http://druggablegenome.net/
http://datascience.unm.edu/

AN ACADEMIC PERSPECTIVE ON DRUG DISCOVERY & REPOSITIONING FOR COVID-19

Tudor I. Oprea

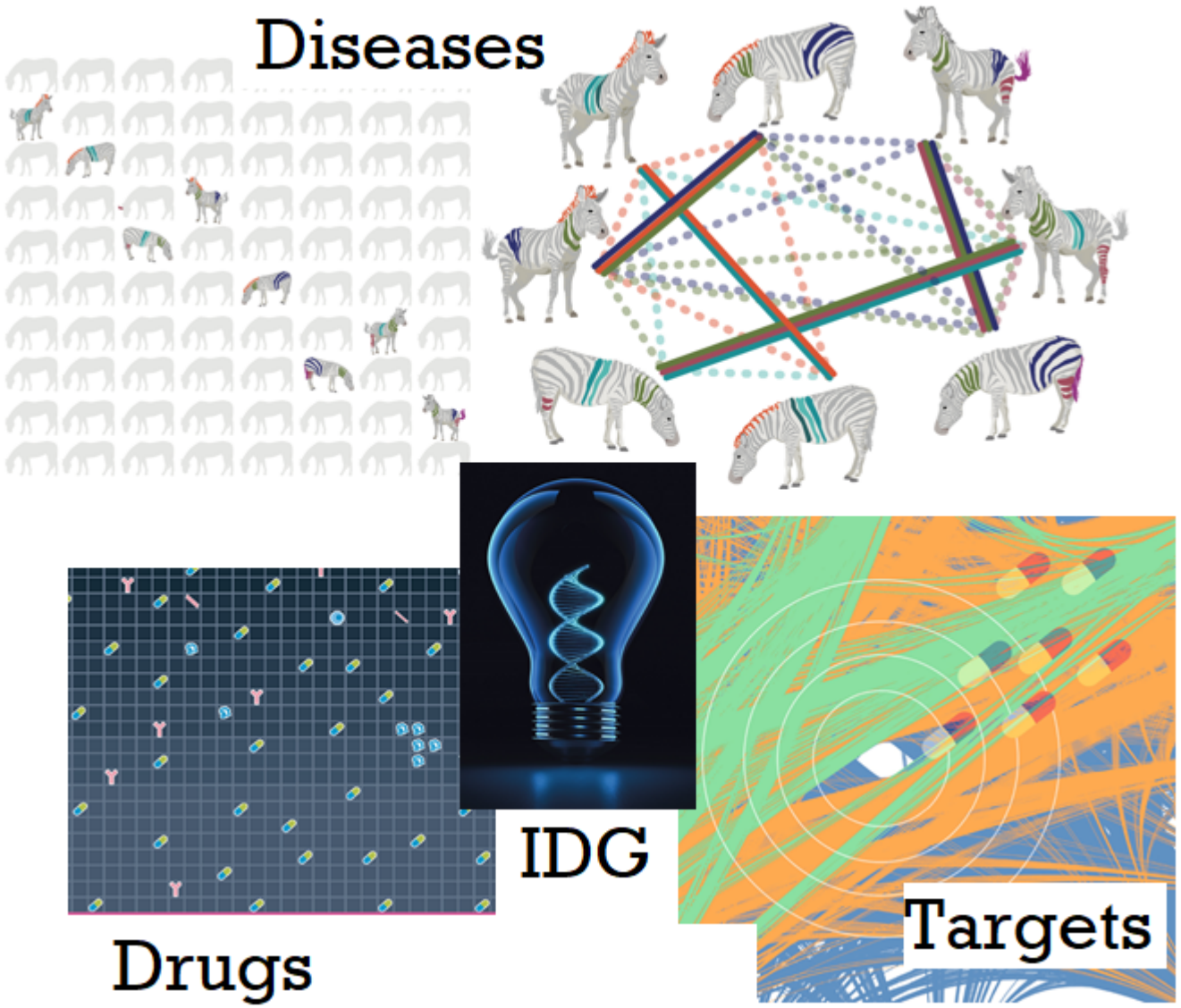
with contributions from Sorin Avram, Giovanni Bocci, Cristian Bologa, Oleg Ursu, Lars Juhl Jensen, Praveen Kumar, Vishal B. Siramshetty, and Gergely Zahoranszky-Kohalmi and many others

Funding: U24 CA224370, U24 TR002278, U01 CA239108 (NIH)

4/14/2020
OHDSI Community Call
Albuquerque, NM, via ZOOM

IDG

THREE PILLARS OF DRUG DISCOVERY



- There are 3 main pillars at the root of successful drug discovery programs.
- Informatics, Data Science and Machine Learning ("AI" according to the hype cycle) are successfully used, as follows:
- **Diseases:** Significant improvements in EMR processing, nosology, ontology, and EMR-based ML for Dx & mechanisms
- **Targets:** Knowledge graph methods, coupled with ML, for biological discovery, target selection & validation
- **Drugs:** From virtual screening to vaccine design, therapeutic modalities benefit from predictive methods across the board
- **IDG** is developing methods applicable to each of these 3 areas

Diseases image credit: Julie McMurry, Melissa Haendel (OHSU).
All other images credit: Nature Reviews Drug Discovery

2/5/20 revision

IDG

COMPREHENSIVE MAP OF MOLECULAR DRUG TARGETS

We curated 667 human genome-derived proteins and 226 pathogen-derived biomolecules through which 1,578 US FDA-approved drugs act.

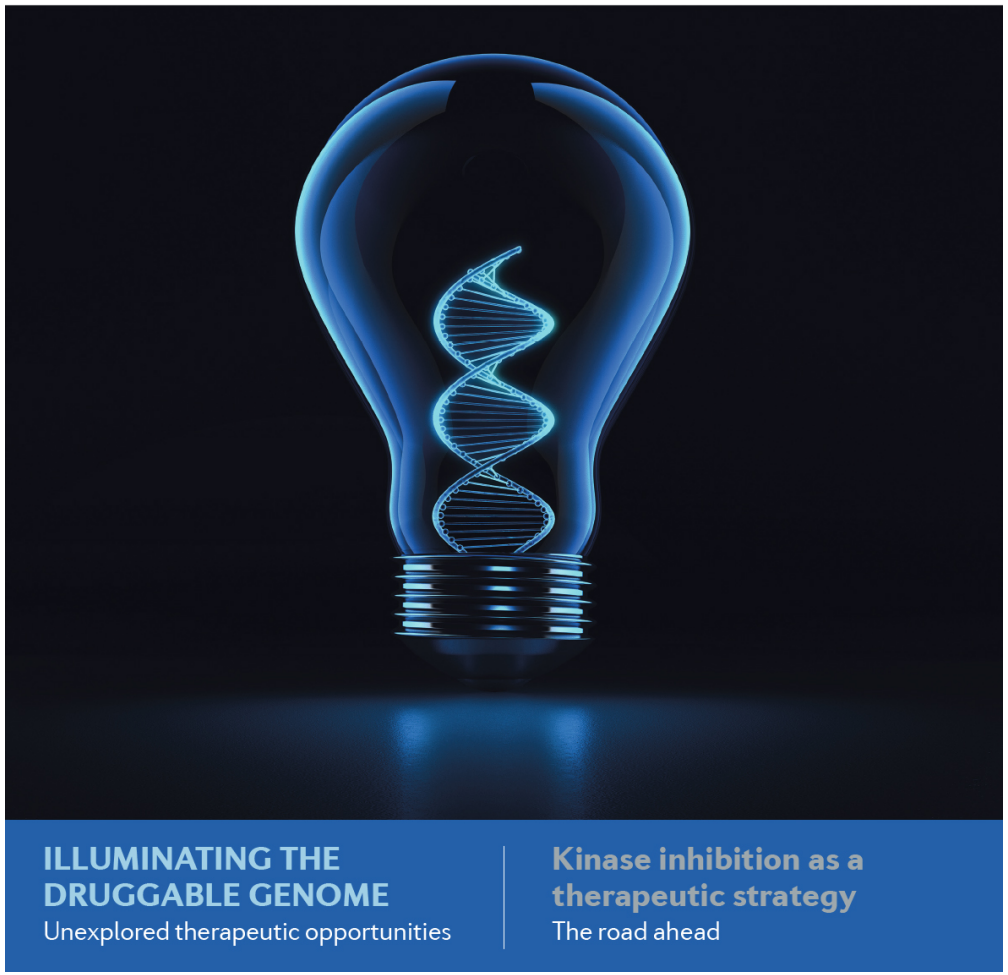
This set included 1004 orally formulated drugs as well as 530 injectable drugs (approved through June 2016).

Data captured in DrugCentral (link)

R. Santos et al., *Nature Rev. Drug Discov.* **2017**, 16:19-34 link

2/4/20 revision

IDG



THE DRUGGABLE

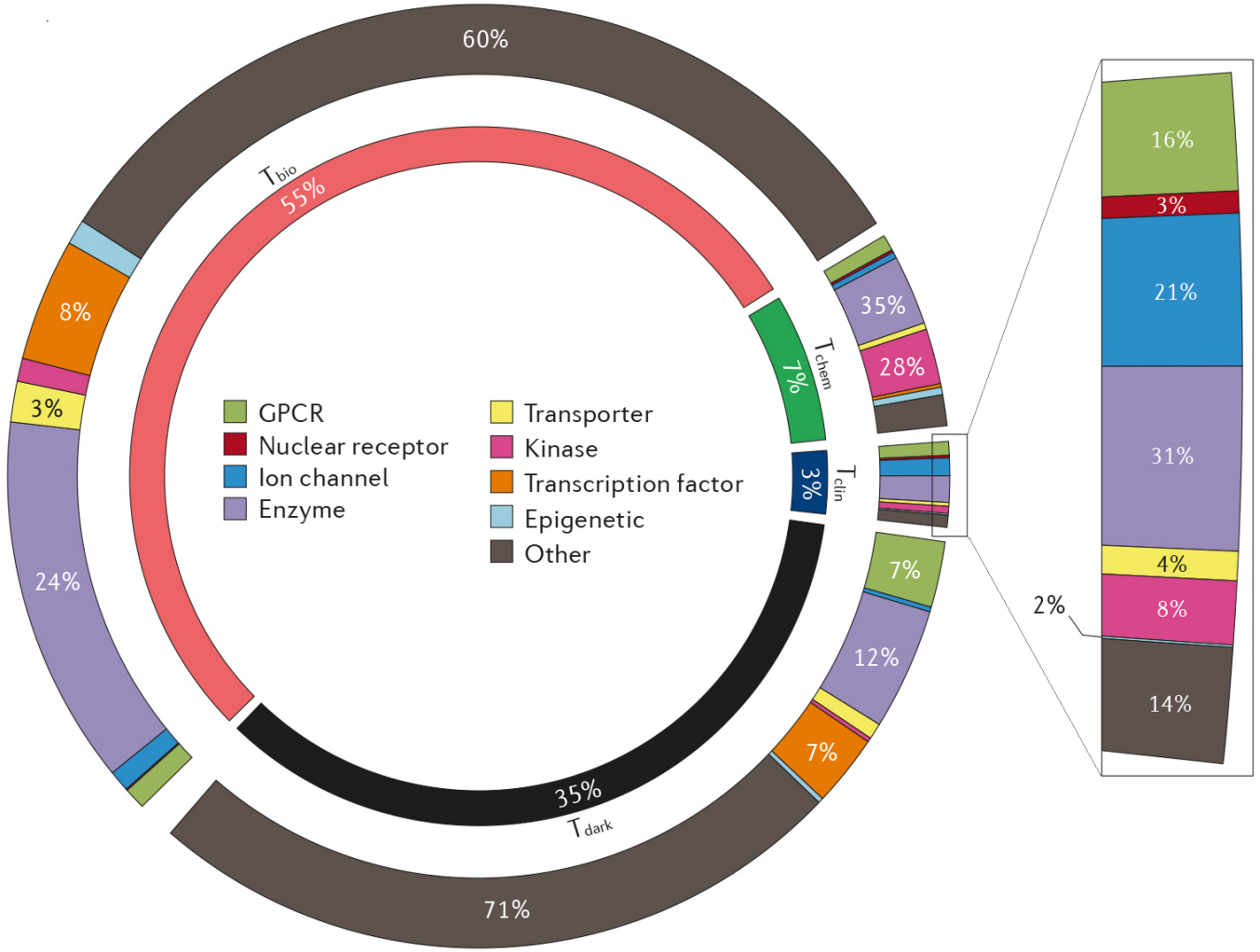
RFA-RM-16-026 (DRGC)	GPCRs	U24 DK116195: Bryan Roth, M.D., Ph.D. (UNC) Brian Shoichet, Ph.D. (UCSF)
	Ion Channels	U24 DK116214: Lily Jan, Ph.D. (UCSF) Michael T. McManus, Ph.D. (UCSF)
	Kinases	U24 DK116204: Gary L. Johnson, Ph.D. (UNC)
RFA-RM-16-025 (RDOC)	Outreach	U24 TR002278: Stephan C. Schürer, Ph.D. (UMiami) Tudor Oprea, M.D., Ph.D. (UNM) Larry A. Sklar, Ph.D. (UNM)
RFA-RM-16-024 (KMC)	Data	U24 CA224260: Avi Ma'ayan, Ph.D. (ISMMS) U24 CA224370: Tudor Oprea, M.D., Ph.D. (UNM)
RFA-RM-18-011 (CEIT)	Tools	U01 CA239106: N Kannan, PhD & KJ Kochut (UGA) U01 CA239108: PN Robinson, MD PhD (JAX), CJ Mungall (LBL), T Oprea (UNM) U01 CA239069: G Wu, PhD (OHSU), PG D'Eustachio PhD (NYU), Lincoln D Stein, PhD (OICR)

T. Oprea et al., *Nature Rev. Drug Discov.* **2018,** 17:317-332 link

2/4/20 revision

IDG

TARGET DEVELOPMENT LEVELS



- Most protein classification schemes are based on structural and functional criteria.
- For therapeutic development, it is useful to understand how much and what types of data are available for a given protein, thereby highlighting well-studied and understudied targets.
- **Tclin:** Proteins annotated as drug targets
- **Tchem:** Proteins for which *potent* small molecules are known
- **Tbio:** Proteins for which biology is better understood
- **Tdark:** These proteins lack antibodies, publications or Gene RIFs

