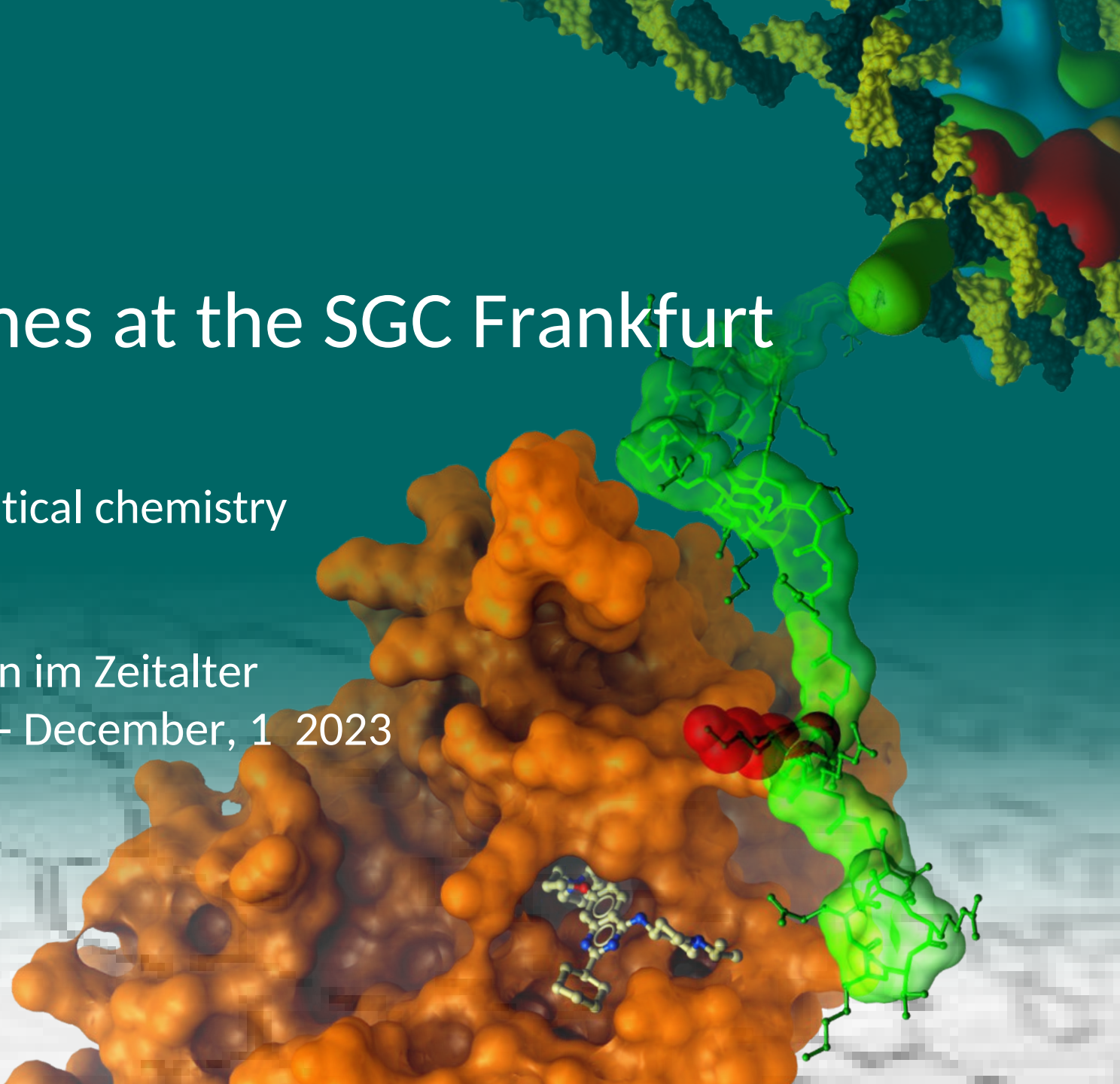
Open Science approaches at the SGC Frankfurt

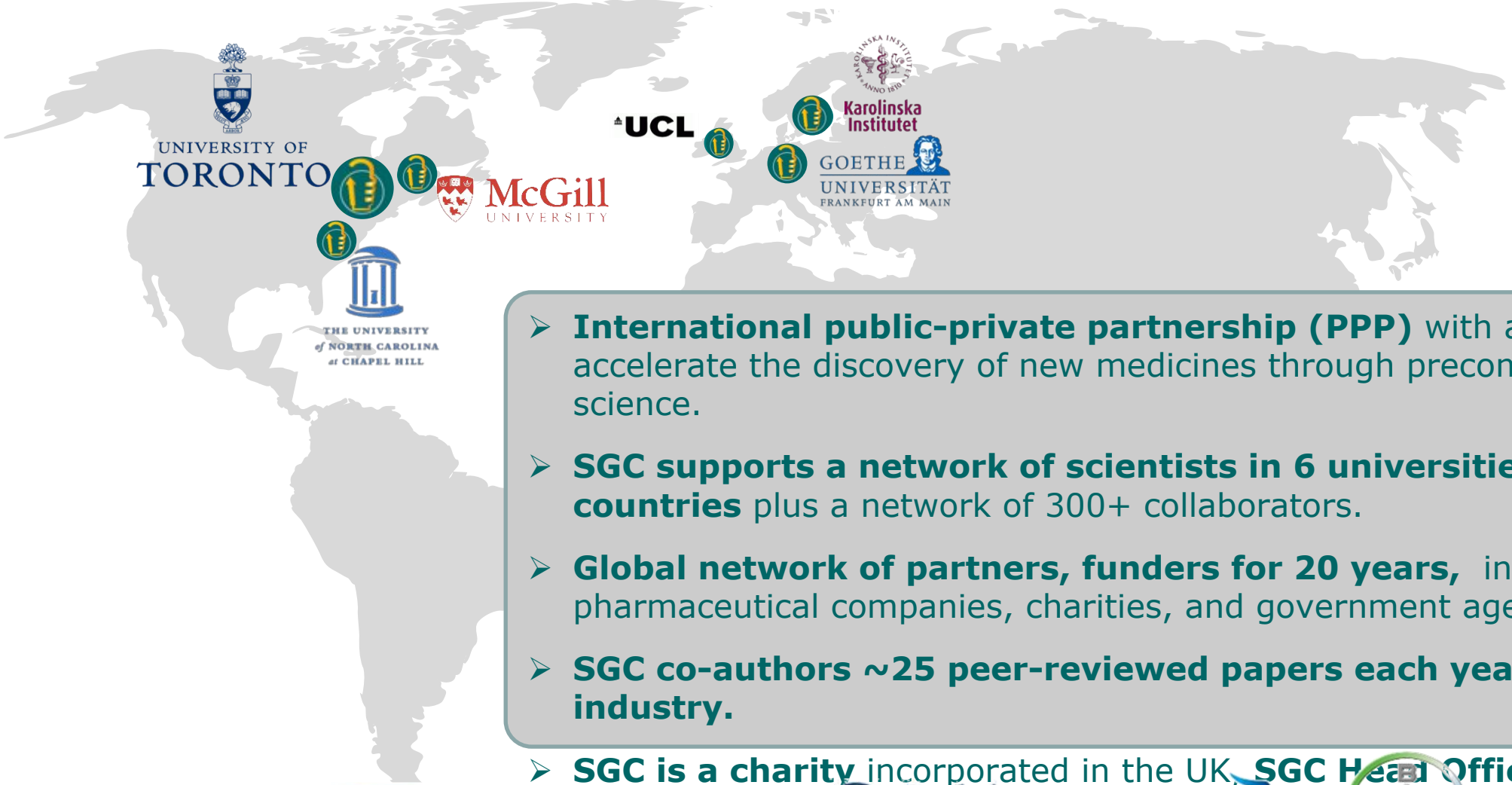
Dr. Susanne Müller-Knapp
SGC Frankfurt, Institut für pharmaceutical chemistry
Goethe University Frankfurt

Wissenschaftliche Publikationskulturen im Zeitalter von Open Access, Jena November, 30- December, 1 2023



THE STRUCTURAL GENOMICS CONSORTIUM (SGC) IS A GLOBAL PUBLIC PRIVATE PARTNERSHIP DEDICATED TO OPEN SCIENCE

SGC



- **International public-private partnership (PPP)** with a mission to accelerate the discovery of new medicines through precompetitive, open science.
- **SGC supports a network of scientists in 6 universities in 5 countries** plus a network of 300+ collaborators.
- **Global network of partners, funders for 20 years,** including pharmaceutical companies, charities, and government agencies.
- **SGC co-authors ~25 peer-reviewed papers each year with industry.**
- **SGC is a charity** incorporated in the UK, **SGC Head Office is in Canada.**

GenomeCanada · Innovative Medicines Initiative · Ontario · imi · BILL & MELINDA GATES foundation · THE ESHELMAN FOUNDATION WILMINGTON • NORTH CAROLINA · Boehringer Ingelheim · Genentech · janssen · Bristol Myers Squibb · BAYER · Pfizer · Takeda · MERCK

Our Ethos: Open Access

Promptly placing results, reagents and know-how in the public domain

We agree **not** to file for patent protection on any of our research outputs
(and encourage our collaborators to do the same)

OPEN ACCESS

www.thesgc.org

Why Do We Need SGC?

