10 YEAR ANNIVERSARY
Open Targets
2014-2024

Pistoria Alliance Conference 2024
24th April 2024

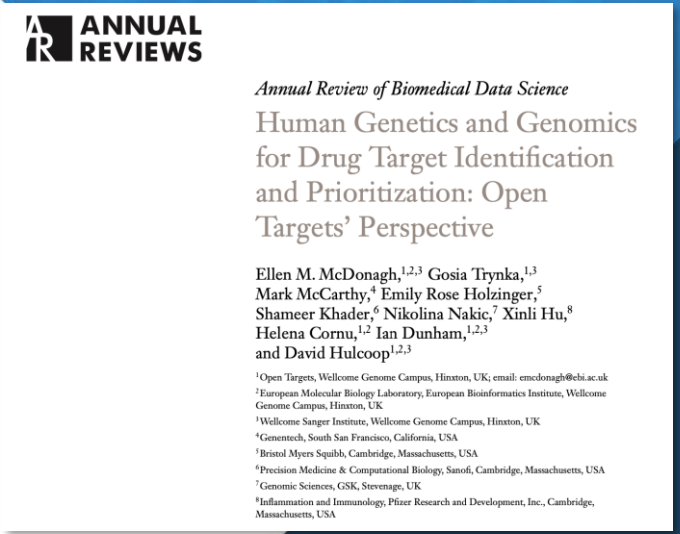
Using human genetics & genomics data for systematic drug target identification and prioritisation

Ellie McDonagh
Translational Informatics Director
emcdonagh@ebi.ac.uk

10 YEAR ANNIVERSARY
Open Targets
2014-2024

Pistoria Alliance Conference 2024
24th April 2024

Using human genetics & genomics data for systematic drug target identification and prioritisation



ANNUAL REVIEWS

Annual Review of Biomedical Data Science

Human Genetics and Genomics for Drug Target Identification and Prioritization: Open Targets' Perspective

Ellen M. McDonagh,[1,2,3] Gosia Trynka,[1,3] Mark McCarthy,[4] Emily Rose Holzinger,[5] Shameer Khader,[6] Nikolina Nakic,[7] Xinli Hu,[8] Helena Cornu,[1,2] Ian Dunham,[1,2,3] and David Hulcoop[1,2,3]

[1]Open Targets, Wellcome Genome Campus, Hinxton, UK; email: emcdonagh@ebi.ac.uk
[2]European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, UK
[3]Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, UK
[4]Genentech, South San Francisco, California, USA
[5]Bristol Myers Squibb, Cambridge, Massachusetts, USA
[6]Precision Medicine & Computational Biology, Sanofi, Cambridge, Massachusetts, USA
[7]Genomic Sciences, GSK, Stevenage, UK
[8]Inflammation and Immunology, Pfizer Research and Development, Inc., Cambridge, Massachusetts, USA

Ellie McDonagh
Translational Informatics Director
emcdonagh@ebi.ac.uk

10 YEAR ANNIVERSARY
Open Targets
2014-2024

Pistoria Alliance Conference 2024
24th April 2024

Harnessing AI to expedite R&D: our approach

Ellie McDonagh
Translational Informatics Director
emcdonagh@ebi.ac.uk

Open Targets

A partnership to transform drug discovery through the systematic identification and prioritisation of targets





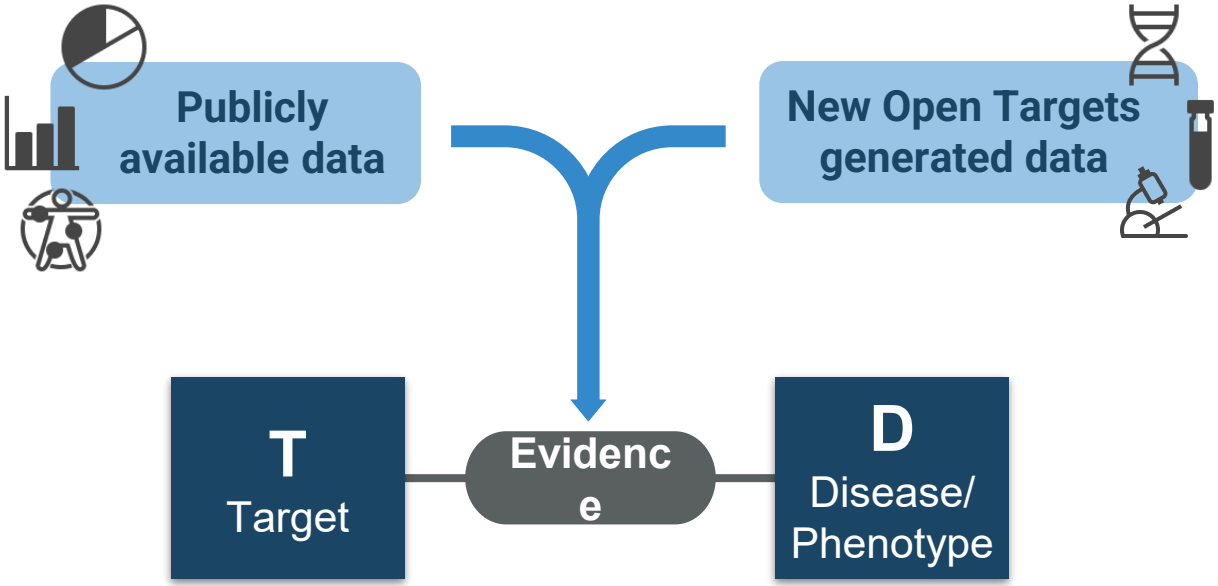
EMBL-EBI	Sanger	GSK	Genentech	?	Pfizer	Sanofi

Academic partners

Industry partners

Our Approach

We systematically use **evidence** to build therapeutic hypotheses between **targets** and **disease**

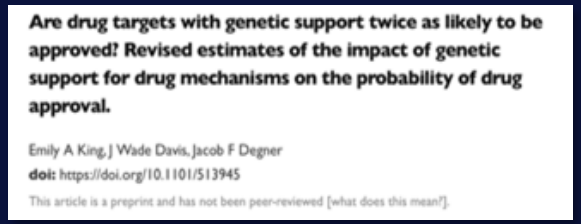


Open Targets

The importance of genetic evidence for drug success

The support of human genetic evidence for approved drug indications

Matthew R Nelson[1], Hannah Tipney[2], Jeffery L Painter[1], Judong Shen[1], Paola Nicoletti[3], Yufeng Shen[3,4], Aris Floratos[3,4], Pak Chung Sham[5,6], Mulin Jun Li[6,7], Junwen Wang[6,7], Lon R Cardon[8], John C Whittaker[2] & Philippe Sanseau[2]



Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval.