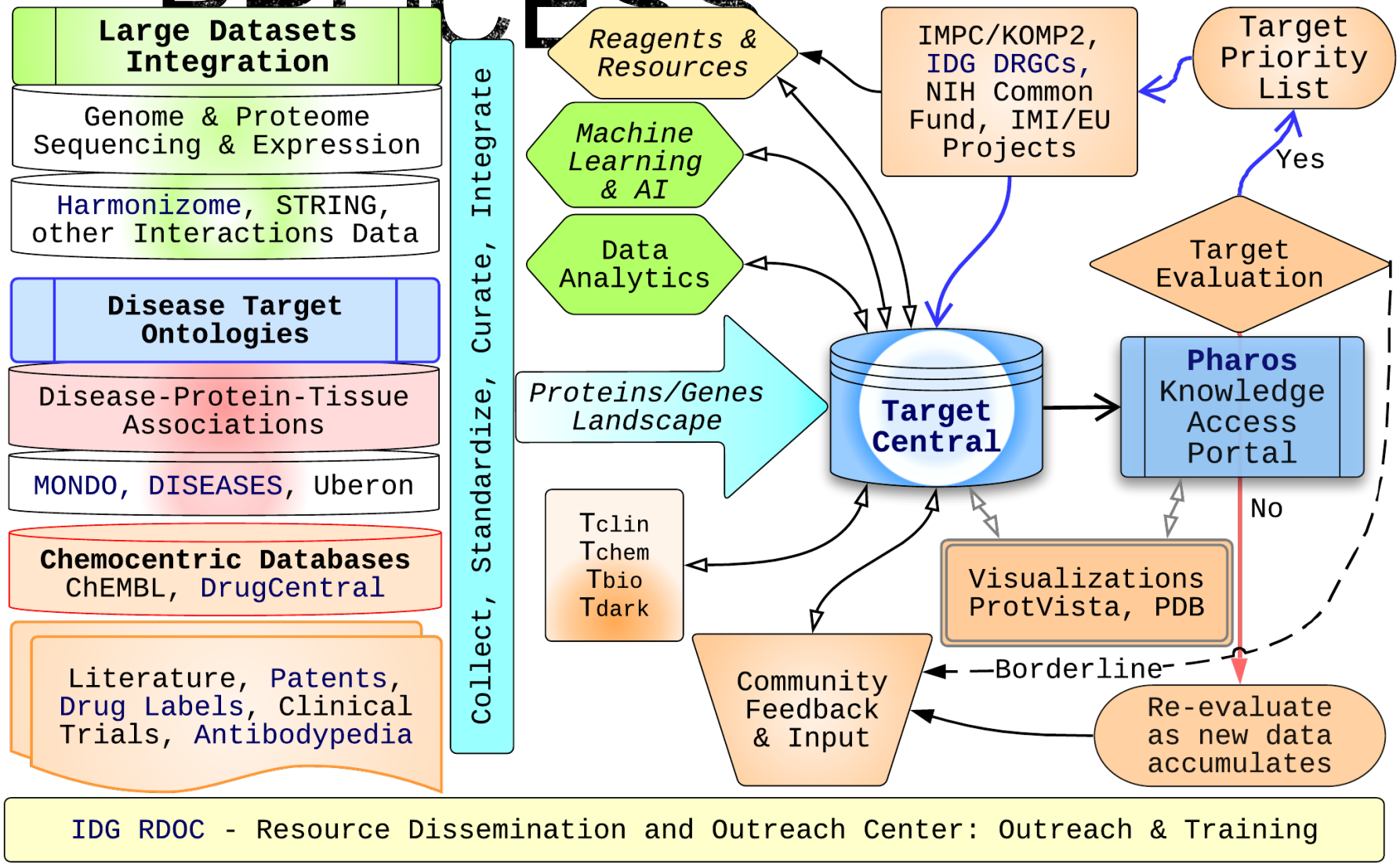
2020 Update: Tdark 31.2%; Tbio 57.7%; Tchem 8%; Tclin 3.1%

T. Oprea et al., *Nature Rev. Drug Discov.* **2018,** 17:317-332 link

2/10/20 revision

IDG

IDG RMC ANNOTATION PROCESS



Further information
Email: idg.rdoc@gmail.com
Follow: @DruggableGenome
URLs:
https://druggablegenome.net/
https://commonfund.nih.gov/idg/

PHAROS

IDG Knowledge User-Interface
Email: pharos@mail.nih.gov
Follow: @IDG_Pharos
URL: https://pharos.nih.gov/

IDG databases are interfaced in UniProt

https://www.uniprot.org/news/2019/09/18/release

GTEx, LINCS, IMPC: Data from 3 CommonFund programs is already in Pharos

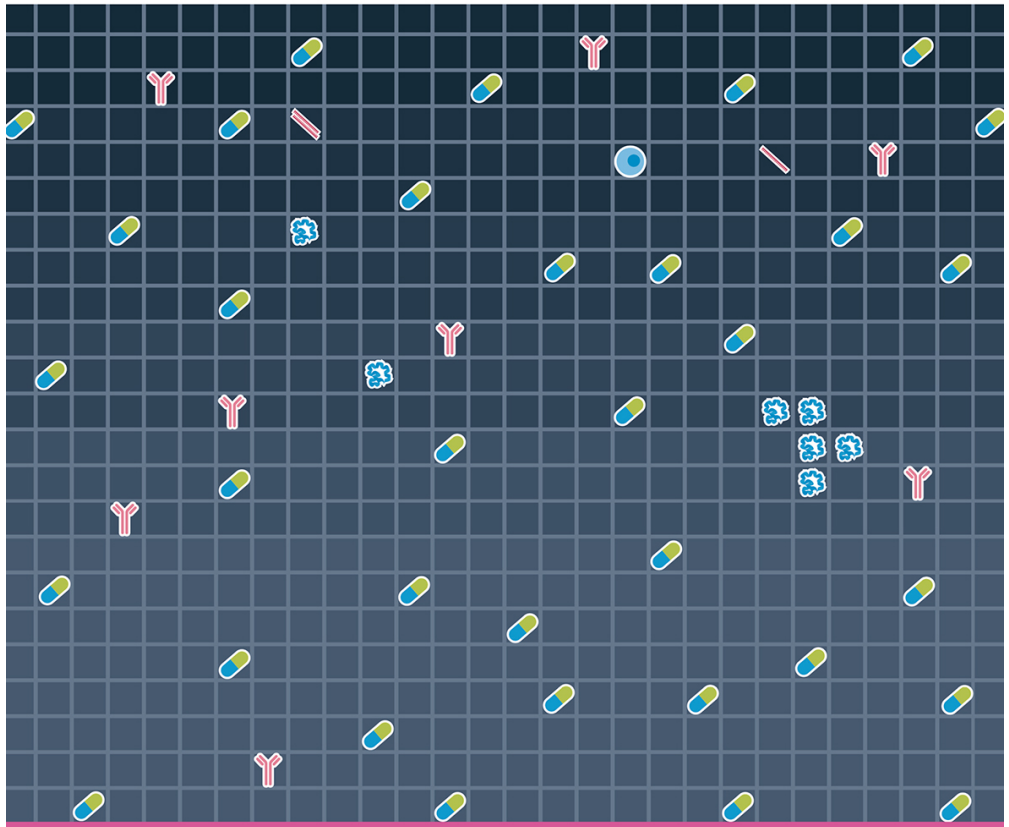
T. Oprea, *Mammalian Genome*, **2019**, 30:192-200 https://bit.ly/2NUK0BK

4/25/19 revision

IDG

February 2020 volume 19 no. 2
www.nature.com/nrd

nature reviews
drug discovery



RARE DISEASES: AN INFORMATICS SURVEY

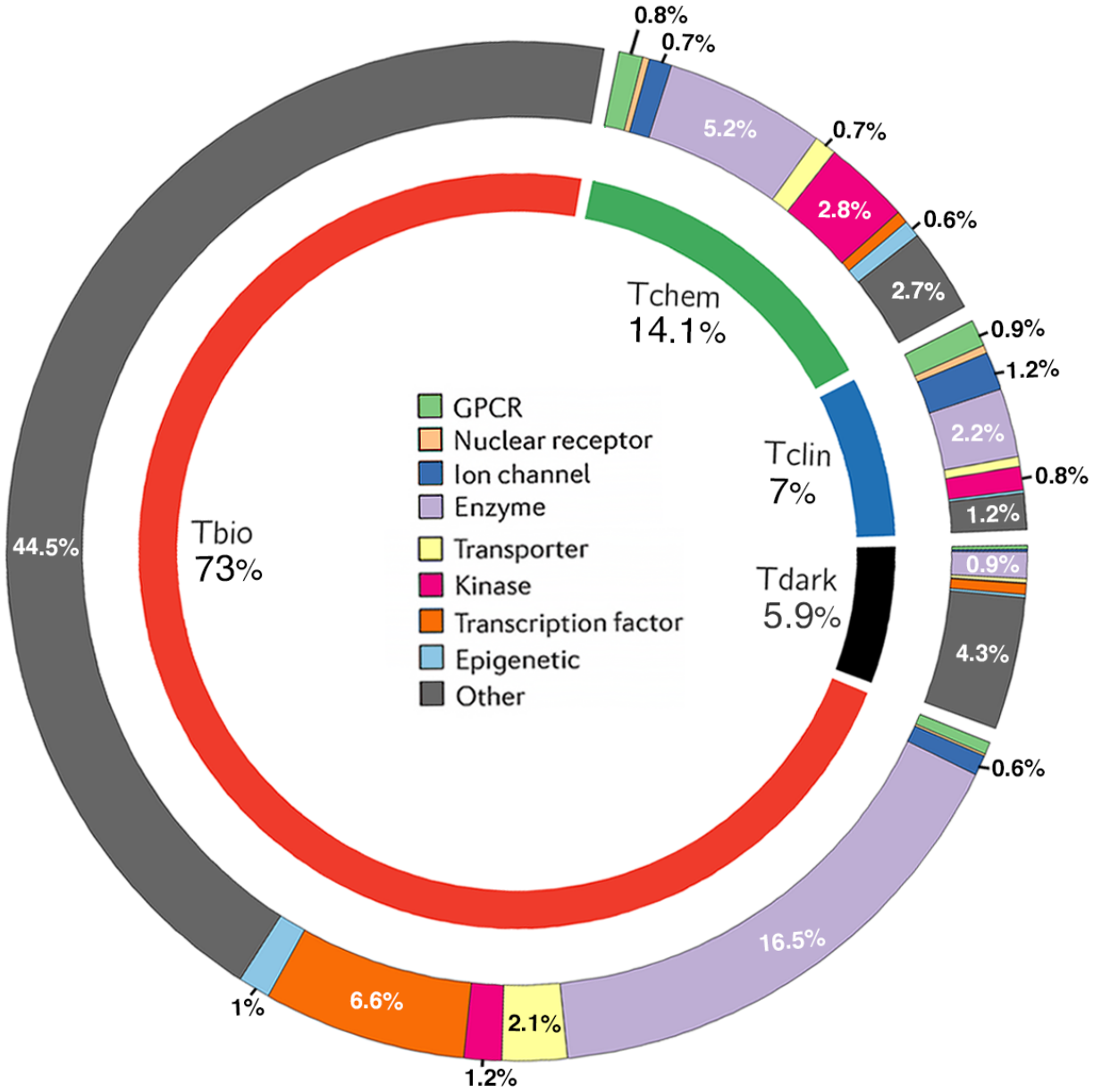
- *We revised the number of RDs from ~7,000 to 10,303* using Disease Ontology, OrphaNet, GARD, NCIT, OMIM and the Monarch Initiative MONDO system
- We also pointed out the lack of a uniform definition for rare diseases, and called for coordinated efforts to precisely define them
- We surveyed therapeutic modalities available to translate advances in the scientific understanding of rare diseases into therapies, and discussed overarching issues in drug development for rare diseases.

Haendel M, et al. *Nature Rev. Drug Discov.* **2020** 19:77-78 link

2/4/20 revision

IDG

THE TARGET SPACE OF RARE DISEASES



- 6077 human proteins are associated with at least one Rare Disease.
- Sources: Disease Ontology (RD-slim), eRAM and OrphaNet
- ~50% agreement (gene level)
- Contrast: Tclin at 3% & Tchem at 7% overall vs. RD subset: 6.94% Tclin and 14.1% for Tchem.
- 20% of the RD proteome is Tclin & Tchem. **This means hope for cures.**
- *Potentially significant opportunities for target & drug repurposing*.

Tambuyzer E, et al. *Nature Rev. Drug Discov.* **2020** 19:93-111 link

2/4/20 revision

IDG

TAKE HOME MESSAGE

THERE IS A KNOWLEDGE DEFICIT

~31% of the proteins remain understudied (Tdark, ignorome)
that number is steadily decreasing

~11.1% of the Proteome (Tclin & Tchem) are currently targeted by small molecule probes and drugs – *that number is slowly increasing*

With help from rare disease patient advocacy groups, rare disease research is likely to witness a significant increase in translation

IDG

COVID-19 CLINICAL TRIALS

- To date, no drug is expressly approved for SARS-CoV-2 infections except the emergency authorization for (hydroxyl)chloroquine.
- Until such time that effective vaccines and/or therapeutics are approved, our "best guess" is "drug repositioning" (aka drug repurposing) followed by drug discovery
- As of 4/13/20, there were 469 clinical trials for "COVID-19".
- Of these 469 clinical trials, 222 are for a "drug" intervention.

https://clinicaltrials.gov/ct2/results?cond=COVID-19

4/14/20 revision

IDG

CURATING COVID-19 CLINICAL TRIALS

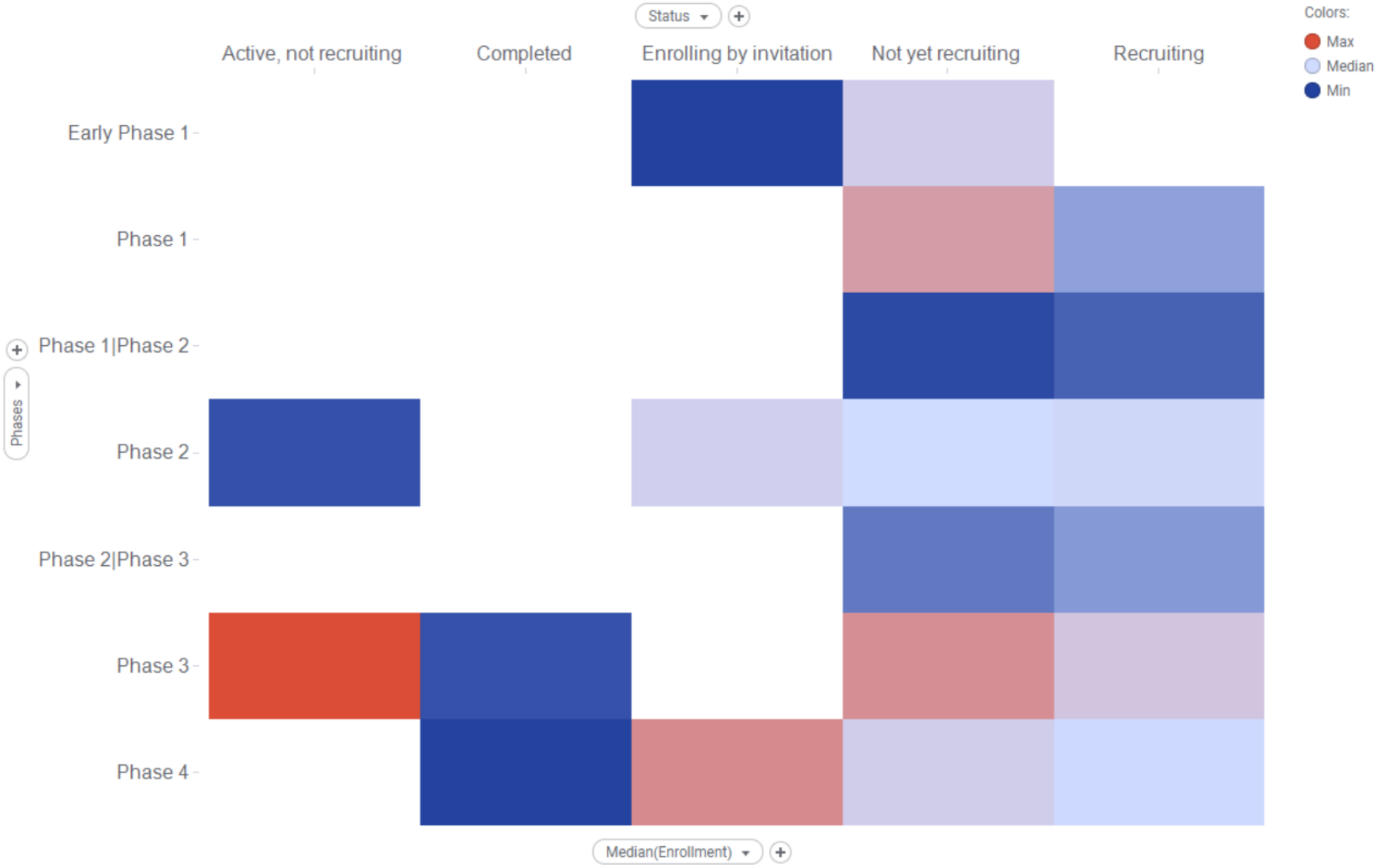
- Manual curation (by intervention), e.g., identify experimental (novel) drug vs. already approved drug; reconcile spelling errorls (e.g., hidroxicloroquin; abidol)
- Five general categories: Placebo (63), Antiviral (40), Experimental (29), Repurposed (165), Biologic (50)
- Twenty-one "specific" categories: HCQ (63), CQ (8), Azithromycin (20), -navir (20), Oseltamivir (4), Favipiravir (3), Umifenovir (5), Remdesivir (9), -tinib (8), RAS drugs (13), NSAIDs (4), Steroid (13), TMPRSS2 (4), Traditional Chinese (8), Colchicine (4), Gases (10), Tocilizumab (14), Anakinra (4), IFN (12), Ig-based (4), Supplements (6)

https://clinicaltrials.gov/ct2/results?cond=COVID-19

4/14/20 revision

IDG

REPOSITIONING CLINICAL



- COVID-19 only
- Filtered out:

Observational studies; "not repurposed" (e.g., experimental); withdrawn; "Phase" N/A, or not applicable).

