

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

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

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

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



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,³ 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}

¹Open Targets, Wellcome Genome Campus, Hinxton, UK; email: emcdonagh@ebi.ac.uk

²European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, UK

³Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, UK

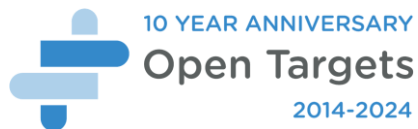
⁴Genentech, South San Francisco, California, USA

⁵Bristol Myers Squibb, Cambridge, Massachusetts, USA

⁶Precision Medicine & Computational Biology, Sanofi, Cambridge, Massachusetts, USA

⁷Genomic Sciences, GSK, Stevenage, UK

⁸Inflammation and Immunology, Pfizer Research and Development, Inc., Cambridge, Massachusetts, USA



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



Genentech
A Member of the Roche Group



sanofi

EMBL-EBI

Sanger

GSK

Genentech

?

Pfizer

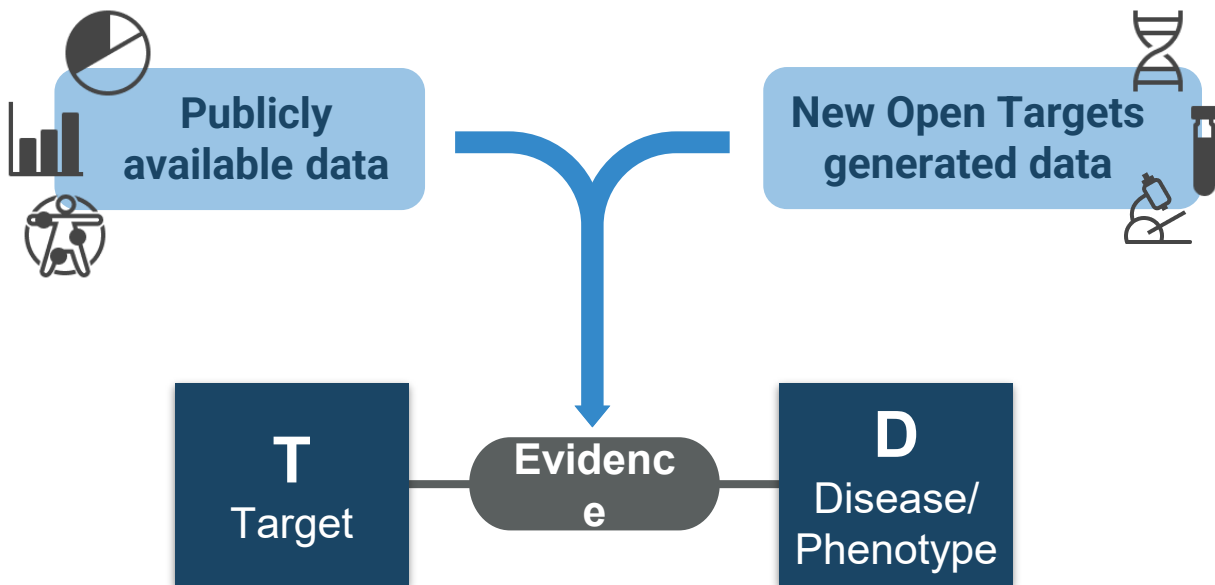
Sanofi

Academic partners

Industry partners

Our Approach

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



The importance of genetic evidence for drug success

The support of human genetic evidence for approved drug indications

Matthew R Nelson¹, Hannah Tipney², Jeffery L Painter¹, Judong Shen¹, Paola Nicoletti³, Yufeng Shen^{3,4}, Aris Floratos^{3,4}, Pak Chung Sham^{5,6}, Malin Jun Li^{6,7}, Junwen Wang^{6,7}, Lon R Cardon⁸, John C Whittaker² & Philippe Sansseau²

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>

This article is a preprint and has not been peer-reviewed [what does this mean?]

- 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

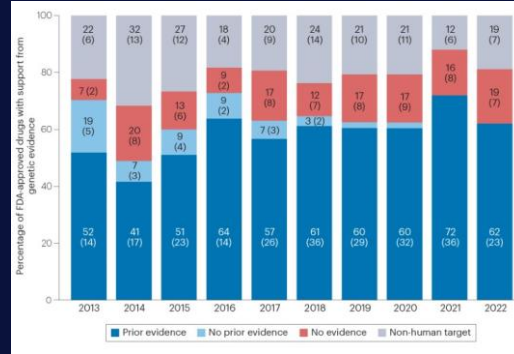
Matthew R Nelson¹, Hannah Tipney², Jeffery L Painter¹, Judong Shen¹, Paola Nicoletti³, Yufeng Shen^{3,4}, Aris Floratos^{3,4}, Pak Chung Sham^{5,6}, Malin Jun Li^{6,7}, Junwen Wang^{6,7}, Lon R Cardon⁸, John C Whittaker² & Philippe Sansseau²

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>

This article is a preprint and has not been peer-reviewed [what does this mean?]