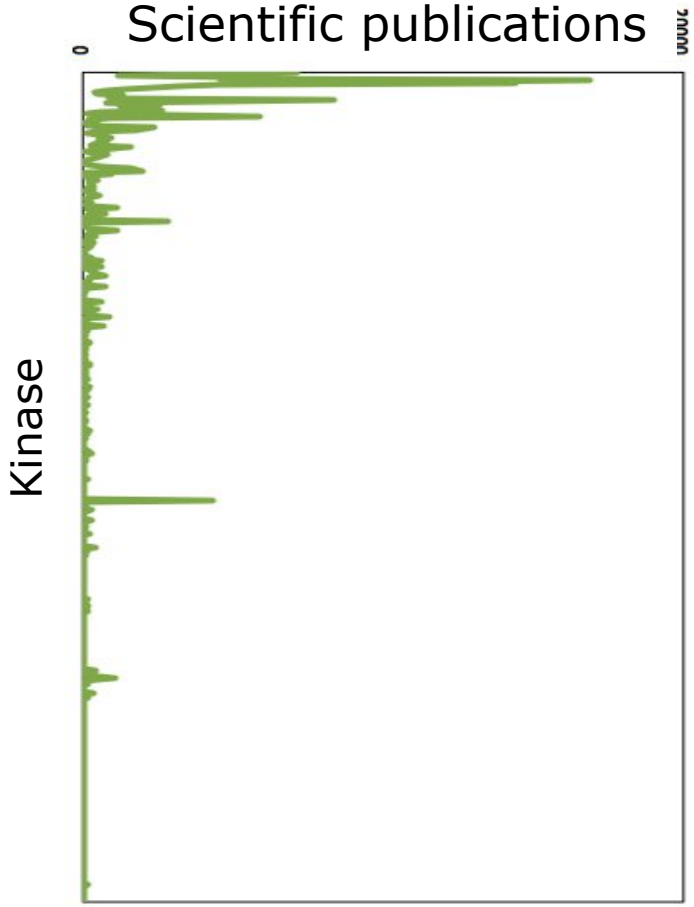
>$250B a year invested in biomedical research

- No new medicines for schizophrenia since 1950's
- No new treatment for Alzheimer's since early 80's

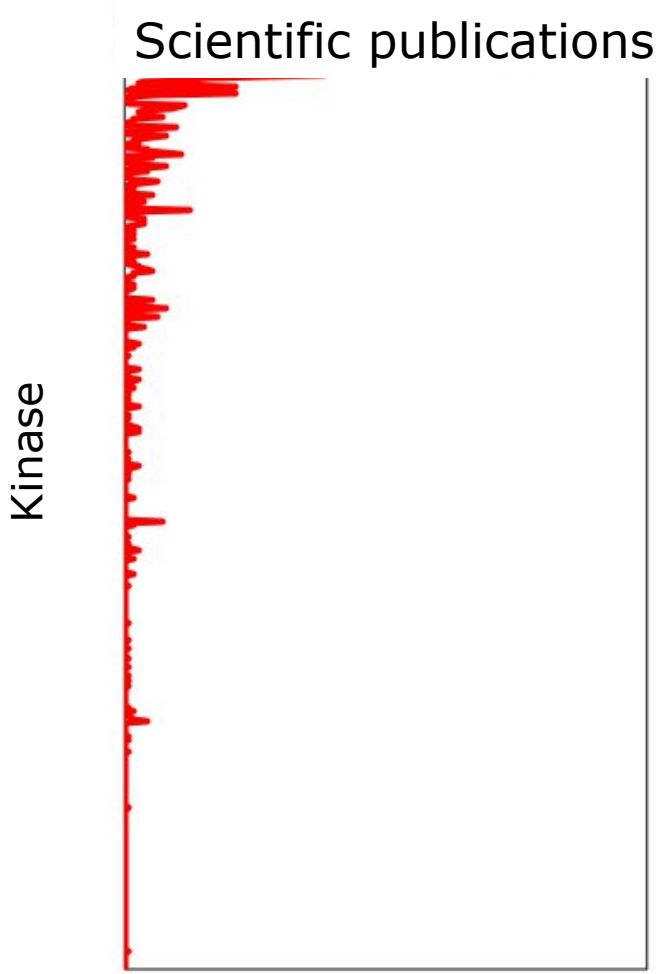
Medicines are not affordable for most people in the world

Most Science Is Redundant

Global Effort (2019)



German Effort (2019)



COMMENT

ANTHROPOLOGY Call for unity in the science of human beings p.160 | GENETICS Reviewed: two primers on personal genomics p.169 | POLICY Sanitation, not vaccination, is most important in Haiti p.175 | OBITUARY Jack Oliver, key player in proof of plate tectonics, remembered p.178

Too many roads not taken

Most protein research focuses on those known before the human genome was mapped. Work on the slew discovered since, urge **Aled M. Edwards** and his colleagues.

When a draft of the human genome was announced in 2000, funders, governments, industry and researchers made grand promises about how genome-based discoveries would revolutionize science. They promised that it would transform our understanding of human biology and disease, and provide new targets for drug discovery. Yet more than 75% of protein research still focuses on the 10% of proteins that were known before the genome was mapped — even though many more have been genetically linked to disease.

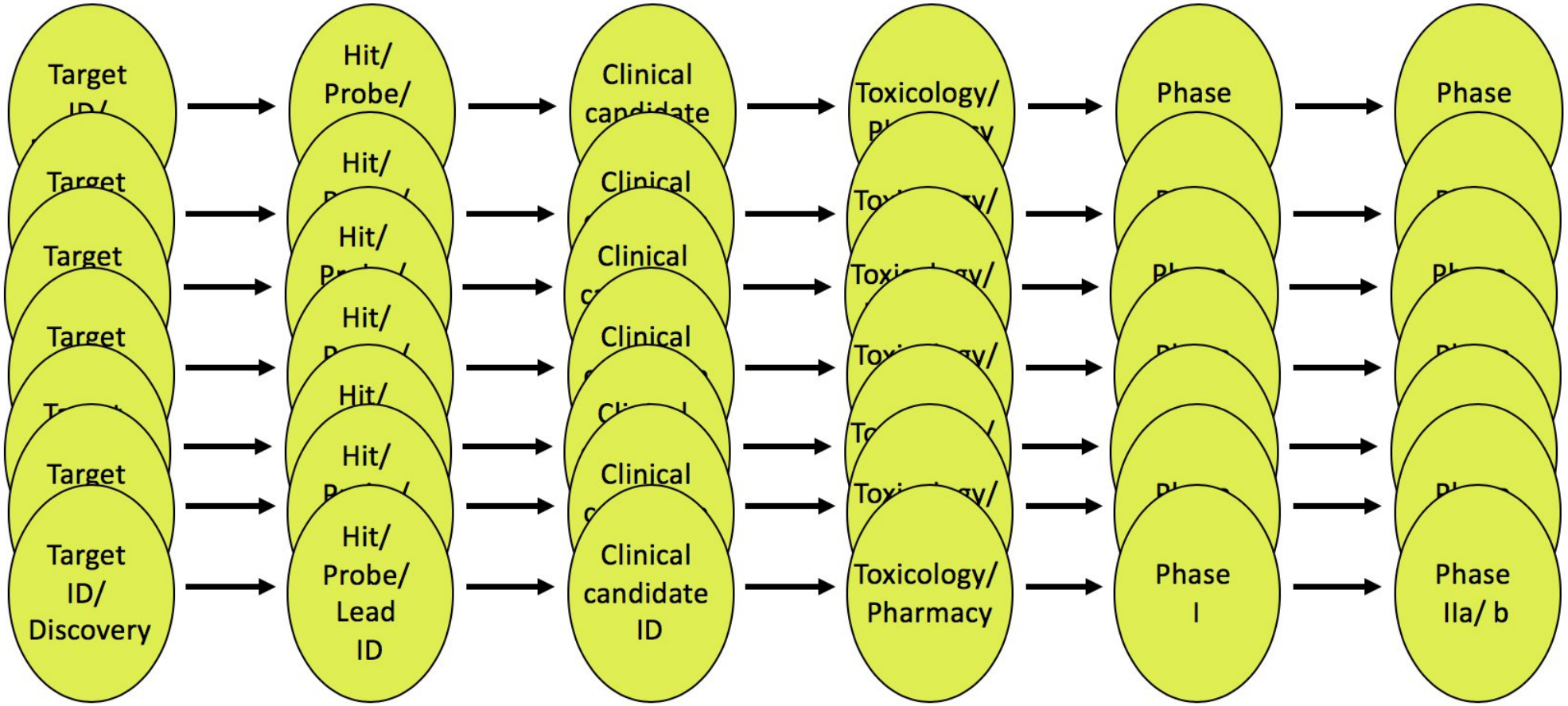
We performed a bibliometric analysis to assess how research activity has altered over time for three protein families that are central in disease and drug discovery: kinases, ion channels and nuclear receptors. For all three, we found very little change in the pattern of research activity — which proteins are associated with the highest number of publications — over the past 20 years[1]. Even those proteins that have been directly associated with disease remain 'hidden in plain sight,' with scientists proving very reluctant to study them.

Where there has been a shift in research activity, it was often spurred by the emergence of tools to study a particular protein, not by a change in the protein's perceived importance. We believe that ensuring high-quality tools are developed for all the proteins discovered may be all that is needed to drive research into the unstudied parts of the human genome — even within funding and peer-review systems that are inherently conservative.

We searched for mention of every human ▶

NATURE.COM Protein mapping gains a human focus: go.nature.com/vbqctf

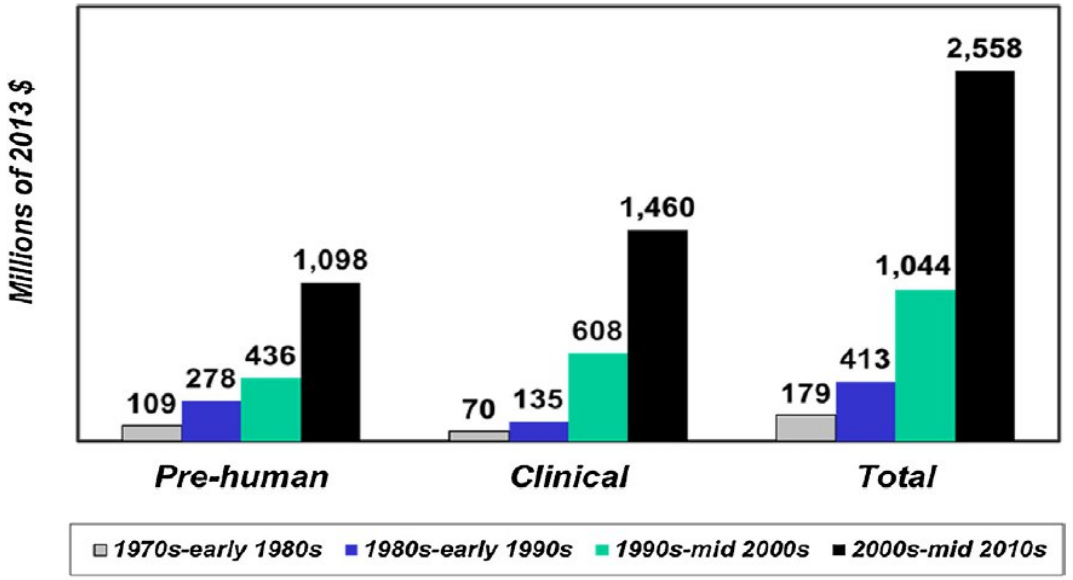
Siloed Proprietary Development Leads to Redundancy



Examples of failed parallel late-stage clinical programs: NK1 receptor antagonists for analgesia; matrix metalloproteases and farnesyltransferase inhibitors for cancer; cholesterol ester transfer protein for CVD; beta-amyloid for AD; aurora kinase inhibitors for breast cancer

What Does Industry Gain?