Emily A King, J Wade Davis, Jacob F Degner

doi: https://doi.org/10.1101/513945

- Drug **~8x** more likely to succeed if target identified in **Mendelian genetic evidence**
- Drug **> 2x** more likely to succeed if target is supported by **GWAS evidence**

- Nelson MR et al, (2015) Nat Genet. Aug;47(8):856-60.
- King EA et al, (2019). PLoS Genet. Dec 12;15(12):e1008489.

The importance of genetic evidence for drug success

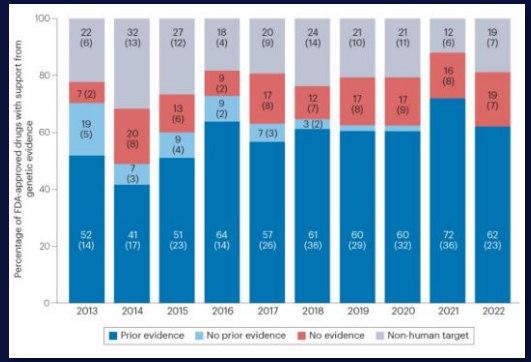
The support of human genetic evidence for approved drug indications

Matthew R Nelson[1], Hannah Tipney[2], Jeffery L Painter[1], Judong Shen[1], Paola Nicoletti[3], Yufeng Shen[3,4], Aris Floratos[3,4], Pak Chung Sham[5,6], Mulin Jun Li[6,7], Junwen Wang[6,7], Lon R Cardon[8], John C Whittaker[2] & Philippe Sanseau[2]

Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval.

Emily A King, J Wade Davis, Jacob F Degner

doi: https://doi.org/10.1101/513945



- Drug **~8x** more likely to succeed if target identified in **Mendelian genetic evidence**
- Drug **> 2x** more likely to succeed if target is supported by **GWAS evidence**

- Human genetics evidence supports **2/3 of the 2021 FDA-approved drugs**
- Genetic evidence supports **63% of new drugs approved in the past decade**

- Nelson MR et al, (2015) Nat Genet. Aug;47(8):856-60.
- King EA et al, (2019). PLoS Genet. Dec 12;15(12):e1008489.

- Ochoa D, et al, (2022) *Nat Rev Drug Discov.* Aug;21(8):551.
- Rusina, PV et al. (2023) *Nat Rev Drug Discov.* 2023 Oct.

The importance of genetic evidence for drug success

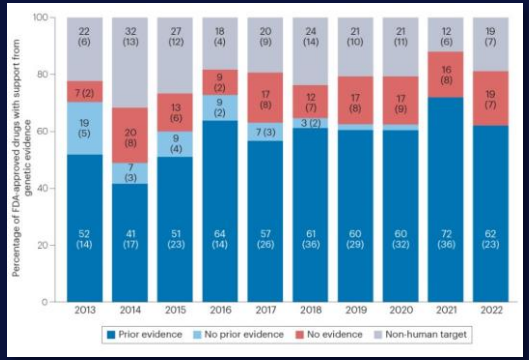
The support of human genetic evidence for approved drug indications

Matthew R Nelson[1], Hannah Tipney[2], Jeffery L Painter[1], Judong Shen[1], Paola Nicoletti[3], Yufeng Shen[3,4], Aris Floratos[3,4], Pak Chung Sham[5,6], Mulin Jun Li[6,7], Junwen Wang[6,7], Lon R Cardon[8], John C Whittaker[2] & Philippe Sanseau[2]

Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval.

Emily A King, J Wade Davis, Jacob F Degner

doi: https://doi.org/10.1101/513945



Analysis

Refining the impact of genetic evidence on clinical success

https://doi.org/10.1038/s41586-024-07316-0
Received: 5 July 2023
Accepted: 14 March 2024
Published online: 17 April 2024
Open access
Check for updates

Eric Vallabh Minikel[1], Jeffery L. Painter[2,5], Coco Chengliang Dong[2] & Matthew R. Nelson[3,4]✉

The cost of drug discovery and development is driven primarily by failure[1], with only about 10% of clinical programmes eventually receiving approval[2–4]. We previously estimated that human genetic evidence doubles the success rate from clinical development to approval[3]. In this study we leverage the growth in genetic evidence over the past decade to better understand the characteristics that distinguish clinical success and failure. We estimate the probability of success for drug mechanisms with genetic support is 2.6 times greater than those without. This relative success varies among therapy areas and development phases, and improves with increasing confidence in the causal gene, but is largely unaffected by genetic effect size, minor allele frequency or year of discovery. These results indicate we are far from reaching peak genetic insights to aid the discovery of targets for more effective drugs.

- Drug **~8x** more likely to succeed if target identified in **Mendelian genetic evidence**
- Drug **> 2x** more likely to succeed if target is supported by **GWAS evidence**

- Nelson MR et al, (2015) Nat Genet. Aug;47(8):856-60.
- King EA et al, (2019). PLoS Genet. Dec 12;15(12):e1008489.

- Human genetics evidence supports **2/3 of the 2021 FDA-approved drugs**
- Genetic evidence supports **63% of new drugs approved in the past decade**

- Ochoa D, et al, (2022) *Nat Rev Drug Discov.* Aug;21(8):551.
- Rusina, PV et al. (2023) *Nat Rev Drug Discov.* 2023 Oct.

- Drug mechanisms with genetic support have 2.6 greater probability of success
- *"These results indicate we are far from reaching peak genetic insights to aid the discovery of targets for more effective drugs."*

- Minikel, E.V. et al (2024) *Nature* https://doi.org/10.1038/s41586-024-07316-0

The challenge of using GWAS for target discovery?

