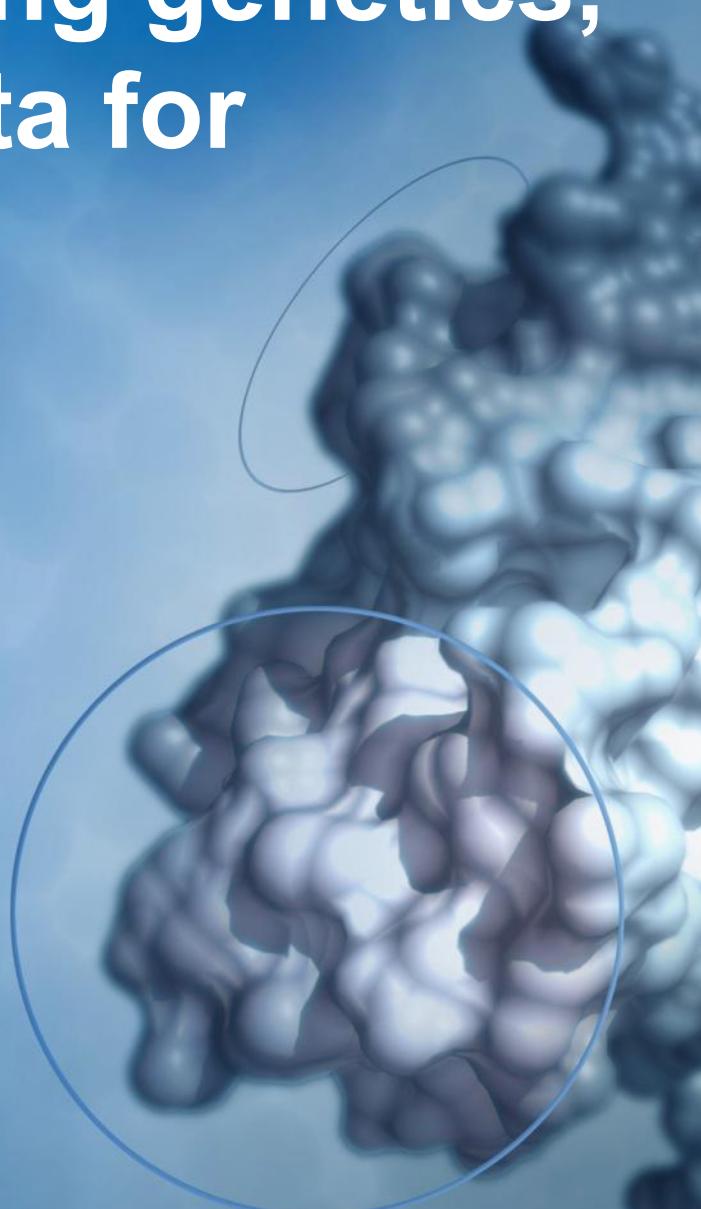


Open Targets: integrating genetics, omics and chemical data for translational research

Kymab
Cambridge
Sep 26th 2019

Denise Carvalho-Silva, PhD
EMBL-EBI | Open Targets
Wellcome Genome Campus
United Kingdom



This session 14:00-17:00

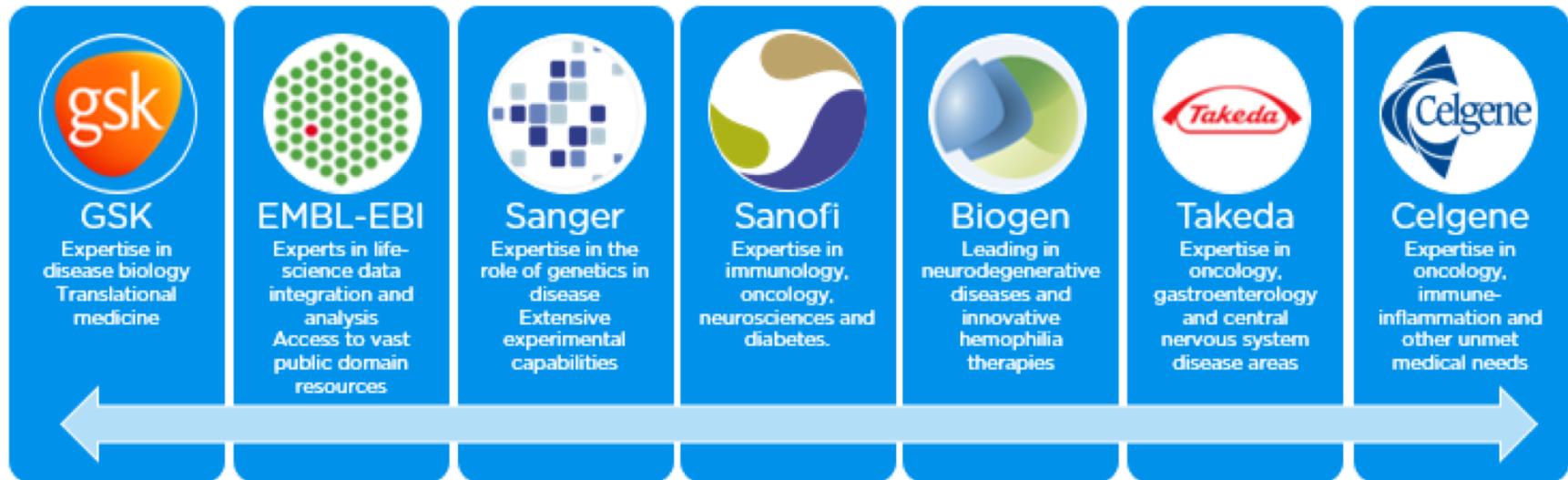
- Introduction to Open Targets
- Open Targets Platform: talk + live demo

Break → 15:15-15:30

- Hands-on demos and exercises
- Wrap up and feedback survey

Open Targets

A partnership to transform drug discovery



Founding partners

2018

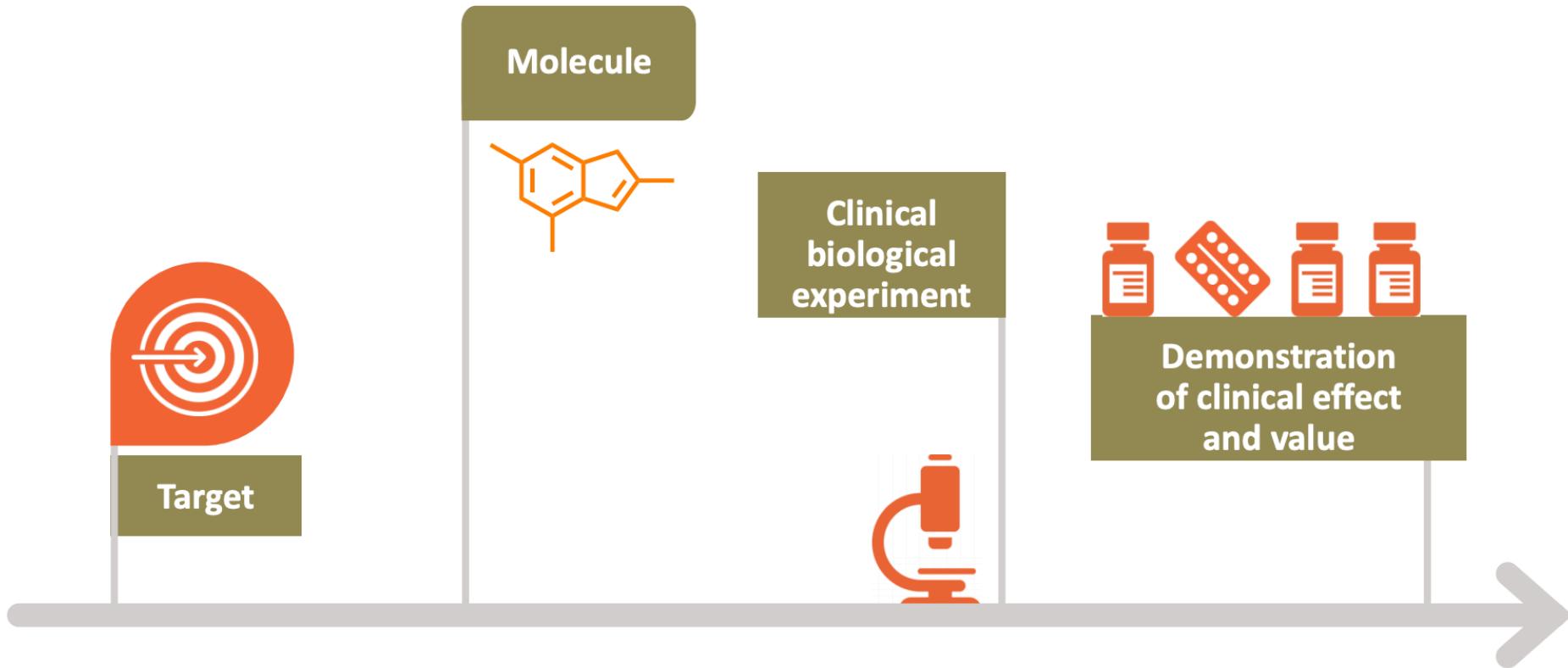
2016

2017

2018

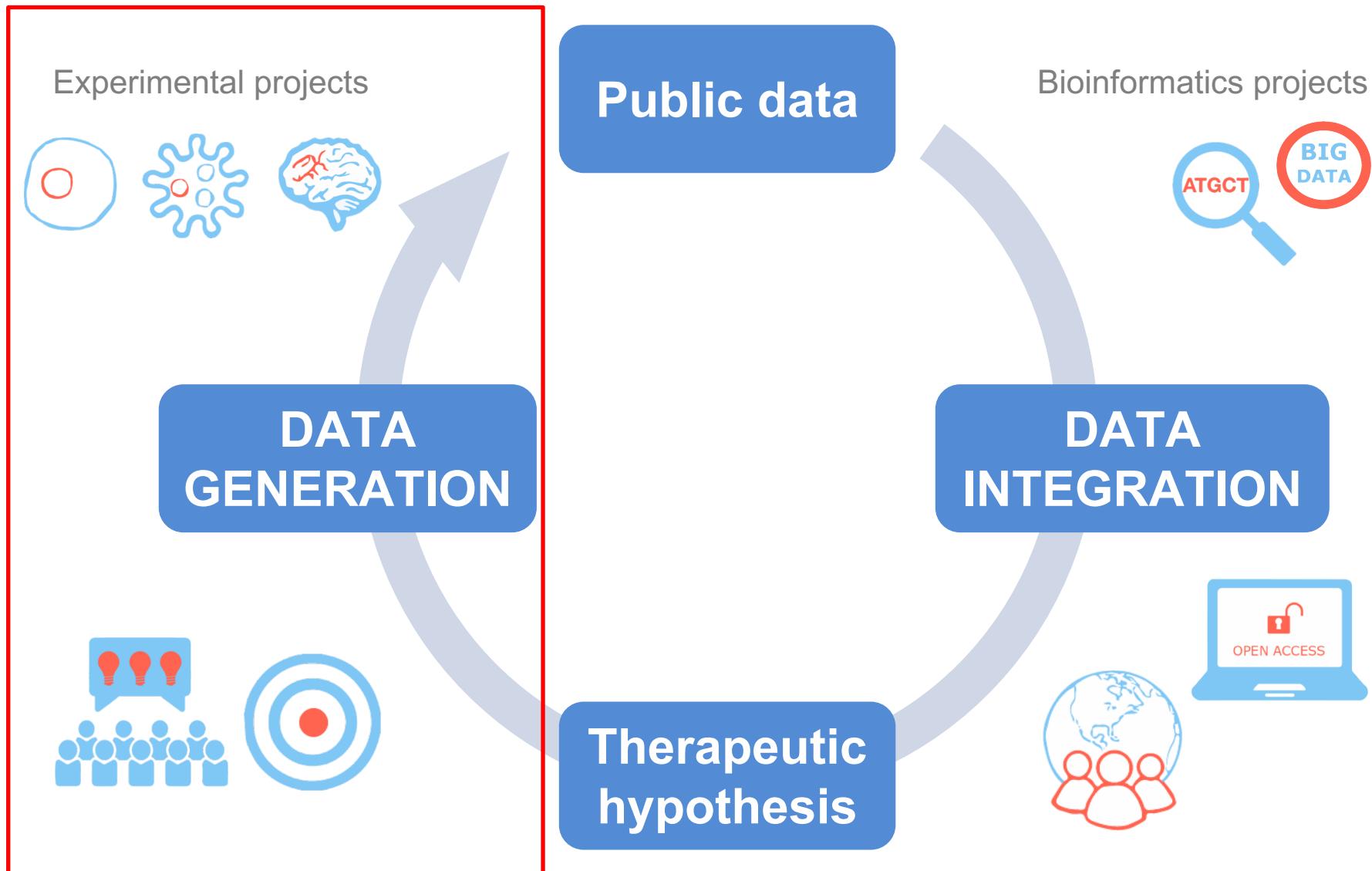
Systematic identification and prioritisation of targets

