π-HuB: The Proteomic Navigator of the Human Body

WHITE PAPER

V1.0

The π-HuB Consortium

2023. 05

We appreciate your time and contribution to the π-HuB project. This is a draft of the white paper requesting for more comments. We look forward to your feedbacks to help us further refine and improve it. If you have any comments, please send to pi-hub@ncpsb.org.cn. Thanks in advance.

Abstract

In the post-Human Genome Project era, the research of human proteome has been one of the most exciting and challenging frontiers of life sciences and medicine. Cutting edge mass spectrometry and other transformative proteomic technologies now provide an opportunity to interrogate the human body and human life at unprecedented resolution and scale. Here we introduce an international big-science project called the π-HuB (The Proteomic Navigator of the Human Body). The π-HuB project will be a 30-year mission with an investment of multi-billion, built upon four key pillars, including human samples, technology innovations, big-science infrastructures and open-access resources. The project has three central goals: 1) to dissect the human body into a hierarchy of digital reference space of the proteome from organs/tissues to single cells; 2) to take the snapshot of individuals' proteomes in lifetime and investigate populations' proteomic adaptation to major exposures on health outcomes; 3) to build an intelligent computational model called the π-HuB navigator which integrates proteomic and other molecule/phenotype data to promote our understanding of human biology, facilitate disease diagnosis and therapy. During the initiation and development stage (2023-2032), the π-HuB project will form an international consortium and achieve the following milestones, including cell-type resolved proteome atlases, life-oriented perturbation proteome atlases and the initial version of the 'proteomic navigator' model. We expect that these efforts will give the whole π-HuB project a significant boost, ushering in an era of proteomics-driven precision medicine. Projecting forward, the π-HuB project will further involve global collaboration and discussion integrating the output from a worldwide community of multi-disciplinary scientists. Taken together, we anticipate that the π-HuB project might greatly contribute to biomedical research over the next several decades.

Highlights

● The π-HuB project will be an ambitious 30-year mission to interrogate the human proteome at unprecedented resolution and scale.

● The π-HuB project will boost technology innovations for the next-gen proteomics and protein sequencing.

● The π-HuB project will accelerate development of novel diagnostic tests and drug discovery.

● The π-HuB project will usher us in an era of Proteomics-Driven Phronesis Medicine.

Introduction

Human biology is tightly linked to the protein 'universe' in the body, or the also called the proteome (the complete set of proteins encoded by a genome), as proteins are primary effectors of biological functions[1]. Moreover, proteins are the key drivers of diseases and the main molecular targets of most therapeutics including drugs and vaccines. Hence, in the post-HGP (the Human Genome Project) era, the research of human proteome has been one of the most exciting and challenging frontiers of life sciences and medicine [2, 3].

In an announcement timed to coincide with the publication of the human genome sequence[4, 5], a group of proteomics researchers launched the international Human Proteome Organization (HUPO)[6]. Since its inception, HUPO has stimulated and coordinated many workshops regarding the Human Proteome Project (HPP). In September 2010, HUPO took the first step towards an international effort by proposing a genetic/chromosome-centric approach to find evidence for all human protein-coding genes using mass spectrometry, serving as the backbone of the HPP[7]. Ten years later, the first high-stringency HPP blueprint was assembled in 2020, which covers >90% of the human proteome, paralleling progress made by the HGP[8]. In addition, many biology/disease-centric initiatives were proposed under the umbrella of HPP to spatially measure and interpret human proteome data under a range of physiological and pathological conditions in terms of abundance, post-translational modifications, interaction partners, and localization. For example, the first proteomics project of a human organ (the Human Liver Proteome Project, HLPP) was launched in 2003[9-11], leading to the characterization of the protein expression profiles and protein-protein interactions in this metabolic organ as well as the discovery of a major role of acetylation in metabolic regulation[12-15]. The proteomes of various tissues or organs (e.g., brain[16], heart[17], stomach[18], skin[19], immunol cells[20], and so on) have since been sequentially characterized, thereby building up the initial version of the organ/tissue-based human proteome map[21-26]. Meanwhile, increasingly more disease proteomes are being analyzed by the Chinese Human Proteome Project (CNHPP) [27], the National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium (CPTAC)[28], the Tumor Profiler Project[29], and so on. These efforts collectively brought the whole HPP project a significant boost, moving us towards the era of Proteomics-Driven Precision Medicine (PDPM)[30-34].

Despite these advances, we are only beginning to scratch the surface. With no doubt, there are still tremendous works to be done pursuing the ultimate goal of the HPP, 'translating the code of life'. It is actually a two-fold challenge. On the aspect of human beings, molecular diversification is particularly pronounced in the human proteome that shapes approximately 37.2 trillion cells with diverse morphologies and functions in the body. On the aspect of human life, an individual proteome is highly dynamics in lifetime and can be diversely remodeled by diverse factors, such as the human microbiome (The microbial ecosystems residing in various habitats of the human body), the type of lifestyle, and the state of

environment, which are intimately linked to human health and well-being.