How do you go from a **variant** associated to disease to the **causal gene**?

- Most disease associated variants are **outside coding regions**
- The **causal gene** is not necessarily the gene closest to the associated variant
- Sometimes variants can affect **several genes**
- The lead variant reported may not be the ***causal* variant**

Open Targets Genetics was established to systematically address this challenge

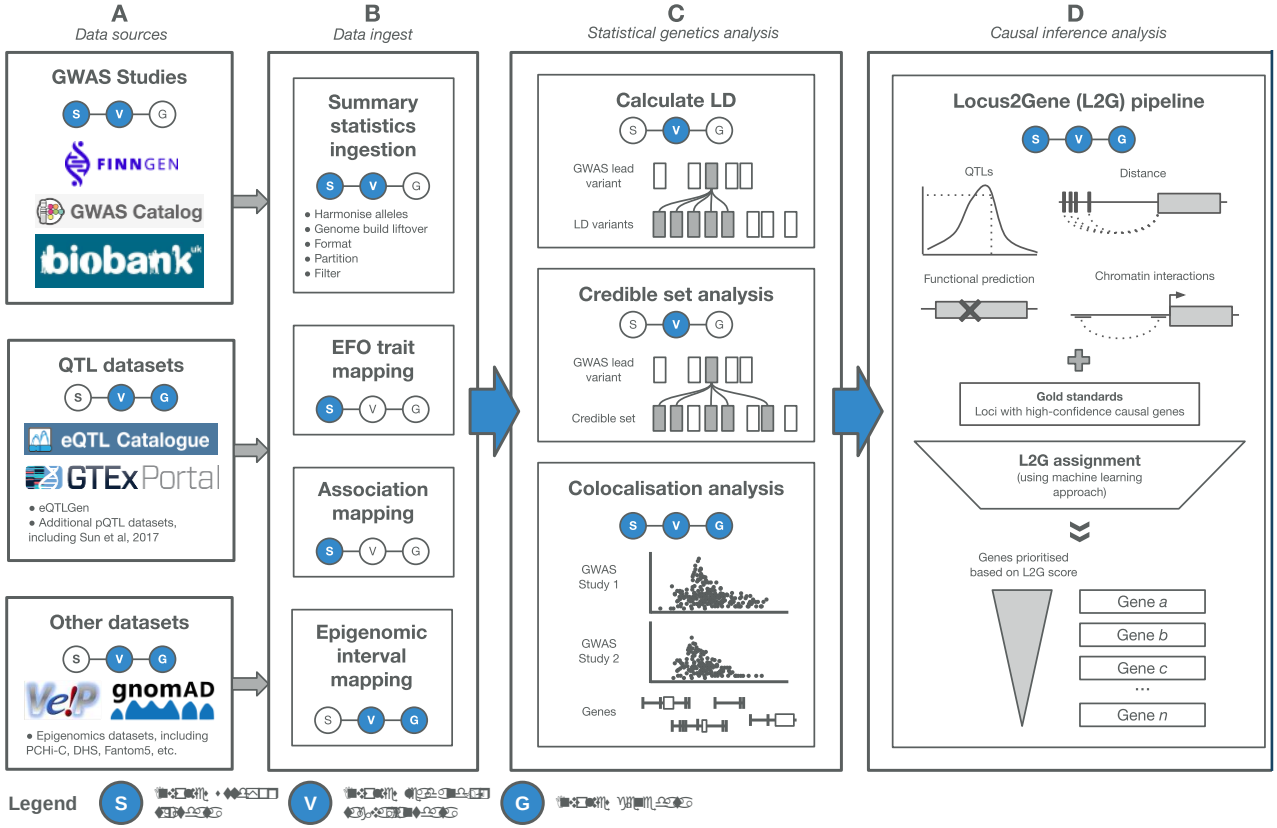
Open Targets

How do we **predict** causal genes (targets) for disease X from GWAS loci?

Locus2Gene Machine Learning pipeline

Goal: provide a predictive score for the most likely causative genes underlying each GWAS association

Nat Genet **53,** 1527–1533 (2021).
PMID: 34711957



Open Targets

How do we **predict** causal genes (targets) for disease X from GWAS loci?

Locus2Gene Machine Learning pipeline

Goal: provide a predictive score for the most likely causative genes underlying each GWAS association

Nat Genet **53,** 1527–1533 (2021).
PMID: 34711957

Open Targets

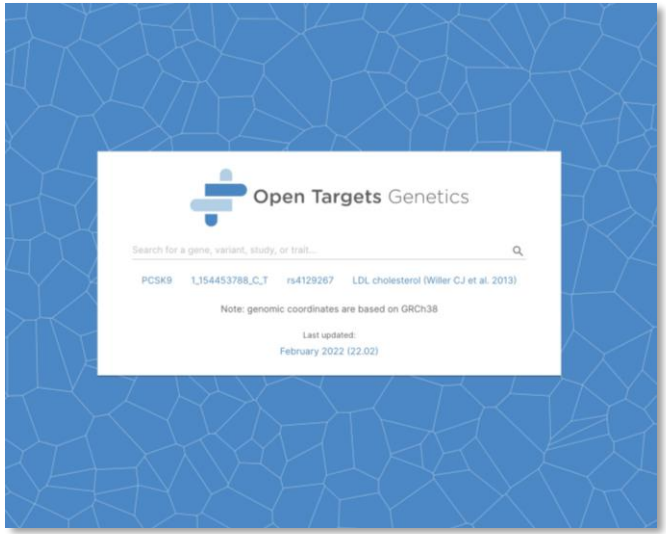
How do we **predict** causal genes (targets) for disease X from GWAS loci?

