTTTGAATTTTTAAGACACGCTGCAACAAAGCAGGTATTGACAAATTT
TATATAACTTTATAAATTACACCGAGAAAGTGTCTTCTAAAAAATGCTTGCTAAAAACCCAGTA
CGTCACAGTGTGCTTAGAACCATAAACTGTTCCCTTATGTGTGTATAAATCCAGTTAACAACAT
AATCATCGTTTGCAGGTTAACCACATGATAAATATAGAACGTCTAGTGGATAAAGAGGAAACTG
GCCCCTTGACTAGCAGTAGGAACAATTACTAACAAATCAGAAGCATTAAATGT
```



NIH U.S. National Library of Medicine NCBI

Genome Data Viewer

www.ncbi.nlm.nih.gov/gdv



www.ensembl.org (vertebrates)



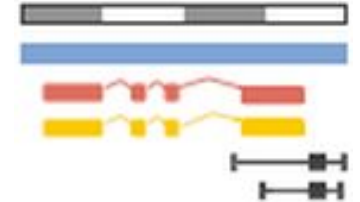
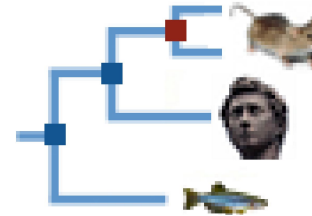
www.ensemblgenomes.org
(non-vertebrates)



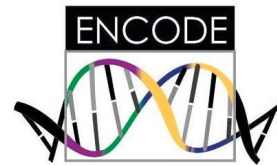
www.genome.ucsc.edu

Ensembl features

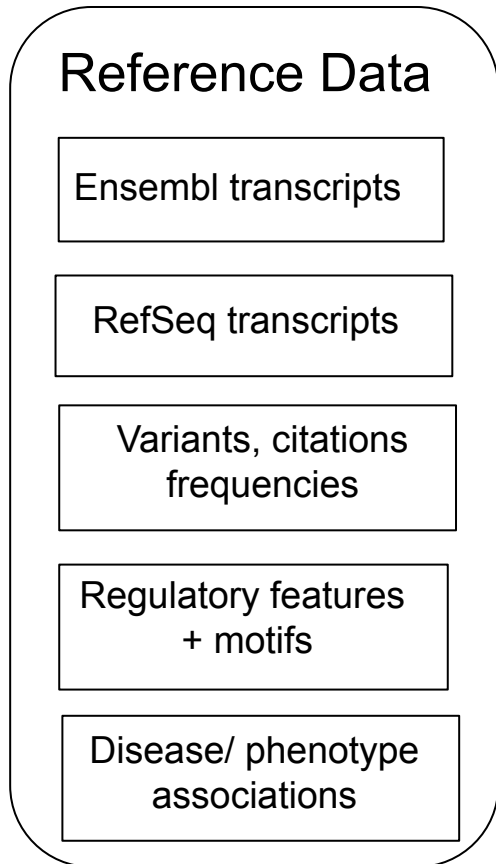
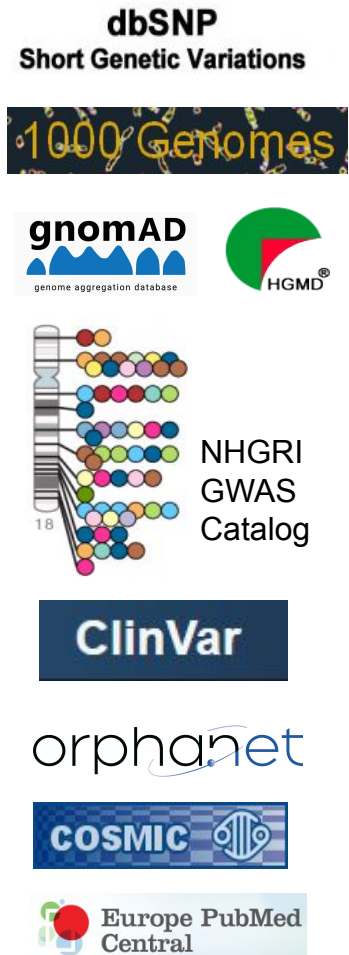
- Genomes and gene builds for >300 species
- Variation data
- Compara (alignments, gene trees, homologues)
- Regulatory builds
- BioMart (data export)
- Tools for data processing, e.g. Ensembl VEP
- Display your own data
- Programmatic access via APIs
- Completely Open Source (FTP, GitHub)



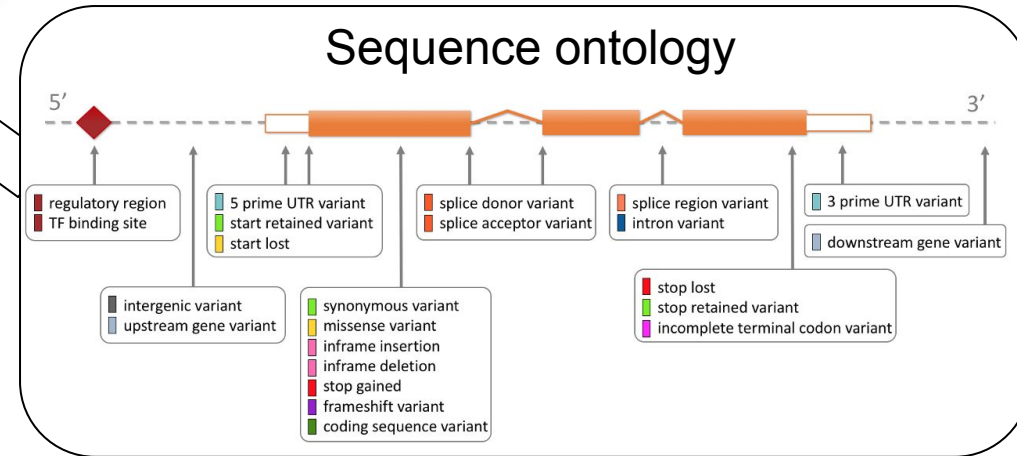
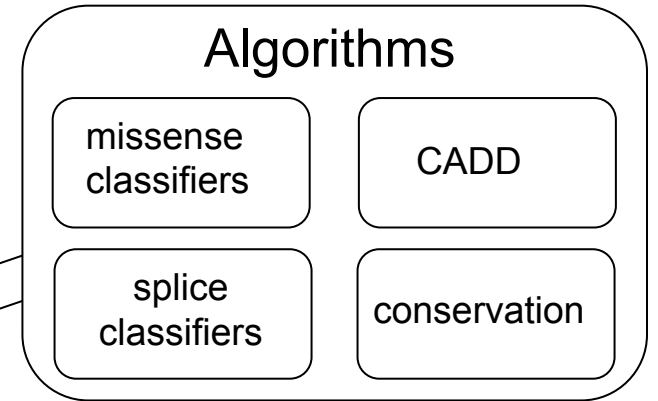
Ve!P



Ensembl Variant Effect Predictor (VEP)



Variants



| Consequence | Symbol | Feature | HGVSp | Codons | SIFT | PolyPhen | gnomAD AFR AF | gnomAD FIN AF | Clinical significance |
|------------------|--------|---------------------------------|-------------------------------|---------|------|----------|---------------|---------------|-----------------------|
| missense_variant | AGT | ENST00000366667 | ENSP00000355627.4:p.Thr207Met | ACG/ATG | 0.01 | 0.997 | 0.06221 | 0.1761 | likely_benign |

Acknowledgements



Funding



National Human
Genome Research
Institute (NHGRI)

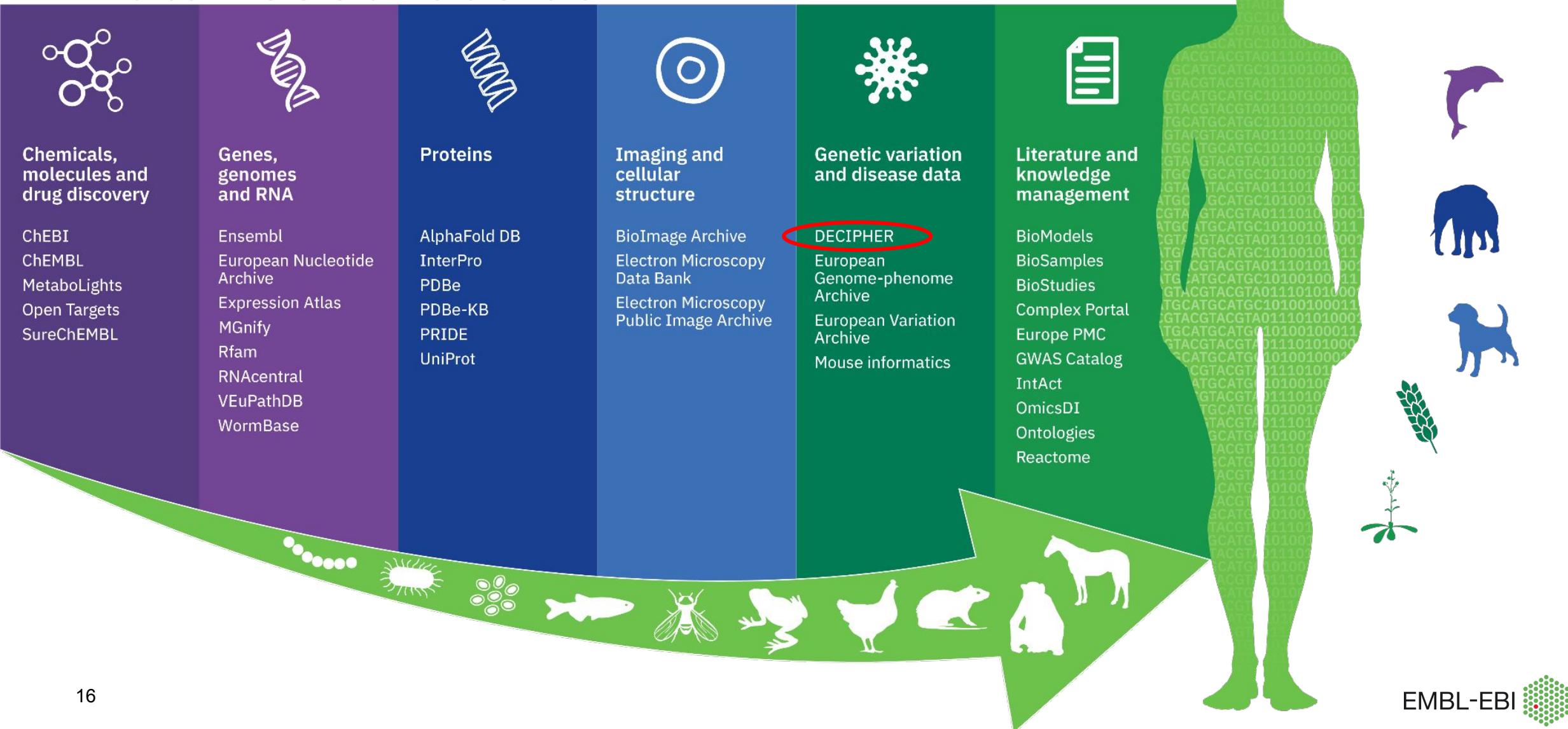
National Institute of
Allergy and Infectious
Diseases (NIAID)



Funded by the
European Union



Data resources at EMBL-EBI



DECIPHER is a web-based platform which helps **clinical** and **research teams** to **assess pathogenicity** and to **share** patient data

Help **integrate genotype & phenotype data** to determine the consequences of variants for health and disease

DECIPHER
GRCh38 Celebrating 20 years

DDD (UK) Search DECIPHER

<https://www.deciphergenomics.org>

Help Join Log in →

Mapping the clinical genome

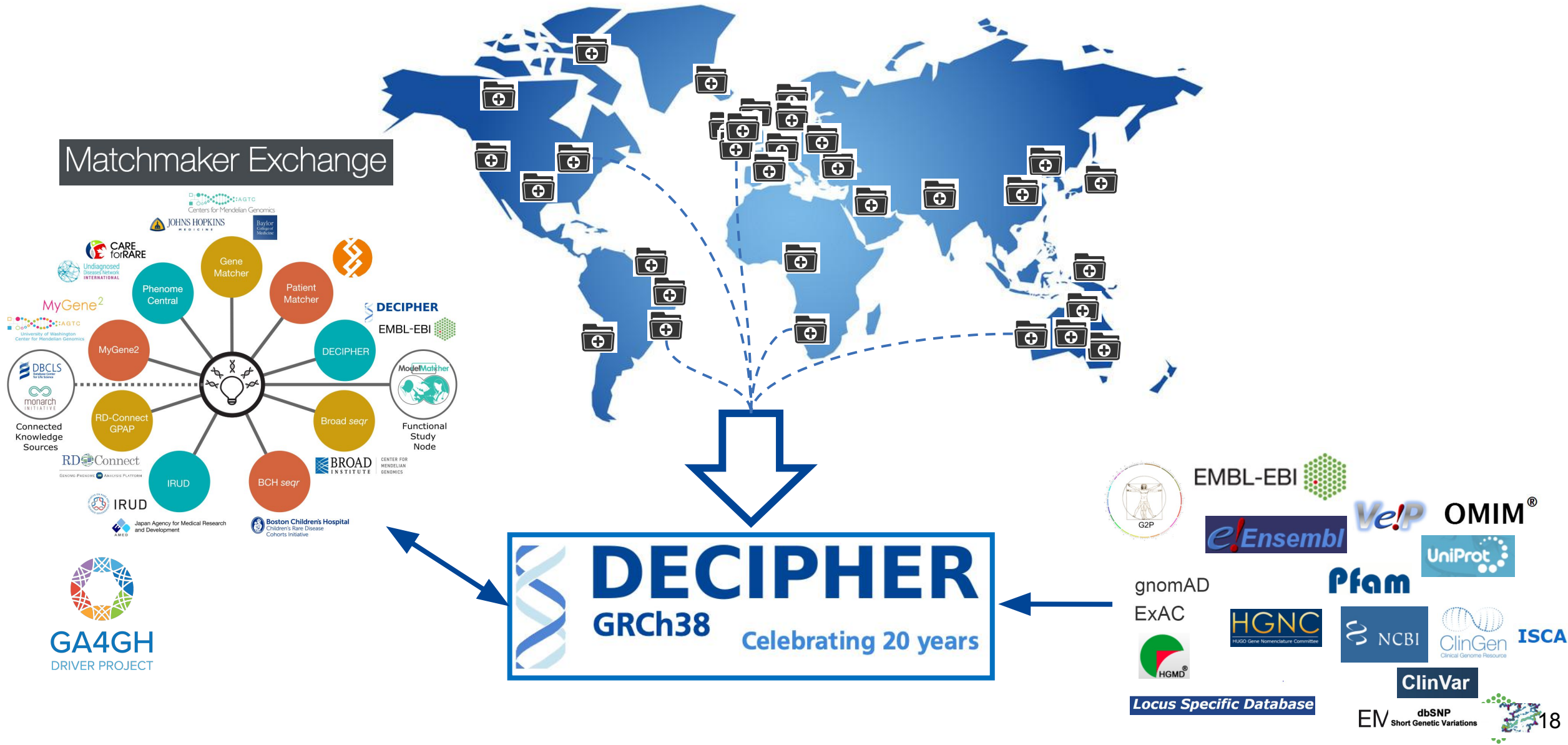
Explore DECIPHER
It's free and you don't need to log in
DECIPHER is used by the clinical community to share and compare phenotypic and genotypic data

Join DECIPHER
Be part of the sharing community
Projects affiliated to DECIPHER can deposit and share patients, variants, and phenotypes to invite

Already a member?