- 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

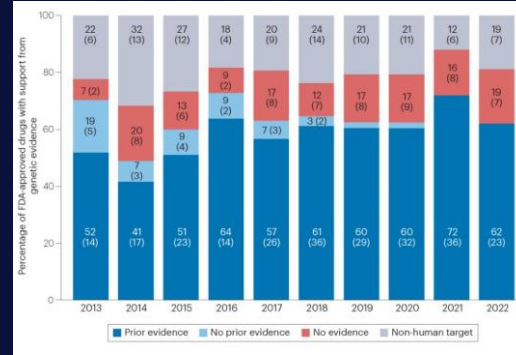
The support of human genetic evidence for approved drug indications

Matthew R Nelson¹, Hannah Tipney², Jeffery L Painter¹, Judong Shen¹, Paola Nicoletti³, Yufeng Shen^{3,4}, Aris Floratos^{3,4}, Pak Chung Sham^{5,6}, Malin Jun Li^{6,7}, Junwen Wang^{6,7}, Lon R Cardon⁸, John C Whittaker² & Philippe Sanssouci²

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>

This article is a preprint and has not been peer-reviewed [what does this mean?]



Analysis

Refining the impact of genetic evidence on clinical success

<https://doi.org/10.1038/s41586-024-07316-0> Eric Vallabh Minikel¹, Jeffery L. Painter^{2,3}, Coco Chengliang Dong² & Matthew R. Nelson^{4,5}

Received: 5 July 2023

Accepted: 14 March 2024

Published online: 17 April 2024

Open access

Check for updates

The cost of drug discovery and development is driven primarily by failure¹, 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 approval⁵. 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**

- Human genetics evidence supports **2/3 of the 2021 FDA-approved drugs**

- Genetic evidence supports **63% of new drugs approved in the past decade**

- 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.”*

- 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.

- 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**?

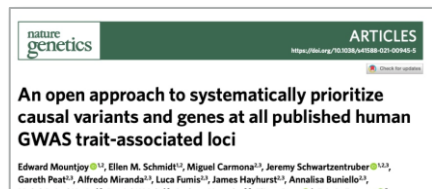
- Most disease associated variants are **outside coding regions**
- The **causal gene** is not necessarily the gene closest to the associated variant
- Sometimes variants can affect **several genes**
- The lead variant reported may not be the **causal variant**

