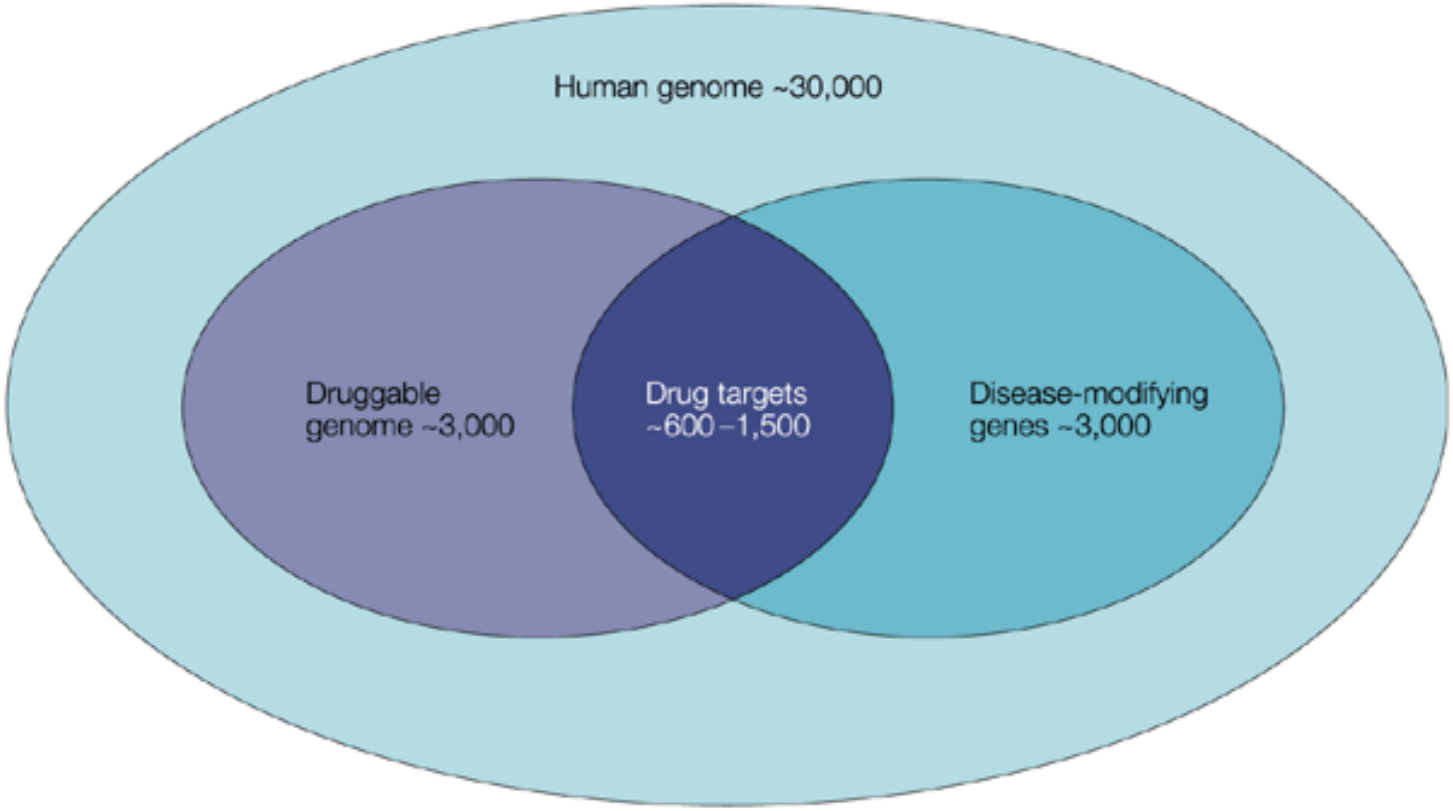
Illuminating the Druggable Genome Program

Council of Councils Meeting
January 25-26, 2024

Karlie Sharma, PhD
Program Director, Office of Drug Development Partnership Programs
National Center for Advancing Translational Sciences

National Institutes of Health
Office of Strategic Coordination–The Common Fund

Illuminating the Druggable Genome (IDG) Program: What is the Druggable Genome?



Druggable Genome - the subset of the human genome that expresses proteins potentially able to bind drug-like compounds.

Hopkins and Groom, NRDD 2002

Illuminating the Druggable Genome (IDG) Program: What is the Druggable Genome?

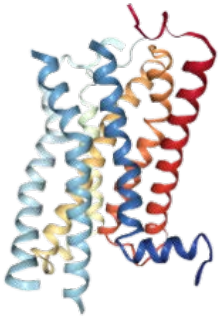
- While the number of proteins in the druggable genome is upwards of 4,500, the existing clinical pharmacopeia is represented by only a few hundred targets, leaving a huge swath of biology that remains unexploited.
- Goal of the IDG Program – catalyze research to improve our understanding of the properties and functions of proteins that are currently not well studied within commonly drug-targeted protein families.

Definition of an Understudied Protein

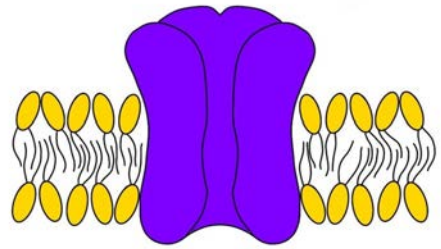
- Understudied Proteins Identified by:
  - Few or no publications
  - Lack of R01 Funding

- Three protein families:
  - Ion Channels
  - G Protein Coupled Receptors
  - Kinases

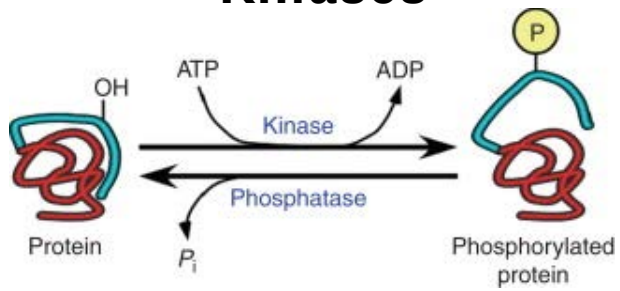
GPCRs



Ion Channels



Kinases



IDG Program Timeline and Budget

2014	2015	2016	2017	2018	2019	2020	2021	2022	2023

Pilot Phase ($25.5M)

Goals:

- Adapt scalable technology platforms for IDG protein families.
- Develop a knowledge/information management platform.

Outputs:

- Core understudied protein datasets.
- Understudied protein interrogation platforms.
- Predictive algorithms.
- Pharos, a protein exploration search engine.