Funded by the Minister of Science and Technology of China, ~40 teams of proteomics researchers worldwide came together in 2020 to brainstorm future HPP projects. We have since established several multi-disciplinary working groups, organized numerous on-site meetings or webinars, and communicated with government and private funders. These fruitful discussions then allowed us to propose a big-science project called the π-HuB (The Proteomic Navigator of the Human Body, Fig. 1). The project aims to form a consortium of Chinese and international scientists to generated unprecedented proteomics data sets from all human tissue types and cells in the body, and build an intelligent computational engine called the π-HuB navigator, which integrates proteomic and other omics data to promote our understanding of human biology, facilitate disease diagnosis and drug discovery.

Ultimate Goals of the π-HuB Project

The π-HuB project will be a 30-year mission with an investment of billions of RMB, aiming to achieve three ultimate goals by scientists all over the world.

- ***Charting A 'Reference-Space' of The Human Proteome*** The π-HuB project will dissect the human body into a hierarchy of digital reference space of the proteome. That is, by harnessing the rapid-evolving techniques like single-cell and spatial proteomics[35-37], to digitalize the complete quantitative composition of a human body, including the cell composition of all major tissues/organs, the protein composition of individual cell types and single cells, and the proteome-centric molecular networks within cells (*e.g.*, proteoform crosstalks, protein complexes, protein-protein interactions, DNA/RNA-protein interactions, protein-metabolite interactions).
- ***Defining Diverse 'State-Spaces' of The Human Proteome*** The π-HuB project will take the snapshot of individuals' proteomes in lifetime and investigate populations' proteomic adaptation to major exposures on health outcomes. Specifically, the π-HuB project aims 1) to trace the proteome-centric trajectories in five prenatal stages including gametogenesis, fertilization, embryonic development, fetal development, and delivery, and the five postnatal stages including childhood, adolescence, pregnancy, menopause, and healthy aging; 2) to profile dynamic changes in complex proteomes during the development and progression of complex diseases (e.g., malignant tumors, cardiovascular diseases, neurodegenerative diseases, metabolic diseases, infectious diseases, etc.); 3) to investigate the effects of non-genetic factors, such as the symbiotic microbiomes, lifestyles (e.g., alcohol, smoking, tea/coffee, diet, food, nutritional supplement, physical activity), and environments (e.g., hot/warm/cold climate, arid, polar/highland climate), on the human proteome as well as the underlying mechanism of short-term acclimatization and long-term adaptation.
- ***Modeling a Meta-Homo Sapiens of The Human Body*** The π-HuB project will build the π-HuB

navigator, a molecule level digital projection of a *homo sapiens which is based on the proteomics and other omics data of human specimens generated in this project. The* π-HuB navigator is in a virtual state-space, created by the convergence of virtually enhanced physiological phenotypes and digital reality (e.g., spatial-temporal biochemical information) on cells, body fluids, tissues and organs. This *meta homo sapiens* model consists of time-sequential frames with each one containing the spatially distributed single-cell proteomic data measured and augmented within a unit period to represent the human body state at a timestamp. The data frame will be formulated according to a 3D anatomical hierarchy that records the digital features of organs, tissues, body fluids and cells at each level respectively. One can instantiate the static data model with real-world data thoroughly measured from healthy adults and therefore generate a few referential state points in the human body state space. Then, the entire human body state space is dissected into multiple subspaces that are dissected through various dimensions, including different growing/aging stages, living environments and disease status.

Pillars for Building the π-HuB Navigator

The π-HuB project will be built upon four key pillars as follows.

- ***Human samples*** Human samples are the fundamental component of the π-HuB project. To achieve the objectives outlined above, samples for π-HuB can be grouped into the following categories. First, anatomy-based cohorts will harvest freshly prepared organs and tissues from ethnically diverse donors. Second, population-based cohorts will cross-sectionally collect high-quality biospecimens (e.g., blood, urine, feces, microbiomes, bone marrow and other bodily fluids) from a large number of individuals in diverse geographical regions with different lifestyle and environment. Third, time-based cohorts will apply non-invasive approaches with relatively high sampling frequency to longitudinally sample healthy individuals or patients with defined exposures on health or therapeutic outcomes over time. Practically, for example, the π-HuB project will first obtain samples from existing resources in China, for example, the National Human Brain Bank for Development and Function, the 500,000-person blood-based Kadoorie Biobank[38], among many others. On the other hand, we will work closely with other clinical centers and biobanks all over the world. In this regard, all samples profiled in the π-HuB project should be well annotated with clinical information obtained from multiple resources, such as the questionnaires, physical measurements, and biochemical tests, wearable device-based records, among others. Furthermore, the π-HuB project will underscore the necessity of rare diseases and large-scale longitudinal follow-up cohorts, as learned from a successful precedent - the UK Biobank[39].
- ***Technology Innovations*** The π-HuB project will first adopt and standardize the combination of several state-of-the-art sample processing methods (e.g., PCT[40], SISPROT[41]), MS data acquiring approaches (e.g., HRMS1-DIA[42], diaPASEF[43]) and freely available computational tools (e.g., MaxQuant[44],