Locus2Gene Machine Learning pipeline

Goal: provide a predictive score for the most likely causative genes underlying each GWAS association

Nat Genet **53,** 1527–1533 (2021).
PMID: 34711957

https://genetics.opentargets.org/



Open Targets

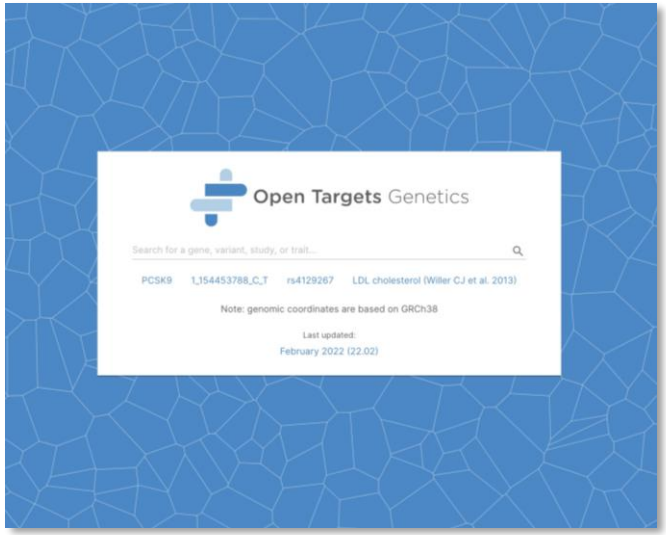
How do we **predict** causal genes (targets) for disease X from GWAS loci?

Locus2Gene Machine Learning pipeline

Goal: provide a predictive score for the most likely causative genes underlying each GWAS association

