

Open Targets Genetics

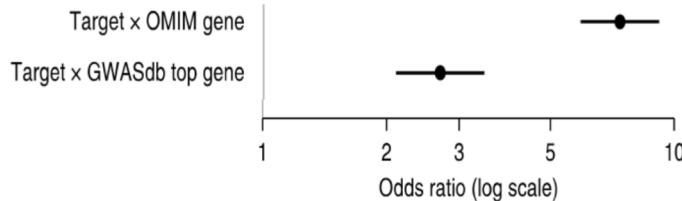
Integrating evidence from genome-wide associations and
functional genomics to identify and prioritise drug targets

Maya Ghoussaini

Genetics increase chances of 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}, Mulin Jun Li^{6,7}, Junwen Wang^{6,7}, Lon R Cardon⁸, John C Whittaker² & Philippe Sanseau²



bioRxiv
THE PREPRINT SERVER FOR BIOLOGY

HOME | ...
| CHAN
Search

New Results

Comment on this paper

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This article is a preprint and has not been peer-reviewed [what does this mean?].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

How can genetics be used to guide therapeutic target identification and prioritisation?

targetvalidation.org

- Rare and complex diseases
- Search for **target** or **disease**
- Germline and somatic mutations, drugs, pathways, expression, text mining, animal models
- Released 2015

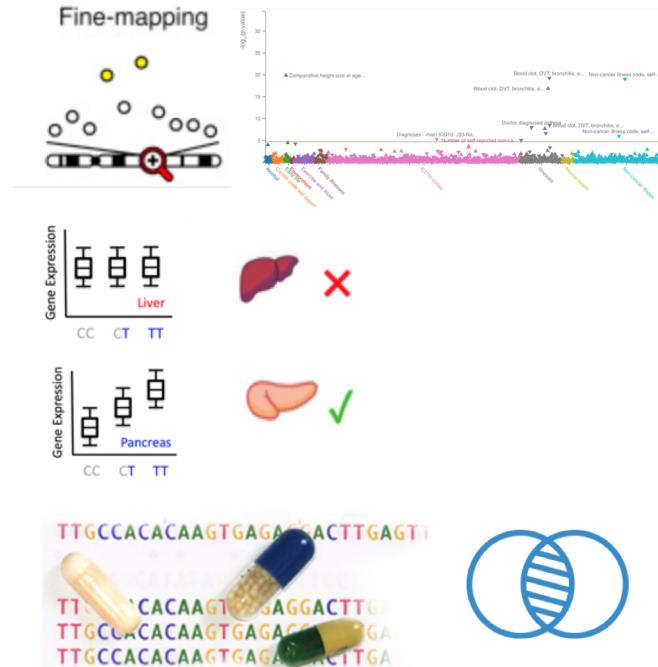
genetics.opentargets.org

- Complex diseases, common genetic variation
- Search for **variant**, **target**, or **disease**
- Variant-to-gene scoring
- Functional genomics, QTLs
- Released 2018



Some of the biological questions

- What is the minimal set of likely causal variants?
- Variants associated with other phenotypes?
- Gene(s) functionally implicated?
Directionality? Tissues/cell types?
- Are there known drugs targeting this gene?
- Which traits share common susceptibility loci?





Open Targets Genetics

genetics.opentargets.org

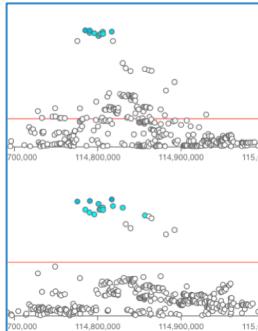


Fine-mapping (UK Biobank & GWAS
Catalog summary statistics repository)



Find shared susceptibility loci

Disease-disease and disease-molecular trait colocalisation
(New feature!)