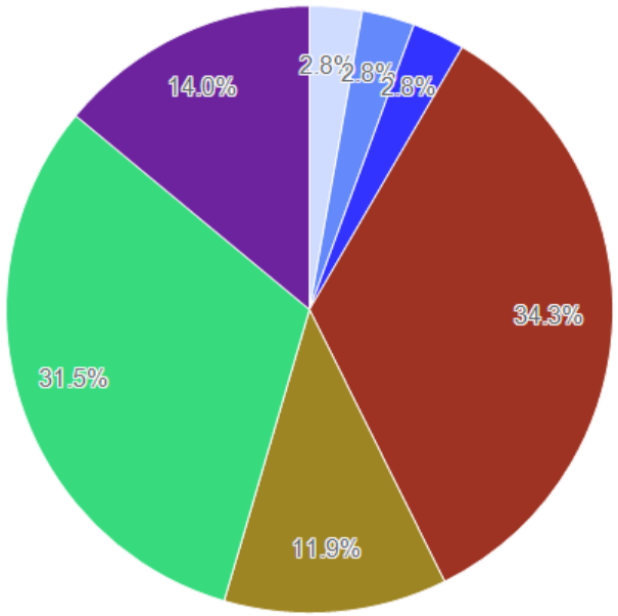
- Median: 205 patients

https://clinicaltrials.gov/ct2/results?cond=COVID-19

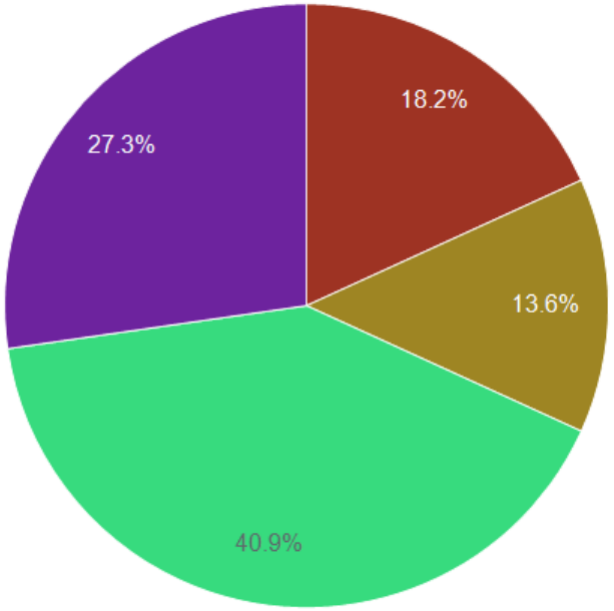
4/14/20 revision

IDG

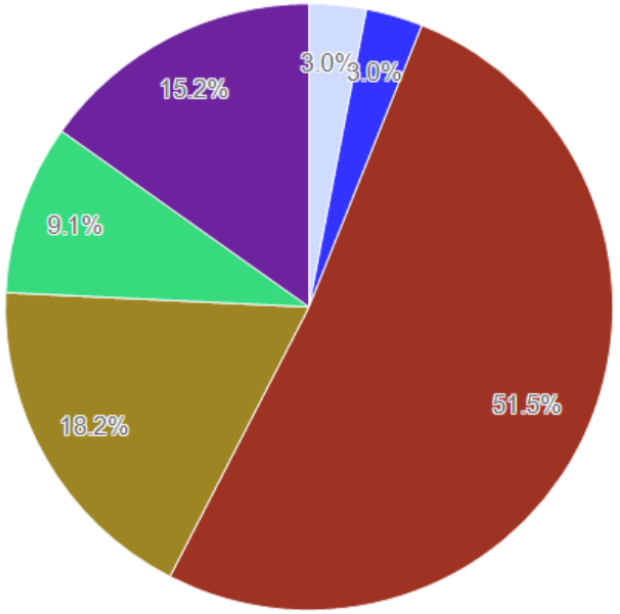
REPOSITIONING CLINICAL TRIALS: SUMMARY



All repurposed; N = 143



Antiviral; N = 22



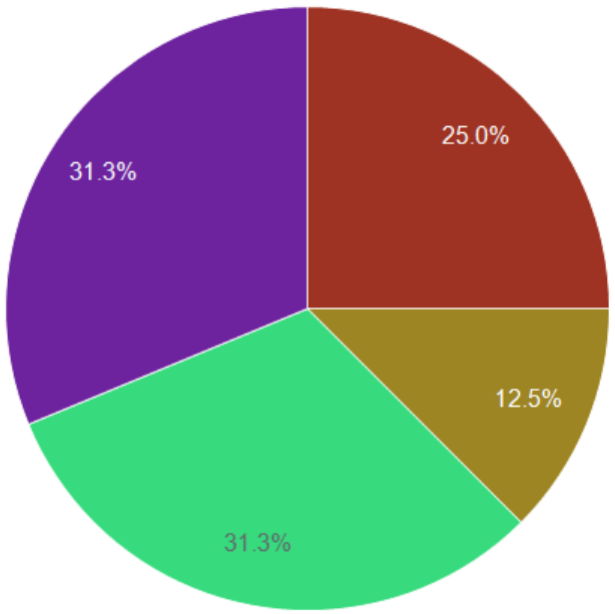
Biologics; N = 33

https://clinicaltrials.gov/ct2/results?cond=COVID-19

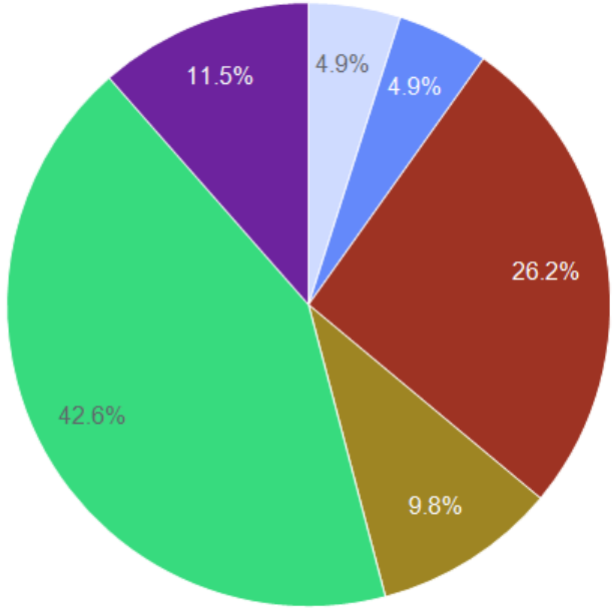
4/14/20 revision

IDG

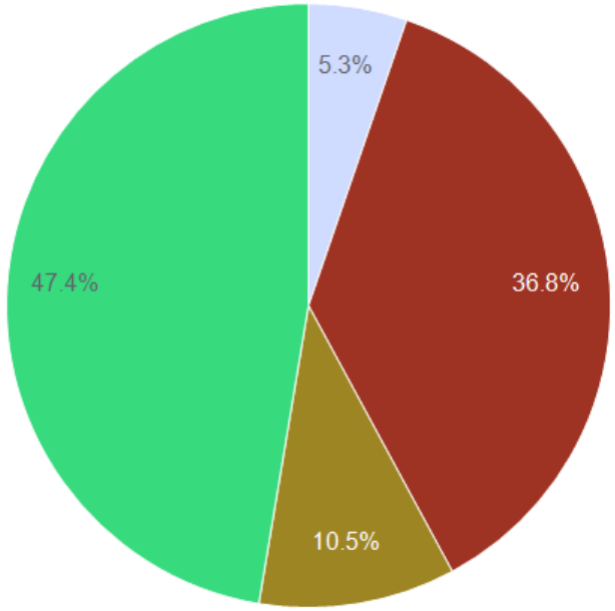
REPOSITIONING TRIALS BY DRUG



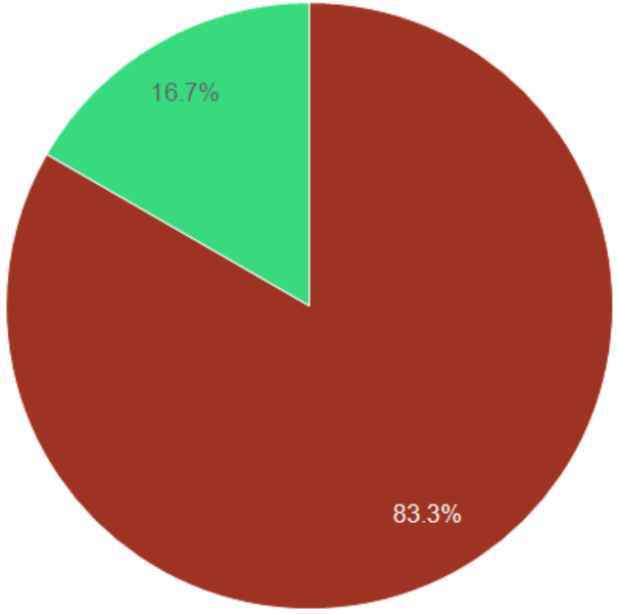
-navir; N = 16



HCQ; N = 61



Azithromycin; N = 19



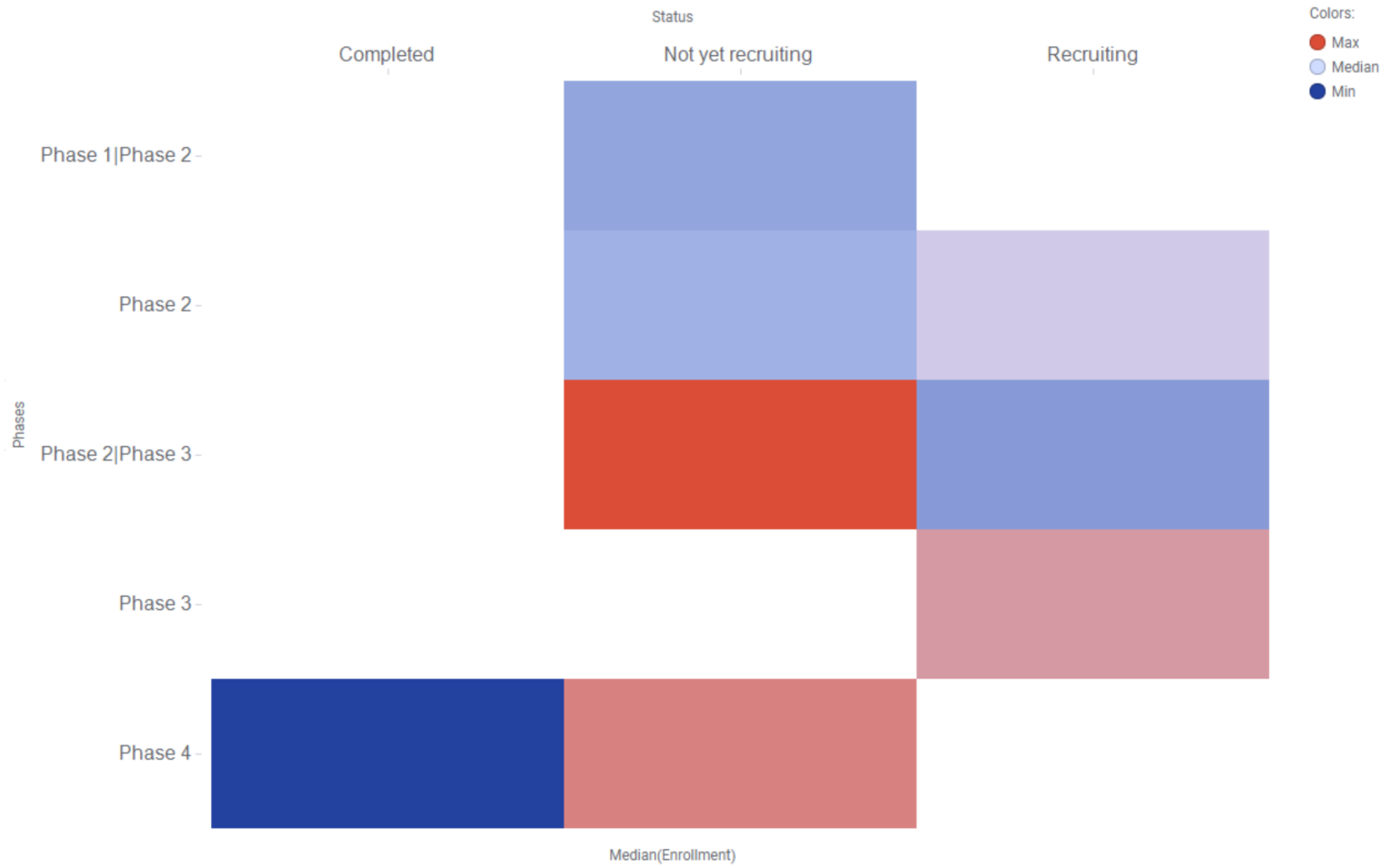
Tocilizumab; N = 12

https://clinicaltrials.gov/ct2/results?cond=COVID-19

4/14/20 revision

IDG

EXPERIMENTAL CLINICAL



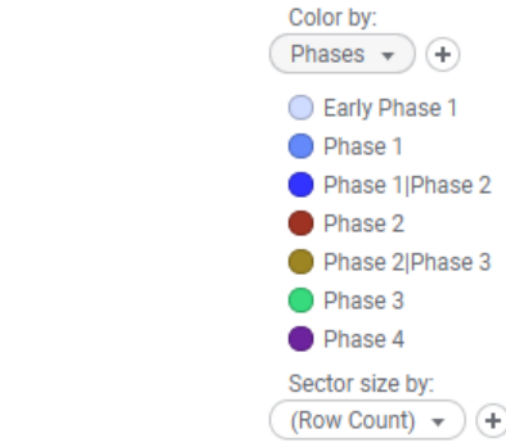
- COVID-19 only
- Filtered out: Observational studies; "not repurposed" (e.g., experimental); withdrawn; "Phase" N/A, or not applicable).
- Experimental drugs shown.
- Median: 216 patients

