

The Human Proteome Project

“Grand Challenge”

A function for every Human protein



September 24-27, 2024
Roscoff - France

Charles Pineau, HPP Chair
charles.pineau@inserm.fr

From genomes to proteomes...



Escherichia coli
4 403 genes



Saccharomyces cerevisiae
6 275 genes



Drosophila melanogaster
13 600 genes



Caenorhabditis elegans
19 099 genes



Mus musculus
18 311 genes



Homo sapiens
~ 20 000 genes

Human genome to proteome



Genome
~ 20 000 genes

Alternative Splicing
of mRNA
x 2-5

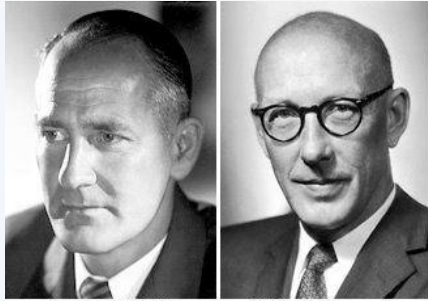
~ 50 - 100 000
transcripts (mRNA)

Post translational
modifications of proteins (PTMs)

x 50-100

Proteome
~ 5 000 000 different
proteins

?



George Wells Beadle
(1903 - 1989)

Edward Lawrie Tatum
(1909 - 1975)

Vol. 27, 1941

GENETICS: BEADLE AND TATUM

499

GENETIC CONTROL OF BIOCHEMICAL REACTIONS IN NEUROSPORA*

BY G. W. BEADLE AND E. L. TATUM

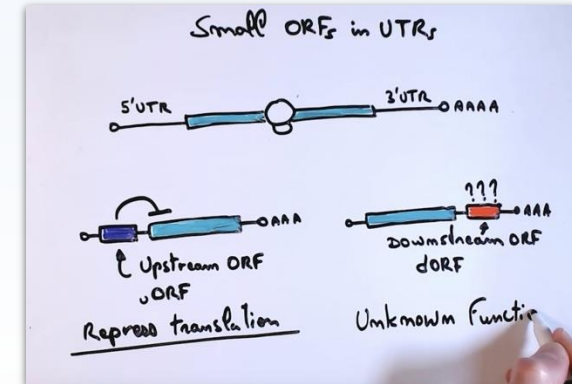
BIOLOGICAL DEPARTMENT, STANFORD UNIVERSITY

Communicated October 8, 1941

One gene (One mRNA) One protein

One gene encodes for about 7 protein in eukaryotes

Protein entries	20359
Isoforms (produced by splicing)	42329
Identifiers	881850
Binary interactions	441858
Post-translational modifications (PTMs)	192917
Natural peptides (seen by MS)	3280422
Variants (including disease mutations)	9720112
Controlled vocabularies and ontology terms	277245
Total number of publications	572783
Links between protein entries and terms	11351467



PROTEOMICS

Proteoforms as the next proteomics currency

Identifying precise molecular forms of proteins can improve our understanding of function

By Lloyd M. Smith¹ and Neil L. Kelleher² | provides invaluable information on protein

The Human microbiome - a treasure trove waiting to be unlocked

Viruses, bacteria and fungi residing in communities within and on the body
1-3% of the body mass - About 46 million genes

Mouth:

- ✓ Assist digestion
- ✓ Ward off pathogens

Colon:

- ✓ Digestion of complex carbohydrates

Sexual organs:

- ✓ Maintain pH and H_2O_2 production to kill microbes

Nose:

- ✓ Mucus production
- ✓ Antimicrobial chemicals

Lung:

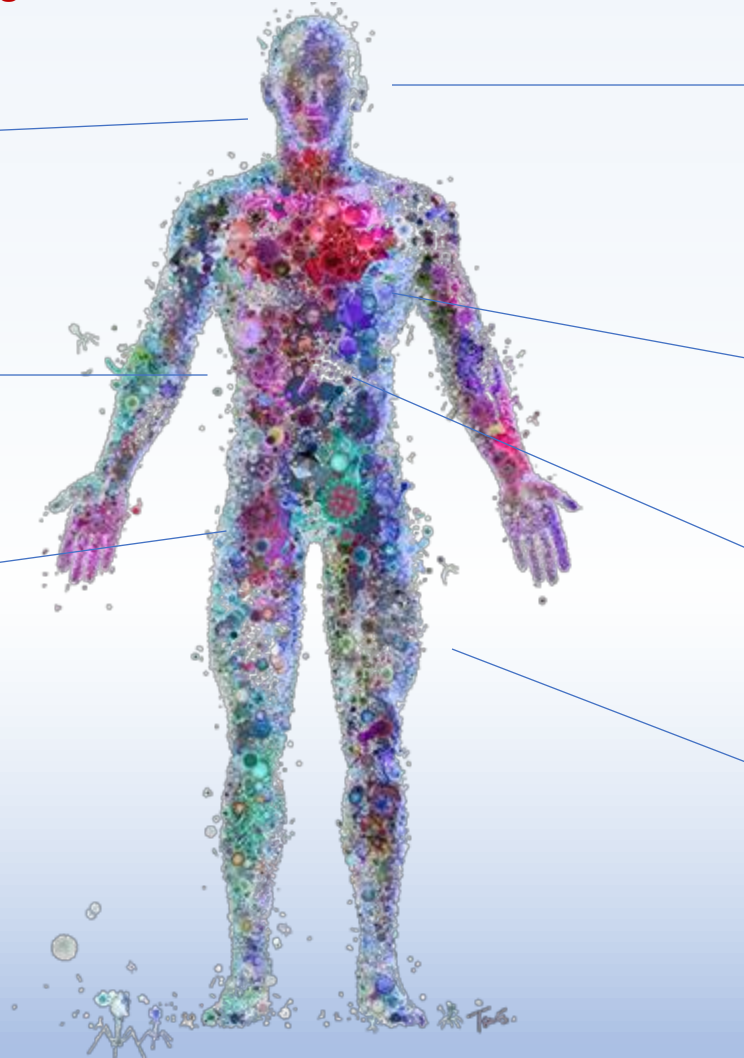
- ✓ Lubricate pulmonary tissues

Stomach:

- ✓ Prevents gastric complications

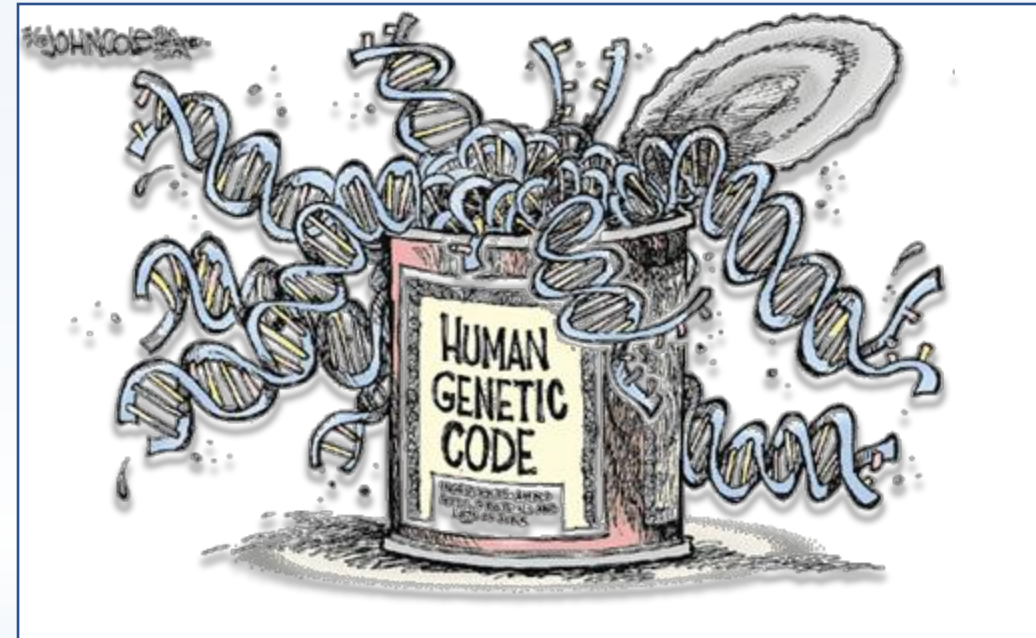
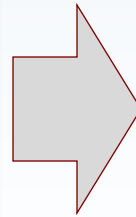
Skin:

- ✓ Fortify immune systems
- ✓ Scent production





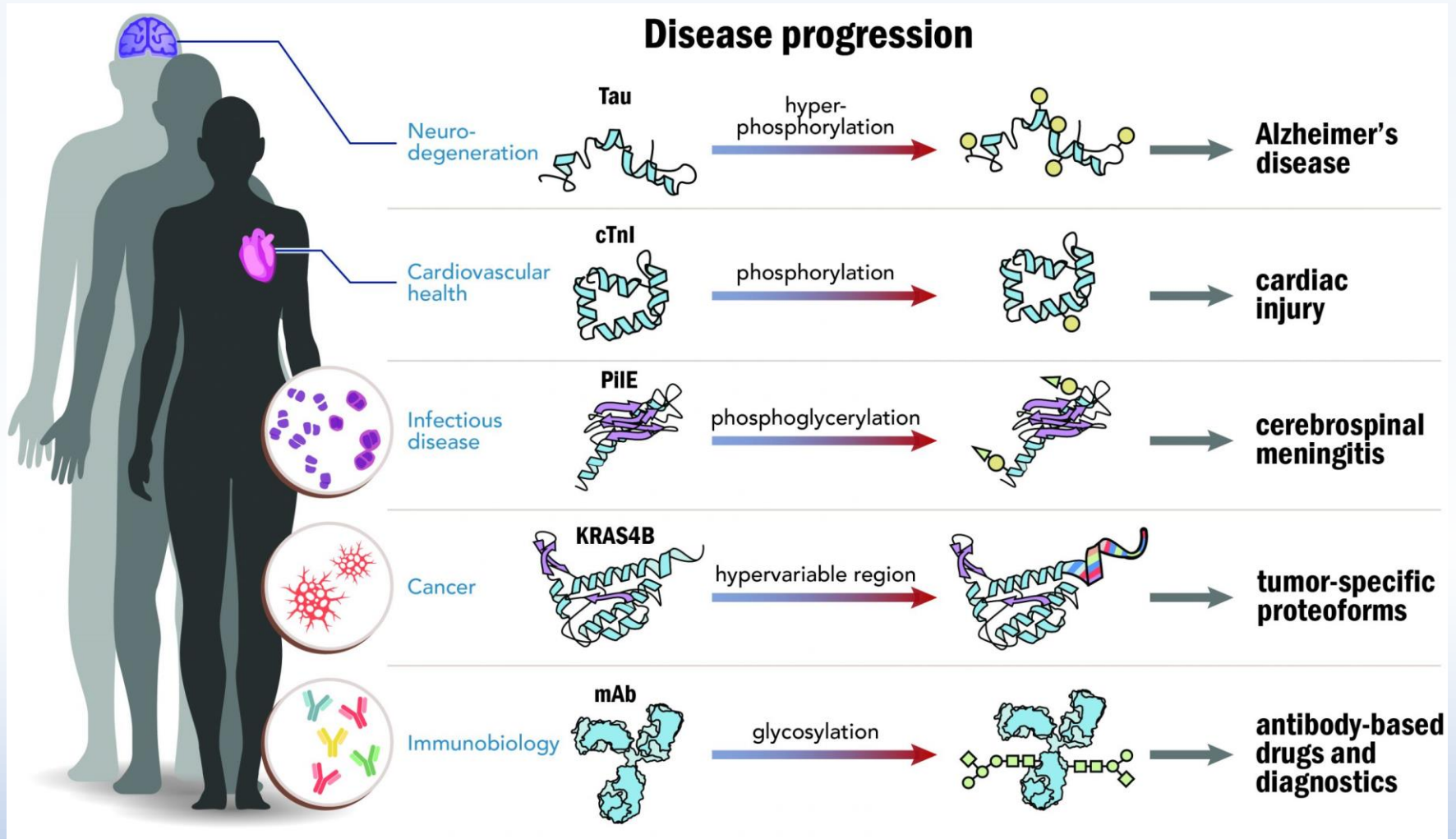
Success of the Human Genome Project



2003
15 years
US\$ 3 billions

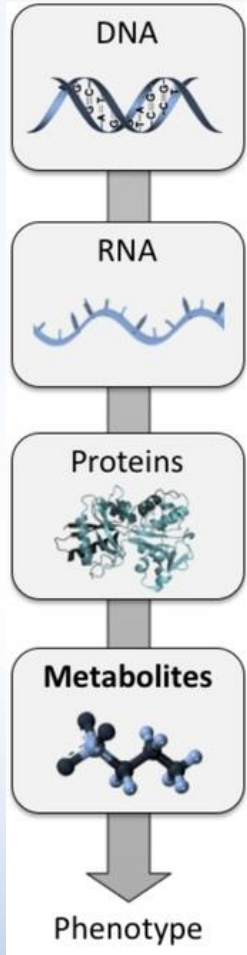
- **Pharmaceutical drugs target proteins** that mainly belong to 4 families: Enzymes, Transporters, Ion channels, Receptors
- **Knowing proteins** in a patient's disease cells will help **understand the pathology**
- **Proteins hold the key to** may be **personalized treatment**

Only Proteins can tell...



Credit: N. Kelleher, 2021, Northwestern Univ, USA

What changes are important and relevant ?



Genome -> Genomics

Transcriptome -> Transcriptomics

Proteome -> Proteomics

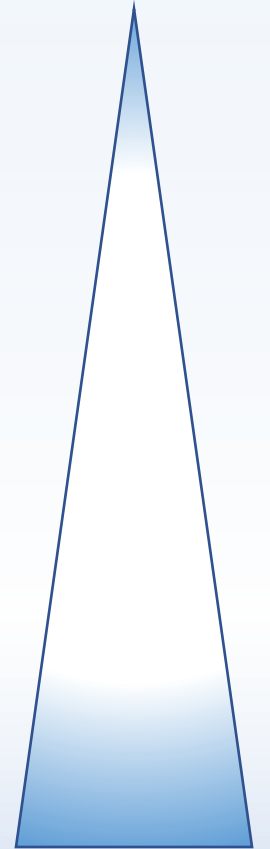
Metabolome -> Metabolomics

The world of possible

What is going to happen *(may be)*

What is happening now

What has happened



Chemical diversity
Dynamic range

Homeostatic principle: composition is constant

HUMAN PROTEOME PROJECT

[DOWNLOAD THE HPP SAB REPORT 2020](#)



[Publications](#)



[News](#)



[Terms](#)



[Resources](#)

[HPP PROGRESS TO DATE](#)

[Wiki](#)



[Get Involved](#)

HPP

TRANSLATING THE CODE OF LIFE

[twitter](#) | [linkedin](#)

ABOUT ▸ HUMAN PROTEOME PROJECT (HPP) ▸ WORLD CONGRESS AND MEETINGS ▸ PUBLICATIONS ▸ MEMBERSHIP ▸ NEWS ▸



HUMAN PROTEOME ORGANIZATION

Founded in 2001, HUPO turned 20 in 2021! We created an anniversary logo to celebrate 20 years of excellent scientific conferences, community engagement, science advocacy, and advanced education! The annual HUPO world congress has been a key event for proteomics scientists since 2002 and has continued through the recent pandemic challenges. HUPO community-centric projects, from the Human Plasma Proteome Project of 2001 to the HPP Grand Challenges of today provide crystallization points for intensive global collaborations, advancing proteomics science and networking.