- New medicines are exorbitantly expensive
- >$2.5B in private sector R&D costs per approved new drug
- # new drugs approved per $1B halved every 9 years since 1950s
- clinical attrition rates



- Launches at $100,000's per patient per year the new normal
- Trends exacerbating pricing:
- Sustainability for public and private payers?

Precompetitive Research – from lab to patients?

SGC

2 Years and 2 months

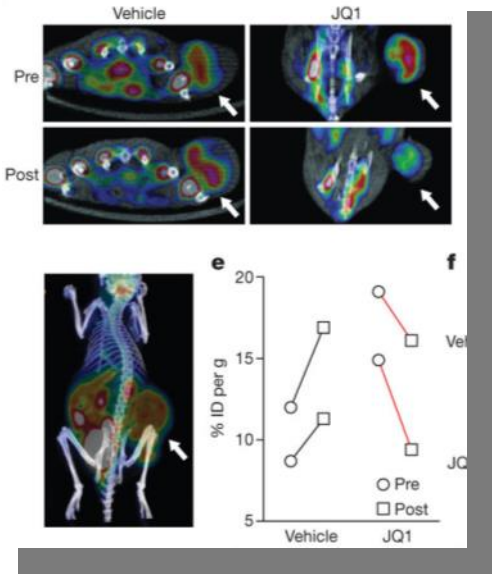
6 Months | 11 Months | 1 Month | 6 Months | 3 Months

July 2009

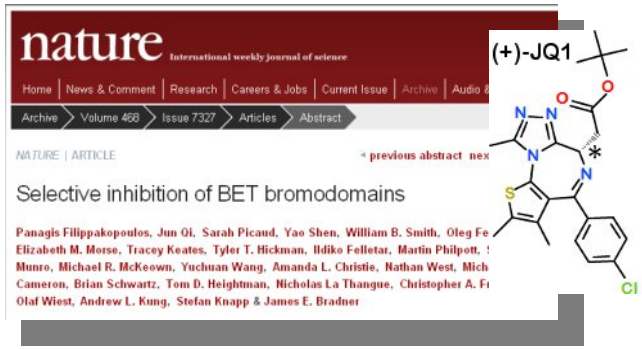
GSK collaboration starts

Jan 2010

SGC and Harvard start collaboration BET (NMC



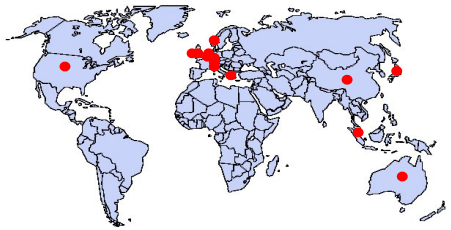
Dec 2010



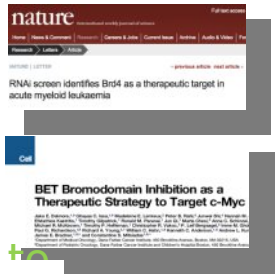
Co -publication of JQ1 probe (SGC; cancer) and I-BET probe (GSK; inflammation)

Jan 2011

JQ1 distributed to



July 2011

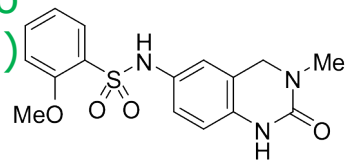


d4 linked to
1L (Nature)
1 (Cell)

Oct 2011

Booming interest in Academia and Industry

Pfizer BET probe announced



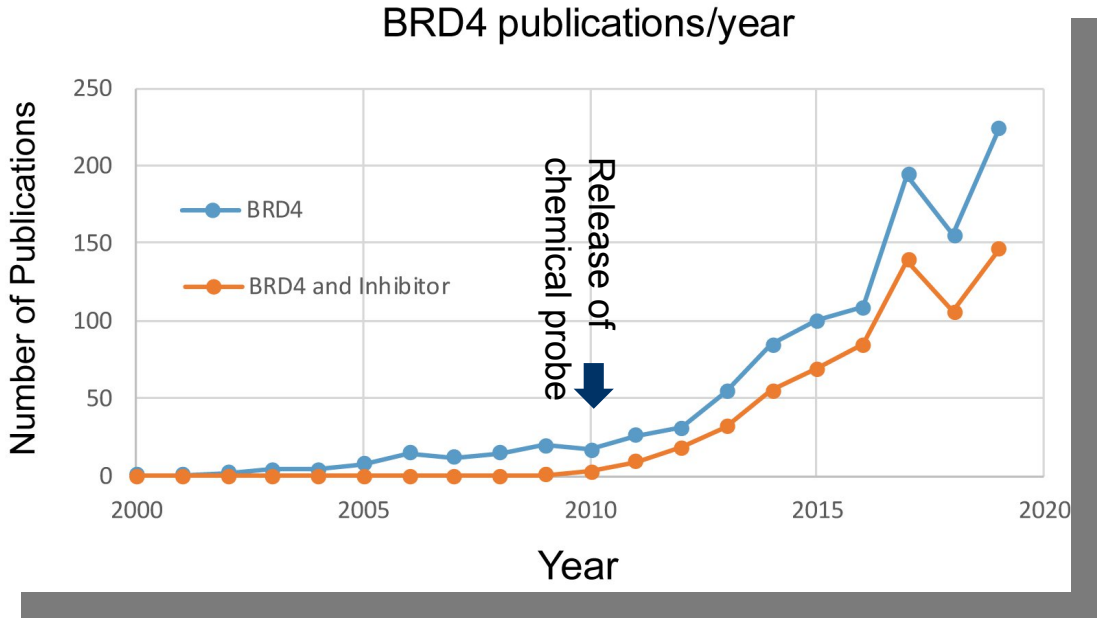
Mar 2012

GSK carries out first in man study (open)

Currently >30 clinical trials

OPEN SCIENCE AS SOLUTION

- Encourage innovation
- Engage industry
- Accelerate science
- Increase reproducibility
- Reduce redundancy
- Engage patients
- Mobilize funding
- Develop new technologies through crowed sourcing



Evolving science and partners to address pressing global needs

SGC

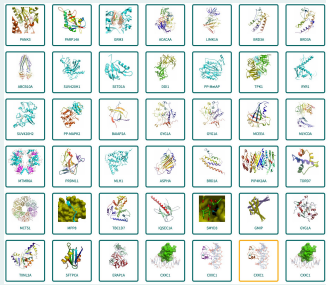
2004 | 2007 | 2011 | 2015 | 2020 | 2025

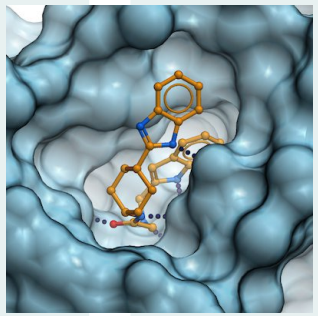
PHASE I | Phase II | Phase III | Phase IV | Phase V

SCIENTIFIC FOCUS

STRUCTURAL GENOMICS

Human Protein Families
Human Parasite Proteins





Bio-active Chemical Tools

CHEMICAL GENOMICS

TARGET 2035

DATA DRIVEN DD

Machine Learning

Patient-derived disease assays

EUbOPEN
Enabling & Unlocking Biology in the OPEN

INDUSTRY PARTNERS

GlaxoSmithKline

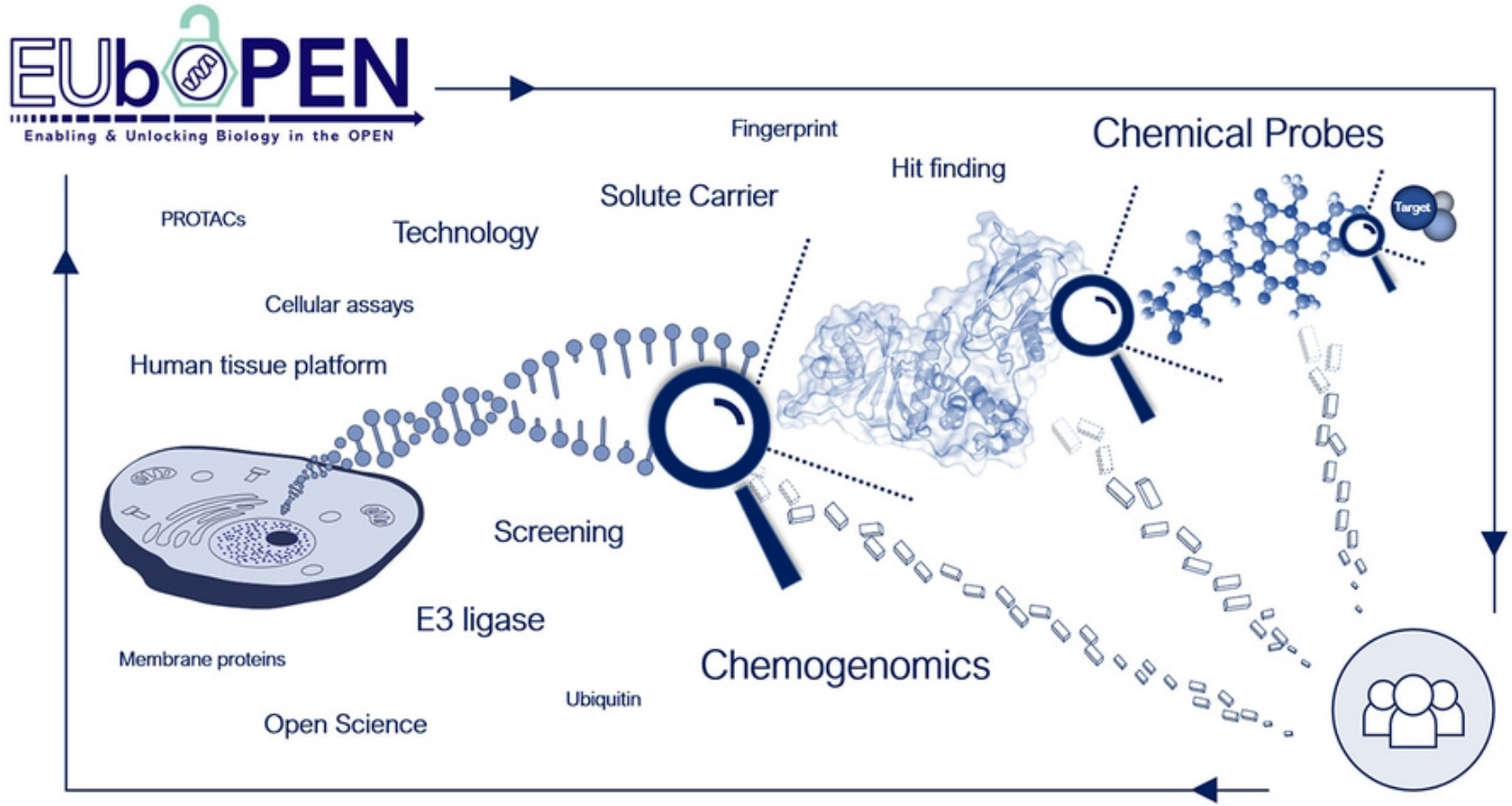
GlaxoSmithKline, Pfizer, MERCK, NOVARTIS