FragPipe[45], pFind[46]) that can process diverse sample types into harmonized collection of metadata. Further, it is important to point out that π-HuB is a long-term mission that will be implemented through the development, integration, and application of cutting-edge single-cell proteomic technologies and the next generation of protein sequencing[47]. First and foremost, there is now a pressing need for both identifying and measuring the minute amount of protein in single cells[48]. In the past few years, we have witnessed tremendous progresses in the development of MS-based single-cell proteomics technologies, such as nanoPOTS/OAD[49, 50], SCoPE-MS[51, 52] and DVP[53]. Meanwhile, novel concepts and technologies for single-molecule protein sequencing also hold a substantial potential to enable broad sequence coverage in single-cell profiling[54]. However, at this time, a real-world, large-scale application of existing technologies for whole-proteome characterization in single cells, particularly in absolute quantification and multiplexed measurements, is yet to be achieved. Just like HGP that drives technology development of nucleic acid sequencing, we envision that the π-HuB project will also boost the development of increasingly powerful and effective single-cell and single-molecule proteomics. In this regard, new approaches and tools developed by the π-HuB researchers will also be integrated into the pipeline as they mature. Second, new challenges in data analysis will become more pronounced as single-cell proteomic measurement technology improves. New computational methods are apparently warranted to specifically address questions that can benefit from multimodal measurements. Data-driven modeling approaches such as deep learning have exhibited powerful performance in approximating many virtual and real-world systems, however, only transforming a biological "black box" system into a digital one does not provide us with any intellectual knowledge or insight and consequently prevents the model from trustworthy clinical practice. In contrast, the π-HuB project will push the boundaries of human reason in biomedicine by unveiling the reconstruction theory of human body. Inspired from the success in mathematical intuition guidance and hypothesis proposal, artificial intelligence (AI) methods and other yet to conceived approaches will be exploited to interpret a well-fitted deep learning model of human body system with upgrading resolutions, from cellular level to molecular level and revealing corresponding reconstruction theories. With these theories, several "white box" prototype *meta homo sapiens* model will be reconstructed and will serve as the critical preliminary condition to build the control method of the system.

- ***Big-Science Infrastructures*** Since very large numbers of human samples (e.g., sorted single cells from the donated human organs and biospecimens from clinical cohorts) will be generally measured in the π-HuB project, ultra-throughput facilities for data collection and processing are required. Ideally, such facilities require the expertise and streamlined pipelines to process human samples, to profile the proteome-centric molecular data in the samples, and to store, transfer, process and interpret the data. Obviously, a large monetary investment is required, which is not feasible or cost effective for

individual academic laboratories or even core facilities at most research institutions. We believe a solution is to establish national facilities/centers as Big-science infrastructures for the π-HuB project that would make ultra-throughput proteomics accessible to all researchers in our Consortium. Minimally, at the first stage of the π-HuB project, such infrastructures should be able to process 1,000 samples per day and 1 TB data per hour. In China, only a few existing programs (e.g., PHOENIX, OMNI) that make such analytical capacities with an automation workstation for 'one-stop' sample preparation, over 40 cutting-edge high-resolution mass spectrometers and a high-performance computing system 'Tianhe-II'. In addition, many other Big-science infrastructures and National Laboratories in China are also committed to support the π-HuB project, further bringing in state-of-the-art single cell technologies and a Cloud-based high-performance AI-computing system 'Pengcheng'. Regardless, the π-HuB project will also include many existing infrastructures or research unites attached to institutions or universities worldwide as research centers.

- ***Open-Access Resources*** To ensure that each π-HuB research center can produce high-quality and unified proteome-centric datasets, the π-HuB consortium is committed to share Standard Operating Procedure (SOPs) for biospecimen collection and annotation, sample handling and preparation, instrument setting and maintenance, data acquiring, processing and analysis. Nonetheless, given the dynamic methodology implementation, the consortium will also develop and share standard, testing samples and benchmarking data to help each research center to develop improved protocols and new techniques. Just as importantly, the π-HuB project will emphasize highly-efficient, international and open-access resources, including samples and data. Like other large community resources with broad utility, the project will also require an open-sharing framework to ensure clear global collaboration between researchers, funding agencies, and stakeholders. In this framework, π-HuB will maximize the importance of re-using collected human samples and re-analyzing data so that people can benefit from scientific advances, while minimizing risks to participant privacy and acknowledging the contributions of researchers. However, specific restrictions may arise across study participants and across jurisdictions. In this regard, we will negotiate different regulatory restrictions on the collection, use, and international sharing of samples and data, especially those related to identifiable individuals. If necessary, sensitive samples or data will be sealed in appropriate access control sections or locally, while allowing fewer sensitive data to be shared openly. Where safety measures are required, simplified access processes will be established to ensure that qualified, trusted researchers can quickly access data for legitimate research purposes without compromising data security. In some jurisdictions, researchers contributing to this project may be required to seek international sharing permission or ethical immunity from individuals. When privacy risks are negligible, it is advisable for researchers to seek public consent from participants so that the international community can access and reuse them

without restrictions. If the privacy risk is low, researchers should seek broad consent for future sharing and use, subject to ongoing and transparent governance mechanisms. Specific consent requirements for restricted areas of use may still pose obstacles to data sharing in some countries. Continued assessment of privacy risks is necessary to ensure that data sharing does not expose participants to inappropriate re-identification risks and the risk of harmful disclosure or misuse of sensitive health information. Security measures should be commensurate with the risk of data breach or misuse and should leverage existing community monitoring and compliance infrastructure. Lastly, the π-HuB project will not only share data to the international scientific community (for instance, all π-HuB-generated raw data will be directly available to the international scientific community through several well-established cloud-based data portals, such as ProteomExchange[55], iProx[56], and so on), but also enable non-experts to freely enquire medical intervention strategies by developing a web-based "*Meta-homo sapiens*" computational framework building upon molecular and spatial data.