TRANSLATING THE CODE OF LIFE

The Human Proteome Organization (HUPO) is an international scientific organization representing and promoting proteomics through international cooperation and collaborations by fostering the development of new technologies, techniques and training.

[Read more](#)



HUMAN PROTEOME PROJECT (HPP)

The Human Proteome Project (HPP) is an international project organized by HUPO that aims to revolutionize our understanding of the human proteome via a coordinated effort by many research laboratories around the world. It is designed to map the entire human proteome in a systematic effort using currently available and emerging techniques. Completion of this project will enhance understanding of human biology at the cellular level and lay a foundation for development of diagnostic, prognostic, therapeutic, and preventive medical applications.

[> More on the HPP](#)

19,778

PREDICTED PROTEINS ENCODED BY THE HUMAN GENOME
(EnoP+PE1+PE2+PE3+PE4)

18,357

FOUND PROTEINS
(EnoP+PE1)

1,421

MISSING PROTEINS
(EnoP+PE2+PE3+PE4)

92.8%

PERCENT HUMAN PROTEOME DISCOVERED
(EnoP+PE1/PE1+PE2+PE3+PE4)*100

MENU

- [Home](#)
- [The HPP](#)
- [About HUPO](#)
- [Initiatives](#)
- [News & Events](#)
- [Membership](#)
- [Contact Us](#)
- [Member Area](#)

SUPPORT

- [Become a Member](#)
- [L&B Membership](#)

FOLLOW HUPO





Copyright 2016 - HUPO

The **Human Proteome Project (HPP)** is a coordinated international project designed to **map the entire human proteome in a systematic effort** using currently available and emerging techniques.

Completion of this project will enhance understanding of human biology at the cellular level and lay a foundation for development of diagnostic, prognostic, therapeutic, and preventive medical applications

The Human Proteome Project

Task forces & Matrix of resources

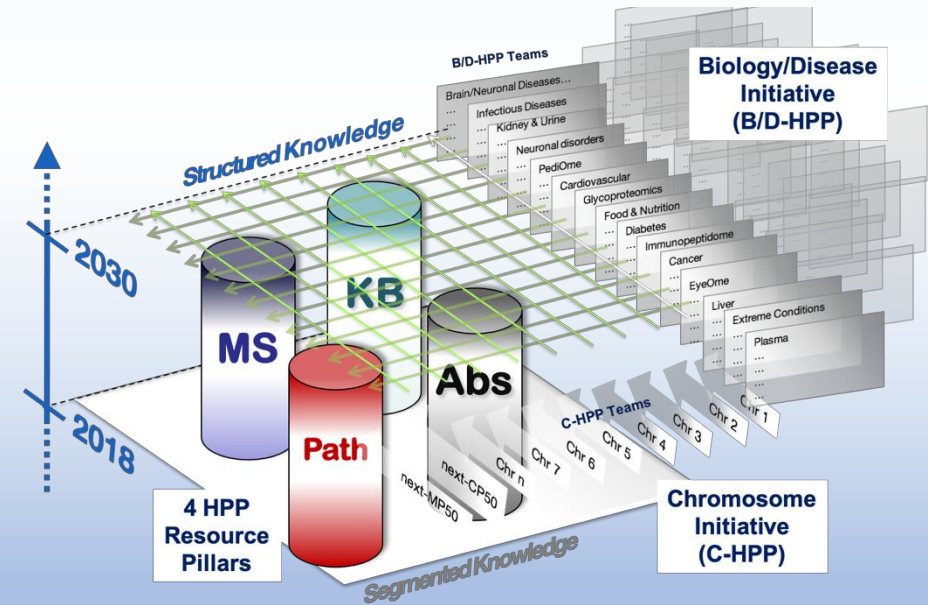
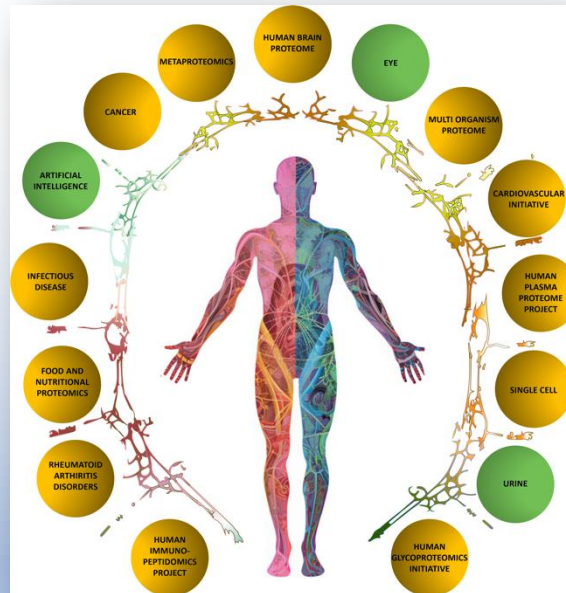
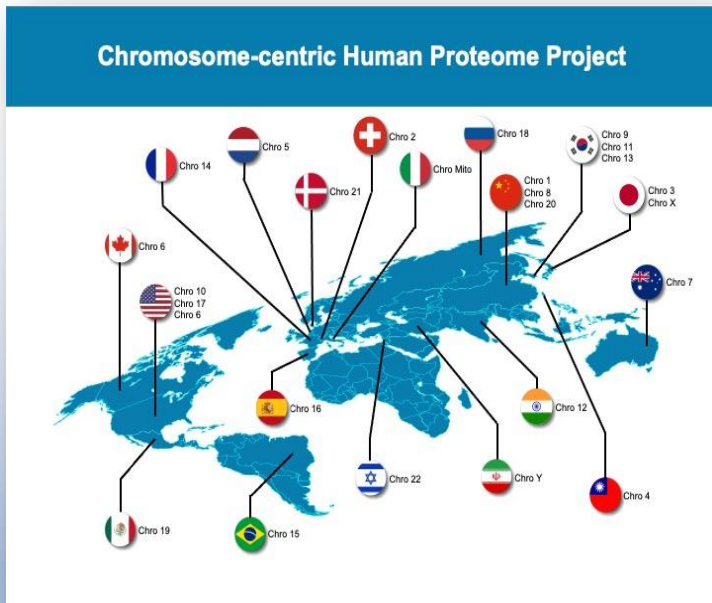


Aims at expanding the understanding of each and every gene on each chromosome

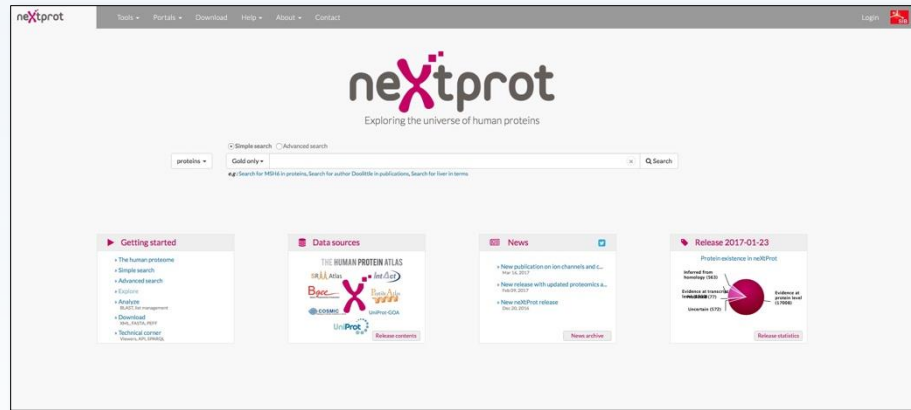


Aims at focusing on important biology research aspects and ongoing disease-focused research