GlaxoSmithKline, Takeda, Pfizer, janssen, Lilly, abbvie, Boehringer Ingelheim, BAYER, NOVARTIS

Takeda, Boehringer Ingelheim, janssen, MERCK, Pfizer, abbvie, BAYER, MSD, NOVARTIS

Genentech, janssen, MERCK, Pfizer, Boehringer Ingelheim, Takeda, BAYER, Bristol Myers Squibb

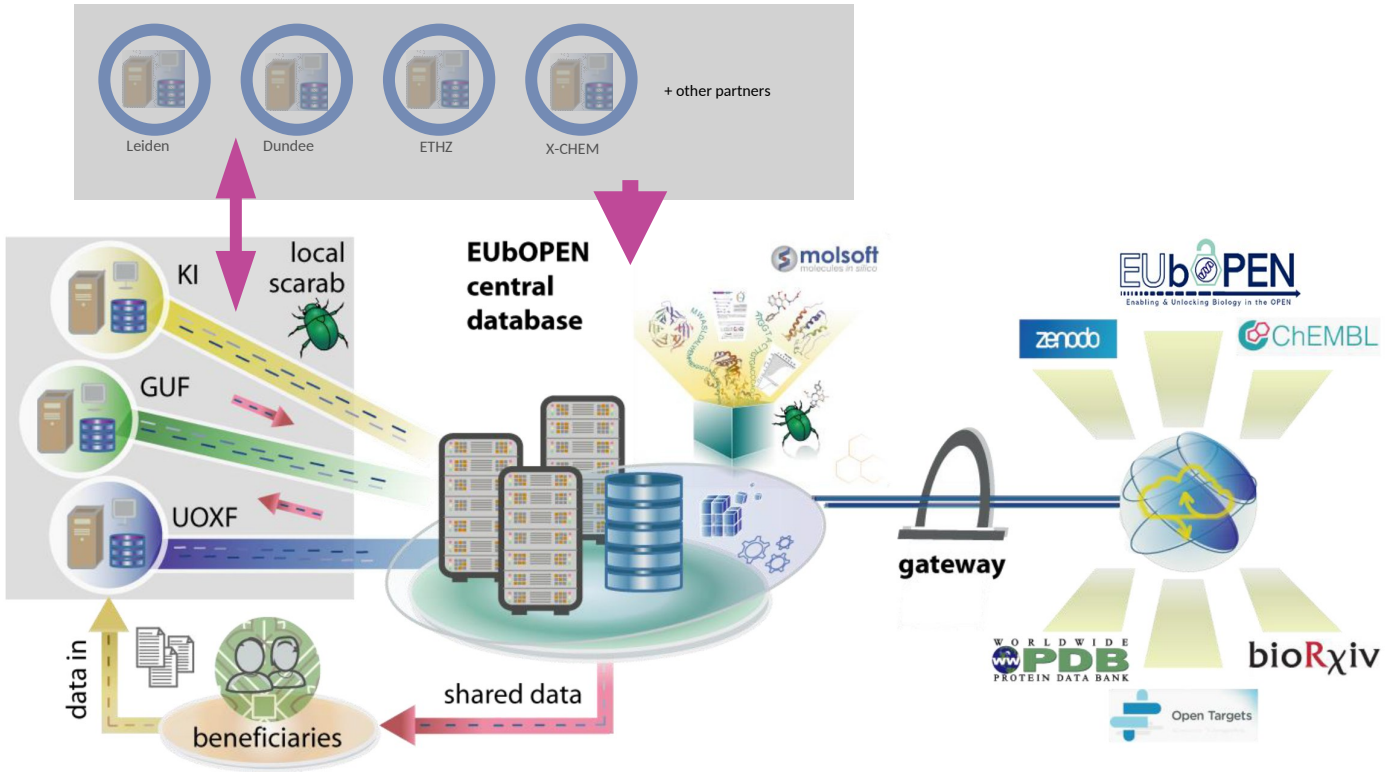
EUbOPEN



- The EUbOPEN consortium is an **Innovative Medicines Initiative** (IMI) funded project to **enable and unlock biology in the open**.
- **22 partners** from academia and industry; five years (2020-2025)
- **Total budget of 65.8 million euros** covered by a grant from the IMI and cash and in-kind contributions from the EFPIA companies, IMI Associated Partners and non-EU partners.
- **Open Science Policy**

GOETHE UNIVERSITÄT FRANKFURT AM MAIN

DATA ACCESSIBILITY



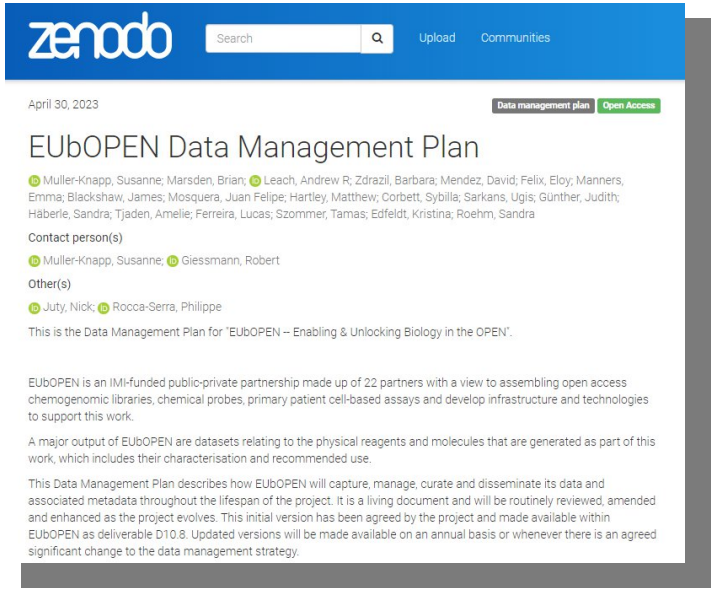
- Open access requirement for **all** EUbOPEN and SGC publications
- Depositing publications in repositories
- Subject-based/thematic repository (e.g., arXiv, Europe PMC), **OR**
- Zenodo the OpenAIRE repository hosted by CERN
- FAIR Guiding Principles
  Findable, Accessible, Interoperable, Reusable

GOETHE UNIVERSITÄT FRANKFURT AM MAIN

Data Repositories

Development and update of Data Management Plan

- DMP has been established and updated on annual basis
- Deposition of DMP in Zenodo



Providing Open Access Research data in sustainable repositories e.g.,

- Protein structures in PDB (315)
- Proteomics sets in Pride (15)
- Images in BioImage Archive (~1.4M images)
- Compounds to ChEMBL and EUbOPEN Gateway

RCSB PDB PROTEIN DATA BANK

PRIDE PRoteomics IDEntifications Database

BioImage Archive

ChEMBL

In addition to data

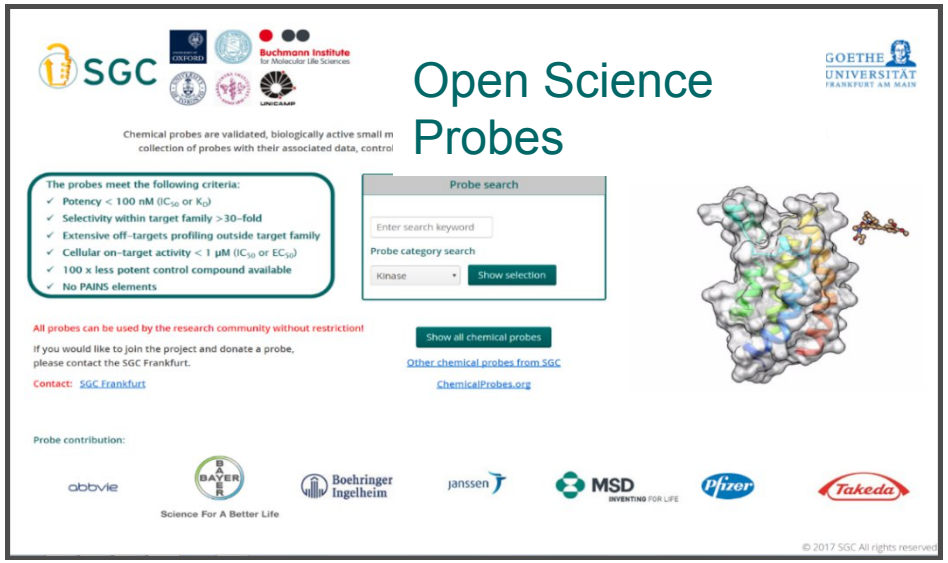
- Plasmids in Addgene

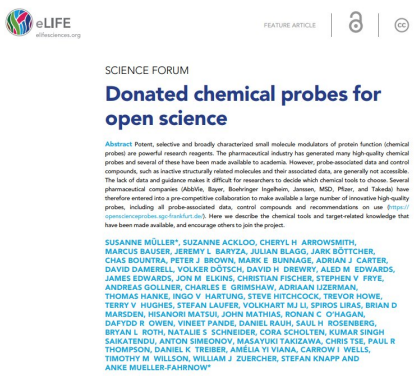
addgene

BEYOND DATA...

