

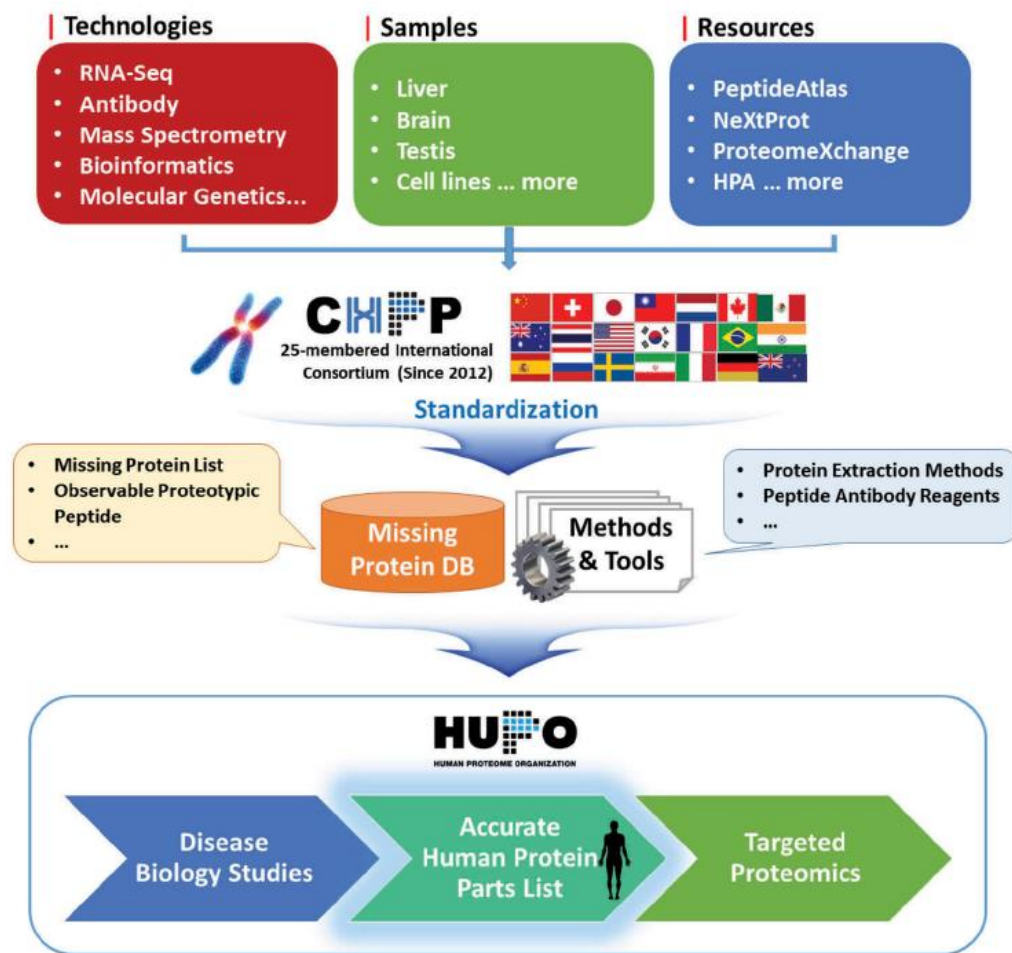
C-HPP: Progress Update and Future Direction

**“Major Achievements during Phase I (2012-2018), Future Plans
in Phase II (2019-2027) and Beyond”**

**All credits are given to 25 International C-HPP Teams, HUPO HPP Investigators,
Resource Pillars and General Proteomics Scientists**

**Presented by Young-Ki Paik, Lydie Lane and Christopher M. Overall, C-HPP Co-Chairs
with Strong Support by Gil Omenn, HPP Chair**

1. Major Achievements during Phase I (2012-2018)



Major Achievements by C-HPP & Global Teams

1. Establishing Data Submission System to PDX ('13)
2. MS Data Guidelines (v2.1)(Deutch/Omenn)
3. Metrix system for progress checking on MP identification (by G. Omenn, Chr17)
4. JPR SI Publications (2013-2018, ACS, Overall)
5. MP IDs (5500 to 2168, neXtProt, Lydie/Eric)
6. Moving From MP Annotation to Function (neXt-MP50 and neXt-CP50)-in Dual Missions
7. Funding Opportunities (HQ, & Several Chr Teams)
8. SOP for C-HPP Operation (JPR, 2012)
9. >20 Scientific Mtgs in Member Countries
10. >200 Scientific Papers in NBT, MCP, JPR SI etc.
11. Annual C-HPP Newsletter Publications (No.1-7)
12. Development of Key Bioinformatics Tools & Databases for Protein Analysis Pipelines (JPR SI+)
13. Propose a co-work model between C- and B/D-HPP for uPE1 study (Overall)