HPP matrix



Examples of major integrated resources



www.nextprot.org



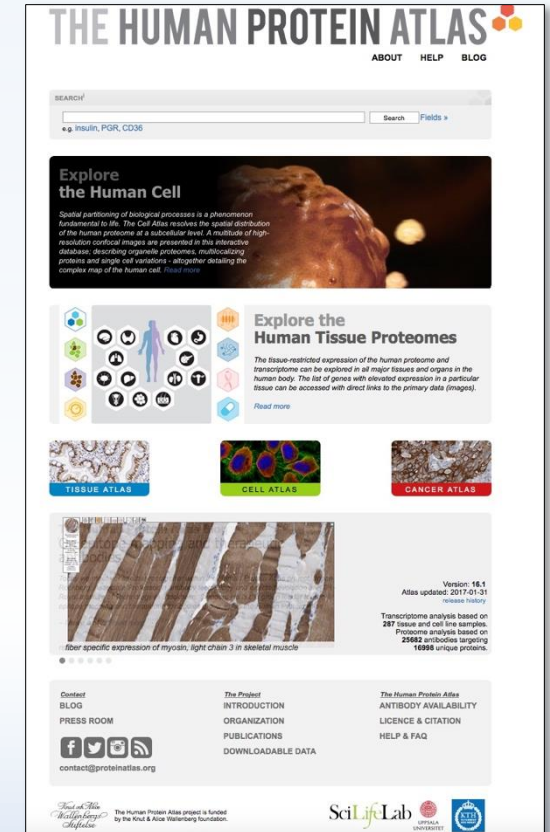
www.massive.ucsd.edu



www.peptideatlas.org



www.proteomexchange.org

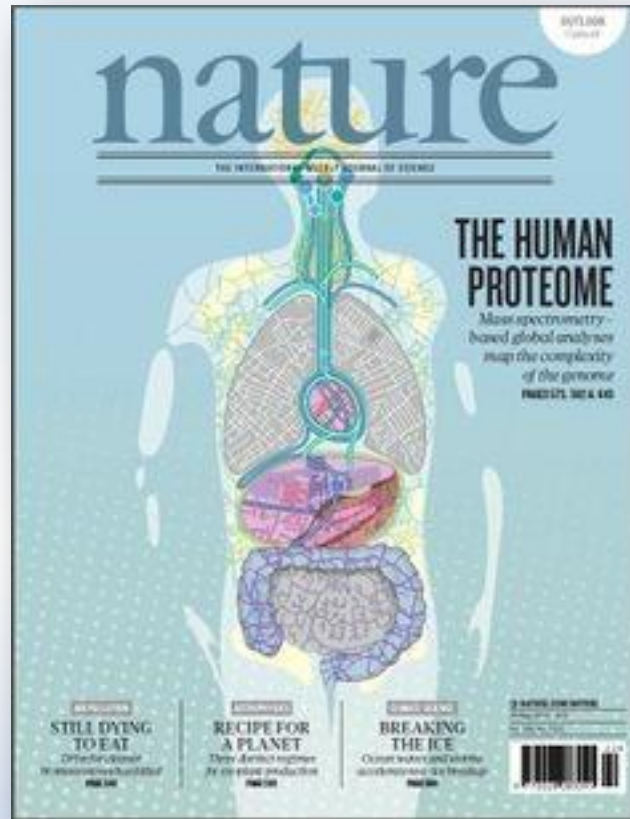


www.proteinatlas.org

First blueprint of the human proteome



Feb 2020: a major milestone



nature COMMUNICATIONS

PERSPECTIVE [Check for updates](#)

<https://doi.org/10.1038/s41467-020-19045-9> OPEN

A high-stringency blueprint of the human proteome

Subash Adhikari et al.[#]

The Human Proteome Organization (HUPO) launched the Human Proteome Project (HPP) in 2010, creating an international framework for global collaboration, data sharing, quality assurance and enhancing accurate annotation of the genome-encoded proteome. During the subsequent decade, the HPP established collaborations, developed guidelines and metrics, and undertook reanalysis of previously deposited community data, continuously increasing the coverage of the human proteome. On the occasion of the HPP's tenth anniversary, we here report a 90.4% complete high-stringency human proteome blueprint. This knowledge is essential for discerning molecular processes in health and disease, as we demonstrate by highlighting potential roles the human proteome plays in our understanding, diagnosis and treatment of cancers, cardiovascular and infectious diseases.

A decade after the release of the draft Human Genome Project (HGP), the Human Proteome Organization (HUPO) leveraged this genomic encyclopedia to launch a visionary international scientific collaboration called the Human Proteome Project (HPP)^{1–3}. Utilizing substantial community data, the HPP connects scientists, clinicians, industry, institutions and knowledgebase (KB) partners to create a framework for collaboration, data sharing and quality assurance—all targeted at discovering credible evidence for the entire complement of human genome-coded proteins (Box 1).

Here we report and discuss HUPO's first high-stringency HPP blueprint (<https://www.nextprot.org/about/updates>, data release 17-01-2020). This blueprint was assembled over 10 years by the HPP and covers >90% of the human proteome, paralleling progress made by the HGP⁴. This effort relied heavily upon community efforts that enabled HPP data inspection and re-analysis, culminating in the creation of a high-stringency human proteome KB. To illustrate the many historical innovations driving growth in proteomics, HUPO has created a historical timeline that will be released coincidentally with this publication (<https://hupo.org/Proteomics-Timeline>).

HPP mission and strategic aims

The HPP mission is to assemble and analyse community data, bringing increased granularity to our molecular understanding of the dynamic nature of the proteome, its modifications and relationships to human biology and disease. This aligns closely with HUPO's aim of 'translating the code of life', providing crucial biochemical and cell biological information that genomics per se cannot deliver, while laying better foundations for diagnostic, prognostic, therapeutic and precision medicine applications.

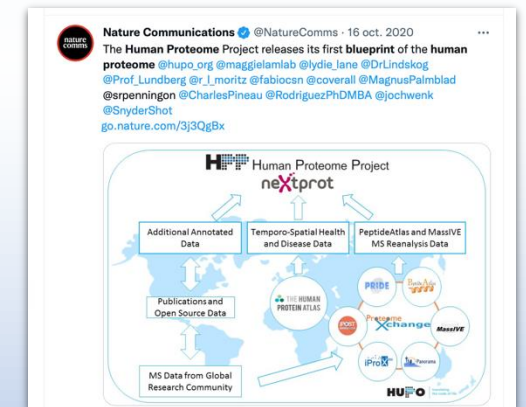
From its inception, the HPP stated two strategic objectives as follows:

1. To credibly catalogue the human proteome parts list and discover its complexity (including posttranslational modifications (PTMs), splice variants, interactions and functions) by:

*A list of authors and their affiliations appears at the end of the paper.

NATURE COMMUNICATIONS | (2020)11:5301 | <https://doi.org/10.1038/s41467-020-19045-9> | www.nature.com/naturecommunications

Completion of 92.8% of the human proteome parts list



HPP progress to date (Phase 1)

April 2023

September 2024

~~19,778~~

19,411

PREDICTED PROTEINS ENCODED BY THE HUMAN GENOME
(neXtProt PE1+ PE2 + PE3 + PE4)



~~18,397~~

18,138

FOUND PROTEINS
(neXtProt PE1)



~~1,381~~

1,273

MISSING PROTEINS
(neXtProt PE2 + PE3 + PE4)