Vision of Proteomics-Driven Phronesis Medicine

"The greatest medicine of all is teaching people how not to need it", says Heppocrates. Although marvel breakthroughs have been achieved in diagnosis and treatment in modern medical paradigms, the progress in disease prevention has gradually and greatly lagged behind. The difficulty of precision prevention stems from the lack of mathematically precise and complete definition of human body state so that there is no accurate predictive model of future states. Such a model for human body state prediction will enable us to identify reliable, easy-to-apply markers for risk diagnostic and progression assessment of preclinical disease, and thus to prevent or reduce morbidity and/or mortality via high-risk population screening, early disease identification and subsequent therapy.

Building a π-HuB navigator model directly empowers clinical applications based upon the proteome-centric 'navigation'. For instance, the prototype model will be transferred from ideal body conditions (healthy adult males and females) to different developing and aging stages, progression of diseases, symbiosis, nutrient and environment conditions to obtain near reality models. It will be followed by creating a state space storing all key states of human body by simulating body dynamics with each model for certain periods. Ultimately, causal inference will be used to distinguish the underlying triggers to induce transforms between adjacent key states from their discrepancies and the state representation of each state are compressed to a minimum number of clinically detectable biomarkers required to distinguish with other states. Each state space can be regarded as a topological navigation map where each node is the key state with a biomarker representation and each edge between two nodes records the triggers to transform from one state to the other. We envisioned that landing this miraculous innovation instantiates a new practice of the reconstruction theory and will grant an opportunity to initiate an unprecedentedly new paradigm of

medicine to provide temporally precise control of human body state so as to prevent disease, ushering in Proteomics-Driven Phronesis Medicine (Fig. 2).

The Phase One π-HuB

The scope of the π-HuB project is obviously unprecedented. To enable the project as a broadly applicable project targeting its ultimate goals, it is necessary to set deliverables on staged programs in a relatively short timeframe to be relevant to the community. During the initiation and development stage (2020-2030, herein referred to as 'the phase one'), the project will build an international cooperation network and achieve the following milestones.

- ***Cell-Type Resolved Proteome Atlases*** The π-HuB project will ultimately generate single-cell resolution atlases of all major human organs or tissues (e.g., liver, stomach, and lung) of different races (e.g., Negroid, Mongoloid, Australoid). Currently, however, proteomic analysis of each single cell in any organ is not feasible due to the technical limitations of existing methods (i.e., relatively high cost and low throughput as compared to single cell transcriptomics). During the first stage, we therefore plan to initially build up the reference cell-type proteome atlases for all major organs using the combination of state-of-the-art flow sorting[57], lase-capture microdissection[58], tissue expansion [cite our ProteomEx paper], paralleling MS acquisition technologies[59]. Moreover, rapidly evolving spatial proteomics technologies will provide additional insights to the secreted proteins in the microenvironment and subcellular localization of the proteome[36]. Compare with many genome/transcriptome-centric projects, these proteome-centric atlases with cell-type resolution will offer unique insights in the structures and functions of different cell types and tissue. In this regard, Synergies of the π-HuB project with the HuBMAP[60], the Human Cell Atlas[61], the Human Tumor Atlas Network[62], and the LifeTime Initiative[63] are obviously natural. In addition, the π-HuB project will intrinsically synergize with other initiatives in the field of proteomics. For instance, the Grand Challenge project, the HUPO's flagship initiative for future HPP, is going to get better understanding of a function for every protein[64]. In fact, few proteins act alone. Instead, they typically exert biological functions by constituting pathways and networks in certain space and time. In this regard, the high-resolution spatial information for every protein in human body should be of great value for understanding protein function.
- ***Life-Oriented Plastic Proteome Atlases*** The proteome of each individual is not only highly 'plastic (i.e., capable of being shaped or reshaped)' across different stages during the lifetime, but also associated with a number of genetic and non-genetic factors, such as lifestyles, environments and the symbiotic microbiomes. Currently, however, the cost of proteomic analysis is not affordable to most populations, even in developed countries. During the first stage, we therefore plan to focus on the most

dominate factors that shape or remodel the human proteome. Specifically, we plan to accumulate the first one million perturbed proteome of biofluids and tissues from large-scale cohort studies, aiming 1) to trace the proteomic trajectory in lifetime by quantifying dynamic changes in the proteomes of five major prebiotic cycles (e.g., gametogenesis, fertilization, embryonic development, fetal development, and delivery) and five major postbiotic cycles (e.g., adolescence, puberty, gestation, menopause, and old age); 2) to analyze the effects of four major dietary nutrition patterns (i.e., Western, Japanese, Mediterranean and their subsistence) on the human proteome; 3) to map the proteomes of populations in six major ecological environments (e.g., hot/warm/cold climate, arid, polar/highland climate), and analyze the trajectory of human proteome to habitat patterns-perception-response-acclimatization-adaptation; 4) to map one hundred types of microbiomes from internal and external environment and its interaction with human proteome, and to construct the adaptation trajectory of human proteome under the symbiotic mode.