Log in to access your patient data

17

DECIPHER supports data sharing and linking



DECIPHER by the numbers

>49,000
anonymised,
open records

>59,000 variants

>197,000
phenotypes
(HPO)



>**3,500** publications citing DECIPHER

>320 depositing
centres

45 countries
with depositing
centres

7 consortia,
>73,000 patient
records

What can I do/see in DECIPHER?



Mapping the clinical genome

Explore DECIPHER

It's free and you don't need to log in

DECIPHER is used by the clinical community to share and compare phenotypic and genotypic data. The DECIPHER database contains data from 48,873 patients who have given consent for broad data-sharing; DECIPHER also supports more limited sharing via consortia. [Have a look at the numbers.](#)

Anyone can browse publicly-available patient data on DECIPHER and request to be put in contact with the responsible clinician. Why? [Because sharing benefits everyone.](#)

Join DECIPHER

Be part of the sharing community

Projects affiliated to DECIPHER can deposit and share patients, variants, and phenotypes to invite collaboration and facilitate diagnosis. Once deposited, you can use DECIPHER to identify and prioritise potential matches, and you can request notifications as soon as new matches arrive.

As well as influencing individual patient outcomes, use of DECIPHER has contributed to over [2600 published articles since 2004](#). It's still free, and you are in control of what data to make public.

Already a member?

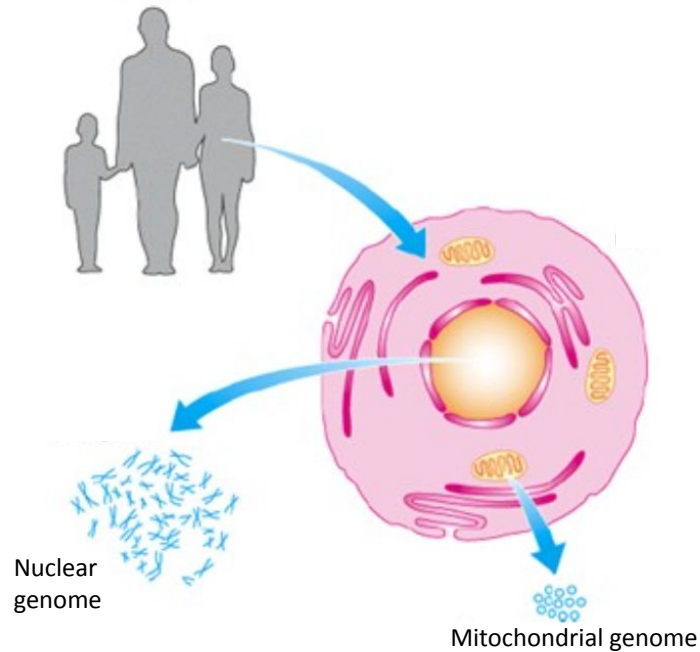
Log in to access your patient data

Email address

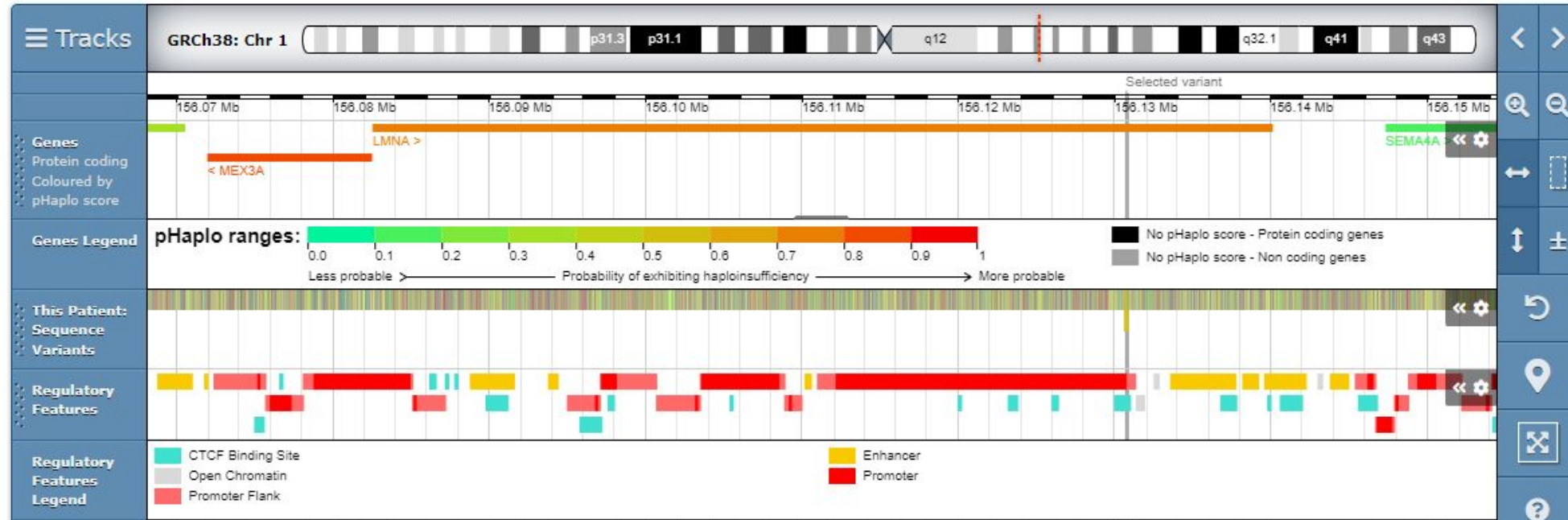
DECIPHER supports many types of genetic variation

| | | |
|--------------------------------|---------------------------------|-----------------------|
| Sequence Variant | Aneuploidy/Segmental aneuploidy | Insertion |
| Sequence Variant | Nullisomy | Mobile Element: Alu |
| Copy-Number Variant | Monosomy | Mobile Element: LINE1 |
| Deletion | Disomy | Mobile Element: SVA |
| Duplication | Trisomy | Retrogene |
| Duplication/Triplication | Tetrasomy | Other Insertion |
| Triplication | Uniparental Disomy | Short Tandem Repeat |
| Amplification (> Triplication) | Isodisomy | Short Tandem Repeat |
| | Heterodisomy | Inversion |
| | Unknown | Inversion |



| Location | Type | Genes | Size | Annotations | Inheritance / Genotype | Pathogenicity / Contribution | Shared | Links |
|--|----------------------------|-------|--|----------------------|--|---|-------------|--------------------------|
| 12 7045880 7045882 C A G Repeated > 50 times | Short Tandem Repeat | ATN1 | > 150 bp | Deposited 2020-06-17 | Unknown Heterozygous | Click here | Person icon | Show more, Share, Delete |
| 15 22749354 28438266 | Uniparental Isodisomy | 119 | 5.69 Mb | Deposited 2020-06-17 | Maternally inherited | Click here | Person icon | Show more, Share, Delete |
| 16 56370773 56370773 G > A | Sequence Variant | GNAO1 | SNV | Deposited 2020-10-07 | De novo Heterozygous | Likely pathogenic Full PM2 PP3 PS2 Last annotated: 2020-10-07 | Person icon | Show more, Share, Delete |
| 5 88100580 88100580 | Insertion of LINE1 element | MEF2C | 6.02 kb insertion Forward strand | Deposited 2020-06-17 | De novo Homozygous | Click here | Person icon | Show more, Share, Delete |
| 11 118374637 118374637 | Insertion | KMT2A | 450.00 kb insertion Sequence from 1:30000-480000 | Deposited 2020-06-17 | Unknown Homozygous | Click here | Person icon | Show more, Share, Delete |
| X 154156796 154227925 | Inversion | F8 | 71.13 kb | Deposited 2020-06-17 | De novo, mosaic Hemizygous | Click here | Person icon | Show more, Share, Delete |
| 6 156392359 158464899 | Duplication/Triplication | 11 | 2.07 Mb | Deposited 2020-06-17 | Dosage sensitivity score: 1.00 Sampling probability: > 5% | De novo, mosaic Homozygous | Person icon | Show more, Share, Delete |

DECIPHER supports visualisation on GRCh38



| Transcript | Protein change | VEP Consequence | Other annotations |
|---|---|------------------|--|
| ENST00000368300.9 - NM_170707.4 MANE Select Selected transcript ENST00000368300.9:c.490G>A | D/N at position 164 of 664 ENSP00000357283.4:p.Asp164Asn | missense_variant | Sift: Deleterious (0.02) PolyPhen: Benign (0.063) CADD: 25.8 REVEL: 0.55 SpliceAI: ≤ 0.2 |
| ENST00000677389.1 - NM_005572.4 MANE Plus Clinical ENST00000677389.1:c.490G>A | D/N at position 164 of 572 ENSP00000503633.1:p.Asp164Asn | missense_variant | Sift: Deleterious (0.01) PolyPhen: Benign (0.037) CADD: 25.8 REVEL: 0.55 SpliceAI: ≤ 0.2 |

DECIPHER supports granular phenotype display

Patient 6 Mother 0 Father 0

Patient phenotypes

Abnormality of head or neck

Cleft palate Dimple chin



Abnormality of limbs

Broad thumb



Abnormality of the nervous system

Delayed speech and language development Specific learning disability



Abnormality of the skeletal system

Clinodactyly of the 5th finger



Milestones Measurements

Social smile

Sat Independently

Walked Independently

First words

Growth

Height

✓ Weight

✓ Head circumference

Visual Function

Visual acuity

Area of preserved field

Area of scotoma

Flicker ERG timing

Fundus imaging

Area of atrophy (AF)

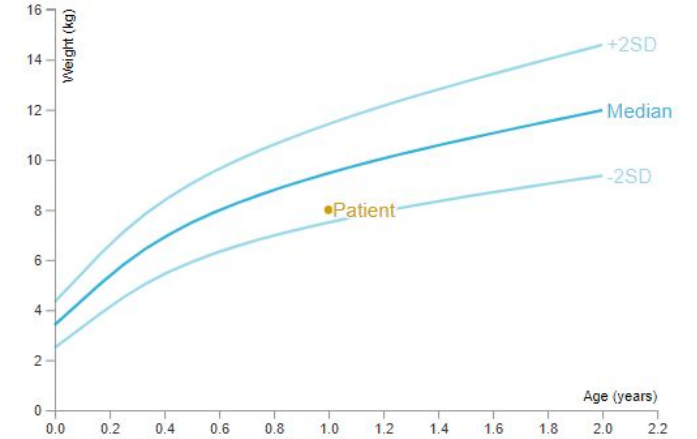
Area within hyperAF ring

Central subfield thickness

Macular thickness

Ellipsoid zone length (H)

Ellipsoid zone length (V)



Weight

| Age | Value | Percentile | SD |
|--------|-------|------------|----------|
| 1 year | 8 kg | 5 | -1.63 SD |

DECIPHER supports protein structure visualisations

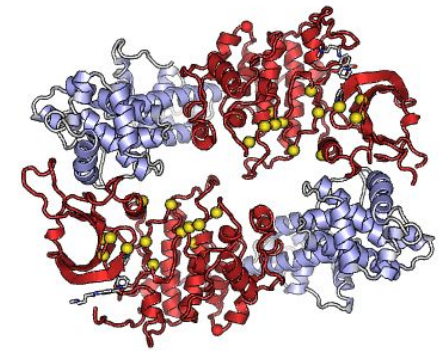
Protein: Q14004

Links Tracks and filters

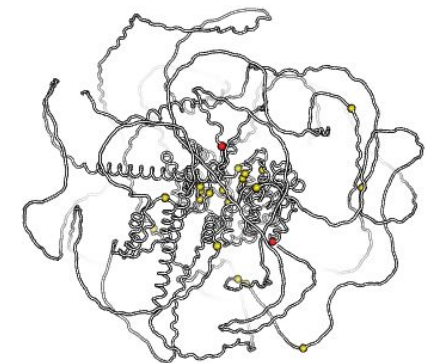
Transcript used in protein view: ENST00000181839.10 14 exons, 1512aa **MANE Select**



PDBe



AlphaFold



DECIPHER supports Ensembl Variant Effect Predictor

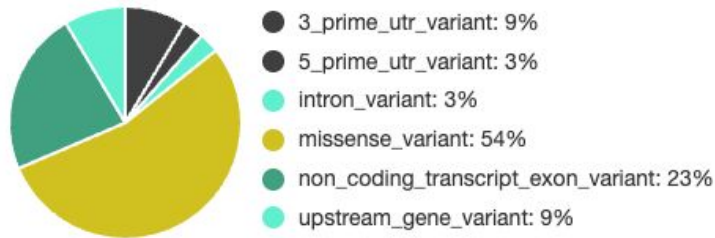
Sequence Variant LMNA 1:156130750 G > A

Browser Gene Protein **Annotation** Matching patient variants **41** Matching DDD research variants **0** Pathogenicity evidence

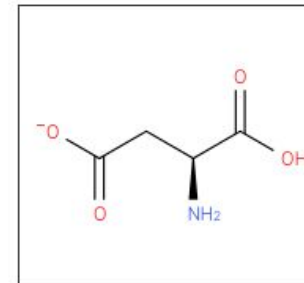
Consequence prediction (VEP) ClinVar Disease cohorts Allele frequency Functional

Ensembl Variant Effect Predictor (VEP) ?

Consequences



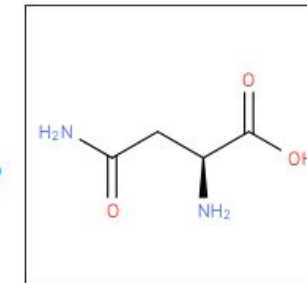
Amino acid substitution



Aspartic acid (D)

Negative, polar, hydrophilic

→
Grantham distance ?
Conservative (23)



Asparagine (N)

Uncharged, polar, hydrophilic

Images generated by ChEBI