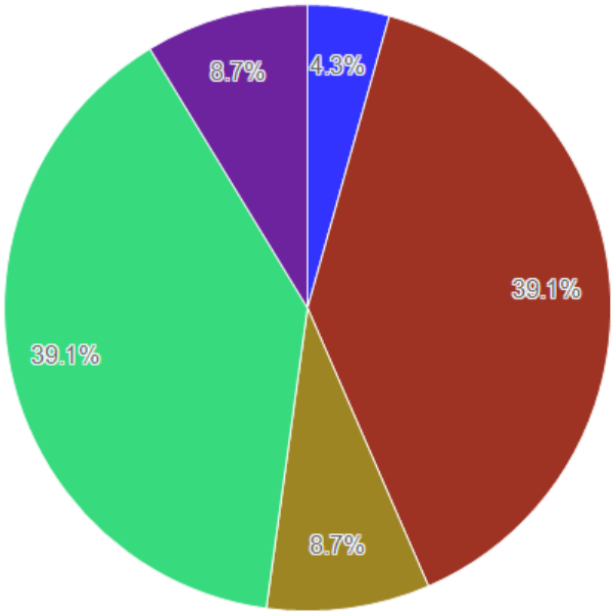
https://clinicaltrials.gov/ct2/results?cond=COVID-19

4/14/20 revision

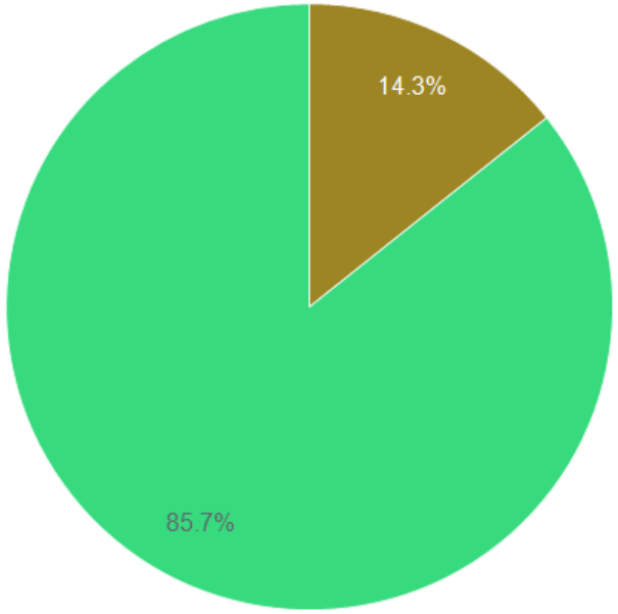
IDG

EXPERIMENTAL CLINICAL TRIALS: SUMMARY

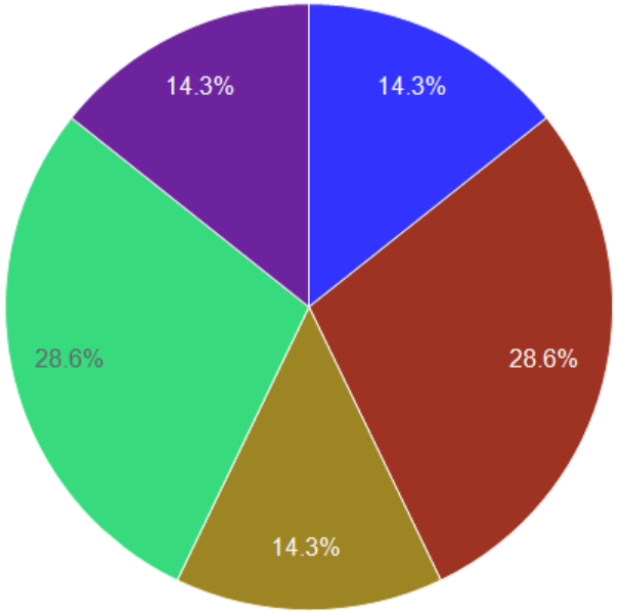




Experimental; N = 23



Remdesivir; N = 7



Biologics; N = 7

https://clinicaltrials.gov/ct2/results?cond=COVID-19

4/14/20 revision

IDG

TAKE HOME MESSAGE

MOST COVID-19 INTERVENTIONAL CLINICAL TRIALS ARE FOCUSED ON DRUG REPOSITIONING

The majority study HCQ (61 out of 143)

2/5/20 revision

IDG

HOW MANY DRUGS FOR REPOSITIONING?

	Drug Product Forms (Patents)*	Drug Products	Type N Drug Products	Drugs (Patents)*	RX Drugs **
On-Patent	12236 (4585)	1057	1057	**785** (738)	762
Off- Patent	22131	11874	1906	**1454**	1404
Discontinued	16963	11388	2801	**680**^	n.a.
All Drugs	51330	22362	5042	2557	1828 (1258)

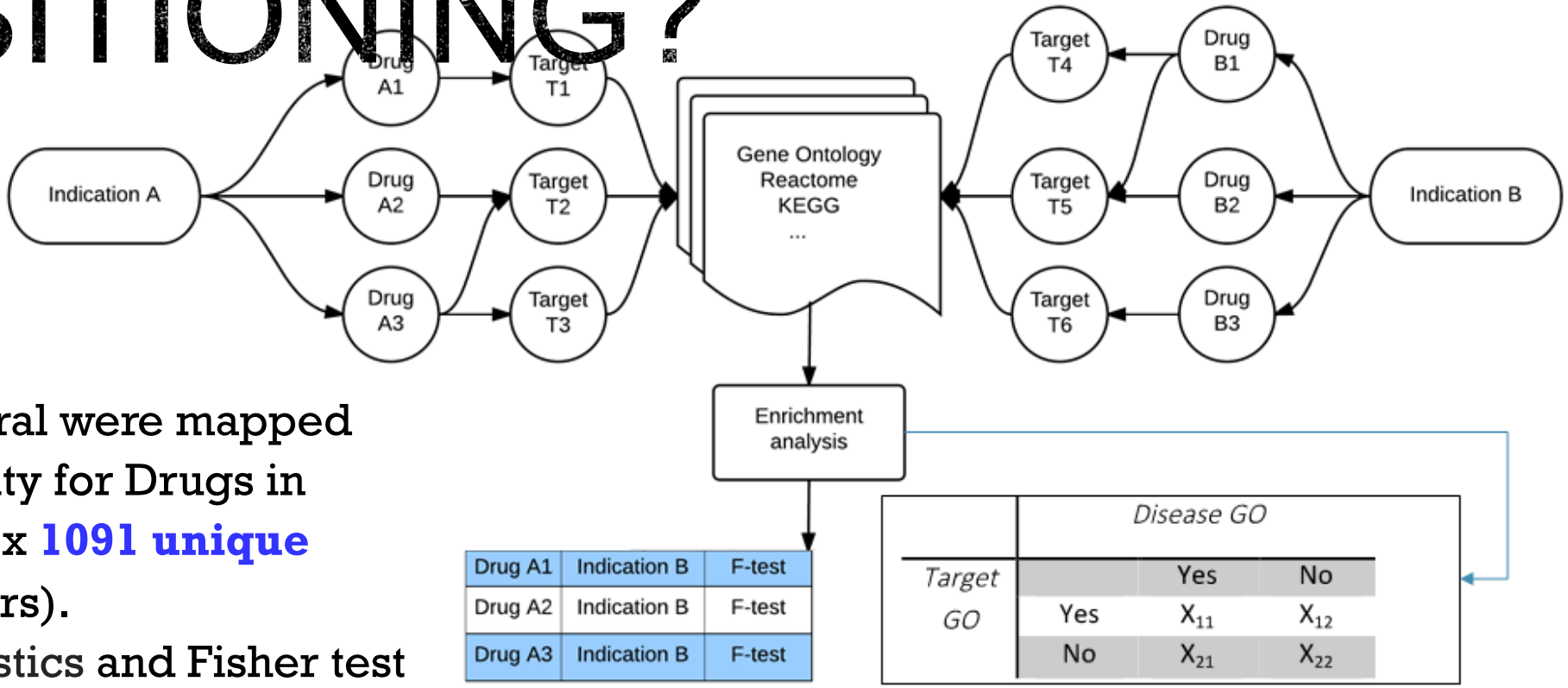
- Analysis based on the archived OrangeBook (2016 – 2019) and the latest **Orange Book Data Files** (*OBDFs*; EOBZIP_2019_10.zip content current as of: 10/18/2019) combined with the lists of **Newly Added Patents and Delisted Patents**
  - * Number of Patents; ** Number of RX drugs as single ingredient; ^ **1833** total, only **680** discontinued
    - Drug Product Forms = # of PIDs (all drug forms, routes, strengths etc)
    - Drug Products = # of Application Numbers
    - Type "N" Drug Products = # of Application Numbers of type "N". i.e., NDA (new drug application). The rest of the Drug products are "ANDA", i.e., abbreviated new drug applications.
    - Drugs = # of (active) ingredients, i.e., actual drug (includes combinations)
    - RX Drugs = # of drugs on prescription. The rest of the drugs are OTCs.
- Up to **1772 active ingredients** may be eligible for "off patent" repurposing

S. Avram & T. Oprea, *unpublished.*

12/04/19 revision

IDG

HOW MANY TARGETS FOR REPOSITIONING?



- Indications from DrugCentral were mapped onto proteins with bioactivity for Drugs in DrugCentral: **881 proteins** x **1091 unique Indications** (over 873k pairs).
- Pearson's chi-squared statistics and Fisher test p-values were applied with the False Discovery Rate correction to p-values (P.adj).
- At $\chi^2 > 1000$ *and* P.adj $\leq 10^{-9}$, **we found up to 60258 novel protein-indication pairs**.

X_{11} – GO terms count annotated for Disease and Target
X_{12} – GO terms count annotated for Target and not Disease
X_{21} – GO terms count annotated for Disease and not Target
X_{22} – GO terms count annotated for not Target and not Disease

Note: this set has not been yet been filtered for off-patent drugs

O. Ursu, C. Bologa & T. Oprea, *unpublished.*

12/04/19 revision

IDG

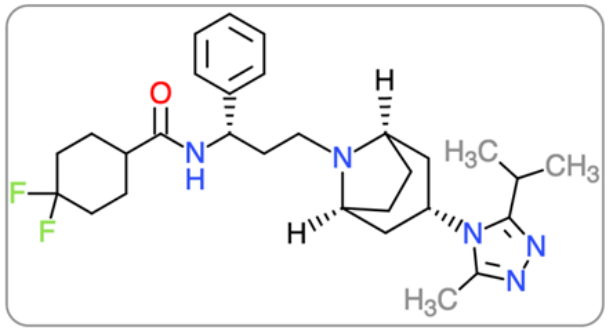
ID MAPPING & THERAPEUTIC DRUG INTENT

IUPHAR International Union of Basic and Clinical Pharmacology
806

U.S. Food and Drug Administration
Protecting and Promoting Your Health
Substance Registration System - Unique Ingredient Identifier (UNII)
MD6P741W8A

PubChem
3002977

RxNav
Navigating RxNorm Drugs
620216



ChEMBL
CHEMBL1201187

MedNet
Communities for Scientific collaboration, Information exchange and sharing
8450

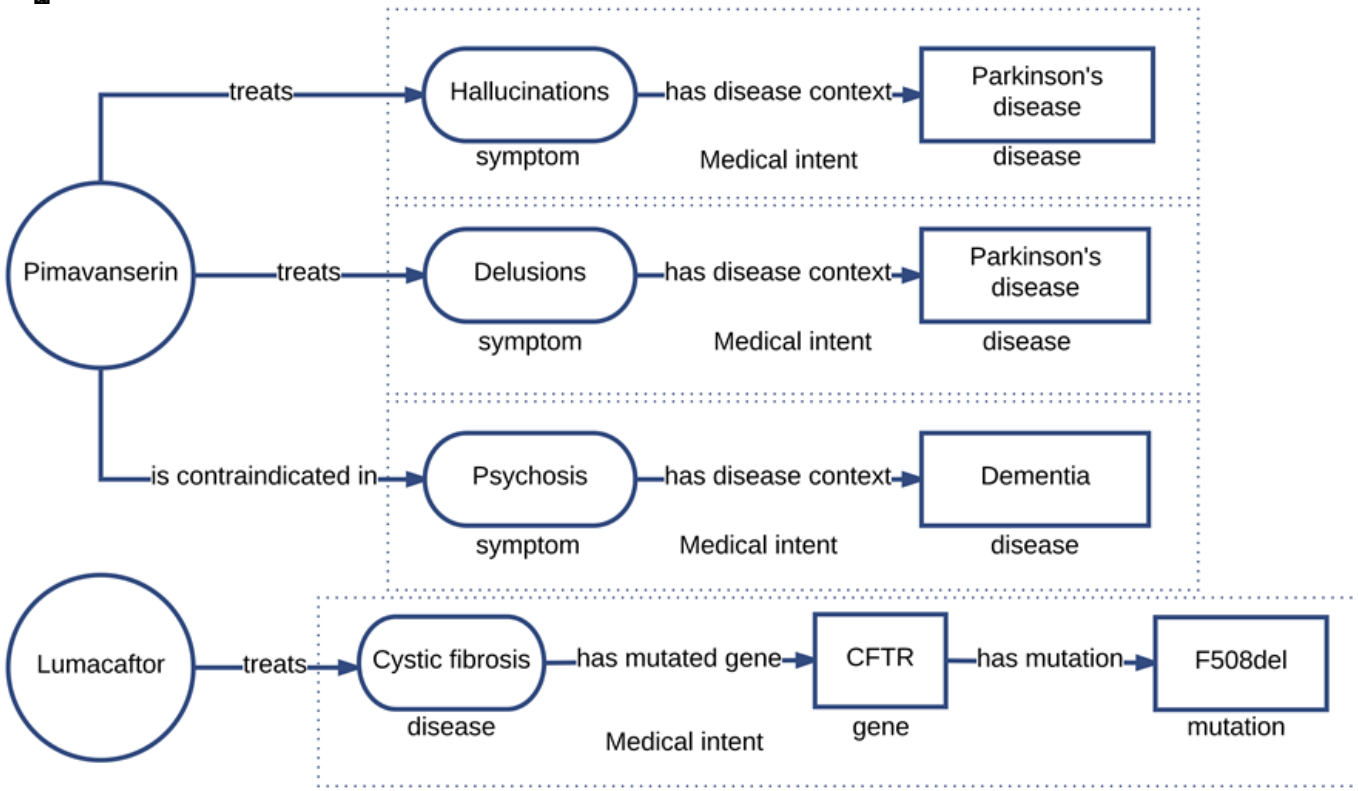
ChEBI
CHEBI:63608

MeSH
C502411

DRUGBANK
DB04835

SciFinder
A CAS SOLUTION
199463-33-7

KEGG
D06670



- Left: Keeping track of multiple identifiers (to navigate across many resources) is a full-time job
- Right: Automated drug repositioning systems are likely to fail because, as of today, *no system* is capable of capturing *therapeutic intent.* If it cannot be captured, it cannot be modeled.