*Ref. Paik, Omenn, Hancock, Lydie & Overall
2017, Expert Rev Proteomics, p1-13*

2. Future Direction for Phase II (2019- beyond)

neXt-CP50, the C-HPP pilot project for functional characterization of 50 uPE1 proteins with no known function

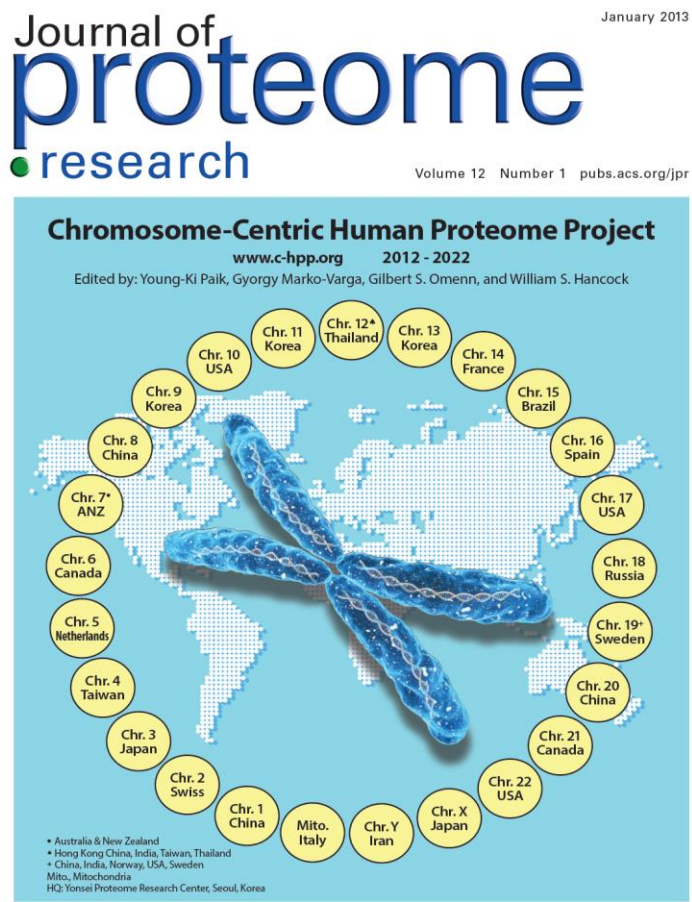
Rationale

- Many disputes over defining the credible number of canonical human proteins (~20,000 protein-coding genes) and their probable biological functions.
- The protein **numbers change** from year to year due to identification of their cognate protein products. Even more, there remain **2,168 missing proteins (MPs)**.
- This **unclosed status** of human protein numbers motivated C-HPP Investigators (e.g., neXt-MP50, neXt-CP50, MS data interpretation Guidelines etc.)
- **Thus**, Identifying **MPs** and characterizing the **proteins of unknown functions** (e.g., uPE1) would be the rate-limiting factors for the completion of the C-HPP

neXt-CP50 Challenge, a Pilot Project for C-HPP

- **Aims:** To test the feasibility of the functional characterization works for 1937 dark proteins (1260 uPE1+677 uMPs), we selected 50 uPE1 proteins that can be investigated within the 3-year term.
- **Scope:** ~50 uPE1 only: uMPs are not included in this initiative since they require a lot more work to establish their protein existence prior to functional characterization. We also excluded **proteoforms** from this pilot study due to their infinitive numbers).(uMPs: unknown function MPs)
- **Funding:** Each team's own grants