Nat Genet **53,** 1527–1533 (2021).
PMID: 34711957

https://genetics.opentargets.org/



Optimising our pipelines & data

Open Targets Gentropy

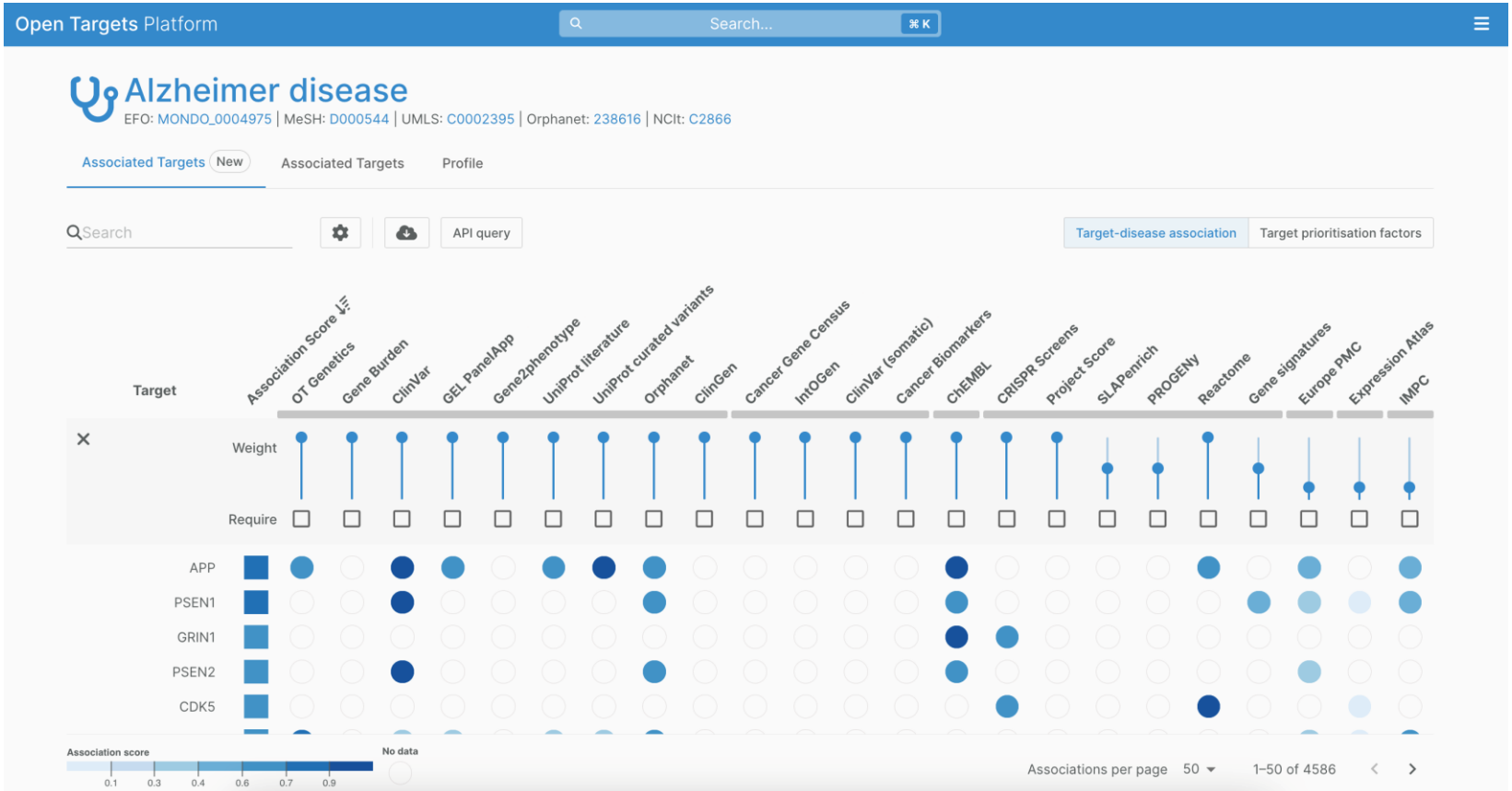
opentargets.github.io/gentropy/

Python package for post-GWAS analysis

Open Targets

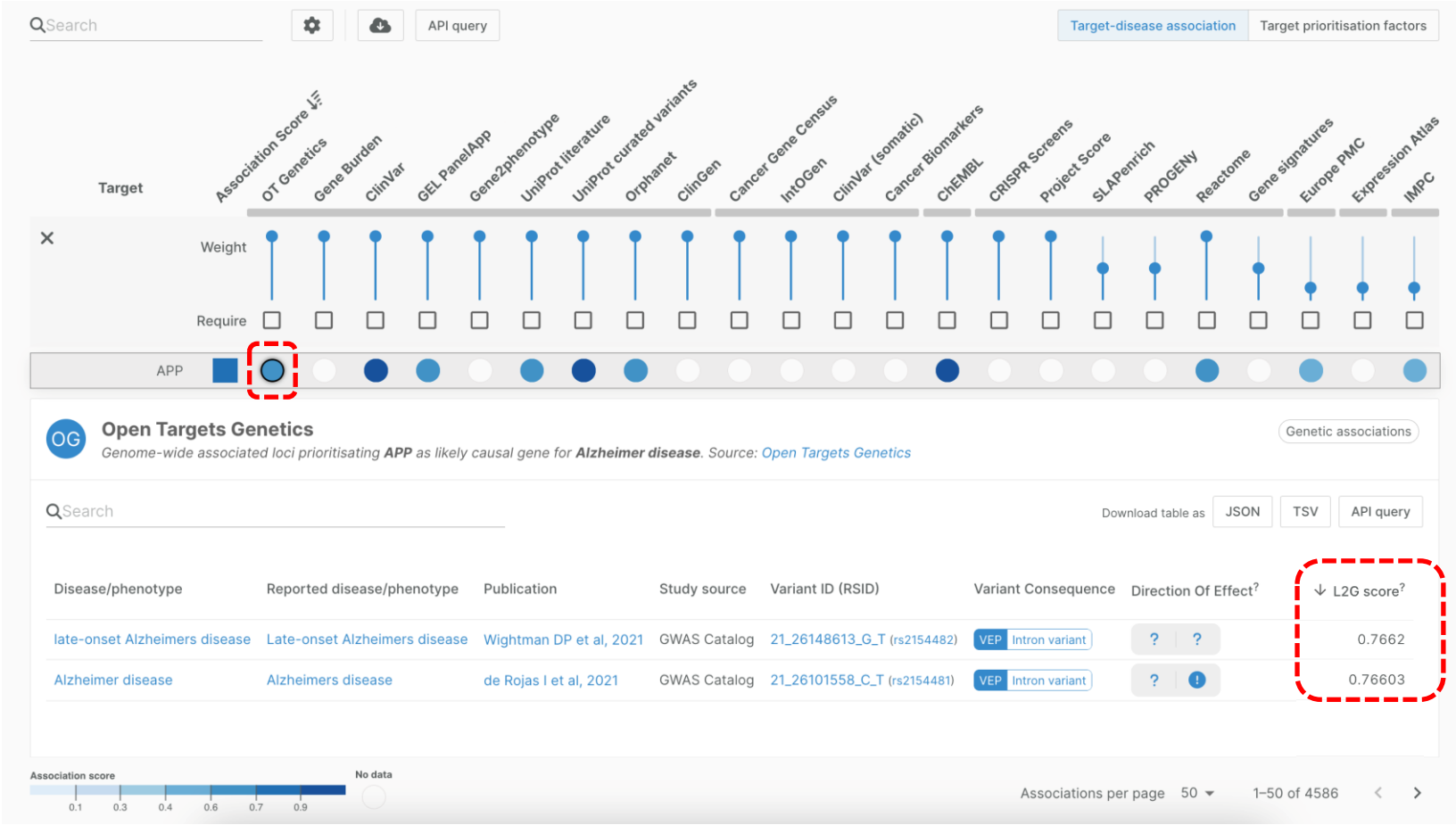
How do we prioritise the best targets for disease X?

https://platform.opentargets.org



Open Targets

How do we prioritise the best targets for disease X?



Open Targets

Named Entity Recognition & Deep Learning

Developed with Europe PMC to extract **target – disease – drug** relationships from the scientific literature

Europe PMC

*Text-mining method evaluating the strength of the association between **APP** and **Alzheimer disease** when they co-occur in the literature. Source: Europe PMC*

Text mining

Download table as JSON TSV API query

Disease/phenotype	Publication	Year	Score
	Gene mutations associated with early onset familial Alzheimer's disease in China: An overview and current status. Qin Q, Yin Y, Wang Y, Lu Y, Tang Y, Jia J Molecular genetics & genomic medicine **2020** 8(10):e1443	2020	54