S.J. Nelson et al, *J Am Med Info Assn*, **2017**, 24:1169-1172. *doi: 10.1093/jamia/ocx064*

9/25/17 revision

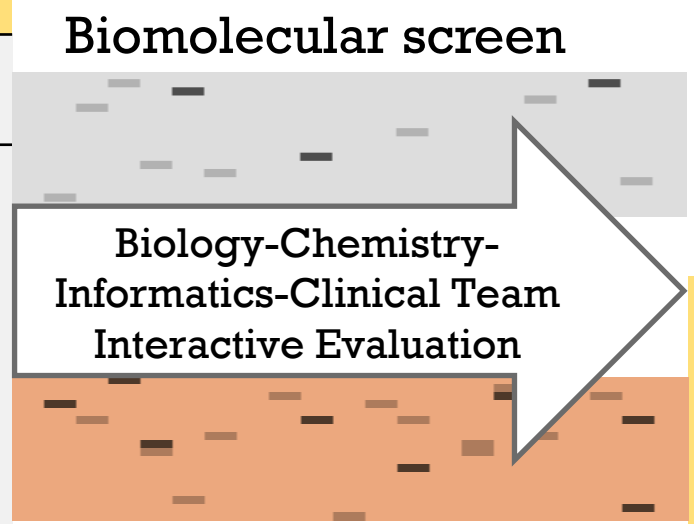
IDG

DRUG REPOSITIONING INTELLIGENCE

Algorithmic Evaluation of Drug Repositioning Opportunities

Physico-Chemical Characteristics	Target & Ligand Based Virtual Screening	Disease-Target-Drug Associations
Chemical fingerprints, chemotypes, derivative chemical descriptors	Shape & electrostatics derived similarity or complementarity	Disease & Drug target GO Annotations from STRING, GO, KEGG, other sources
Solubility, Permeability, Distribution, relevant PK properties	3D models for intended targets, followed by molecular docking	Contigency Tables, e.g., Fisher test for associations *augmented via AI/ML*
Metabolism, efflux transporters, toxicity end-points	Multiple tautomers / protomers / conformers / binding modes per protein	*Prioritize novel disease-target associations for off-patent drugs*

Iteratively Compare, Integrate, Cross-check, Prioritize, Evaluate



Phase II / Phase III clinical trials

Confirmation of clinical effects

New Target Activity:
- *In vitro* animal
- *In vivo* animal
- Human studies
- Disease relevance
- Efficacy/Safety

Implement exact annotations for drug indications and off label uses.
Mandate rigorous validation for computational models.
Support community-based therapy-area specific research.

Turn DR into an international effort, preferably focused on diseases that lack cure

12/04/19 revision

IDG

SOME REPOSITIONING PAPERS

ACS central science

IN FOCUS

Research and Development on Therapeutic Agents and Vaccines for COVID-19 and Related Human Coronavirus Diseases

Cite This: *ACS Cent. Sci.* 2020, 6, 315−331 | Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

Cynthia Liu,* Qiongqiong Zhou, Yingzhu Li, Linda V. Garner, Steve P. Watkins, Linda J. Carter, Jeffrey Smoot, Anne C. Gregg, Angela D. Daniels, Susan Jervey and Dana Albaiu

Since the outbreak of the novel coronavirus disease COVID-19, caused by the SARS-CoV-2 virus, this disease has spread rapidly around the globe. Considering the potential threat of a pandemic, scientists and physicians have been racing to understand this new virus and the pathophysiology of this disease to uncover possible treatment regimens and discover effective therapeutic agents and vaccines. To support the current research and development, CAS has produced a special report to provide an overview of published scientific information with an emphasis on patents in the CAS content collection. It highlights antiviral strategies involving small molecules and biologics targeting complex molecular interactions involved in coronavirus infection and replication. The drug-repurposing effort documented herein focuses primarily on agents known to be effective against other RNA viruses including SARS-CoV and MERS-CoV. The patent analysis of coronavirus-related biologics includes therapeutic antibodies, cytokines, and nucleic acid-based therapies targeting virus gene expression as well as various types of vaccines. More than 500 patents disclose methodologies of these four biologics with the potential for treating and preventing coronavirus infections, which may be applicable to COVID-19. The information included in this report provides a strong intellectual groundwork for the ongoing development of therapeutic agents and vaccines.

https://pubs.acs.org/doi/pdf/10.1021/acscentsci.0c00272

Review

ONLINE FIRST FREE

April 13, 2020

Pharmacologic Treatments for Coronavirus Disease 2019 (COVID-19)

A Review

Abstract

Importance The pandemic of coronavirus disease 2019 (COVID-19) caused by the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) presents an unprecedented challenge to identify effective drugs for prevention and treatment. Given the rapid pace of scientific discovery and clinical data generated by the large number of people rapidly infected by SARS-CoV-2, clinicians need accurate evidence regarding effective medical treatments for this infection.

Observations No proven effective therapies for this virus currently exist. The rapidly expanding knowledge regarding SARS-CoV-2 virology provides a significant number of potential drug targets. The most promising therapy is remdesivir. Remdesivir has potent in vitro activity against SARS-CoV-2, but it is not US Food and Drug Administration approved and currently is being tested in ongoing randomized trials. Oseltamivir has not been shown to have efficacy, and corticosteroids are currently not recommended. Current clinical evidence does not support stopping angiotensin-converting enzyme inhibitors or angiotensin receptor blockers in patients with COVID-19.

Conclusions and Relevance The COVID-19 pandemic represents the greatest global public health crisis of this generation and, potentially, since the pandemic influenza outbreak of 1918. The speed and volume of clinical trials launched to investigate potential therapies for COVID-19 highlight both the need and capability to produce high-quality evidence even in the middle of a pandemic. No therapies have been shown effective to date.

https://jamanetwork.com/journals/jama/fullarticle/2764727

04/14/20 revision

IDG

THE CANCER HUMAN PROTEIN INTERACTOME

04/14/20 revision

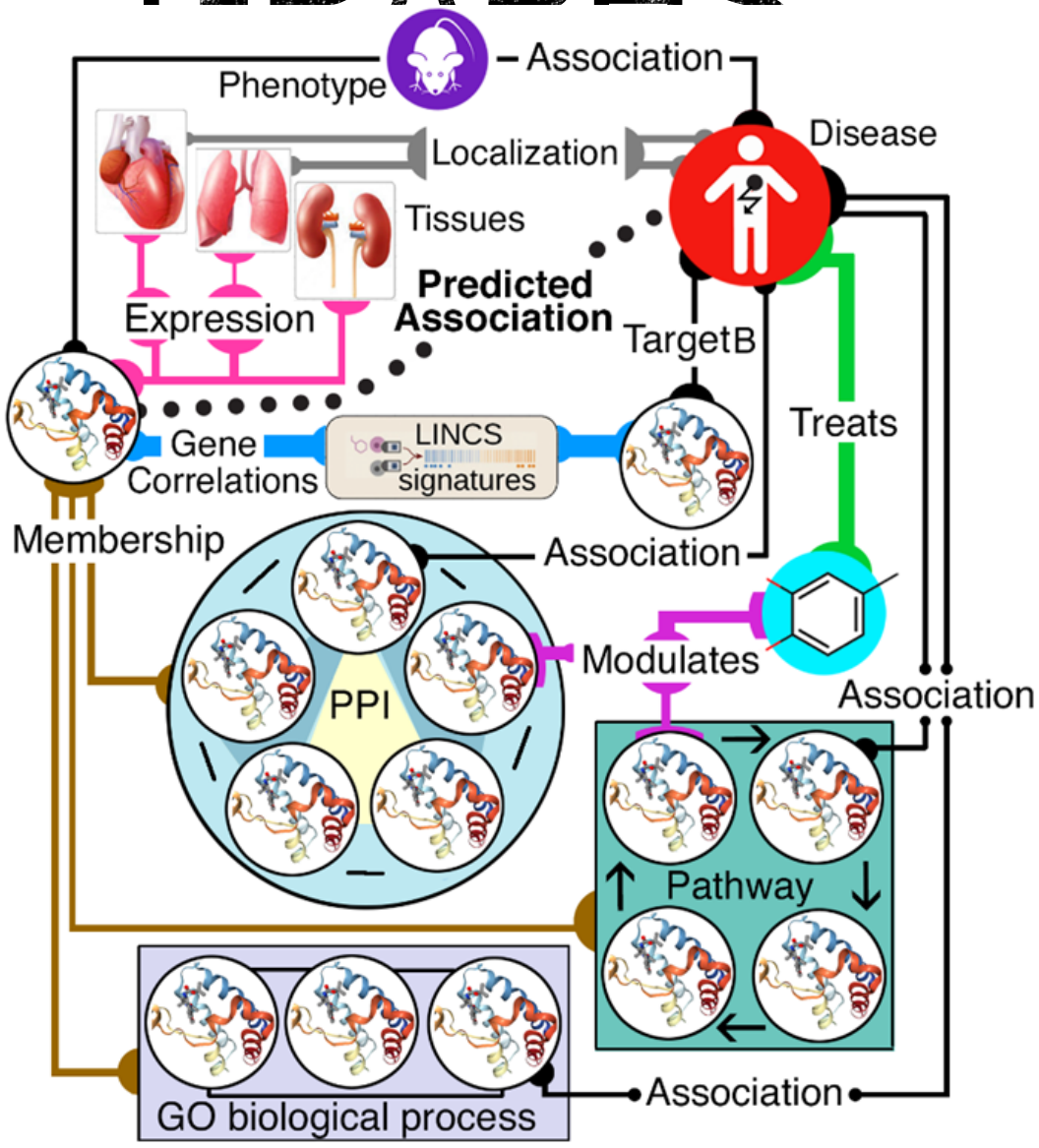
IDG

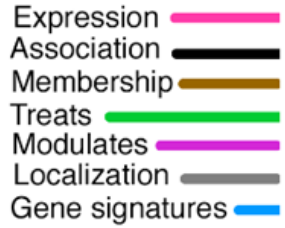
DATA SOURCES AND METHODS

- Krogan et al., preprint (cleaned-up by Lars Jensen, PhD)
- P-HIPSTER predictions
- Metapath/XGBoost AIML predictions
- STRING analysis (StringApp, Cytoscape)

IDG

PROTEIN KNOWLEDGE GAPS





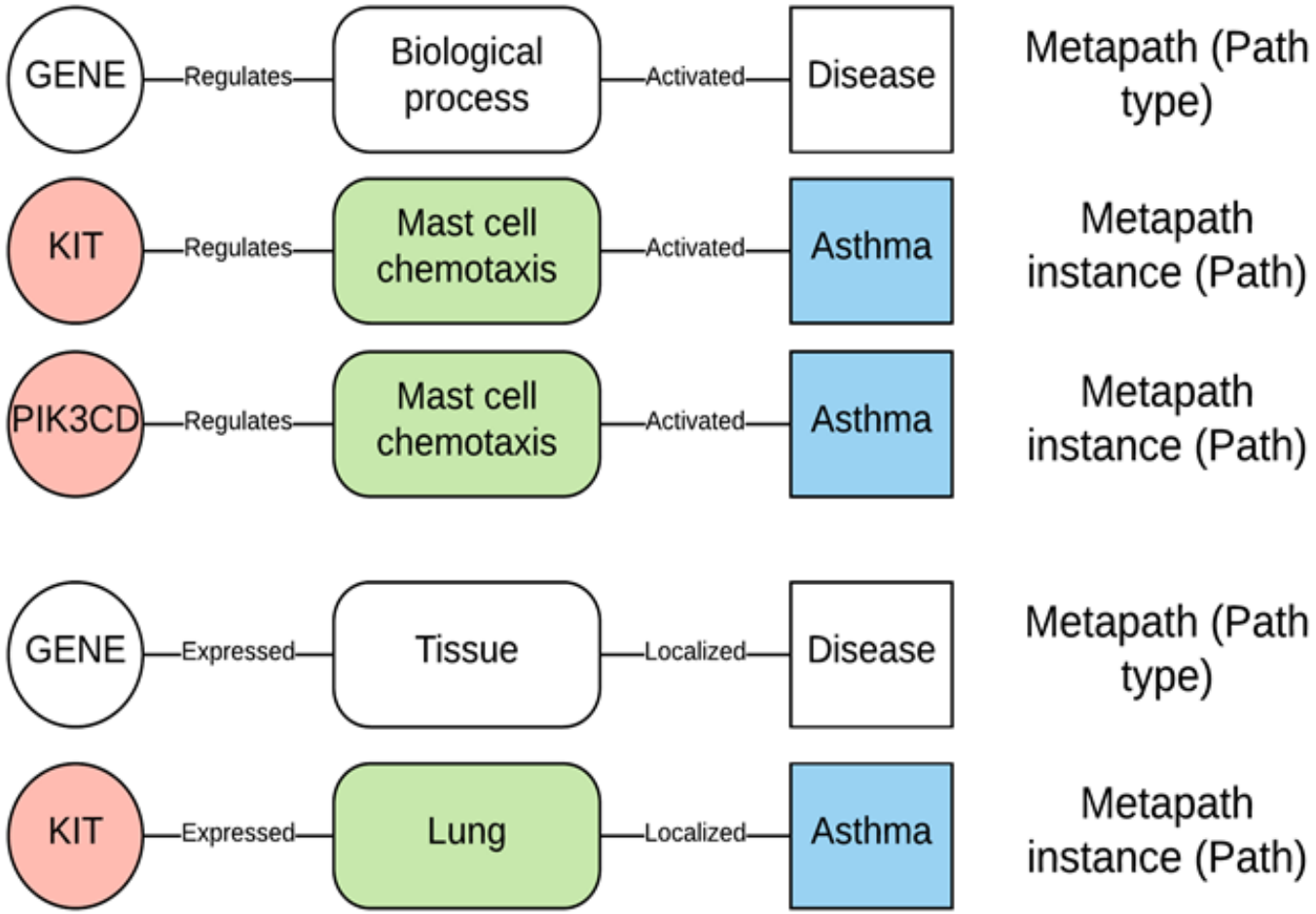
- IDG KMC2 seeks knowledge gaps across the five branches of the "knowledge tree":
- Genotype; Phenotype; Interactions & Pathways; Structure & Function; and Expression, respectively.
- We can use biological systems network modeling to infer novel relationships based on available evidence, and infer new "function" and "role in disease" data based on other layers of evidence
- Primary focus on **Tdark** & **Tbio**

O. Ursu, T Oprea et al., IDG2 KMC

2/1/18 revision

IDG

THE METAPATH APPROACH



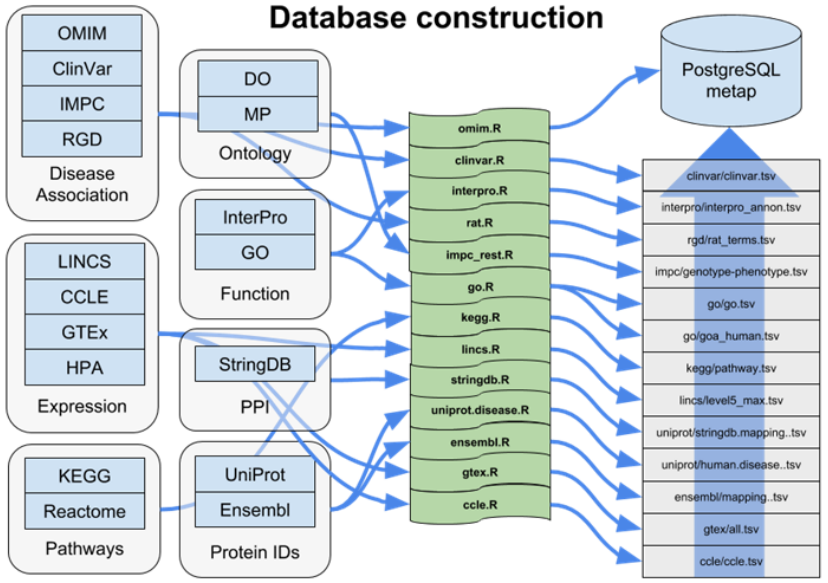
- a *meta-path* is a path consisting of a sequence of relations defined between different object types (*i.e.,* structural paths at the meta level)
- Our metapaths encode type-specific network topology between the source node (e.g., Protein) and the destination node (e.g., Disease).
- This approach enables the trans-formation of assertions/evidence chains of heterogeneous biological data types into a ML ready format.

