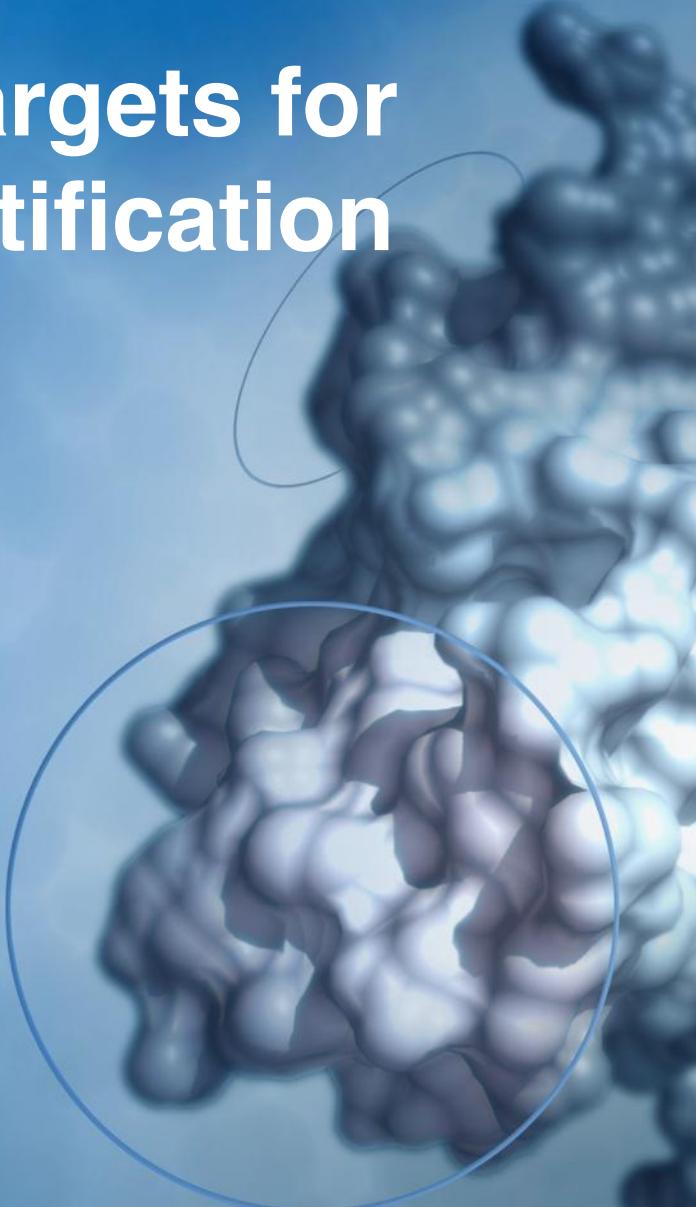


# Mining gene and disease associations with Open Targets for improved drug target identification

**Webinar**  
Bristol-Myers Squibb

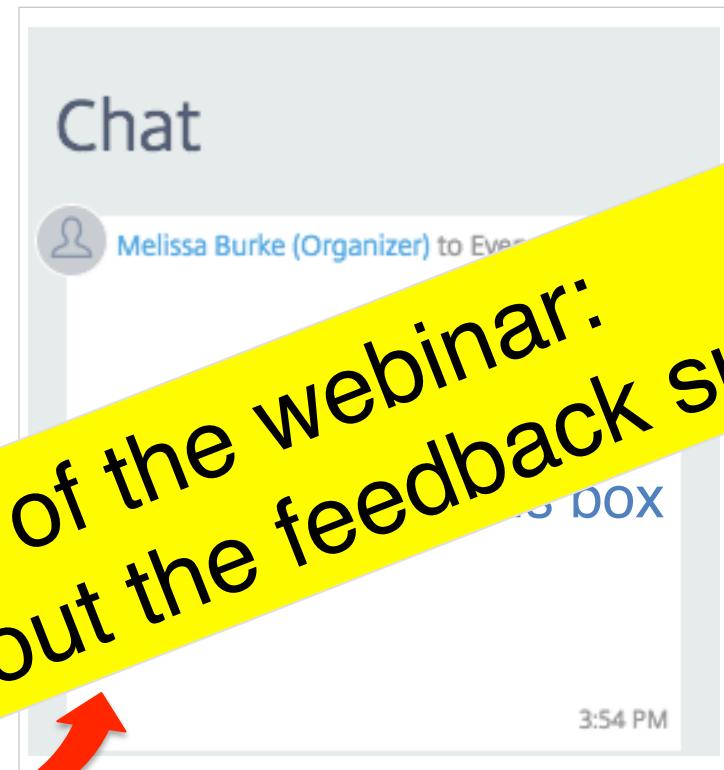
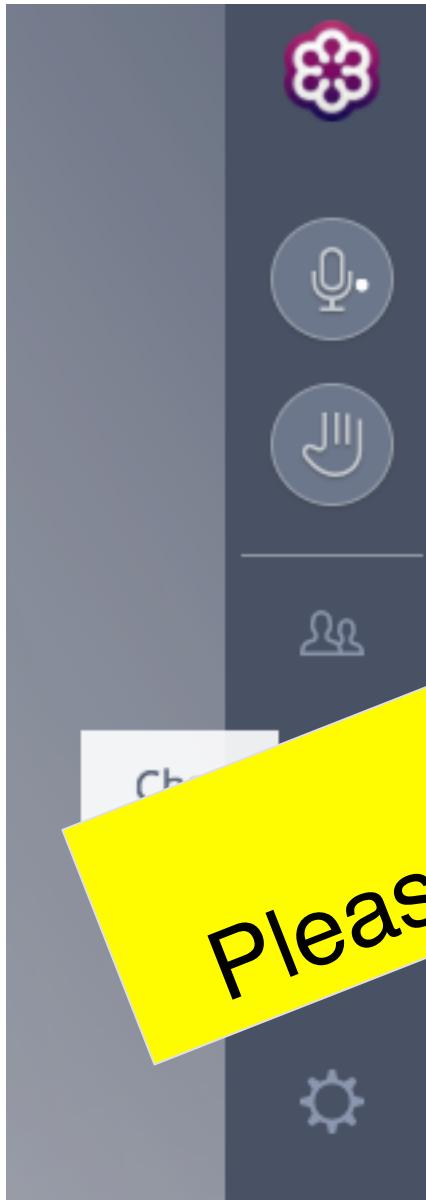
**Denise Carvalho-Silva**  
Wellcome Genome Campus, United Kingdom  
Open Targets Consortium  
Core Bioinformatics team



# The logistics for the next hour

- ‘Materials’ section: presentation.pdf, coursebook.pdf
-  muted
- Take questions at the end: look for the chat box!

# Chat box



Please fill out the feedback survey

End of the webinar:

Address chat to 'Everyone'

# Webinar's objectives

What is the Open Targets Consortium?

What is the Open Targets Platform?

How to navigate the Platform?