Selecting the right target

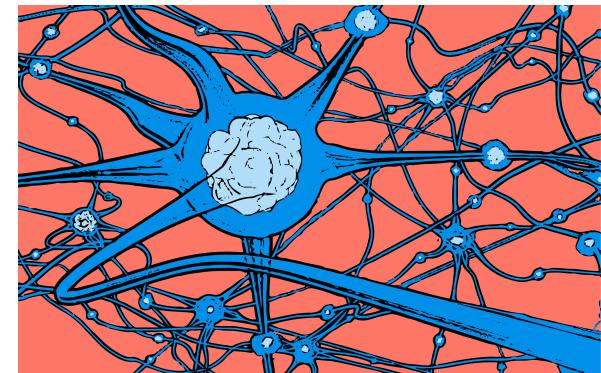
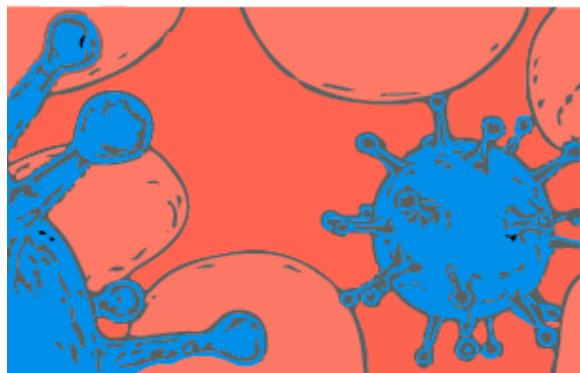
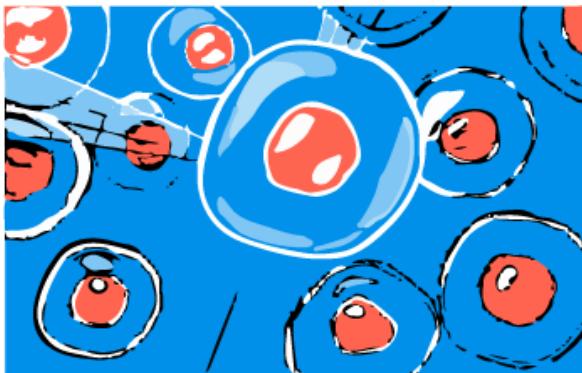


Drugs that fail are neither safe nor efficacious

Knowledge cycle



Data generation: therapeutic areas



5 ongoing projects

3 complete projects

1 publication (2019)

7 ongoing projects

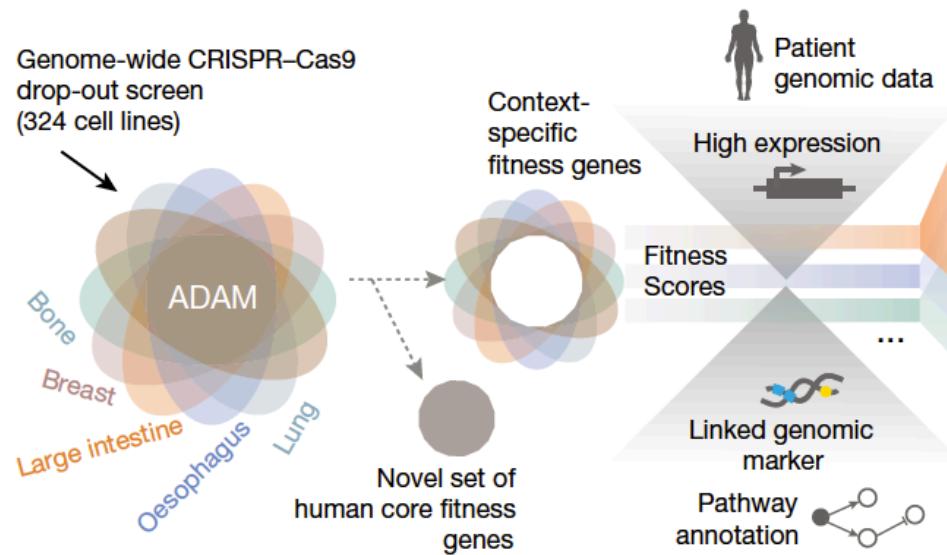
2 complete projects

3 manuscripts on bioRxiv

9 ongoing projects

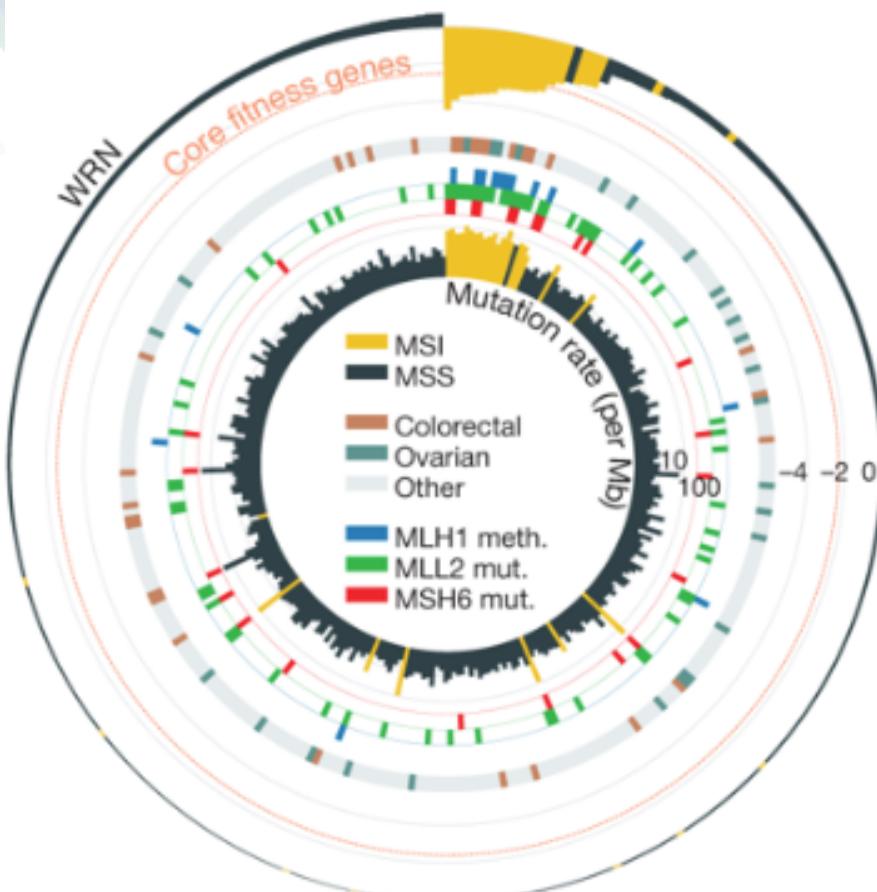
1 publication (2019)

Behan et al (2019)

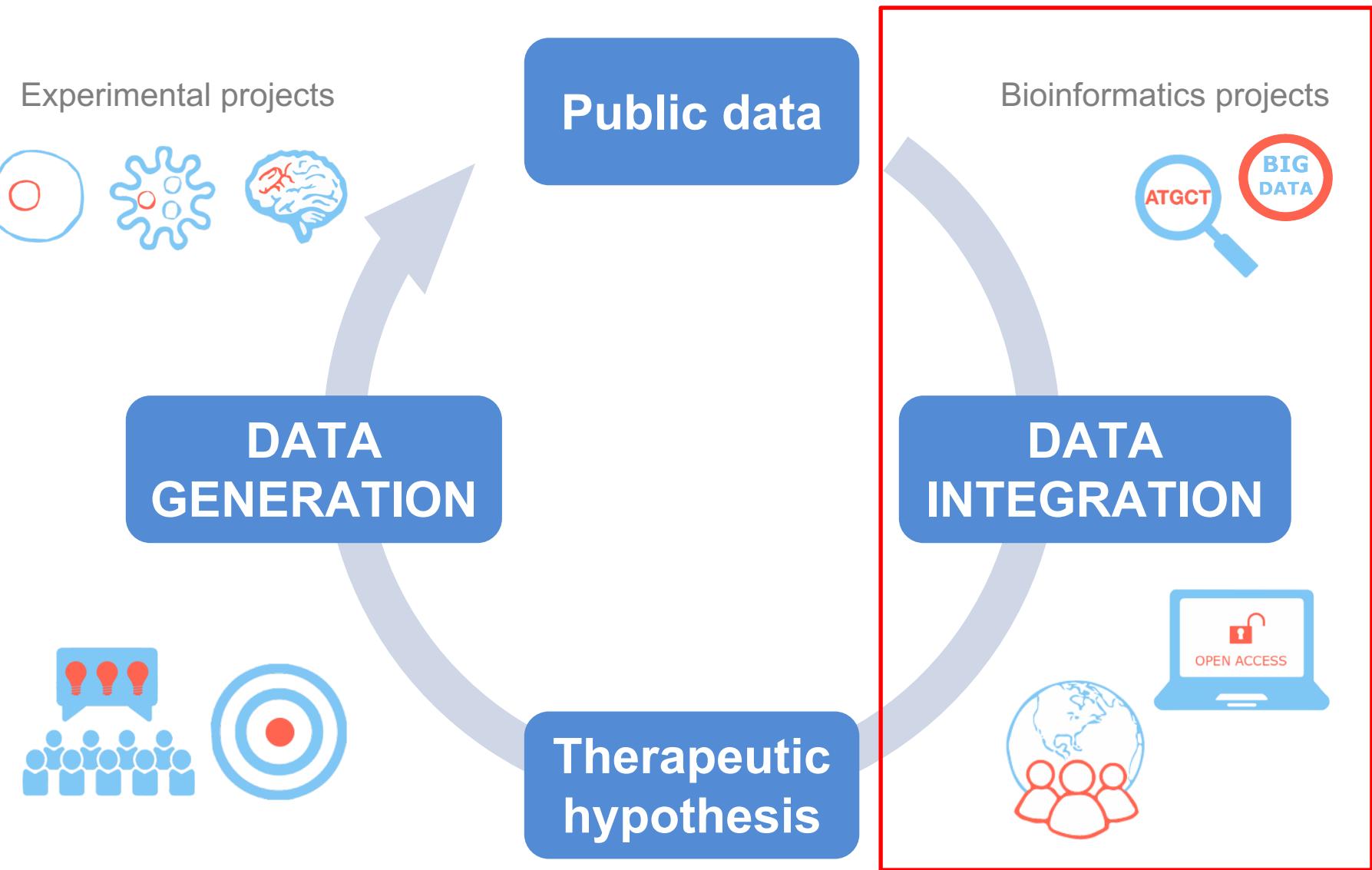


WRN sustains *in vivo* growth in:

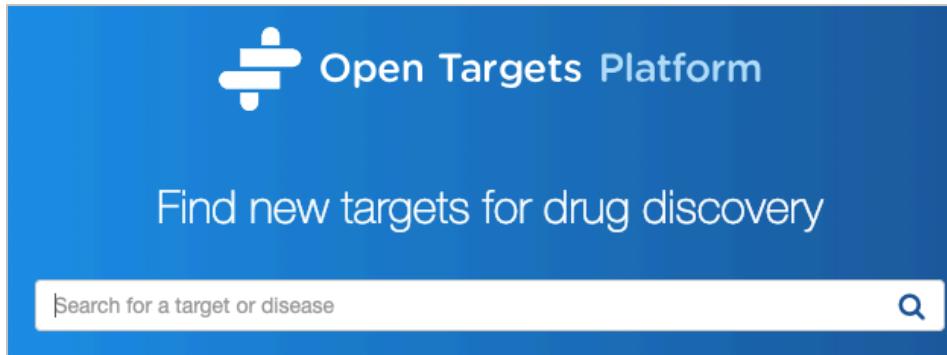
- colorectal
 - ovarian
 - endometrial
 - gastric
-
- New candidate target for tumours with MSI (WRN antagonists)



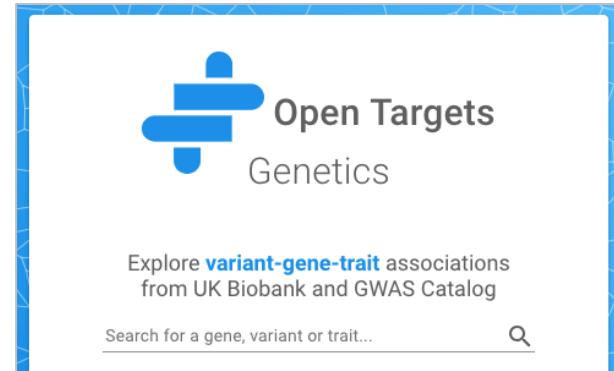
Knowledge cycle



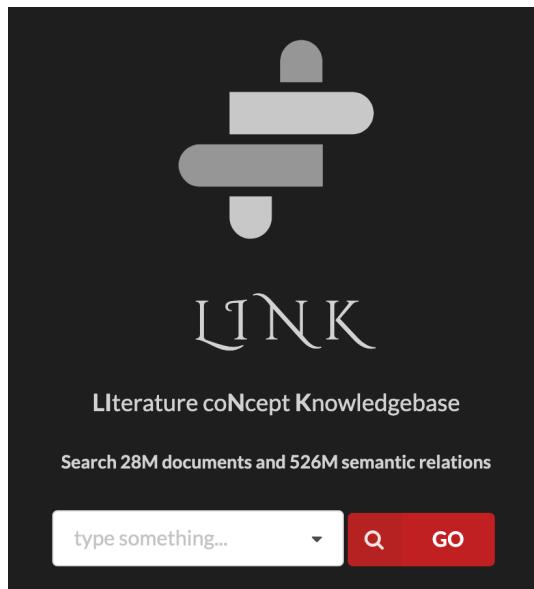
Data integration: web resources



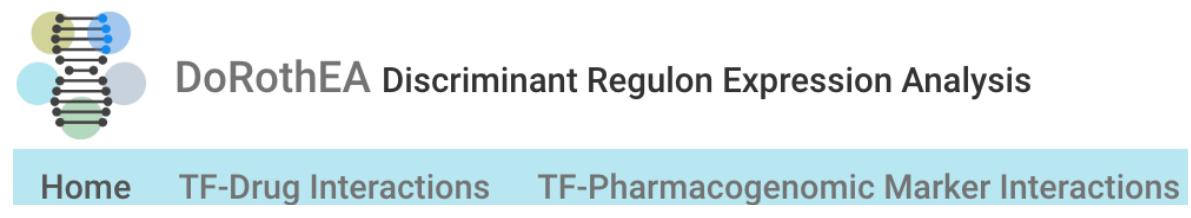
The Open Targets Platform homepage features a blue header with the logo and text "Open Targets Platform". Below the header is the tagline "Find new targets for drug discovery". A search bar at the bottom contains the placeholder "Search for a target or disease" and a magnifying glass icon.