Hide summary | Show abstract | Show 12 match details

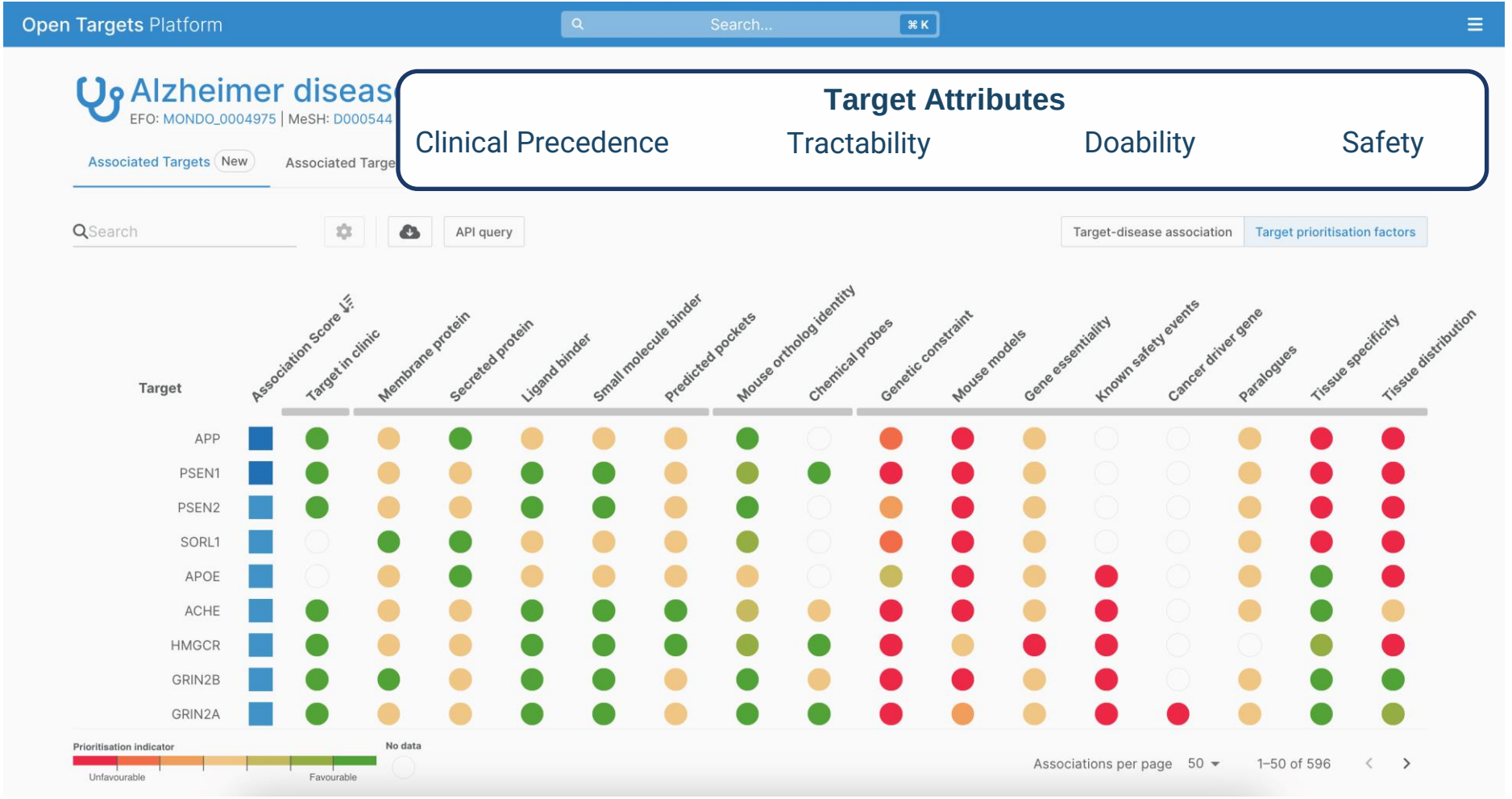
Evidence summary

The relationship between APP and Alzheimer's disease (AD) is significant, as mutations in the APP gene can lead to early-onset familial AD (EOFAD). Abnormal cleavage of APP by secretases can result in the production of amyloid beta (Aβ) peptides, which are a key component of amyloid plaques in the AD brain. Mutations in the APP gene have been identified in Chinese patients with EOFAD, leading to increased Aβ production, altered Aβ42/Aβ40 ratio, and neurotoxicity, contributing to the development of AD.

Powered by OpenAI

Open AI: Exploration of LLM for conversational summaries of evidence

Open Targets



Open Targets

Known Drugs

*Clinical precedence for drugs with investigational or approved indications targeting **APP** according to their curated mechanism of action. Source: ChEMBL.*

Search

Download table as JSON TSV API query

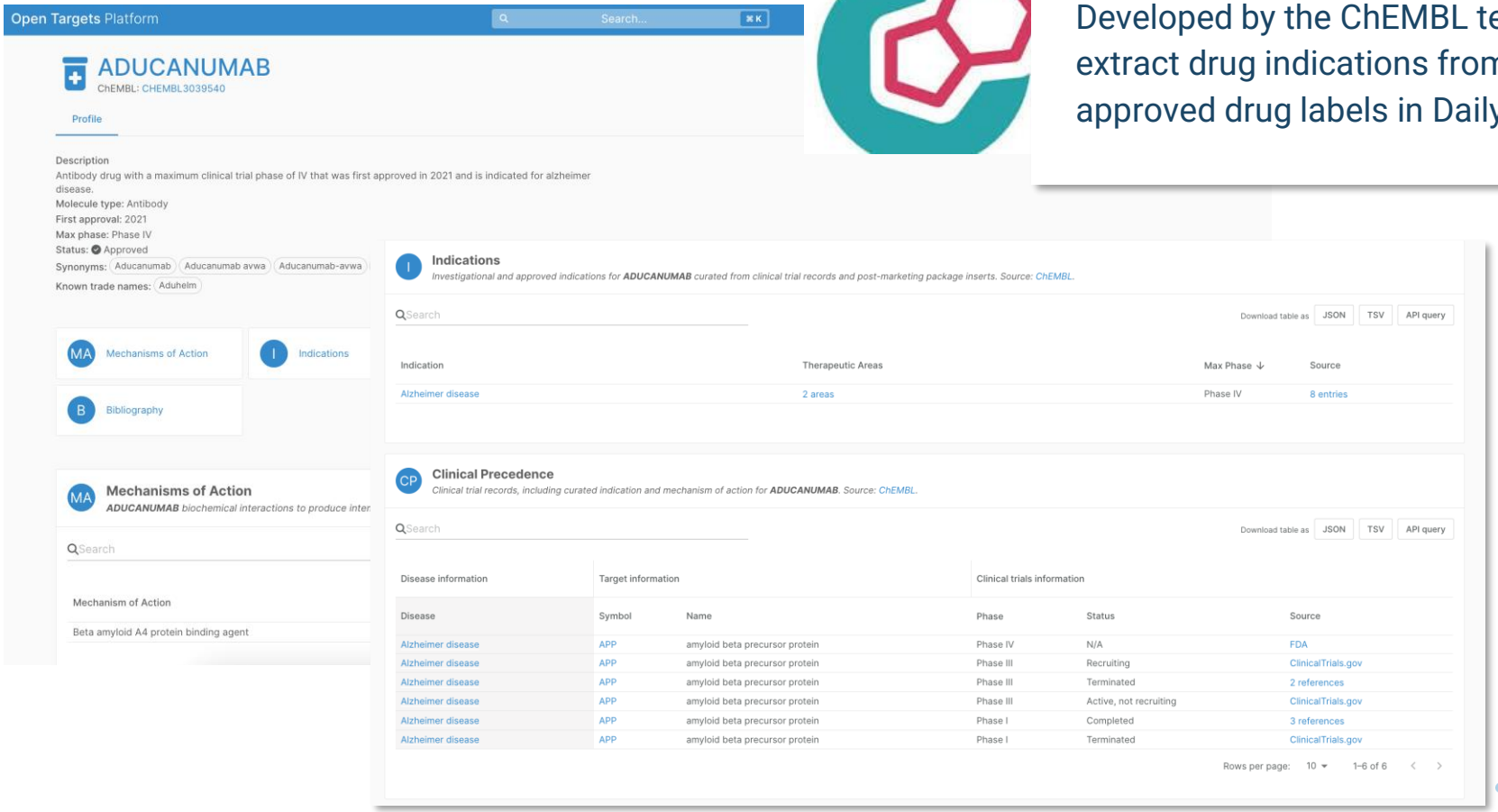
Drug information				Disease information	Clinical trials information		
Drug	Type	Mechanism Of Action	Action Type	Disease	Phase	Status	Source
ADUCANUMAB	Antibody	Beta amyloid A4 protein binding agent	Binding agent	Alzheimer disease	Phase IV	N/A	FDA
TRAMIPROSATE	Small molecule	Beta amyloid A4 protein stabiliser	Binding agent	Cognitive impairment	Phase IV	Recruiting	ClinicalTrials.gov
TRAMIPROSATE	Small molecule	Beta amyloid A4 protein stabiliser	Binding agent	Alzheimer disease	Phase III	Unknown status	ClinicalTrials.gov
GANTENERUMAB	Antibody	Beta amyloid A4 protein binding agent	Binding agent	Alzheimer disease	Phase III	Terminated	5 references
DONANEMAB	Antibody	Beta amyloid A4 protein disrupting agent	Binding agent	Alzheimer disease	Phase III	Recruiting	ClinicalTrials.gov
CRENEZUMAB	Antibody	Beta amyloid A4 protein inhibitor	Binding agent	Alzheimer disease	Phase III	Terminated	3 references
ADUCANUMAB	Antibody	Beta amyloid A4 protein binding agent	Binding agent	Alzheimer disease	Phase III	Recruiting	ClinicalTrials.gov
LECANEMAB	Antibody	Beta amyloid A4 protein inhibitor	Binding agent	Alzheimer disease	Phase III	Active, not recruiting	ClinicalTrials.gov
SOLANEZUMAB	Antibody	Beta amyloid A4 protein binding agent	Binding agent	Cognitive impairment	Phase III	Active, not recruiting	ClinicalTrials.gov
BAPINEUZUMAB	Antibody	Beta amyloid A4 protein binding agent	Binding agent	Alzheimer disease	Phase III	Terminated	5 references