- ***Organ-centric Disease Proteome Atlases.*** In the past decade, mounting evidence have already demonstrated that MS-based proteomic approaches can facilitate mechanistic understanding of diseases as well as its related biomarker discovery and therapy development. For instance, MS-based proteomics has joined forces with genomics and transcriptomics for comprehensive molecular characterization of numerous human tumors[28, 65, 66]. Moreover, MS-based proteomics has solely identified potential biomarkers or potential therapeutic targets for many tumors as well (e.g., early-stage hepatocellular carcinoma[30], melanoma[67], lung adenocarcinoma[31], pancreatic ductal adenocarcinoma[34], ovarian cancer[68]) and a variety of other diseases, such as Coronavirus disease 2019[32], non-alcoholic fatty liver disease[69], Parkinson's disease[70], type 2 diabetes[71], Alzheimer's disease[72] and so on. During the first phase of this project, establishing paths to further implement proteome-driven approaches for the better understanding of complex diseases is therefore to be a natural extension. As such, we plan to map the proteomes of all major organs and corresponding biofluids at different pathophysiological stages, with a focus on 3-5 representative diseases for each related organ. Such analysis, together with aforementioned tissue proteome atlases with cell-type resolution and biofluid proteome atlases with perturbation, is able to construct the proteomic evolutionary trajectories of the occurrence and development of these diseases and their association pathways with specific life stages, life states, and survival conditions. Furthermore, the π-HuB project will actively collaborate with clinicians, policy makers and industrial partners to catalyze the discovery of novel protein-based biomarkers that can be applied in clinic to diagnose diseases, and discover novel drug targets.
- ***The π-HuB Navigator 1.0.*** All the proteome data generated in this project will be projected into a virtual space and processed with cutting-edge AI-based algorithms that have emerged as one of the

most promising methods in helping explain the complex relationships between molecular layers and phenotypes[73, 74]. These efforts will result in an initial version of the 'navigator' (herein referred to as the π-HuB navigator 1.0) which, to certain extent, could depict and predict the physiological and pathological processes inside the human body. Newly obtained proteomic data could therefore be fed into the π-HuB navigator 1.0, leading to in-depth interpretation of the data, facilitating the discovery and development of intervention and therapeutic strategies for proteomics-driven precision medicine.

Outlook

Since its inception in 2020, the π-HuB consortium has grown to be a 100-member international joint force through mobilizing scientists around the world across academic, industrial, and government sectors in protein sciences. Projecting forward, undoubtedly, the π-HuB project will further involve global collaboration and discussion (even between competing entities) integrating the output from a worldwide community of multi-disciplinary scientists, while attracting young scientists to this exciting project. We anticipate that the π-HuB project may greatly contribute to biomedical research over the next several decades, facilitating disease prevention and diagnosis, accelerating drug discovery, and ultimately ushering in an era of proteomics-driven phronesis medicine.

References

1. Aebersold, R. & Mann, M. Mass-spectrometric exploration of proteome structure and function. *Nature* **537**, 347-355 (2016).
2. Abbott, A. And now for the proteome. *Nature* **409**, 747 (2001).
3. Fields, S. Proteomics. Proteomics in genomeland. *Science* **291**, 1221-1224 (2001).
4. Lander, E.S. et al. Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921 (2001).
5. Venter, J.C. et al. The sequence of the human genome. *Science* **291**, 1304-1351 (2001).
6. The proteome isn't genome II. *Nature* **410**, 725 (2001).
7. A gene-centric human proteome project: HUPO--the Human Proteome organization. *Mol Cell Proteomics* **9**, 427-429 (2010).
8. Adhikari, S. et al. A high-stringency blueprint of the human proteome. *Nat Commun* **11**, 5301 (2020).
9. Cyranoski, D. China takes centre stage for liver proteome. *Nature* **425**, 441 (2003).
10. Jia, H. & Louet, S. China pushes liver proteomics. *Nat Biotechnol* **22**, 136 (2004).
11. Service, R.F. Proteomics. Public projects gear up to chart the protein landscape. *Science* **302**, 1316-1318 (2003).
12. Zhao, S. et al. Regulation of cellular metabolism by protein lysine acetylation. *Science* **327**, 1000-1004 (2010).
13. Wang, Q. et al. Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. *Science* **327**, 1004-1007 (2010).
14. He, F. Human liver proteome project: plan, progress, and perspectives. *Mol Cell Proteomics* **4**, 1841-1848 (2005).
15. Wang, J. et al. Toward an understanding of the protein interaction network of the human liver. *Mol Syst Biol* **13**, 965 (2017).
16. Sharma, K. et al. Cell type- and brain region-resolved mouse brain proteome. *Nat Neurosci* **18**, 1819-1831 (2015).
17. Doll, S. et al. Region and cell-type resolved quantitative proteomic map of the human heart. *Nat Commun* **8**,