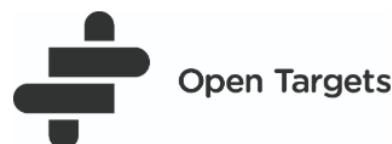
The Open Targets Genetics homepage features a white header with the logo and text "Open Targets Genetics". Below the header is the tagline "Explore variant-gene-trait associations from UK Biobank and GWAS Catalog". A search bar at the bottom contains the placeholder "Search for a gene, variant or trait..." and a magnifying glass icon.



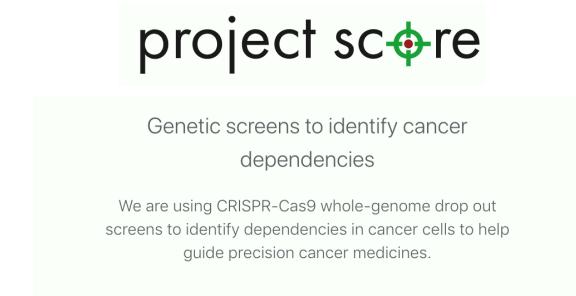
The LINK homepage features a dark background with the logo and text "LINK". Below the logo is the tagline "Literature coNcept Knowledgebase". A search bar at the bottom contains the placeholder "type something..." and a magnifying glass icon.



The DoRothEA Discriminant Regulon Expression Analysis homepage features a logo of a DNA helix and the text "DoRothEA Discriminant Regulon Expression Analysis". Below the logo is a navigation bar with links "Home", "TF-Drug Interactions", and "TF-Pharmacogenomic Marker Interactions".



The Open Targets logo, featuring a stylized "T" shape composed of three horizontal bars of decreasing length from left to right, with the text "Open Targets" to its right.



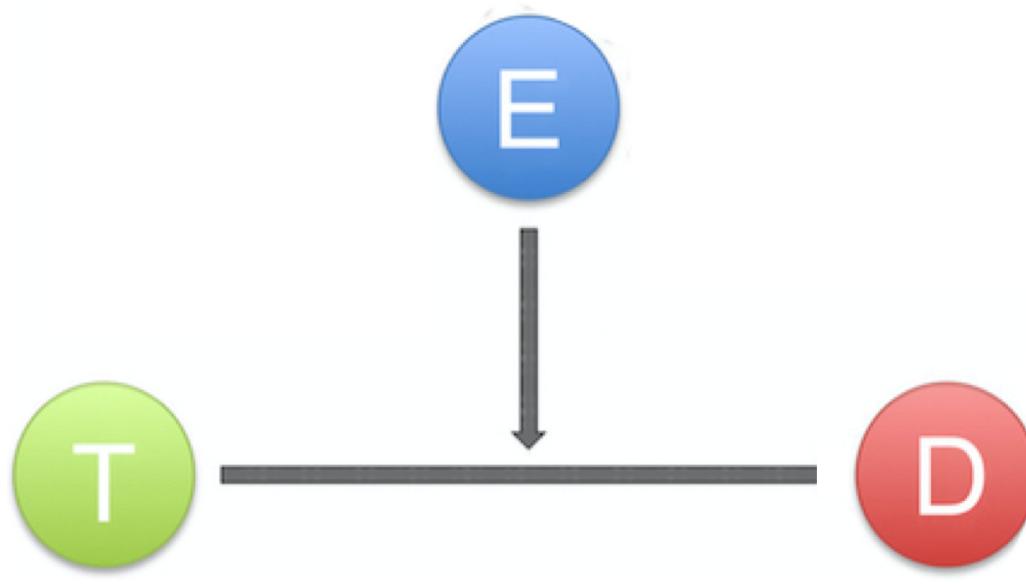
The project score homepage features the logo "project score". Below the logo is the tagline "Genetic screens to identify cancer dependencies". A paragraph at the bottom states: "We are using CRISPR-Cas9 whole-genome drop out screens to identify dependencies in cancer cells to help guide precision cancer medicines."

<https://www.opentargets.org/resources/#open-targets-platform>

Open Targets Platform: data model

Evidence

(variants, drugs, literature, etc)



Target

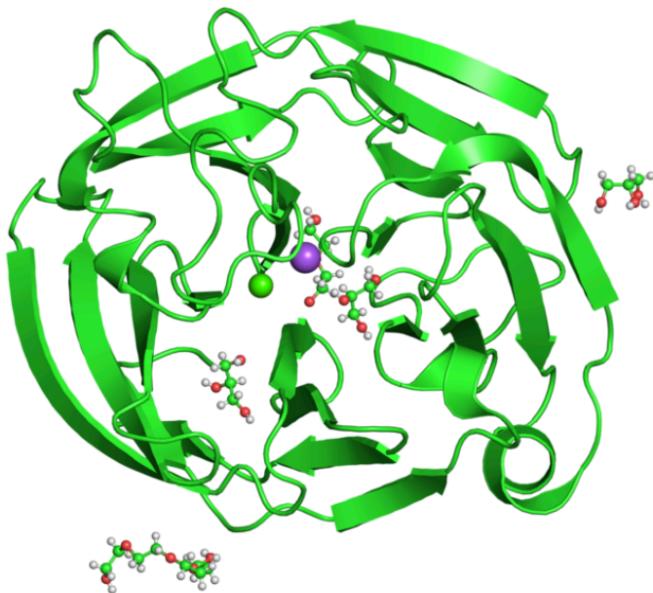
(protein and non-protein coding genes)

Disease

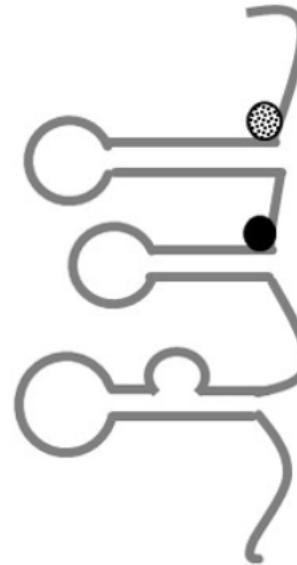
(rare and common)

<https://www.targetvalidation.org>

What is a target?



MYOC - glaucoma
(PDB 4wx5)

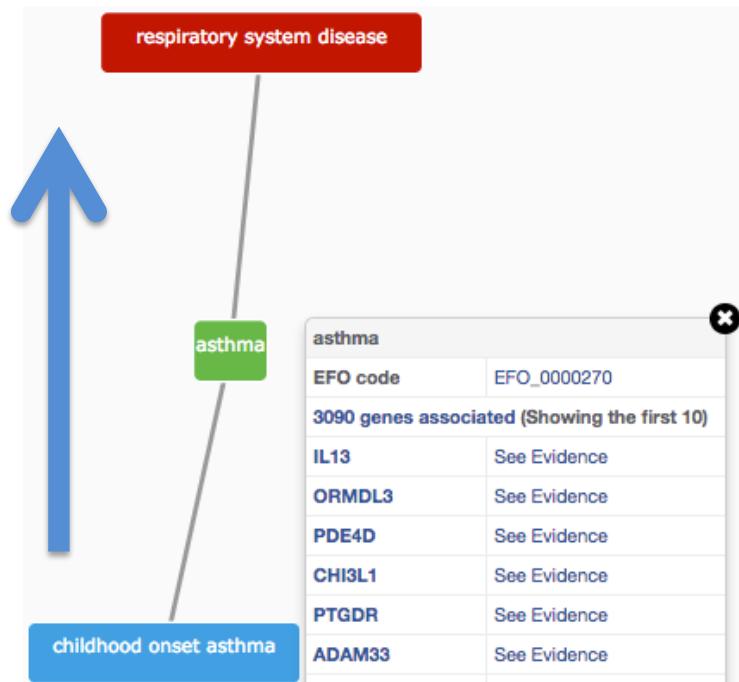


Xist – Turner syndrome
(lncRNA)

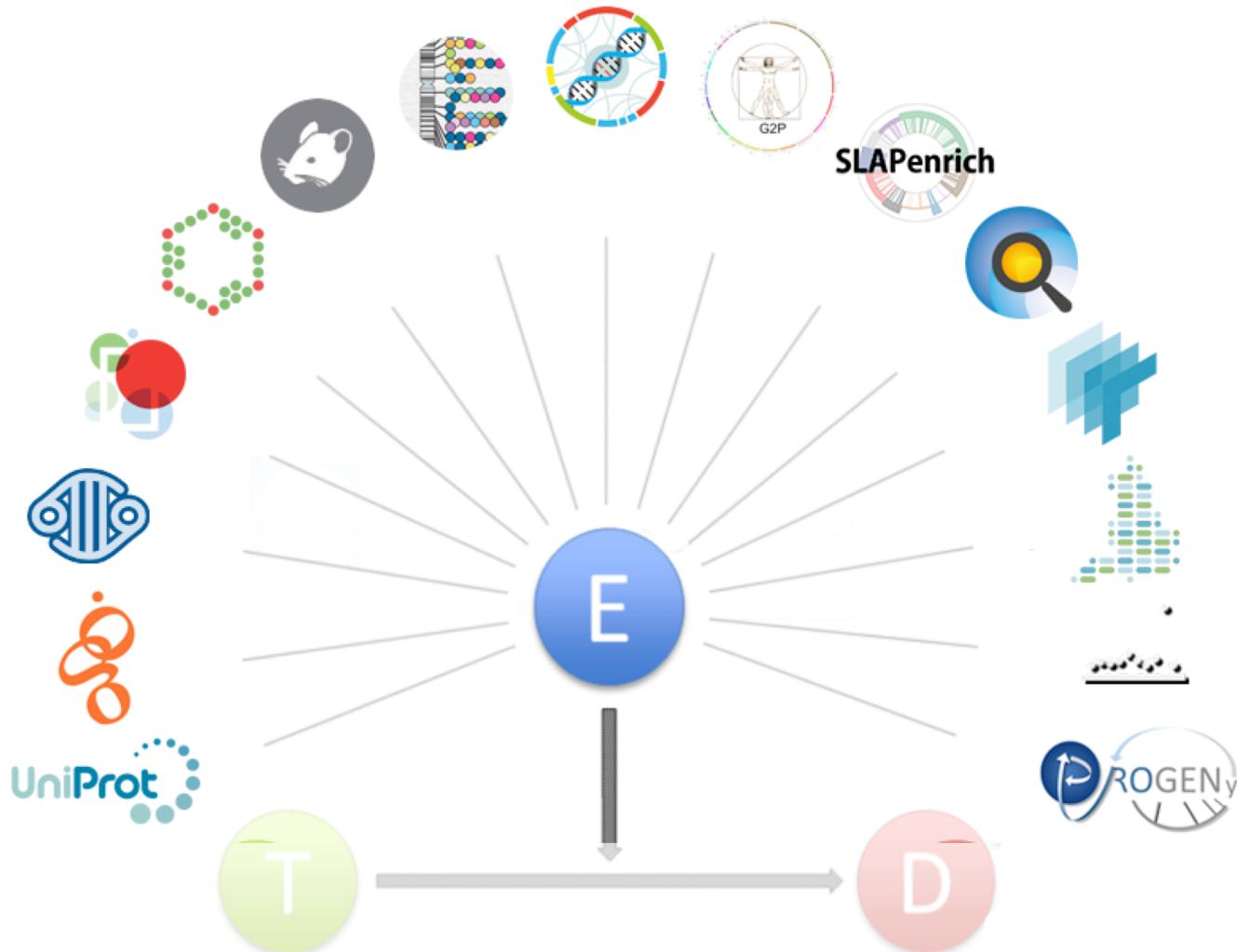
T

How do we describe our diseases?