How to connect with the team



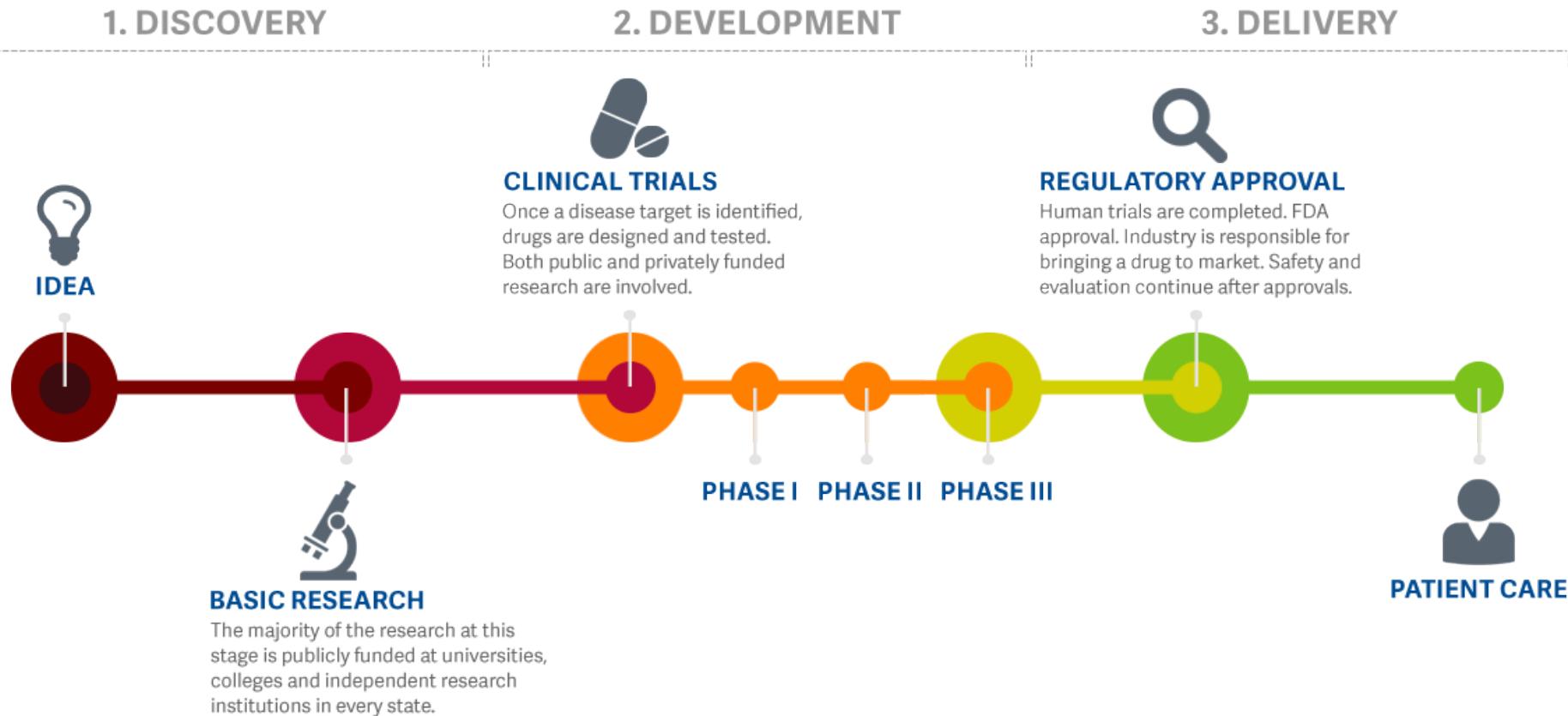
# Outline

- Introduction
- Live demos
- Get in touch

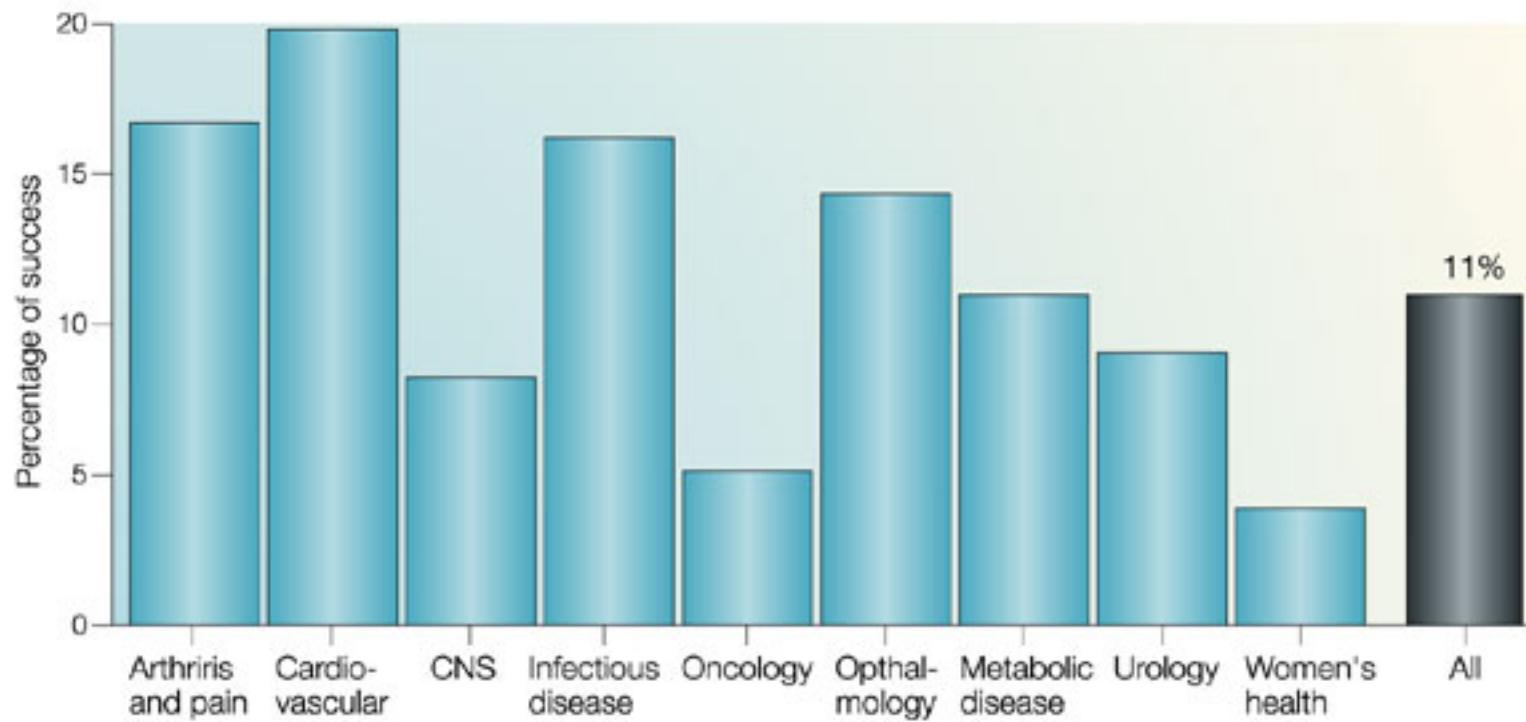
# Outline

- Introduction
- Live demos
- Get in touch

# Drug discovery path: timeline



# Drug discovery: the challenges



Lengthy, costly, low success rate, **HIGH ATTRITION RATES**



Professor Sir  
Mike Stratton  
Director, Sanger Institute

Can we improve  
target identification?