Connect genes with known drugs in Open Targets Platform



Key Features

1_10596022_C_T rs2506889	
Gene	Overall V2G ↓
PEX14	●
PGD	●
APITD1	●
CASZ1	•

Prioritise genes using an integrated functional score



Visualise associations between traits, variants, and genes

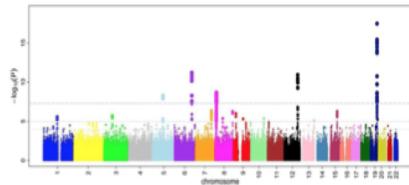


Open Targets Genetics

genetics.opentargets.org

Data Sources

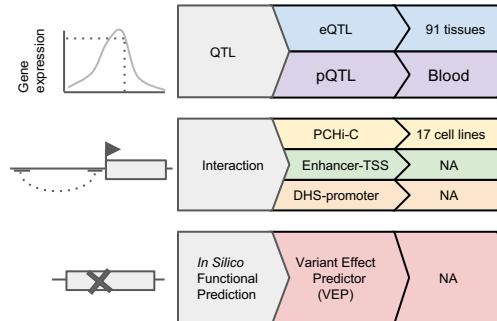
Human Genetics Variant – Disease



biobank^{uk}

1.7 billion
evidence
strings

Functional Genomics Variant – Gene

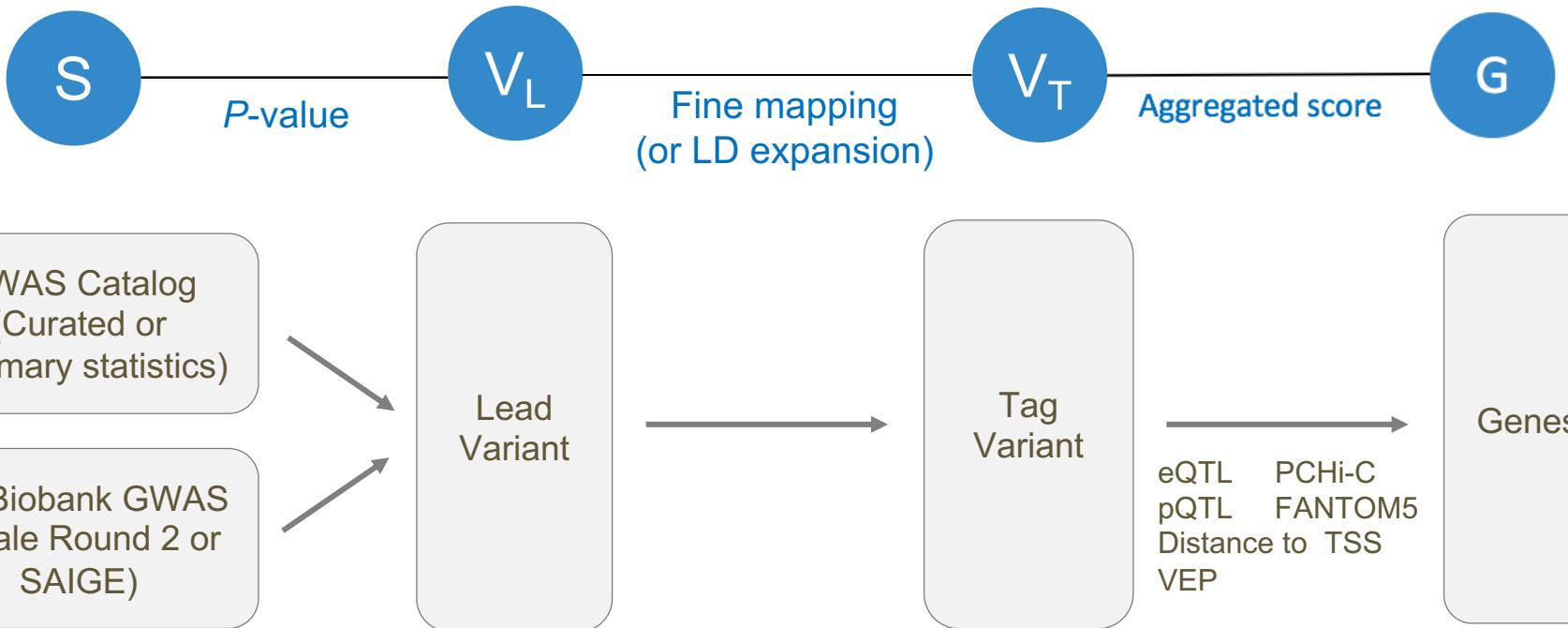


- ~5,000 GWAS Catalog studies (curated)
- 201 GWAS Catalog sum stats
- 2,139 UK Biobank sum stats (Neale Round 2)
- 1,281 UK Biobank binary sum stats (SAIGE)