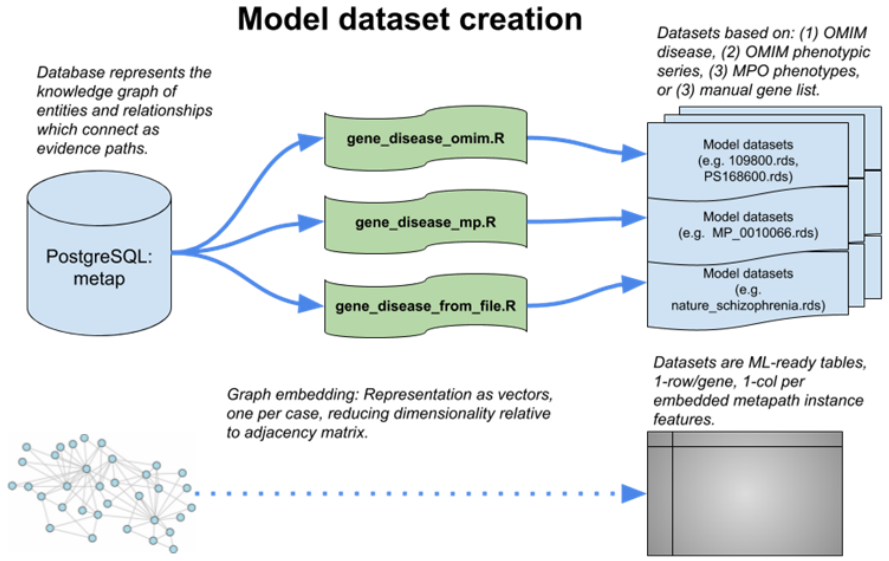
Similar assertions or evidence form metapaths (white).
Instances of metapath (paths) are used to determine the strength of the evidence linking a gene to disease/phenotype/function.

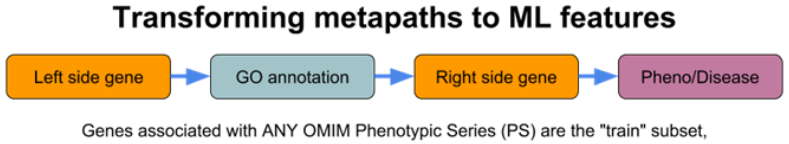
G. Fu et al., *BMC Bioinformatics* **2016**, 17:160 is an early example for drug discovery

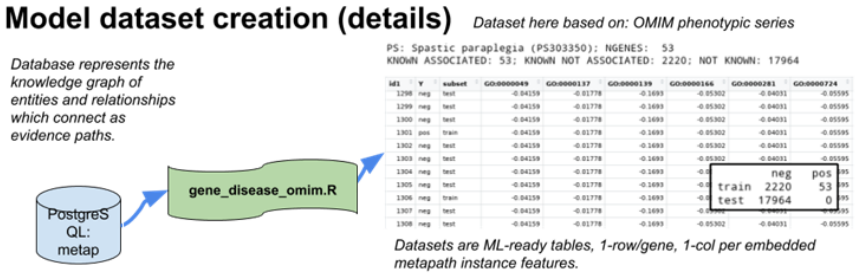
IDG

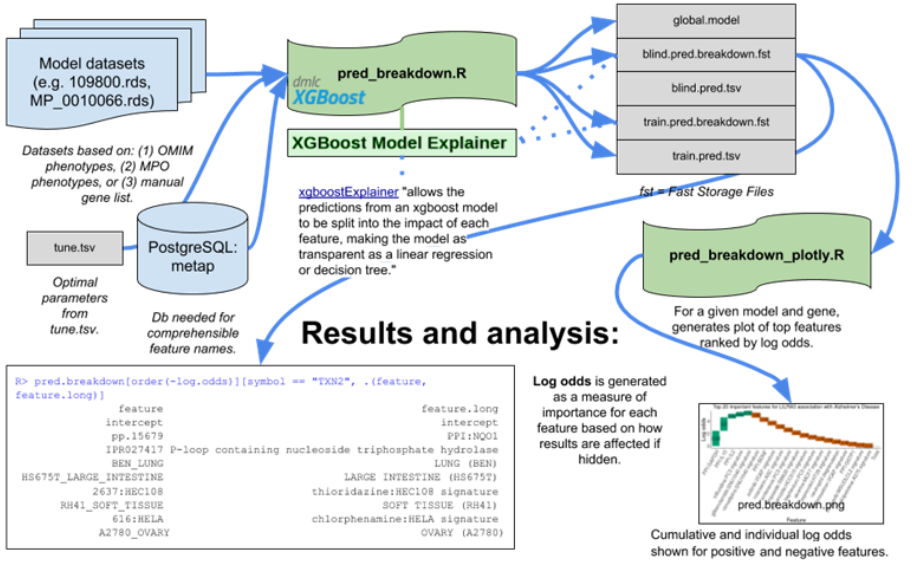
METAPATH/XGBOOST ML WORKFLOW











All datasets are merged, via R scripts, into a PostgreSQL. Python under development.

Graph embedding transforms evidence paths into vectors, converting data into matrices.

Input genes are positive labels. OMIM (not input) are negative labels (we prefer *true negatives* where possible).

XGBoost runs 100 models. The "median model" (AUC, F1) is then selected for analysis and prediction to avoid overfitting.

J.J. Yang, P. Kumar, D. Byrd et al., IDG2 KMC

10/15/19 revision

IDG

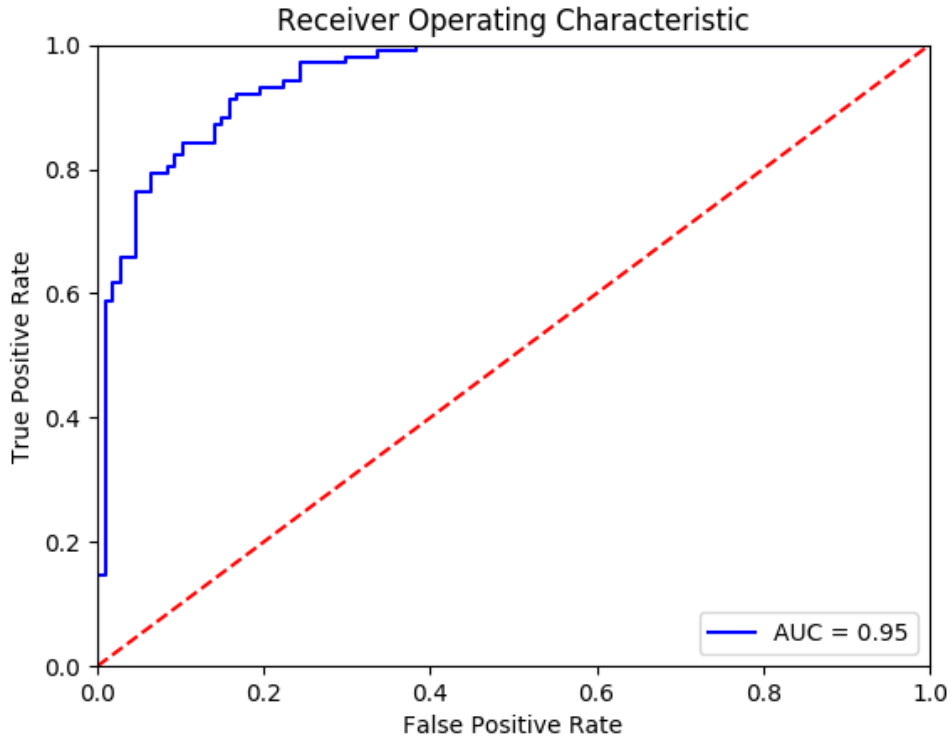
METAPATH / XGBOOST MODEL INPUT

- PHIPSTER ACE2 (experimental) plus 25 other predicted proteins
- CD147 (experimental)
- 71 proteins (mass proteomic pull-down, Krogan et al paper)
- Total 98 positives
- Another 120 negatives from the Krogan paper
- 6 models built, based on variations of this input

IDG

METAPATH / XGBOOST MODEL OUTPUT

- 986 proteins were predicted with "high confidence" by the 6 models
- 136 are predicted by 3 or more models.
- 99 of the 3x predicted proteins were Tbio/Tchem/ Tclin
- These were used in combination with the input proteins plus the viral proteins from Krogan et al to examine the network models.



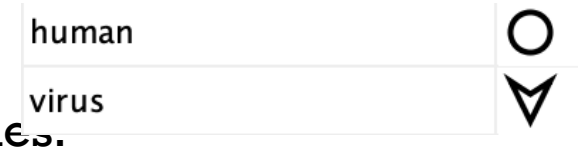
NETWORK WORKFLOW

- Data consolidation
- Building the core network (analysis/network/COVID_19_Merged_Virus_Human_PPI_Network.cys)
  - Virus-protein PPIs extracted
    - Preprint: as is
    - P-HIPSTER: as is
    - Tudor's AIML: assumed that predictions were made for the spike protein only (?)
  - PPIs merged, deduplicated, annotated by data source
    - Direction of PPIs: virus -> human (for consistency)
  - Node attributes added:
    - NodeType: virus/human
    - Tdl: TCRD v64 Yellow-colored proteins, virus proteins were assigned to value of "virus", rest to "unk" (meaning that I did not have the information, did not check Pharos)
- Assembling the extended network (analysis/network/STRING_Extended_COVID_19_PPI_Network.cys)
  - Human nodes of core network used as query for STRINGApp in Cytoscape
  - STRINGApp settings: min. PPI confidence: 0.90, max interactors: 100
  - Network edges and all attributes were exported
  - Network nodes and edges merged with that of "core network

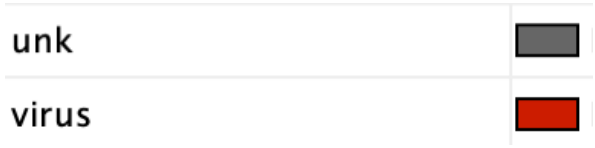
IDG

NETWORK VISUALIZATION

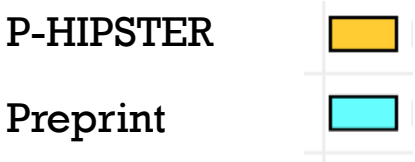
- Custom Cytoscape style created (analysis/network/covid19_styles.xml)
- Nodes:
  - NodeType



  - TDL categories:



- Edges



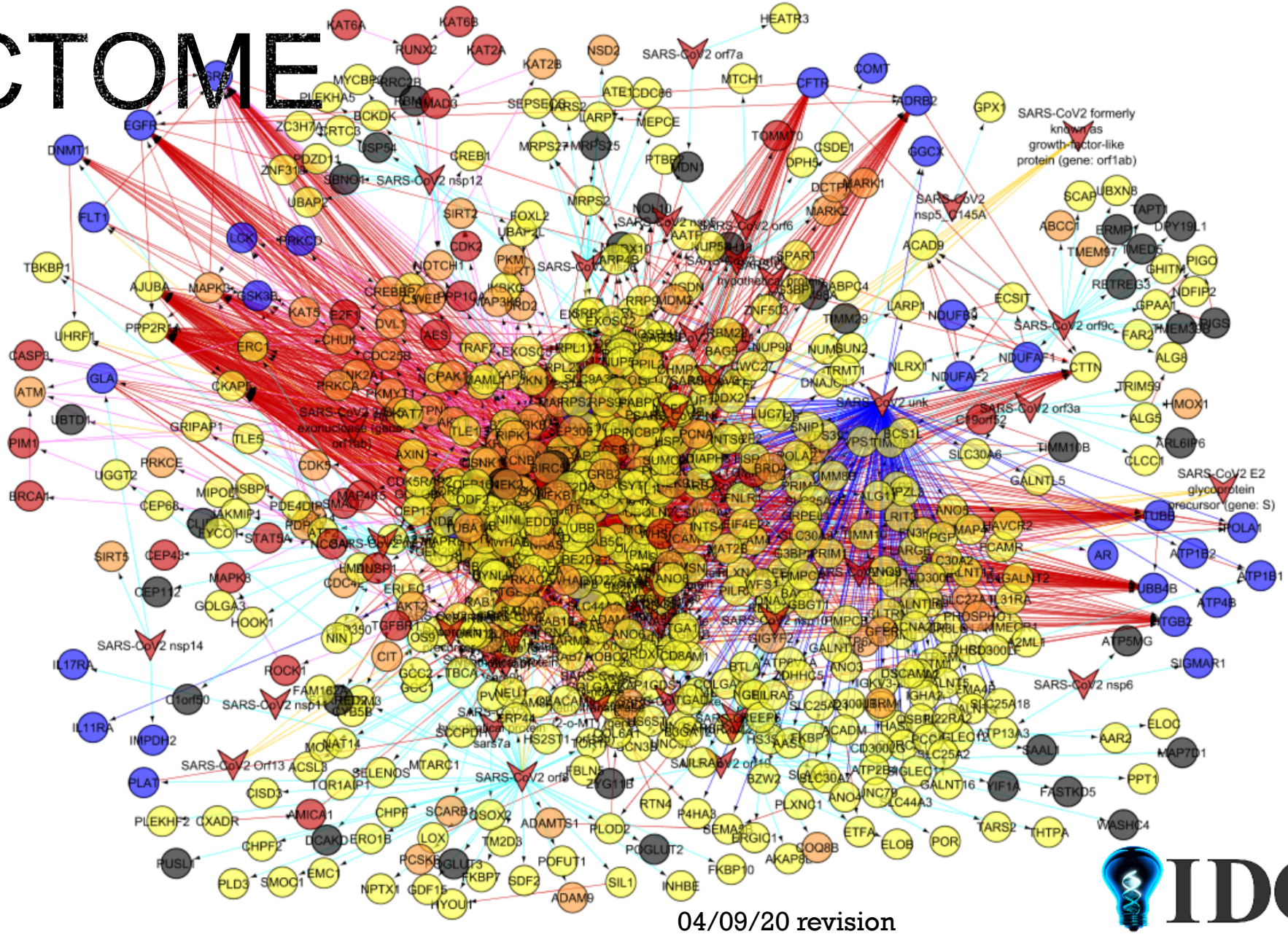
IDG

SARS-COV-2 HUMAN PROTEIN INTERACTOME

SARS-CoV-2 proteins interact with multiple Tclin targets (blue).