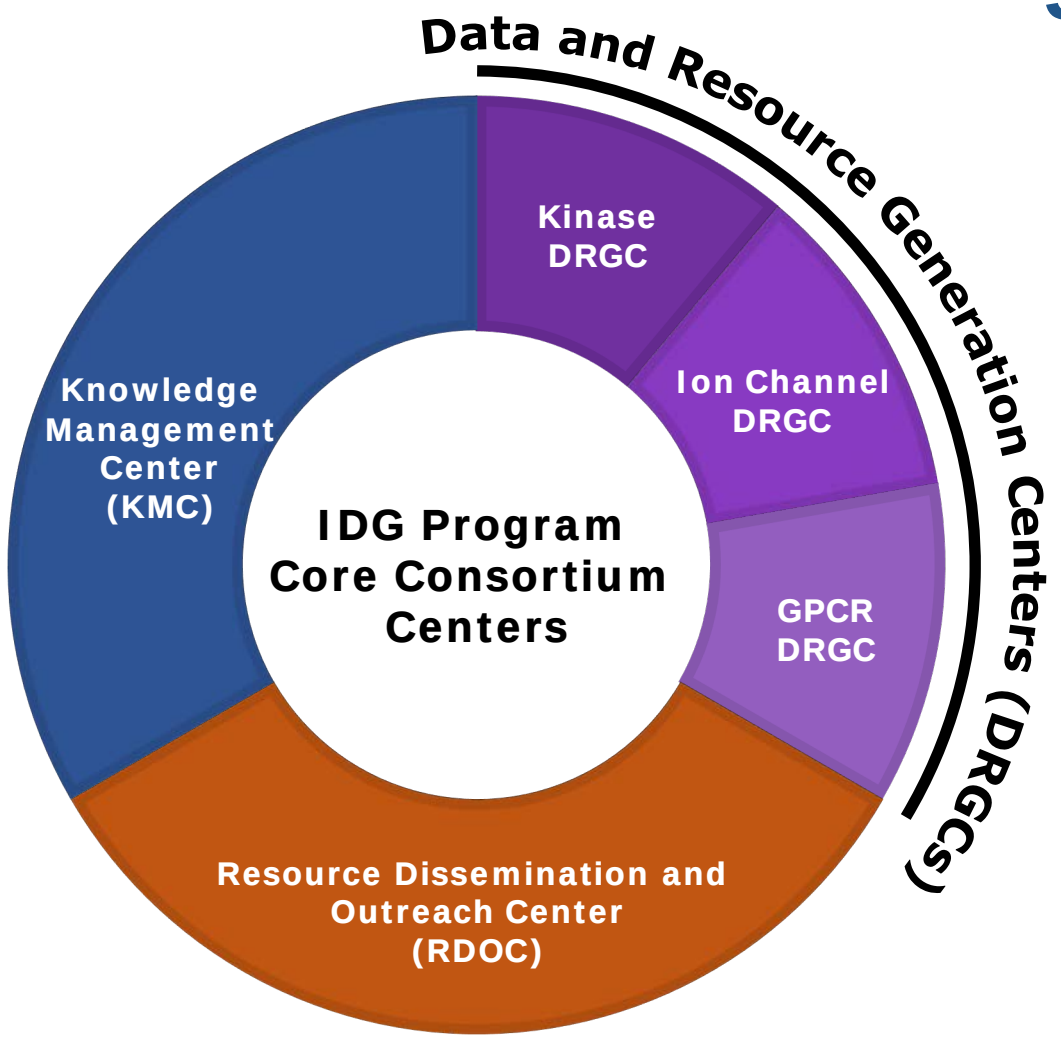
Implementation Phase ($71M)

Goals:

- Identify phenotypes of understudied proteins.
- Provide reagents and tools.
- Create an enriched, minable knowledge base (Pharos).

- **Outputs:**
- Pharos interface and underlying database.
- Reagents available through several repositories.
- Improved understanding of nearly 100 understudied proteins and their role in human disease
- Updated and improved Expanded high-throughput platforms for interrogating the understudied genome.

IDG Program: Implementation Phase Consortium Structure

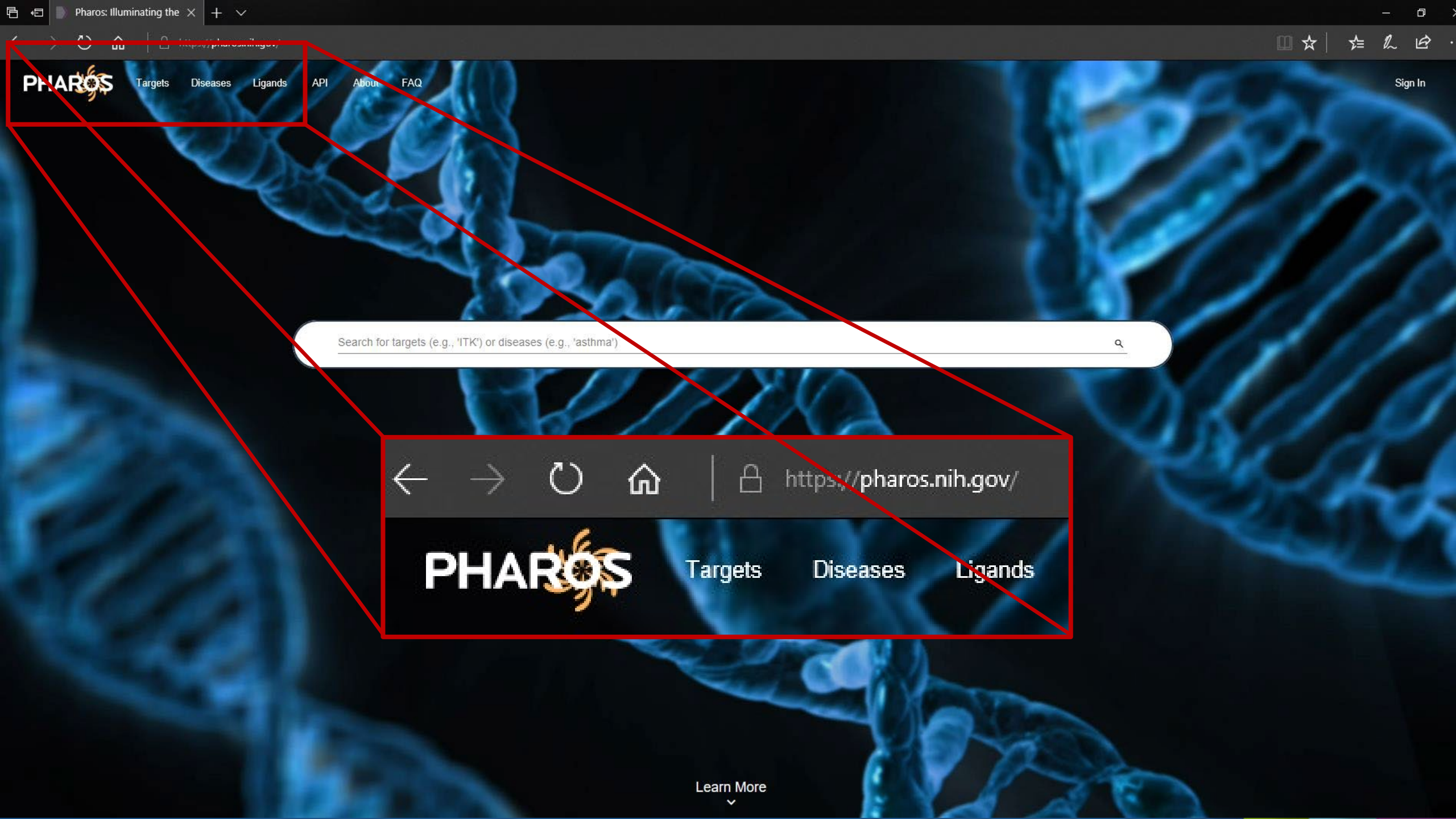


Cutting Edge Informatics Tools (CEIT) Awards - NOFO issued in 2018 and 2021 to deploy tools to enhance the community's ability to process, analyze, and visualize data around the understudied proteins.