Annotations for LMNA: 1 to 35 of 35

Filter...

| Transcript | Protein change | VEP Consequence | Other annotations |
|--|---|-------------------------|---|
| ENST00000368300.9 - NM_170707.4 MANE Select Selected transcript | D/N at position 164 of 664 ENSP00000357283.4:p.Asp164Asn | missense_variant | Sift Deleterious (0.02) PolyPhen Benign (0.066) CADD 25.8 REVEL 0.55 AlphaMissense Likely Benign (0.2051) SpliceAI ≤ 0.2 phyloP 10.003 |
| NM_001406991.1 NM_001406987.1 NM_001406986.1 NM_001406985.1 NM_001406983.1 ENST00000368300.9:c.490G>A | | | |

DECIPHER supports matching patients

GRIN2B 12:13437942-13982002

Reverse strand gene: glutamate ionotropic receptor NMDA type subunit 2B

DECIPHER holds 60 sequence variants in this gene, in 60 open-access patients

Overview **Matching patient variants 131** Matching DDD research variants 0 Phenotypes Phenotype browser Transcripts 10 Browser

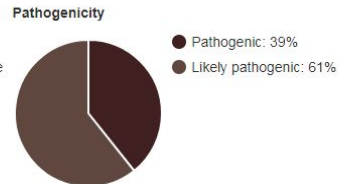
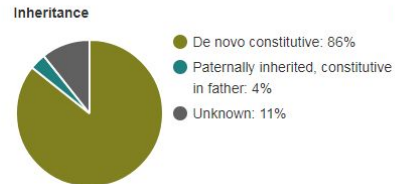
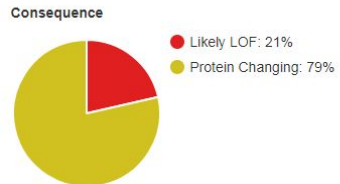
Sequence variants 60 Copy-number variants 60 Other variants 11

Patients with sequence variants matching this gene

Filters: Functional Similarity Sex Consequence **Pathogenicity** Contribution Inheritance Genotype Phenotypes Enter text...

Displaying data for 28 out of 45 variants, filtered by: [Clear all filters](#)

Pathogenicity [Pathogenic](#) [Likely pathogenic](#)



Phenotypes present in multiple matching patients

- 12 Global developmental delay
- 7 Severe global developmental delay
- 5 Seizures
- 4 Delayed speech and language development
- 3 Cerebral visual impairment
- 3 Generalized hypotonia
- 3 Intellectual disability
- 3 Strabismus
- [Show more](#)

This patient's phenotypes present in matching patients

- 12 Global developmental delay
- 5 Seizures
- 3 Intellectual disability
- 1 Hypertelorism

This patient's phenotypes absent in matching patients

- Microcephaly
- Short stature

| Patient | Sex | Transcript / Location (GRCh37) | Functional Similarity | Consequence | Pathogenicity / Contribution | Inheritance / Genotype | Shared Phenotypes |
|---------|------|--|-----------------------|--|------------------------------|---|--|
| 262325 | 46XY | ENST00000609686 c.2459G>C 12:13720098-13720098 | 0.70 | missense_variant (820 G/A) p.Gly820Ala | Pathogenic Full | De novo constitutive Heterozygous | 1 of 3: Global developmental delay ; Delayed CNS myelination; Neonatal hypotonia |
| 277140 | 46XX | ENST00000609686 c.2065G>A 12:13724844-13724844 | 0.70 | missense_variant (689 G/S) p.Gly689Ser | Pathogenic Full | De novo constitutive Heterozygous | 0 of 10: Absent speech; Cerebral visual impairment; Dysphagia; Exaggerated startle response; Gastrostomy tube feeding in infancy; Infantile axial hypotonia; Myoclonus; Postnatal microcephaly; Severe global developmental delay; Widely spaced teeth |

- Inform diagnosis
- New phenotypes associated with disease
- New variant-disease pairings

DECIPHER links to disease management resources

ATP7B 13:51930436-52012125

Reverse strand gene: ATPase copper transporting beta

Formerly known as: **WND**

Also known as: **ENSG00000123191**

Function: "Copper ion transmembrane transporter involved in the export of copper from hepatocytes into the bile in response to copper overload." Source: UniProt

DECIPHER holds 4 sequence variants in this gene, in 4 individuals

Overview Matching patient variants **70** Matching DDD research variants

Clinical Protein / Genomic

Gene/disease association

Gene2Phenotype ?

-

OMIM ?

606882

Morbid ?

Wilson disease (Autosomal recessive)

GeneReviews ?

- Hereditary Dystonia
- Wilson Disease

ClinGen gene/disease ?

- Wilson disease **Definitive; AR**

ClinGen Dosage Sensitivity ?

- Haploinsufficiency: Gene associated with autosomal recessive phenotype (30)
- Triplosensitivity: No evidence for dosage pathogenicity (0)

GenCC ?

Definitive: **1**

Strong: **2**

Supportive: **1**

(Assessed by ClinGen, Genomics England PanelApp, Orphanet, Invitae, with respect to Autosomal recessive inheritance)

ClinGen Clinical Actionability

ATP7B has been curated by ClinGen, with respect to clinical actionability.

Filter...

| Disease | MONDO ID | Status / Consensus Assertion | Context | Report | Outcome / Intervention |
|----------------|---------------|-------------------------------------|-----------|------------------|---|
| Wilson Disease | MONDO:0010200 | Complete / Strong Actionability | Adult | 2022-02-14 AC028 | Morbidity and mortality from copper deposition / Referral to specialist for evaluation and to guide appropriate therapy including diet, copper chelation and zinc therapy |
| Wilson Disease | MONDO:0010200 | Complete / Definitive Actionability | Pediatric | 2022-02-14 AC028 | Morbidity and mortality from copper deposition / Copper chelation and zinc therapy |

LOEUF **1.19** ?

sHet **0.004** ?

pHaplo **0.16** ?

pTriplo **0.88** ?

- Gene Tests
- Genomics England PanelApp ?
- LSDB
- Entries in DECIPHER for this gene
- Open Targets Platform

Interpretation and management resources

Treatable ID ?

- Copper-transporting ATPase subunit beta deficiency (synonym: Wilson disease) (AR)

IEMbase ?

- Copper-transporting ATPase subunit beta deficiency (Autosomal recessive)

ACMG Secondary Findings ?

- Wilson disease (AR)

ClinGen Clinical Actionability Report ?

- Wilson Disease

DECIPHER: Data integration for Genomic Medicine

Management resources

Treatable ID

- Ionotropic glutamate receptor NMDA type subunit 1 dysregulation (AD)

IEMBase

- Ionotropic glutamate receptor NMDA type subunit 1 dysregulation (Autosomal dominant, Autosomal recessive)

neXtProt

MSH2-iso1-p.Gly322Asp - See data in neXtProt

| Effect | Intensity | Quality | Subject protein origin |
|--|-----------|---------|------------------------|
| Decreases mismatch repair | Severe | Gold | Homo sapiens |
| Decreases protein abundance | Moderate | Gold | Homo sapiens |
| Does not cause phenotype increased cellular sensitivity to alkylating agents | No impact | Gold | Non homo sapiens |
| Has no impact on binding to EXO1 | No impact | Silver | Non homo sapiens |

Patient: 265083

Overview Genotype Phenotypes Assessments Karyotype Citations Contact

GRIN1 mutation associated with intellectual disability alters NMDA receptor trafficking and function.

Chen W, Shieh C, Swanger SA, Tankovic A, Au M, McGuire M, Tagliati M, Graham JM, Madan-Khetarpal S, Traynelis SF, Yuan H, Pierson TM.

