

Chemical biology resources at EMBL-EBI, 2016-2024



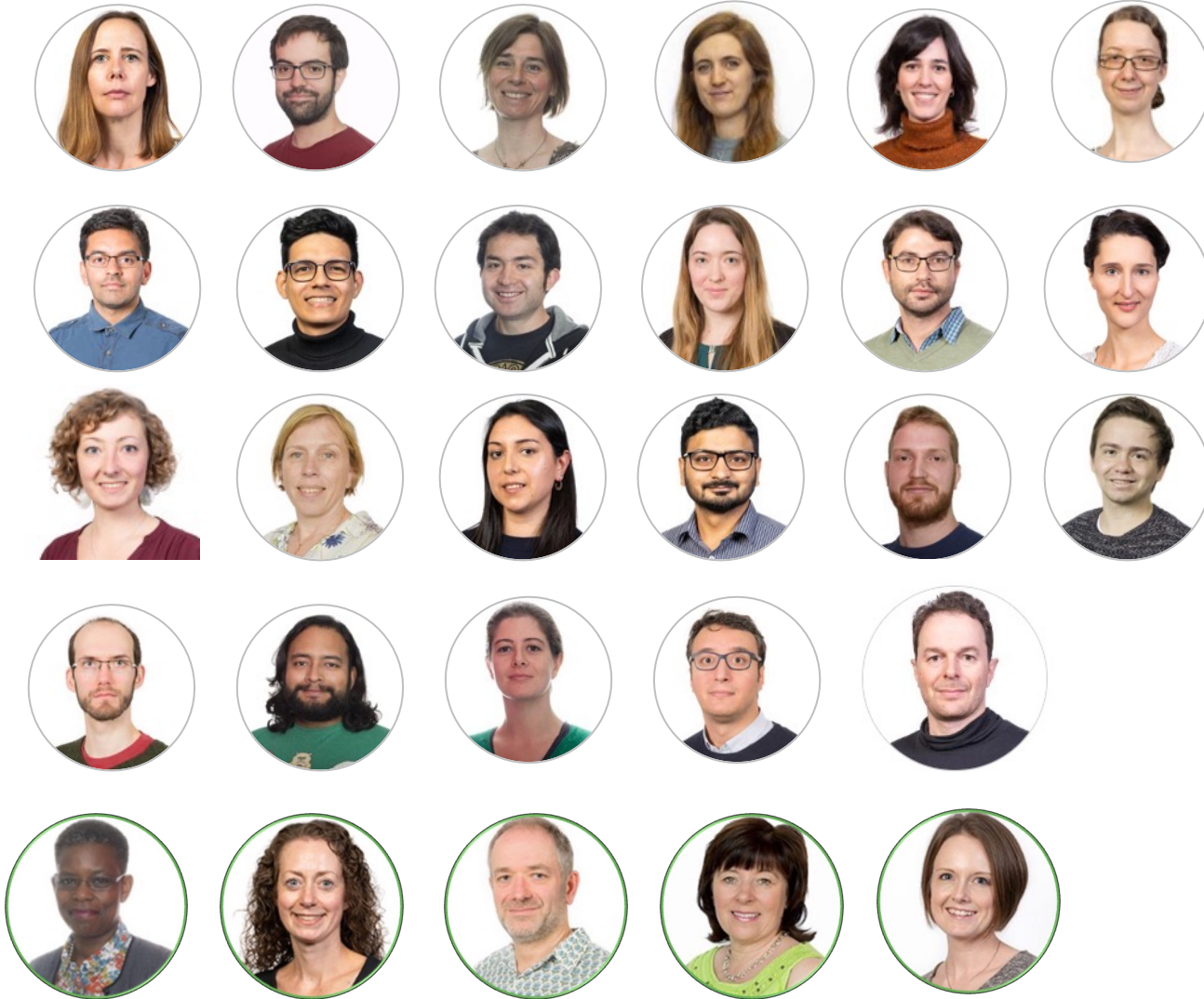
Andrew Leach

Formerly

Head of Chemical Biology and
Head of Industry Partnerships



None of the following would be possible without an amazing team



EBI's chemical biology resources



ChEBI
~60K structures



2004



ChEMBL
>2m structures



2009



UniChem
Links ~40 resources
>175m structures

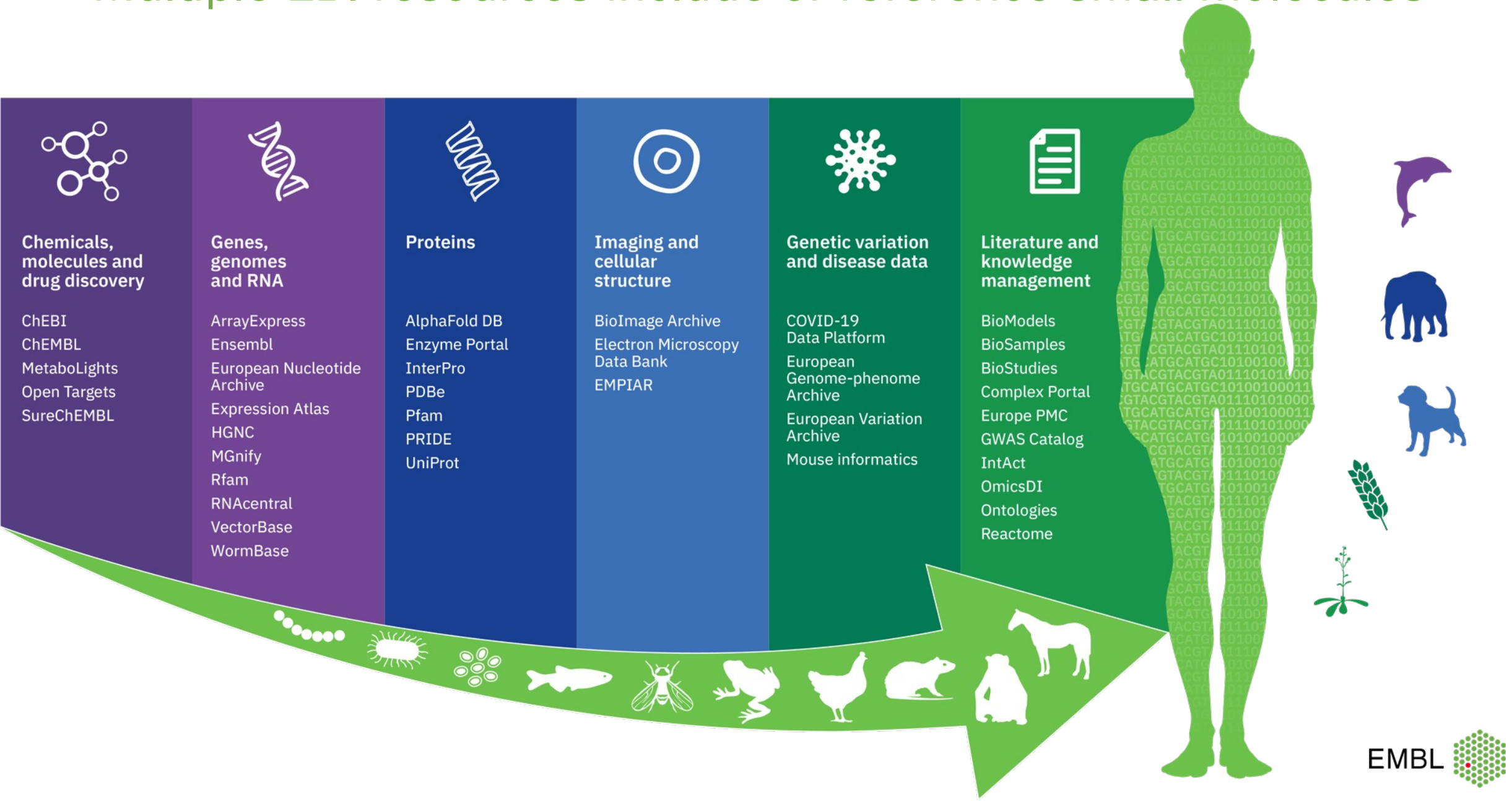
2013



SureChEMBL
>20m structures

2014

Multiple EBI resources include or reference small molecules



A few “headline” achievements

>50 Team
members

£10m in new
funding

Significant
growth in data
resources

Major software
overhaul and
rewrite

GBC and Elixir
accreditation
for ChEMBL &
ChEBI

>40
publications



















Two service
reviews, one
research
review

Open Targets
impact incl.
tractability and
safety

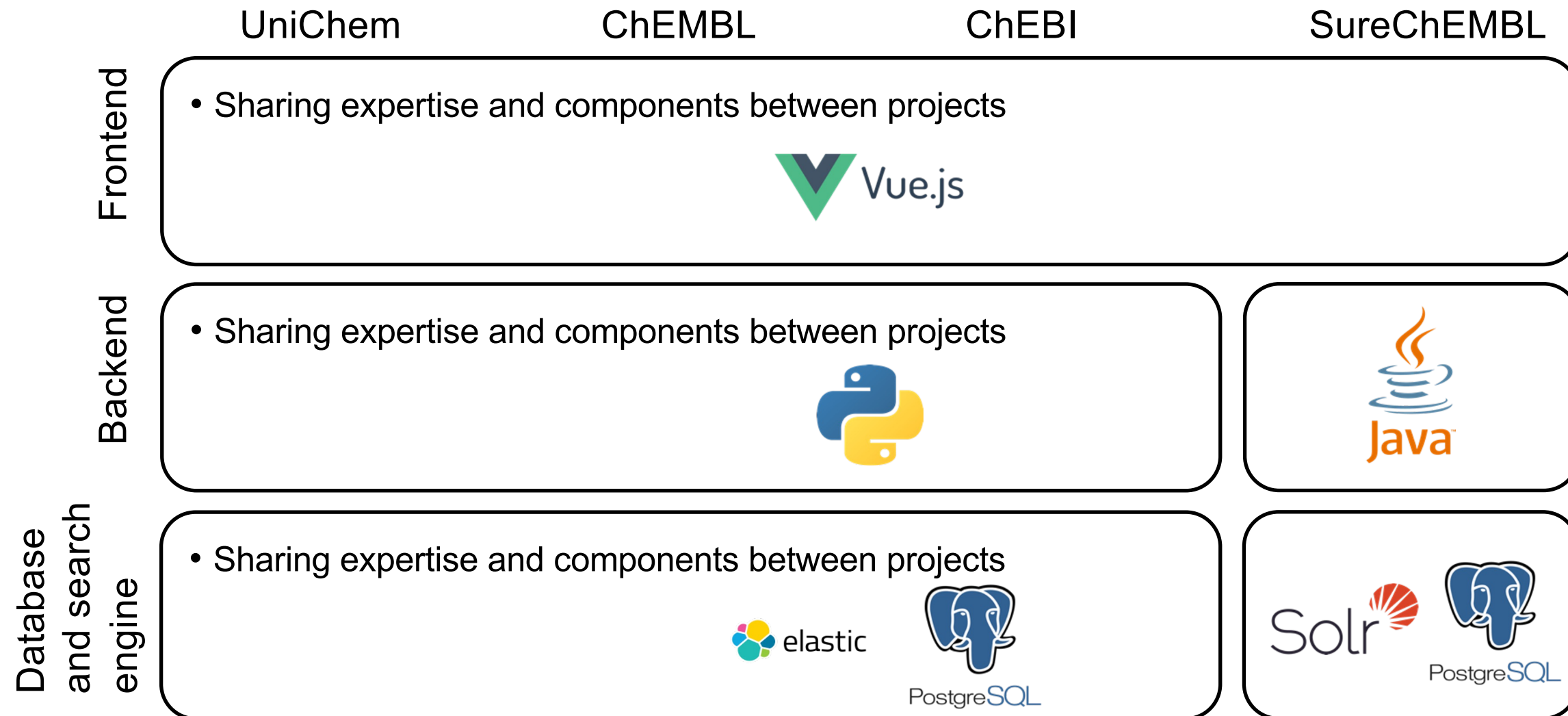
Multiple
collaborations

Survived 1
pandemic

Software infrastructure: legacy

	UniChem	ChEMBL	ChEBI	SureChEMBL
Frontend	Vanilla JS, no framework used 	 BACKBONE.JS  Vue.js	Vanilla JS, no framework used 	Vanilla JS, no framework used 
Backend	 			  
Database and search engine		  ORACLE PostgreSQL	 ORACLE	  Solr PostgreSQL

Software infrastructure: future



New interfaces for all our resources

The screenshot shows the ChEMBL search results page. At the top, there's a navigation bar with the ChEMBL logo and links to Home, Search, Downloads, Web Services, and More. A search bar contains the text 'herg'. Below the navigation bar, a yellow banner asks users to take a 10-minute survey. The main content area is titled 'Search Results' and shows a list of results for 'herg'. The results are displayed in a table with columns for ChEMBL ID, Name, and Uniprot Accession. The first result is ChEMBL240, which is a hERG channel. The table also includes a 'Type' column and a 'By Std. Type' column. The page is filtered by 'Organism Taxonomy L1' and 'Protein Classification L1'.