R03 Pilot Projects – NOFO issued each year from 2018-2022 to support the generation of preliminary data and tools around eligible understudied protein(s) to:

- elucidate function in the context of human disease
- support R01 applications and/or drug discovery projects

Commercializing Understudied Proteins from the Illuminating the Druggable Genome Project (SBIR and STTR) – NOFO to initiate early research leading to the commercialization of assays or products.


Pharos: Illuminating the
PHAROS
Targets
Diseases
Ligands
API
About
FAQ
Sign In
Search for targets (e.g., 'ITK') or diseases (e.g., 'asthma')
https://pharos.nih.gov/
PHAROS
Targets
Diseases
Ligands
Learn More

Pharos: Target Development Levels

Tdark	These are targets about which virtually nothing is known. They do not have known drug or small molecule activities and satisfyT two or more of the following criteria: very low publication number (usually <5), gene function annotations <3 and <50 antibodies available.
Tbio	These targets do not have known drug or small molecule activities but do have some publications and data. They satisfy all conditions for Tdark and satisfy one or more of the following criteria: known molecular-level activity and/or a known molecular relationship between a genetic variation and phenotypic expression.
Tchem	These targets have at least one bioactive, drug-like compound with an activity cutoff of <30 nM and satisfy all preceding conditions (Tdark and Tbio).
Tclin	These targets have at least one approved drug and satisfy all preceding conditions (Tdark, Tbio and Tchem).

Tclin

TNF

Tumor necrosis factor

Jump to section:

Protein Summary

Uniprot Accession IDs

P01375 O43647 Q9P1Q2 Q9UIV3

Gene Name

TNF

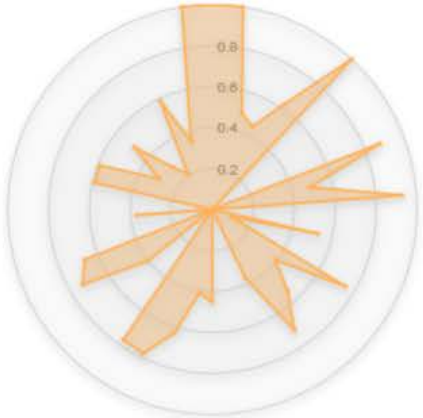
Ensembl ID

ENST00000376122	ENSP00000365290	ENSG00000204490
ENST00000383496	ENSP00000372988	ENSG00000206439
ENST00000412275	ENSP00000392858	ENSG00000228321
ENST00000420425	ENSP00000410668	ENSG00000228849
ENST00000443707	ENSP00000389492	ENSG00000230108
ENST00000448781	ENSP00000389490	ENSG00000223952
ENST00000449264	ENSP00000398698	ENSG00000232810

Symbol

TNFA TNFSF2 DIF TNFA TNFSF2 TNLG1F TNF-alpha

Illumination Graph



Knowledge Table

Most Knowledge About	Knowledge Value (0 to 1 scale)
biological process	1
biological term	1
chemical	1
virus perturbation	1
drug	0.93

IDG Development Level Summary

Tdark

These are targets about which virtually nothing is known. They do not have known drug or small molecule activities
- AND -
satisfy two or more of the following criteria:

- Pubmed score: 49978.25 *(req: > 5)*
- Gene RIFs: 4125 *(req: > 3)*

Tbio

These targets do not have known drug or small molecule activities
- AND -
satisfy the preceding conditions
- AND -
satisfy one or more of the following criteria:

- Gene Ontology Molecular Functions: 156
- OMIM Phenotypes: 4

Tchem

Target has at least one ChEMBL compound with an activity cutoff of < 30 nM
- AND -
satisfies the preceding conditions

- Active Ligands: 352

Tclin

Target has at least one approved drug
- AND -
satisfies the preceding conditions

- Active Drugs: 5

TNF

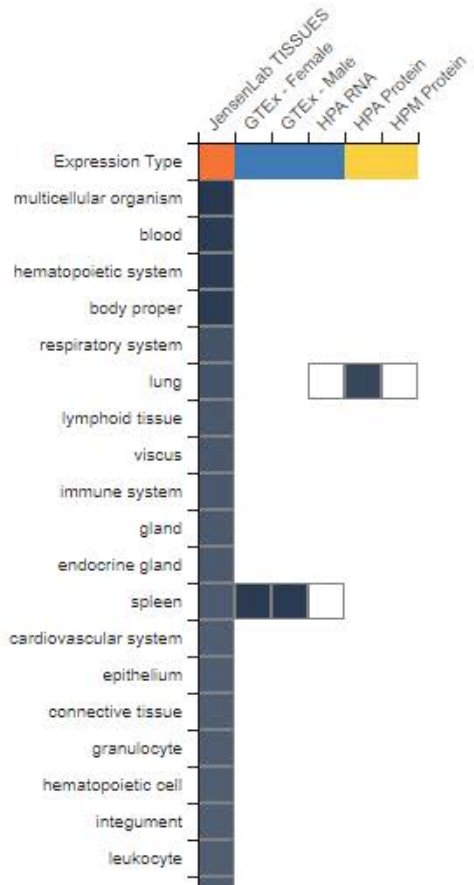
Expression Data (1089 Tissues)