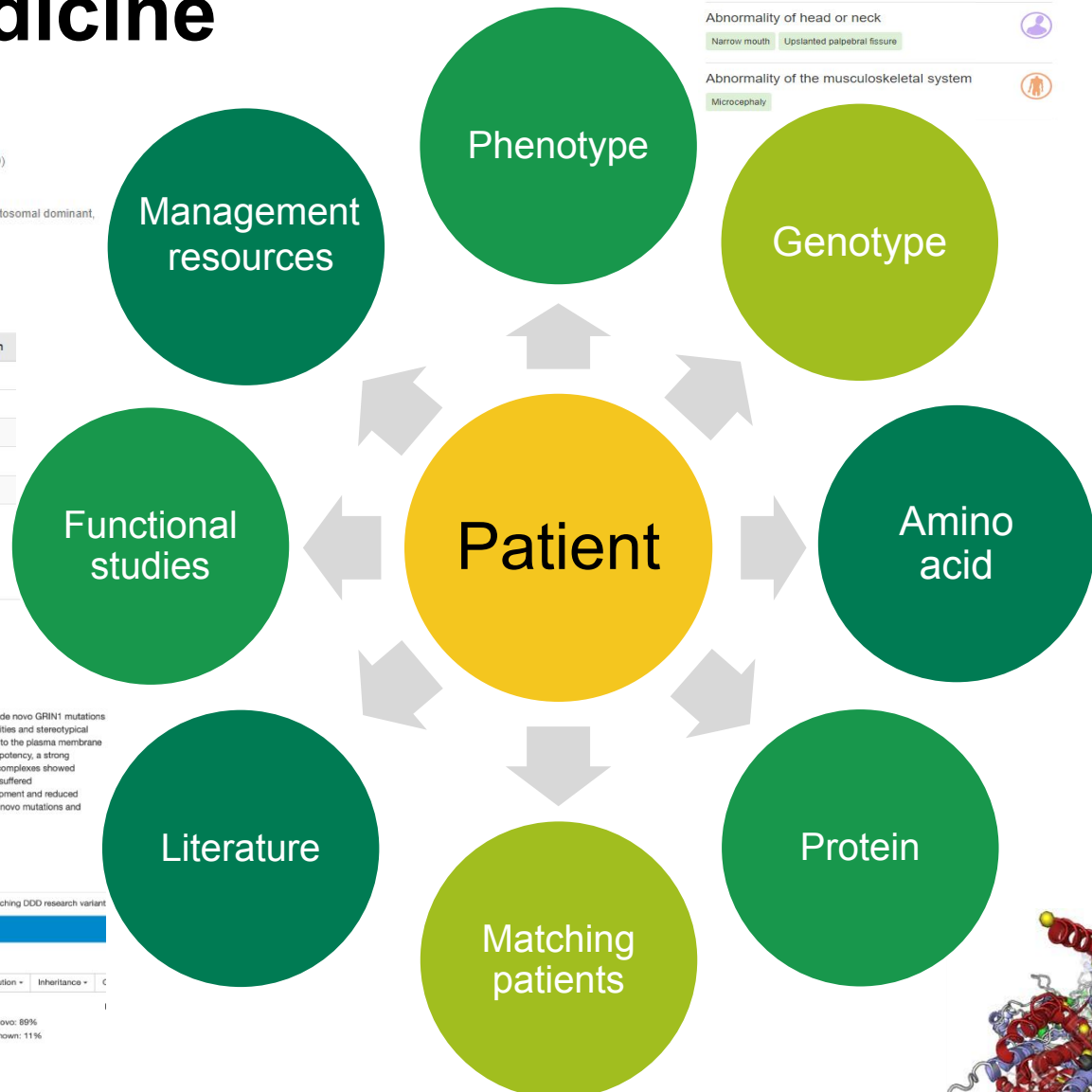
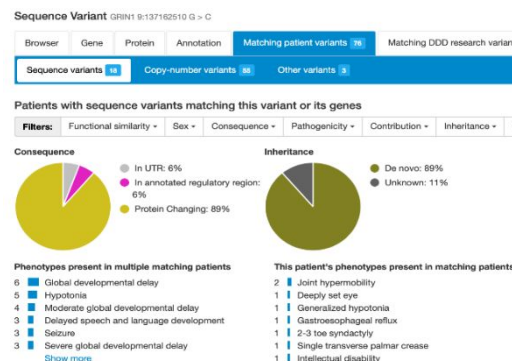
Department of Pharmacology, Emory University School of Medicine, Atlanta, GA, USA.

NCATS NIH HHS: UL1 TR000454, NICHD NIH HHS: R01 HD082373, NINDS NIH HHS: R01 NS036654, R24 NS092989

Journal of human genetics 2017; 82: 6: 589-597

N-methyl-D-aspartate receptors (NMDARs) play important roles in brain development and neurological disease. We report two individuals with similar dominant de novo GRIN1 mutations (c.1858 G>A and c.1858 G>C, both p.G620R). Both individuals presented at birth with developmental delay and hypotonia associated with behavioral abnormalities and stereotypical movements. Recombinant NMDARs containing the mutant GluN1-G620R together with either GluN2A or GluN2B were evaluated for changes in their trafficking to the plasma membrane and their electrophysiological properties. GluN1-G620R/GluN2A complexes showed a mild reduction in trafficking, a ~2-fold decrease in glutamate and glycine potency, a strong decrease in sensitivity to Mg²⁺ block, and a significant reduction of current responses to a maximal effective concentration of agonists. GluN1-G620R/GluN2B complexes showed significantly reduced delivery of protein to the cell surface associated with similarly altered electrophysiology. These results indicate these individuals may have suffered neurodevelopmental deficits as a result of the decreased presence of GluN1-G620R/GluN2B complexes on the neuronal surface during embryonic brain development and reduced current responses of GluN1-G620R-containing NMDARs after birth. These cases emphasize the importance of comprehensive functional characterization of de novo mutations and illustrates how a combination of several distinct features of NMDAR expression, trafficking and function can be present and influence phenotype.

PUBMED: 28228639; PMC: 5637823; DOI: 10.1038/jhg.2017.19



Patient 6 Mother 0 Father 0

Phenotype phenotypes

Growth abnormality

Obesity

Abnormality of head or neck

Narrow mouth Upslanted palpebral fissure

Abnormality of the musculoskeletal system

Microcephaly

9 137162510 137162510 Sequence Variant GRIN1 SNV missense_variant

GRCh38 G > C

ENST00000371561.8

c.1858G>C

p.Gly620Arg (620 G/R)

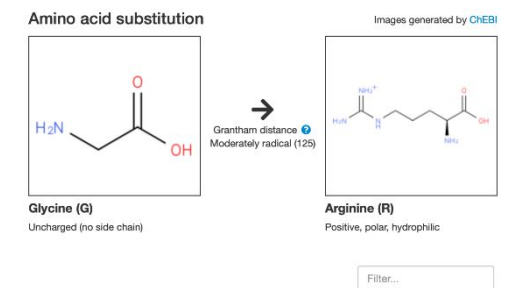
Not in gnomAD: see coverage

Sequence Variant GRIN1 9:137162510 G>C

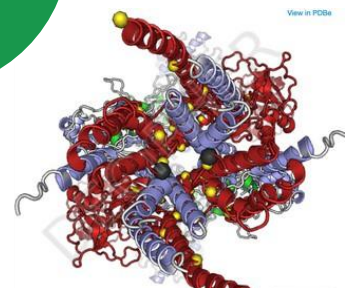
Browser Gene Protein Annotation Matching patient variants Matching DDD research variants

Consequence prediction (VEP) ClinVar Allele frequency Functional

- ### ClinVar assertions for 9:137162510 G > C
- 1★ Likely pathogenic for Neurodevelopmental disorder with or without hyperkinetic movements and seizures, autosomal dominant
 - 1★ Pathogenic for not provided
 - 1★ Likely pathogenic for Inborn genetic diseases
 - 1★ Pathogenic for Seizures



| VEP Consequence | Other annotations |
|------------------|--------------------------------|
| missense_variant | Sift Deleterious (0) |
| | PolyPhen Probably damaging (1) |
| | CADD 31 |
| | REVEL 0.909 |
| | SpliceAI <= 0.2 |



Acknowledgements: DECIPHER

- ★ Patients and families for permission to include their data in DECIPHER
- ★ Members of the global DECIPHER community
- ★ Resources which DECIPHER uses

Website: <https://www.deciphergenomics.org>

X/Twitter: @deciphergenomic

Email: contact@deciphergenomics.org

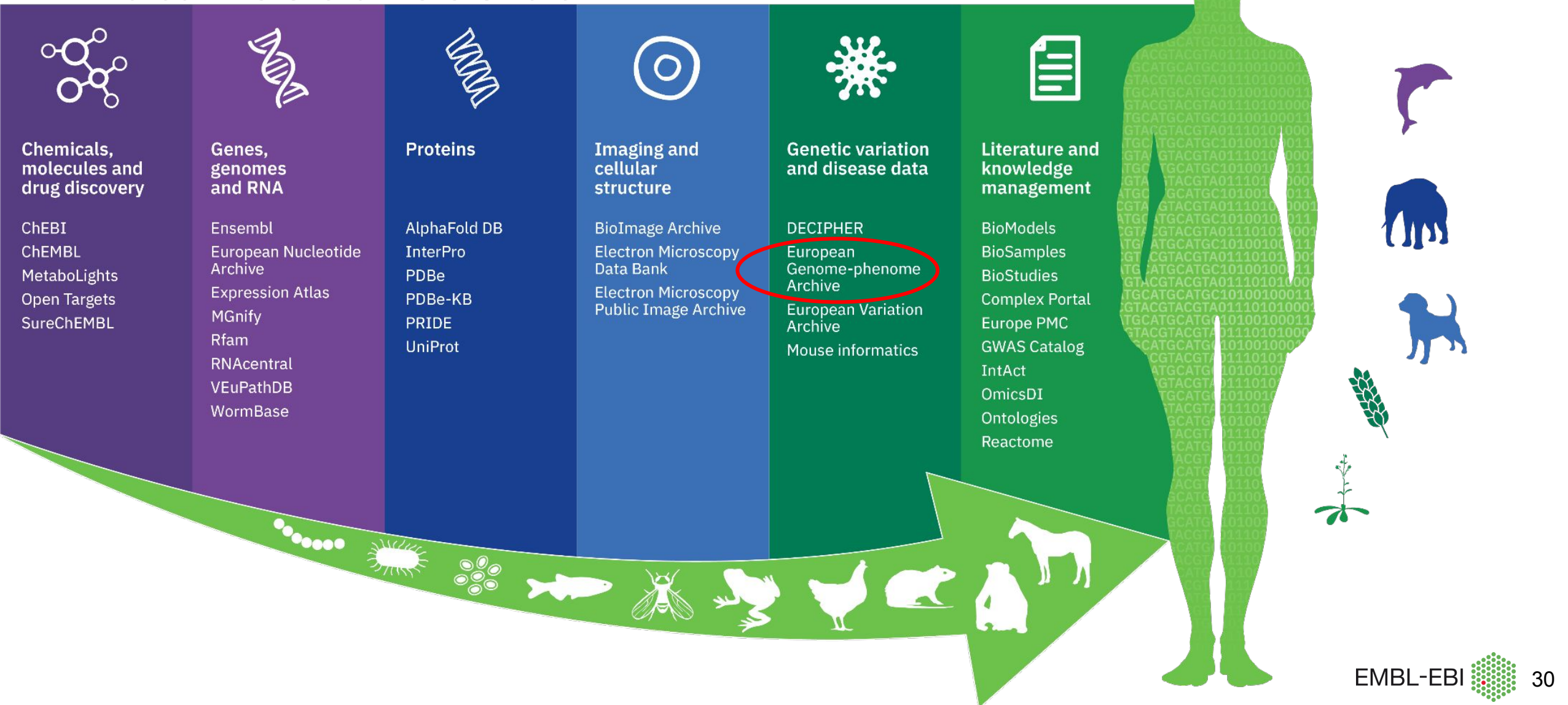


DECIPHER Leadership & Core Team



Recent alumni: Daniel Perrett, Simon Brent, Ben Hutton

Data resources at EMBL-EBI



Sharing data about humans requires care



Exposure of personal or sensitive data is **potentially harmful**

Individuals have the **right to protection** of their personal data

Governments pass legislation that **codifies data protection rights**

Support aims of research & health care initiatives

EGA manages sensitive human data



Mission: Permanent archiving and sharing of personally identifiable genetic, phenotypic, and clinical data generated for the purposes of biomedical research projects or in the context of research-focused healthcare systems



[Login](#) [Register](#) [Need Help?](#)

[ABOUT](#) [DISCOVERY](#) [SUBMISSION](#) [ACCESS](#)

Search...