~~93.01%~~

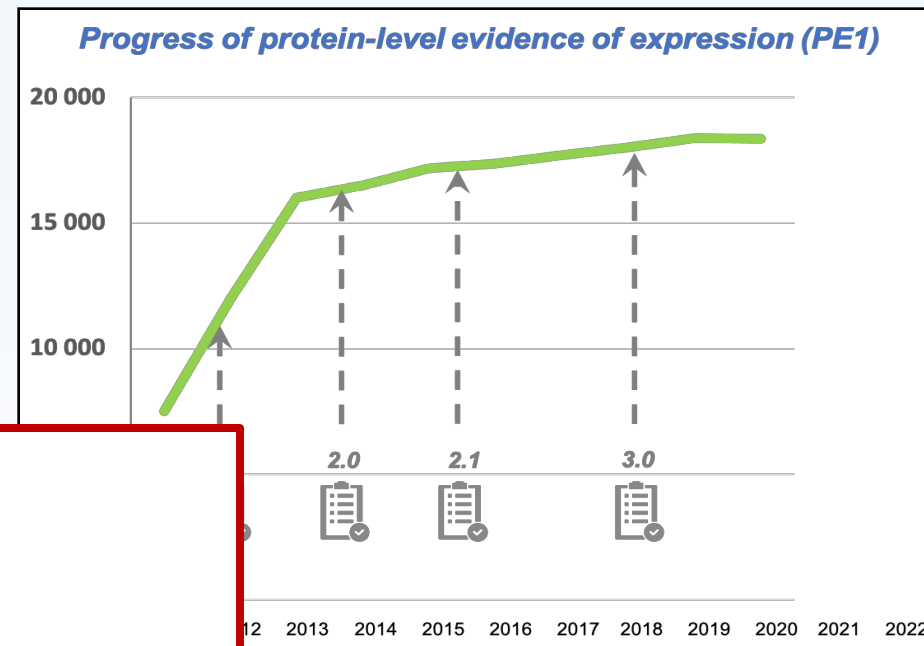
93.5%

PERCENT HUMAN PROTEOME DISCOVERED
(neXtProt PE1/(PE1 + PE2 + PE3 + PE4)) *100



Unseen proteins: why?

- ✓ Low abundance
- ✓ Time- or stress-dependent
- ✓ Tissue- or cell-specific expression
- ✓ Particular physicochemical properties
- ✓ Protein instability



> *J Proteome Res.* 2019 Dec 6;18(12):4108-4116. doi: 10.1021/acs.jproteome.9b00542. Epub 2019 Oct 21.

Human Proteome Project Mass Spectrometry Data Interpretation Guidelines 3.0

Eric W Deutsch ¹, Lydie Lane ², Christopher M Overall ³, Nuno Bandeira ⁴, Mark S Baker ⁵, Charles Pineau ⁶, Robert L Moritz ¹, Fernando Corrales ⁷, Sandra Orchard ⁸, Jennifer E Van Eyk ⁹, Young-Ki Paik ¹⁰, Susan T Weintraub ¹¹, Yves Vandenbrouck ¹², Gilbert S Omenn ^{1 13}

Affiliations + expand

PMID: 31599596 PMID: PMC6986310 DOI: 10.1021/acs.jproteome.9b00542

2021

Announcement of the Human Proteome “Grand Challenge” | *a function for every protein*

The Human Proteome “Grand Challenge”

| *a function for every protein*



- **Build upon the large resources constructed through the significant efforts of the HPP**
- **Understanding the Proteome in the context of Networks of Networks**
- **Partner with complementary consortia to expand the reach of the HPP and increase its relevance**
- **Develop forward thinking strategies to provide functional information of each protein**
- **Be a focal point for translational researchers for mechanisms, diagnostics, and therapeutics to improve human health**

HPP Grand Project | “A Function for Every Protein” White Paper:

In 2020, the Human Proteome Organization's (HUPO) Human Proteome Project (HPP) celebrated its 10th Anniversary with major achievements resulting in the high stringency blueprint of the human proteome. The HPP detailed the detection of over 90% of all predicted human proteins, and has increased this effort to 92.8% in 2021. Coming 20 years after the release of the first draft of the Human Genome, these efforts by international consortia have provided substance to the collection of human genes and highlighted proteins that are extremely difficult to reliably detect by any current technology platform.

With these unprecedented developments providing rich volumes of data and resources to interrogate the proteome, we now turn our focus to understanding the proteome in its fullest with the formation of the “HPP Grand Project”. This project will focus on the exploration of the human proteome by understanding proteins in their network environment and will enable the elucidation of its biological significance, and ultimately its function. Here, we present the HPP Grand Project: “A function for every protein” to the community. By partnering with the significant enabling initiatives worldwide to provide resources that will enhance participation by interested parties, we look towards the community to self-associate and propose HPP projects that advance the knowledge of proteins and their network of networks to provide rich biological information on their actions and ultimately their function.

Concept
The HPP is a collective membership of voluntary researchers wanting to contribute to the knowledge of the human proteome, and has recently produced the draft of the human proteome at over 90% completion. The HPP has delivered significantly on the knowledge of the human proteome and provided resources for the entire community to access and utilize in their own interest outside of the HPP. These include knowledgebases such as neXtProt and resources of the chromosome-centric, and biological/disease focused projects. The HPP has organized itself into specific focus groups that span discovery of proteins and provide credible evidence for their spatial existence. The HPP has developed a framework to disseminate and provide focused biological knowledge to understand various aspects of human biology and disease. These encompassing efforts provide an ever-expanding set of knowledgebases that is the first in the world and provides goalposts as well as community metrics for data integrity and biological relevance. Given this background the HPP has developed over the last 20 years, starting off with the Human Plasma Proteome project, and the milestones achieved recently, it is time for the Human Proteome Organization, through the HPP, to focus its efforts to collectively contribute to addressing important biological problems on a “Grand scale”.

To this end, the HPP Grand Project will provide a framework for the HPP, and its various initiatives and pillars, to join forces in order to integrate and direct their various proteome focused efforts to collectively contribute to a grand program important to human health and well-being. The goal of the HPP Grand Project is to advance our understanding of a complex biological problem in a consortia framework where individuals, or groups of individuals, can propose various work plans to contribute to the grand project and access available resources as well as support grant applications at their domicile.

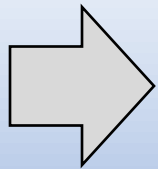
*Robert Moritz - Nov 2021
Chair Human Proteome Project*

The Human Proteome “Grand Challenge”

| *a function for every protein*

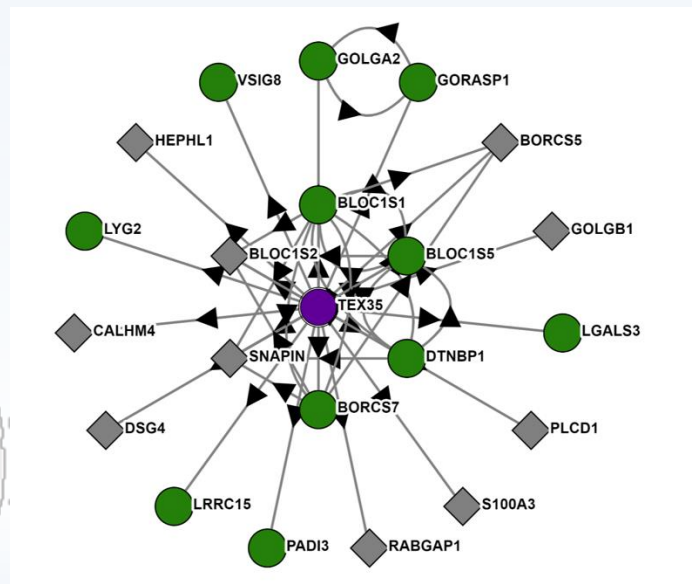
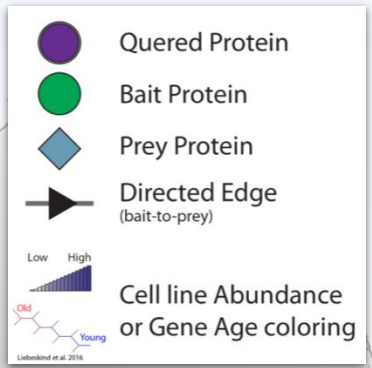


- Open to all interested groups or individuals
- Agnostic to disease and biological context
- Utilize resources already familiar to the research group
- Advance the knowledge of systems being worked on through the access to resources to perturb these systems
- Build networks and understand the context of each protein to build a KB of overlapping networks