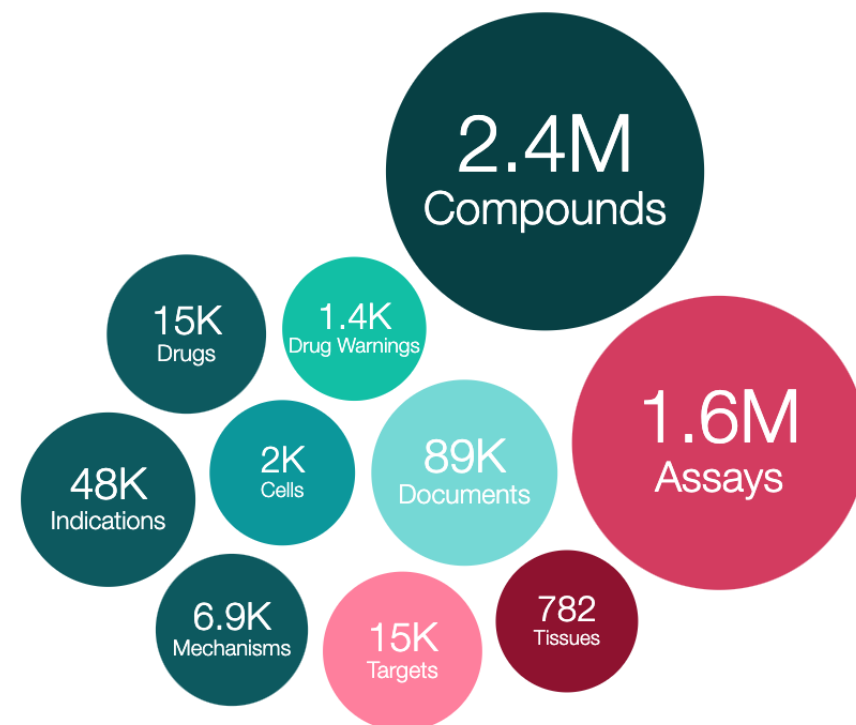
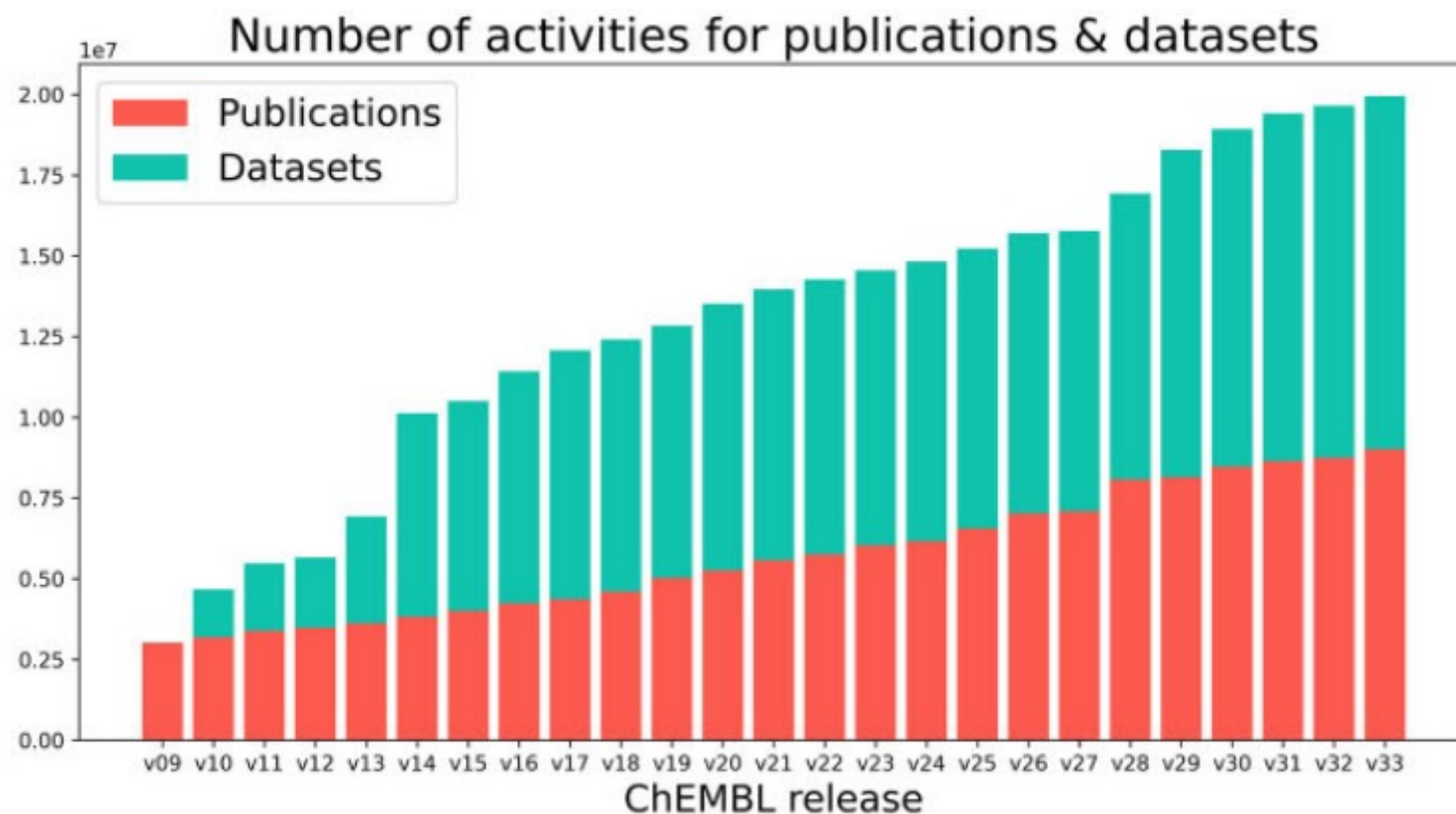
The screenshot shows the SureChEMBL Beta search page. At the top, there's a navigation bar with the SureChEMBL Beta logo and links to Search, Downloads, Wiki, and Contact Us. A search bar contains the text 'diabet*'. Below the search bar, there are three toggle switches: 'All chemically annotated authorities' (checked), 'Specify dates' (unchecked), and 'Structure search' (unchecked). A yellow banner at the top right asks users to take a 10-minute survey.

Coming soon!

The screenshot shows the ChEBI homepage. At the top, there's a navigation bar with the ChEBI logo and links to Advanced Search, Submit, Downloads, Tools, Docs, About, and Contact. The main content area is titled 'Chemical Entities of Biological Interest' and describes it as a manually curated database and ontology of chemical entities. Below the title, there's a search bar and a 'Search' button. The page is divided into six sections, each with an icon and a description: 'Advanced Search' (magnifying glass icon), 'Submit' (submit icon), 'Downloads' (download icon), 'Tools' (wrench icon), 'Documentation' (document icon), and 'About ChEBI' (ChEBI logo icon).

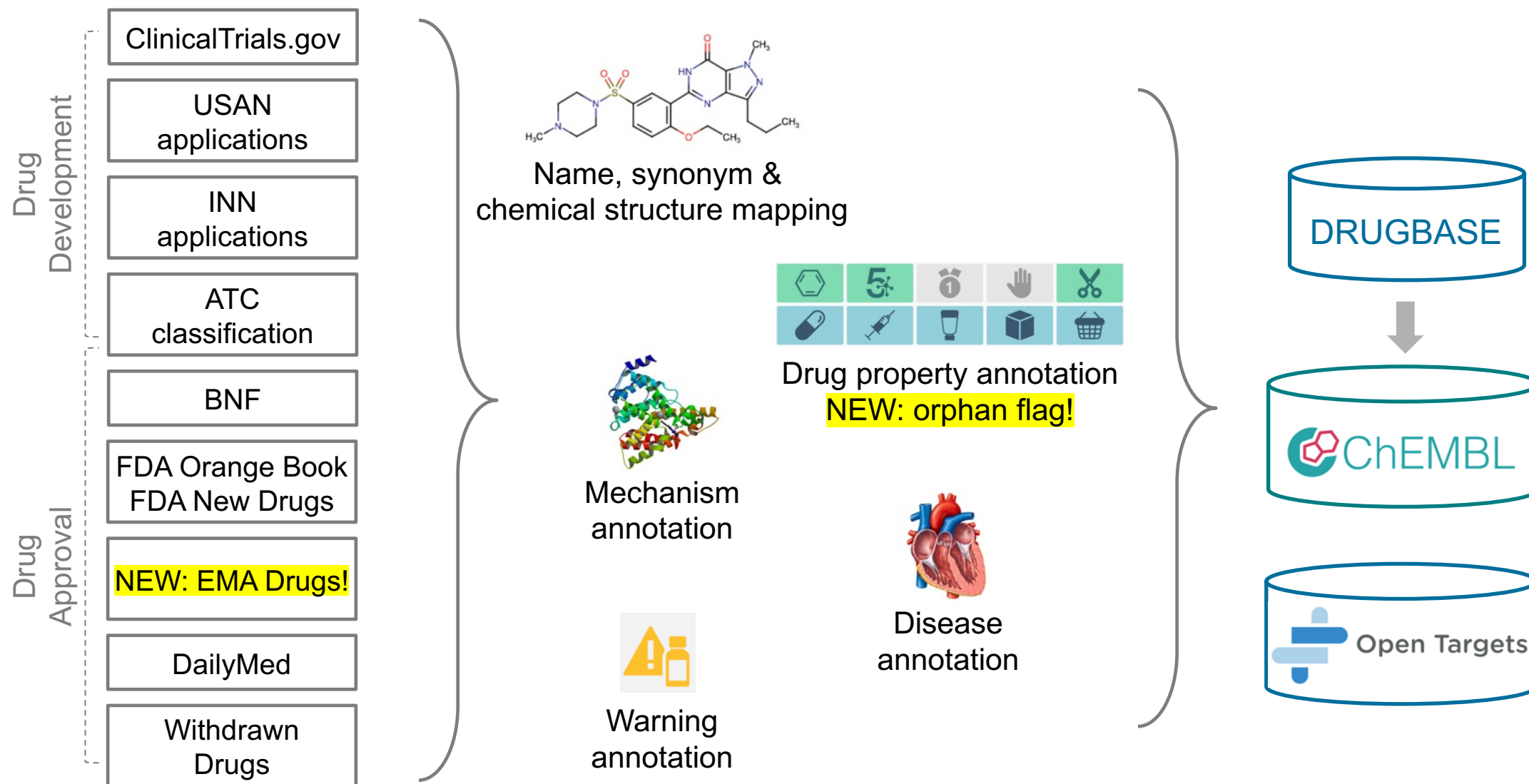
We have also stopped and retired some things