1469 (2017).
18. Ni, X. et al. A region-resolved mucosa proteome of the human stomach. *Nat Commun* **10**, 39 (2019).
19. Dyring-Andersen, B. et al. Spatially and cell-type resolved quantitative proteomic atlas of healthy human skin. *Nat Commun* **11**, 5587 (2020).
20. Rieckmann, J.C. et al. Social network architecture of human immune cells unveiled by quantitative proteomics. *Nat Immunol* **18**, 583-593 (2017).
21. Kim, M.S. et al. A draft map of the human proteome. *Nature* **509**, 575-581 (2014).
22. Wilhelm, M. et al. Mass-spectrometry-based draft of the human proteome. *Nature* **509**, 582-587 (2014).
23. Eraslan, B. et al. Quantification and discovery of sequence determinants of protein-per-mRNA amount in 29 human tissues. *Mol Syst Biol* **15**, e8513 (2019).
24. Wang, D. et al. A deep proteome and transcriptome abundance atlas of 29 healthy human tissues. *Mol Syst Biol* **15**, e8503 (2019).
25. Jiang, L. et al. A Quantitative Proteome Map of the Human Body. *Cell* **183**, 269-283 e219 (2020).
26. Uhlen, M. et al. Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).
27. Cyranoski, D. China pushes for the proteome. *Nature* **467**, 380 (2010).
28. Rodriguez, H., Zenklusen, J.C., Staudt, L.M., Doroshow, J.H. & Lowy, D.R. The next horizon in precision oncology: Proteogenomics to inform cancer diagnosis and treatment. *Cell* **184**, 1661-1670 (2021).
29. Irmisch, A. et al. The Tumor Profiler Study: integrated, multi-omic, functional tumor profiling for clinical decision support. *Cancer Cell* **39**, 288-293 (2021).
30. Jiang, Y. et al. Proteomics identifies new therapeutic targets of early-stage hepatocellular carcinoma. *Nature* **567**, 257-261 (2019).
31. Xu, J.Y. et al. Integrative Proteomic Characterization of Human Lung Adenocarcinoma. *Cell* **182**, 245-261 e217 (2020).
32. Shen, B. et al. Proteomic and Metabolomic Characterization of COVID-19 Patient Sera. *Cell* **182**, 59-72 e15 (2020).
33. Nie, X. et al. Multi-organ proteomic landscape of COVID-19 autopsies. *Cell* **184**, 775-791 e714 (2021).
34. Shi, Y. et al. Targeting LIF-mediated paracrine interaction for pancreatic cancer therapy and monitoring. *Nature* **569**, 131-135 (2019).
35. Kelly, R.T. Single-cell Proteomics: Progress and Prospects. *Mol Cell Proteomics* **19**, 1739-1748 (2020).
36. Mund, A., Brunner, A.D. & Mann, M. Unbiased spatial proteomics with single-cell resolution in tissues. *Mol Cell* **82**, 2335-2349 (2022).
37. Li, L. et al. Spatially resolved proteomics via tissue expansion. *Nat Commun* **13**, 7242 (2022).
38. Chen, Z. et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *Int J Epidemiol* **40**, 1652-1666 (2011).
39. Collins, R. What makes UK Biobank special? *Lancet* **379**, 1173-1174 (2012).
40. Guo, T. et al. Rapid mass spectrometric conversion of tissue biopsy samples into permanent quantitative digital proteome maps. *Nat Med* **21**, 407-413 (2015).
41. Chen, W. et al. Simple and Integrated Spintip-Based Technology Applied for Deep Proteome Profiling. *Anal Chem* **88**, 4864-4871 (2016).
42. Xuan, Y. et al. Standardization and harmonization of distributed multi-center proteotype analysis supporting precision medicine studies. *Nat Commun* **11**, 5248 (2020).
43. Meier, F. et al. diaPASEF: parallel accumulation-serial fragmentation combined with data-independent acquisition. *Nat Methods* **17**, 1229-1236 (2020).
44. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**, 1367-1372 (2008).
45. da Veiga Leprevost, F. et al. Philosopher: a versatile toolkit for shotgun proteomics data analysis. *Nat Methods* **17**, 869-870 (2020).
46. Chi, H. et al. Comprehensive identification of peptides in tandem mass spectra using an efficient open search engine. *Nat Biotechnol* **36**, 1059-1061 (2018).
47. MacCoss, M.J. et al. Sampling the proteome by emerging single-molecule and mass spectrometry methods. *Nat Methods* **20**, 339-346 (2023).
48. Perkel, J.M. Single-cell proteomics takes centre stage. *Nature* **597**, 580-582 (2021).
49. Zhu, Y. et al. Nanodroplet processing platform for deep and quantitative proteome profiling of 10-100 mammalian cells. *Nat Commun* **9**, 882 (2018).
50. Li, Z.Y. et al. Nanoliter-Scale Oil-Air-Droplet Chip-Based Single Cell Proteomic Analysis. *Anal Chem* **90**, 5430-5438 (2018).