Mark Allalun-Smith,  
Pharmaceuticals R&D  
GlaxoSmithKline



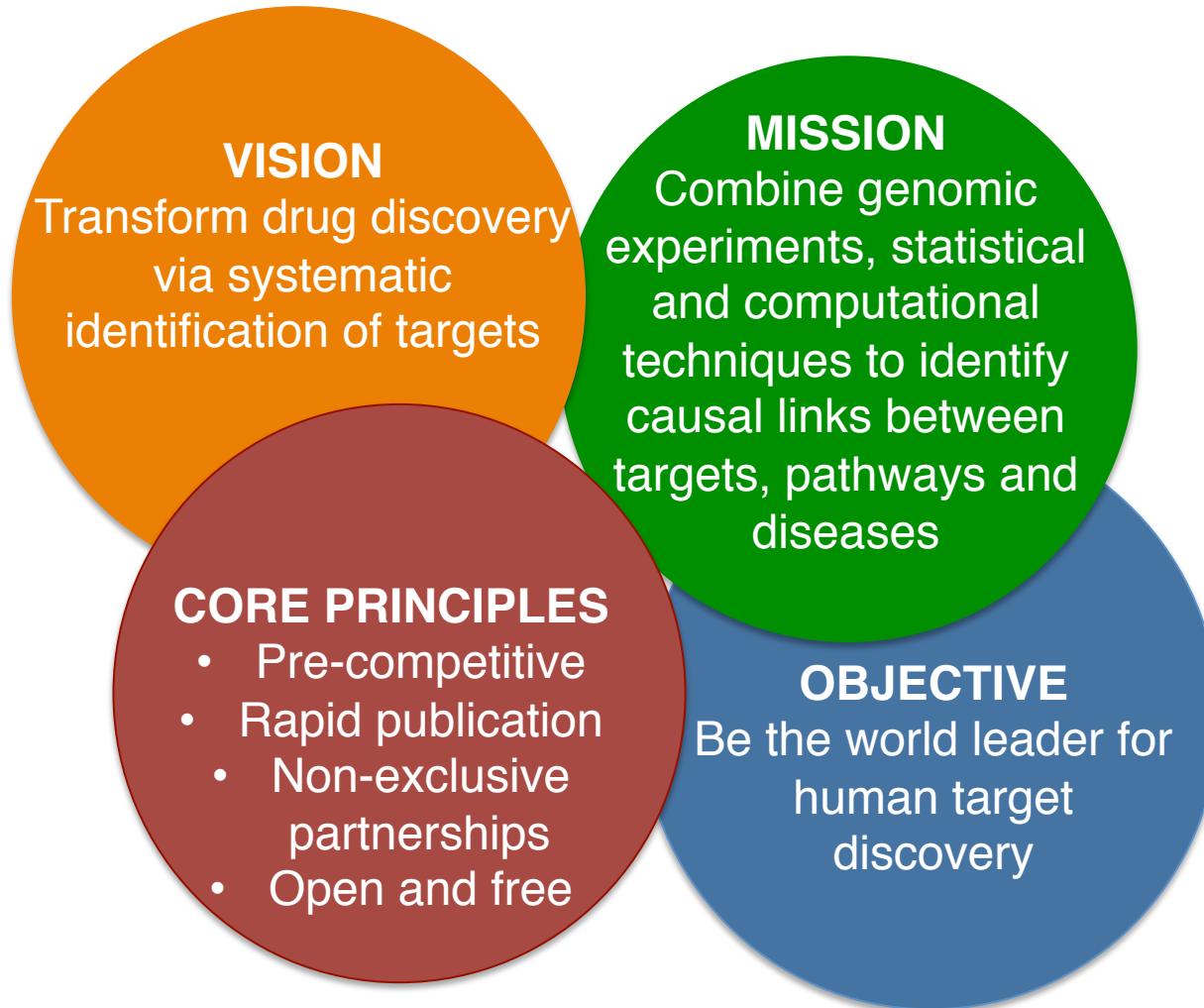
Professor Dame  
Janet Thornton  
former Director, EMBL-EBI

# Open Targets founded (formerly CTTV)

But one institution  
can not do it alone.



# Open Targets Consortium\*



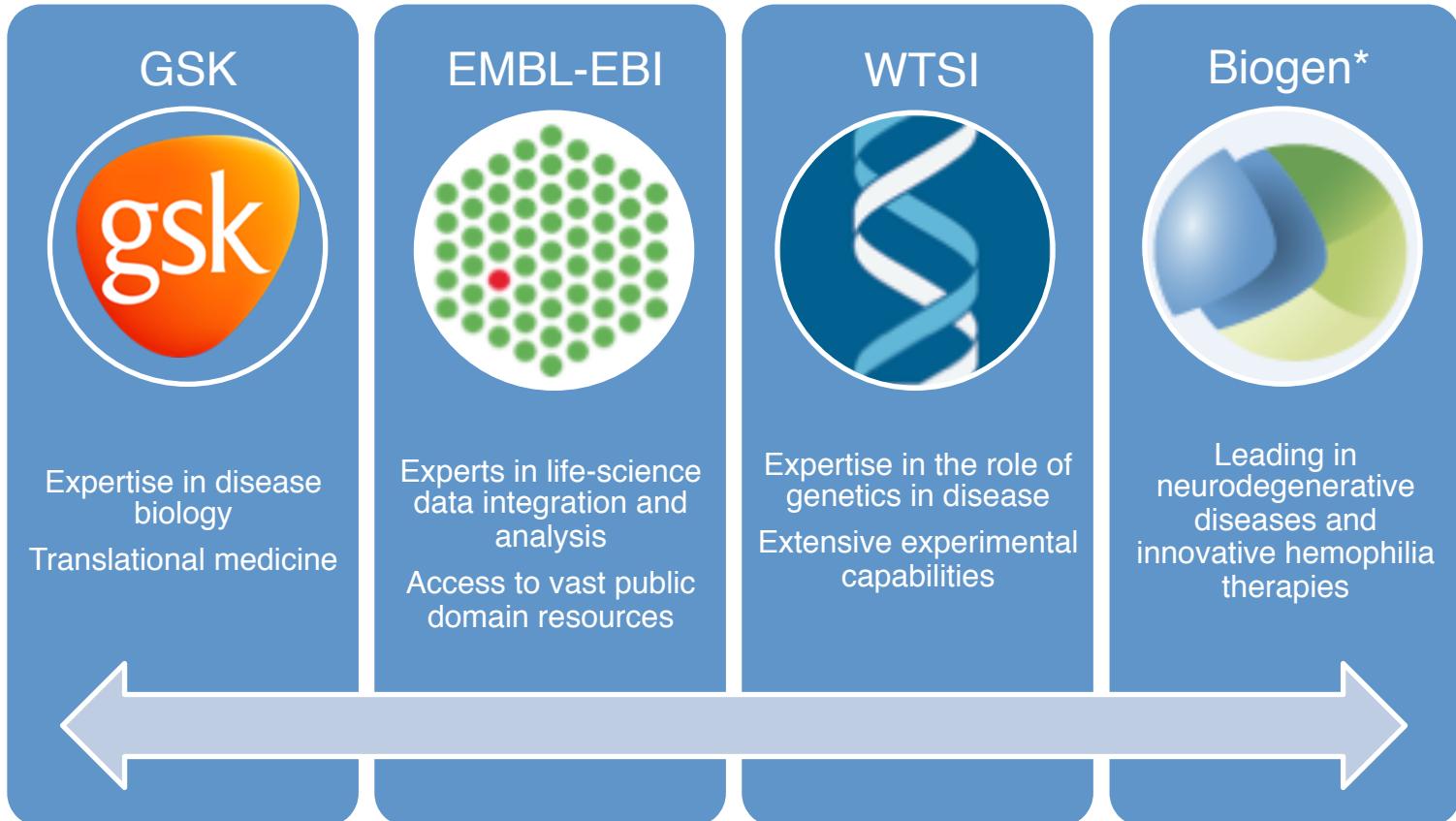
\* Launched in March 2014  
Three founding partners