Tissue Search On

Search Uberon Hierarchy

Filter: **none**

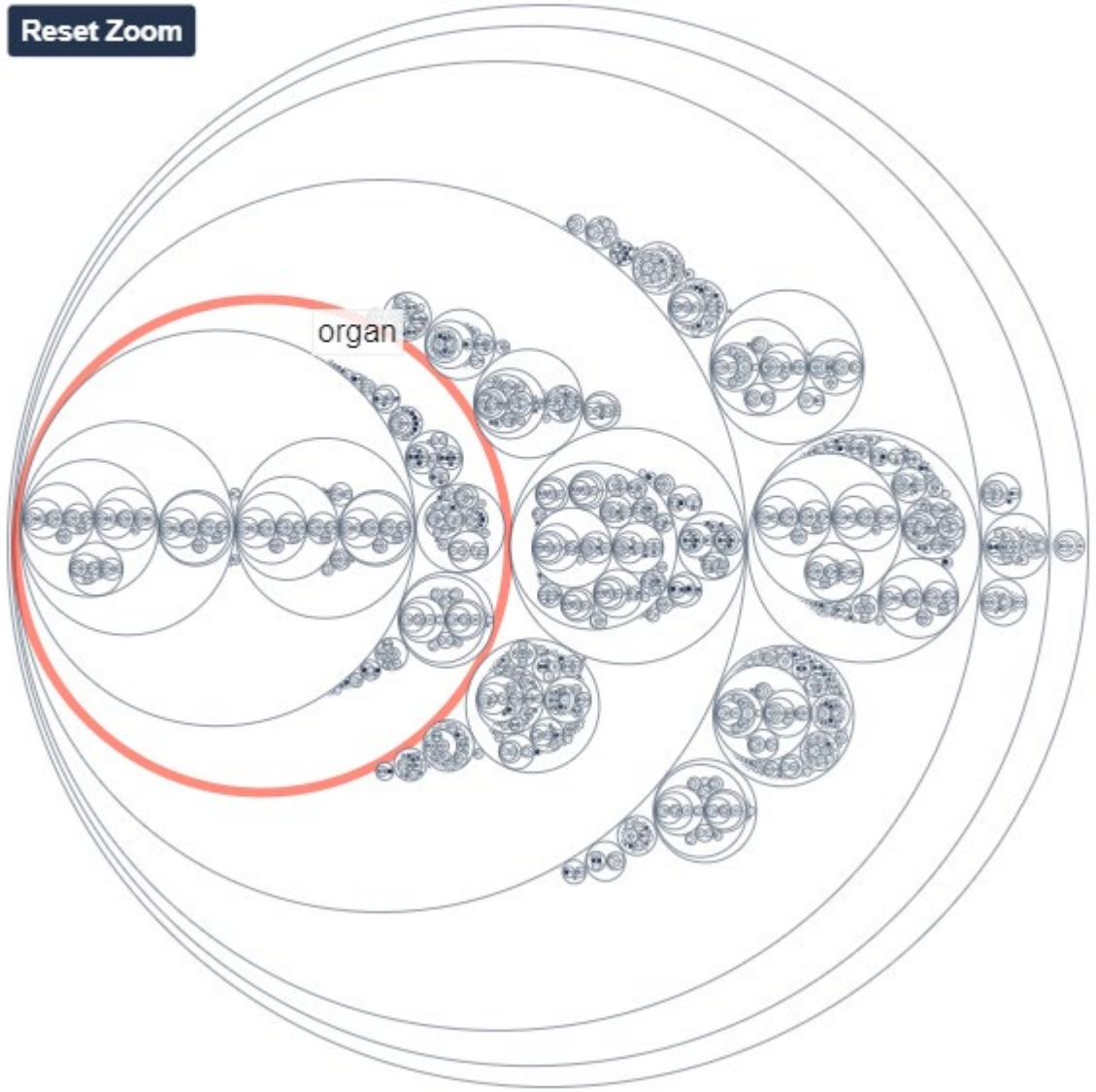
Sort Column: **JensenLab TISSUES**



Circle Plot (5) Anatomogram

root: anatomical entity (UBERON:0001062)

Reset Zoom

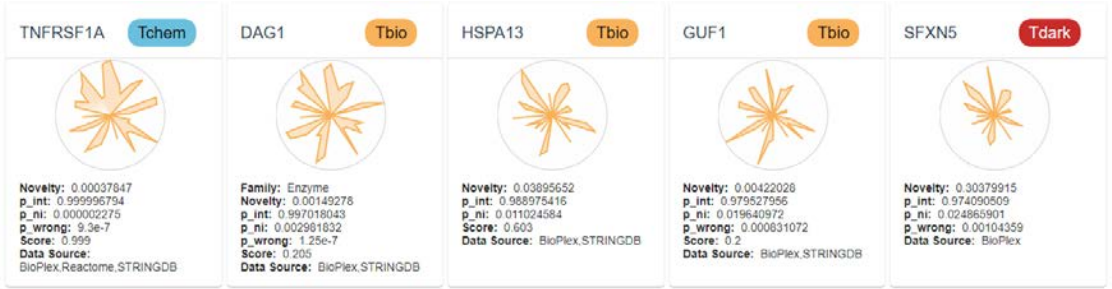


Pharos: Other Notable Features

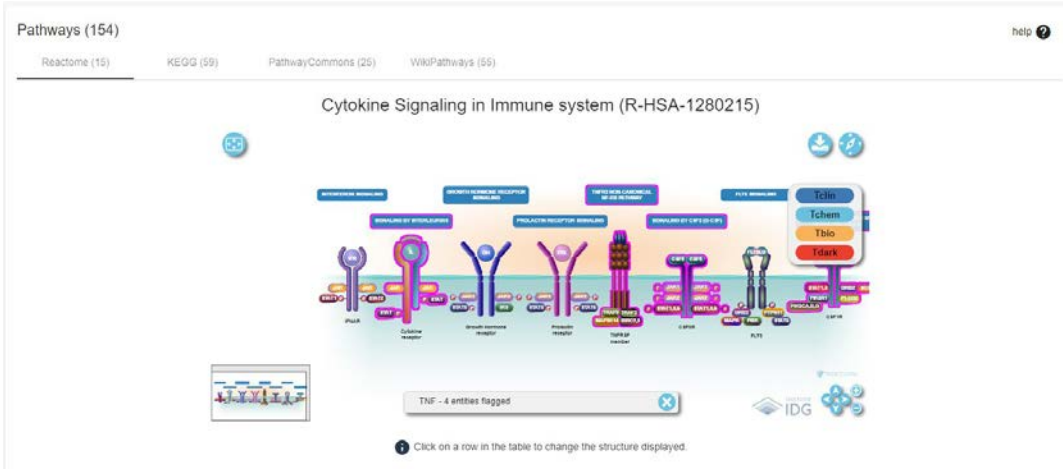
Approved Drugs

Active Ligands

Protein-Protein Interactions



Pathways



Structures

IDG Program Generated Resources

Tdark **PAN3**

PAN2-PAN3 deadenylation complex subunit PAN3 Download

Protein Summary

help

Description

Regulatory subunit of the poly(A)-nuclease (PAN) deadenylation complex, one of two cytoplasmic mRNA deadenylases involved in general and miRNA-mediated mRNA turnover. PAN specifically shortens poly(A) tails of RNA and the activity is stimulated by poly(A)-binding protein (PABP). PAN deadenylation is followed by rapid degradation of the shortened mRNA tails by the CCR4-NOT complex. Deadenylated mRNAs are then degraded by two alternative mechanisms, namely exosome-mediated 3'-5' exonucleolytic degradation, or deadenlyation-dependent mRNA decaping and subsequent 5'-3' exonucleolytic degradation by XRN1. PAN3 acts as a positive regulator for PAN activity, recruiting the catalytic subunit PAN2 to mRNA via its interaction with RNA and PABP, and to miRNA targets via its interaction with GW182 family proteins.