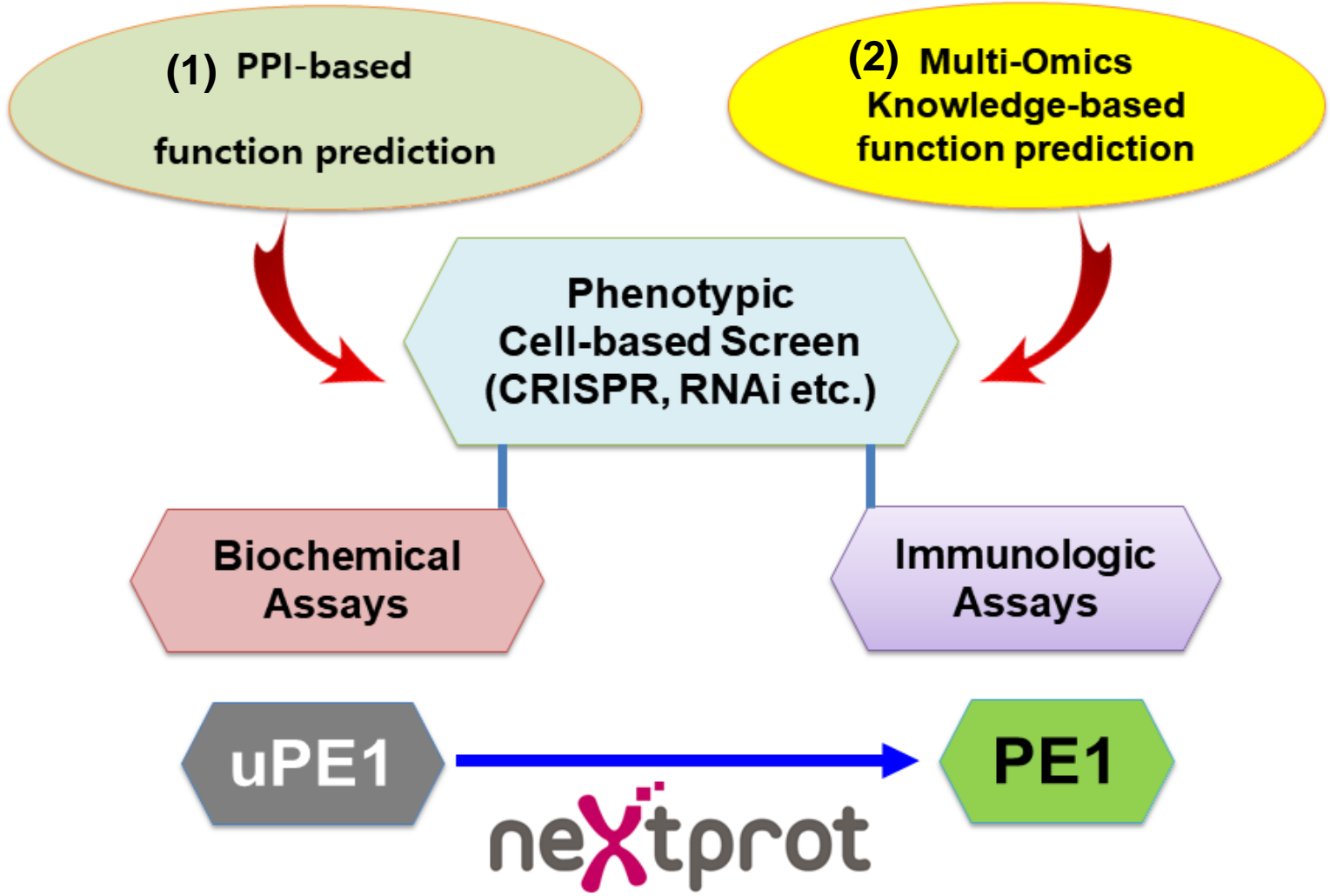
Voluntary Participants to neXt-CP50 Pilot Project (14 out of 25 Chromosome Teams: 11 countries from the 2017 Survey)