EMBL-EBI



# Who is Open Targets now?



\* Biogen joined the consortium in February 2016

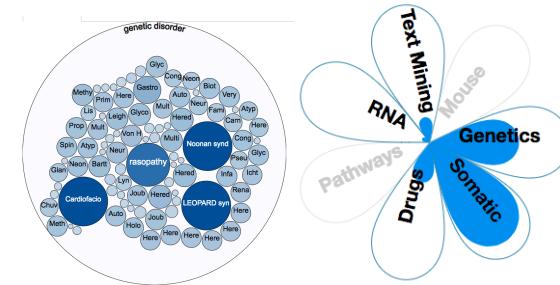
# Two major areas of work in Open Targets

## Experimental projects



Generate new evidence  
CRISPR/Cas9, Organoids

## Core bioinformatics pipelines



Integration of available data  
Web interface, API and data dumps

Concurrent  
[www.opentargets.org/projects](http://www.opentargets.org/projects)

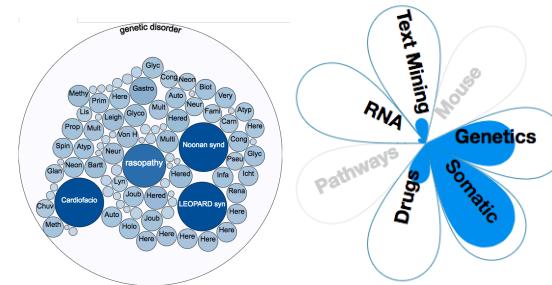
# Two major areas of work in Open Targets

## Experimental projects



Generate new evidence  
CRISPR/Cas9, Organoids

## Core bioinformatics pipelines



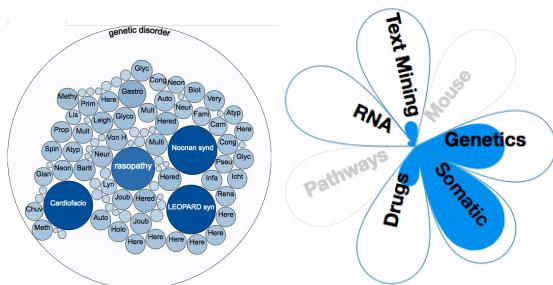
Database for data integration  
Web interface, API and data dumps

Concurrent  
[www.opentargets.org/projects](http://www.opentargets.org/projects)

# Open Targets Platform\*

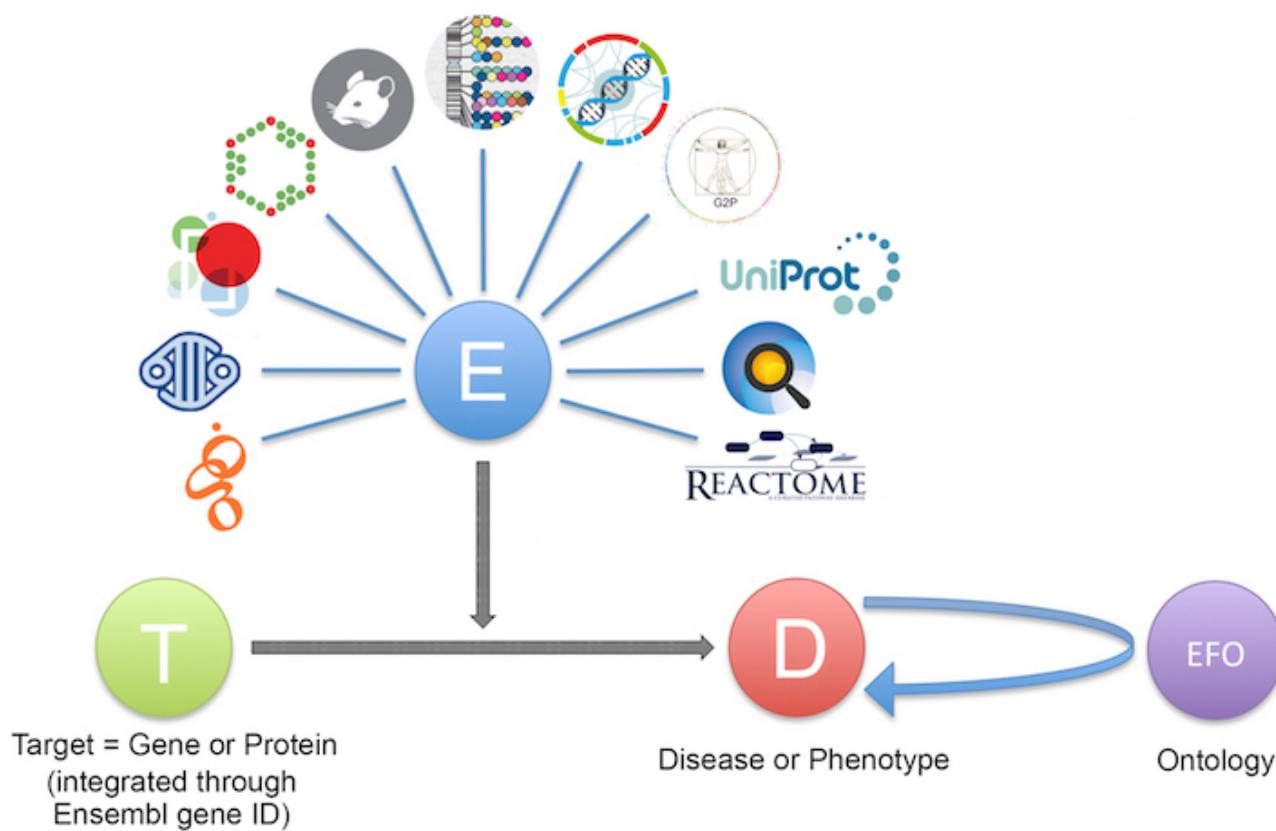
- Developed by the Core Bioinformatics team at EMBL-EBI
- Allow users to identify target–disease associations
- Improvements driven by you

