DisGeNET knowledge platform on disease genomics

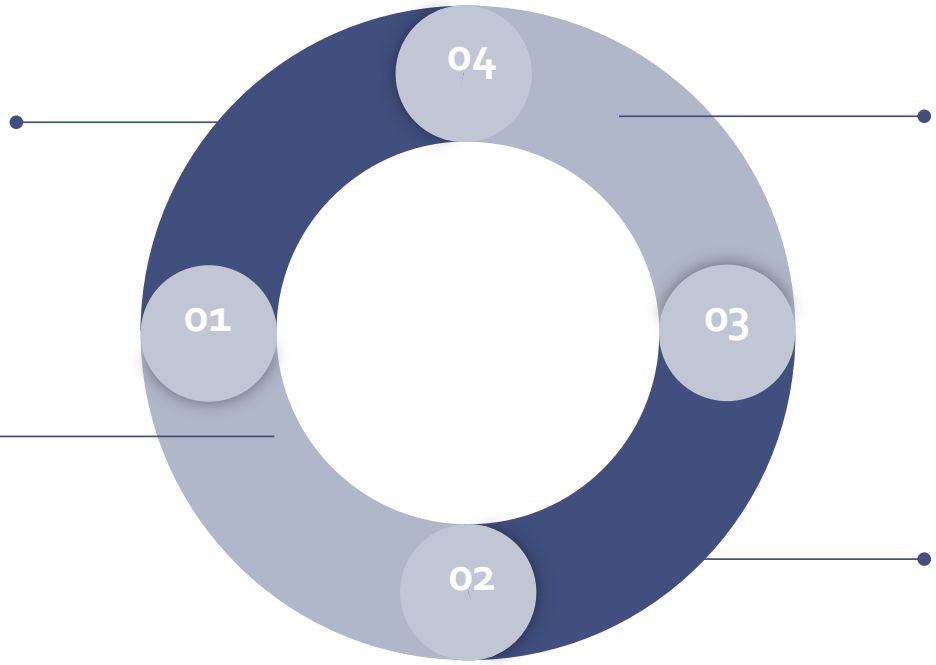
Research Programme on Biomedical Informatics (GRIB), Hospital del Mar Research Institute (IMIM), Universitat Pompeu Fabra (UPF)

February 2021

Current challenges in exploiting disease gen-phen data

Large volume
More than 500,000 associations between genomic variants and common diseases have been discovered thanks to NGS.

Data silos
Data is scattered across numerous, often disease-specific, repositories that do not communicate with one another.

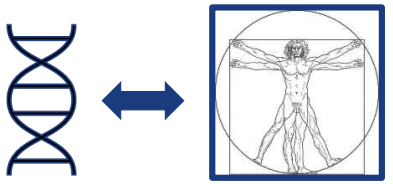


Innovation locked in publications
More than 95,000 papers published each year on disease genomics (more than 8,000 publications/month), leading to an information overload.

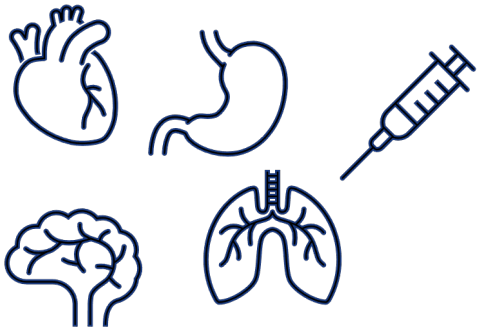
Lack of interoperability
Computers need standardized data to make proper connections among datasets.

DisGeNET

Comprehensive knowledge database **integrating** and **standardizing** information on disease associated genes and variants.



Coverage of the **full spectrum of human diseases** as well as normal and abnormal traits, and **adverse drug events**.



elixir Recommended Interoperability Resources

Interoperable resource supporting a variety of applications in genomic medicine and drug R&D.