Less priority given to ATP- and tubulin- type related targets.

Exploration in progress…



04/09/20 revision

IDG

EARLY RESULTS

- **HDAC2**
- Interacts with SARS-CoV2 nsp5 and with SARS-CoV2-Spike.
- HDAC inhibitors: *"HDIs have a long history of use in psychiatry and neurology as mood stabilizers and anti-epileptics. More recently they are being investigated as possible treatments for cancers, parasitic and inflammatory diseases."*

- HDACs or HDAC inhibitors can be used to treat viral infections including coronavirus infections:
- https://www.ncbi.nlm.nih.gov/pubmed/?term=28780424
- https://www.ncbi.nlm.nih.gov/pubmed/?term=23807710

WAKE UP CALL ON HDAC2

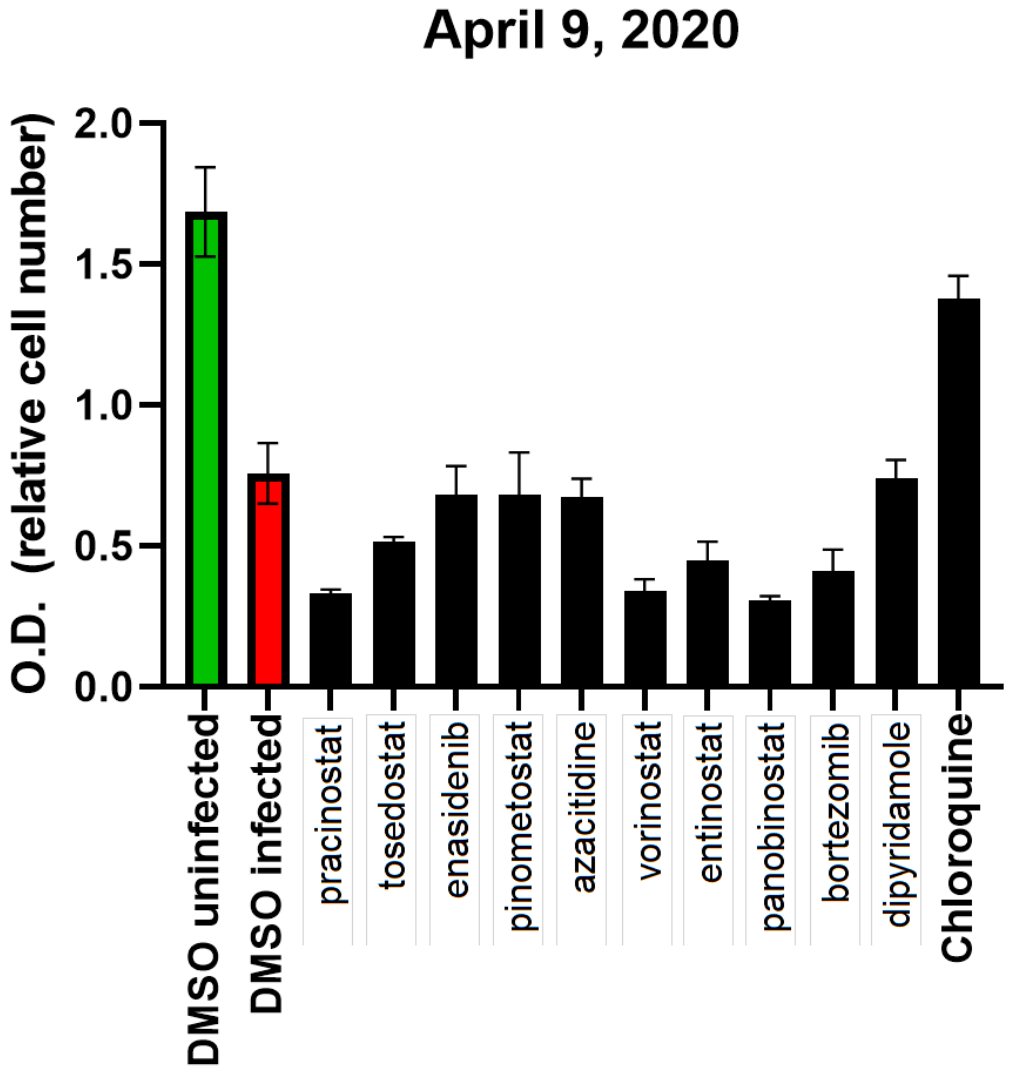
- **From:** Willson, Tim [mailto:tim.willson@unc.edu]
  Sent: Thursday, March 26, 2020 4:38 PM
  Reference 1 says in its abstract "Surprisingly, the antiviral activity of U18666A was suppressed by the histone deacetylase inhibitor (HDACi), Vorinostat"
- Reference 2 says in the abstract "Not surprisingly, viruses have evolved a wide array of mechanisms to subvert HDAC functions."
- I have not read the papers, but sounds like an HDAC inhibitor is likely to promote replication of the virus

- Viruses appear to HDAC activity so they can initiate their own replication. So the association is real, it just favors viruses. Tim suggested the reverse effect by blocking HATs, histone acetyl transferases.

IDG

INDIRECT(?) CONFIRMATION

- Pracinostat, vorinostat and panobinostat appear to *accelerate* the virus-induced killing process
- Alternatively, these drugs directly kill VEROE6 cells.
- This preliminary finding encourages us to further pursue the histone deacetylase hypothesis



04/09/20 revision

IDG

TCLIN AND ASSOCIATED DRUGS

PreyGene	Drug	Test
XPO1	**selinexor**	++++
IDH2	enasidenib	+++
GLA	migalastat	++
IMPDH2	mycophenolate mofetil	+++
UMPS	oteracil	++++
FDPS	zoledronic acid	++++
PSMB2	bortezomib	++++
NDUFA10	**metformin**	+++++
NDUFB10	metformin	
MT-ND1	metformin	
MT-ND3	metformin	
MT-ND5	metformin	
DNMT1	**azacitidine**	+++++
ITGB1	Natalizumab	++++
SLC29A1	Dipyridamole	+++
CRBN	**lenalidomide**	+++++

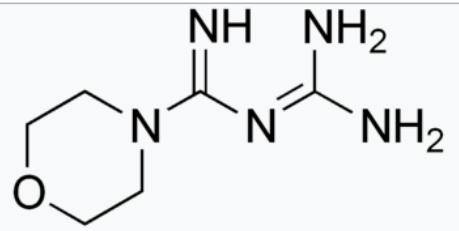
Moroxydine

From Wikipedia, the free encyclopedia

Moroxydine is an antiviral drug that was originally developed in the 1950s as an influenza treatment. It has potential applications against a number of RNA and DNA viruses.[1] Structurally moroxydine is a heterocyclic biguanidine.

It was reported in March 2014 that three kindergartens in two provinces of China had been found to be secretly dosing their students with moroxydine hydrochloride to try to prevent them from becoming ill. The kindergartens are paid only for the days that pupils attend and wanted to ensure that they maximised their earnings.[2]

Moroxydine



Names
IUPAC name *N*-(Diaminomethylidene)morpholine-4-carboximidamide
Other names Moroxidine
Identifiers

Metformin is remarkably similar to moroxydine.
From Krogan et al data, metformin targets ~20 proteins. Other potential mechanisms of action to be explored (based on the summary Table).

04/09/20 revision

IDG

SMARTGRAPH ANALYSIS OF HATS AND COVID-19 IMPLICATED PROTEINS

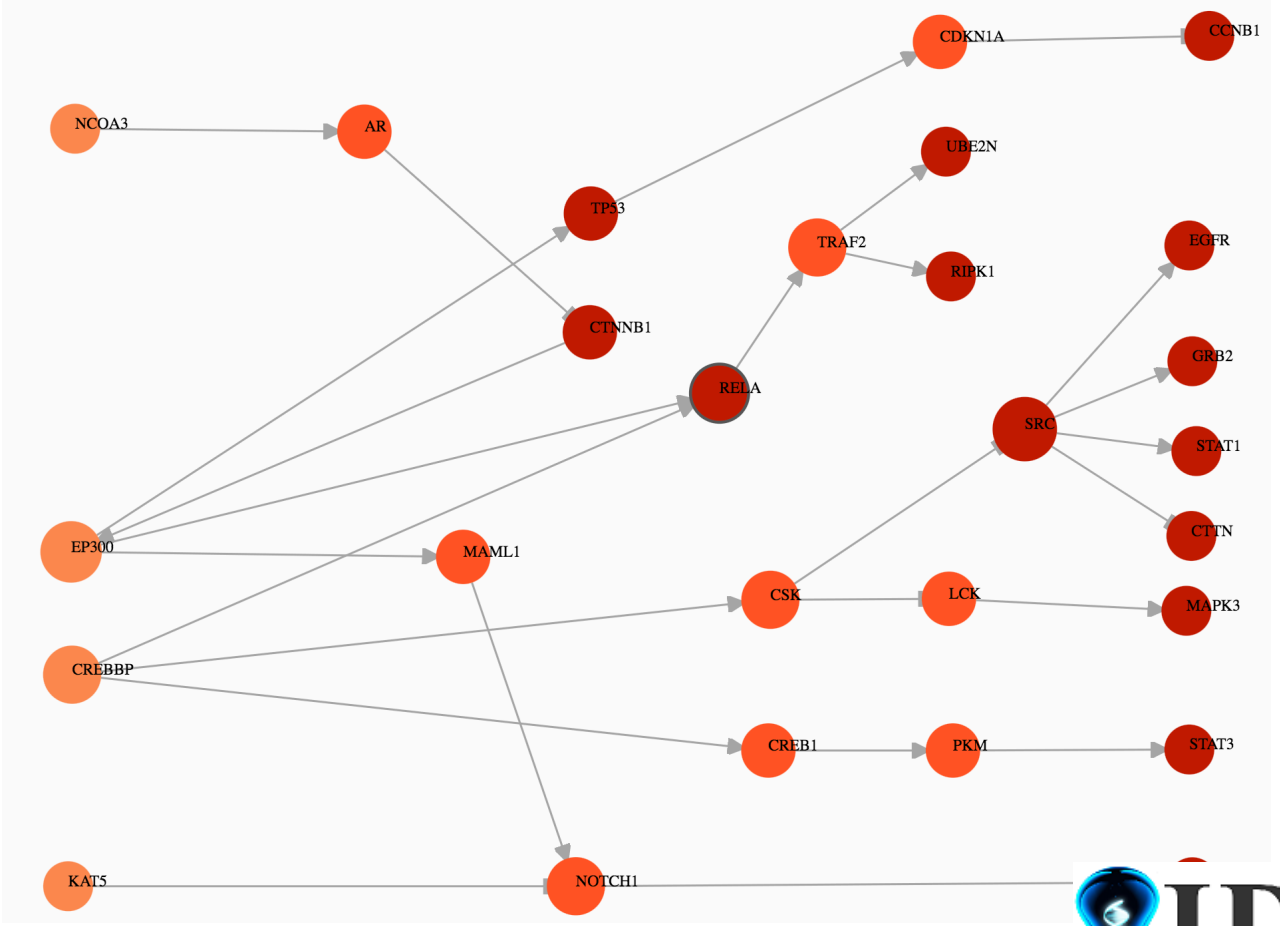
- SmartGraph available at: https://smartgraph.ncats.io/
- Source
  - HATs provided by Tudor
  - Other proteins from the *stringified PPI* network.
  - Parameters
    - distance <= 3
    - confidence >= 0.15
  - Directionality
    - HATs were used as start or end nodes (“forward”/“reverse” networks, respectively)
    - “*Bidirectional HATs PPI*” by merging forward and end networks.

IDG

HATS

EP300	Q09472	Tchem
HAT1	O14929	Tbio
KAT2A	Q92830	Tbio
KAT2B	Q92831	Tchem
KAT5	Q92993	Tchem
KAT6A	Q92794	Tbio
KAT6B	Q8WYB5	Tbio
KAT7	O95251	Tbio
KAT8	Q9H7Z6	Tchem
NAA60	Q9H7X0	Tbio
RBBP7	Q16576	Tbio
CREBBP	Q92793	Tchem
ATF2	P15336	Tbio
TAF1	P21675	Tchem
NAA40	Q86UY6	Tbio
NCOA1	Q15788	Tchem
NCOA3	Q9Y6Q9	Tbio

- One of the HATs, RBBP7, is consistently overexpressed when interacting w/ SARS-CoV2. Two others, HAT1 and NAA40, are relatively under-expressed (data from Krogan et al)
- No other HATs are on the list.
- Graph below: "chemicalizing" the HATs network