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



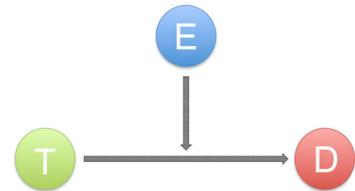
# Open Targets Platform

## Evidence model for target and disease associations



# Evidence from publicly available data

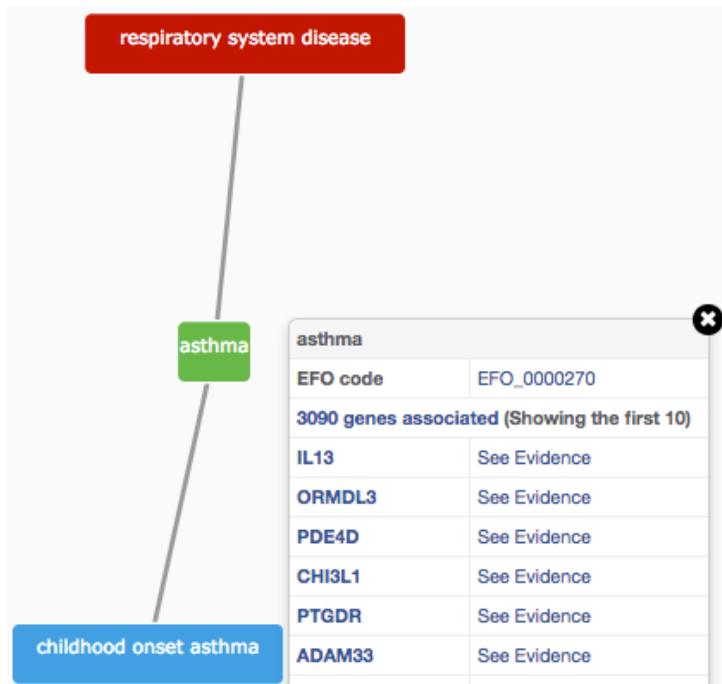
- Similar data sources are grouped into data types



Data sources	Data types
GWAS catalog, UniProt, EVA, G2P	Genetic associations
Cancer Gene Census, EVA, IntOgen	Somatic mutations
Expression Atlas	RNA expression
ChEMBL	Drugs
Reactome	Affected pathways
Europe PMC	Text mining
PhenoDigm	Animal models
<b>Your favourite data?</b>	<b>Let us know!</b>

# Experimental Factor Ontology\* (EFO)

- Ontology: dictionary of relationships between entities
- EFO: way to organise experimental variables (e.g. diseases)

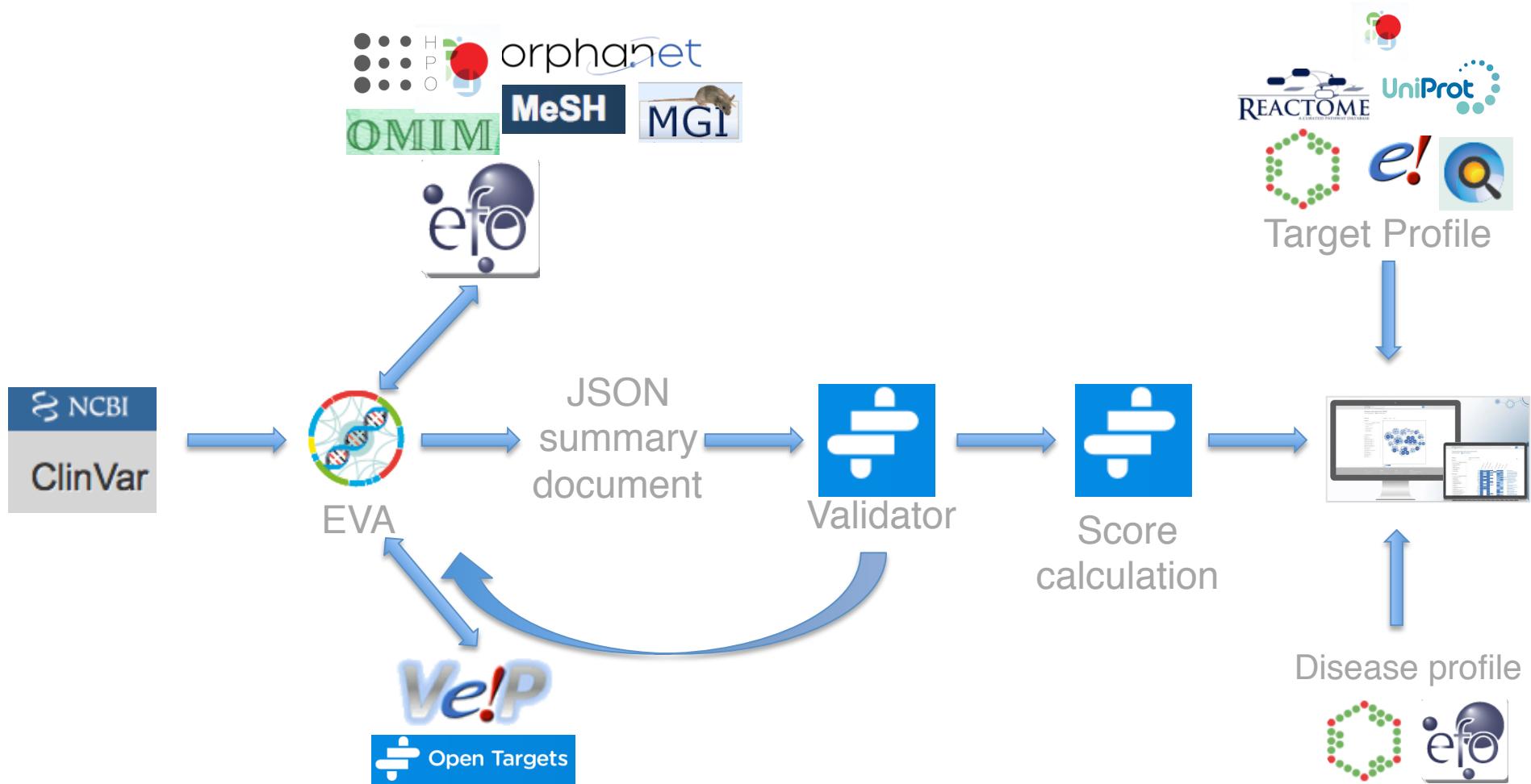


controlled vocabulary  
+  
hierarchy (relationship)