After publication of data, 70% of reagents remainnot accessible

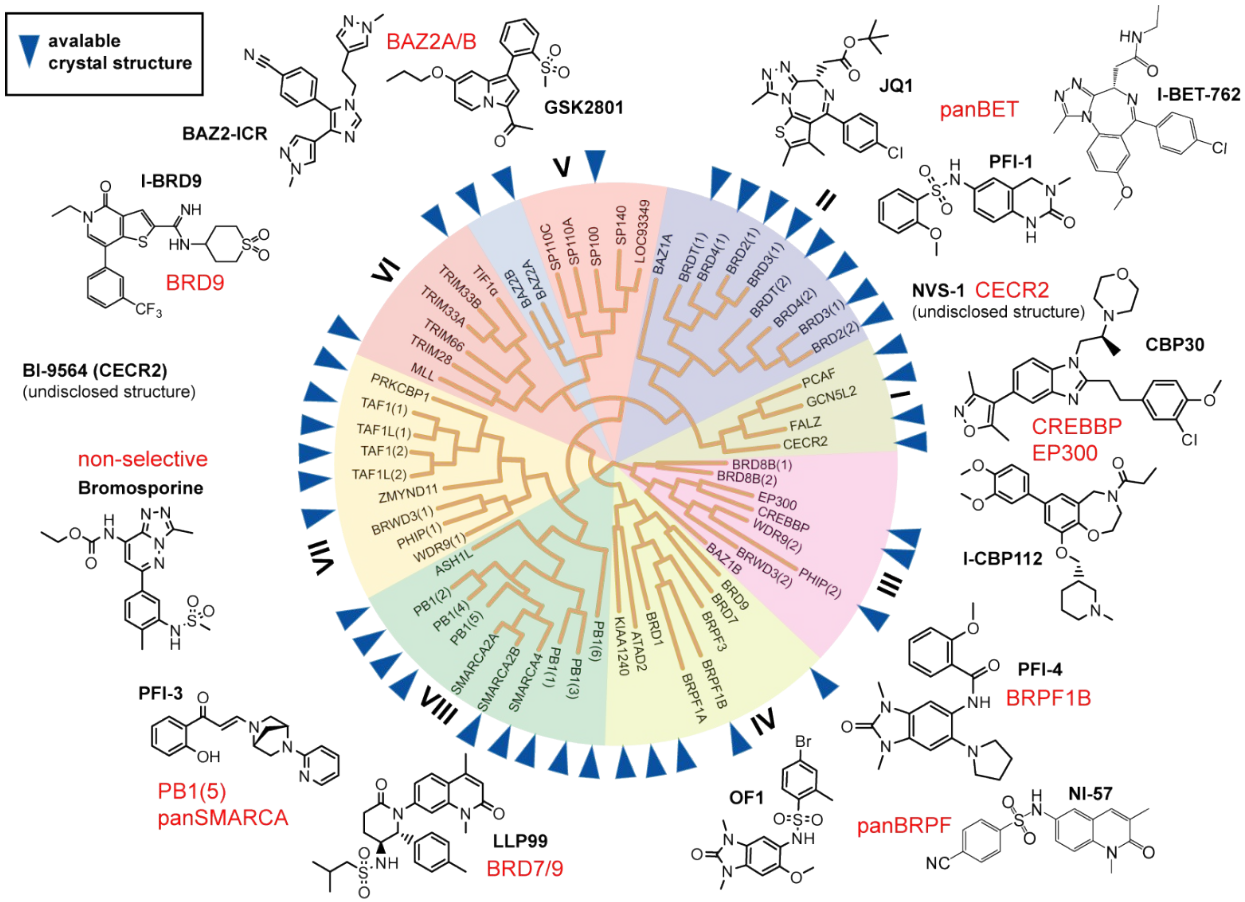
- Reproducibility of results?
- Quality control?





Open Science Probes

http://www.sgc-ffm.uni-frankfurt.de/



https://www.thesgc.org/chemical-probes

SGC CHEMICAL PROBES

- Lack of tools, which help understand biology and disease-relevant processes
- In a bibliometric analysis, we found that chemical probes were the most impactful tools to enable researchers to work on new genomics targets
- Freely available probes to human proteins will enable discovery of new medicines
- Collaboration with vendors for sustainability

Chemical probe = a drug-like small molecule that selectively modulates the activity of a specific protein in cells

IMPACT OF SGC CHEMICAL PROBES

DISCOVERED

190+

Chemical probes discovered by SGC or with pharma or academics

DISTRIBUTED

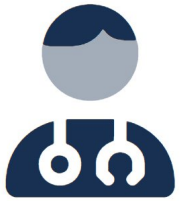
42,662+

Samples of chemical probes distributed globally by SGC and trusted vendors

CITATIONS

7,285+

SGC chemical probes used by scientists around the world



CLINICAL TRIALS

25+

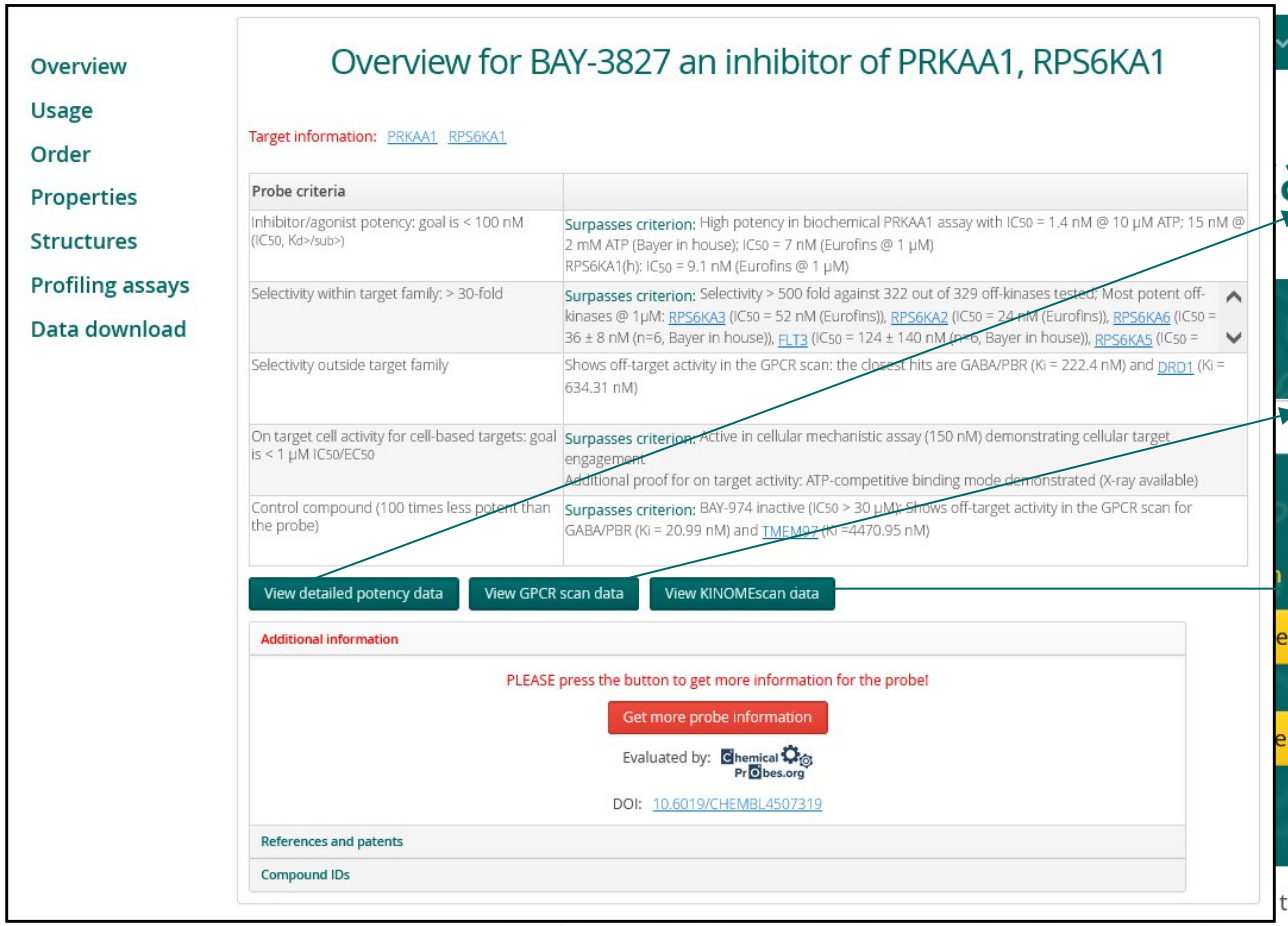
Clinical trials and late-stage preclinical programs based on therapeutic hypotheses generated with SGC chemical probes

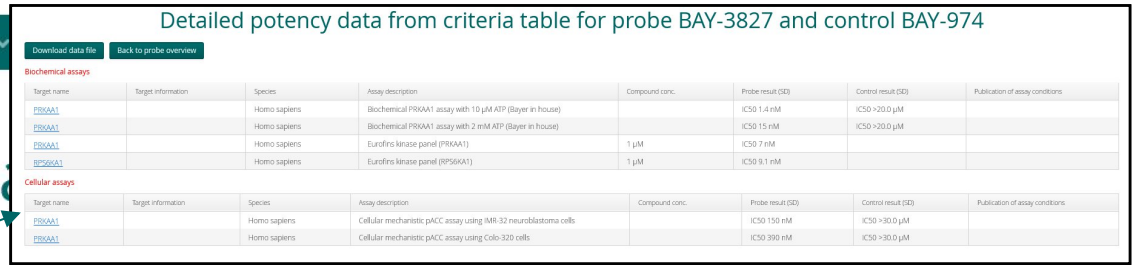
HUNDREDS OF PAPERS USING SGC PROBES RESULTING IN THERAPEUTIC HYPOTHESES

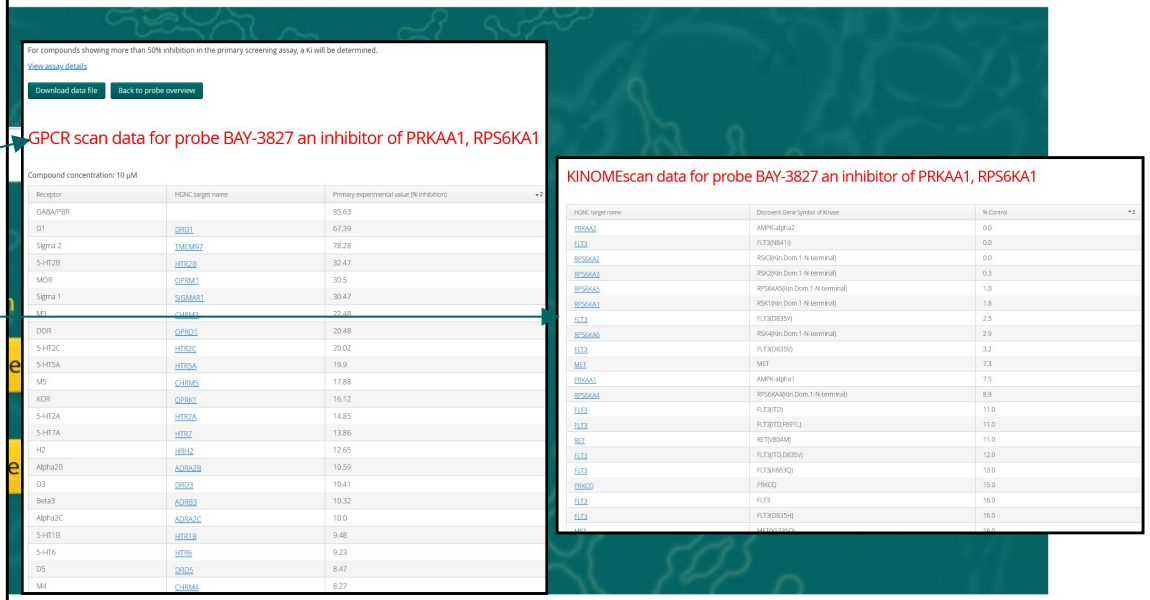
Database - Website

- All data in a publicly available database
- Order whole set, subsets or single probe pairs

https://www.sgc-ffm.uni-frankfurt.de/







their target protein(s). This website provides a unique collection of probes with their associated data, control compounds and recommendations on their use as well as a way to order the molecules.