Uniprot Accession IDs

Q58A45 A0N0X1 A1A4Y8 A1A4Y9 B1ALF1 B7Z3W7 Q0D2P2 Q5HYG6 Q5T515 Q5T516 Q5TBA0 Q76E13 Q8NBA6

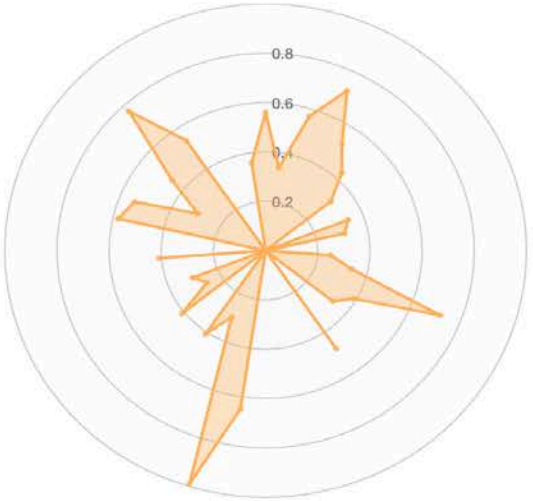
Gene Name

PAN3

Ensembl ID

ENST00000380958 ENSP00000370345 ENSG00000152520

Illumination Graph



Knowledge Table

Most Knowledge About	Knowledge Value (0 to 1 scale)
molecular function	0.99
transcription factor binding site profile	0.77
cell type or tissue	0.72
histone modification site profile	0.72
microRNA	0.65

Protein Classes

help

DTO Classes

Protein / Kinase / Protein Kinase / Other Group / Pan3 Family / Pab-dependent Poly(a)-specific Ribonuclease Subunit Pan3

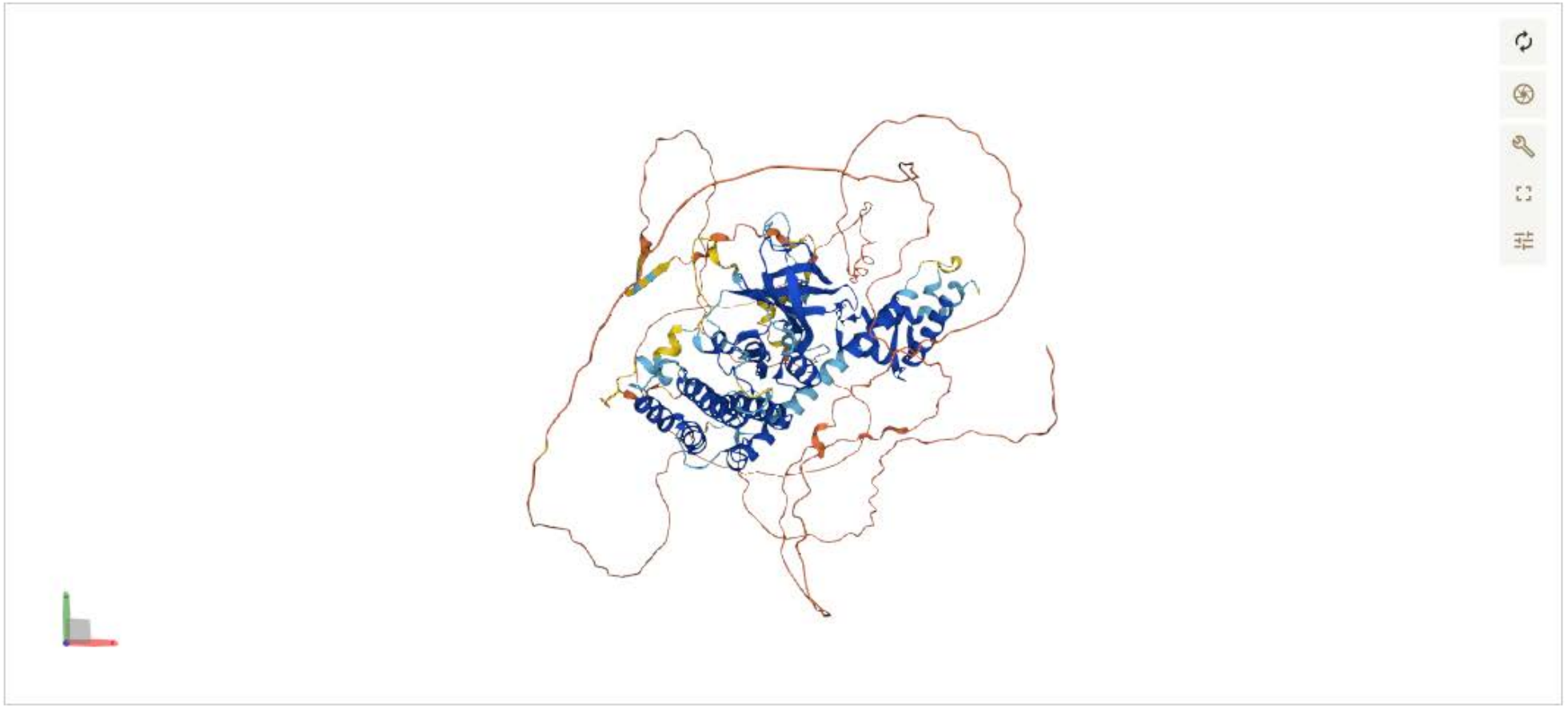
PAN3 (Tdark)

200 400 600 800

Model Confidence:

- Very high (pLDDT > 90)
- Confident (90 > pLDDT > 70)
- Low (70 > pLDDT > 50)
- Very low (pLDDT < 50)

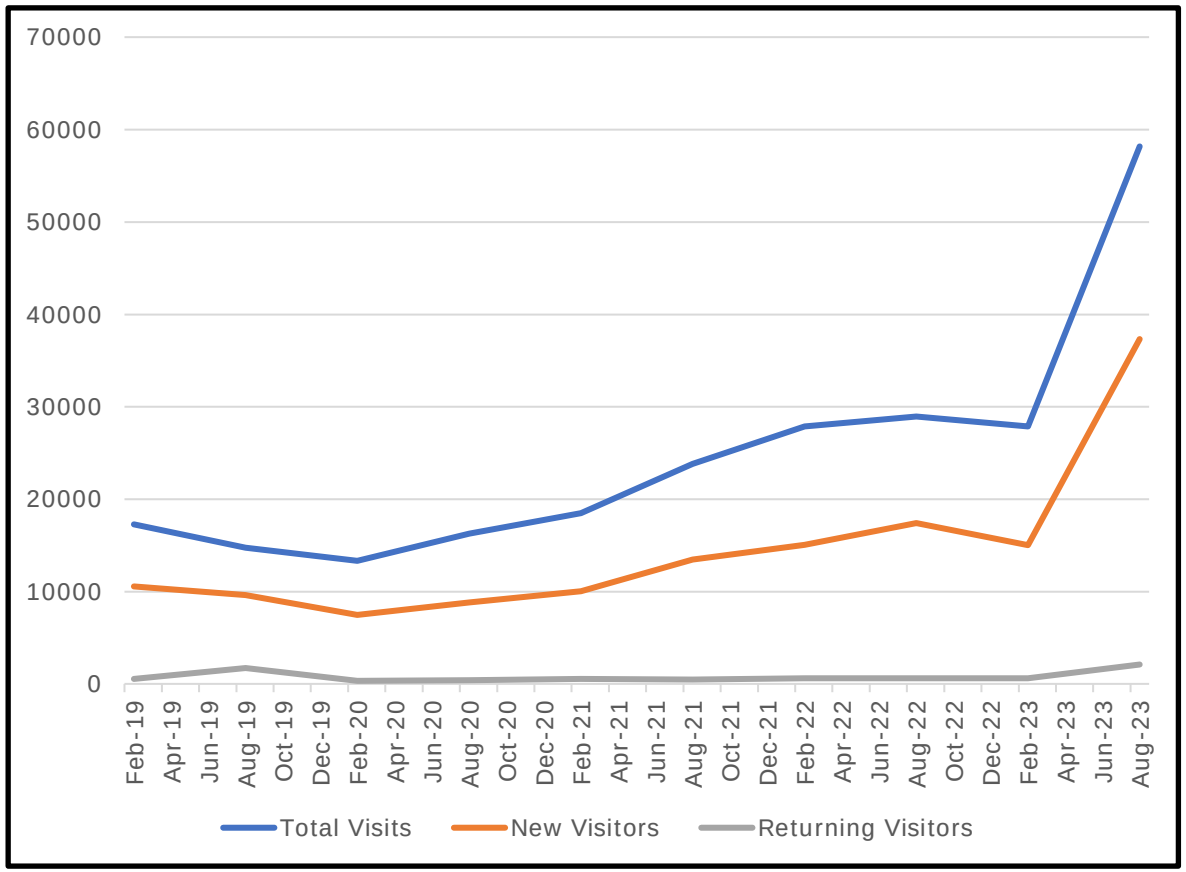
AlphaFold produces a per-residue confidence score (pLDDT) between 0 and 100. Some regions with low pLDDT may be unstructured in isolation.