- Modified version of Experimental Factor Ontology (EFO)
- Controlled vocabulary (Coeliac versus Celiac)
- Hierarchy (relationships)
 - Promotes consistency
 - Increases the richness of annotation
 - Allow for easier and automatic integration

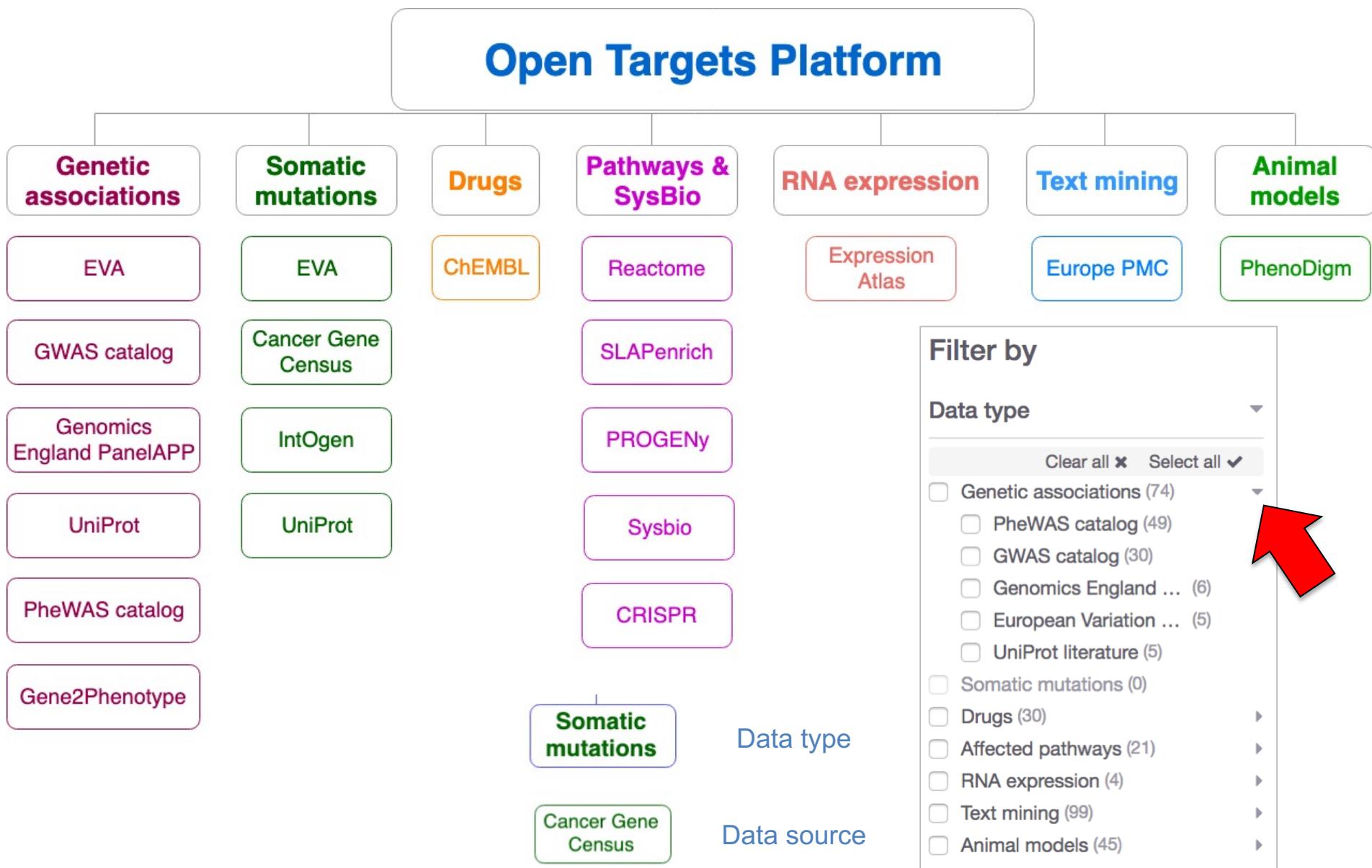


Evidence for target-disease associations



<https://docs.targetvalidation.org/data-sources/data-sources>

Data sources grouped into data types



What can you do with the Open Targets Platform?



- Target-disease associations (+ evidence + score)

https://www.targetvalidation.org/evidence/ENSG00000141510/EFO_0000228

- Target annotations

<http://www.targetvalidation.org/target/ENSG00000141510>

- Disease annotations

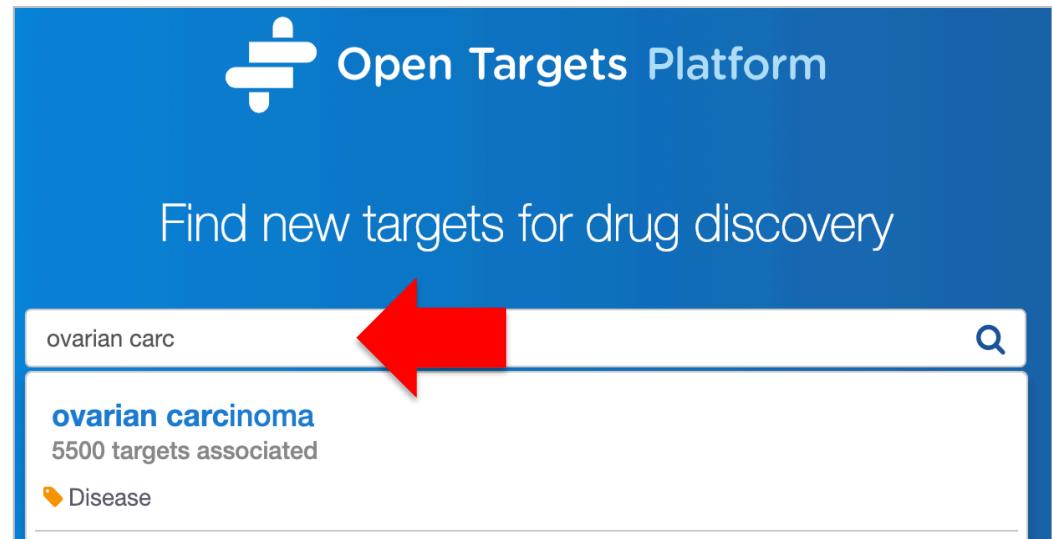
http://www.targetvalidation.org/diseaset/EFO_0000228

Demo 1

Which targets are associated with my disease?

What is the evidence for the association?

Is there any data to help to prioritise the list of associated targets?



Coursebook: pages 7-16

What can you do with the Open Targets Platform?



- Target-disease associations (+ evidence + score)
https://www.targetvalidation.org/evidence/ENSG00000141510/EFO_0000228
- Target annotations
<https://www.targetvalidation.org/target/ENSG00000141510>
- Disease annotations
https://www.targetvalidation.org/diseaset/EFO_0000228

Target annotations → target profile

⚠ Target safety

Drugs

Target tractability

Protein information

Pathways

Similar targets (based on diseases in common)

Variants, isoforms and genomic context

Protein interactions

RNA and protein baseline expression

Mouse phenotypes

Protein structure

Gene ontology

Gene tree

Bibliography

Cancer hallmarks

Cancer biomarkers

Target profile page



A Target Enabling Package (TEP) [?](#) has been developed for this target. Get more details on [CDK12 TEP](#)



TEPs are provided by the [Structural Genomics Consortium](#)



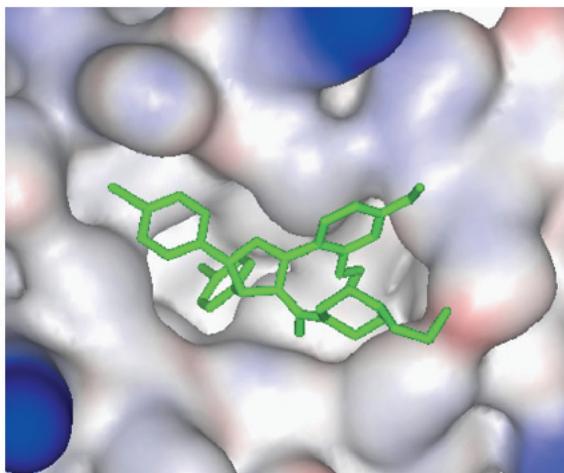
Potential chemical probes can be explored with [Probe Miner](#).

[Project Score ADORA2A profile page](#)

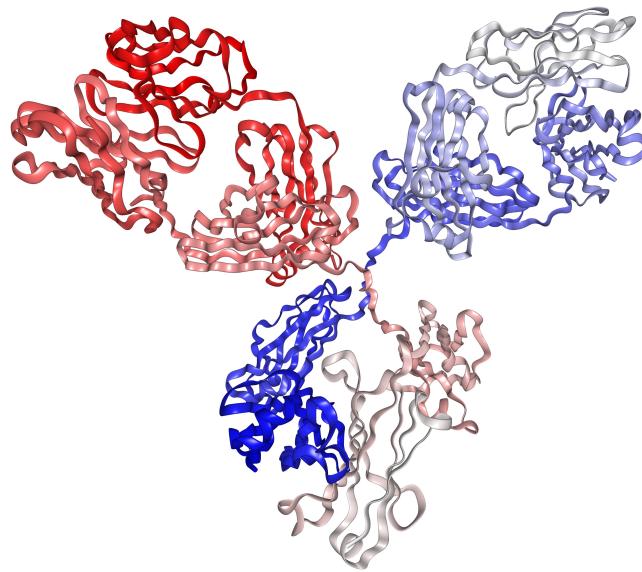
Is my target expressed in the relevant tissue?



Is my target tractable?



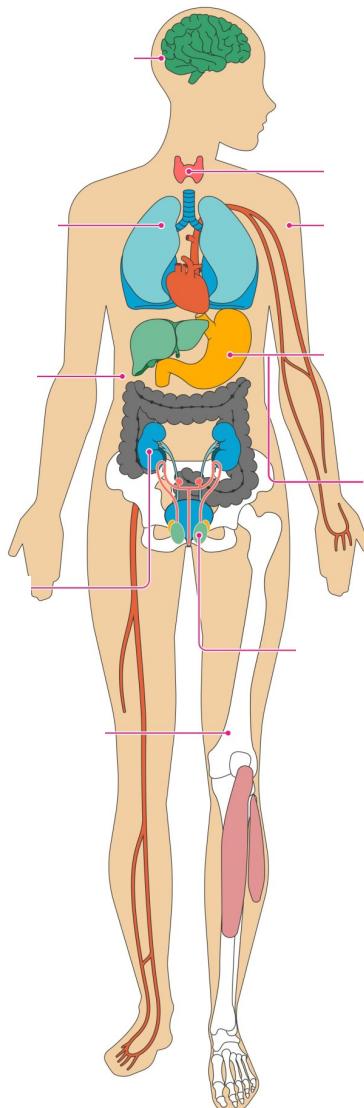
DOI:10.1517/17460441.3.4.391



Antibody

Small molecule			Antibody						
Clinical precedence			Discovery precedence		Predicted tractable - high confidence				
Phase 4	Phase 2 or 3	Phase 0 or 1	PDB targets with ligands	Active compounds in ChEMBL	Phase 4	Phase 2 or 3	Phase 0 or 1		
Predicted tractable					Predicted tractable - medium to low confidence				
DrugEBIliy score > 0.7	DrugEBIliy score 0 to 0.7				UniProt location - low or unknown confidence	UniProt predicted signal peptide or transmembrane region	GO cell component - medium confidence		

Is my target safe to be modulated?



⚠ Target safety New

The following effects have been observed and noted in one or more publications on drug target safety assessment.

Showing 1 to 2 of 2 entries

Search:

Main organs & systems affected **Agonism or activation effects** **Antagonism or inhibition effects** **Publications**

Main organs & systems affected	Agonism or activation effects	Antagonism or inhibition effects	Publications
• Endocrine	General <ul style="list-style-type: none">Increased prostate carcinomaOedemaAndrogenicity femalesIncreased muscle massIncreased hostilitySleep apnoeaLiver complications	General <ul style="list-style-type: none">Decreased spermatogenesisImpotenceGynecomastiaMastodyniaIncreased breast carcinoma	Bowes et al. (2012)

193 genes with target safety information

- Drowsiness
- Increased heart rate
- Hepatotoxicity

Disease
profile page

Disease annotations

Similar diseases (based on targets in common)



Phenotypes



Drugs



Bibliography



Classification

<https://docs.targetvalidation.org/getting-started/getting-started/disease-profile>

All drugs for a disease

Drugs

Drugs in clinical trials or approved for inflammatory bowel disease.

Source: ChEMBL

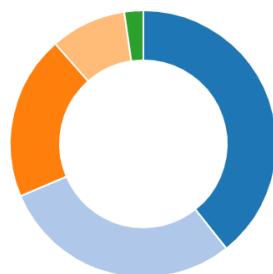
Table summary

108 unique drugs

6 associated disease

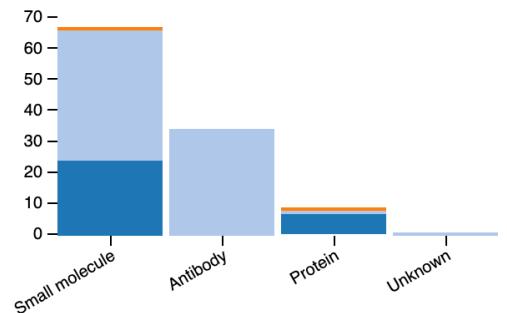
158 associated targets

Trials



- Phase IV (565)
- Phase III (421)
- Phase II (286)
- Phase I (132)
- Phase 0 (33)

Type vs Activity



- drug_positive_modulator
- drug_negative_modulator
- up_or_down

Showing 1 to 10 of 1,437 entries

Search:

[Download .csv](#)

[Download .tsv](#)

[Download .json](#)