\* <https://www.ebi.ac.uk/efo/>

Increases the richness of annotation  
Promotes consistency  
Allow for easier and automatic integration

# Data flow pipeline

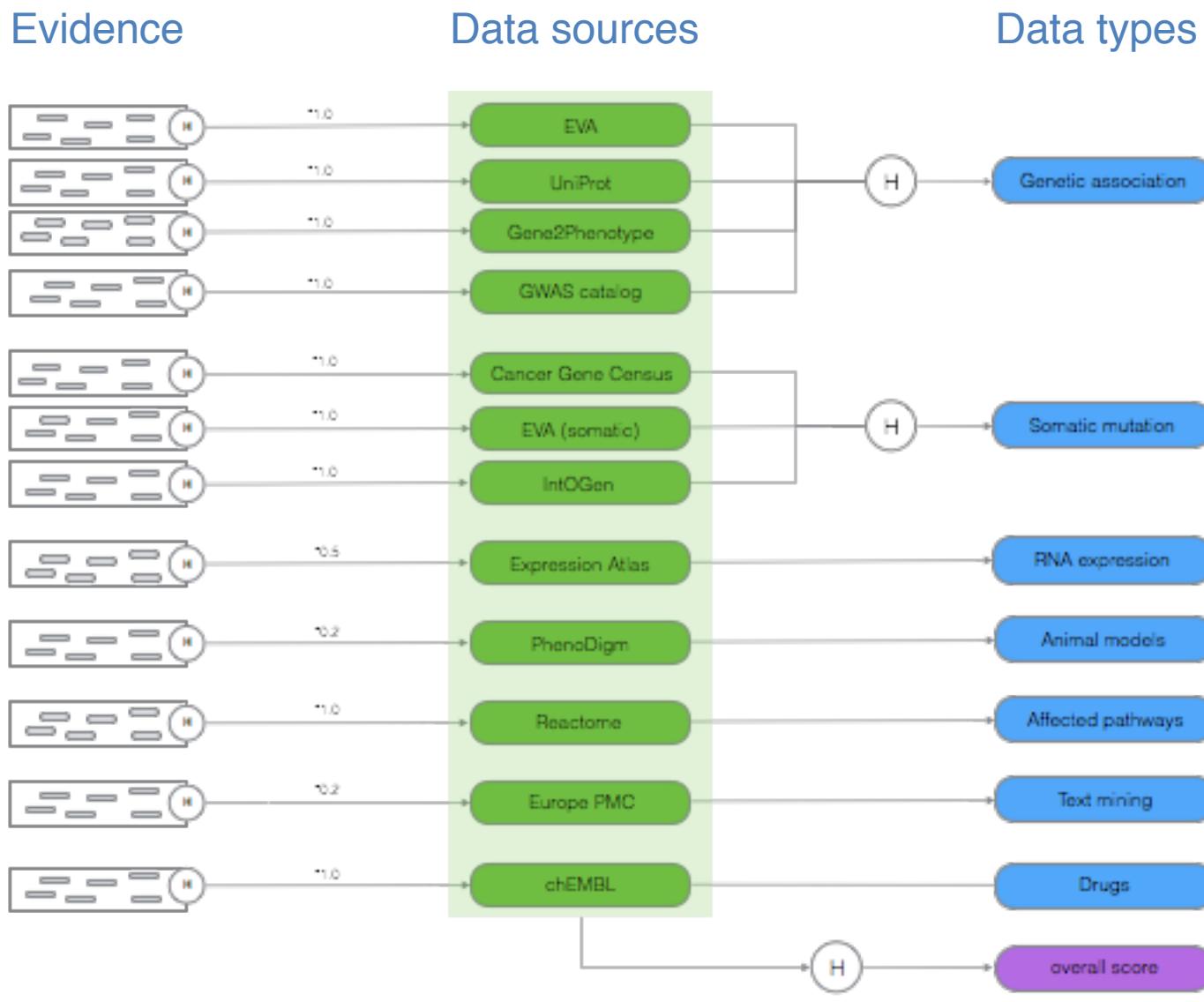


\* JSON summary document = IDs (gene, disease, papers) + curation (e.g. manual) + evidence + source + stats for the score

# JSON summary document

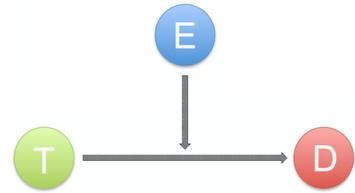
```
ArrayExpress Experiment overview"}], "experiment_overview": "Transcription profiling of human intestinal epithelium in patients suffering from inflammatory bowel disease.", "access_level": "public", "id": "866f4a506d5a01ebcea8bd84bb5485c8", "unique_association_fields": {"comparison_name": "'active ulcerative colitis' vs 'normal'", "geneID": "http://identifiers.org/ensembl/ENSG00000073605", "study_id": "http://identifiers.org/gxa.expt/E-MEXP-2083"}, "sourceID": "expression_atlas", "target": {"gene_info": {"symbol": "GSDMB", "geneid": "ENSG00000073605", "name": "gasdermin B"}, "id": "ENSG00000073605", "activity": "increased_transcript_level", "target_type": "transcript_evidence"}, "disease": {"id": "EFO_0000729", "efo_info": {"path": [[{"id": "EFO_0000405", "label": "ulcerative colitis"}, {"id": "EFO_0003767", "label": "therapeutic area"}, {"id": "EFO_0000729", "label": "immune system disease"}, {"id": "EFO_0000540", "label": "digestive system disease"}], "label": "ulcerative colitis", "therapeutic_area": {"labels": ["immune system disease", "digestive system disease"]}, "codes": [{"id": "EFO_0000405", "label": "ulcerative colitis"}, {"id": "EFO_0000540", "label": "digestive system disease"}]}, "biosample": {"name": "intestinal epithelial cell", "id": "http://purl.obolibrary.org/obo/CL_0002563"}, "data_release": "16.12", "type": "rna_expression", "scores": {"association_score": 0.014397535094125722}}  
{"validated_against_schema_version": "1.2.3", "evidence": {"confidence_level": "low", "log2_fold_change": {"value": 1, "percentile_rank": 97}, "unique_experiment_reference": "STUDYID_E-GEOID-48258", "resource_score": {"value": 0.0328, "type": "pvalue", "method": {"description": "Moderated t-statistics computed with limma version 3.16.8. Independent filtering performed using genefilter version 1.46.1, with gene variances as filter statistic. p-values adjusted using Benjamini & Hochberg (1995) FDR correction. By default, significant differential expression called if adjusted p <= 0.05."}}, "date_asserted": "2015-07-20T22:01:23Z", "reference_replicates_n": 3, "evidence_codes_info": [{"eco_id": "ECO_0000356", "label": "differential gene expression evidence from microarray experiment"}], "test_replicates_n": 3, "is_associated": true, "provenance_type": {"database": {"version": "18-11-2016", "id": "Expression_Atlas"}, "evidence_codes": [{"id": "ECO_0000356"}], "reference_sample": "control", "comparison_name": "'fludarabine, 10 micromolar' vs 'control'", "test_sample": "fludarabine, 10 micromolar", "url": [{"url": "http://www.ebi.ac.uk/gxa/experiments/E-GEOID-48258?geneQuery=ENSG00000179889", "nice_name": "Gene expression in Expression Atlas"}, {"url": "http://www.ebi.ac.uk/gxa/genes/ENSG00000179889", "nice_name": "Basel"}]}
```