Resources on disease genomics

ClinVar
Clinically relevant variation

ClinGen

isyGeNet

MGI

GWA db

ctd

RGD

GWAS Catalog

Genomics england

orphanet

UniProt

Scientific literature

Text Mining

DisGeNET

Findable Accessible Interoperable Reusable

Data enriched with

- Scores, provenance
- Standards
- Information from other resources

Data available through

- Web interface
- REST & SPARQL API
- Cytoscape App
- R package
- Beacon
- Datasets download

DisGeNET statistics

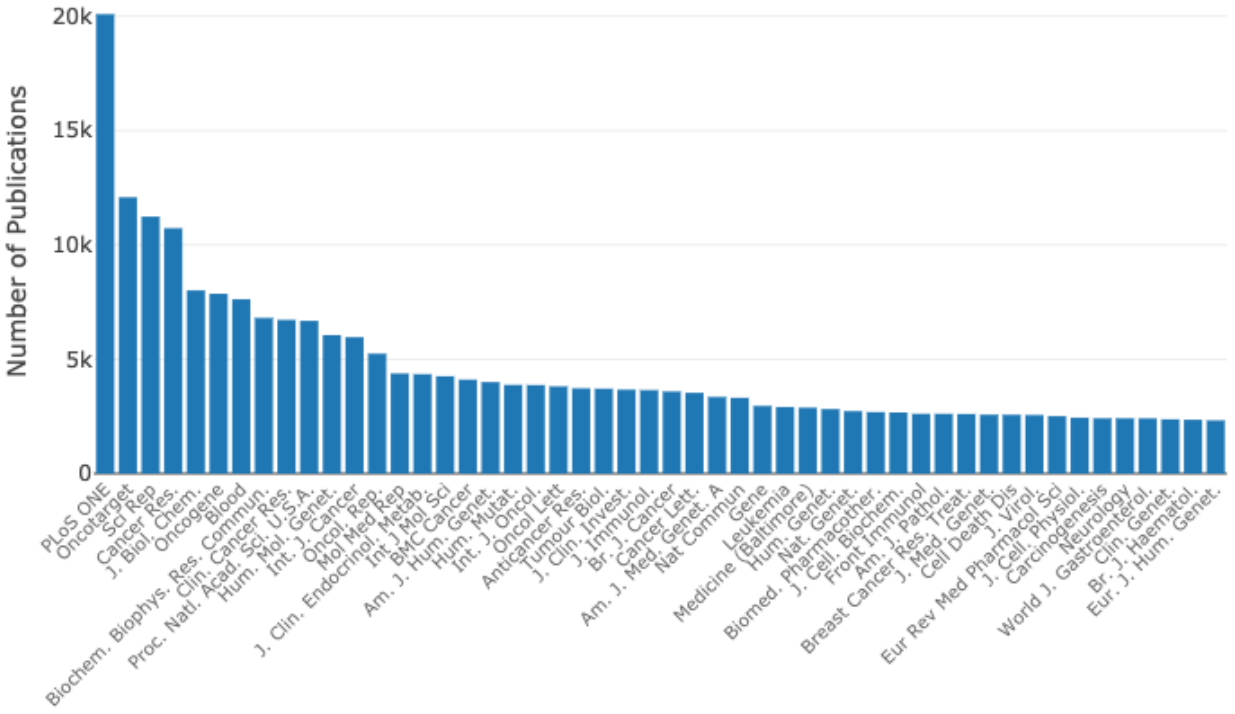
1,13M GDAs
370K VDAs

GENOTYPE ⟷ **PHENOTYPE**

22,000 genes
195,000 variants

30,000 diseases & traits

830,000 publications



Exploring genotype-phenotype information from different perspectives

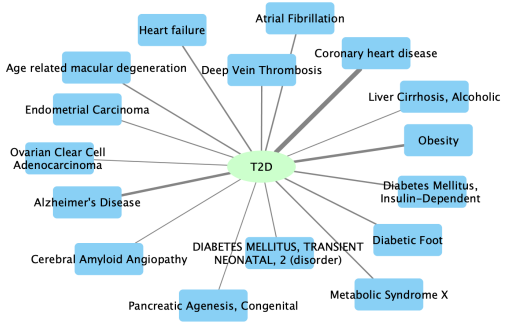
- Drug target identification
- Selection of genes for sequencing panels