- Gene expression (eQTL, 91 cell types)
- Protein abundance (pQTL, Sun et al. Nature, 2018)
- PChI-C (Javierre et al. Cell, 2016)
- Enhancer-TSS corr (Andersson et al. Nature, 2014)
- DHS-promoter corr (Thurman et al. Nature, 2012)
- Ensembl VEP score

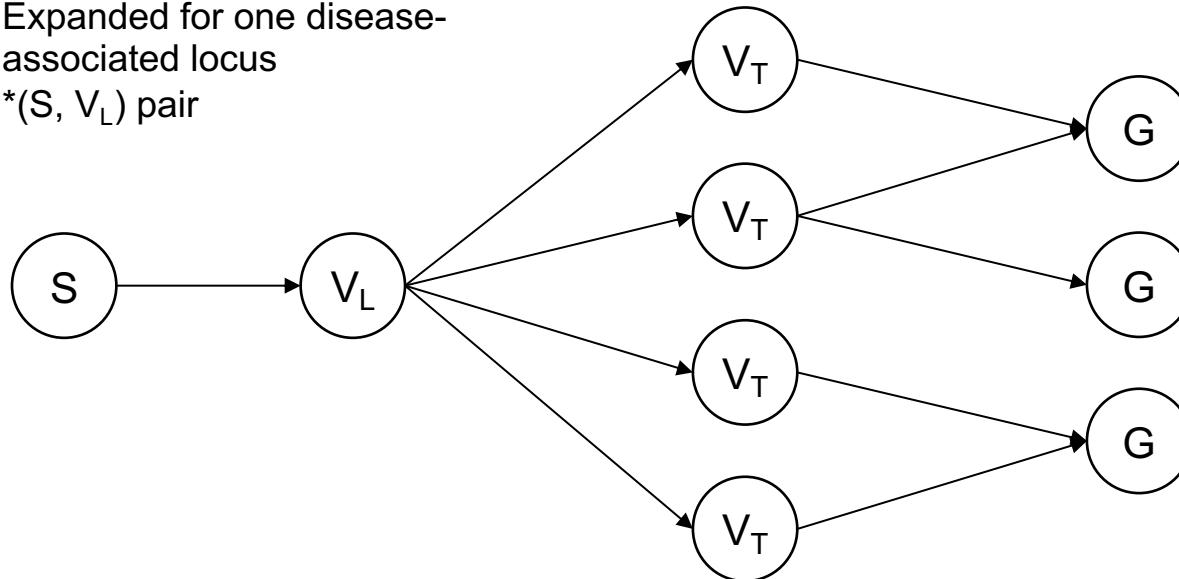


Data Model





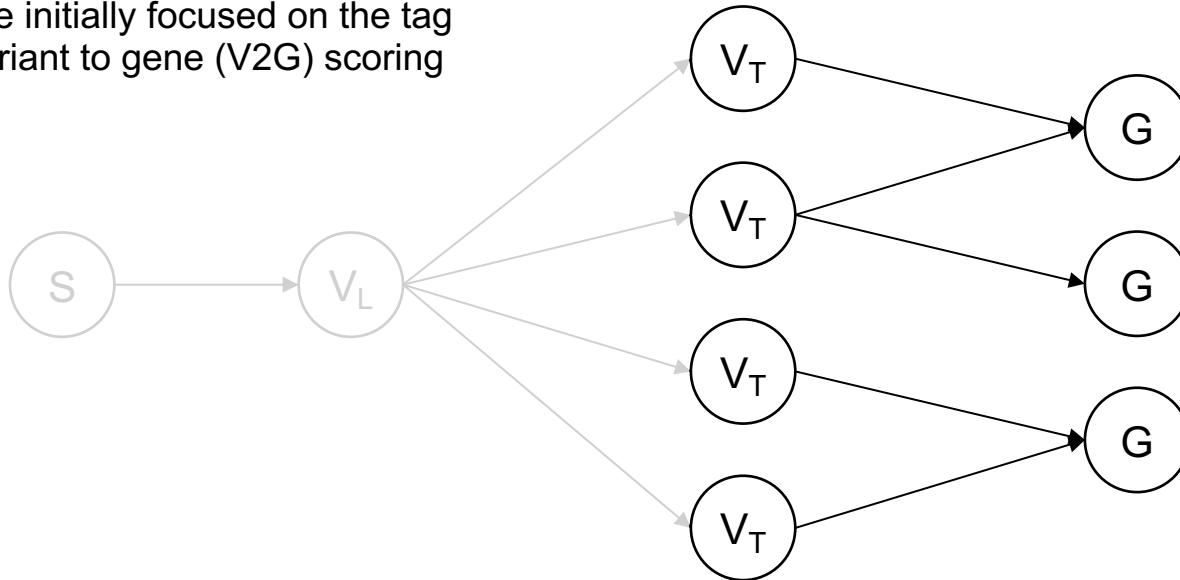
Expanded for one disease-associated locus
*(S, V_L) pair