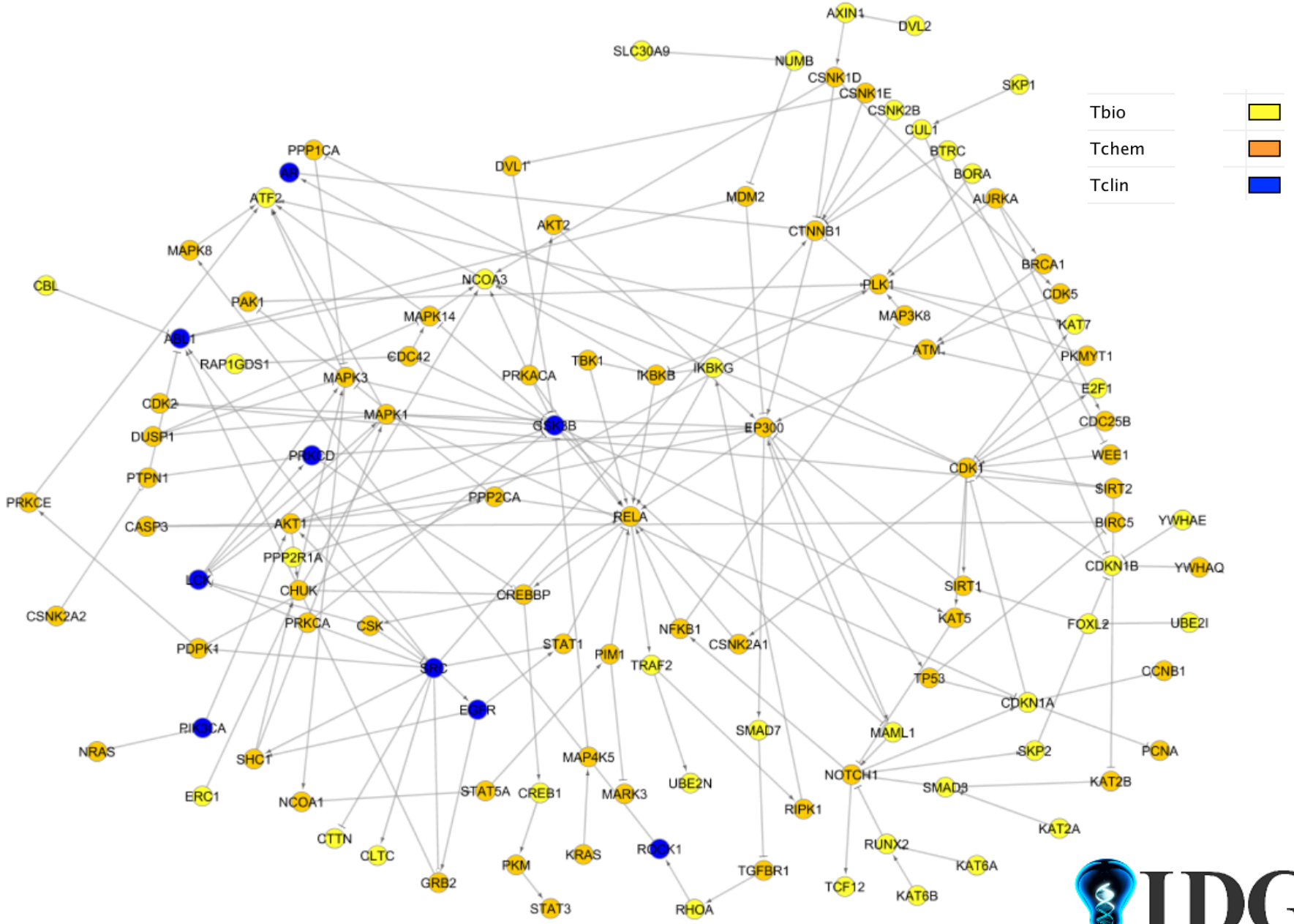


04/02/20 revision

BIDIRECTIONAL NETWORK CENTERE

Merging the "forward" and "reverse" networks illustrates how Tclin end-nodes could be used to modulate HAT function

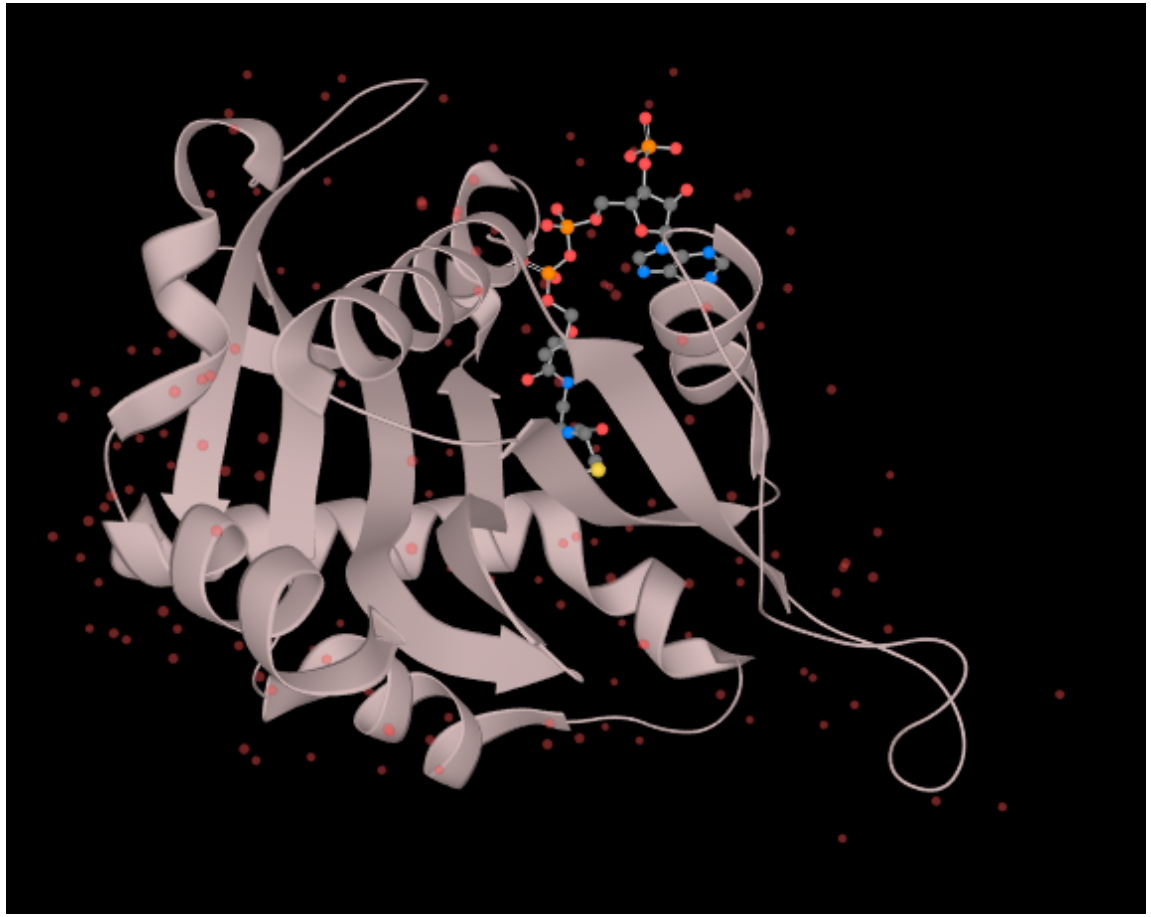
Exploration in progress…



04/02/20 revision

NAA40

4U9V, shown here liganded to acetyl CoA, is very likely to be *ligandable* with an inhibitor. Starting with similarity to the SGRG tetrapeptide could serve as basis for virtual screening.



N-alpha-acetyltransferase 40 specifically mediates the acetylation of the N-terminal residues of histones H4 and H2A (PubMed: 21935442, PubMed:25619998).

- In contrast to other N-alpha-acetyltransferase, has a very specific selectivity for histones H4 and H2A N-terminus and specifically recognizes the 'Ser-Gly-Arg-Gly sequence' (PubMed: 21935442, PubMed:25619998).
- This enzyme is consistently under-expressed when exposed to SARS-CoV-2, suggesting that its inhibition.

Note: Cristian Bologa is running virtual screening for NAA40 & HAT1 (drug repositioning)

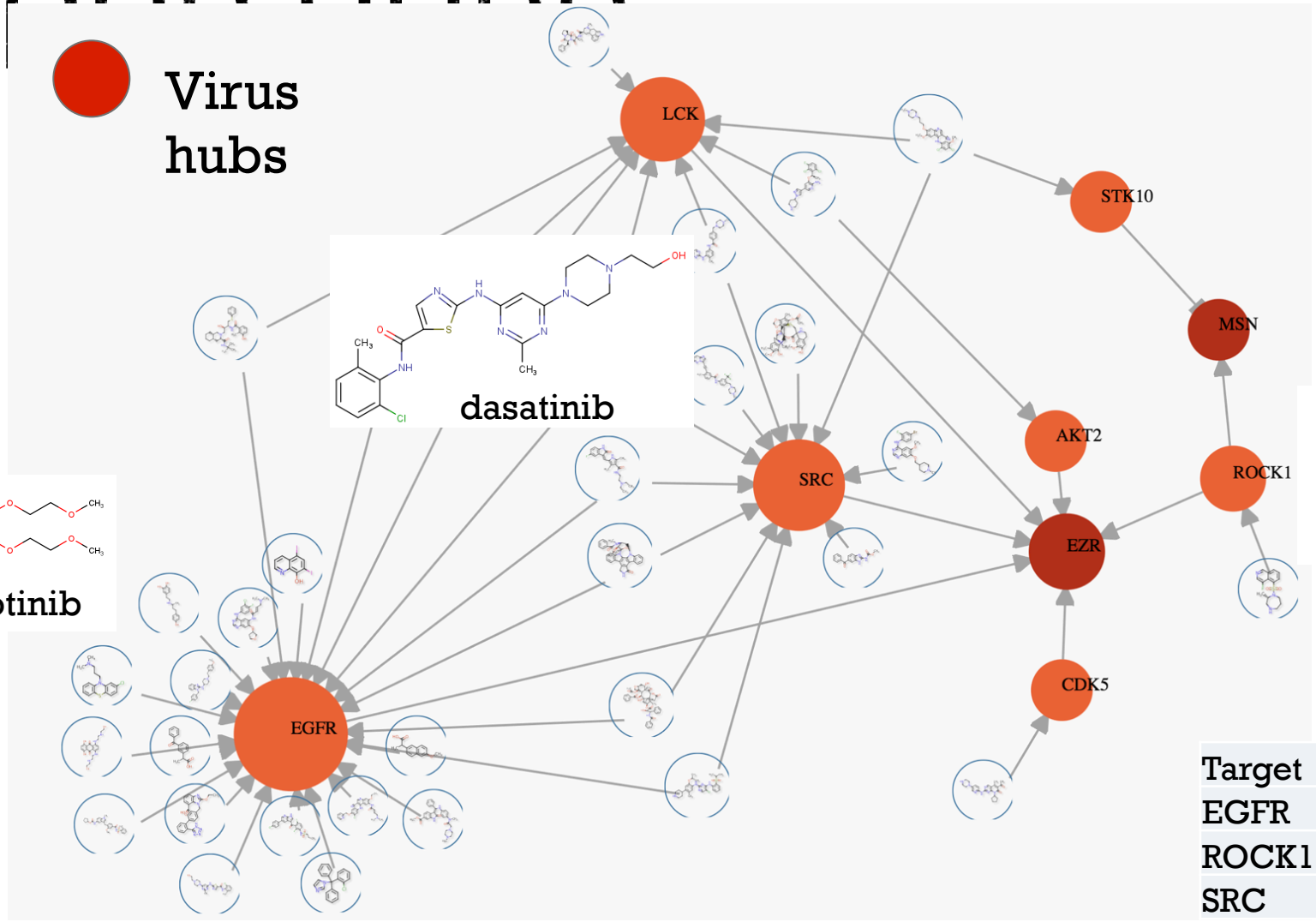
04/02/20 revision

VIRUS-HUMAN BIPARTITE PPI NETWORK

- Rationale: Are there human proteins acting as hubs for virus proteins?
- Derived from merged network
  - Preprint PPIs, Tudor's AIML, P-HIPSTER, STRING, HATs SmartGraph
  - All human PPIs removed, but virus-human PPIs retained
  - ("SARS-CoV2 unk" artificial node also removed)
- "High-degree" (>=3) human proteins in the resultant bipartite network hypothesized to be hubs, thus of potential importance.
- "Virus hubs": **ISG15, UBC, UBB, EZR, NEDD8, UBA52, UBD, MSN, RAD23B**
- Network: analysis/network/Bipartite_COVID-19_PPI.cys
- "CytoHubba" plugin utilized for analysis

IDG

DTIS AND PPIS IMPLICATED BY VIRUS HUBS



Target	Drug	Test
EGFR	erlotinib	++++
ROCK1	fasudil	+++
SRC	dasatinib	++++

04/09/20 revision

IDG

SARS-COV-2 INTERACTOME: TCLIN/TCHEM

- **Tclin**
- There are 169 significant (Fold change above 10) between 23 viral and 56 human proteins. Some occur multiple times, and are likely to be crucial in the way the virus subverts intracellular machinery.
- These are currently evaluated for potential repurposing.
- **Tchem**
- The Spike protein only had one significant (Fold change above 10) Tclin target, so we added 8 Tchem proteins for Spike and 6 for E-protein.
- Looking for drugs hitting these targets as well.

- *Note:* Giovanni Bocci is virtual screening viral targets for drug repositioning

IDG

VIRTUAL SCREENING AGAINST SARS-COV-2 VIRAL TARGETS

Drug	CMax (uM)	PDB Template
tegaserod	0.01	5E6J
triamterene	0.33	6NUR
meloxicam	5.41	5E6J|6NUR
ibuprofen	295.71	5E6J|6NUR
naproxen	408.24	5E6J|6NUR
hydrochlorothiazide	0.25	6NUR
baclofen	0.75	5E6J
trimethoprim	4.13	5E6J|6NUR
ethambutol	17.13	5E6J|6NUR
cidofovir	70.20	5E6J|6NUR
ixazomib	0.17	5E6J
safinamide	3.31	5E6J|6NUR
avibactam	55.04	6NUR

Start: 3,981 small molecules drugs from DrugCentral. Enumerated tautomeric and protomeric forms available at pH 7.4 with minimum abundance of 25% (total, 6057 structures). These were docked into the main SARS-CoV targets:
5E6J: SARS Coronavirus Papain-like Protease
6NUR: SARS-CoV nsp12 polymerase

Note: BDDCS Class 2 and 3 (matching antivirals) present in screening libraries in-house, with half-life of 2 hours or more, and with known CMax, were selected and shown in this Table (the last 3 lack BDDCS category).

04/13/20 revision

IDG

TAKE HOME MESSAGE

DRUG REPOSITIONING FOR COVID-19 IS AN ITERATIVE EXERCISE

Experiments need to inform future experiments

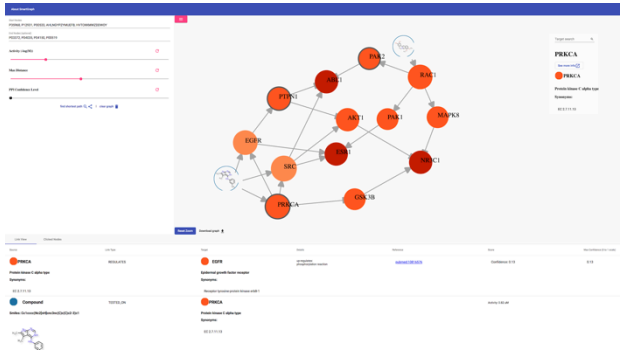
4/14/20 revision

IDG

WHAT CAN WE IMPROVE?

DATA INTEGRATION VIA GRAPH DATABASE (NEO4J)

- SmartGraph has a Neo4j backend with a simple yet powerful data structure.



Zahoranszky-Kohalmi *et al.*, *J Cheminform* 12, 5 (2020)

https://smartgraph.ncats.io/

- Focused on COVID-19: 'Holistic network' as core.
- Needs a more comprehensive PPI data.
- Needs additional layer of biomedical data:
  - Metabolome
  - IDG/Pharos: TDLs, tissue expression data, etc.
  - OMICs
  - Pharmacological action (NCATS - Inxight Drugs, UNM TID - DrugCentral)

NOTE ON SCREENING DRUGS FOR REPOSITIONING

- A Prestwick library screen (pre-print) suggests metformin (VeroE6 cells, SARS-CoV-2 strain BavPat1, 10 μM drug concentration) is inactive.
- Metformin maximum recommended therapeutic dose in man is 2500 mg/day
- https://www.mayoclinic.org/drugs-supplements/metformin-oral-route/proper-use/drg-20067074
- MW 165.62 for metformin HCl formulation
- Average 70 L human body and 1250 mg dose ➔ *test metformin at 100 μM*

- Screening concentration of approved drugs needs to take into account MRTD & Cmax where available
- Of 845 drugs, 101 drugs >= 20 μM, and 169 drugs <= 0.1 μM.

IDG