Continued growth in database size and content: ChEMBL

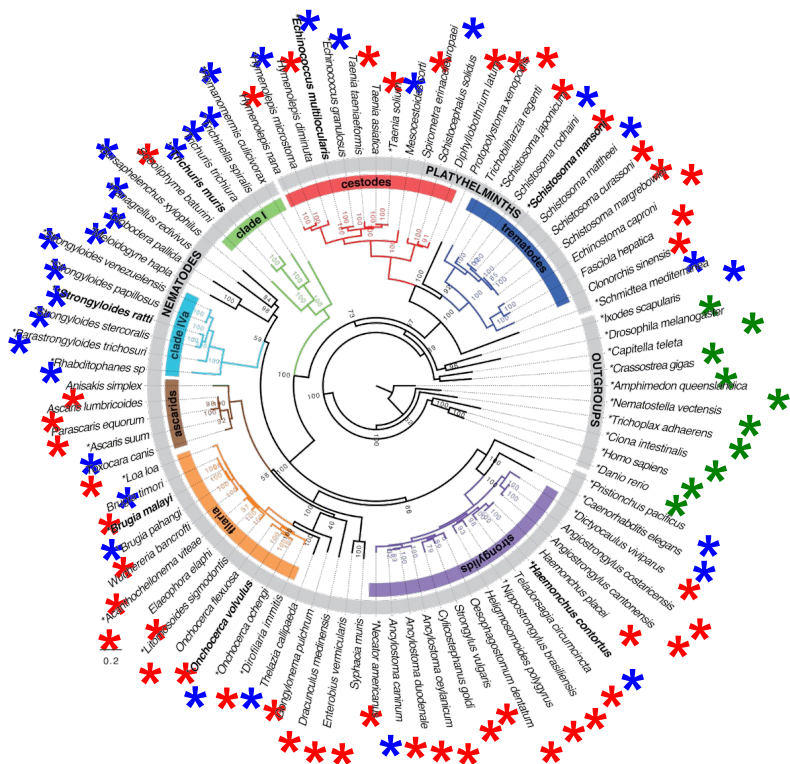


Data depositions are very welcome!

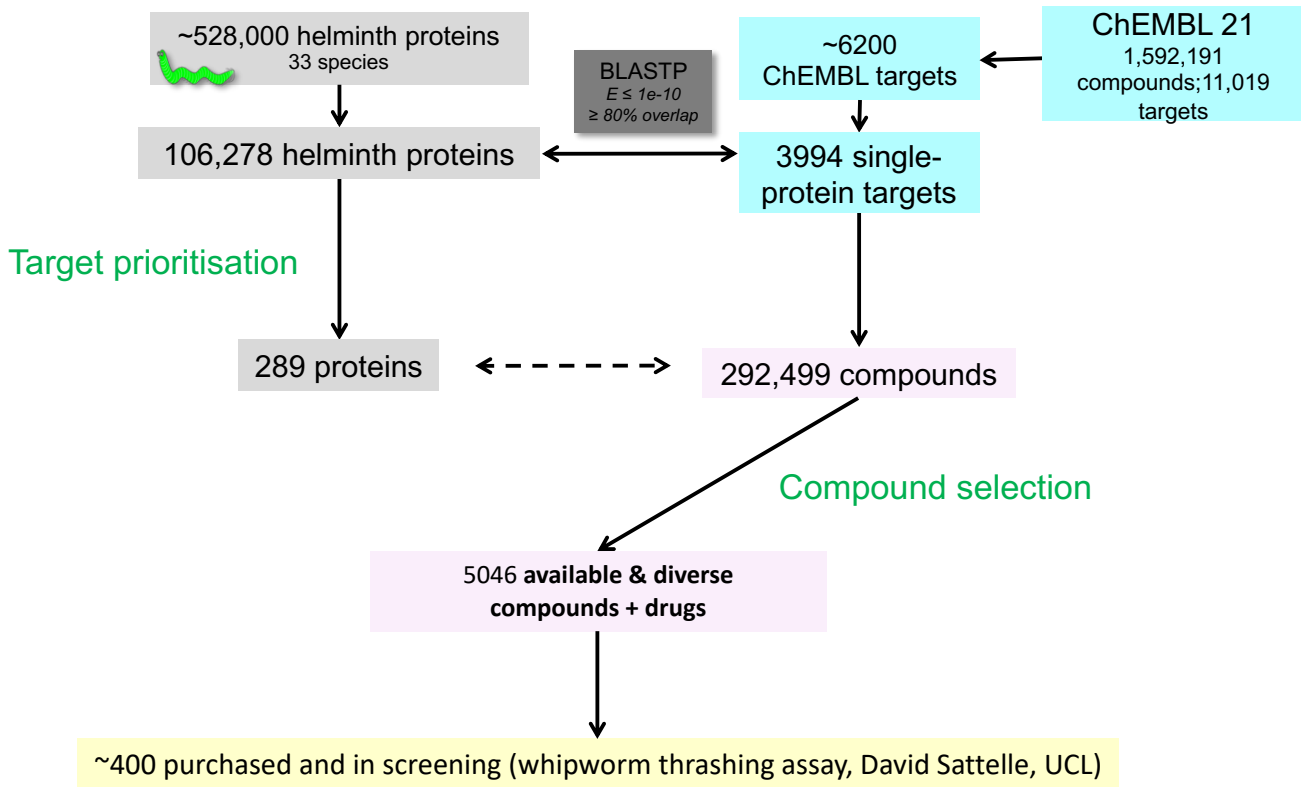
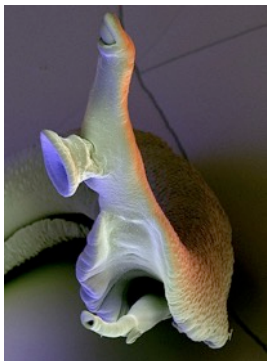
ChEMBL Drug and Clinical Candidate data: a unique resource!



Helminth chemogenomics analysis and compound selection



*33 published Helminth sequences; image by James Cotton & Bhavana Harsha



ARTICLES

<https://doi.org/10.1038/s41588-018-0262-1>

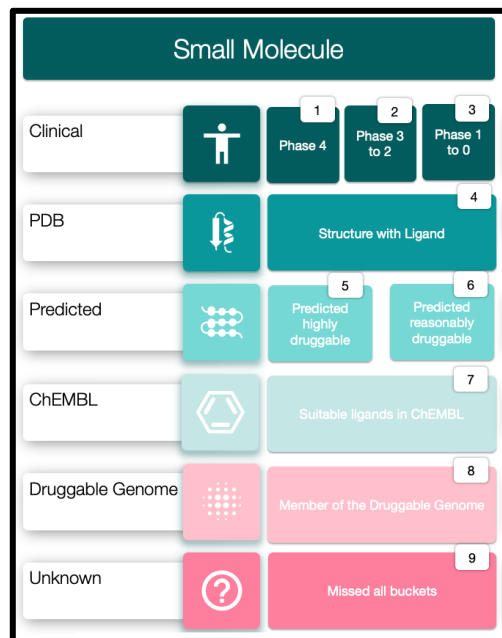
OPEN

Comparative genomics of the major parasitic worms

International Helminth Genomes Consortium*

Parasitic nematodes (roundworms) and platyhelminths (flatworms) cause debilitating chronic infections of humans and animals, decimate crop production and are a major impediment to socioeconomic development. Here we report a broad comparative study of 81 genomes of parasitic and non-parasitic worms. We have identified gene family births and hundreds of expanded gene families at key nodes in the phylogeny that are relevant to parasitism. Examples include gene families that modulate host immune responses, enable parasite migration through host tissues or allow the parasite to feed. We reveal extensive lineage-specific differences in core metabolism and protein families historically targeted for drug development. From an in silico screen, we have identified and prioritized new potential drug targets and compounds for testing. This comparative genomics resource provides a much-needed boost for the research community to understand and combat parasitic worms.