Open Targets
Genetics was
established to
systematically
address this
challenge

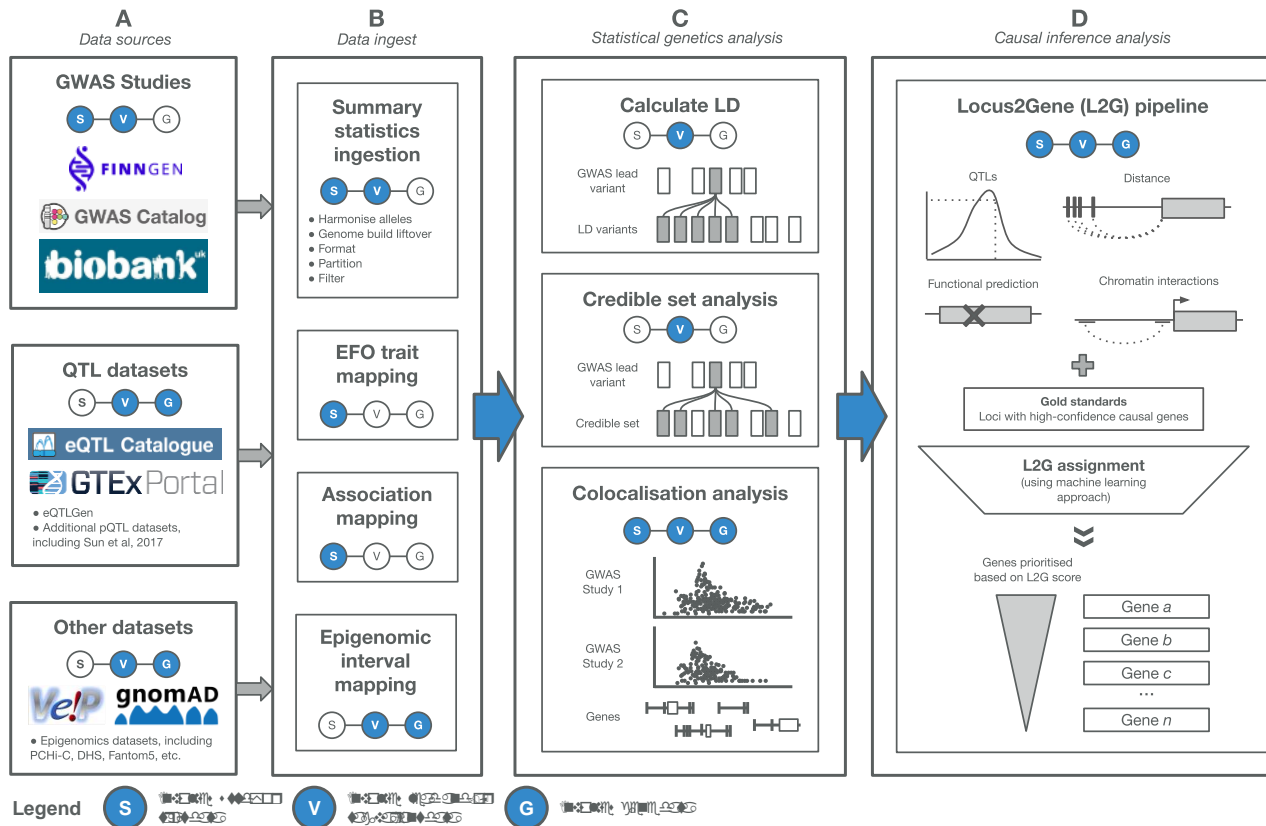
How do we predict causal genes (targets) for disease X from GWAS loci?

Locus2Gene Machine Learning pipeline

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



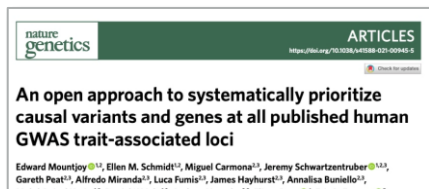
Nat Genet 53, 1527–1533 (2021).
PMID: 34711957



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).

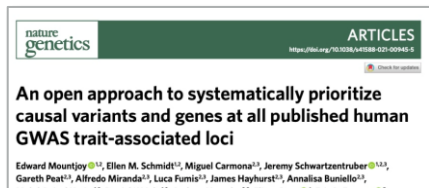
PMID: 34711957

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

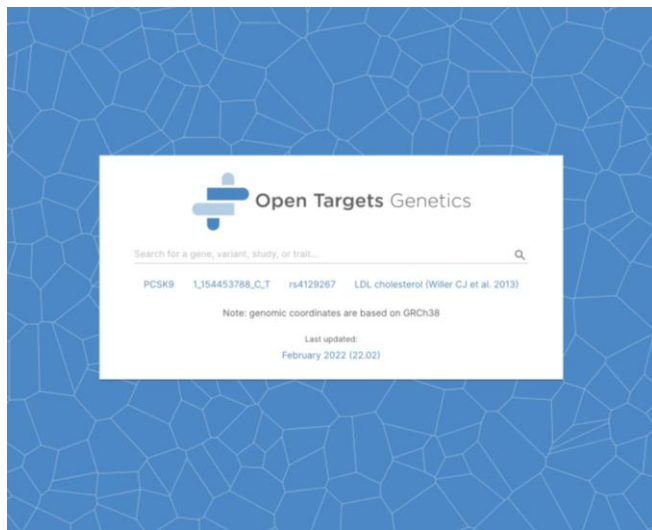
Locus2Gene Machine Learning pipeline

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

<https://genetics.opentargets.org/>



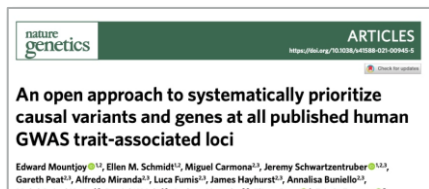
Nat Genet **53**, 1527–1533 (2021).
PMID: 34711957