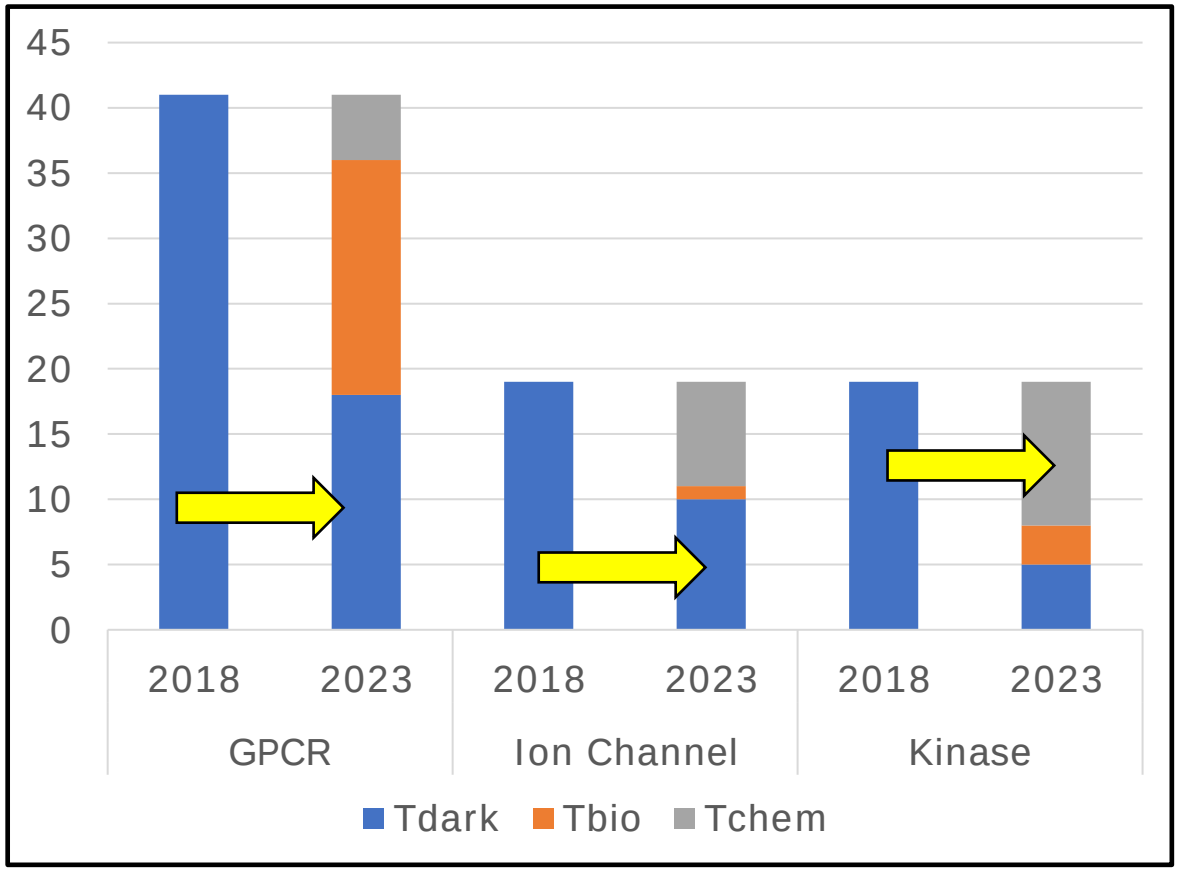
SOURCE -- Select --	IDENTIFIER	METHOD -- Select --	RESOLUTION	CHAIN	POSITIONS	LINKS
AlphaFold	AF-Q58A45-F1	Predicted			1-887	**AlphaFold**

Pharos: Impacts on Evolution of Target Development Levels

Pharos Usage – New and Returning Users



IDG Program Protein Families – Target Development Levels 2018 vs. 2023



Utilization of Pharos for New Opportunities

FONDAZIONE

Telethon

We've been investing in scientific research since 1990

We've funded more than 2700 projects and 1600 researchers to give people affected by a rare genetic condition a new hope.

01 / 02

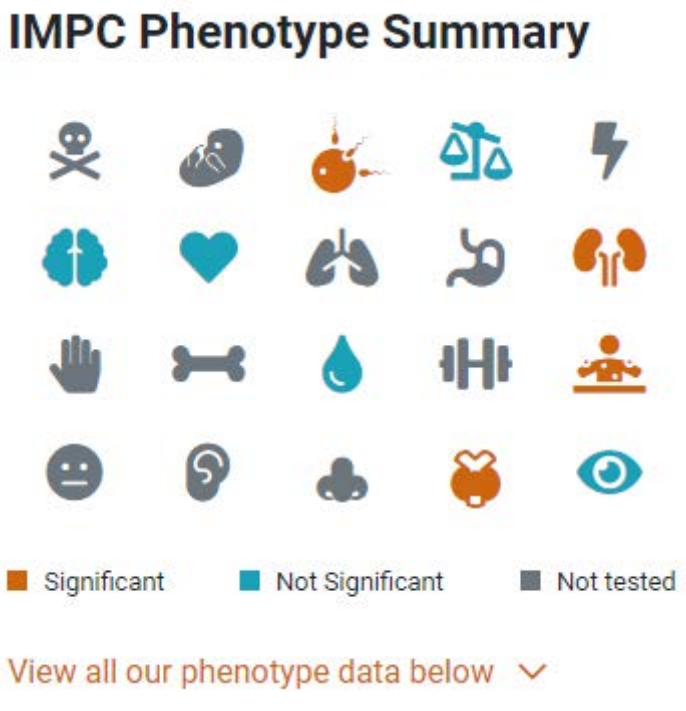
LEARN MORE

NIH National Institutes of Health
Office of Strategic Coordination–The Common Fund

Fondazione CARIPLO

IDG Partners with KOMP2

- IDG partnership with the Knockout Mouse Phenotyping Program (KOMP2) initiated September 2022
- Working together to develop twenty-seven understudied ion channel knockout mouse models



Phenotype	System	Allele	Zyg	Sex	Life Stage
small kidney		Pllp$^{tm1(KOMP)Vlcg}$	HOM	♀♂	Early adult
abnormal kidney morphology		Pllp$^{tm1(KOMP)Vlcg}$	HOM	♀♂	Early adult
enlarged testis		Pllp$^{tm1(KOMP)Vlcg}$	HOM	♂	Early adult
abnormal testis morphology		Pllp$^{tm1(KOMP)Vlcg}$	HOM	♂	Early adult
abnormal skin morphology		Pllp$^{tm1(KOMP)Vlcg}$	HOM	♂	Early adult