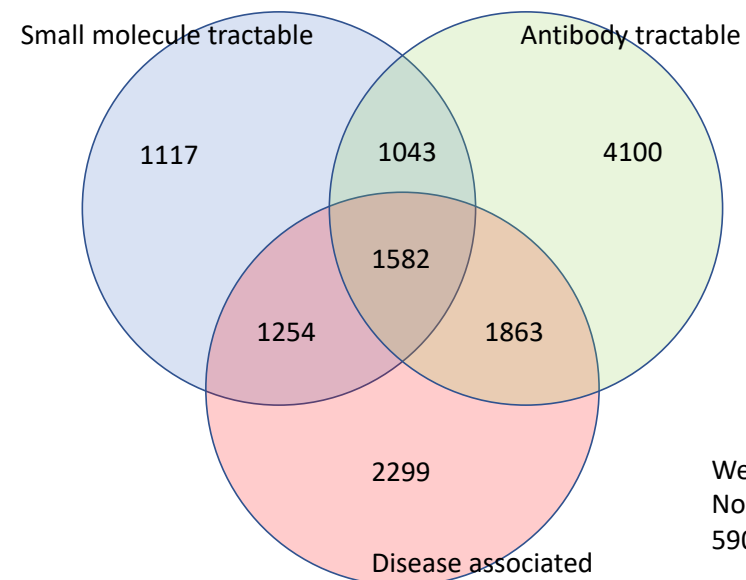
Tractability assessment: a key factor in target selection



Summary of tractability information for PSEN1 for small molecule and antibody modalities.

Source: [Open Targets](#)

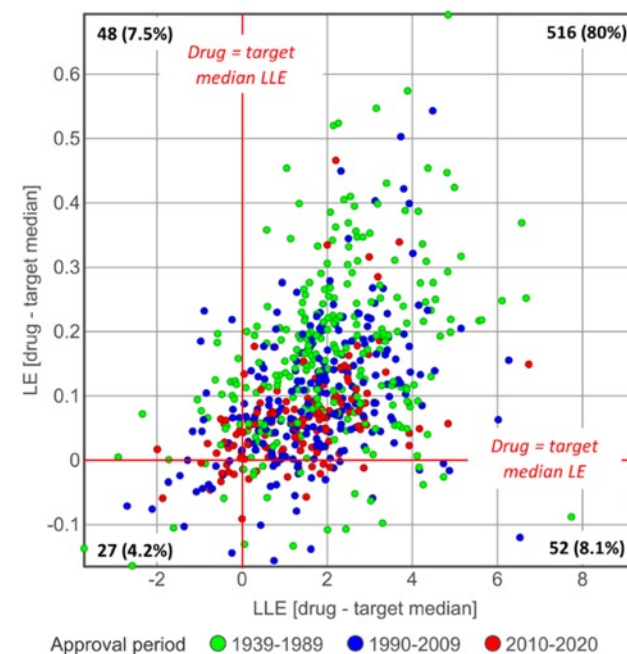
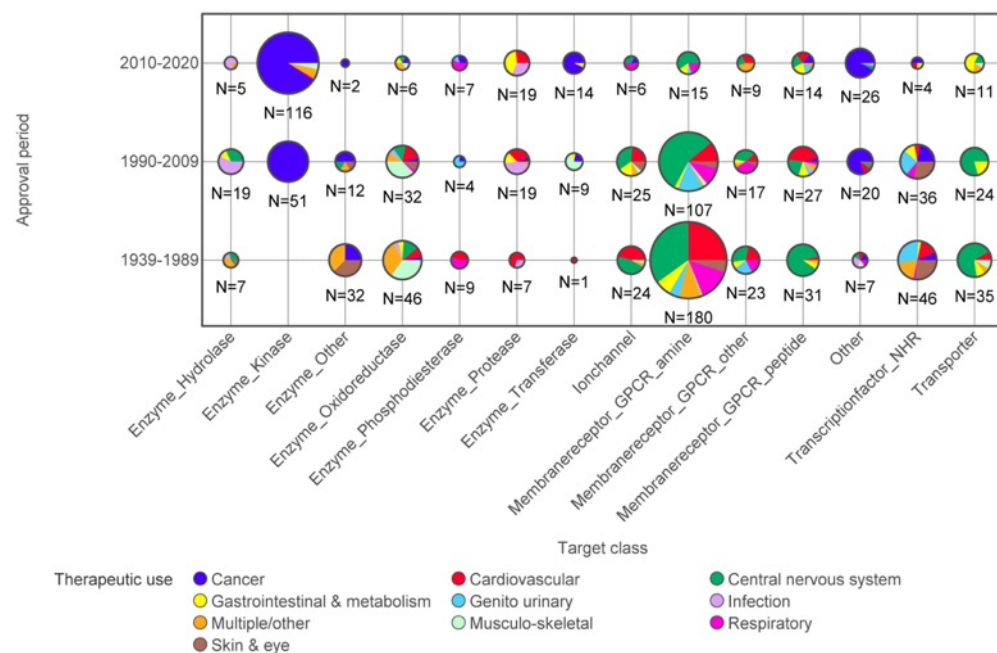
Small molecule			Antibody		
Clinical precedence	Discovery precedence	Predicted tractable	Clinical precedence	Predicted tractable - high confidence	Predicted tractable - medium to low confidence
Phase 4	Phase 2 or 3	Phase 0 or 1	PDB targets with ligands	Active compounds in ChEMBL	DrugEBLity score > 0.7
Phase 4	Phase 2 or 3	Phase 0 or 1	UniProt location - high confidence	GO cell component - high confidence	UniProt location - low or unknown confidence
			UniProt predicted signal peptide or transmembrane region	GO cell component - medium confidence	Human Protein Atlas - high confidence