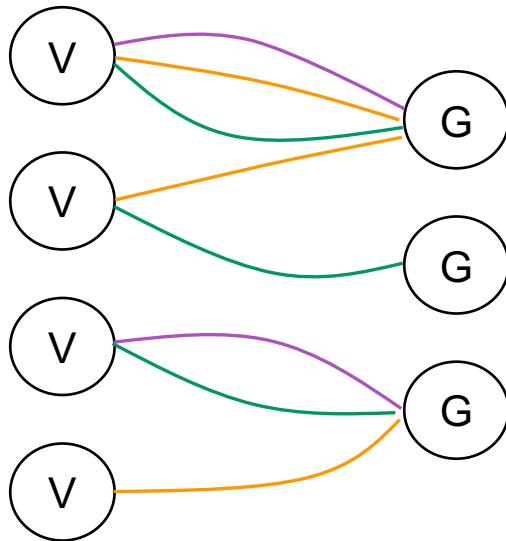
V2G Scoring

We initially focused on the tag variant to gene (V2G) scoring



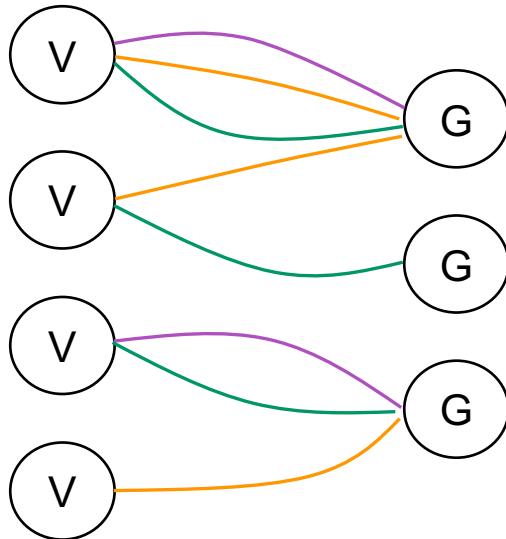


Evidence from many functional
genomics sources and tissues
for each (V, G) pair





Evidence from many functional genomics sources and tissues for each (V, G) pair

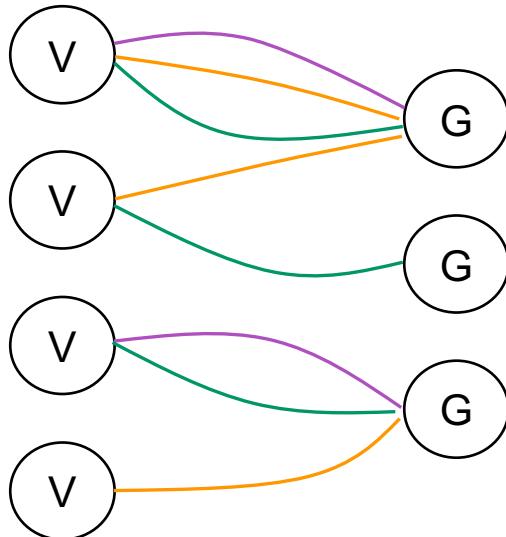


Quantitative trait loci (QTL) evidence

- Variant-centric association score
- Required cols: *chrom, pos, ref, alt, ensembl_id, beta, se, pval*
- Score used: $-\log_{10}(pval)$



