Published in final edited form as:

Curr Pharmacogenomics Person Med. 2010 December 1; 8(4): 240–244. doi:
10.2174/187569210793368230.

Nutriproteomics and Proteogenomics: Cultivating Two Novel Hybrid Fields of Personalized Medicine with Added Societal Value

Vural Ozdemir^{1,*}, Jean Armengaud², Laurette Dubé³, Ramy Karam Aziz⁴, and Bartha M. Knoppers¹

¹ Centre of Genomics and Policy, Department of Human Genetics, Faculty of Medicine, McGill University, Montreal, QC, Canada

² Laboratoire de Biochimie des Systèmes Perturbés (LBSP), Institut de Biologie Environnementale et Biotechnologie (IBEB), Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA), Bagnols-sur-Cèze, F-30207, France

³ Desautels Faculty of Management, McGill University and the McGill World Platform (MWP) for Health and Economic Convergence, Montreal, QC, Canada

⁴ Department of Microbiology and Immunology, Faculty of Pharmacy, Cairo University, Cairo, Egypt

Keywords

Biomarkers; biotechnology foresight; gene-centric human proteome project; nutriproteomics; personalized nutrition; population biobanks; prospective policy; proteogenomics

1. NUTRIPROTEOMICS: A NEW SUBSPECIALTY IN PERSONALIZED MEDICINE

Personalized medicine requires diagnostics that enable customization of health interventions such as drugs, vaccines, stem cell therapy and nutrition [1–6]. It is important to approach personalized medicine with this broader outlook, rather than a narrow focus on drug therapy, as nutrition and other health interventions influence health [7]. Such global vision is also needed in order to integrate information from multiple levels of the biological hierarchy from genome to proteome to metabolome, and ways in which these biological parts interact with each other, the environment, and society more generally [8].

In the past, nutrition research and public health programs have focused on adequate access to food or alleviation of nutritional deficiencies. This framework, however, has shifted considerably over the past decade. Recognition of population heterogeneity in nutritional and adverse responses to food led to a greater emphasis on understanding the molecular

*Correspondence to this author at the Centre of Genomics and Policy, Department of Human Genetics, Faculty of Medicine, McGill University, 740 Dr. Penfield, Suite 5200, Montreal, QC, Canada H3A 1A4. vural.ozdemir@mcgill.ca; Telephone: (514) 398-6920; Fax: (514) 398-8954.

basis of this variability and by extension, on the possibility of nutritional interventions customized at a subpopulation level.

In the December issue of the CPPM, Kussmann introduces a new subfield of personalized medicine research: nutritional proteomics or nutriproteomics [9]. This emerging application of proteomics involves the characterization and quantification of food-derived bioactive peptides and proteins, and discerning the mechanisms by which proteome variation impacts nutrition related health outcomes. Kussmann notes that nutriproteomics science sits on a metaproteomics approach representing a synthesis of information from three different proteome levels: host, food and resident microbes. While food (from animal, plant, or microbial sources) and host proteomes have been intensively studied in the past, proteomes of human-associated resident microbes, i.e., the intestinal microbial metaproteome, have received relatively less attention [10]. Hence, nutriproteomics presents an opportunity for personalized medicine to integrate proteomics information from microbiome, host and food.

2. ANTICIPATING THE NUTRIPROTEOMICS FUTURE(S)

Nutrition represents a fundamental and ever present environmental exposure that is essential to sustain cellular life and population health. Hence, anticipating the future trajectories of nutriproteomics is of substantial interest both to innovators and to the end-users of this knowledge in personalized medicine. Health technology assessment (HTA) is a widely used tool to support decisions on the future of technologies in health care, public policy and business since the 1970s. More recently, HTA evolved beyond deterministic analyses of “impacts” of a new technology. Modern HTA does not assume that technologies or their social environment are necessarily static [11]. Instead, we are increasingly witnessing a model based on a dynamic co-evolution of technology and society, or real-time technology assessment [12]. Traditional HTA questions such as “Do we adopt/reject a technology, given that it is now well developed and mature?” are being replaced with “How can a new technology and its applications be co-designed and governed collaboratively early on, together with innovators and anticipated end-users?” This signals an upstream shift in HTA to the stage of technology design before its applications enter the society. In effect, this presents at the same time an important opportunity to negotiate a social contract between science and society, and to steer the biotechnology future(s) towards those that are likely to have desirable and sustainable impacts on global public health. Hence, modern HTA efforts underscore the co-construction of innovations and their social context.