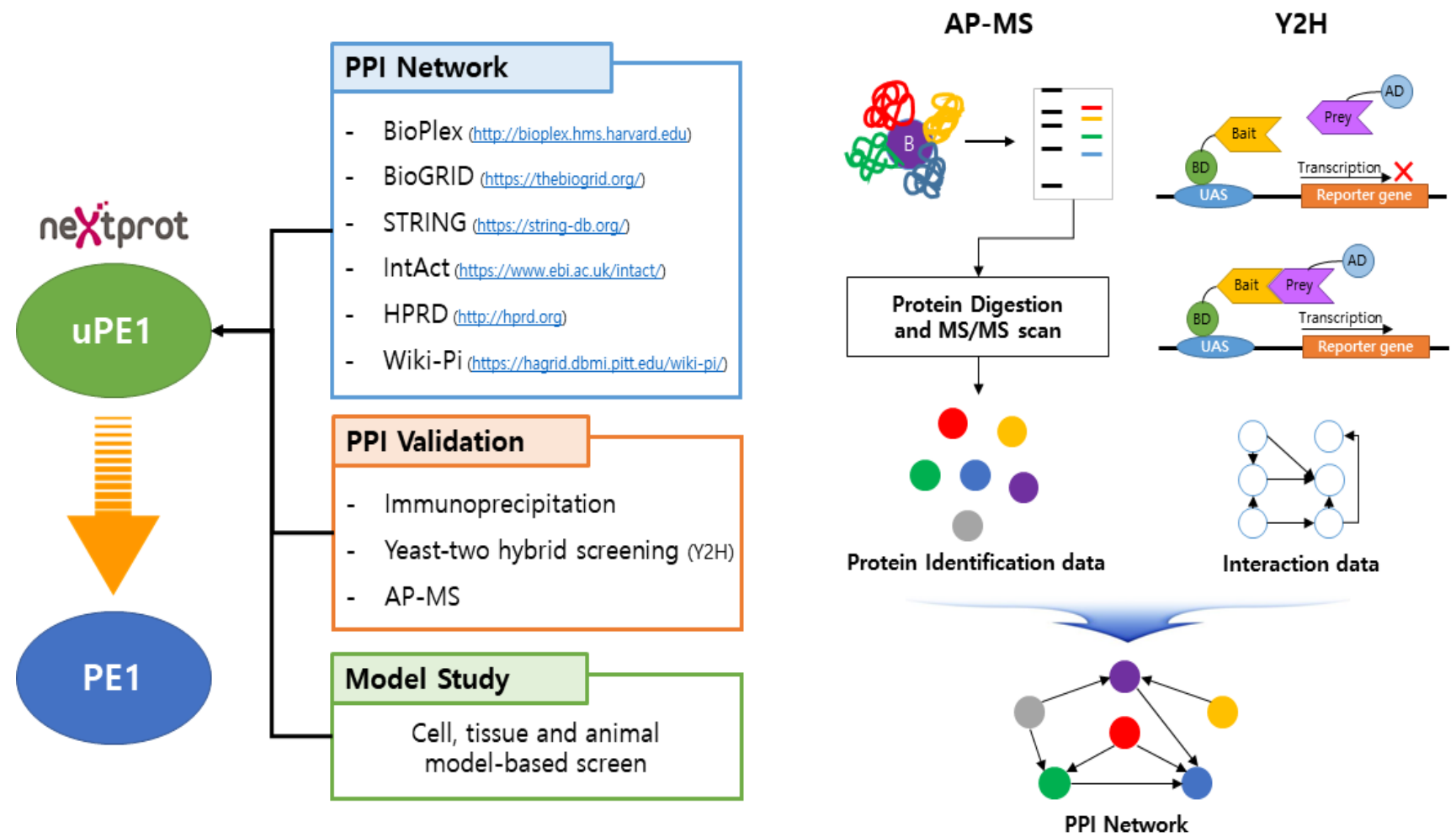
Initial Target uPE1s for the neXt-CP50 Teams

Chr.	PI	Total*	1 st priority	1 st period	2 nd priority	2 nd period	Comments
2	Lydie Lane	80	5	2018~2019	N/A	N/A	Three of target protein got funded, all target from other chromosomes (Not chr 2)
3	Takeshi Kawamura	56	2	2018~2020	N/A	2021~	
4	Yu-Ju Chen	53	2	2018~2020	8	2021~	
9	Je-Yoel Cho	57	2	2018~2019	8	N/A	
10	Joshua LaBaer	61	2	2018~2020	8	2021~	
11	Jong Shin Yoo	70	2	2018~2020	8	2021~	
13	Young-Ki Paik	26	5	2018~2021			
15	Gilberto B Domont	44	2	2018~2020	6	2021~	
16	Fernando J Corrales	54	2	2018~2021	4	N/A	
17	Gilbert S Omenn	66	~17?	2018-2021			*Dev of bioinformatics tool for functional predictions
18	Alexander Archakov	14	2	2018~2020	14	2021~	
19	Sergio Encarnación-Guevara	68	2	2018~2019	6	2020~	
20	Siqi Lui	35	4	2018~2020	11	2021~	
Y	Ghasem Hosseini Salekdeh	1	1	2018~2020	N/A	N/A	
Total	neXt-CP50		33-50	2018~2021	73	>2021	