# Score approach and aggregation



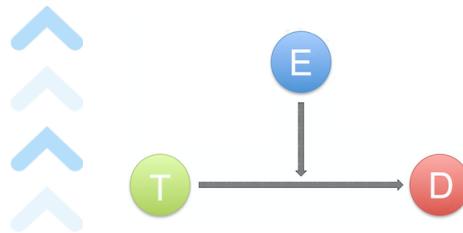
Evidence score and aggregation with the harmonic sum

$$H = S_1 + S_2/2^2 + S_3/3^2 + S_4/4^2 + S_i/i^2$$



# Currently: Integration of existing data

## Public Databases and Pipelines



Open  
Targets  
Platform

Target symbol	Association score	Genetic association	Somatic mutations	Drugs	Affects
TNFRSF1A					
KCNB2					
IL2RA					

Open Targets experimental data: NEW  
Physiologically relevant and at scale

Oncology



Immunology



Neurodegeneration



Generated  
as we speak

# Outline

- Introduction
- Live demos
- Get in touch

Which targets are associated with a disease?



## Demo 1:

Disease centric workflow

Open Targets Platform

Find new targets for drug discovery

Search for a target or disease

Try: BRAF PTEN Asthma Inflammatory bowel disease

Feedback

Follow us



multiple sclero

**multiple sclerosis**  
2697 targets associated

Disease

An autoimmune disorder mainly affecting young adults and characterized by destruction of myelin in the central nervous system. Pathologic findings include multiple sharply demarcated areas of demyelination throughout the white matter of the central nervous system. Clinical manifestations include vis...

Targets

MBP myelin basic protein

Diseases

relapsing-remitting multiple sclerosis  
autoimmune disease > multiple sclerosis > relapsing-remitting multiple ...

chronic progressive multiple sclerosis  
autoimmune disease > multiple sclerosis > chronic progressive multiple...

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

# Choose your favourite internet browser\*

\*Supported ones: Internet Explorer 11 (and above), Chrome, Firefox and Safari

What is the evidence supporting the  
*CD86* - multiple sclerosis association?



## Demo 2:

Evidence for *CD86* in multiple sclerosis

**CD86**  
CD86 molecule  
Synonyms: B7.2, B7-2, CD28LG2  
Receptor involved in the costimulatory signal essential for T-lymphocyte proliferation and interleukin-2 production, by binding CD28 or CTLA-4. May play a critical role in the early events of T-cell a...

**multiple sclerosis**  
Synonyms: MS (Multiple Sclerosis), MS, MULTIPLE SCLEROSIS ACUTE FULMINATING, Disseminated Sclerosis, Sclerosis...  
An autoimmune disorder mainly affecting young adults and characterized by destruction of myelin in the central nervous system. Pathologic findings include multiple sharply demarcated areas of demyelin...

Evidence for a T-D association



[https://www.targetvalidation.org/evidence/ENSG00000114013/EFO\\_0003885](https://www.targetvalidation.org/evidence/ENSG00000114013/EFO_0003885)



# Demo 3: Searching for many targets at once



We have a list of 26 possible targets for IBD (inflammatory bowel disease).

Are these targets represented in other diseases?

What are the pathways represented in this set of targets?

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



# Alternative ways to access the data

The screenshot shows a web browser window with the URL <https://www.targetvalidation.org/download> in the address bar. The page itself has a blue header with the Open Targets Platform logo and navigation icons. The main content area is titled "Data Download" and contains text explaining that all data from targetvalidation.org is available for download as compressed JSON files. It describes the availability of associations and evidence objects via API methods. Below this, a section titled "2017 Feb (Latest)" lists two download links: "Association objects (2016-12-09, 215MB, md5sum)" and "Evidence objects (2016-12-09, 4.35Gb, md5sum)".

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.

**2017 Feb (Latest)**

- Association objects (2016-12-09, 215MB, md5sum)
- Evidence objects (2016-12-09, 4.35Gb, md5sum)

# Alternative ways to access the data



Open/Hide | List operations | Expand operations

public : Publicly supported stable API.

GET /public/evidence

POST /public/evidence

GET /public/evidence/filter

POST /public/evidence/filter

GET /public/association

GET /public/association/filter

POST /public/association/filter

GET /public/search

GET /public/auth/request\_token

GET /public/auth/validate\_token

GET /public/utils/ping

GET /public/utils/version

GET /public/utils/stats

- Paste the URL in a location bar in a browser
- Use the terminal window (e.g. with CURL)
- Use our clients (i.e. R and Python)

<https://www.targetvalidation.org/documentation/api>

# Wrap up

## Open Targets Platform:

For drug target ID and selection in drug delivery

Rank target and disease association sources

Integrated information

Intuitive interface

Oh Yes!  
And all is 100% free  
and open source

# We support decision-making

Which targets are associated with a disease?

Can I find out about the mechanisms of the disease?

Are there FDA drugs for this association?

...



Open Targets

# How to cite us

Published online 8 December 2016

*Nucleic Acids Research*, 2017, Vol. 45, Database issue D985–D994  
doi: 10.1093/nar/gkw1055