Evidence from many functional genomics sources and tissues for each (V, G) pair

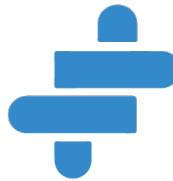


Quantitative trait loci (QTL) evidence

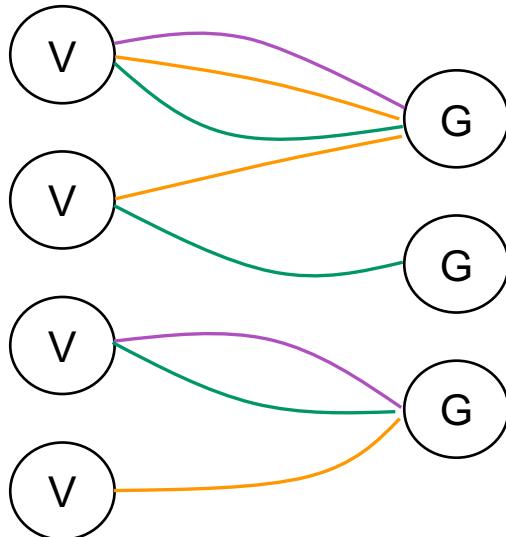
- Variant-centric association score
- Required cols: *chrom, pos, ref, alt, ensembl_id, beta, se, pval*
- Score used: $-\log_{10}(pval)$

Interval based evidence

- Association between two genomic ranges
- One range mapped to gene TSS, the other mapped to variants
- Required cols: *chrom, start, end, ensembl_id, score*
- Can be different score types, e.g.: CHiCAGO, R², q-value



Evidence from many functional genomics sources and tissues for each (V, G) pair



Quantitative trait loci (QTL) evidence

- Variant-centric association score
- Required cols: *chrom, pos, ref, alt, ensembl_id, beta, se, pval*
- Score used: $-\log_{10}(pval)$

Interval based evidence

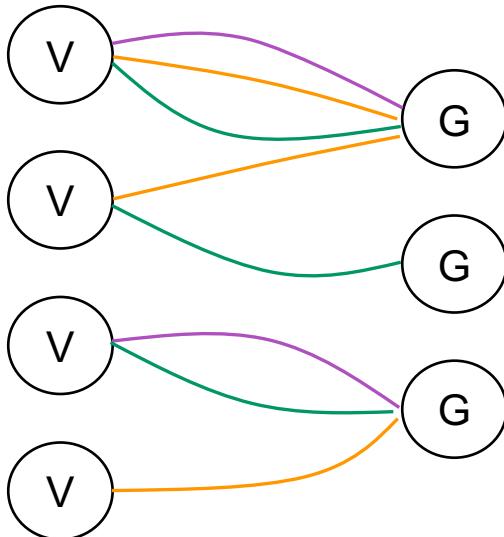
- Association between two genomic ranges
- One range mapped to gene TSS, the other mapped to variants
- Required cols: *chrom, start, end, ensembl_id, score*
- Can be different score types, e.g.: CHiCAGO, R², q-value

Functional prediction

- Variant-centric annotation
- VEP output in VCF format
- Transcripts mapped to genes
- Consequence impact mapped to scores: High → 1.0; Medium → 0.66; Low → 0.33; None → 0



Evidence from many functional genomics sources and tissues for each (V, G) pair



Quantitative trait loci (QTL) evidence

- Variant-centric association score
- Required cols: *chrom, pos, ref, alt, ensembl_id, beta, se, pval*
- Score used: $-\log_{10}(pval)$

Interval based evidence

- Association between two genomic ranges
- One range mapped to gene TSS, the other mapped to variants
- Required cols: *chrom, start, end, ensembl_id, score*
- Can be different score types, e.g.: CHiCAGO, R², q-value

Functional prediction

- Variant-centric annotation
- VEP output in VCF format
- Transcripts mapped to genes
- Consequence impact mapped to scores: High → 1.0; Medium → 0.66; Low → 0.33; None → 0