Experimental Strategy



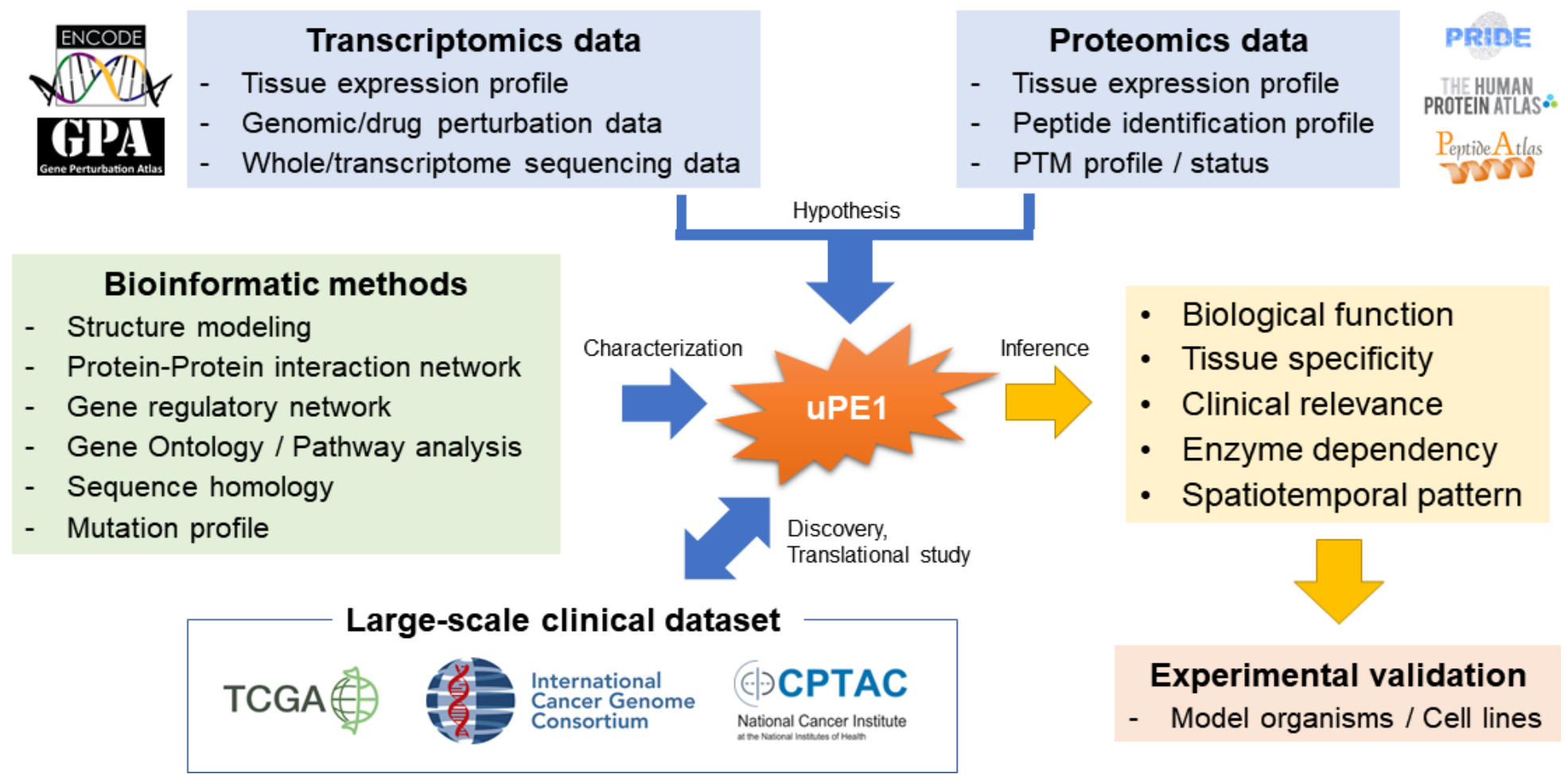
(1) The Workflow of PPI-based Function Prediction of uPE1



This workflow utilizes the currently well-established PPI-network DBs in order to predict probable functions of uPE1 based on those protein-protein interaction data. Several components of this workflow are shown, which are interchangeably exercised singly or combined where appropriate.