GA4GH
DRIVER PROJECT



I want to...



<https://ega-archive.org/>

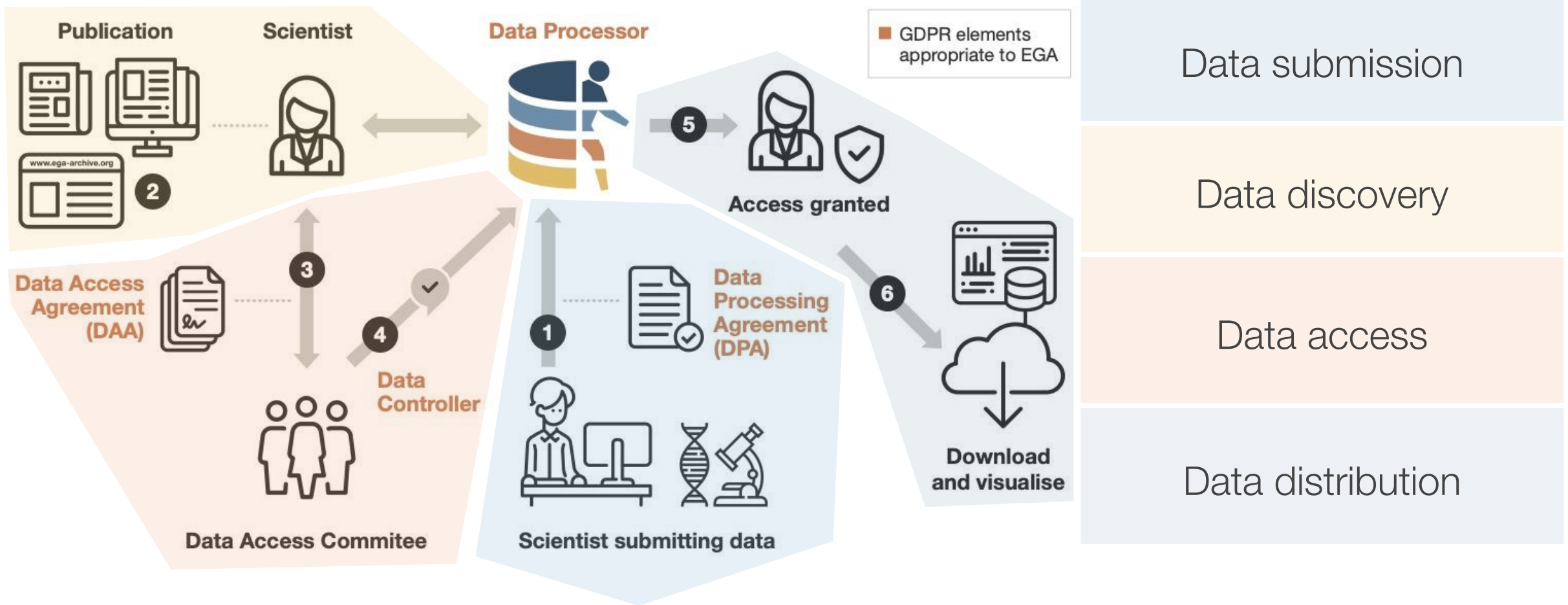


2008: EGA launched at EMBL-EBI

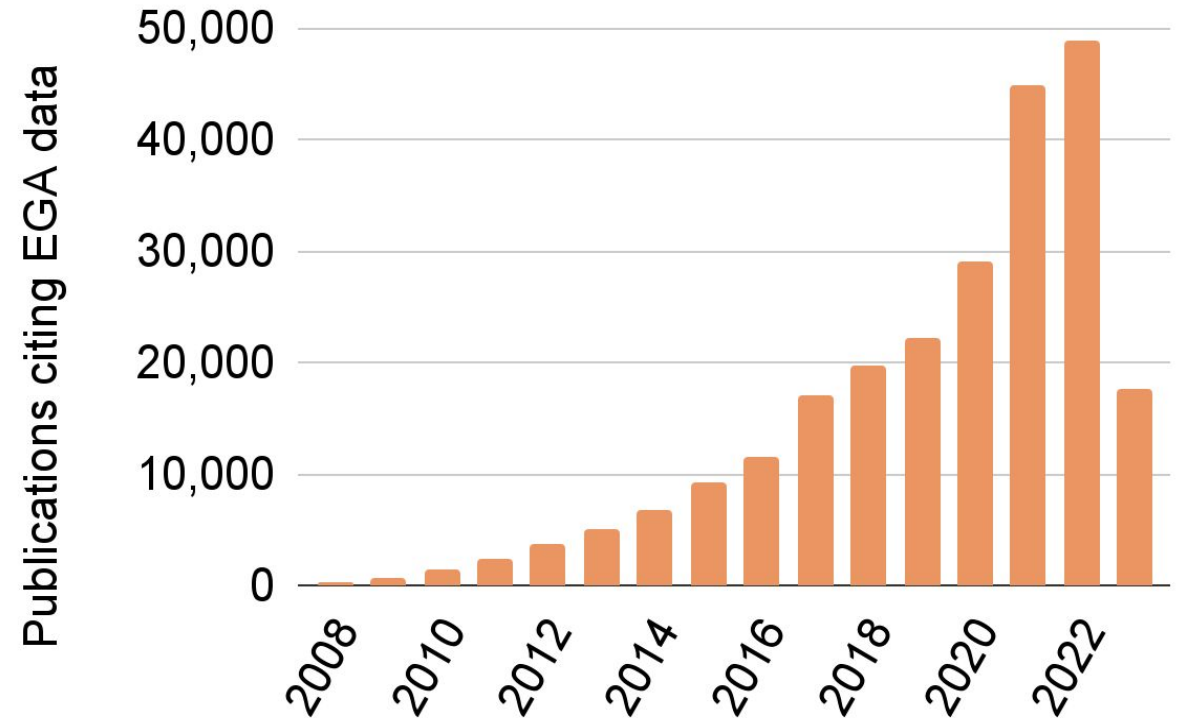
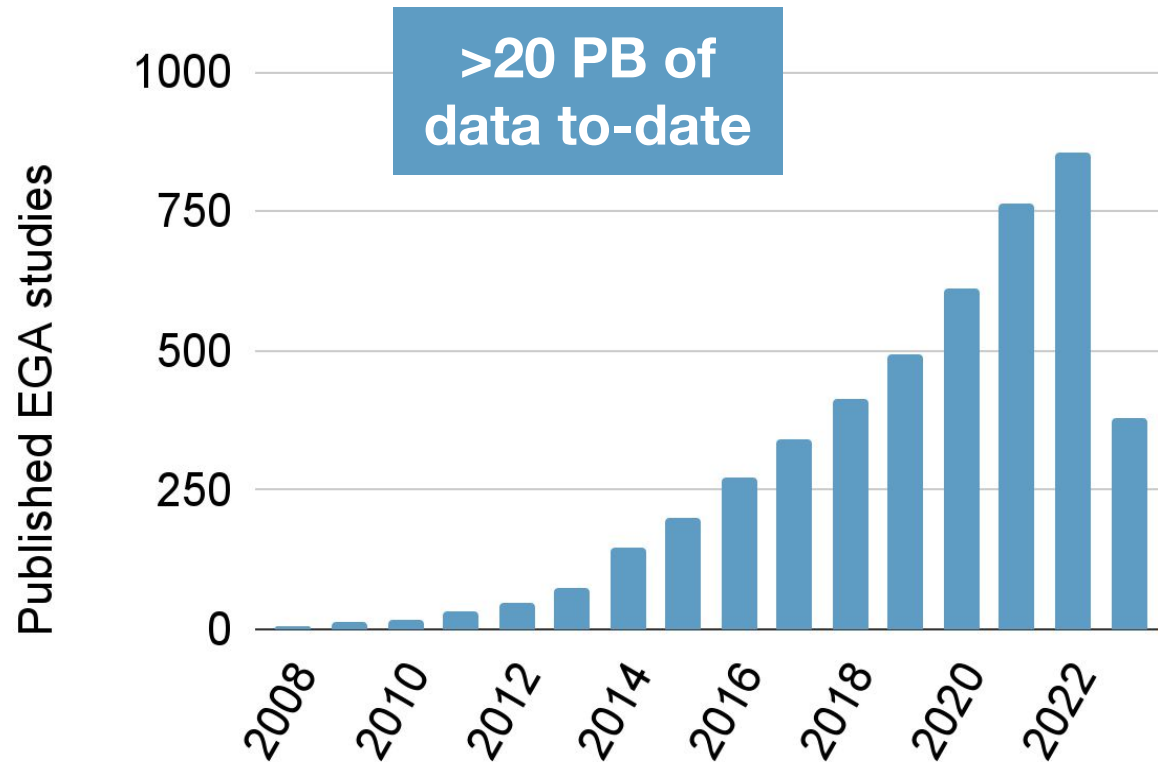


2013: CRG & EMBL-EBI enter formal collaboration to co-manage EGA

EGA data access model



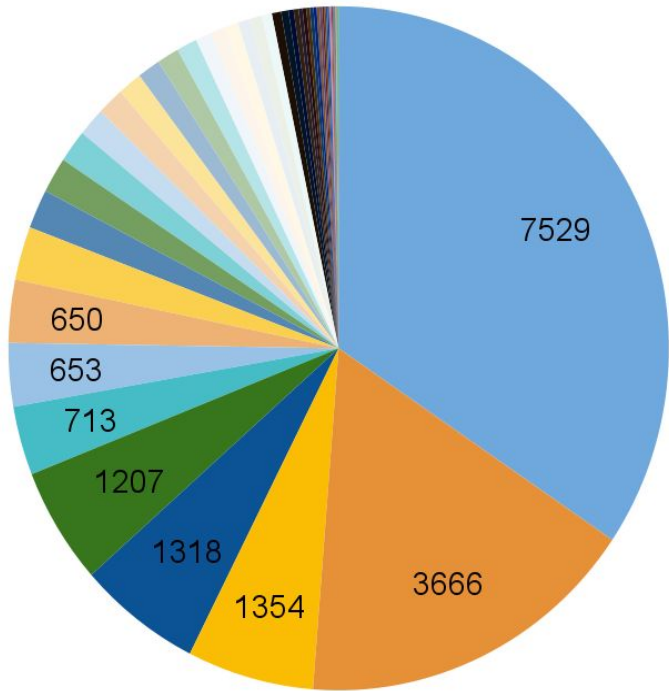
EGA grows and remains key to life science research



Data deposited and requested by global community

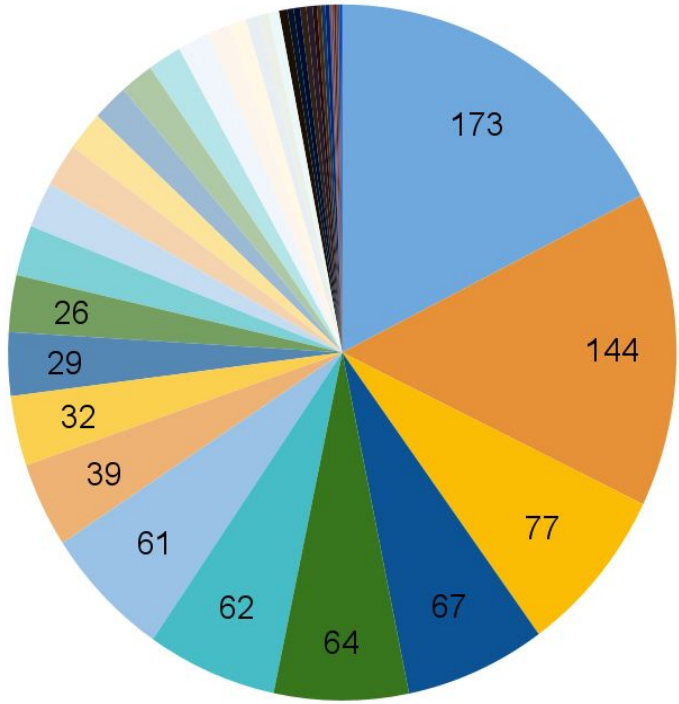
Data requesters by country

- United States
- United Kingdom
- China
- Canada
- Germany
- Australia
- Spain
- Netherlands
- France
- Sweden
- Korea, Republic of
- Italy
- 45 more



Data submitters by country

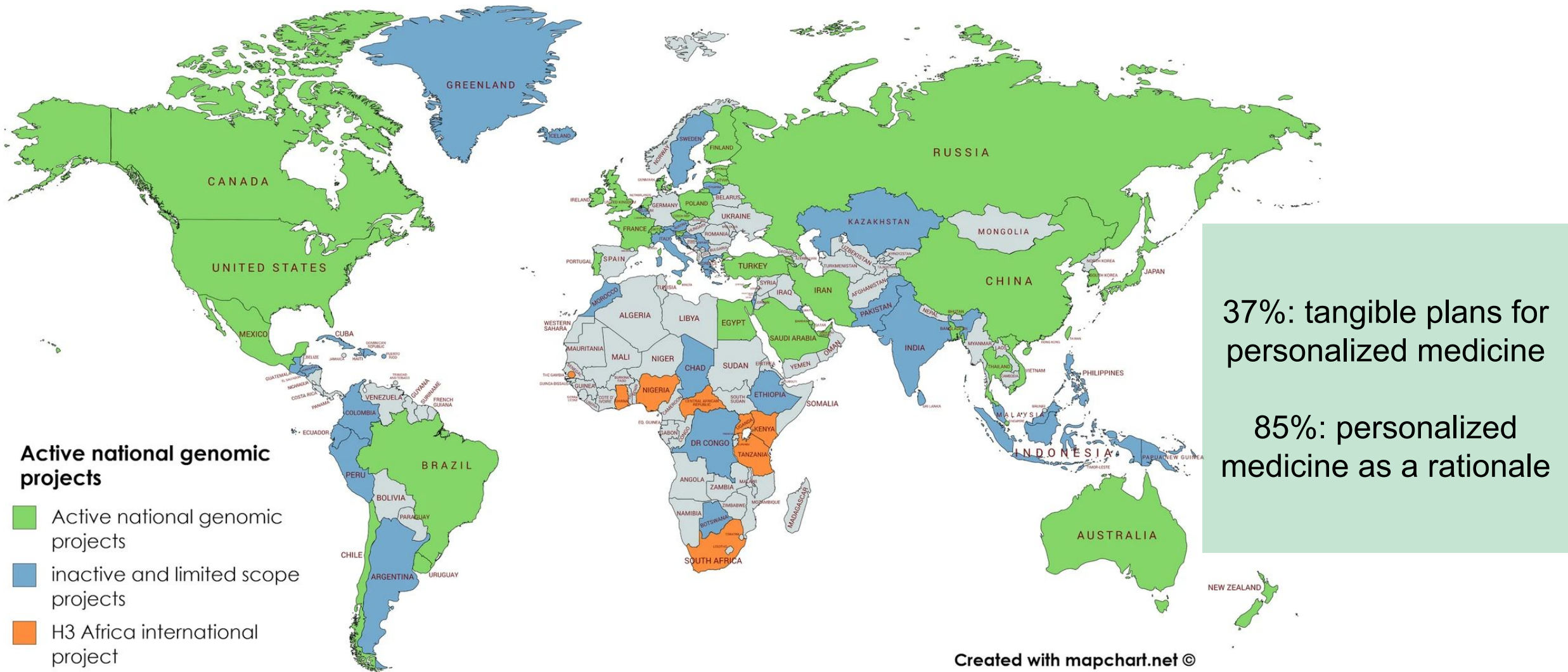
- United Kingdom
- USA
- Germany
- China
- Canada
- France
- Netherlands
- Spain
- Australia
- Sweden
- Japan
- Finland
- 27 more



Last Updated : November 2023

Image credit :Aina Jene Cortada

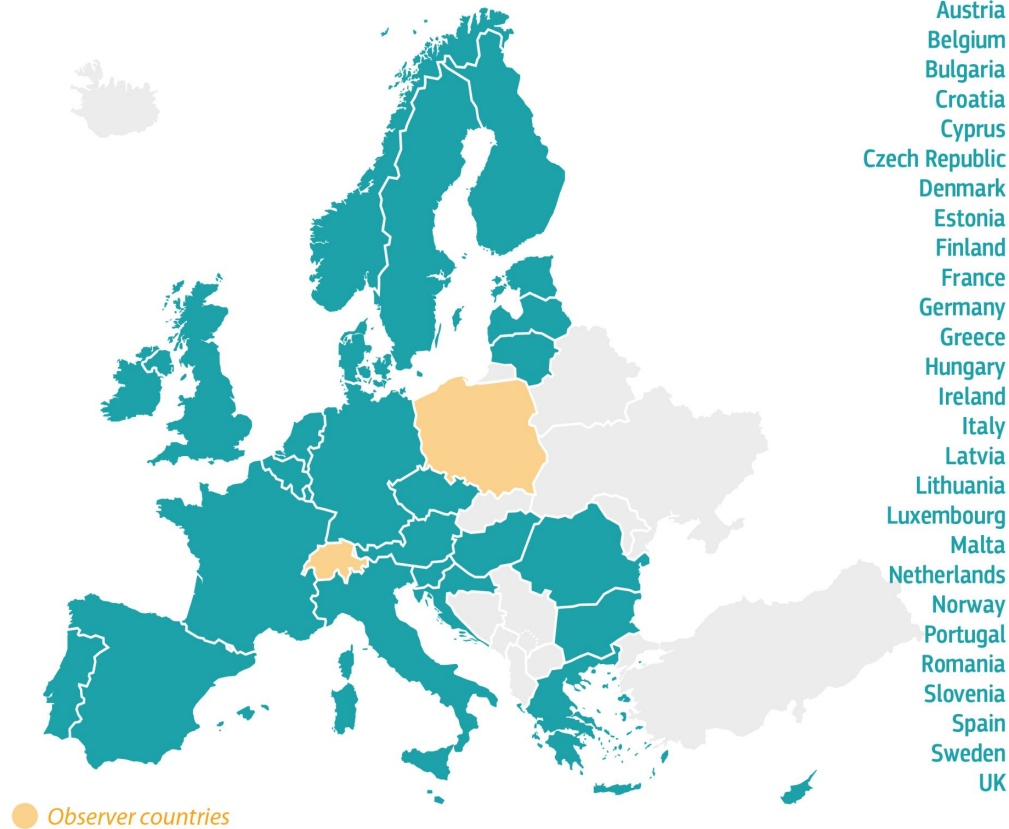