Gene-Disease Association

Variant-Disease Association

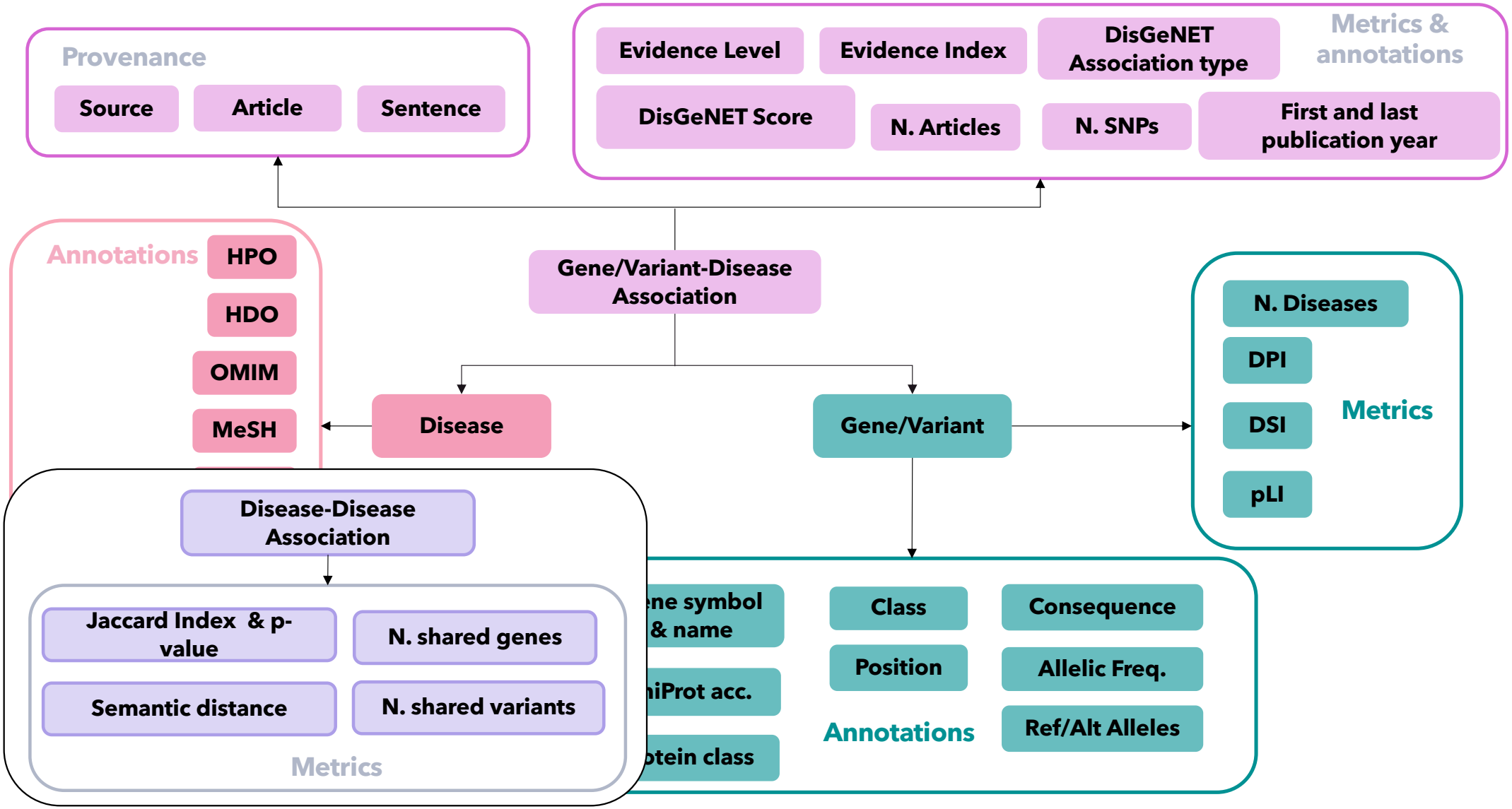
- Variant interpretation
- Analysis of GWAs and sequencing studies

DisGeNET



Disease-Disease Association

- Comorbidity studies
- Finding similar diseases


Provenance
Source
Article
Sentence
Evidence Level
Evidence Index
DisGeNET Association type
Metrics & annotations
DisGeNET Score
N. Articles
N. SNPs
First and last publication year
Annotations
HPO
HDO
OMIM
MeSH
Gene/Variant-Disease Association
Disease
Gene/Variant
N. Diseases
DPI
DSI
pLI
Metrics
Disease-Disease Association
Jaccard Index & p-value
N. shared genes
Semantic distance
N. shared variants
Metrics
ene symbol & name
iProt acc.
otein class
Class
Position
Annotations
Consequence
Allelic Freq.
Ref/Alt Alleles