R03 Pilot Projects – A Blueprint for Success

R03 Pilot Program: Pilot Projects Investigating Understudied G Protein-Coupled Receptors, Ion Channels, and Protein Kinases

- Over 120 specific dark proteins studied via 98 awards, with over 60 publications to date
- 29 Early Stage and/or New Investigators funded

R03 Pilot Projects – A Blueprint for Success

- Select awardee achievements:
  - R01 projects focused on:
    - Expanding our understanding of the role for the branched-chain a-ketoacid dehydrogenase kinase (BCKDK) in cognition and vocal communication.
    - Interrogation of the NEK family of kinases in cancer (NEK1-11, NEK4-7, 10, 11 considered understudied)
  - R21 awards to:
    - Identify endogenous peptide ligands for orphan GPCRs in order to explore their roles in CNS disorders.
    - Analyze GWAS loci associated with hearing loss – understudied kinases
  - An award to further explore the role of the understudied GPR83 protein in opioid addiction.
  - An award made by the American Heart Association to study a dark kinase in cardiac hypertrophy.

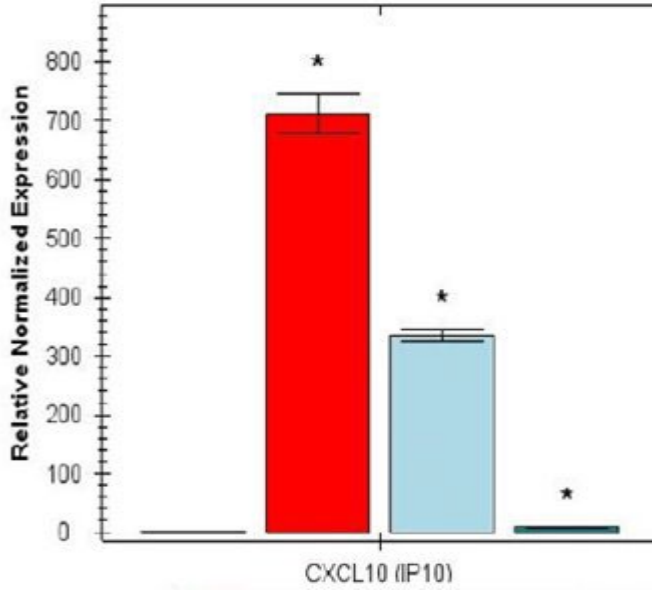
R03 Pilot Projects – A Blueprint for Success

The Atypical Kinase RIOK3 Limits RVFV Propagation and Is Regulated by Alternative Splicing

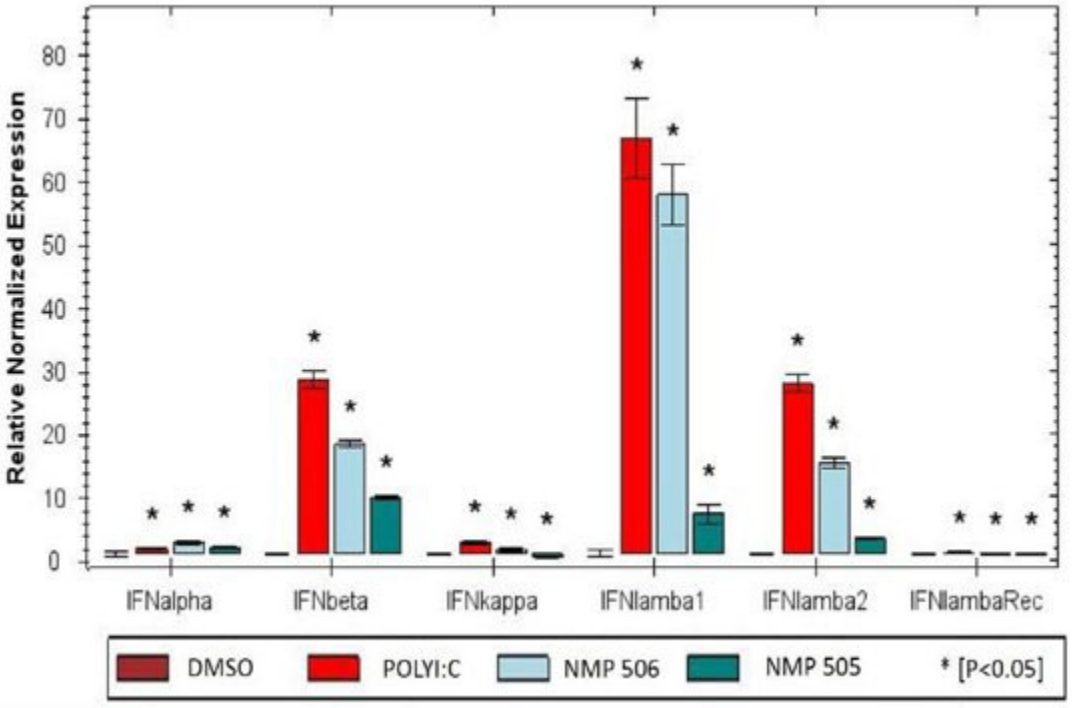
by Katherine E. Havranek [1,†], Luke Adam White [1,†], Thomas C. Bisom [2], Jean-Marc Lanchy [1] and J. Stephen Lodmell [1,3,*]

Viruses **2021**, *13*(3), 367; https://doi.org/10.3390/v13030367

Received: 16 January 2021 / Revised: 16 February 2021 / Accepted: 22 February 2021 / Published: 26 February 2021