More genomics data generated in healthcare context



European 1+MG initiative sets vision for data sharing



Cross-border access to genomic data, implementation of genomics-based health. Supported by 27 countries to date



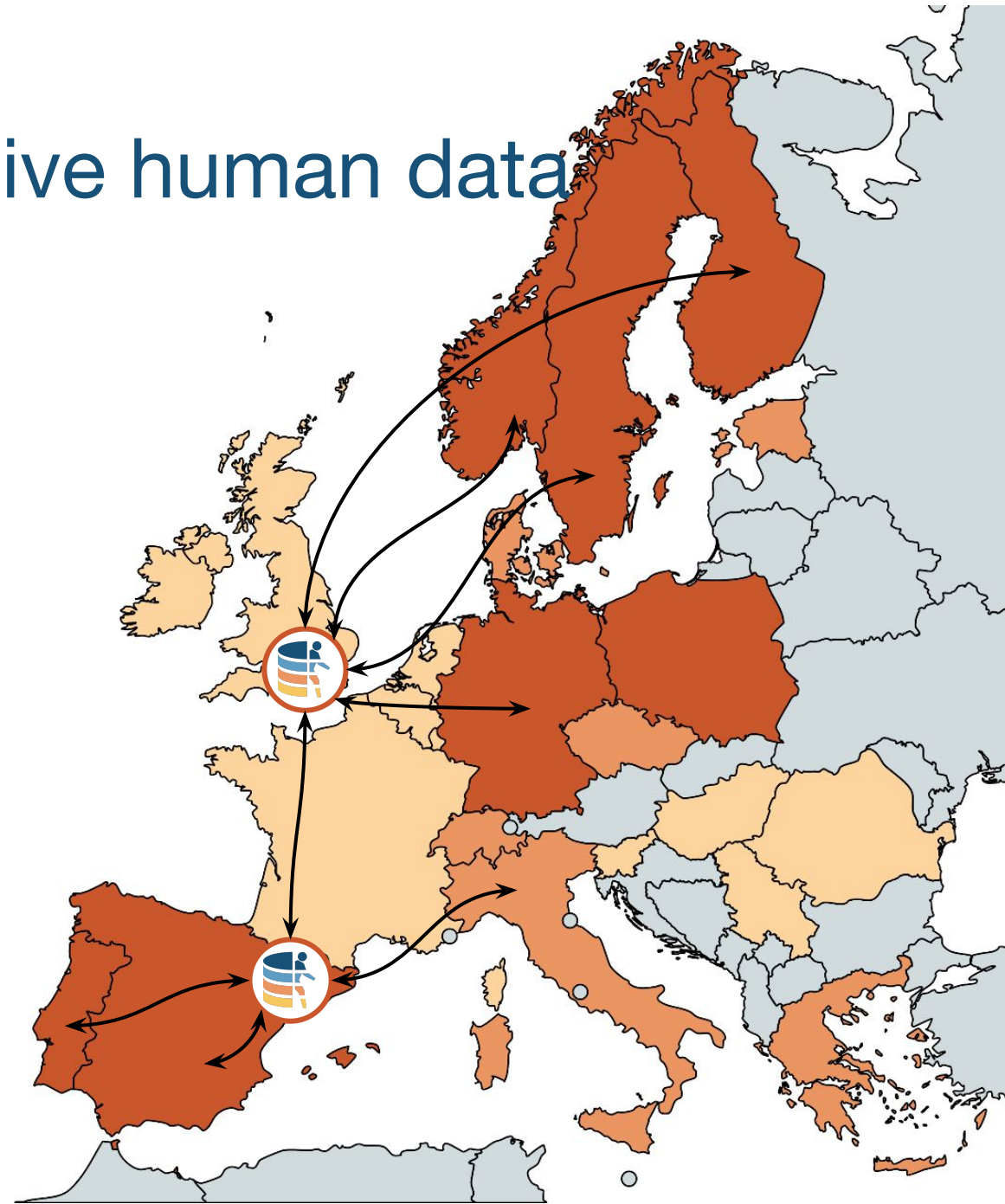
- Austria
- Belgium
- Bulgaria
- Croatia
- Cyprus
- Czech Republic
- Denmark
- Estonia
- Finland
- France
- Germany
- Greece
- Hungary
- Ireland
- Italy
- Latvia
- Lithuania
- Luxembourg
- Malta
- Netherlands
- Norway
- Portugal
- Romania
- Slovenia
- Spain
- Sweden
- UK

Ambition

Enable secure access to high-quality genomics and the corresponding clinical data across Europe for better research, personalised healthcare and health policy making

Federated EGA: discovery & access for sensitive human data







Federated EGA strives to support the discovery of and secure access to human data globally, while respecting national data protection regulations, with the goal of accelerating disease research and understanding and improving human health.