Drug Information

Disease

Drug

Phase

Status

Type

Mechanism
of action

Activity

Gene-Drug Evidence

Target

Target class

Evidence
curated
from

Crohn's disease

PREDNISONE

Phase IV

N/A

Small molecule

Glucocorticoid
receptor agonist
 1 publication
EuropePMC

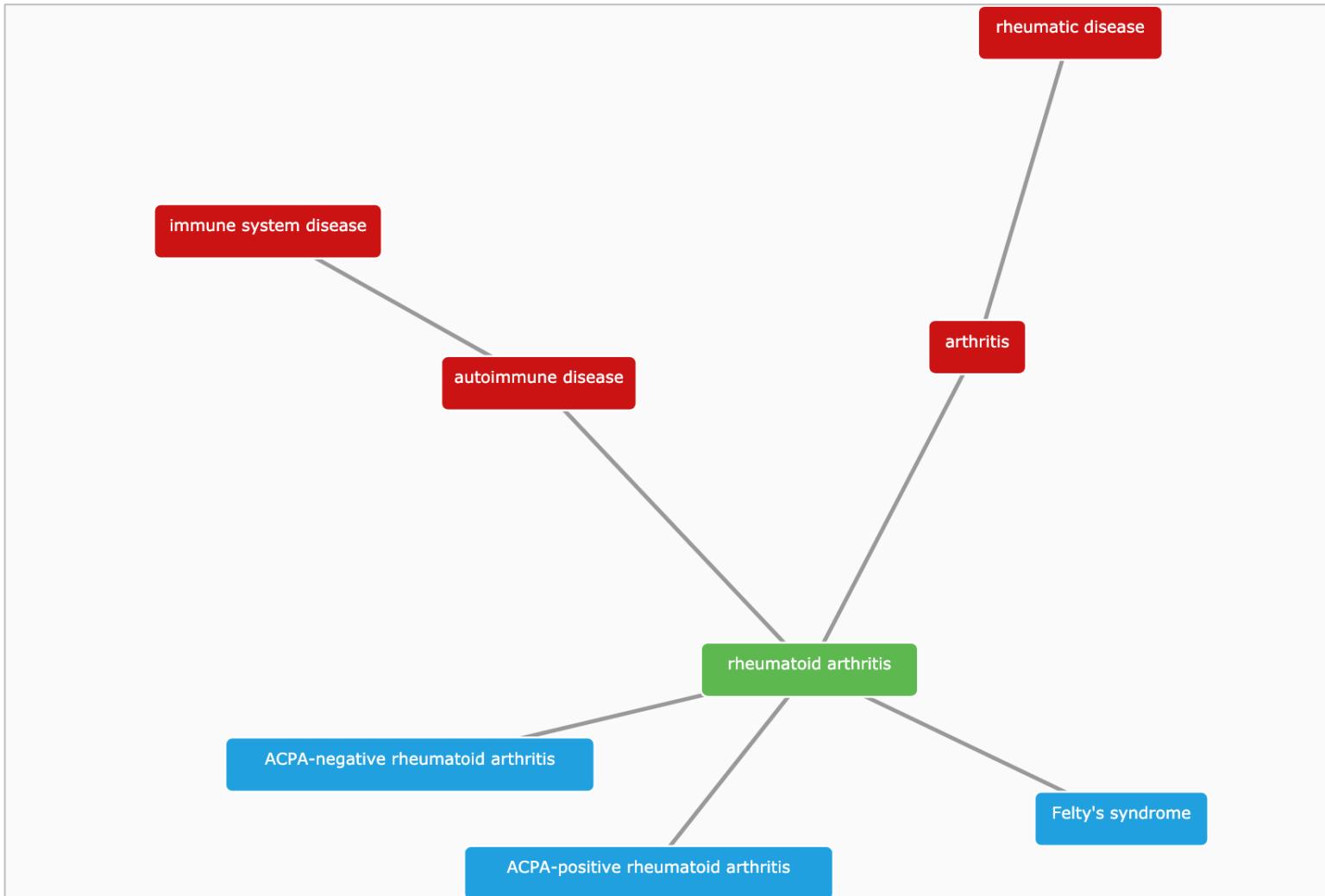
agonist

NR3C1

Nuclear hormone
receptor
subfamily 3
group C member
1

[DailyMed
Information](#)

(Disease) Classification



https://www.targetvalidation.org/disease/EFO_0000685

Hands-on time

Demos

Exercises

Target and disease annotations

Durvalumab

Searching for a target

ICOS

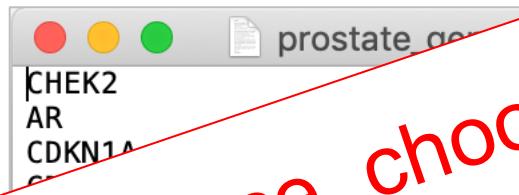
Pages 16-25

Pages 26-27

Have you got a list of targets?

	A	B	C
1	ADCY7		
2	ATG16L1		
3	CARD9		
4	CD6		
5	ERAP2		
6	FCGR2A		
7	FUT2		
8	ICAM1		
9	IFIH1		
10	IL10		
11	IL18B		
12			
13			
14	I		
15	IT		
16	ITG		
17	MST		
18	NOD2		
19	NXPE1		
20	PLCG2		
21	PTPN22		
22	SLAMF8		
23	SMAD3		
24	SP140		
25	TNFSF8		
26	TYK2		

Using the web interface, choose
• Batch search or
“Your target list” filtering



AD_gene_IDs
ENSG000000000815
ENSG00000000143801
ENSG00000000130203
ENSG0000000010704
ENSG00000000169252
ENSG00000000186318
ENSG0000000091513
ENSG00000000175899

Batch search



About ▾ Help ▾ API ▾ Downloads Blog



Search for a target or disease



1. Upload your target list:

List format is one target per line (supported identifier formats: HGNC, UniProt, ENSG IDs, synonyms). A maximum of 200 targets is currently allowed.

[Load sample list](#)

Load your list from a file:

No file chosen

Copy and paste your list:

Give your list a name

Your target list

Feedback

<https://www.targetvalidation.org/batch-search>

Your target list

https://www.targetvalidation.org/disease/EFO_0001075/associations

Open Targets Platform About Help API Downloads Blog Search

5532 targets associated with ovarian carcinoma

[View disease profile](#)

Filter by

- Data type
- Pathway types
- Target class
- RNA tissue specificity [?](#)
- Target reactivity
- Your target list

Associations view Prioritisation view

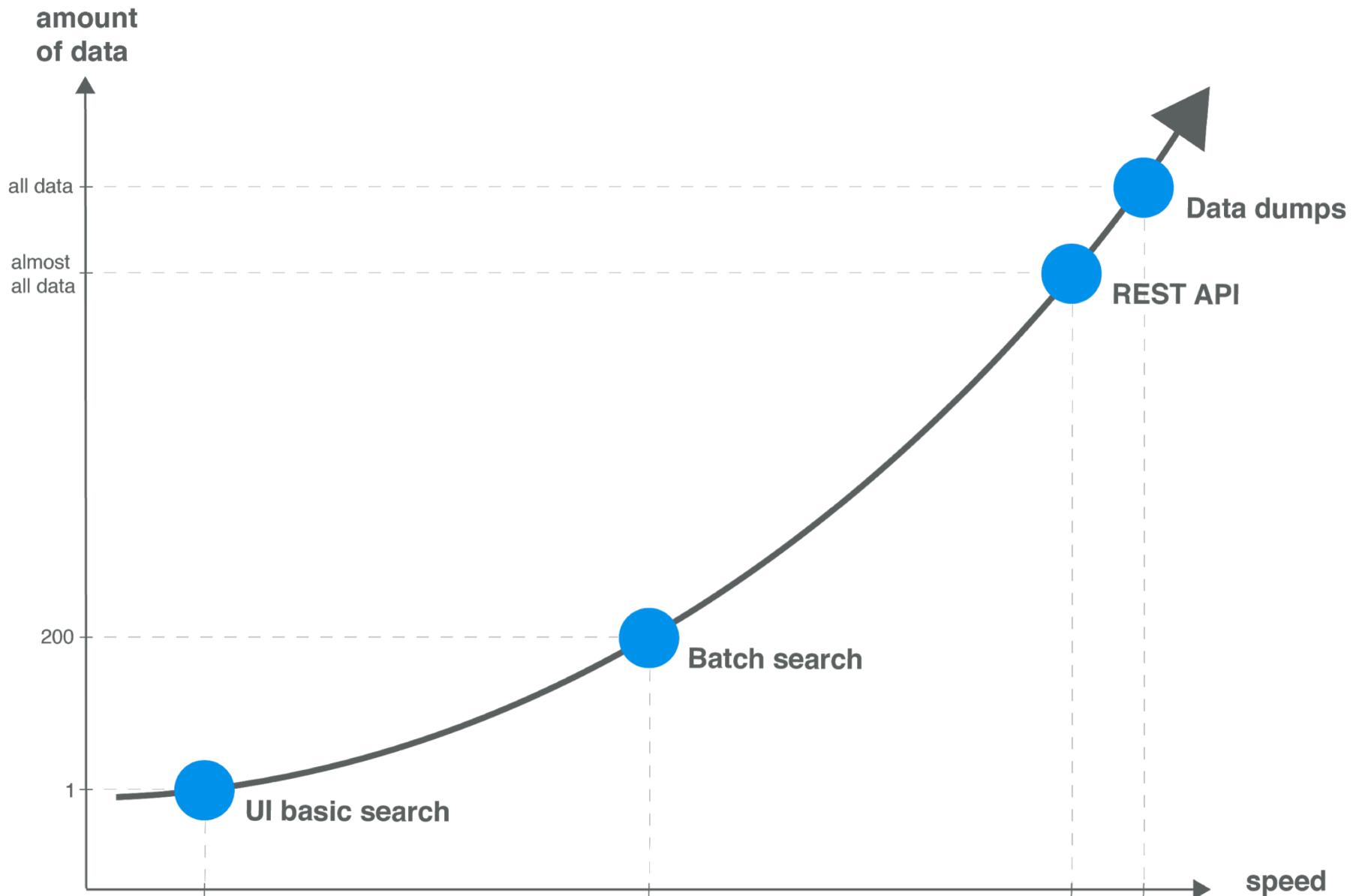
Showing 1 to 10 of 5,532 targets

Search:

Target symbol	Overall association score	Genetic associations	Somatic mutations	Drugs	Pathways & biology
BRCA1	High	Medium	Low	Low	Low
BRCA2	High	Medium	Low	Low	Low
RAD51C	Medium	Medium	Low	Low	Low
RAD51D	Medium	Medium	Low	Low	Low
FGFR2	Medium	Medium	Low	Low	Low

Choose File No file chosen

Modes of access → data volume



Open Targets Platform REST API*

Server <https://platform-api.opentargets.io/v3/platform/>

Endpoint public/association/filter

Parameters ?target=ENSG00000163914&size=10000&fields=target.id&fields=disease.id



Open Targets
Python client **

* <https://docs.targetvalidation.org/tutorials/rest-api>

** <https://docs.targetvalidation.org/tutorials/python-client>

Looking for the entire data?



Open Targets Platform

About ▾ Help ▾ API ▾ Downloads Blog

Data Download

All data from targetvalidation.org is available for download as compressed JSON files.

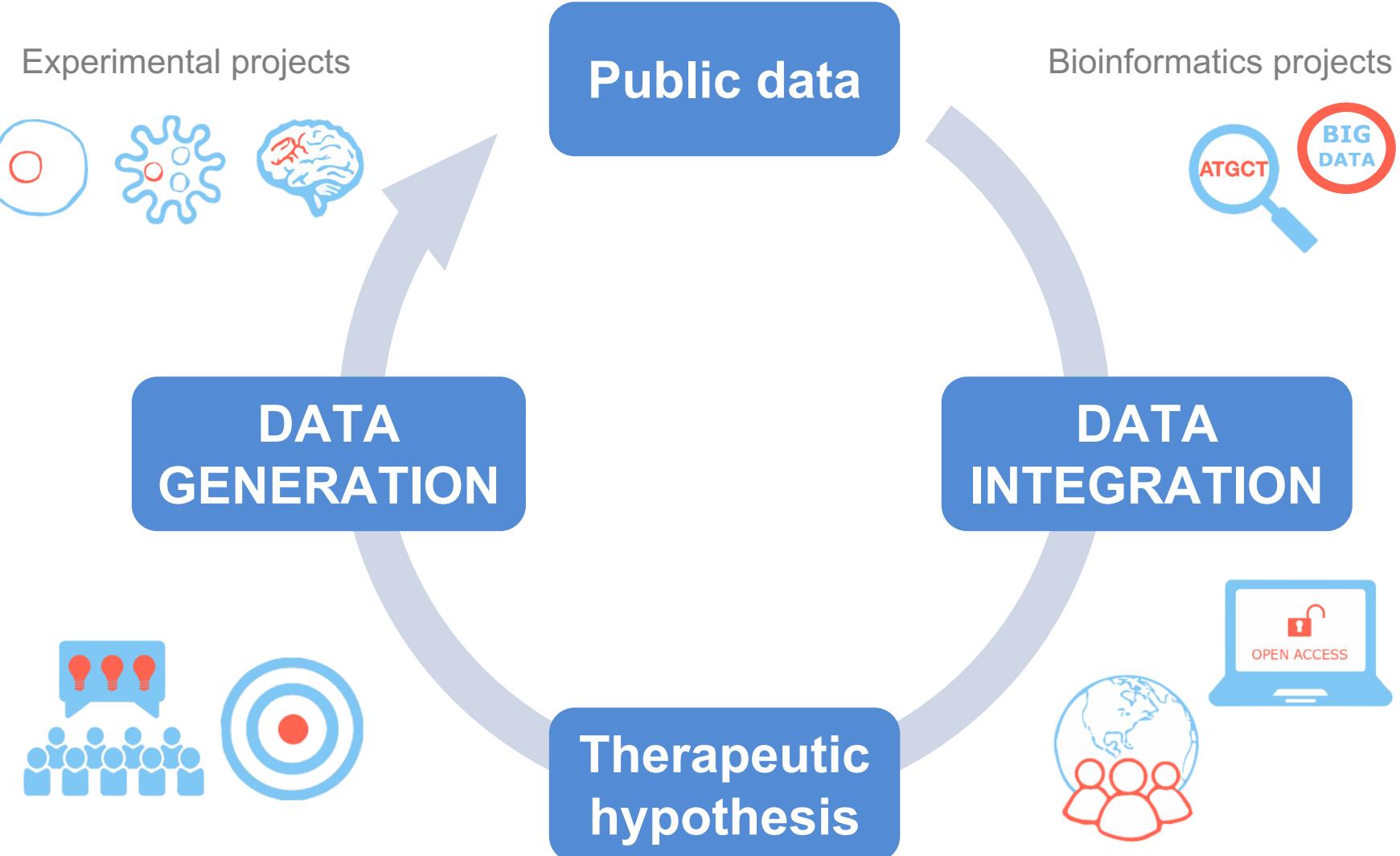
We provide downloads of all associations between target and disease calculated by the platform, as well as all the evidence used in calculating each associations. These are the same objects returned by the corresponding [/public/associations](#) and [/public/evidence](#) API methods. Head to the [API documentation](#) for further details.

NOTE: the files below are useful only if you want to analyze the data. They are not a database dump and cannot be easily used to replicate the platform locally/somewhere else

2019 Jun

- [Association objects](#) (2019-06, 270MB)
- [Evidence objects](#) (2019-06, 2.72Gb)

Wrapping up



Wrapping up

- Resource of integrated multiomics data
- Target identification and prioritisation
- Graphical web interface: easy to use
- REST-API for larger, more flexible queries

27K
targets

10K
diseases

3.3 M
associations

7.2 M
evidence

September 2019 release, 19.09

Want to learn more?