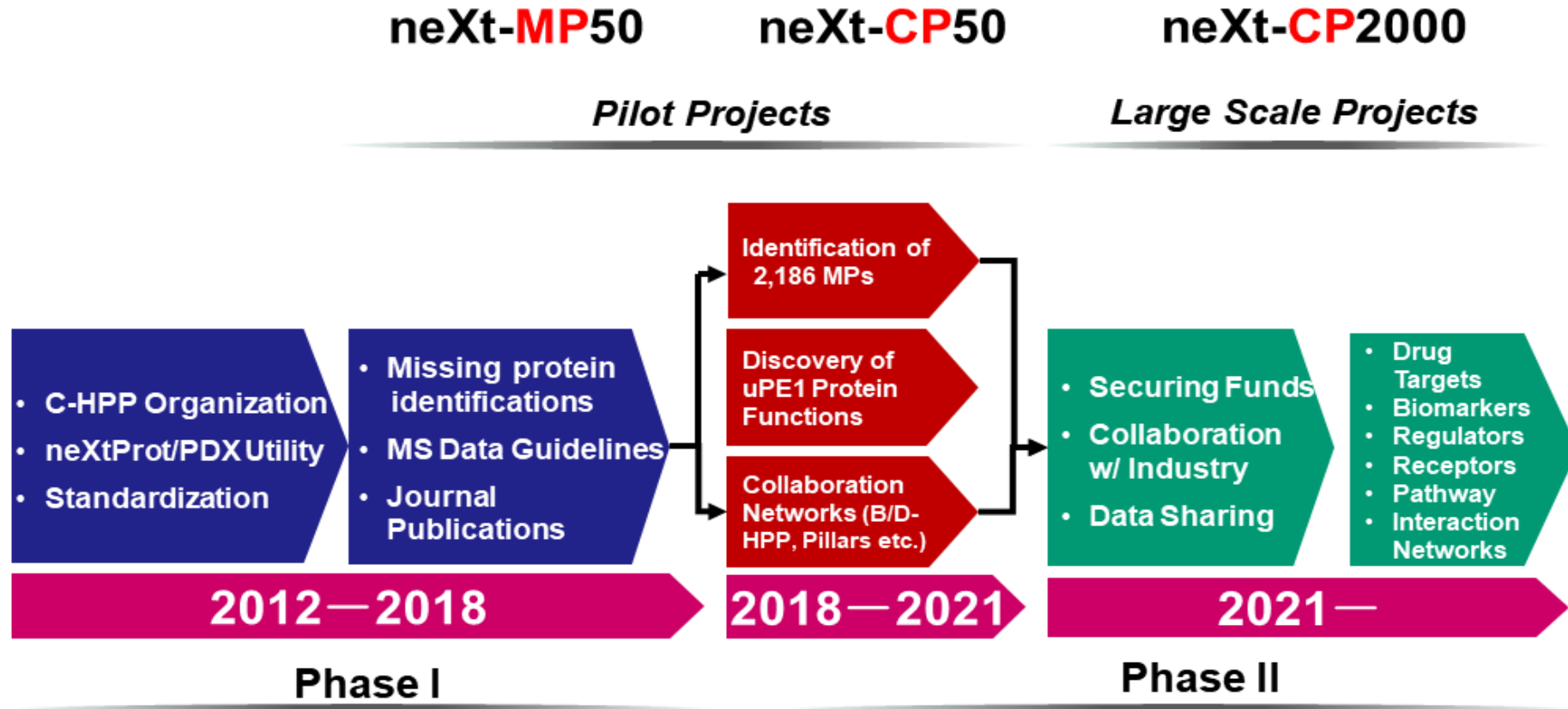
(Paik et al., 2018, JPR in revision)

(2) Multi-Omics Knowledge-based function prediction of uPE1 protein targets



This workflow utilizes the well-established publicly available large-scale or global profiling knowledge base which may well cover diverse functions of proteins or genes at the molecular levels (Baird et al, 2018, *NPJ Precision Oncology*)

A proposed short- and long-term plan to characterize human dark proteins of unknown functions (with C-HPP 2.0)



The terms, neXt-CP50 and neXt-CP2000 stand for characterizing 50 uPE1, a small set of dark proteins in 3 years (a pilot) and then ~2000 dark proteins (1260 uPE1 plus 677 uMPs) over a longer period of time. Term ‘CP’ stands for characterization of proteins with unknown functions. (Paik et al., 2018, JPR in revision)

Summary and Conclusions

- **Present:** Our C-HPP efforts during Phase it seems quite fair to say C-HPP made a notable progress in several categories. We are about to put the C-HPP 2.0 plan in place for enhancing the ongoing projects.
- **Future:** The C-HPP pilot project, neXt-CP50, will move forward with full support from the HUPO community (e.g., 4 papers to be published in JPR SI 2018) despite of its very short history.
- **New Opportunity:** In the long term, as neXt-CP50 or neXt-CP2000 project is moving well with help from B/D-HPP and Resource Pillars, we anticipate more opportunities will be given to study on the basic biology, resulting in a good impact on proteomics and biomedical sciences.

Thank you!