51. Petelski, A.A. et al. Multiplexed single-cell proteomics using SCoPE2. *Nat Protoc* **16**, 5398-5425 (2021).
52. Huffman, R.G. et al. Prioritized mass spectrometry increases the depth, sensitivity and data completeness of single-cell proteomics. *Nat Methods* (2023).
53. Mund, A. et al. Deep Visual Proteomics defines single-cell identity and heterogeneity. *Nat Biotechnol* **40**, 1231-1240 (2022).
54. Alfaro, J.A. et al. The emerging landscape of single-molecule protein sequencing technologies. *Nat Methods* **18**, 604-617 (2021).
55. Deutsch, E.W. et al. The ProteomeXchange consortium at 10 years: 2023 update. *Nucleic Acids Res* **51**, D1539-D1548 (2023).
56. Chen, T. et al. iProX in 2021: connecting proteomics data sharing with big data. *Nucleic Acids Res* **50**, D1522-D1527 (2022).
57. Shapiro, E., Biezuner, T. & Linnarsson, S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet* **14**, 618-630 (2013).
58. Espina, V. et al. Laser-capture microdissection. *Nat Protoc* **1**, 586-603 (2006).
59. Goncalves, E. et al. Pan-cancer proteomic map of 949 human cell lines. *Cancer Cell* **40**, 835-849 e838 (2022).
60. Hu, B.C. The human body at cellular resolution: the NIH Human Biomolecular Atlas Program. *Nature* **574**, 187-192 (2019).
61. Elmentaite, R., Dominguez Conde, C., Yang, L. & Teichmann, S.A. Single-cell atlases: shared and tissue-specific cell types across human organs. *Nat Rev Genet* **23**, 395-410 (2022).
62. Rozenblatt-Rosen, O. et al. The Human Tumor Atlas Network: Charting Tumor Transitions across Space and Time at Single-Cell Resolution. *Cell* **181**, 236-249 (2020).
63. Rajewsky, N. et al. LifeTime and improving European healthcare through cell-based interceptive medicine. *Nature* **587**, 377-386 (2020).
64. Omenn, G.S. et al. The 2022 Report on the Human Proteome from the HUPO Human Proteome Project. *J Proteome Res* (2022).
65. Mani, D.R. et al. Cancer proteogenomics: current impact and future prospects. *Nat Rev Cancer* **22**, 298-313 (2022).
66. Zhang, B. et al. Clinical potential of mass spectrometry-based proteogenomics. *Nat Rev Clin Oncol* **16**, 256-268 (2019).
67. Harel, M. et al. Proteomics of Melanoma Response to Immunotherapy Reveals Mitochondrial Dependence. *Cell* **179**, 236-250 e218 (2019).
68. Eckert, M.A. et al. Proteomics reveals NNMT as a master metabolic regulator of cancer-associated fibroblasts. *Nature* **569**, 723-728 (2019).
69. Niu, L. et al. Noninvasive proteomic biomarkers for alcohol-related liver disease. *Nat Med* **28**, 1277-1287 (2022).
70. Virreira Winter, S. et al. Urinary proteome profiling for stratifying patients with familial Parkinson's disease. *EMBO Mol Med* **13**, e13257 (2021).
71. Wigger, L. et al. Multi-omics profiling of living human pancreatic islet donors reveals heterogeneous beta cell trajectories towards type 2 diabetes. *Nat Metab* **3**, 1017-1031 (2021).
72. Johnson, E.C.B. et al. Large-scale deep multi-layer analysis of Alzheimer's disease brain reveals strong proteomic disease-related changes not observed at the RNA level. *Nat Neurosci* **25**, 213-225 (2022).
73. Kang, M., Ko, E. & Mersha, T.B. A roadmap for multi-omics data integration using deep learning. *Brief Bioinform* **23** (2022).
74. Mann, M., Kumar, C., Zeng, W.F. & Strauss, M.T. Artificial intelligence for proteomics and biomarker discovery. *Cell Syst* **12**, 759-770 (2021).