Sharing data and Protocols via Website

EUbOPEN
Enabling & Unlocking Biology in the OPEN

efpia imi

About | Collaborate | Data Gateway | Protocols | Chemical Tools | Chemogenomics | Antibodies | Tissue Assays | News & Events | Publications

Tissue Assays

Within EUbOPEN, our aim is to develop open access cell assay protocols and data from well-characterised human disease tissue and blood-d
profile chemical probes and chemogenomic libraries, guiding identification of biomarkers and novel targets to drive drug discovery. To this er
human tissue and blood-derived assays within areas of high medical need in inflammatory diseases, fibrosis, oncology and neurodegenerativ
current network of collaborations consists of hospitals, research institutes and universities in Toronto and Montreal (Canada), Frankfurt (Ger
Stockholm (Sweden).

EUbOPEN Tissue Assay data sets:

Hepatotoxicity screening in primary human hepatocyte spheroid cultures

Viability screening in matched tumor and normal organoids from colorectal cancer patients

Therapy re-sensitization of drug resistant organoids

Toxicity screen in isolated B cell assay

Cytokine secretion screen in isolated B cell assay

NASH model to assess anti-steatosis effects of chemical probes

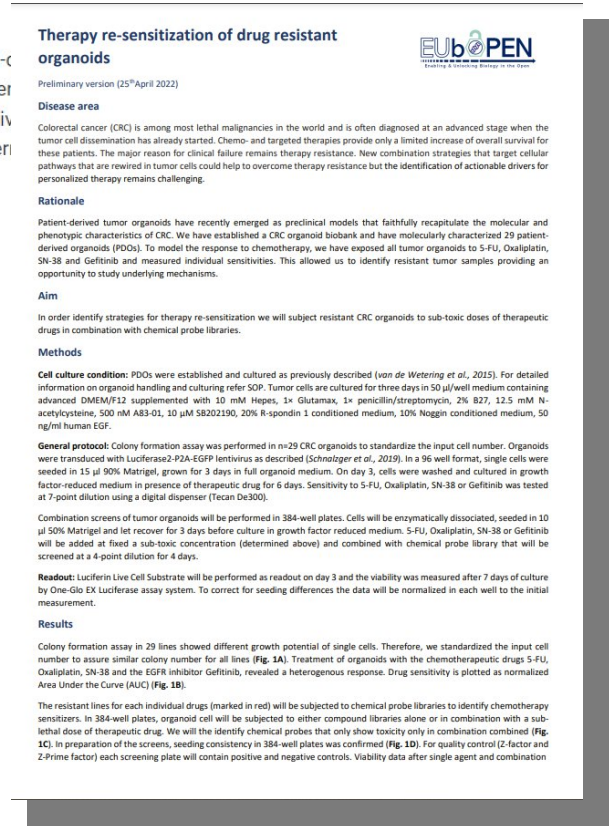
NASH model to assess fibrosis effects of chemical probes

NASH model to assess inflammatory effects of chemical probes

Targeting of CAF-induced therapy resistance in CRC organoid-stroma co-cultures

Identification of compartment-specific drug sensitivities in CRC organoid-stroma co-cultures

Protocols and results



Therapy re-sensitization of drug resistant organoids

EUbOPEN

Preliminary version (25th April 2022)

Disease area

Colorectal cancer (CRC) is among most lethal malignancies in the world and is often diagnosed at an advanced stage when the tumor cell dissemination has already started. Chemo- and targeted therapies provide only a limited increase of overall survival for these patients. The major reason for clinical failure remains therapy resistance. New combination strategies that target cellular pathways that are rewired in tumor cells could help to overcome therapy resistance but the identification of actionable drivers for personalized therapy remains challenging.

Rationale

Patient-derived tumor organoids have recently emerged as preclinical models that faithfully recapitulate the molecular and phenotypic characteristics of CRC. We have established a CRC organoid biobank and have molecularly characterized 29 patient-derived organoids (PDOs). To model the response to chemotherapy, we have exposed all tumor organoids to 5-FU, Oxaliplatin, SN-38 and Gefitinib and measured individual sensitivities. This allowed us to identify resistant tumor samples providing an opportunity to study underlying mechanisms.

Aim

In order identify strategies for therapy re-sensitization we will subject resistant CRC organoids to sub-toxic doses of therapeutic drugs in combination with chemical probe libraries.

Methods

Cell culture condition: PDOs were established and cultured as previously described (*van de Wetering et al., 2015*). For detailed information on organoid handling and culturing refer SOP. Tumor cells are cultured for three days in 50 µl/well medium containing advanced DMEM/F12 supplemented with 10 mM Hepes, 1× Glutamax, 1× penicillin/streptomycin, 2% B27, 12.5 mM N-acetylcysteine, 500 nM A83-01, 10 µM SB202190, 20% R-spondin 1 conditioned medium, 10% Noggin conditioned medium, 50 ng/ml human EGF.

General protocol: Colony formation assay was performed in n=29 CRC organoids to standardize the input cell number. Organoids were transduced with Luciferase2-P2A-EGFP lentivirus as described (*Schnalzger et al., 2019*). In a 96 well format, single cells were seeded in 15 µl 90% Matrigel, grown for 3 days in full organoid medium. On day 3, cells were washed and cultured in growth factor-reduced medium in presence of therapeutic drug for 6 days. Sensitivity to 5-FU, Oxaliplatin, SN-38 or Gefitinib was tested at 7-point dilution using a digital dispenser (Tecan De300).

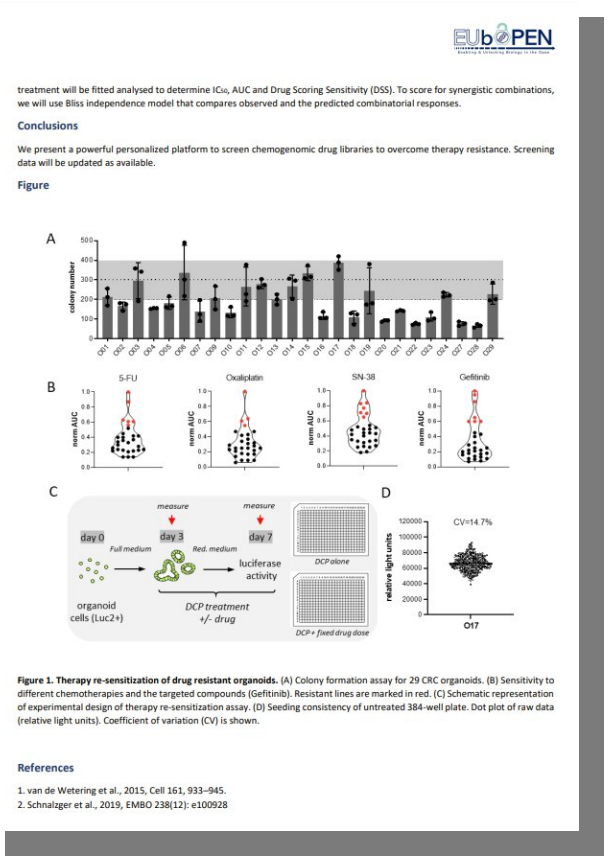
Combination screens of tumor organoids will be performed in 384-well plates. Cells will be enzymatically dissociated, seeded in 10 µl 50% Matrigel and let recover for 3 days before culture in growth factor reduced medium. 5-FU, Oxaliplatin, SN-38 or Gefitinib will be added at fixed a sub-toxic concentration (determined above) and combined with chemical probe library that will be screened at a 4-point dilution for 4 days.

Readout: Luciferin Live Cell Substrate will be performed as readout on day 3 and the viability was measured after 7 days of culture by One-Glo EX Luciferase assay system. To correct for seeding differences the data will be normalized in each well to the initial measurement.

Results

Colony formation assay in 29 lines showed different growth potential of single cells. Therefore, we standardized the input cell number to assure similar colony number for all lines (**Fig. 1A**). Treatment of organoids with the chemotherapeutic drugs 5-FU, Oxaliplatin, SN-38 and the EGFR inhibitor Gefitinib, revealed a heterogenous response. Drug sensitivity is plotted as normalized Area Under the Curve (AUC) (**Fig. 1B**).

The resistant lines for each individual drugs (marked in red) will be subjected to chemical probe libraries to identify chemotherapy sensitizers. In 384-well plates, organoid cell will be subjected to either compound libraries alone or in combination with a sub-lethal dose of therapeutic drug. We will the identify chemical probes that only show toxicity only in combination combined (**Fig. 1C**). In preparation of the screens, seeding consistency in 384-well plates was confirmed (**Fig. 1D**). For quality control (Z-factor and Z-Prime factor) each screening plate will contain positive and negative controls. Viability data after single agent and combination



EUbOPEN

treatment will be fitted analysed to determine IC50, AUC and Drug Scoring Sensitivity (DSS). To score for synergistic combinations, we will use Bliss independence model that compares observed and the predicted combinatorial responses.

Conclusions

We present a powerful personalized platform to screen chemogenomic drug libraries to overcome therapy resistance. Screening data will be updated as available.