Impact

- More than 2,000 citations
- More than 56,000 web users in 2020
- Linked by several ELIXIR core resources and other databases: UniProt, Reactome, PMC Central, NIH Pharos, NextProt, EMBL VEP, etc

Examples of use in disease genomics

Shared genetic architecture between blood pressure and other traits & diseases.

nature genetics

Explore our content ⌄ Journal information ⌄

nature › nature genetics › articles › article

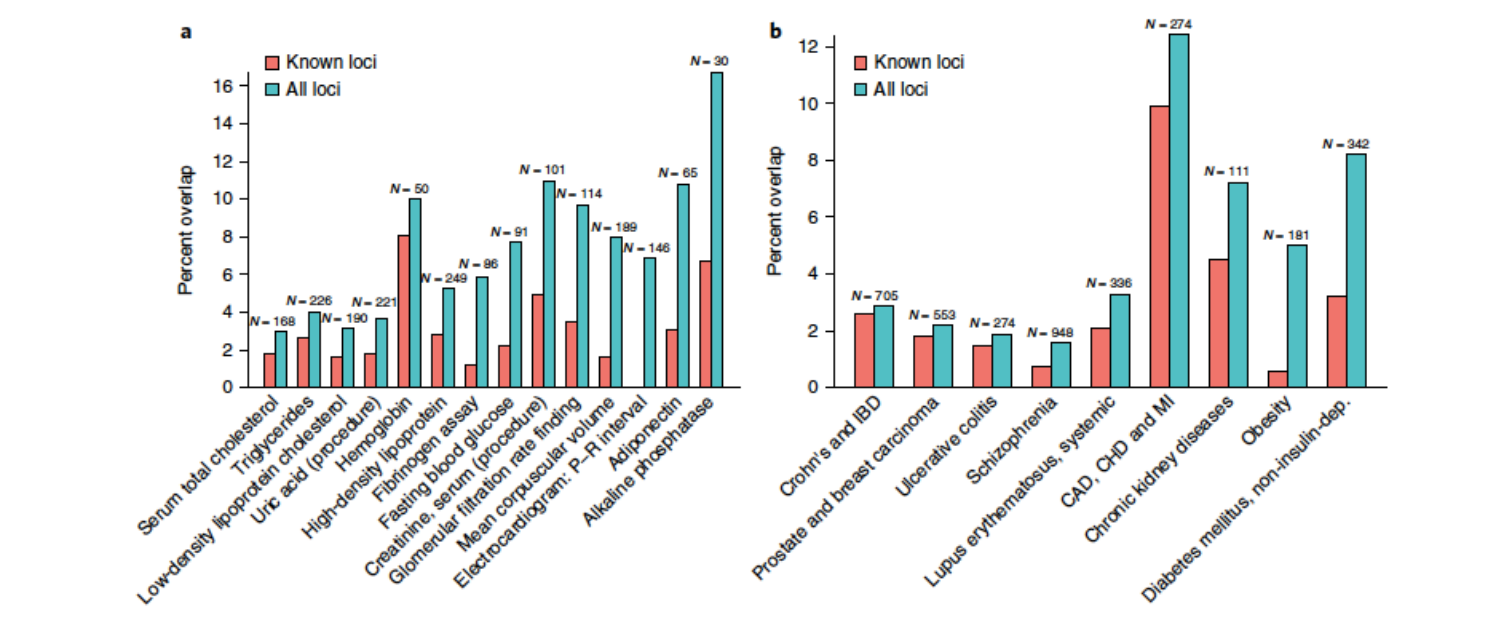
Article | Published: 17 September 2018

Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits

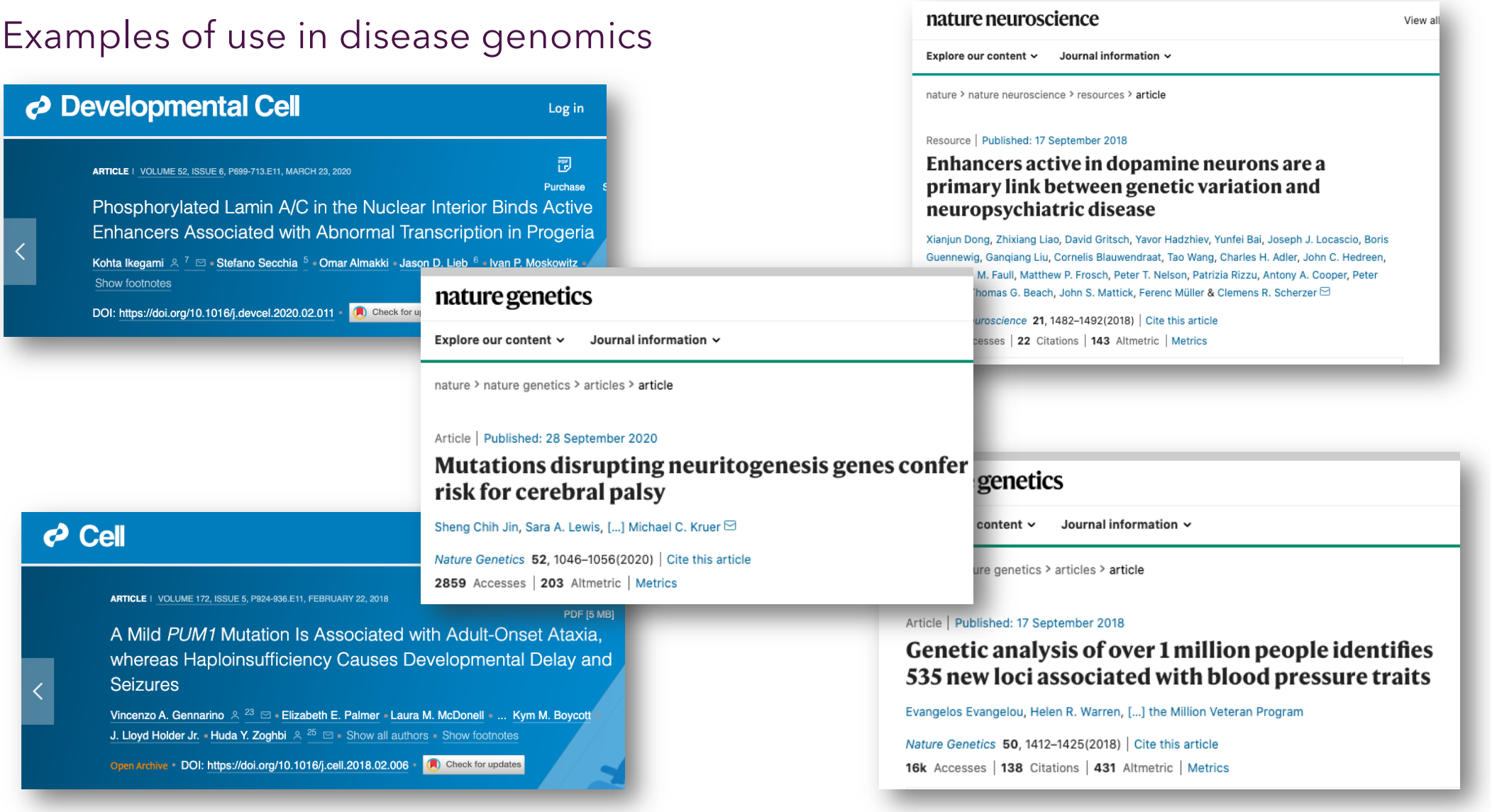
Evangelos Evangelou, Helen R. Warren, [...] the Million Veteran Program