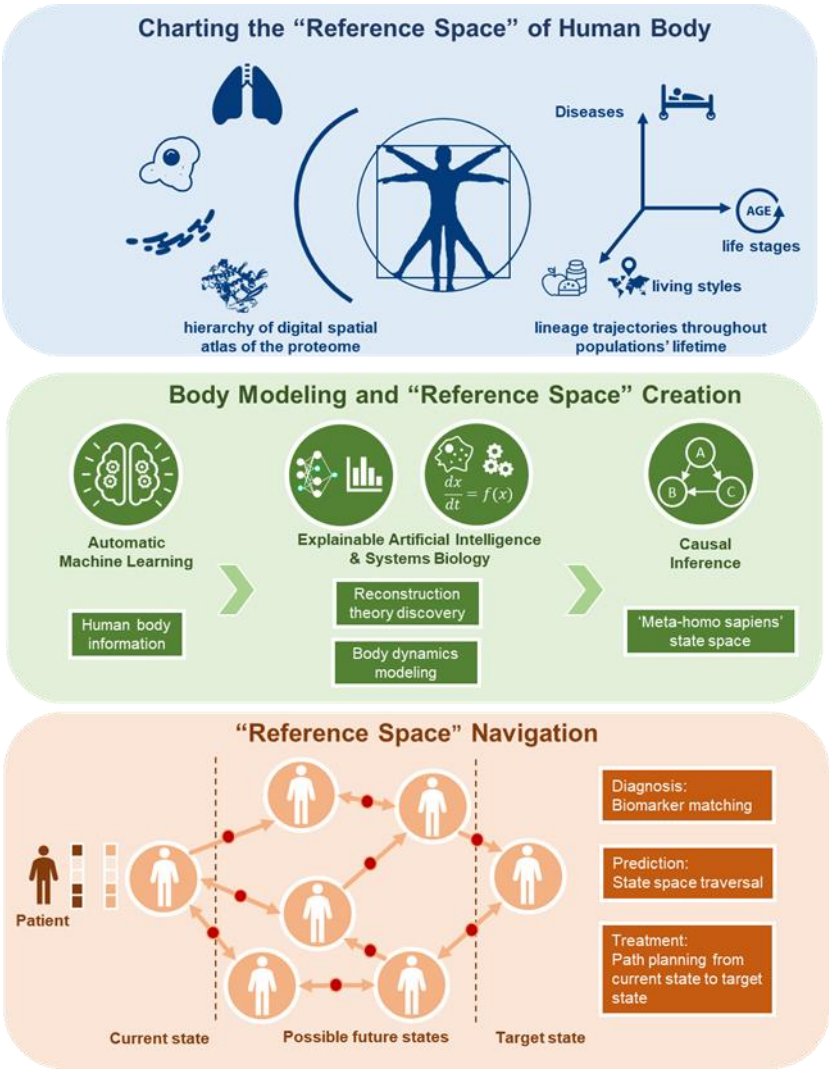

Fig. 1. Building a Proteomic Navigator of the Human Body.

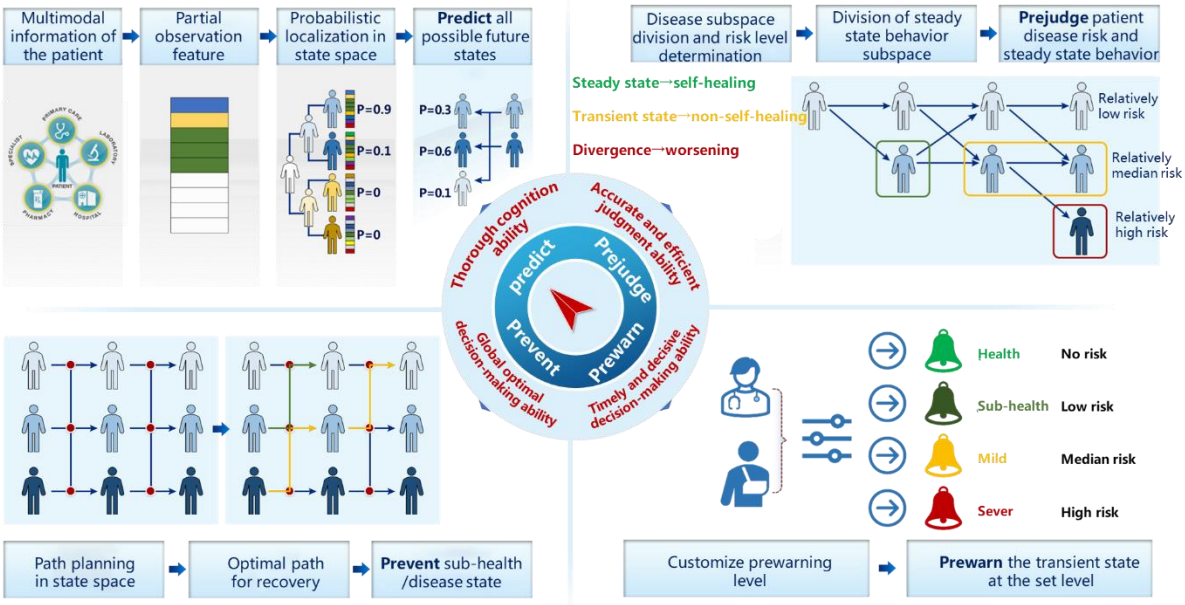

Fig. 2. Vision of Proteomics-Driven Phronesis Medicine. With the accumulation of data-driven research instead of hypothesis-driven research, the enrichment of multi-dimensional dynamic data rather than one-dimensional static data, the integration of information between human cells and internal and external environment and the new cognition of human individuals, human and the environment, and human and society will be obtained, thus ushering in the phronesis medicine. Here we tentatively define phronesis medicine: Phronesis medicine shall realize the accurate, efficient diagnosis and treatment capabilities and highly robust decision-making capabilities for disease prevention, control and health care, which will eventually establish a medical model of popularization and normalization of diagnosis and treatment decisions and health management. It contains the ultimate observation of information on diagnosis, treatment, prevention and control of diseases as well as health care, cognition of disciplinary knowledge and summarization of general rules, accurate and efficient diagnosis and treatment decision-making ability based on big data and artificial intelligence technology to process massive information, and highly robust disease prevention and health management decision-making ability. In order to realize phronesis medicine, it is necessary to predict and prevent events such as instability, imbalance, disorder, and illness, so as to ensure the health of entire population and life cycle and to enhance the overall interests of society. Functional molecular big data of proteome is the most important driving force.