Prioritisation indicator
Unfavourable — Favourable
No data

Associations per page 50 1–50 of 596

Open Targets

Drug Label Extraction



ML model to data mine drug labels

Developed by the ChEMBL team to extract drug indications from FDA-approved drug labels in DailyMed

Open Targets

Why were Clinical trials terminated?

Phase	Status	Start Date	Source
Phase IV	Terminated?	2018	ClinicalTrials.gov
Phase IV	Terminated?	2009	ClinicalTrials.gov
Phase IV	Terminated?	2008	ClinicalTrials.gov
Phase IV	Terminated?	2009	ClinicalTrials.gov
Phase IV	Terminated?	2004	ClinicalTrials.gov
Phase IV	Terminated?	2004	ClinicalTrials.gov
Phase III	Terminated?	2015	ClinicalTrials.gov
Phase III	Terminated?	2012	ClinicalTrials.gov
Phase III	Terminated?	2006	ClinicalTrials.gov
Phase III	Terminated?	2014	ClinicalTrials.gov

Study stop reason: Rofecoxib was withdrawn from the market due to safety concerns.

Safety or side effects | Insufficient enrollment

Rows per page: 10 11-20 of 38

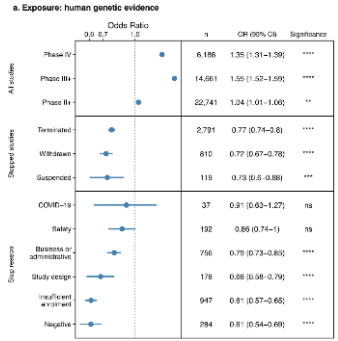
Example: SCN9A and Pain
https://platform.opentargets.org/evidence/ENSG00000169432/EFO_0003843

Natural Language Processing (NLP) to classify why clinical trials were terminated