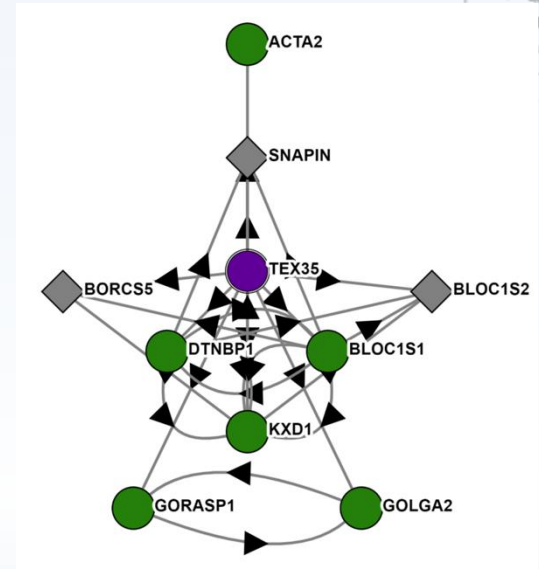


Need for a central repository

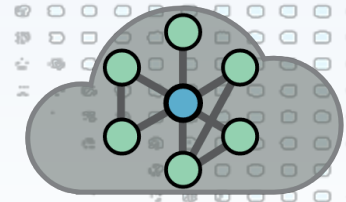
Providing context to networks through perturbation



HEK293T
human embryonic kidney cells
with SV40 large T antigen



HCT116
Adult male human
colorectal carcinoma cell line



BioPlex
bioplex.hms.harvard.edu

Approach protein function through structures

The Human-Proteome “Grand Challenge”

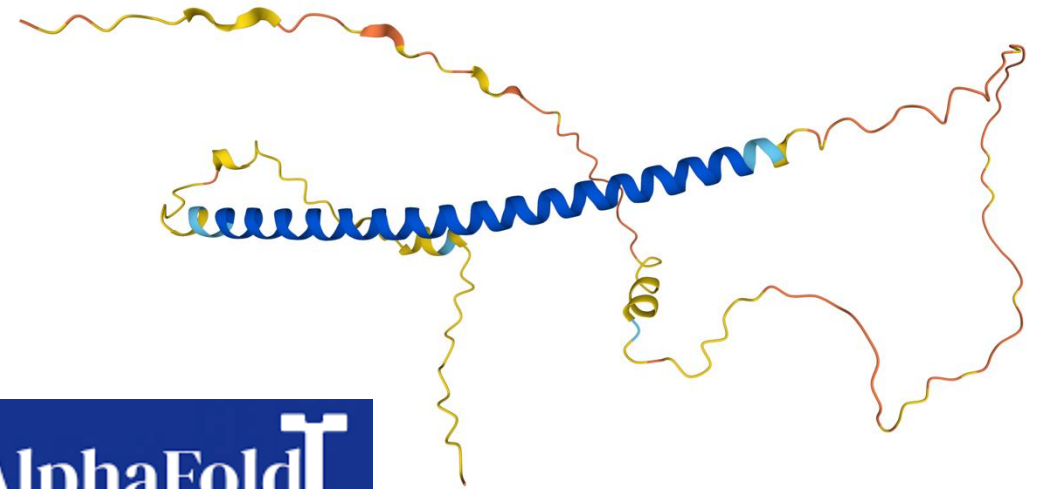
a function for every protein

Tracking the progress of the project

In the current neXtProt release there are:

- 20,379 entries
- 1,669 proteins with no function annotated
- 25 entries with a function prediction

Testis-expressed protein 35 (Gene TEX35)



AlphaFold

https://www.nextprot.org/entry/NX_Q5T0J7/

Use small molecules to perturbate the proteome of a cell

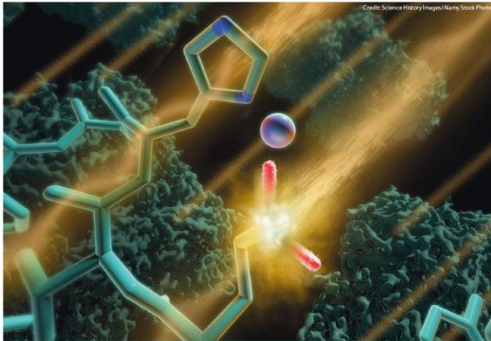
The Human Proteome "Grand Challenge"

a function for every protein



Target 2035 Chemical Probes

nature DRUG DISCOVERY NEWS & ANALYSIS



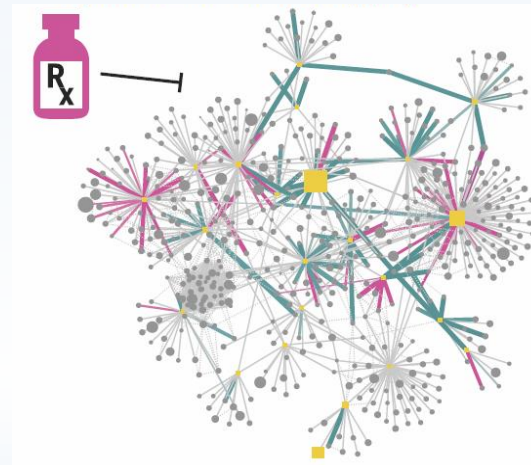
A probe for every protein



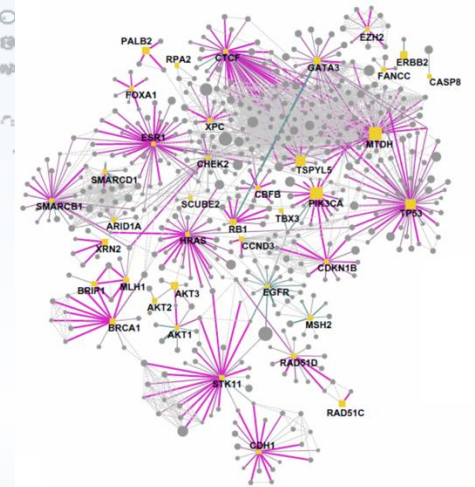
Feature
Target 2035: probing the human proteome

Adrian J. Carter¹, Oliver Kraemer¹, Matthias Zwick², Anke Mueller-Fahnrow³, Cheryl H. Arrowsmith⁴, Aled M. Edwards⁵

"Target 2035 — an ambitious open- science proposal to develop a suite of chemical genomics tools to modulate every protein in the proteome"
- Asher Mullard



Comparative mapping

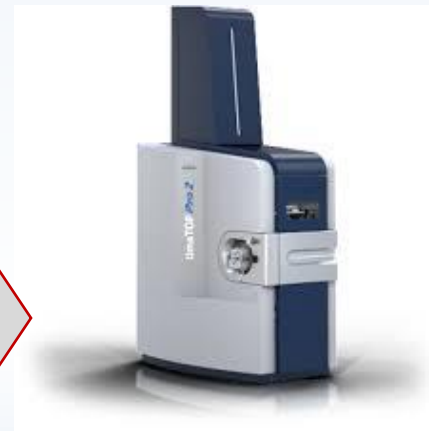
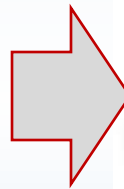
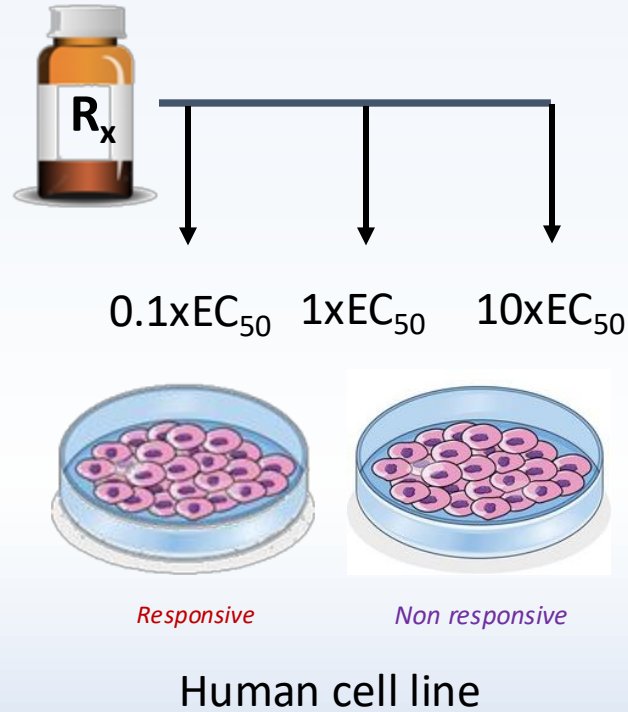