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

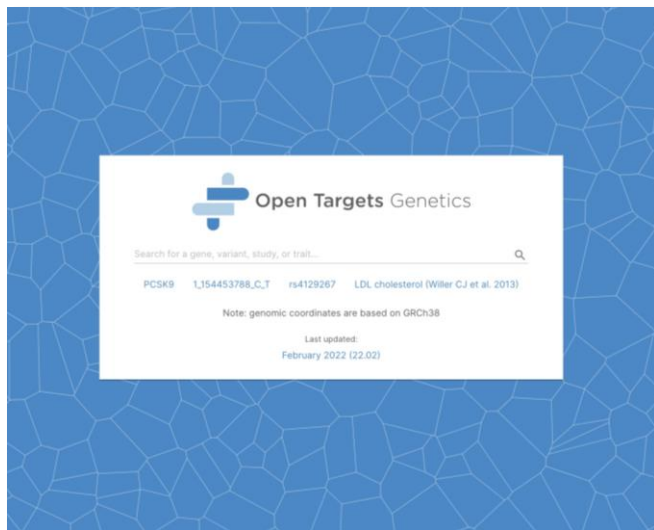
Locus2Gene Machine Learning pipeline

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



opentargets.github.io/gentropy/

Python package for post-GWAS analysis

How do we prioritise the best targets for disease X?

<https://platform.opentargets.org>

Open Targets Platform

Search... ⌘ K

Alzheimer disease

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

Associated Targets New Associated Targets Profile

Q Search ⚙️ 📄 API query Target-disease association Target prioritisation factors

Target	Association Score	OT Genetics	Gene Burden	ClinVar	GEL PanelApp	Gene2phenotype	UniProt literature	UniProt curated variants	Orphanet	ClinGen	Cancer Gene Census	IntOGen	ClinVar (somatic)	Cancer Biomarkers	ChEMBL	CRISPR Screens	Project Score	SLAPenrich	PROGENY	Reactome	Gene signatures	Europe PMC	Expression Atlas	IMPC
X	Weight	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
	Require	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
APP	■	●	○	●	●	○	●	●	●	○	○	○	○	○	●	○	○	○	○	●	○	●	○	●
PSEN1	■	○	○	●	○	○	○	○	●	○	○	○	○	○	●	○	○	○	○	○	●	●	○	●
GRIN1	■	○	○	○	○	○	○	○	○	○	○	○	○	○	●	●	○	○	○	○	○	○	○	○
PSEN2	■	○	○	●	○	○	○	○	●	○	○	○	○	○	●	○	○	○	○	○	○	●	○	○
CDK5	■	○	○	○	○	○	○	○	○	○	○	○	○	○	●	○	○	○	○	●	○	○	○	○

Association score: 0.1 0.3 0.4 0.6 0.7 0.9 No data

Associations per page: 50 1-50 of 4586

Open Targets



Named Entity Recognition & Deep Learning

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

Target

Association Score ↓

OT Genetics

Gene Burden

ClinVar

GEL Pane.

Gene2phen

Unip

APP

PR Screens

Project Score

SLAPenrich

PROGENy

Reactome

Gene signatures

Europe PMC

Expression Atlas

IMPC



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

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

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 (A β) 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 A β production, altered A β 42/A β 40 ratio, and neurotoxicity, contributing to the development of AD.

Powered by OpenAI

2020 54



Open AI:
Exploration of LLM for conversational summaries of evidence

Associated Targets New

Associated Target

Clinical Precedence

Target Attributes

Tractability

Doability

Safety

QSearch



API query

Target-disease association

Target prioritisation factors

Target	Association Score ↓	Target in clinic	Membrane protein	Secreted protein	Ligand binder	Small molecule binder	Predicted pockets	Mouse ortholog identity	Chemical probes	Genetic constraint	Mouse models	Gene essentiality	Known safety events	Cancer driver gene	Paralogues	Tissue specificity	Tissue distribution
APP	Blue	Green	Orange	Green	Orange	Orange	Orange	Green	White	Red	Red	Orange	White	White	Orange	Red	Red
PSEN1	Blue	Green	Orange	Orange	Green	Green	Orange	Green	Green	Red	Red	Orange	White	White	Orange	Red	Red
PSEN2	Blue	Green	Orange	Orange	Green	Green	Orange	Green	White	Orange	Red	Orange	White	White	Orange	Red	Red
SORL1	Blue	White	Green	Green	Orange	Orange	Orange	Green	White	Red	Red	Orange	White	White	Orange	Red	Red
APOE	Blue	White	Orange	Green	Orange	Orange	Orange	Orange	White	Green	Red	Orange	Red	White	Orange	Green	Red
ACHE	Blue	Green	Orange	Orange	Green	Green	Green	Orange	Orange	Red	Red	Orange	Red	White	Orange	Green	Orange
HMGCR	Blue	Green	Orange	Orange	Green	Green	Green	Green	Green	Red	Orange	Red	Red	White	White	Green	Red
GRIN2B	Blue	Green	Green	Orange	Green	Green	Orange	Green	Orange	Red	Red	Orange	Red	White	Orange	Green	Green
GRIN2A	Blue	Green	Orange	Orange	Green	Green	Orange	Green	Green	Red	Orange	Orange	Red	Red	Orange	Green	Green