Nature Genetics **50**, 1412–1425(2018) | Cite this article

16k Accesses | **138** Citations | **431** Altmetric | Metrics

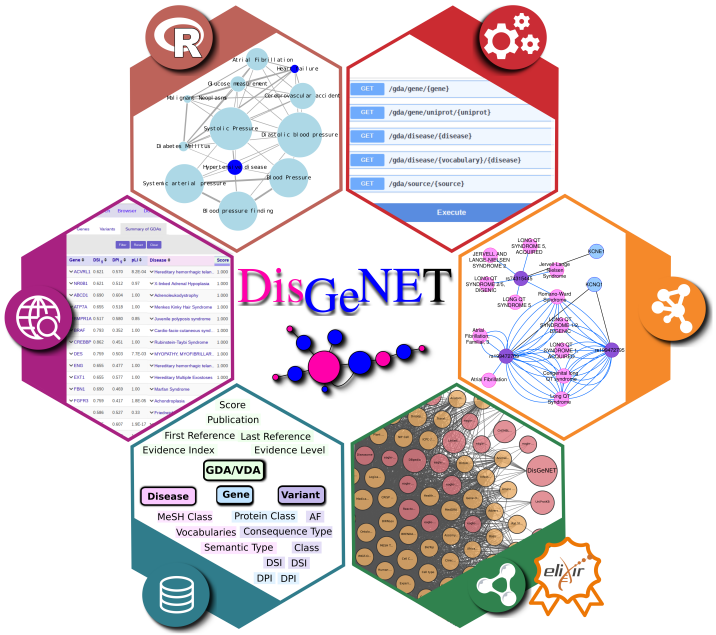


Examples of use in disease genomics

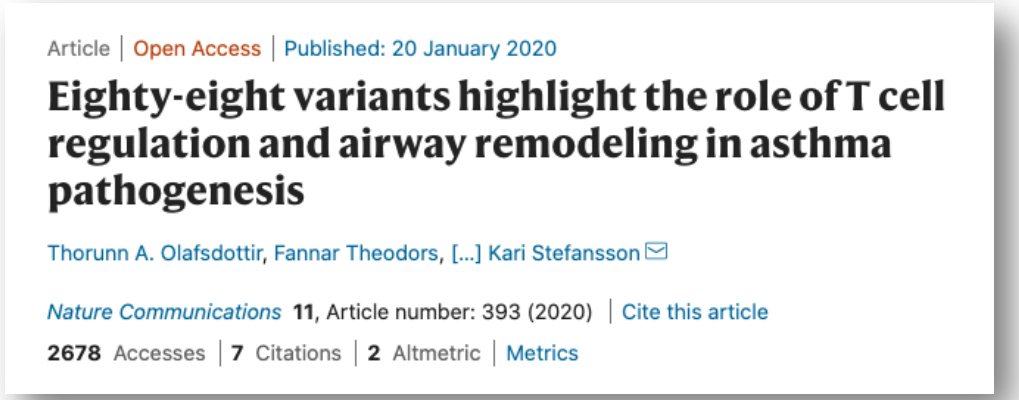


DisGeNET tools

- New REST API available, including a disease enrichment function for genes and variants
- New release of the DisGeNET Cytoscape App, including :
  - disease enrichment function for genes and variants
  - Exposing functionalities through Cytoscape Automation



Application in Precision Medicine: demo



Article | Open Access | Published: 20 January 2020

Eighty-eight variants highlight the role of T cell regulation and airway remodeling in asthma pathogenesis

Thorunn A. Olafsdottir, Fannar Theodors, [...] Kari Stefansson

Nature Communications **11**, Article number: 393 (2020) | Cite this article

2678 Accesses | **7** Citations | **2** Altmetric | Metrics

- Genome-wide association meta-analysis of 69K cases and 702K controls from Iceland and UK biobank on asthma
- Report 88 asthma risk variants at 56 loci (19 previously unreported)

DisGeNET

- Are the risk variants associated to asthma? What is the evidence for each association?
- What asthma subtypes are they associated to?
- Are the risk variants associated to other diseases and phenotypes?
- Are the genes to which these variants map also associated to asthma?

IMIM Institut de Recerca Hospital del Mar

Parc de Salut MAR Barcelona

GRIB RESEARCH PROGRAMME ON BIOMEDICAL INFORMATICS

upf. Universitat Pompeu Fabra Barcelona

Integrative Biomedical Informatics Group

http://www.disgenet.org/

@DisGeNET

support@disgenet.org

RIS3CAT - VEIS
VALORIZACIÓN de EGA para la INDUSTRIA y la SOCIEDAD

PRB2

transQST

elixir SPAIN

Unión Europea
Fondo Europeo de Desarrollo Regional

INB Spanish National Bioinformatics Institute

excelerate

Generalitat de Catalunya
Departament d'Empresa i Coneixement

Instituto de Salud Carlos III

eTRANSAFE

Unió Europea
Fons Europeu de Desenvolupament Regional