Nucleic Acids Research

<http://bit.ly/cite-us>

Issues Section browse ▾ Advance articles Submit ▾ Purchase About ▾



Nucleic Acids Research, 2018 **1**
doi: 10.1093/nar/gky1133

Open Targets Platform: new developments and updates two years on

Denise Carvalho-Silva^{1,2,*}, Andrea Pierleoni^{1,2}, Miguel Pignatelli^{1,2}, ChuangKee Ong^{1,2}, Luca Fumis^{1,2}, Nikiforos Karamanis^{1,2}, Miguel Carmona^{1,2}, Adam Faulconbridge^{1,2}, Andrew Hercules^{1,2}, Elaine McAuley^{1,2}, Alfredo Miranda^{1,2}, Gareth Peat^{1,2}, Michaela Spitzer^{1,2}, Jeffrey Barrett^{2,3}, David G. Hulcoop^{2,4}, Eliseo Papa^{2,5}, Gautier Koscielny^{2,4} and Ian Dunham^{1,2,*}

¹European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Hinxton, Cambridgeshire CB10 1SD, UK, ²Open Targets, Wellcome Genome Campus, Hinxton, Cambridgeshire CB10 1SD, UK, ³Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SA, UK, ⁴GSK, Medicines Research Center, Gunnels Wood Road, Stevenage, SG1 2NY, UK and ⁵Biogen, Cambridge, MA 02142, USA

Received September 14, 2018; Revised October 22, 2018; Editorial Decision October 23, 2018; Accepted October 26, 2018

Need help?



<https://docs.targetvalidation.org/>



support@targetvalidation.org



<https://tinyurl.com/opentargets-youtube>



[@targetvalidate](https://twitter.com/targetvalidate)



<http://tinyurl.com/opentargets-in>



blog.opentargets.org/

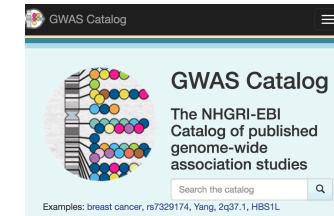
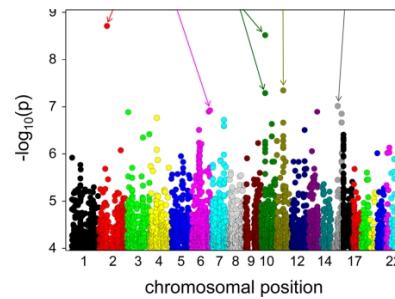
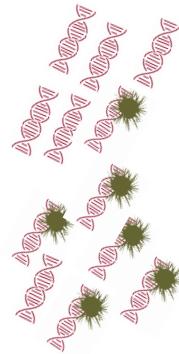
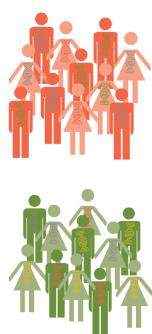
Extra extra extra

Details on data sources to associate
targets and diseases

Other Open Targets resources

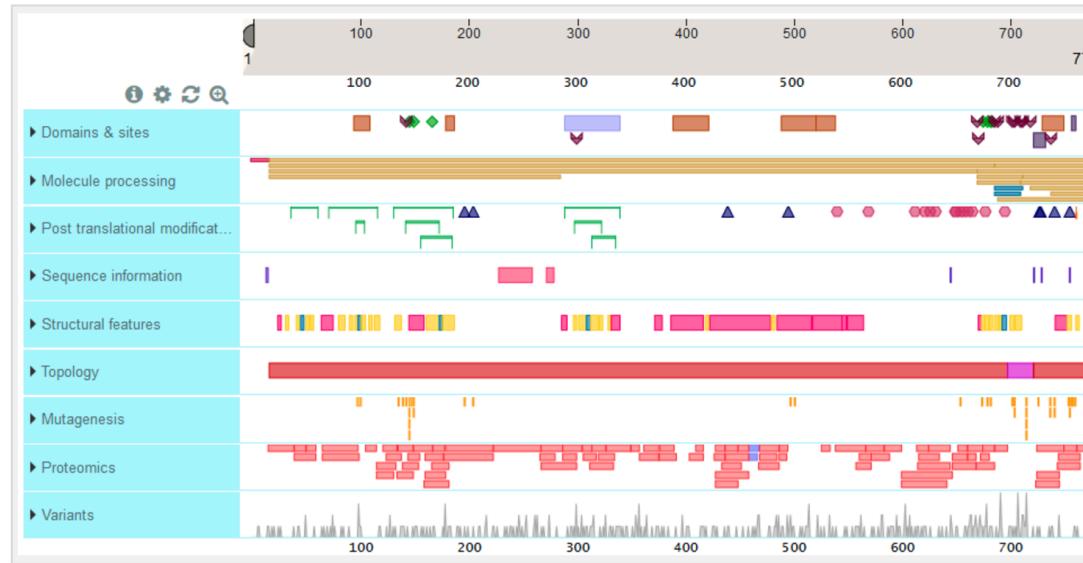
Data sources: GWAS catalog

- Genome Wide Association Studies
- Array-based chips → genotyping 100,000 SNPs genomewide



Data sources: UniProt

- Protein: sequence, annotation, function



- Manual curation of coding variants in patients



Data sources: Gene2Phenotype

The screenshot shows the homepage of the Gene2Phenotype database. At the top left is the "Gene2Phenotype" logo, which features a circular emblem with a Vitruvian Man inside and the acronym "G2P" at the bottom. To the right of the logo is the word "Downloads". Below these are search and search results sections. The search section includes a "Search panel" with dropdown menus set to "ALL" and "for:", a search input field containing placeholder text, and a blue "Search" button. Below the search bar is a descriptive text example.

Gene2Phenotype Downloads

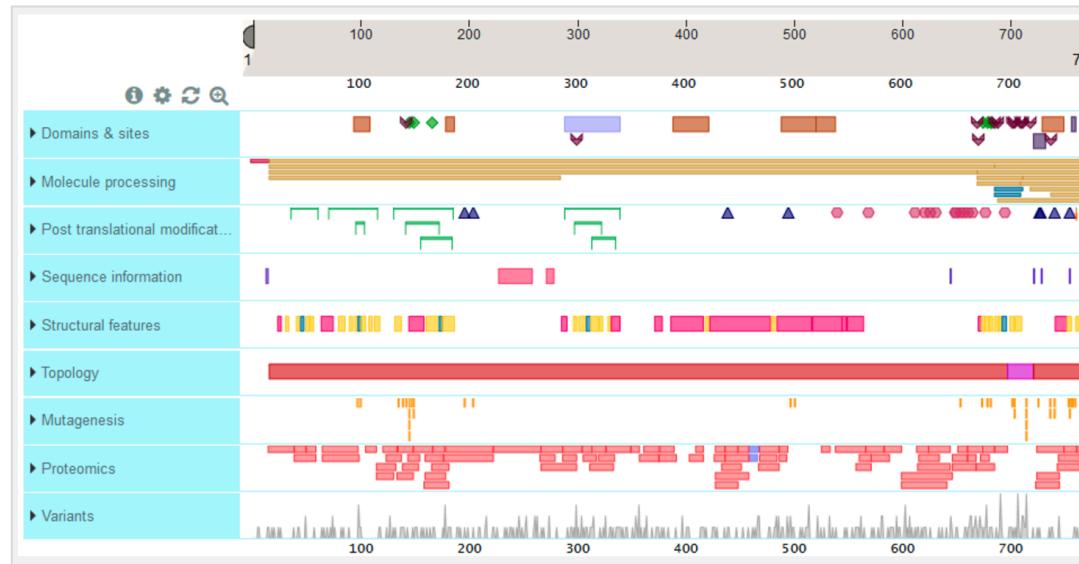
Search panel ALL for: Search

For example: **CRYBA1, ZEB2, TBX1, CHANARIN-DORFMAN SYNDROME or MITOCHONDRIAL COMPLEX III DEFICIENCY, NUCLEAR TYPE 1**

- Variants, genes, phenotypes in rare diseases
- Literature curation → consultant clinical geneticists in the UK

Data sources: UniProt

- Protein: sequence, annotation, function

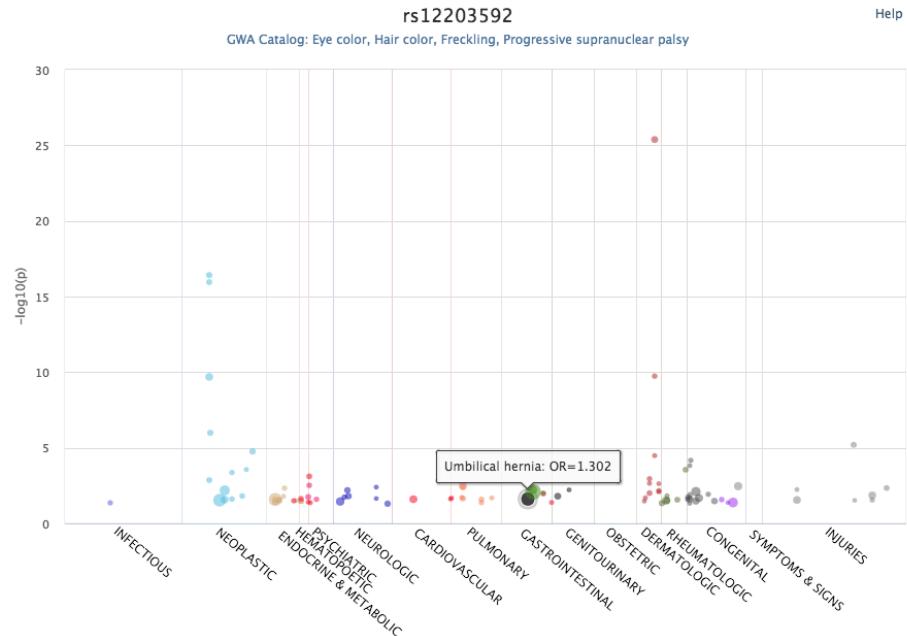


- Manual curation of coding variants in patients



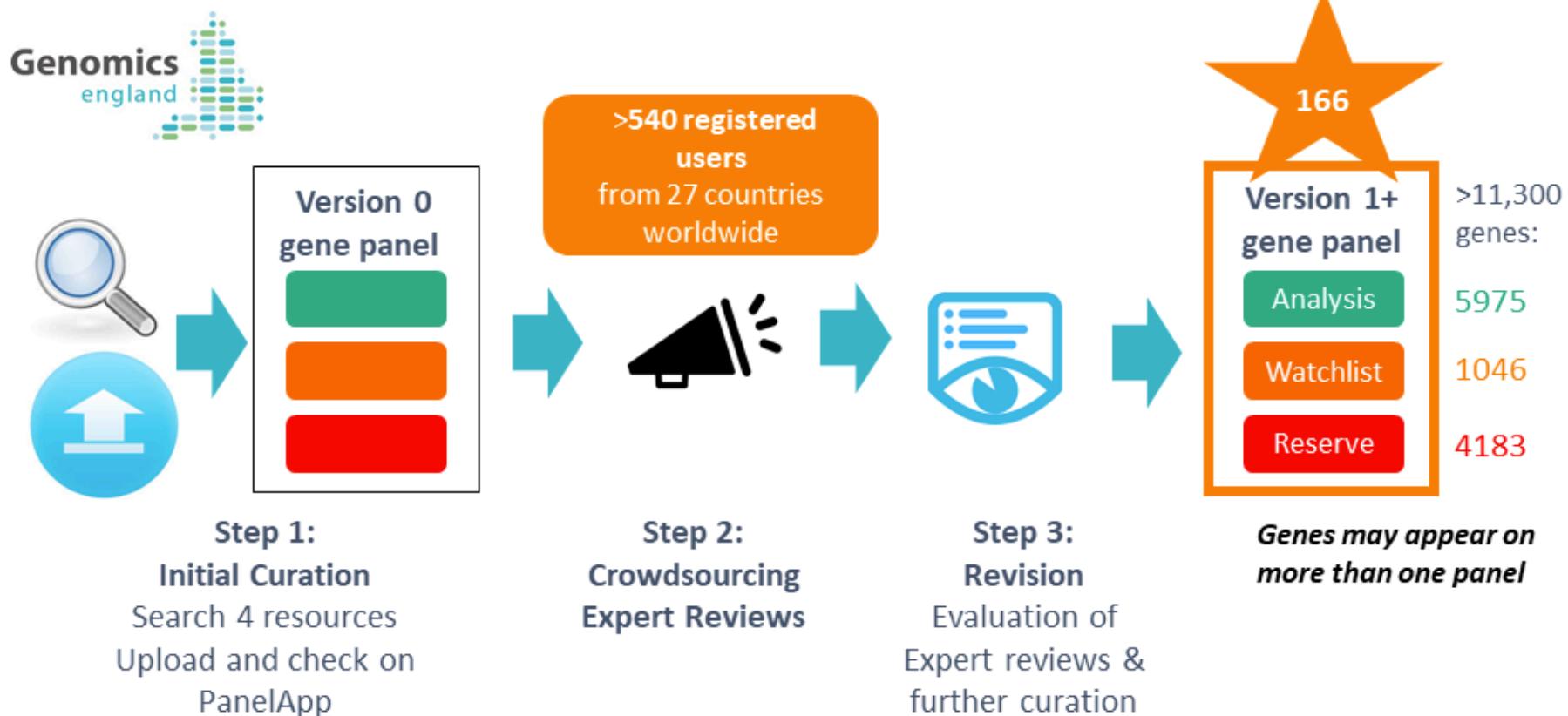
Data sources: PheWAS

- Phenome Wide Association Studies
- A variant associated with multiple phenotypes
- Clinical phenotypes derived from EMR-linked biobank BioVU
- ICD9 codes mapped to EFO



Data sources: GE PanelApp

- Aid clinical interpretation of genomes for the 100K project
- We include ‘green genes’ from version 1+ and phenotypes



**Germline
variants**

**Somatic
mutations**

Data sources: EVA

- With ClinVar information for rare diseases
- Clinical significance: pathogenic, protective

The screenshot shows the European Variation Archive (EVA) website. The header features the EVA logo and the text "European Variation Archive". Below the header is a navigation bar with links: Home, Submit Data, Study Browser, Variant Browser, Clinical Browser (which is highlighted in black), GA4GH, API, FAQ, and Feedback. The main content area is titled "ClinVar Browser" with a help icon. On the left, there is a "Filter" section with buttons for "Reset" and "Sub...", and dropdown menus for "Position" (set to GRCh37), "Assembly" (set to GRCh37), "Filter By:" (set to Chromosomal), and a specific position entry "2:48000000-49000000". The main table displays 10 records out of 960, showing columns for Position, Affect., Most Severe Consequence, Trait, Clinical Significance, and ClinVar ID. The first few rows show variants for the MSH6 gene at position 2:48000000, with various clinical significances like Benign, Lynch syndrome, and Uncertain significance.