Nature Genetics - Accepted

Why Clinical Trials Stop: The role of genetics

Olesya Razuvayevskaya, Irene Lopez, Ian Dunham and David Ochoa



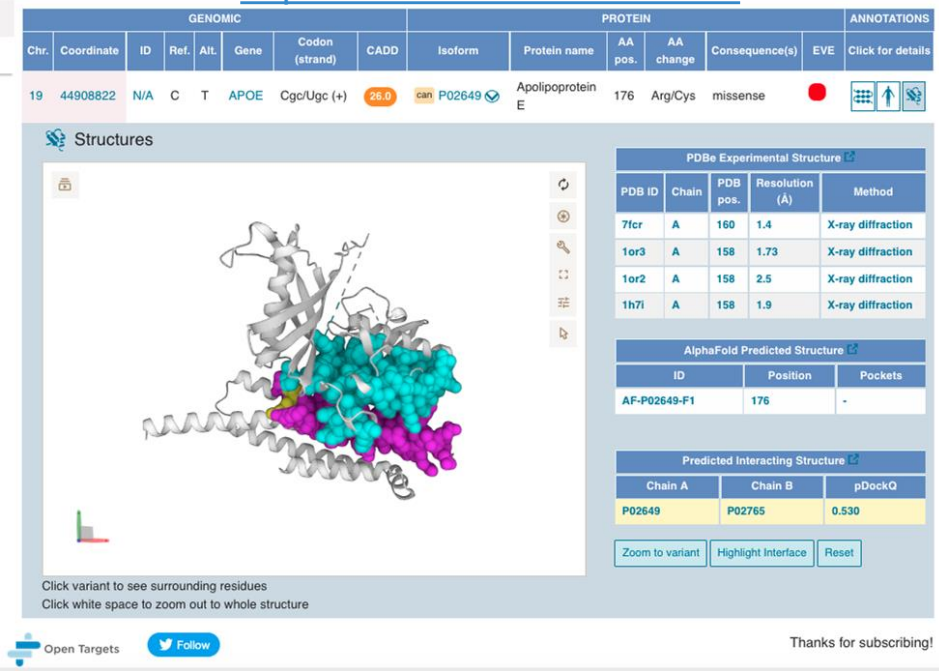
Open Targets

Alzheimer disease

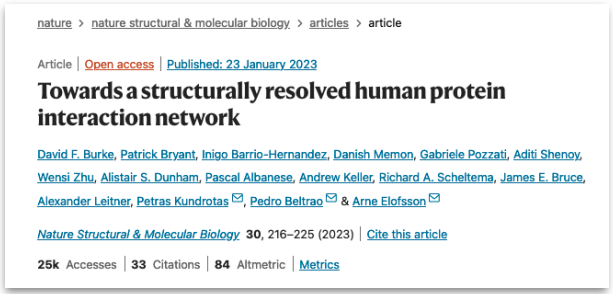
EFO: MONDO_0004975 | MeSH: D000544 | UMLS: C0002395 | Orphanet: 238616 | NCIt: C2866

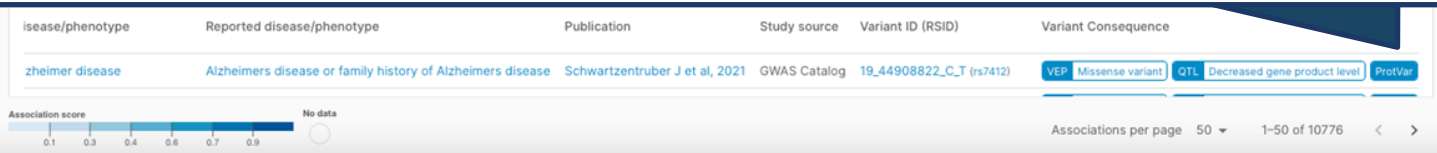
Associated Ta...

https://www.ebi.ac.uk/ProtVar/



Launch of **ProtVar resource** contextualizing human missense variation integrated predictions of the consequence of missense variation on stability using **FoldX**, protein-protein interactions based on **AlphaFold** structures and predicted binding pockets

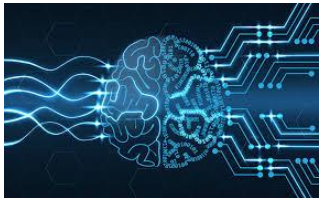


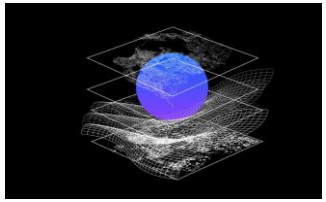


Some Open Targets AI areas of research interest

Recruiting!

Knowledge Extraction, Representation, & usage
Project Leads: Barbara Zdrazil (EMBL-EBI), Sebastian Lobentanzer (UKHD)





Recruiting!

Developing foundation models to analyse single-cell perturbation data
Project Lead: Mo Lotfollahi (Sanger)

Recruiting!

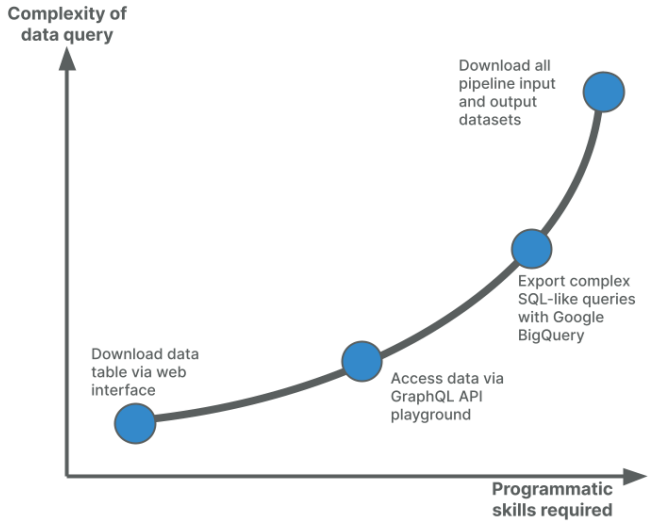
Using AI for predicting combination targets in cancer
Project Leads: Evangelia Petsalaki (EMBL-EBI) & Mathew Garnett (Sanger)

Digital Twins for disease modelling
Project Leads: Sheriff Rahuman (EMBL-EBI) & Ellie McDonagh (EMBL-EBI/OT)

Exciting opportunities!: https://www.opentargets.org/jobs

How we enable AI

- Provide integrated, standardised, harmonised, open source data
- Cloud compatible (Google, AWS)



A few external examples:

- Han, Y. *et al.* Empowering the discovery of novel target-disease associations via **machine learning approaches** in the open targets platform. *BMC Bioinformatics* **23**, 232 (2022).
- Gogleva, A. *et al.* **Knowledge graph**-based recommendation framework identifies drivers of resistance in EGFR mutant non-small cell lung cancer. *Nat. Commun.* **13**, 1667 (2022).
- Ye, C., *et al.* **Knowledge Graph**-Enhanced Tensor Factorisation Model for Discovering Drug Targets. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **PP**, (2022).
- Raies, A. *et al.* DrugnomeAI is an ensemble **machine-learning framework for predicting druggability** of candidate drug targets. *Commun Biol* **5**, 1291 (2022).

Pistoia Alliance

- KG & LLM project (led by Vladimir Makarov)

Open Targets

How we harness AI

Generating new knowledge

- Addressing particular scientific challenges (e.g. predicting causal genes underlying GWAS associations)
- Applying to scientific research questions

Extracting knowledge

- Associations from the Scientific literature
- Clinical precedence
- Learning from clinical trial failures
- Enhancing automation and reducing manual curation time

Enabling others to harness AI/ML

- Provide integrated, standardised, harmonised, open source data
- Open Source software
- Cloud compatible

Enabling Therapeutic Hypothesis Generation For More Effective Disease Treatments

Acknowledgements

EMBL-EBI

wellcome sanger institute

sanofi

Pfizer

GSK

Genentech
A Member of the Roche Group

emcdonagh@ebi.ac.uk

Open Targets