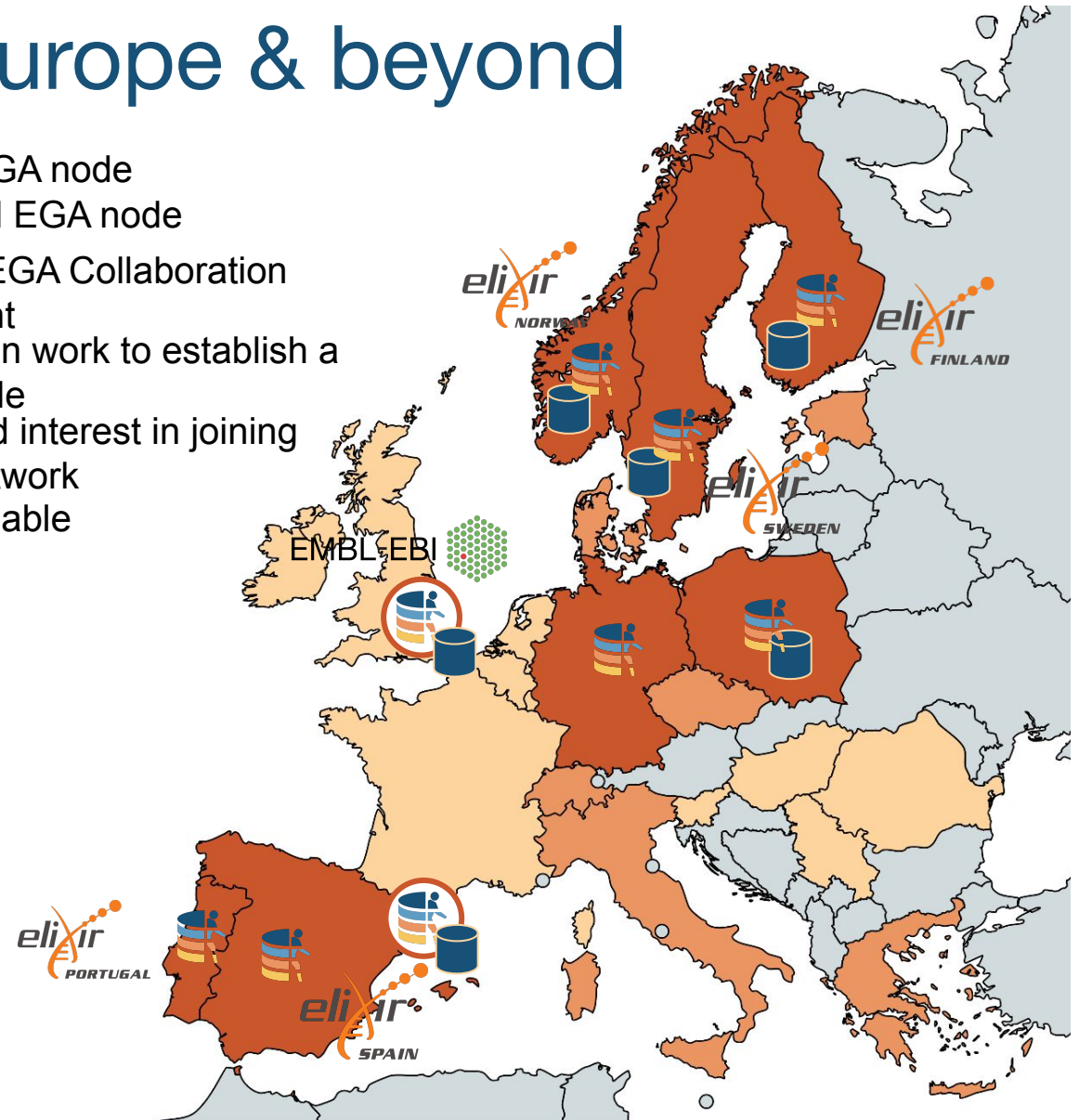
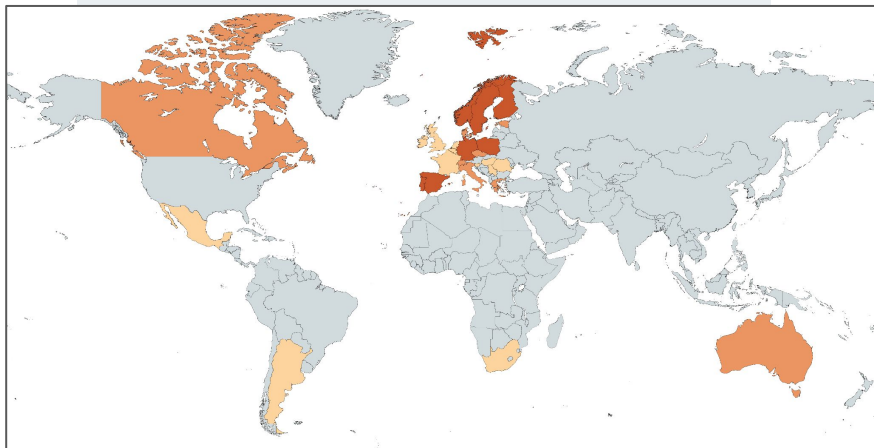
Federated EGA collaboration officially launched in September 2022



FEGA engagement across Europe & beyond

-  Central EGA node
-  Federated EGA node
-  Signed FEGA Collaboration Agreement
-  Engaged in work to establish a FEGA node
-  Expressed interest in joining FEGA Network
-  Data available

>24 nodes engaged and counting!



Standards are key to interoperability



Global
for Ge
Collabor

Standards

The EGA is a long-standing supporter of the [Global Alliance for Genomics and Health](#) (GA4GH) and has contributed to the development and implementation of GA4GH standards and APIs that are currently used by Driver Projects and has contributed to the development and implementation of GA4GH standards and APIs that are currently used by Driver Projects.

Below is a list of the GA4GH standards and APIs that are currently used by Driver Projects.

| Technical Standards | Purpose |
|----------------------------------|--|
| Large Scale Genomics | |
| htsget | A protocol for secure, efficient, and reliable access to sequencing read and variant data. |
| Read File Formats (SAM/BAM/CRAM) | Specifications for storing next-generation sequencing read data. |
| Variation File Formats (VCF/BCF) | The specifications for Variant Call Format (VCF) and its binary counterpart BCF. |
| Crypt4GH v1.0 | Enables direct byte-level compatible random access to encrypted genetic data stored in community standards (e.g. CRAM, VCF). |

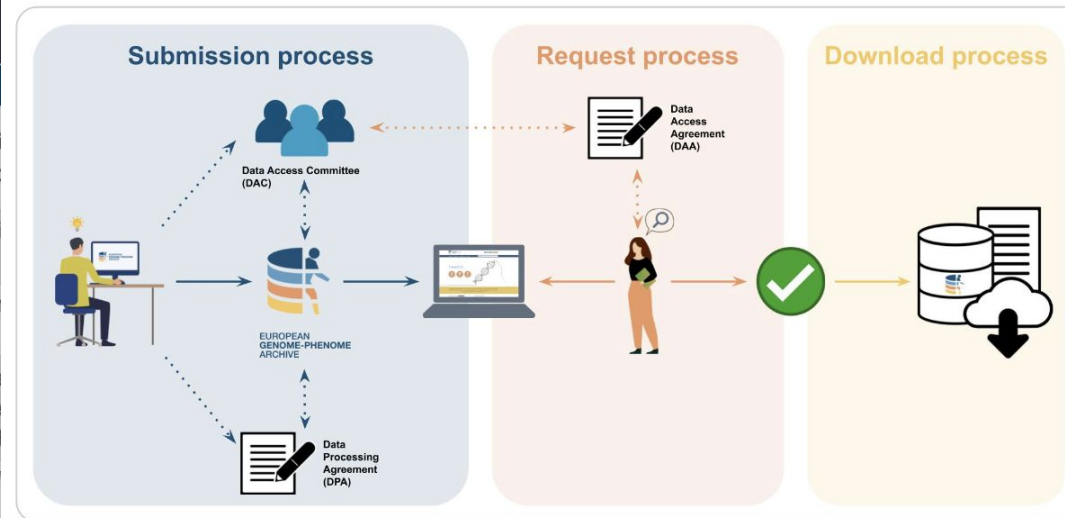
Our community | Driver Projects

European Genome-phenome Archive (EGA)

Provides a secure service for storing and sharing human genetic, phenotypic, clinical, and other “omics” data for research projects

The EGA is a global network for the permanent archiving and sharing of personally identifiable genetic, phenotypic, and clinical data generated for biomedical research projects or in the context of research-focused healthcare systems. Since its launch in 2008, the EGA has collected data from over 6,000 different research projects at nearly 1,000 institutions worldwide. The EGA makes this data discoverable, accessible, and reusable for researchers who need it, adhering to the findable, accessible, interoperable, and reusable (FAIR) principles and maximising the value of the collected data.

EGA is funded by ongoing [institutional support](#) from EMBL-EBI and the Centre for Genomic Regulation (CRG) and support through collaborative [projects and partnerships](#).



European Genome-phenome Archive (EGA) information

[DRIVER PROJECT WEBSITE](#)

| | |
|---------------|---|
| LOCATION | Europe |
| THEMATIC AREA | Rare Disease Cancer Complex Traits Basic Biology |
| CHAMPIONS | Thomas Keane Jordi Rambla Aina Jené |

Acknowledgements: Federated EGA



Acknowledgements: EGA



Roderic Guigó



Thomas Keane



Arcadi Navarro



Helen Parkinson



Jordi Rambia



Ana Alonso



Silvia Bahena



Aldar Cabelles



Àngel Carreño



Marcos Casado



Amy Curwin



Teresa D'Altri



Abeer Fadda



Raül García



Teresa García



Sara Gregorio



Csaba Halmagyi



Aina Jené



Bela Juhasz



Oriol Lopez-Doriga



Mireia Marin



Óscar Martínez



Andrea Mero



Gemma Milla



Akiris Moctezuma



Liina Nagirnaja



Francesc de Puig



Santiago Rensonnet



Aravind Sankar



Andres Silva



Coline Thomas



Sabela de la Torre



Claudia Vasallo

Core organizations:



Additional sources:



And critical support from the following sources:

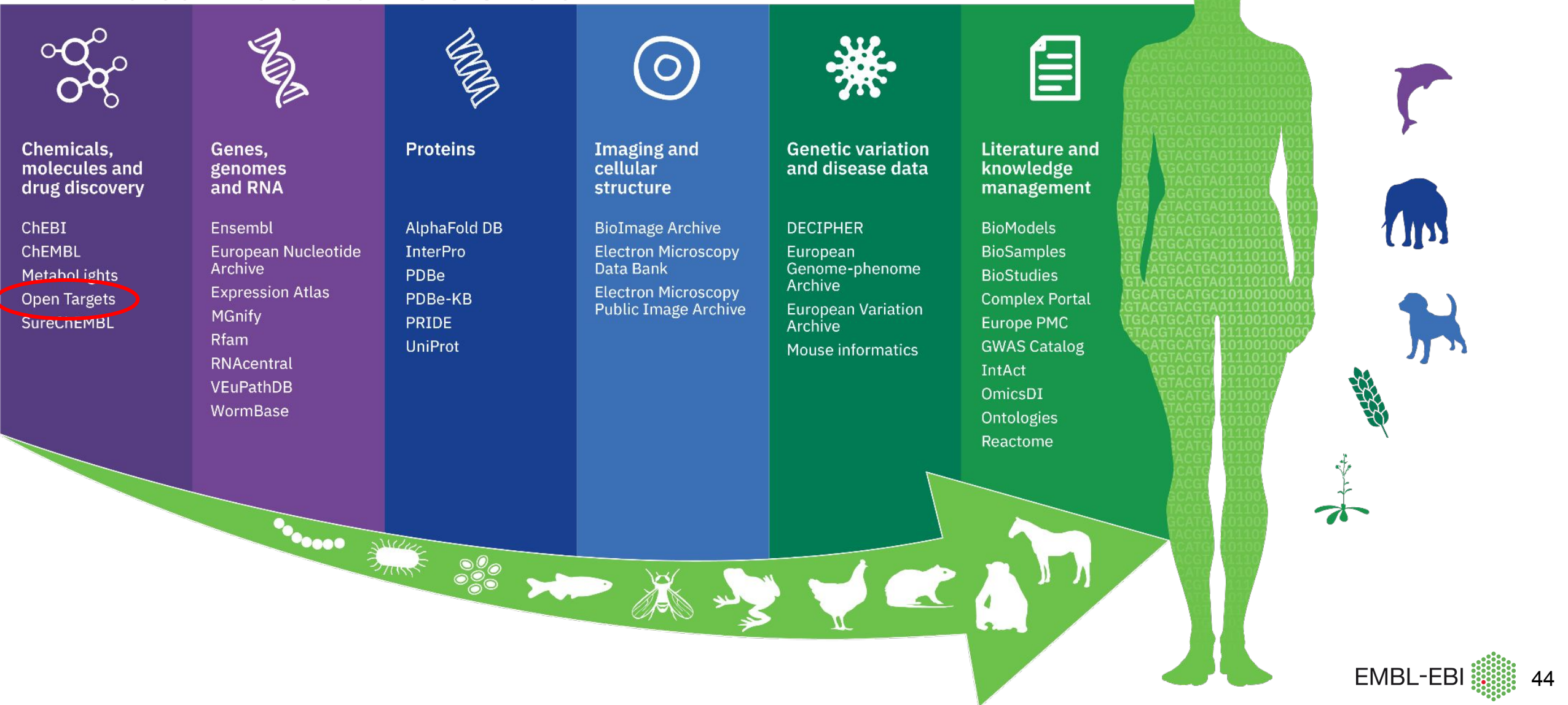


LinkedIn: <https://www.linkedin.com/company/european-genome-phenome-archive-ega/>

Web: <https://ega-archive.org/>

Tw: @EGAarchive

Data resources at EMBL-EBI



What is Open Targets?