...	Posi...	Affecte. i	A...	Most Severe Consequence...	Trait	Clinical Significance	ClinVar ...
2	480...	MSH6	T/G	upstream_gen...	Lynch synd...	Benign	RCV000...
2	480...	MSH6	G/A	upstream_gen...	Lynch synd...	Benign	RCV000...
2	480...	MSH6	C/T	upstream_gen...	Lynch synd...	Benign	RCV000...
2	480...	MSH6	C/T	upstream_gen...	Lynch synd...	Benign	RCV000...
2	480...	MSH6	G/T	5_prime_UTR...	Lynch synd...	Uncertain s...	RCV000...
2	480...	MSH6	G/T	5_prime_UTR...	Hereditary ...	conflicting ...	RCV000...



Data sources: The Cancer Gene Census

Census

Breakdown

Abbreviations

The cancer Gene Census is an ongoing effort to catalogue those genes for which mutations have been causally implicated in cancer. The original census and analysis was published in [Nature Reviews Cancer](#) and supplemental analysis information related to the paper is also available.

- Genes with mutations causally implicated in cancer
- Gene associated with a cancer plus other cancers associated with that gene

Data sources: IntOGen

The screenshot shows the homepage of the intOGen website. At the top is a navigation bar with an orange gradient background. From left to right, it contains: the intOGen logo (a stylized orange 'i' icon followed by the word 'intOGen'), a 'Search' button with a magnifying glass icon, a 'Downloads' button with a download icon, an 'Analysis' button with a gear icon, an 'About' button with a speech bubble icon, and a 'Sign In' button with a user profile icon.

Below the navigation bar is the main content area. On the left is a large, stylized orange 'i' icon composed of three nested circles. To its right, the word 'intOgen' is written in a lowercase, sans-serif font. To the right of 'intOgen', the words 'Integrative', 'Onco', and 'Genomics' are stacked vertically in a smaller, orange sans-serif font.

- Genes and somatic (driver) mutations, 28 cancer types
- Involvement in cancer biology
- Rubio-Perez et al. 2015

Data sources: ChEMBL

EMBL-EBI

ChEMBL

EBI > Databases > Small Molecules > ChEMBL Database > Home

Search ChEMBL... Compounds Targets Assays

Ligand Search Target Search Browse Targets Browse Drugs Browse Drug Targets

- Known drugs linked to a disease and a known target
- FDA approved for clinical trials or marketing



Data sources: Reactome



The image shows the Reactome homepage. At the top left, there is a logo with three rounded rectangles connected by arrows. Below it, the word "REACTOME" is written in large, white, serif capital letters, with "A CURATED PATHWAY DATABASE" in smaller letters underneath. To the right of the logo is a decorative graphic featuring a red sphere, blue curved lines, and green and blue abstract shapes. At the bottom of the page is a navigation bar with links: "About", "Content", "Documentation", "Tools", "Community", "Download", and "Contact". To the right of these links is a search bar containing the placeholder text "e.g. O95631, NTN1, signalin" and a "Search" button.

- Biochemical reactions and pathways
- Manual curation of pathways affected by mutations



Data sources: SLAPenrich

METHOD

Dissecting the genomic heterogeneity of cancer hallmarks' acquisition with SLAPenrich

Francesco Iorio^{1,5†}, Luz Garcia-Alonso^{1,5}, Jonathan Brammell², Iñigo Martincorena², David R Wille^{3,5}, Ultan McDermott^{2,5} and Julio Saez-Rodriguez^{1,4,5*†}

- Statistical framework to identify pathways significantly mutated in cancer patients from TCGA
- 374 pathways curated and mapped to hallmarks
- 25 cancer types

Data sources: PROGENy

ARTICLE

DOI: 10.1038/s41467-017-02391-6

OPEN

Perturbation-response genes reveal signaling footprints in cancer gene expression

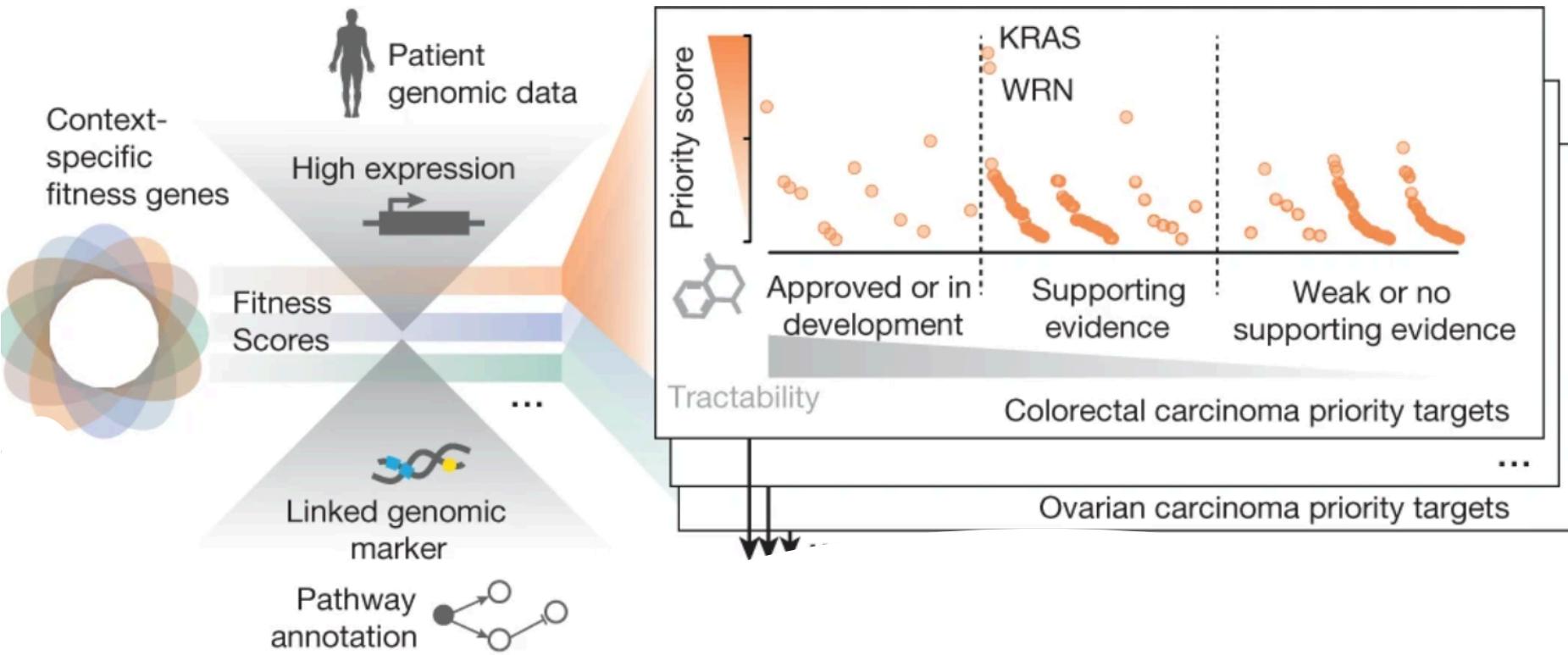
Michael Schubert¹, Bertram Klinger^{2,3}, Martina Klünemann^{2,3}, Anja Sieber^{2,3}, Florian Uhlitz^{2,3}, Sascha Sauer⁴, Mathew J. Garnett⁵, Nils Blüthgen^{2,3} & Julio Saez-Rodriguez^{2,3}

- Comparison of pathway activities between normal and primary samples from The Cancer Genome Atlas (TCGA)
- Inferred from RNA-seq: 9,250 tumour and 741 normal samples
- EGFR, hypoxia, JAK/STAT, MAPK, NFkB, PI3K, TGFb, TNFa, Trail, VEGF, and p53

Data sources: SysBio

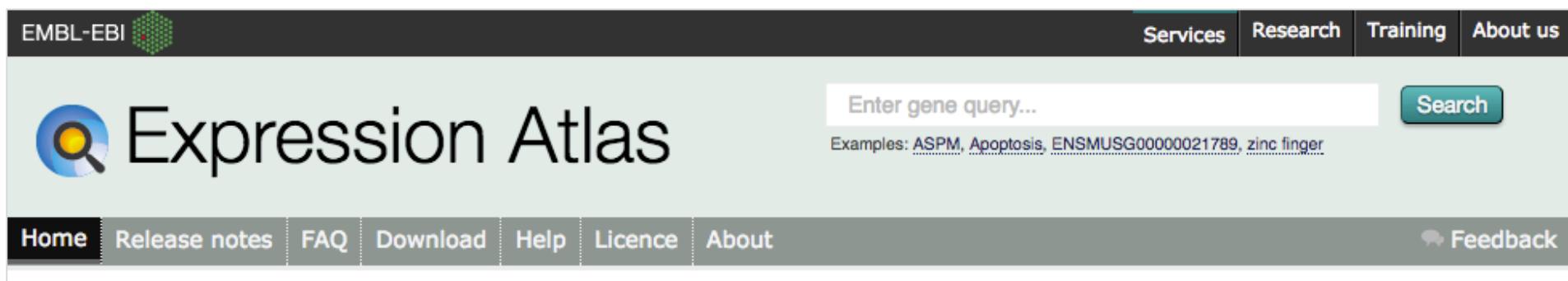
- Curation of four systems biology papers
- Available for six gene lists: ~ 400 genes
- Late onset Alzheimers, cognitive decline, CHD, and IBD
- Score: p-values or rank-based scores if available, otherwise 0.5 is assigned.

Data sources: CRISPR



- 624 genes prioritised following gene essentiality, tractability and biomarker information
- 1846 associations with diseases

Data sources: Expression Atlas



The screenshot shows the Expression Atlas website. At the top, there's a dark header bar with the EMBL-EBI logo on the left and navigation links for Services, Research, Training, and About us on the right. Below the header is a search bar with a placeholder "Enter gene query..." and a "Search" button. To the left of the search bar is a magnifying glass icon. The main title "Expression Atlas" is prominently displayed in large black font next to the search bar. Below the title is a navigation bar with links for Home, Release notes, FAQ, Download, Help, Licence, and About. On the far right of this bar is a "Feedback" link with a speech bubble icon.

- Baseline expression for human genes
 - target profile page
- Differential mRNA expression (healthy *versus* diseased):
 - target-disease associations



EMBL-EBI train online

Data sources: Europe PMC



Europe PMC

About Tools Developers Help Europe PMC plus

Search worldwide, life-sciences literature

- Mining titles, abstracts, full text in research articles
- Target and disease co-occurrence in the same sentence
- Dictionary (not NLP)