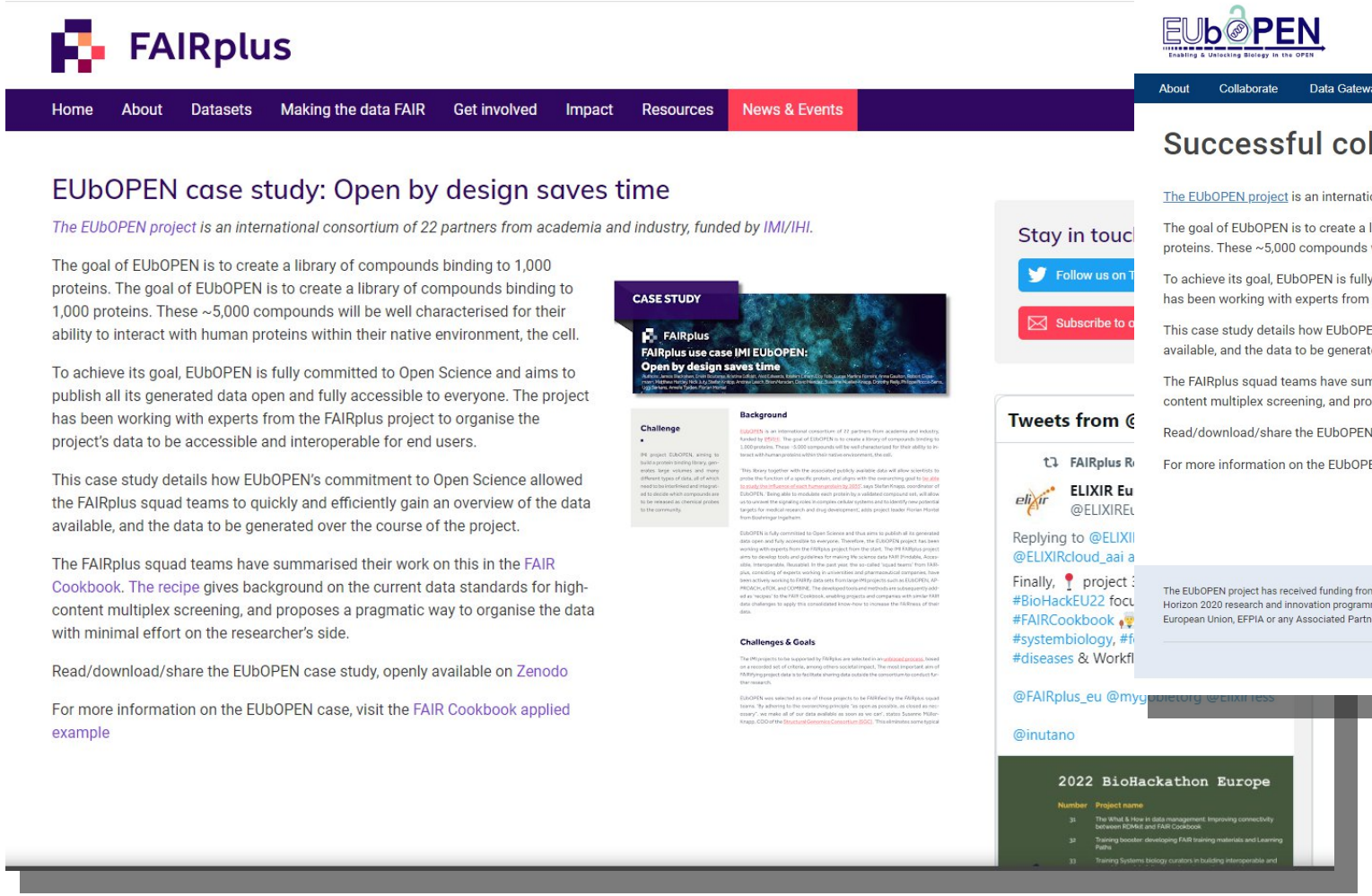
Figure

Figure 1. Therapy re-sensitization of drug resistant organoids. (A) Colony formation assay for 29 CRC organoids. (B) Sensitivity to different chemotherapies and the targeted compounds (Gefitinib). Resistant lines are marked in red. (C) Schematic representation of experimental design of therapy re-sensitization assay. (D) Seeding consistency of untreated 384-well plate. Dot plot of raw data (relative light units). Coefficient of variation (CV) is shown.

References

1. van de Wetering et al., 2015, Cell 161, 933–945.
2. Schnalzger et al., 2019, EMBO 238(12): e100928

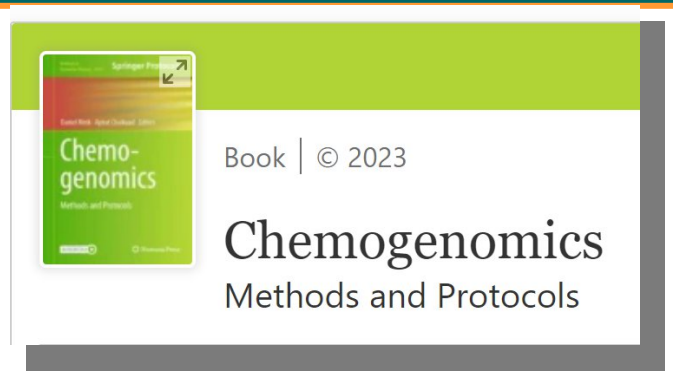
EUbOPEN - FAIRPlus Collaboration enabled data upload of Metadata to BioImage archive



Availability of metadata

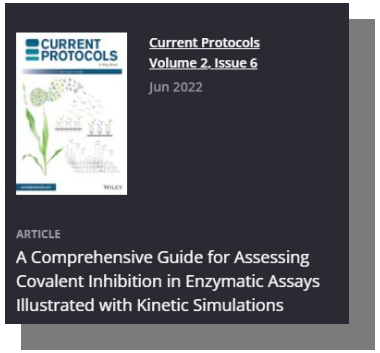
TB of data

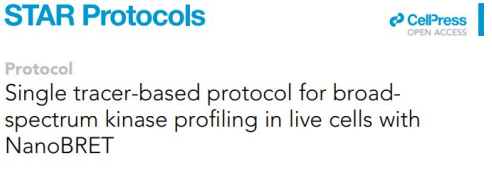
LONG-LIVED ASSAY PLATFORMS

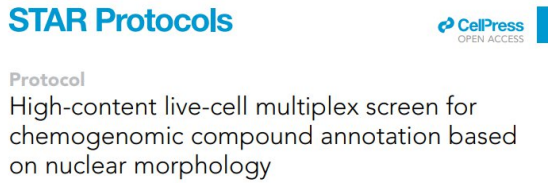


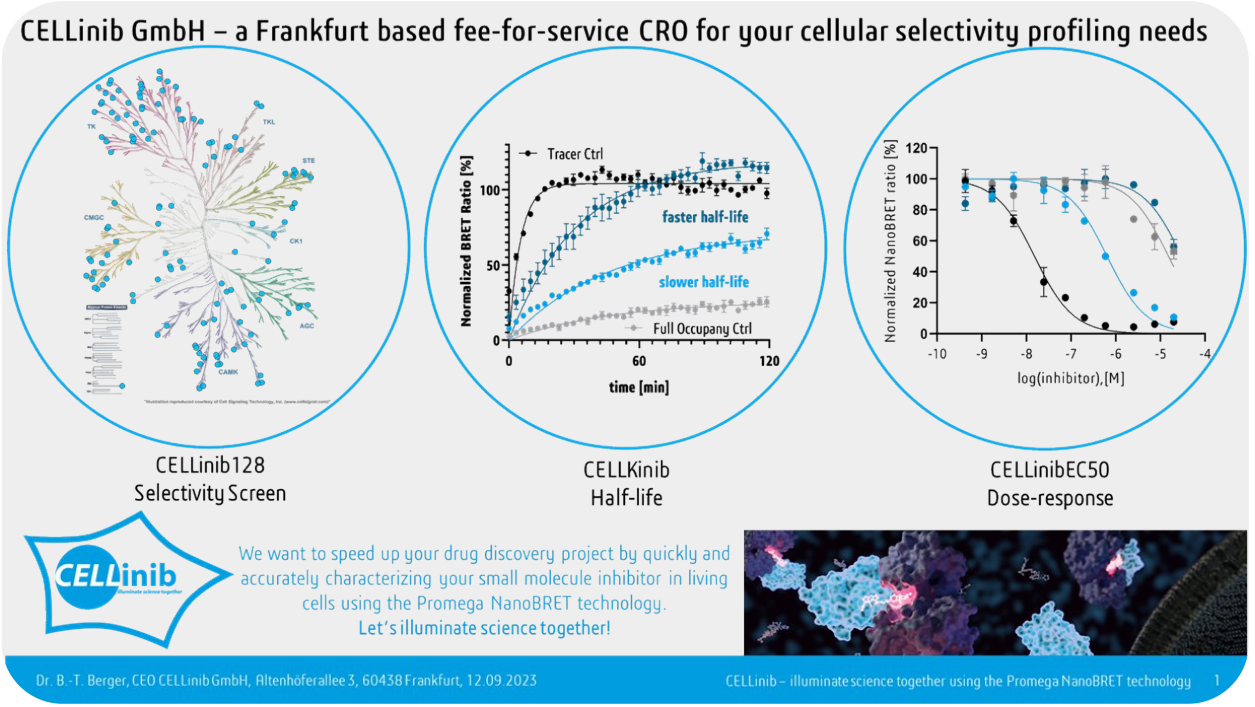
Merk, D., Chaikuad, A. (eds) Chemogenomics. Methods in Molecular Biology, vol 2706. Humana, New York, NY https://doi.org/10.1007/978-1-0716-3397-7_1

- Protocols and methods established for Chemogenomic published in articles and as book
- Protocols on EUbOPEN Website
- All tools available
- Company founded offering established assay platform









EXTREME OPEN SCIENCE

SGC

SGC

Search

News Careers Publications Open Lab Notebooks Contact Log in

About Science Reagents & Resources People News & Outreach

Open Lab Notebooks

Extreme Open Science Initiative: SGC scientists around the world are starting to post their lab notebook online in real time.

University of Toronto, Canada

Heng Zhang Jolene Ho Megha Abbey Rachel Harding

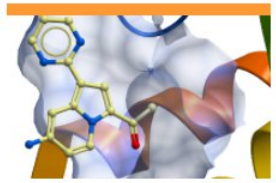
University of Oxford, U.K

Jong Fu Wong Roslin Adamson Elizabeth Brown

Mandeep Mann David Dilworth Genna Luciani

University of North Carolina, U.S.A.

Alfredo Picado Carla Alamillo Nirav Kapadia



SGC

Latest Structures

MBD3: Human methyl-CpG binding domain protein 3, MBD domain, in complex with methylated CpG DNA

CARM1: Human

New Publications

Structural basis for the ability of MBD domains to bind methyl-CG and TG sites in DNA.
J. Biol. Chem.