Target

Association Score

Target in clinic

Membrane protein

Secreted protein

Ligand binder

Small molecule binder

Predicted pockets

Mouse ortholog identity

Chemical probes

Genetic constraint

Mouse models

Gene essentiality

Known safety events

Cancer driver gene

Paralogues

Tissue specificity

Tissue distribution

APP



Known Drugs

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

Q Search

Download table as

JSON

TSV

API query

Drug information

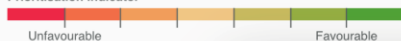
Disease information

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

No data



Associations per page 50

1-50 of 596



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 Platform

Search...

ADUCANUMAB
ChEMBL: CHEMBL3039540

Profile

Description

Antibody drug with a maximum clinical trial phase of IV that was first approved in 2021 and is indicated for alzheimer disease.

Molecule type: Antibody

First approval: 2021

Max phase: Phase IV

Status: Approved

Synonyms:

Known trade names:

MA Mechanisms of Action

I Indications

B Bibliography

MA Mechanisms of Action

ADUCANUMAB biochemical interactions to produce inter

Search

Mechanism of Action

Beta amyloid A4 protein binding agent

I Indications

Investigational and approved indications for **ADUCANUMAB** curated from clinical trial records and post-marketing package inserts. Source: ChEMBL.

Search

Download table as

Indication	Therapeutic Areas	Max Phase ↓	Source
Alzheimer disease	2 areas	Phase IV	8 entries

CP Clinical Precedence

Clinical trial records, including curated indication and mechanism of action for **ADUCANUMAB**. Source: ChEMBL.

Search

Download table as

Disease information	Target information		Clinical trials information		
Disease	Symbol	Name	Phase	Status	Source
Alzheimer disease	APP	amyloid beta precursor protein	Phase IV	N/A	FDA
Alzheimer disease	APP	amyloid beta precursor protein	Phase III	Recruiting	ClinicalTrials.gov
Alzheimer disease	APP	amyloid beta precursor protein	Phase III	Terminated	2 references
Alzheimer disease	APP	amyloid beta precursor protein	Phase III	Active, not recruiting	ClinicalTrials.gov
Alzheimer disease	APP	amyloid beta precursor protein	Phase I	Completed	3 references
Alzheimer disease	APP	amyloid beta precursor protein	Phase I	Terminated	ClinicalTrials.gov

Rows per page: 1-6 of 6 < >

Why were Clinical trials terminated?

Phase ↓	Status	Start Date	Source
Phase IV	Terminated?	2018	ClinicalTrials.gov
Phase IV	Terminated?	2009	ClinicalTrials.gov
Phase IV	Study stop reason: Rofecoxib was withdrawn from the market due to safety concerns.	2008	ClinicalTrials.gov
Phase IV	Safety or side effects	2009	ClinicalTrials.gov
Phase IV	Insufficient enrollment	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

Rows per page: 10 ▼ 11-20 of 38 |< >

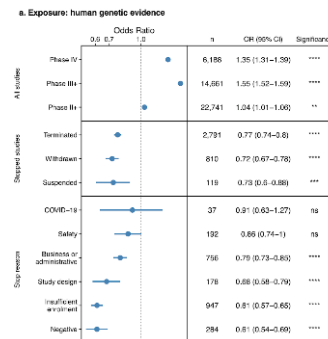
Example: SCN9A and Pain

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



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

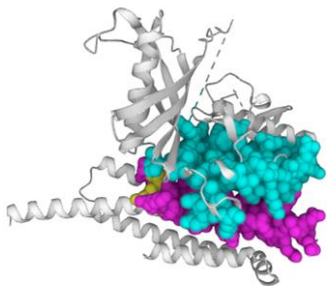
GENOMIC							PROTEIN					ANNOTATIONS		
Chr.	Coordinate	ID	Ref.	Alt.	Gene	Codon (strand)	CADD	isoform	Protein name	AA pos.	AA change	Consequence(s)	EVE	Click for details
19	44908822	N/A	C	T	APOE	Cgc/Ugc (+)	26.0	can P02649	Apolipoprotein E	176	Arg/Cys	missense	●	  



tion factors

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

Structures



PDB Experimental Structure				
PDB ID	Chain	PDB pos.	Resolution (Å)	Method
7fcr	A	160	1.4	X-ray diffraction
1or3	A	158	1.73	X-ray diffraction
1or2	A	158	2.5	X-ray diffraction
1h7i	A	158	1.9	X-ray diffraction