## Open Targets: a platform for therapeutic target identification and validation

Gautier Koscielny<sup>1,2,\*</sup>, Peter An<sup>1,3</sup>, Denise Carvalho-Silva<sup>1,4</sup>, Jennifer A. Cham<sup>1,4</sup>, Luca Fumis<sup>1,4</sup>, Rippa Gasparyan<sup>1,3</sup>, Samiul Hasan<sup>1,2</sup>, Nikiforos Karamanis<sup>1,4</sup>, Michael Maguire<sup>1,4</sup>, Eliseo Papa<sup>1,3</sup>, Andrea Pierleoni<sup>1,4</sup>, Miguel Pignatelli<sup>1,4</sup>, Theo Platt<sup>1,3</sup>, Francis Rowland<sup>1,4</sup>, Priyanka Wankar<sup>1,3</sup>, A. Patrícia Bento<sup>1,4</sup>, Tony Burdett<sup>1,4</sup>, Antonio Fabregat<sup>1,4</sup>, Simon Forbes<sup>1,5</sup>, Anna Gaulton<sup>1,4</sup>, Cristina Yenyxe Gonzalez<sup>1,4</sup>, Henning Hermjakob<sup>1,4,6</sup>, Anne Hersey<sup>1,4</sup>, Steven Jupe<sup>1,4</sup>, Şenay Kafkas<sup>1,4</sup>, Maria Keays<sup>1,4</sup>, Catherine Leroy<sup>1,4</sup>, Francisco-Javier Lopez<sup>1,4</sup>, Maria Paula Magarinos<sup>1,4</sup>, James Malone<sup>1,4</sup>, Johanna McEntyre<sup>1,4</sup>, Alfonso Munoz-Pomer Fuentes<sup>1,4</sup>, Claire O'Donovan<sup>1,4</sup>, Irene Papatheodorou<sup>1,4</sup>, Helen Parkinson<sup>1,4</sup>, Barbara Palka<sup>1,4</sup>, Justin Paschall<sup>1,4</sup>, Robert Petryszak<sup>1,4</sup>, Naruemon Pratanwanich<sup>1,4</sup>, Sirarat Sarntivijal<sup>1,4</sup>, Gary Saunders<sup>1,4</sup>, Konstantinos Sidiropoulos<sup>1,4</sup>, Thomas Smith<sup>1,4</sup>, Zbyslaw Sondka<sup>1,5</sup>, Oliver Stegle<sup>1,4</sup>, Y. Amy Tang<sup>1,4</sup>, Edward Turner<sup>1,4</sup>, Brendan Vaughan<sup>1,4</sup>, Olga Vrousou<sup>1,4</sup>, Xavier Watkins<sup>1,4</sup>, Maria-Jesus Martin<sup>1,4</sup>, Philippe Sanseau<sup>1,2</sup>, Jessica Vamathevan<sup>4</sup>, Ewan Birney<sup>1,4</sup>, Jeffrey Barrett<sup>1,4,5</sup> and Ian Dunham<sup>1,4,\*</sup>

<sup>1</sup>Open Targets, Wellcome Genome Campus, Hinxton, Cambridge, CB10 1SD, UK, <sup>2</sup>GSK, Medicines Research Center, Gunnels Wood Road, Stevenage, SG1 2NY, UK, <sup>3</sup>Biogen, Cambridge, MA 02142, USA, <sup>4</sup>European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Hinxton, Cambridge, CB10 1SD, UK, <sup>5</sup>Wellcome Trust Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge, CB10 1SA, UK and <sup>6</sup>National Center for Protein Research, No. 38, Life Science Park Road, Changping District, 102206 Beijing, China

Received August 19, 2016; Revised October 19, 2016; Editorial Decision October 20, 2016; Accepted November 03, 2016

# Acknowledgements



# Outline

- Introduction
- Live demos
- Get in touch

# Support, dissemination, GIFs



[support@targetvalidation.org](mailto:support@targetvalidation.org)



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



@targetvalidate



[blog.opentargets.org/](http://blog.opentargets.org/)



[www.facebook.com/OpenTargets/](http://www.facebook.com/OpenTargets/)



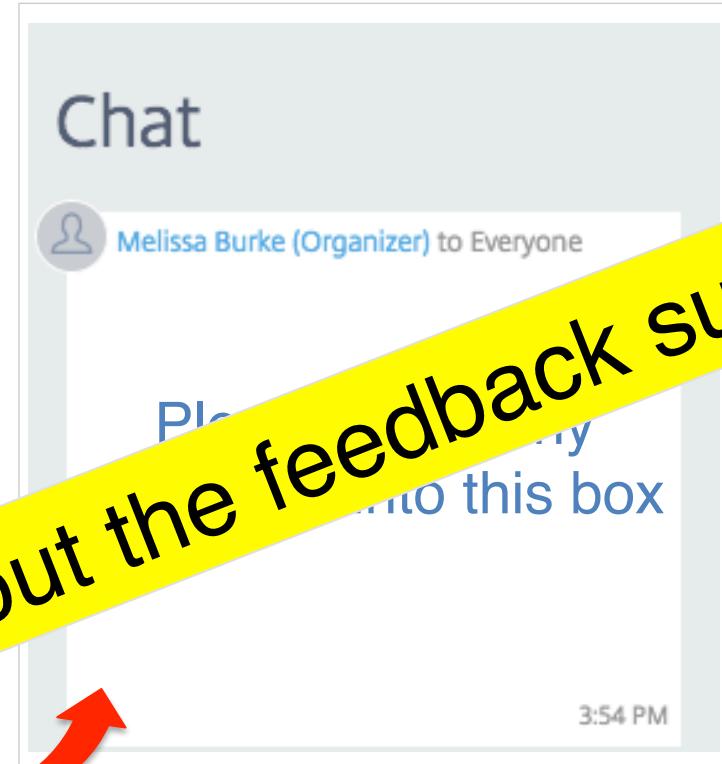
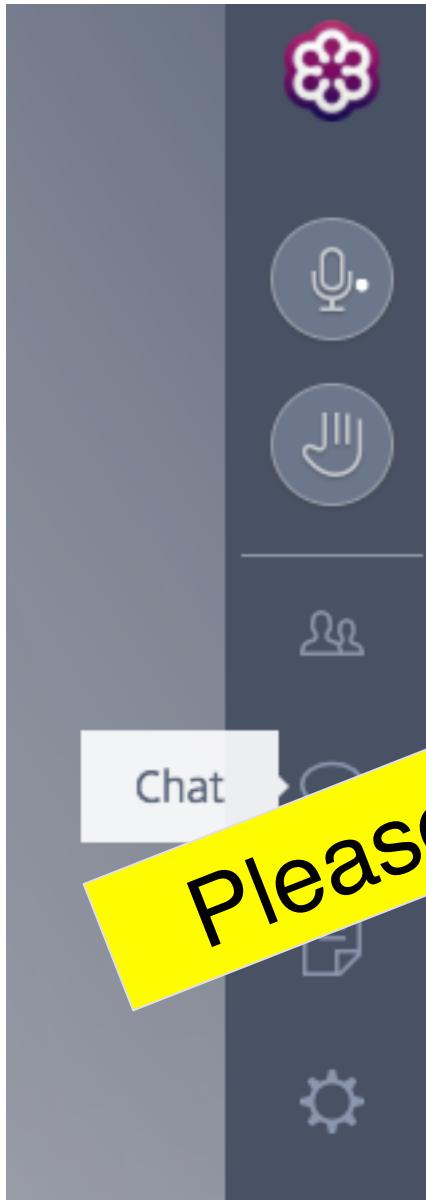
<http://imgur.com/a/JIDCP>

<http://imgur.com/a/LKDhp>



Open Targets

# Chat box



Address chat to 'Everyone'



Open Targets