Data sources: PhenoDigm

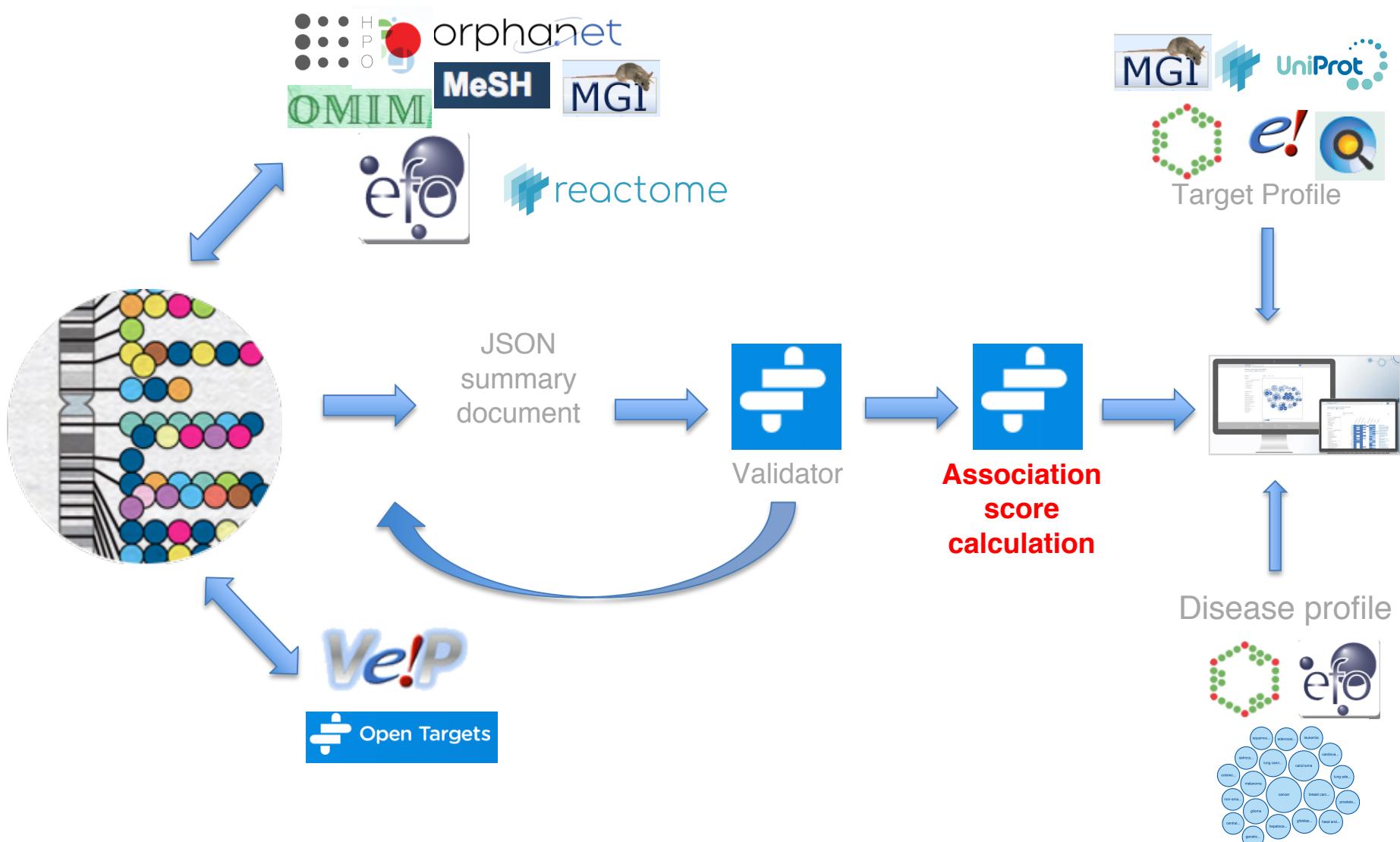
The screenshot shows the homepage of the PhenoDigm website. At the top, there is a dark header bar with the Wellcome Trust Sanger Institute logo on the left. To the right of the logo is a blue navigation bar with the following links: "ABOUT" (with a dropdown arrow), "Who we are", "Careers", "Study", "Sex in Science", "Groups", and "Campus". On the far right of the blue bar is a magnifying glass icon representing a search function. Below the header, the main title "Welcome to PhenoDigm (PHENOtype comparisons for DIsease and Gene Models)" is displayed in large, bold, black font. Underneath the title, there is a horizontal menu bar with three items: "Diseases" (which is highlighted in blue and underlined), "Tissue phenotype associations", and "Secondary phenotypes".

Welcome to PhenoDigm (PHENOtype comparisons for DIsease and Gene Models)

Diseases Tissue phenotype associations Secondary phenotypes

- Semantic approach to associate mouse models with diseases

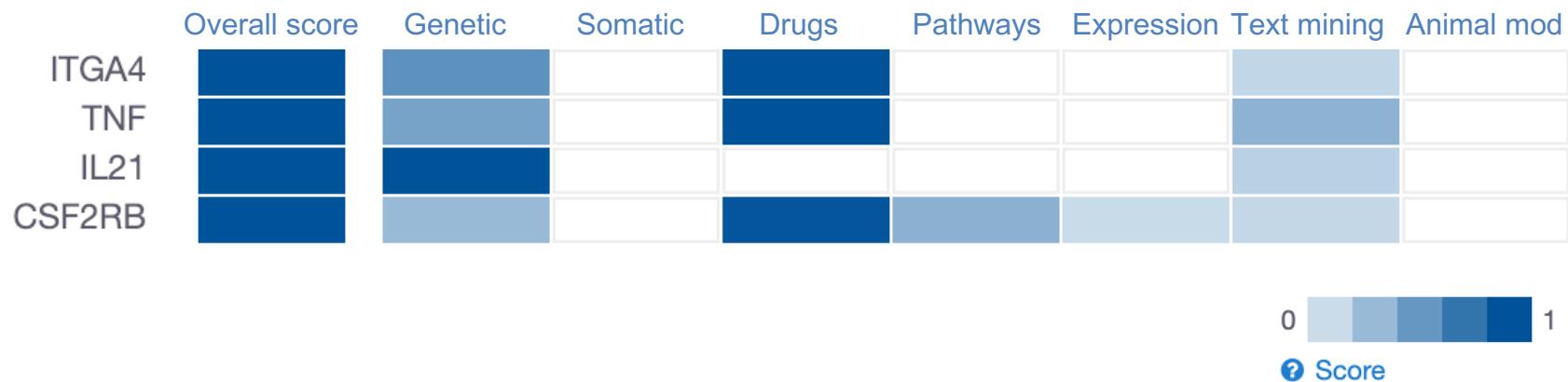
How the data* flows



* e.g. genetic variants from NHGRI-GWAS catalog

How well supported are the associations?

- Association score to convey confidence
- More evidence → stronger association score
- Different evidence are scored and weighted in different ways



Evidence score to overall score

Statistical integration, aggregation and scoring

- 1) Evidence score (e.g. one SNP from a GWAS paper)
- 2) Data source score (e.g. all SNPs from the GWAS catalog)
- 3) Data type score (e.g. all sources of Genetic associations)
- 4) Overall association score

<https://docs.targetvalidation.org/getting-started/scoring>

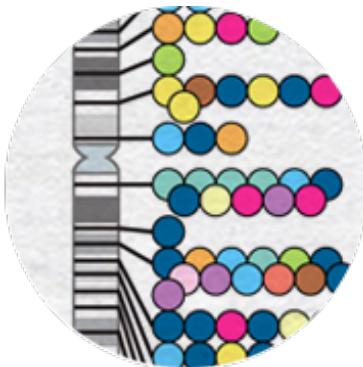
Computing the score for one evidence

$$\text{score} = f * s * c$$

f, relative occurrence of a target-disease evidence

s, strength of the effect of the variant

c, confidence of the observation for the target-disease evidence



f = sample size (cases and controls)

s = predicted functional consequence (VEP)

c = p value reported in the paper

(Factors affecting the relative strength of GWAS Catalog evidence)

Factors affecting the relative strength of the evidence

Somatic mutations	Cancer Gene Census (functional consequence score of variants); European Variation Archive (functional consequence score of variants); IntOGen (binned score based on tumour type categories. If the gene has several signals of positive selections in the tumour, the score will be 0.25. If the gene is already described as a cancer gene and exhibits a signal of positive selection in a tumor type, the score will be 0.5. If in addition to a signal of positive selection, the gene is functionally connected to other genes in the same tumor type, the score will be 0.75)
Drugs	ChEMBL (Clinical trials phase binned score. Scores will be 0.09 for phase 0, 0.1 for phase I, 0.2 for Phase II, 0.7 for Phase III, and 1 for Phase IV drugs)
Affected pathways	Reactome (functional consequence of 1 for a pathway inferred by a curator). SLAPenrich evidence is scored according to Iorio F et al 2018 followed by quantifying, in large cohorts of cancer patients, the divergence of the total number of samples with genomic alterations in a Reactome-pathway from its expectation, accounting for mutational burdens and total exonic block lengths of genes in that pathway. PROGENy evidence is scored per sample and pathway following a modifications of the original implementation described by Schubert et al. 2016 . Further details can be found elsewhere .
RNA Expression	Expression Atlas score (normalised p-value, normalised expression fold change and normalised percentile rank)
Text mining	Europe PMC (weighting document sections, sentence locations and title for full text articles and abstracts (Kafkas et al 2016))
Animal models	PhenoDigm (similarity score between a mouse model and a human disease described by Smedley et al 2013)

Harmonic sum

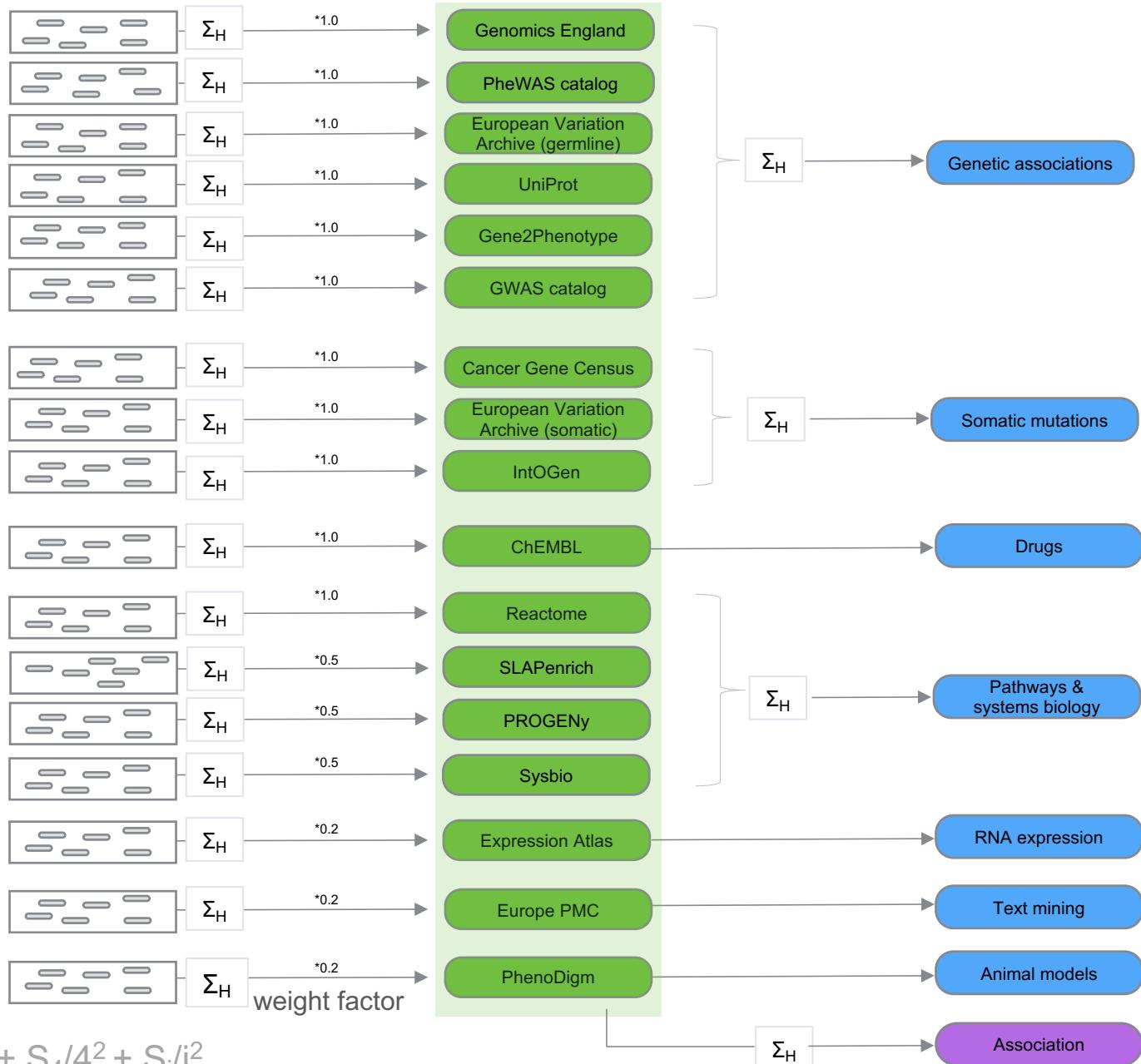
Score: 0 to 1 (max)

Calculated at four levels:

- Evidence
- Data source
- Data type
- Overall

Aggregation with Σ_H

Note: Each data set has its own scoring and ranking scheme



$$\Sigma_H = S_1 + S_2/2^2 + S_3/3^2 + S_4/4^2 + \dots + S_i/i^2$$

Open Targets Platform REST API*

The image shows a YouTube video player interface. In the top left corner, there is a 'YouTube GB' logo. To its right is a search bar with the placeholder text 'Search'. The main content area displays a screenshot of the Open Targets Platform website. The screenshot shows a search result for 'PTEN' with the title '674 diseases associated with PTEN'. Below this, there is a 'Filter by' section with dropdown menus for 'Data type' (Genetic associations, Somatic mutations, Drugs, Affected pathways, RNA expression, Test mining, Animal models) and 'Therapeutic areas'. A large heatmap visualization is shown, with a legend indicating 'Affected cancer predisposing & PTEN hamartoma syndromes' and 'Overgrowth syndromes'. The bottom of the screenshot shows a URL: 'http://api.opentargets.io/v3/platform/public/association/filter?target=ENSG00000171862&select=' followed by a timestamp '28:56 / 40:43'. To the right of the screenshot, a large block of JSON code is displayed, representing the API response for the query. The JSON includes fields like 'from', 'took', 'next', 'data_version', 'therapeutic_areas', 'query', 'total', and 'data', which contains detailed information about the target gene PTEN.

```
{
  "from": 0,
  "took": 25,
  "- next": [
    1.4207987,
    "ENSG00000171862-Orphanet_210548"
  ],
  "data_version": "17.12",
  "+ therapeutic_areas": [],
  "+ query": {},
  "total": 674,
  "- data": [
    {
      "- target": {
        "- gene_info": {
          "symbol": "PTEN",
          "name": "phosphatase and tensin homolog"
        },
        "id": "ENSG00000171862"
      },
      "- association_score": {
        "- datatypes": {
          "literature": 0.3241324475302135,
          "rna_expression": 0,
          "genetic_association": 1,
          "somatic_mutations": 1,
          "known_drug": 0,
          "animal_model": 0,
          "affected_pathway": 1
        },
        "overall": 1
      },
      "- datasources": {
        "slapenrich": 0.817215326415924,
        "expression_atlas": 0,
        "europcpmc": 0.3241324475302135
      }
    }
  ]
}
```

<https://youtu.be/KQbfhwpeEvc>