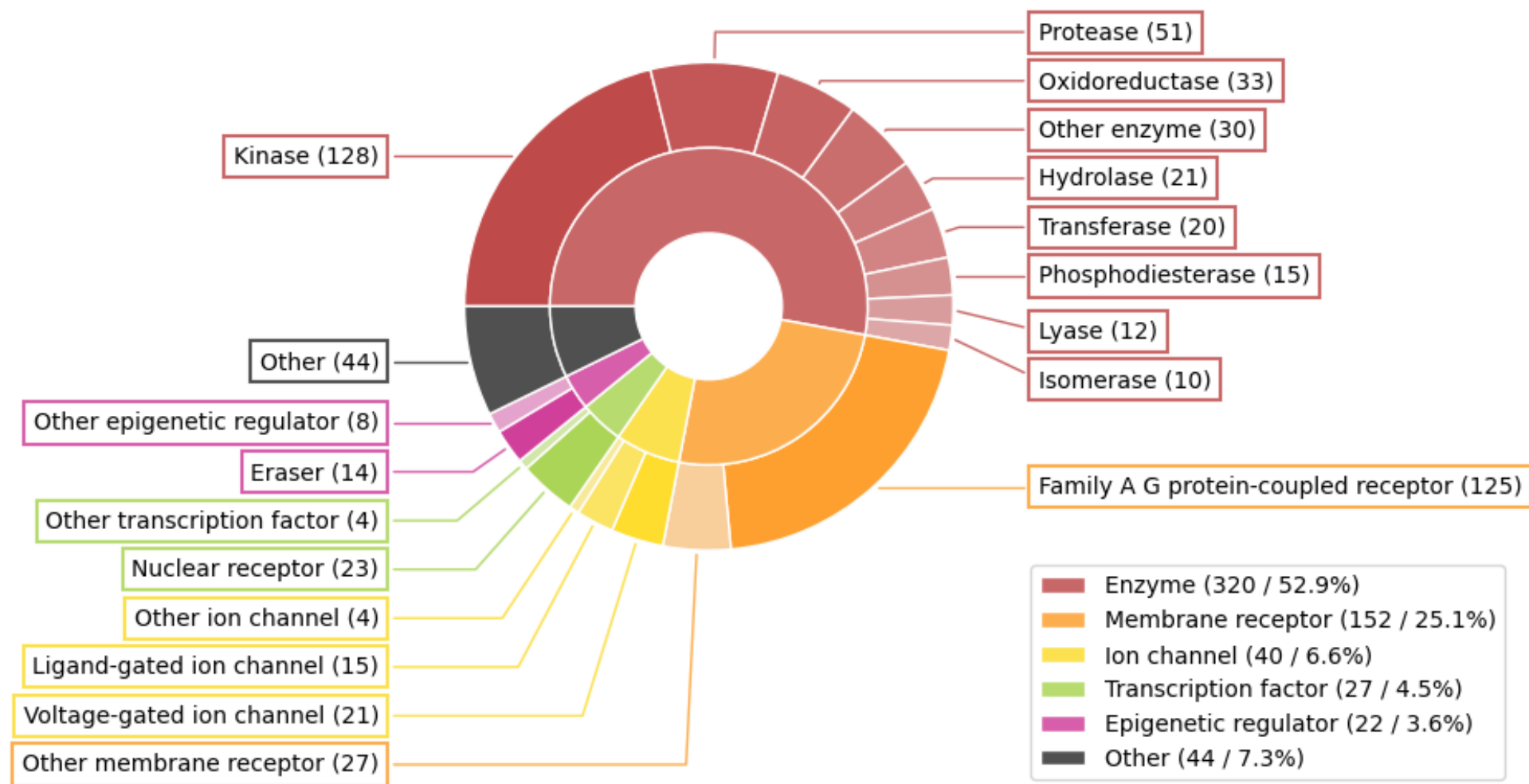
Weak disease association
No tractability evidence
5908

ChEMBL data differentiates drugs from “bioactive molecules”

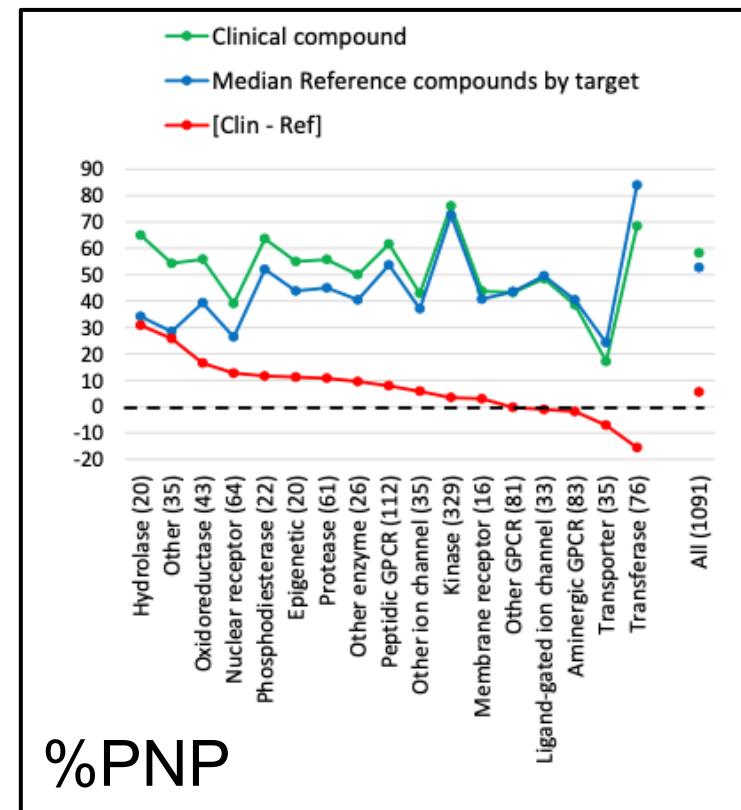
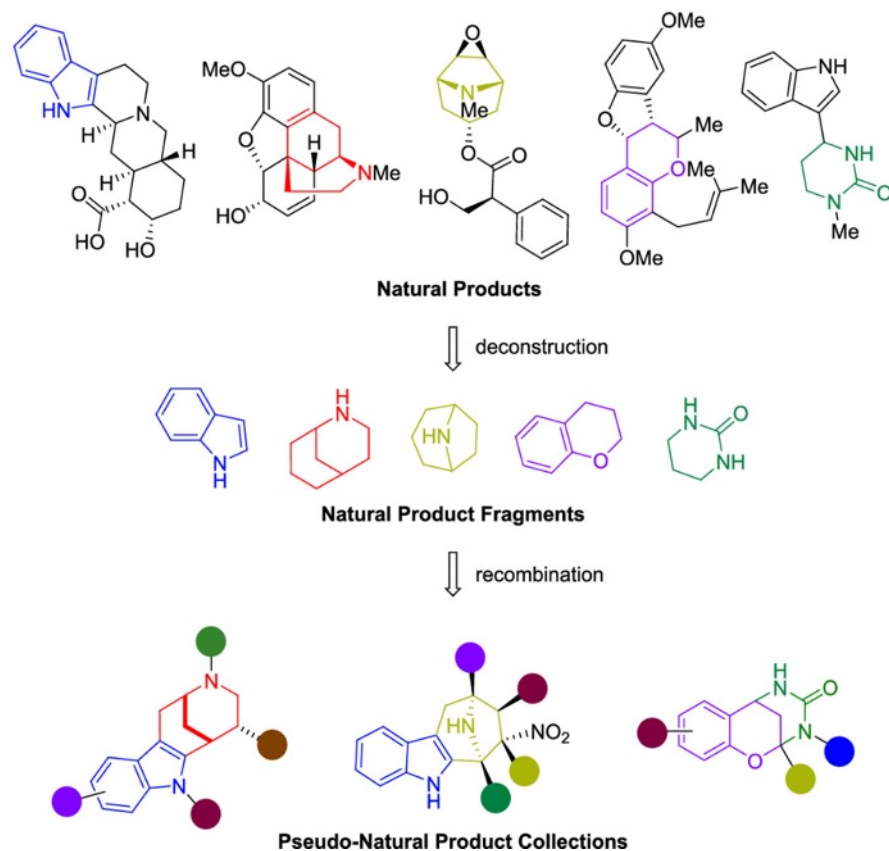
- 643 drugs against 271 targets
- 360K comparator compounds
- All with quantitative activity data and calculated properties
- Drugs continue to be more efficient than their comparator compounds