Define interactors and perturbers to understand impact and function



translating the code of life

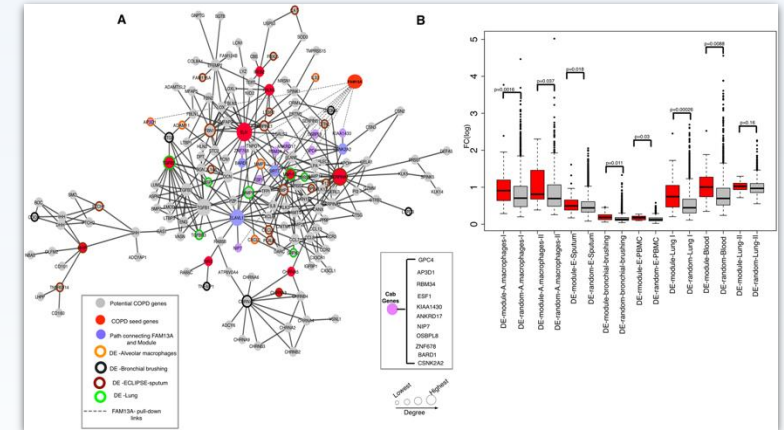
2023 Call with HPP



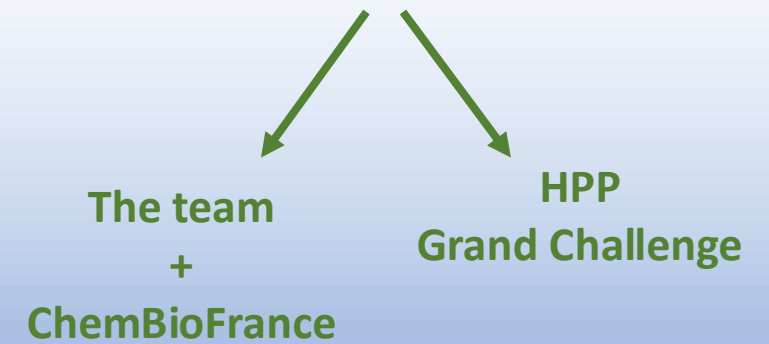
Proteomics analysis

Pilot project will help define strategies and standards for:

- MS method and data acquisition ✓
- Data exploitation ✓
- Setting up a central Repository
-



- Known perturbation
- Non-expected perturbation to understand possible function



How to contribute to the HPP and Grand Challenge ?

- ✓ **A team, an individual**
- ✓ **Specialist of a protein, a protein family, a pathology, a cellular model, etc.**
- ✓ **Mastering a cutting-edge technological approach**
- ✓ **Willing to share results with the HUPO community**
- ✓ **Interested in joining HPP initiatives or to launch a new one**

The Human Proteome “Grand Challenge”

| *a function for every protein*

Ongoing actions :

- Computation of a “Function Evidence” scoring
- Release of the annual PE numbers now carried out by Uni-Prot
- Creation of an HPP Portal

Computation of a “Function Evidence” scoring

Summary from the HPP Day @  **HUPO**
BUSAN
2023 SEP 17-21
22nd Human Proteome Organization World Congress *Emerging Strategies to address protein function(s)*

The proposal to score each protein with an FE score similar to a PE score

- FE 1:** At least one function for the protein is very well characterized. Essentially textbook quality understanding of function. There’s always more to understand, of course.
- FE 2:** Quite a lot is known about the protein’s function, but it is lacking really good manual assertion information of both biological process and molecular function.
- FE 3:** Somewhat specific ideas of the biological process that the protein participates in is annotated, but still a long way from where we want to be.
- FE 4:** Some vague function information is known, a decent basis for more research.
- FE 5:** Essentially no function is known. There may be some vague ideas like “metal ion binding” or some localization information, but this is not deemed to be a useful function.



Working Group on Functional scoring (*Eric Deutsch*)

Update on the Protein Evidence scoring now carried out by Uni-Prot and based on GENCODE references

How many proteins remain to be identified in the Human Proteome?

From: Omenn *et al.*, Submitted to J Proteome Res - Sept20234

Protein Category	2012-02	2013-09	2014-10	2017-01	2019-01	2020-01	2021-02	2022-02	2023-04	2024-05
PE1: Evidence at protein level	13,975	15,646	16,491	17,008	17,694	17,874	18,357	18,407	18,397	18,138
PE2: Evidence at transcript level	5205	3570	2647	1939	1548	1596	1265	1135	1151	928
PE3: Inferred from homology	218	187	214	563	510	253	147	195	215	168
PE4: Predicted	88	87	87	77	71	50	9	13	15	62
PE5 & no PE: Uncertain										115
Missing Proteins	5511	3844	2948	2579	2129	1899	1421	1343	1381	1273
PeptideAtlas canonical	12,509	13,377	14,928	15,173	16,293	16,655	16,702	16,957	17,245	17,416

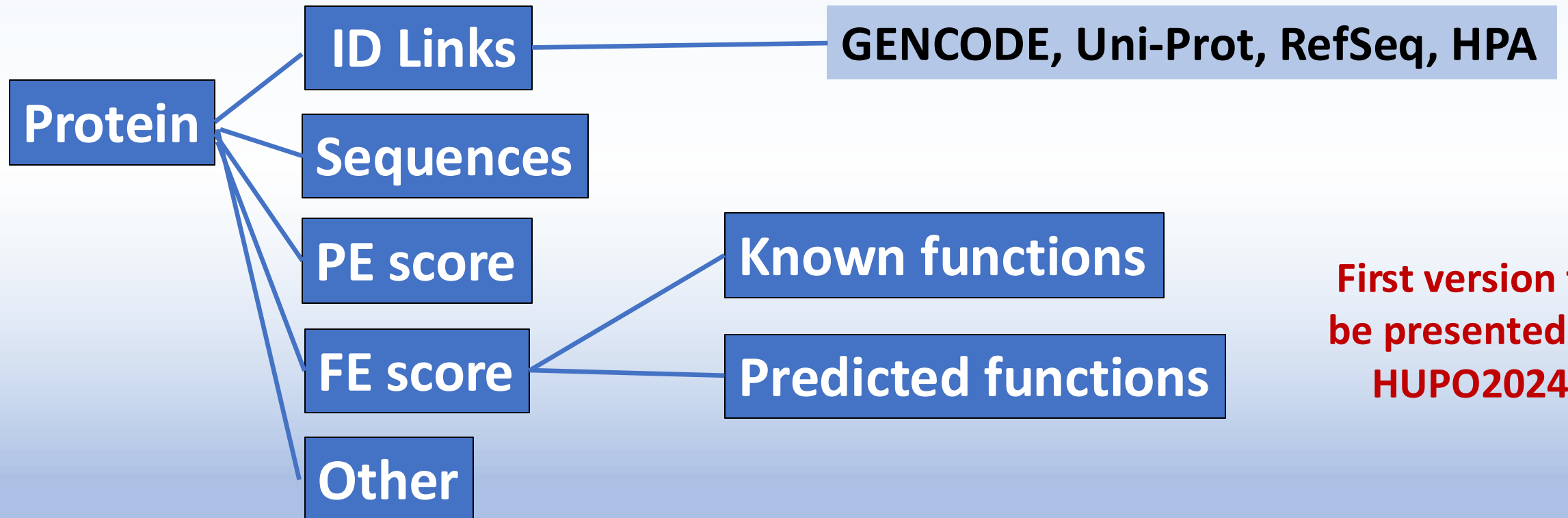
Numbers of proteins by PE level from neXtProt for 2012-02 to 2023-04, plus 2024-05 from GENCODE+UniProtKB, showing the progress in identifying missing proteins to become PE1 proteins, and PeptideAtlas canonical proteins by mass spectrometry.^{a,b}

a) *PE1/entire target list = 18138/19411 (93.5%)*

b) *More stringent HPP Guidelines (v2.1) were imposed in 2016 and enhanced in 2019 (v3.0)*