Structural and functional analysis of the DOT1L-AF10

Chemical Probes

SGC3027 - for PRMT7
22nd March 2018

SGC-GAK-1 - a chemical probe for GAK
27th February 2018

SGC-AAK1-1 - a dual inhibitor of

News from SGC

Conference announcement: BMP Signalling in Cancer II
Posted on *16th March 2018*

ALS Reproducible Antibody Platform: Open Science to enable consistent data and accelerate

scientificdataupdates

a blog from *Scientific Data*

Go

Advanced search

Scientific Data › Blog › Post

Previous post
Data Matters: Interview with Ben Lehner

Next post
Expanding our generalist data repository options

SCIENTIFIC DATA | SCIENTIFIC DATA

An open approach to Huntington's disease research

October 19, 2016 | 1:11 pm | Posted by Andrew Hufton | Category: Featured, Guest Posts

Guest post by Rachel Harding, postdoctoral fellow at the Structural Genomics Consortium, University of Toronto, Canada

RACHEL HARDING

Huntington's disease (HD) is a fatal neurodegenerative disorder caused by a mutation in the *huntingtin* gene[1]. The progressive break down of brain neuronal cells in HD patients leads to deteriorating mental and physical abilities over a 10-20 year period prior to death, the symptoms often described as having Parkinson's disease, Alzheimer's disease and amyotrophic lateral sclerosis (ALS) simultaneously[2]. At the start of the *huntingtin* gene there is a CAG trinucleotide repeat region that encodes a stretch of poly-glutamine residues in the amino-terminus of the encoded protein. This repeat tract is expanded in HD patients. The repeat length of this region correlates with the age of symptom onset[3]. Affecting approximately 1 in 10,000 of the population[4], rare juvenile forms of the disease exist in patients with the longest CAG expansions, although adult-onset HD patients typically have between 40-50 CAG repeats with symptom onset beginning between the ages of 35-50.

What interests many scientists most in the field, and would be imperative to developing targeted and effective

About this blog

Scientific Data is an online-only, peer-reviewed publication for descriptions of scientifically valuable datasets. Follow this blog for news about *Scientific Data*, as well as commentary from our editors and the diverse set of researchers, funders, and data managers who are supporting us.
Find out more

E-alert RSS Twitter

nature.com blogs home

Featured posts from this blog

Calls for Submissions, Featured
Call for submissions: Rescue your data
add a comment

Calls for Submissions, Featured, Repository Highlights
Call for submissions: Open research data resources
add a comment

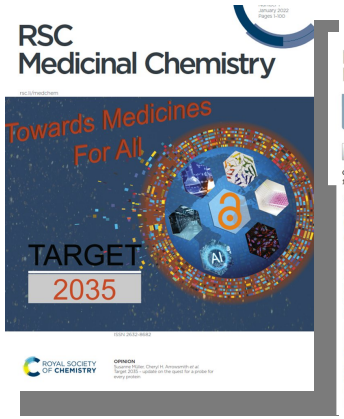
Featured, Guest Posts
An open approach to Huntington's disease research
add a comment

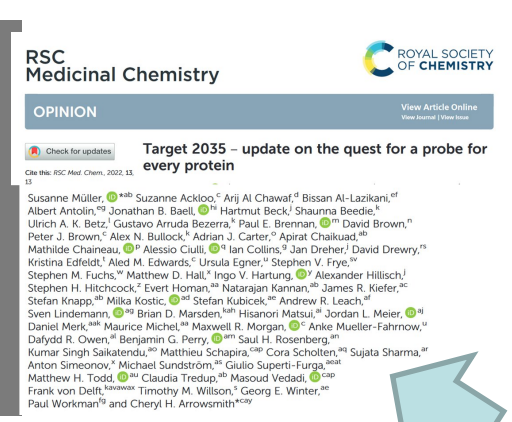
Data Matters, Featured
Data Matters: Interview with Ben Lehner
add a comment

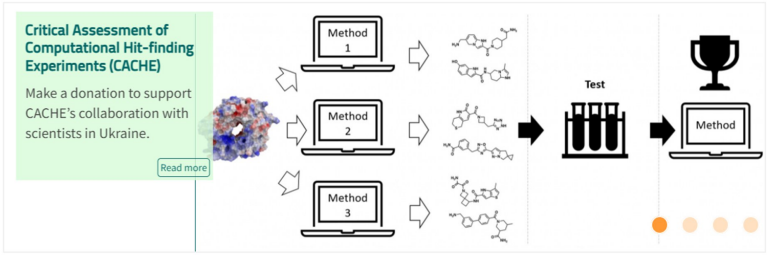
Editor Posts, Featured
Data citations at *Scientific Data*
add a comment

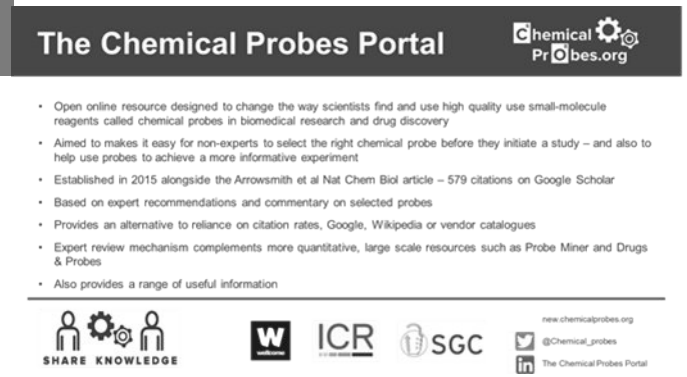
PUSHING THE OPEN SCIENCE BOUNDARIES

SGC



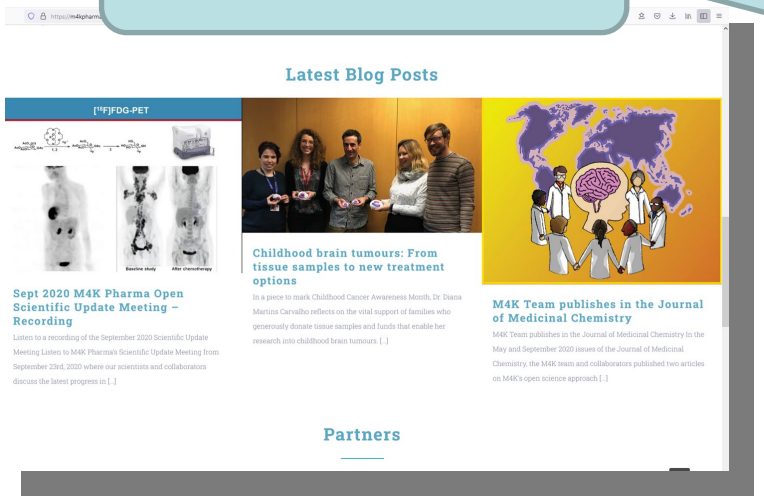


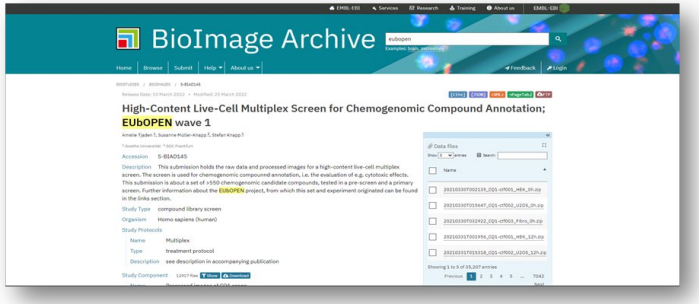




… crowd-sourced evaluation of research tools

Agora Open Science Trust (Charity)





… deposition of metadata

Open Lab Notebooks

Extreme Open Science Initiative: SGC scientists around the world are starting to post their lab notebook online in real time.

… University of British Columbia, Emory University, Sick Kids Hospital…

ACKNOWLEDGEMENTS

SGC Frankfurt
Stefan Knapp
Apirat Chaikuad
Thomas Hanke
Vaclav Nemec
Krishna Saxena
Vladimir Rogov
Sebastian Matthea
Andreas Jörger
Andreas Krämer
Benedikt-T. Berger
Lena Berger
Amelie Tjaden
Yufeng Pan
Lewis Elson
Tim Krapoth
Theresa Ehret
Saran Aswathaman
Sivashanmugam
Claudia Tredup
Natalie Schneider
Sandra Röhm
Dinh Nguyen To
Hanna Holtmann
Martin Schwalm
Jelena Bozilovic
Theresa Mensing
Jennifer Arnheim
Marcel Rak
Nadine Russ
Joshua Gerninghaus
Viktoria Morasch
Nicolas Bauer
Nebojsa Miletic
Guiqun Wang
Thomas Hanke
Dimitrios-Ilias Balourdas
Francesco Greco
Nicolas Bauer
Frederic Farges
Sandra Häberle
Christopher Lenz
Guiqun Wang
Anton Hamann
Sebastian Mande
Florian Born
Adreas Krämer
Lasse Hoffmann
Yeojin Kim
Adarsh Kumar
Christian Sosa
Nicolai Raig
Rezart Zhubi
Krishna Saxena
Vladimir Rogov
Christopher Lenz
Adrian Haag
Andreas Jörger
Aleksandar Lucic
Sebastian Mathea
Verena Dedered
Václav Němec

Academic and Industry Partners
Anton Simeonov (NCATS)
Bryan Roth (UNC)
Benardina Ndreshkjana (GSH)
Henner Farin (GSH)
Yves Janin (Pasteur)
Erik Kowarz (Univ Frankfurt)
Rolf Marschalek (Univ Frankfurt)
Tobias Weiss (Univ Zürich)
Aurino M. Kemas (KI)
Sonia Youhanna (KI)
Volker M. Lauschke (KI)
Andrew Leach (ChEMBL)
Ugis Sarkans (BioImage Archive)
Robert Giessmann (Bayer, Pfizer)
Mark Rolers, Kristin Huwiler, Caroline End

Bayer team
BI team
BMS team
Genentech team
Janssen team
Merck KgaA team
Pfizer team
Takeda team

SGC team
Aled Edwards
Cheryl Arrowsmith
Michael Sundström
Tim Willson

www.thesgc.org

Contact

Susanne Müller-Knapp

COO, SGC-Frankfurt

Susanne.mueller-knapp@bmls.de

www.thesgc.org

FUNDING PARTNERS

The Structural Genomics Consortium is a registered charity (no: 1097737) that receives funds from Bayer AG, Boehringer Ingelheim, Bristol Myers Squibb, Genentech, Genome Canada through Ontario Genomics Institute [OGI-196], EU/EFPIA/OICR/McGill/KTH/Diamond Innovative Medicines Initiative 2 Joint Undertaking [EUbOPEN grant 875510], Janssen, Merck KGaA (aka EMD in Canada and US), Pfizer and Takeda. The Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA and Ontario Institute for Cancer Research, Royal Institution for the Advancement of Learning McGill University, Kungliga Tekniska Hoegskolan, Diamond Light Source Limited.