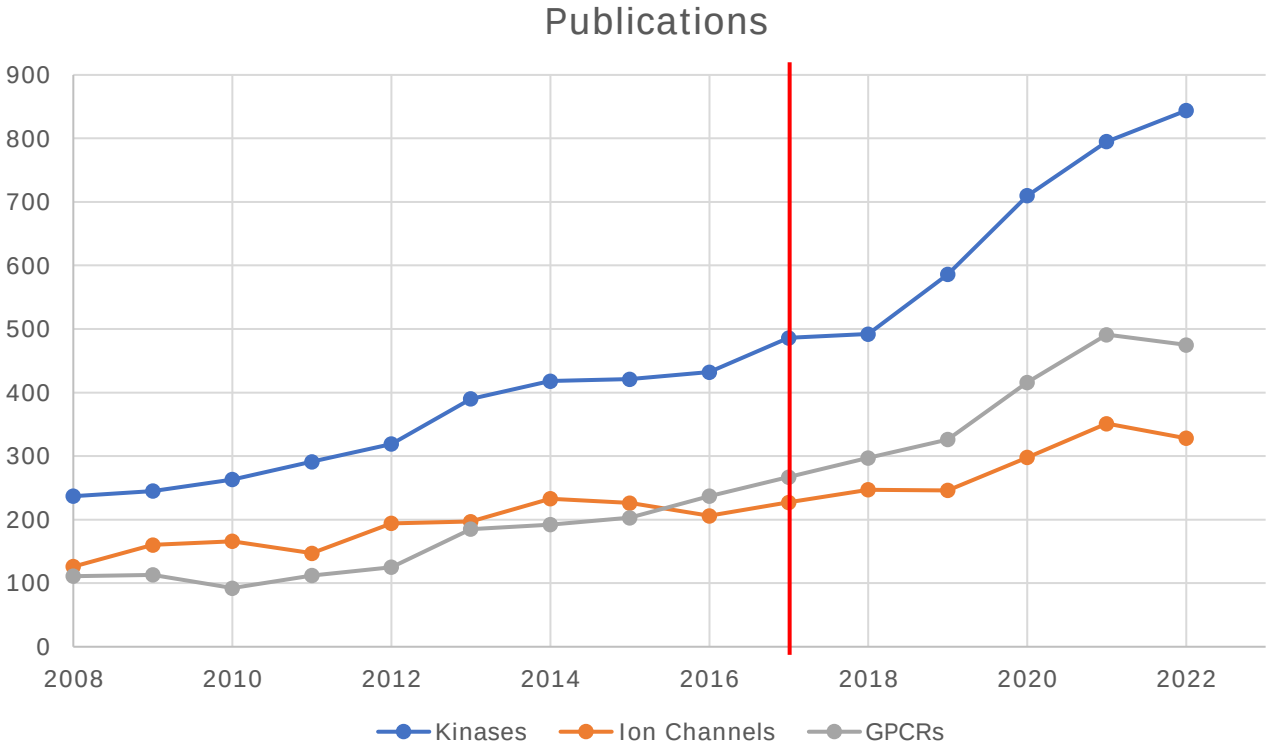
- Identification of role for RIOK3 in anti-viral immunity and the inflammatory reaction
- Paired with Dermaxon, LLC on SBIR award to develop RIOK3 as novel therapeutic target for cutaneous lupus
- Now have small molecules and PROTACs that bind to RIOK3 with nanomolar affinity (NMP505 and NMP506).
  - Blockade of RIOK3 also prevents inflammation from spreading to bystander cells



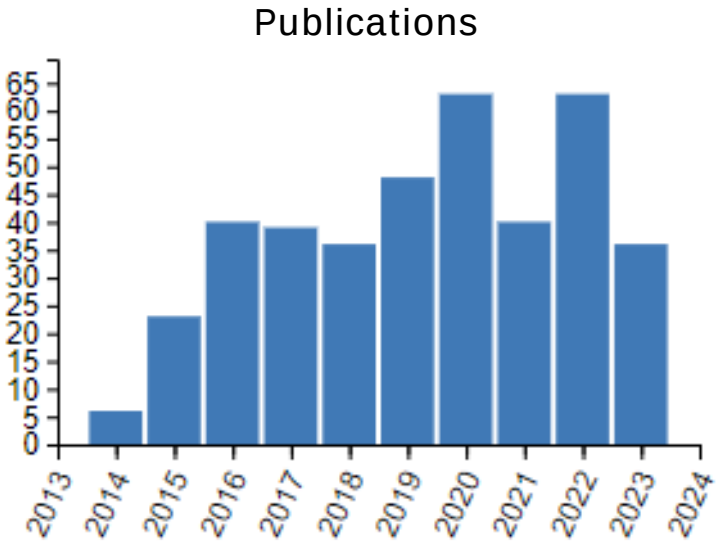
Dermaxon

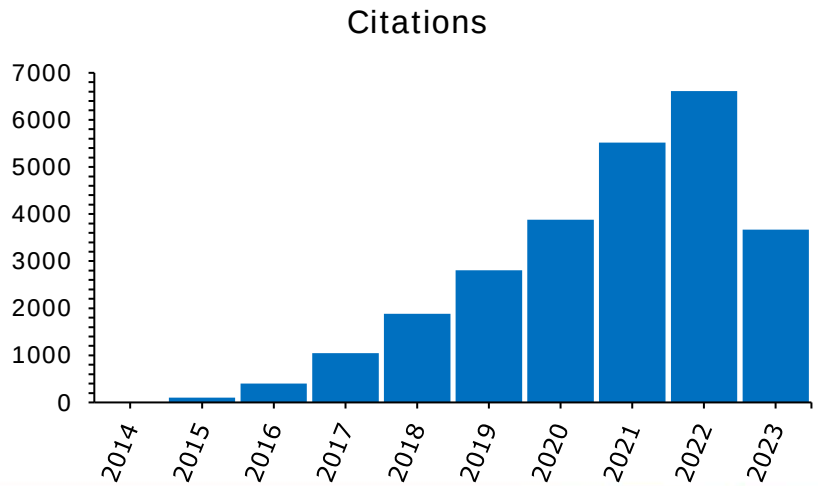
mmon Fund

IDG Program – Impacts on the Field



***The above analysis does not include publications from awardees of the IDG Program





IDG Program Outputs: Impacts on the Future

- Pharos sustainability
- R03 Pilot award successes still coming in – much impact of this initiative not yet felt
- Common Fund and NIH Institutes/Centers incorporating pilot program format into other initiatives
- Sustainability of many resources through utilization of public repositories
- Increased interest around understudied protein families – demonstration of value of these targets to human disease

Druggable Genome Special Series

Keynote (green)

Illuminating the druggable genome: Pathways to progress

Karlie R. Sharma [1], Christine M. Colvis [1], Griffin P. Rodgers [2], Douglas M. Sheeley [3]

Additional five articles in series:

- Overview of the Knowledge Management Center for Illuminating the Druggable Genome
- Informatic Challenges and Advances in Illuminating the Druggable Proteome
- Illuminating the Understudied GPCR-ome
- Illuminating Function of the Understudied Druggable Kinome
- Best Practices for Managing and Disseminating Resources, Outreach, and Evaluating Impact from IDG Consortium

https://www.sciencedirect.com/journal/drug-discovery-today/vol/29/issue/3

IDG Program Leadership

IDG Co-Chairs
Griffin Rodgers, NIDDK
Christine Colvis, NCATS

IDG Common Fund Lead
Becky Miller (OD)

IDG Program Coordinator
Karlie Sharma (NCATS)

IDG Program Analyst
Amber Peters (NCATS)

IDG Working Group

Sam Ananthan (NIDA)
Steve Benowitz (NCATS)
Mark Caprara (CSR)
Marc Charette (NHLBI)
Hemin Chin (NIAAA)
Ben Churn (NINDS)
David Dzamashvili (OD)
Colin Fletcher (NHGRI)
Zorina Galis (NHLBI)
Brionna Hair (OD)
Jerry Li (NCI)
Chris Lindsey (NICHD)
Tristan McClure-Begley (NIDA)
Mehdi Mesri (NCI)
Enrique Michelotti (NIMH)
Laurie Nadler (NIMH)
Antonio Noronha (NIAAA)
Michael O'Neil (NIAID)
Aaron Pawlyk (NICHD)
Suzana Petanceska (NIA)
Ajay Pillai (NHGRI)
Steve Pittenger (NCATS)
Lu Wang (NIDDK)
Ashley Xia (NIDDK)
Zuoyu Xu (NIAID)
Yong Yao (NIMH)
Jean Yuan (OPA)

commonfund.nih.gov

@NIH_CommonFund

National Institutes of Health
Office of Strategic Coordination–The Common Fund