The Human Proteome Project Portal

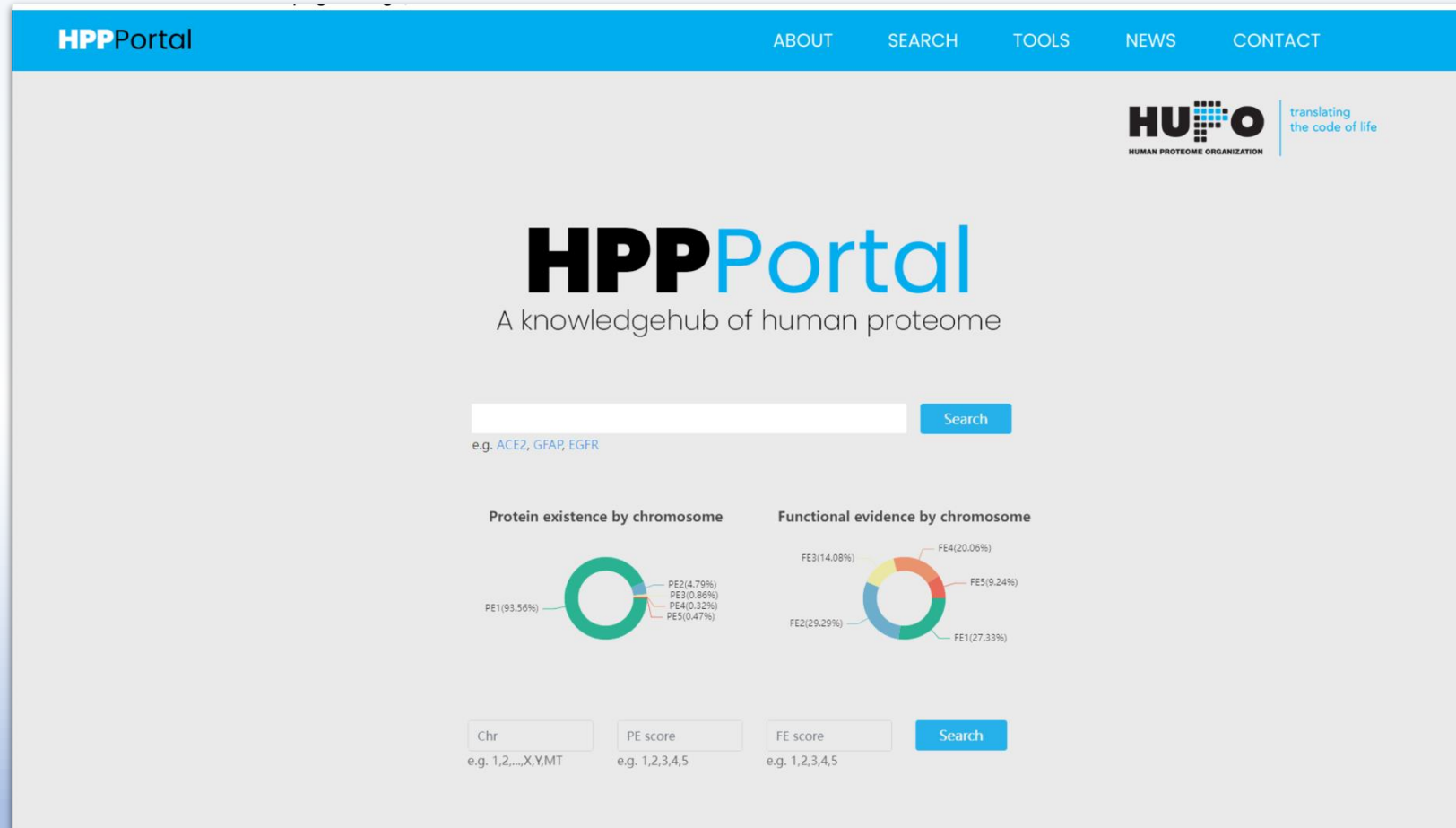
- ✓ A light construction
- ✓ Ongoing construction by Gong Zhang et coll. (*Jina Univ & Chi-Biotech, China*)
- ✓ Will allow full HPP and HUPO branding



**First version to
be presented at
HUPO2024**

Under construction... First view of the HPP Portal

Official knowledge hub for the HPP Grand Challenge



The screenshot shows the HPP Portal website interface. At the top, there is a blue navigation bar with the text "HPPPortal" on the left and "ABOUT SEARCH TOOLS NEWS CONTACT" on the right. Below the navigation bar, the HUPRO logo is displayed with the tagline "translating the code of life". The main heading "HPPPortal" is prominently featured, followed by the subtitle "A knowledgehub of human proteome". A search bar is present with a "Search" button and an example input "e.g. ACE2, GFAP, EGFR". Below the search bar, there are two donut charts: "Protein existence by chromosome" and "Functional evidence by chromosome". The first chart shows data for PE1 (93.56%), PE2 (4.79%), PE3 (0.86%), PE4 (0.32%), and PE5 (0.47%). The second chart shows data for FE1 (27.33%), FE2 (29.29%), FE3 (14.08%), FE4 (20.06%), and FE5 (9.24%). At the bottom, there are three input fields for "Chr", "PE score", and "FE score", each with a "Search" button and an example input "e.g. 1,2,...,X,Y,MT", "e.g. 1,2,3,4,5", and "e.g. 1,2,3,4,5" respectively.

**Will be launched
at HUPO2024**

Submitting proposals to get funded nationally and join the HPP Grand Challenge

1. Download project form template

2. Submit completed form hpp@hupo.org

3. Obtain validation of proposal

+/- Recommendations



Charles Pineau – France	Susan Weintraub – USA
Cecilia Lindskog – Sweden	Gilbert S. Omenn – USA
Robert Moritz – USA	Ed Nice - Australia
Chris Overall – Canada	Bogdan Budnik – USA
Eric Deutsch – USA	Gong Zhang - China
Nicolle H. Packer – Australia	Jennifer Van Eyk – USA
Michael H. Roehrl - USA	Livia Rosa - USA

Grand Challenge Task Force



4. Submit with grant application to your national funding agency



List of validated projects forwarded to national funding agencies

THE π -HuB PROJECT

The proteomic navigator of the human body

人体蛋白质组导航计划

Led by Dr Fuchu He



DRESDEN, GERMANY 2024

OCT 20-24

23rd HUMAN PROTEOME ORGANIZATION WORLD CONGRESS



JUNE
16-20,
2025



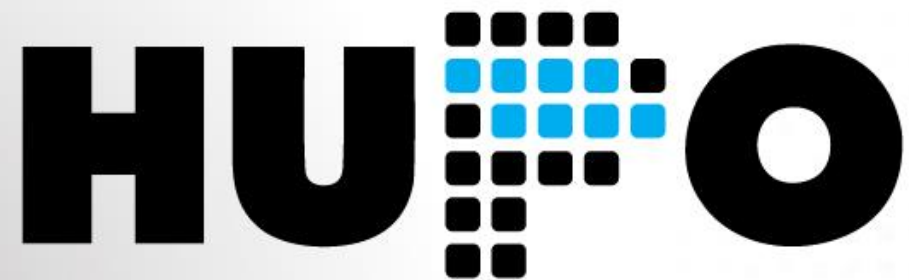
SAINT-MALO
France

LE GRAND LARGE
Palais des Congrès
SAINT-MALO

EUPA

2025





HUMAN PROTEOME ORGANIZATION

translating
the code of life