AlphaFold Predicted Structure		
ID	Position	Pockets
AF-P02649-F1	176	-

Predicted Interacting Structure		
Chain A	Chain B	pDockQ
P02649	P02765	0.530

[Zoom to variant](#) [Highlight interface](#) [Reset](#)

Click variant to see surrounding residues
Click white space to zoom out to whole structure

 Open Targets  Follow

Thanks for subscribing!

[nature](#) > [nature structural & molecular biology](#) > [articles](#) > [article](#)

Article | [Open access](#) | [Published: 23 January 2023](#)


Towards a structurally resolved human protein interaction network

[David F. Burke](#), [Patrick Bryant](#), [Inigo Barrio-Hernandez](#), [Danish Memon](#), [Gabriele Pozzati](#), [Aditi Shengy](#), [Wensi Zhu](#), [Alistair S. Dunham](#), [Pascal Albanese](#), [Andrew Keller](#), [Richard A. Scheltema](#), [James F. Bruce](#), [Alexander Leitner](#), [Petras Kundrotas](#), [Pedro Beltrao](#) & [Arne Elofsson](#)

Nature Structural & Molecular Biology **30**, 216–225 (2023) | [Cite this article](#)

25k Accesses | 33 Citations | 84 Altmetric | [Metrics](#)

disease/phenotype	Reported disease/phenotype	Publication	Study source	Variant ID (RSID)	Variant Consequence
Alzheimer disease	Alzheimers disease or family history of Alzheimers disease	Schwartzentruber J et al, 2021	GWAS Catalog	19_44908822_C_T (rs7412)	VEP: Missense variant QTL: Decreased gene product level ProtVar

Association score:  No data

Associations per page 50 | 1–50 of 10776

Recruiting!

Knowledge Extraction, Representation, & usage

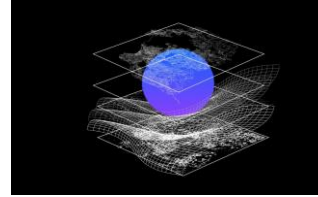
Project Leads: Barbara Zdrzil
(EMBL-EBI), Sebastian
Lobentanzer (UKHD)



Recruiting!

Developing foundation models to analyse single- cell perturbation data

Project Lead: Mo Lotfollahi
(Sanger)



Some Open Targets AI areas of research interest

Recruiting!

Using AI for predicting combination targets in cancer

Project Leads: Evangelia
Petsalaki (EMBL-EBI) &
Mathew Garnett (Sanger)



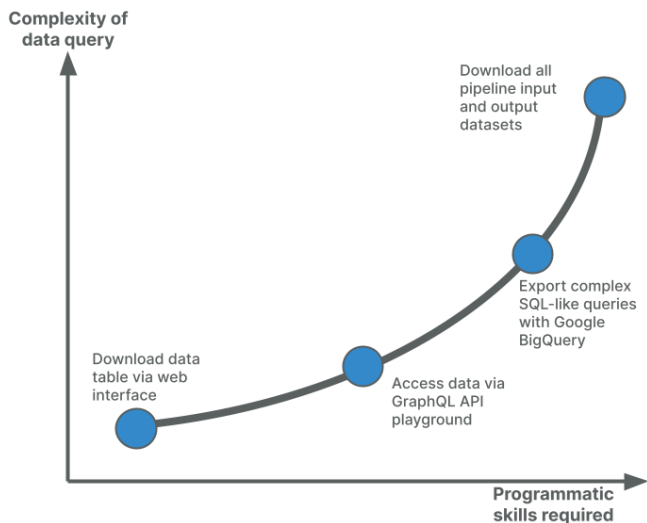
Digital Twins for disease modelling

Project Leads: Sheriff
Rahuman (EMBL-EBI) & Ellie
McDonagh (EMBL-EBI/OT)



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)

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

Acknowledgements



emcdonagh@ebi.ac.uk