Distance variant-TSS



V2G Weighting

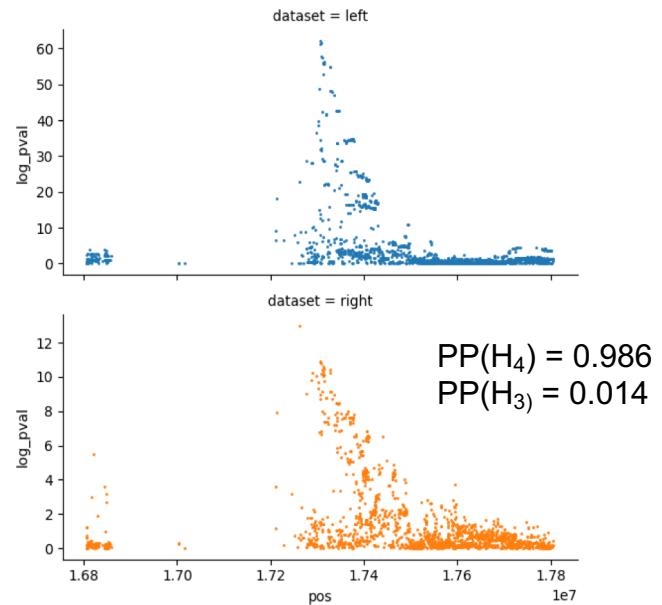
Data type	Experiment type	Source	Weighting
<i>In silico</i> functional prediction	Transcript consequence	VEP	1.0
QTL	eQTL	GTEX v7	0.66
QTL	pQTL	Sun <i>et al.</i> (Nature, 2018)	0.66
Interaction	PCHi-C	Javierre <i>et al.</i> (Cell, 2016)	0.33
Interaction	Enhancer-TSS correlation	Andersson <i>et al.</i> (Nature, 2014)	0.33
Interaction	DHS-promoter correlation	Thurman <i>et al.</i> (Nature, 2012)	0.33
Distance	Canonical TSS		0.33

Colocalisation Analysis

Colocalisation analysis: Used to test whether two independent association signals at a locus are consistent with having a shared causal variant

Used to test the hypotheses:

- H_3 : Independent association with trait 1 and trait 2
- H_4 : Shared association with trait 1 and trait 2



Open Targets-EMBL EBI eQTL Database + GTEx

Consortia	Cell types
GTEX	44 tissues
ALASOO	Macrophage naïve/IFNG/IFNG_Salmonella
BLUEPRINT	Monocyte, neutrophil, T_Cells
CEDAR	B cells, ileum, monocyte, neutrophil, platelet, rectum, transverse colon, Tcell_CD4/CD8
FAIRFAX	Monocyte naïve/IFN24/LPS2/LPS24, neutrophil, T_cells, B cell
GENCORD	Fibroblast, LCL, T cell
GEUVADIS	LCL
HIPSCI	iPSC
NARANBHAI	Neutrophil_CD16
NEDELEC	Macrophage naïve/Listeria/Salmonella
QUACH	Monocyte_IAV/LPS/Naïve/PAM3CSK4/R848
SCHWYZENTRUBER_2018	Sensory neuron
TWINSUK	Skin, fat, LCL
VAN_DE_BUNT_2015	Pancreatic islet



Future plans

New Data

Human Genetics

- UK Biobank exome data
- Generating a community gold standard repository

New Features

- Locus to gene score
- Therapeutic-area specific profiles for gene assignment
- Feedback from today's workshop



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

GraphQL API: <http://genetics-api.opentargets.io/>

Google BigQuery (SQL): <https://console.cloud.google.com/bigquery?p=open-targets-genetics>

Bulk Download

- **Google Cloud Storage:** <https://console.cloud.google.com/storage/browser/genetics-portal-output>
- **FTP via European Bioinformatics Institute (EBI):**
<ftp://ftp.ebi.ac.uk/pub/databases/opentargets/genetics/>



Google BigQuery



Google Cloud



Open Targets Genetics



Open Targets Genetics



Ellen Schmidt
Maya Ghousaini
Edward Mountjoy

Contact us:

✉ geneticsportal@opentargets.org
🐦 @targetvalidate

Core Platform Team



Ian Dunham



David Hulcoop



David Ochoa



Elaine McAuley



Denise Carvalho Silva



Adam Faulconbridge



Michaela Spitzer



Luca Fumis



Asier Gonzalez



Gareth Peat



Eirini Petsalaki



Cinzia Malangone



Andrew Hercules



Alfredo Miranda



Miguel Carmona

GWAS Catalog



Annalisa Buniello



James Hayhurst

POSTGAP/eQTL DB



Daniel Zerbino



William Newell



Open Targets Genetics

genetics.opentargets.org

Demo

genetics.opentargets.org



Open Targets Genetics

genetics.opentargets.org

Questions