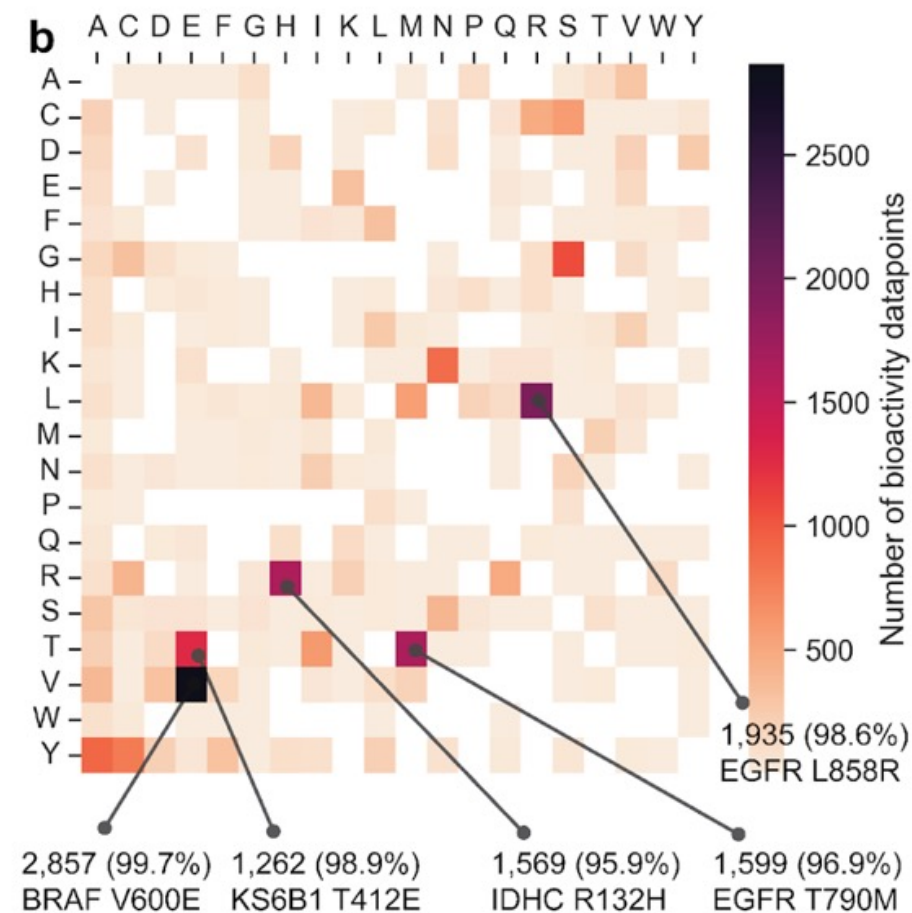
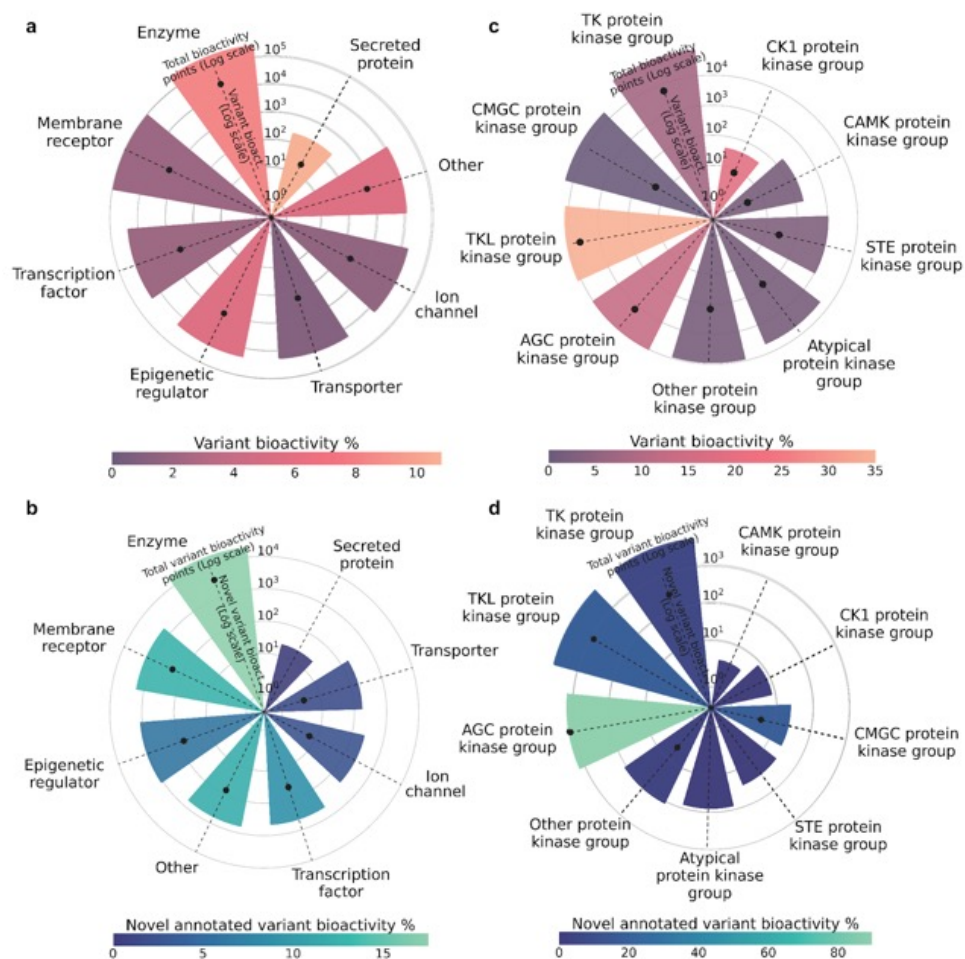
Robust & reliable datasets for community use



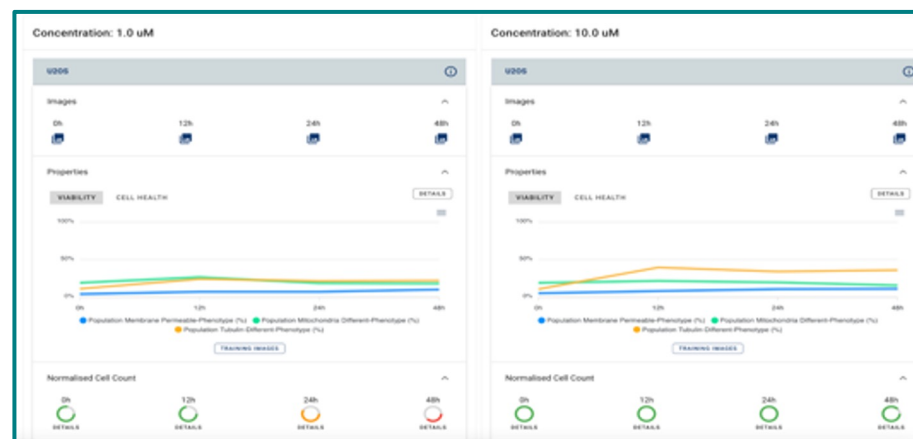
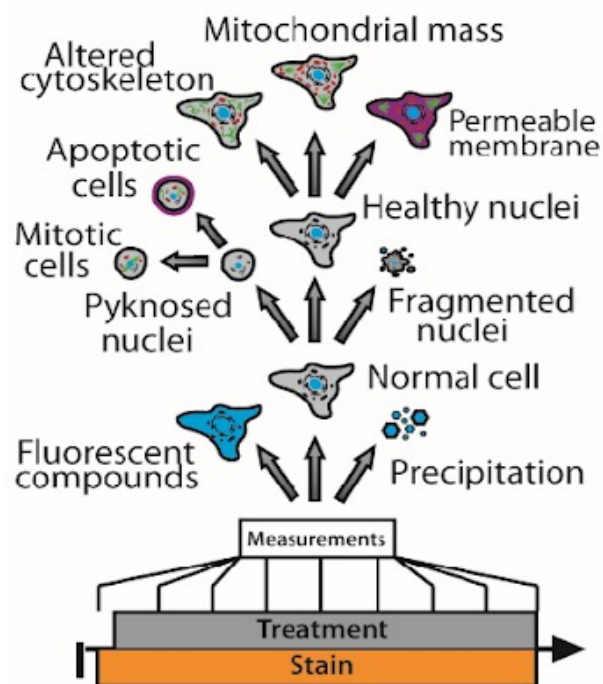
“Natural selection” in small molecule drug discovery



Protein variants & bioactivity data (Gerard's talk)



ChEMBL underpins multiple external collaborations (Brian's talk)