Developing safe and effective drugs is difficult and expensive

We are dedicated to changing this with innovative experimental and informatics approaches

Open Targets is an innovative, large-scale, multi-year, public-private partnership that uses human genetics and genomics data for systematic drug target identification and prioritisation.

Visit the [Open Targets Platform](#) which integrates public domain data to enable target identification and prioritisation, or [Open Targets Genetics](#) which identifies targets based on GWAS and functional genomics. We complement data integration with large scale systematic experimental approaches to support target identification, prioritisation and validation. Check out [our latest papers](#) describing our experimental target identification approaches in oncology, neurodegeneration, and immunity and inflammation.

Find targets for a given disease

Visit the Open Targets Platform

Discover genetic evidence for targets

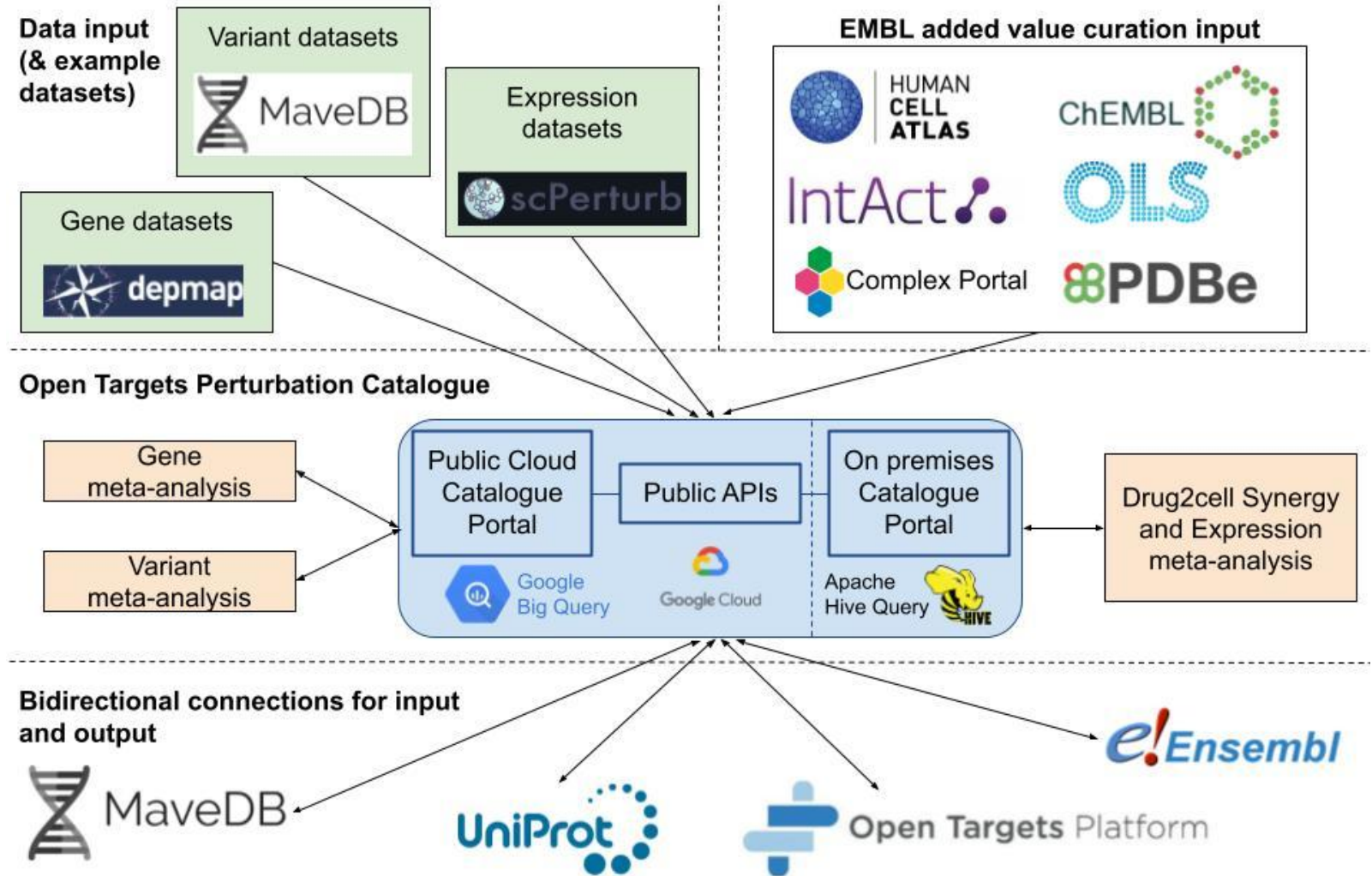
Visit Open Targets Genetics

Explore our research projects

Read our publications

Open Targets Perturbation Catalogue

“Integrate data from functional genetics screens and post-perturbational experiments, including data from CRISPR, MAVE, and Perturb-seq, to deliver gold standard machine learning-ready datasets, and a highly customisable public cloud platform with distributed data warehousing technology for cross metadata queries”



Acknowledgements: Open Targets Perturbation Catalogue

EMBL-EBI

Alexey Sokolov
James Stephenson
Prabhat Tootoo
Maria-Jesus Martin
Mallory Freeberg

Genentech

David Richmond
Hector Corrada Bravo
Tim Sterne-Weiler

GSK

Aleix Lafita
Daniel Seaton
Philippe Sanseau

Human Technopole

Francesco Iorio

Open Targets

Daniel Suveges

Pfizer

Charul Jani
Isac Lee
Jason Arroyo
Paul Yenerall

Sanofi

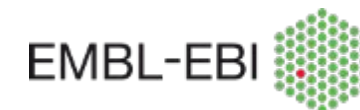
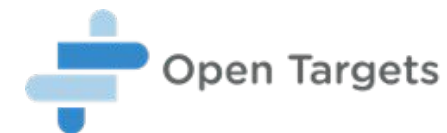
Charlie Fulco
TJ Iyyanki

MaveDB

Alan Rubin (WEHI Australia)
Doug Fowler (University of
Washington)

Wellcome Sanger Institute

Ally Dunham
Leopold Parts
Mo Lotfollahi



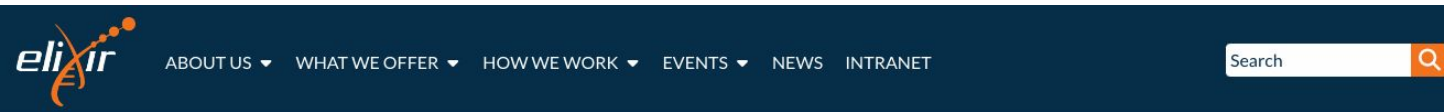
Website: Coming soon!

What makes these human
genomics and health data
resources **global**?

Active
participants
in **global**
initiatives



ELIXIR infrastructure for life sciences



ELIXIR is a European life sciences infrastructure, bringing together scientists from 21 countries and over 250 research institutes.

We enable researchers to access and analyse life science data, to improve the value and impact of life science research on public health, the environment and the economy.



What we offer



Guidelines

Guidelines and best practices to help you manage life science data, run training courses, develop software and more.



Web portals

Find the right software, training courses, standards and more in our interlinked portals to life science resources.



All resources

Find compute services, databases, and the full list of resources ELIXIR coordinates.



Partnerships with Industry and SMEs



Opportunities to work together



For ELIXIR members

If you work at an institute that is part of

ELIXIR Core Data Resources

Ensembl, EGA




International collaborations

- Australian BioCommons
- Global Organisation for Bioinformatics Learning, Education and Training (GOBLET)
- Research Data Alliance (RDA)
- GA4GH Strategic Partnership

<https://elixir-europe.org/>



Global Alliance for Genomics and Health


Global Alliance
 for Genomics & Health
 Collaborate. Innovate. Accelerate.

[About us](#) ▾ [Our community](#) ▾ [What we do](#) ▾ [Our products](#) ▾ [Get involved](#) ▾ [News and events](#) ▾

Unlocking the power of genomic data to benefit human health

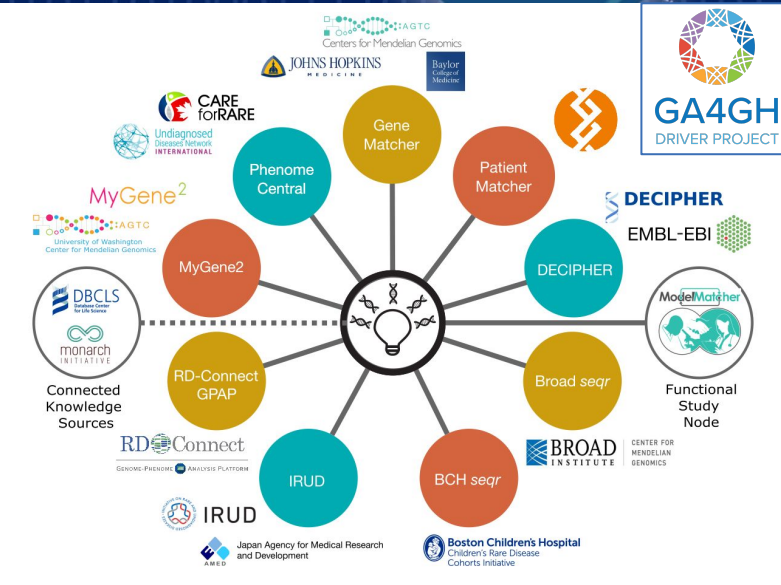


GA4GH builds the foundation for broad use of genomic data.

The not-for-profit Global Alliance for Genomics and Health (GA4GH) sets standards for genomic data use within a human rights framework.

[LEARN ABOUT GA4GH](#)

[READ OUR FRAMEWORK](#)



Driver Projects
EGA, Matchmaker Exchange (DECIPHER)

Communities of Interest
Rare Disease, Cancer, Clinical Genomics Laboratory

Strategic Partners
ELIXIR, RDA, NIH, Health Data Research UK



- ★ Patients and families for permission to include their data in DECIPHER
- ★ Members of the global DECIPHER community
- ★ Resources which DECIPHER uses

DECIPHER Leadership & Core Team



Recent alumni: Daniel Perrett, Simon Brent, Ben Hutton

Website: <https://www.deciphergenomics.org>

X/Twitter: @deciphergenomic

Email: contact@deciphergenomics.org



Funding



National Human Genome Research Institute (NHGRI)

National Institute of Allergy and Infectious Diseases (NIAID)



Funded by the European Union



EMBL-EBI

Alexey Sokolov
James Stephenson
Prabhat Tooto
Maria-Jesus Martin
Mallory Freeberg

Genentech

David Richmond
Hector Corrada Bravo
Tim Sterne-Weiler

GSK

Alex Lafita
Daniel Seaton
Philippe Sanseau

Human Technopole

Francesco Iorio

Open Targets

Daniel Suveges

Pfizer

Charul Jani
Isac Lee
Jason Arroyo
Paul Yenerall

Sanofi

Charlie Fulco
TJ Iyyanki

MaveDB

Alan Rubin (WEHI Australia)
Doug Fowler (University of Washington)

Wellcome Sanger Institute

Ally Dunham
Leopold Parts
Mo Lotfollahi



55

Website: Coming soon!

ADDITIONAL SOURCES:



And critical support from the following sources:



Acknowledgements: Federated EGA



Acknowledgements: EGA



LinkedIn: <https://www.linkedin.com/company/european-genome-phenome-archive-ega/>

Web: <https://ega-archive.org/>

Tw: @EGAarchive



52



Australian BioCommons, especially
Bernie Pope ~ Melissa Burke ~ Christina Hall ~ Jess Holliday



**Happy to take
Questions**