Key data in ChEMBL,
also visible via
EuBOPEN gateway



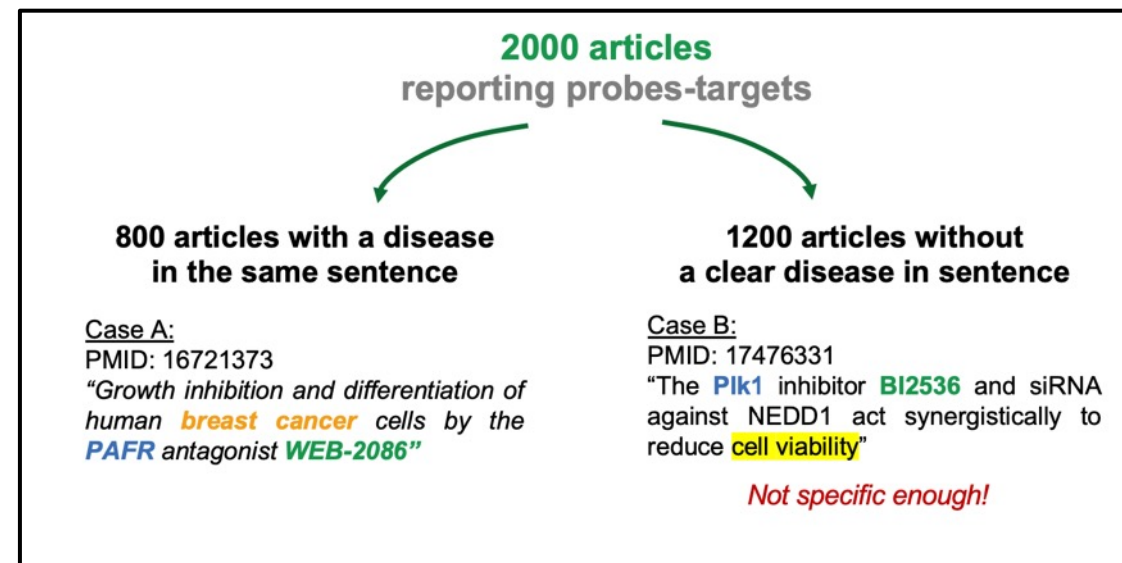
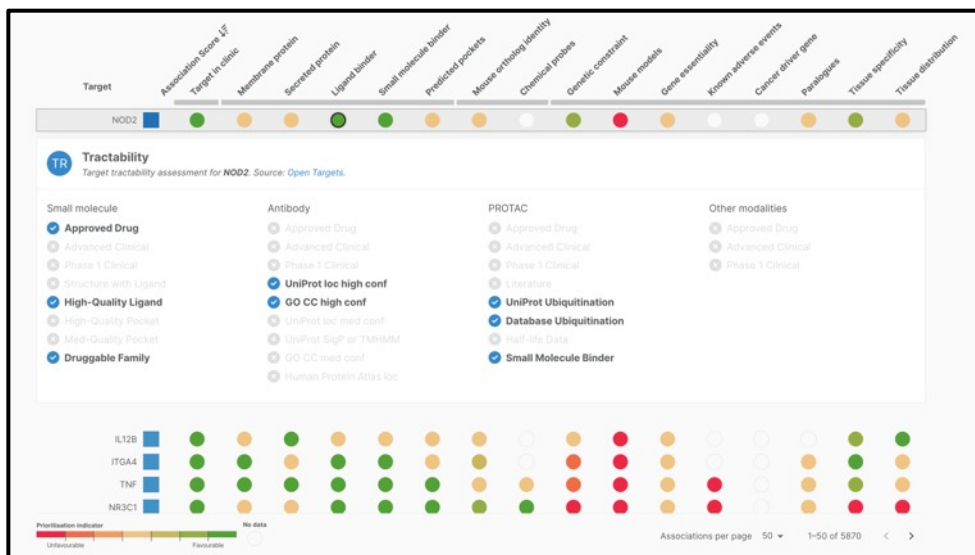
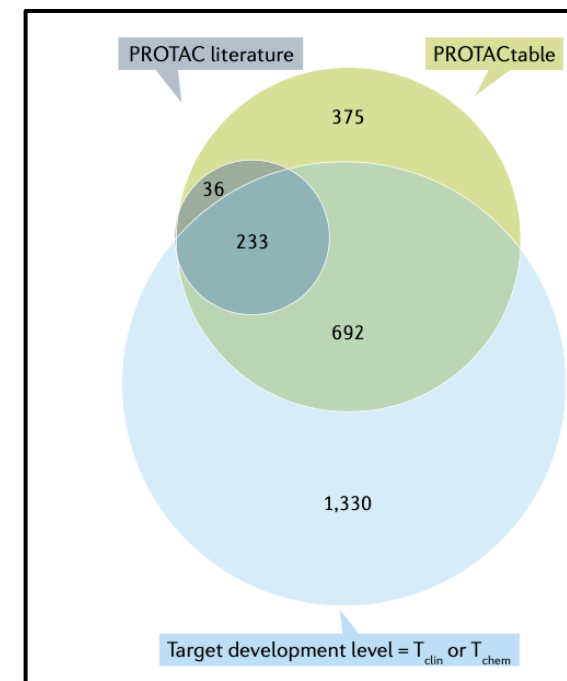
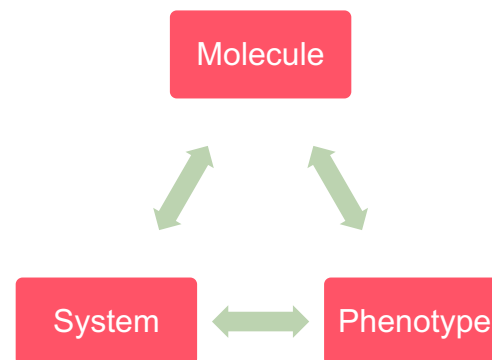
Plate images per time
point & replicate.
**Hosted by the BioImage
Archive.**

Open Targets is a key strategic partner

DW Drug Warnings
Manually curated withdrawn and black box warnings for **ROSIGLITAZONE**. Source: [ChEMBL](#)

Q Search

Warning type	Adverse event	ChEMBL warning class	Country / region
Withdrawn	cardiotoxicity?	cardiotoxicity	Armenia; Guatemala; Egypt
Withdrawn	edema?	cardiotoxicity	European Union
Withdrawn	heart failure?	cardiotoxicity	European Union
Black Box Warning	N/A	cardiotoxicity	United States



What will the next 15 years bring?



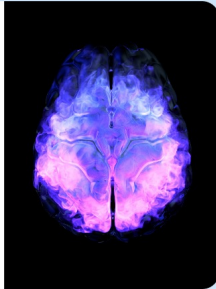
For more meaningful insights, see the talks, view the posters, meet the team!

My next chapter: LifeArc, a self-financing charity addressing unmet patient need

Our Translational Challenges

Motor Neuron Disease

Our vision is a world where motor neuron disease is preventable and treatable.

[Read more](#)

Chronic Respiratory Infection

Our vision is improved quality of life for those living with cystic fibrosis and bronchiectasis.

[Read more](#)

Global Health

Our vision is a world with affordable and accessible solutions to better understand, treat, and prevent infectious diseases.

[Read more](#)

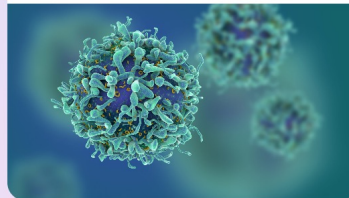
Rare Disease

Our vision is an ecosystem working together to get discoveries and treatments to rare disease patients faster.

[Read more](#)

Childhood Cancer

Our vision is to drive life-changing innovations for children with cancer.

[Read more](#)

16 May 2024

Britain's largest health research programme, Our Future Health, forms new £10m partnership with LifeArc

News releases



Chemical Biology resources: funders & collaborators



NIH



Above all, a big thank you to all colleagues past & present