Insofar as the current context of personalized medicine innovations is concerned, the rise of nutriproteomics coincides with two related, and potentially transformational, data intensive biotechnology initiatives:

1. the recent proposal for a gene-centric Human Proteome Project (HPP) [13, 14] and,
2. the new field of proteogenomics [15].

In this concise editorial commentary, we discuss how nutriproteomics, the HPP and proteogenomics might together cultivate a favourable ground for a productive alliance

between genomics and proteomics [16], and for novel strategies to study complex nutritional endpoints such as obesity and healthy eating [17].

3. TOWARDS A GENE-CENTRIC HUMAN PROTEOME PROJECT

At the protein level, we still have scant knowledge of the 20,300 predicted protein-coding human genes discovered through the Human Genome Project (HGP). For the proteins encoded by most of these genes, the abundance, distribution, subcellular localization, post-translational characteristics, interaction networks, and function are still poorly understood. Recently, a group of researchers under the umbrella of the Human Proteome Organization (HUPO) proposed a federated international effort to map the protein complement of the human genome, a gene-centric HPP [13]. Given the enormous complexity of the human proteome and its variation in each person over time, normal physiology and disease in different cell types, the HPP is being conceived as a systematic effort to deliver on reasonable and achievable end-points. Accordingly, this group of proteomics researchers aims “to ensure that, for each predicted protein-coding gene, at least one of its major representative proteins will be characterized in the context of its major anatomical sites of expression, its abundance, and its interacting protein partners” [13]. The information from the HPP will be made publicly available with no restrictions, as was done with the HGP [13]. An independent paper, coinciding with the HPP proposal, further argued that the mass spectrometry, a significant driver of high-throughput proteomics analysis, has now advanced to a stage that can produce data with improved reproducibility and should be ready for such a large international project [18]. Indeed, novel breakthroughs in mass spectrometry technology are anticipated over the coming years. As a result, this will accelerate the intensive research program presently being conceived to enable the HPP. Moreover, systems biology and synthetic biology concepts that are presented by Armengaud in this issue of the CPPM offer new perspectives for exploiting proteomics data that will be generated through the HPP [19].

While vigorous debates on whether the HPP should be gene- or protein-centric are expected [14, 20], the call made for an international HPP brings about formidable traction to the field of proteomics and nutriproteomics. This project might also draw the genomics and proteomics communities much closer. Both communities share a complementary vision on data intensive 21st century science [8, 21, 22] that aims to understand the human biology and the interacting environmental-societal factors at a systems level, for advances in diagnostic, preventive, prognostic and therapeutic end-points. Conceivably, the HPP will also stimulate spinoff novel proteomics platforms capable of higher throughput analyses, as with genomics technologies after the HGP, and thus drive down the cost of proteomics applications for personalized medicine in the future.

4. PROTEOGENOMICS: A PRODUCTIVE ALLIANCE OF PROTEOMICS AND GENOMICS

Genome annotation involves finding and delineating structural genes and attributing a biological function to them. This is important not only for the human genome but also for genomes from microbial pathogens and non-pathogens that are becoming available as efforts

to map the human microbiome are accelerating [10]. Genome annotation, especially assigning functions to genes and consistently classifying these functions using a controlled vocabulary, is an arduous and dynamic process that improves over time as knowledge and understanding of gene products (e.g., proteins) accumulate. To this end, proteogenomics is a new alliance of genomics and proteomics that substantially benefited the annotation of genomes [15, 16, 23]. Proteogenomics involves high-throughput identification and characterization of proteins by extra-large shotgun mass spectrometry approaches and the integration of these data with genomic data. In essence, proteomics provides *orthogonal* data for genome annotation, as a complement to the DNA-centric evidence used to predict protein-coding sequences and gene function, or the RNA-centric evidence revealing transcribed loci. These orthogonal data will be integrated as essential parts of automated genome annotation pipelines (e.g., RAST), especially as the number of sequenced genomes increases exponentially [24]. Notably, proteogenomics analyses of the bacterium *Deinococcus deserti* VCD115, isolated from the Sahara surface sand, allowed identification of 15 unpredicted genes, and importantly, reversal of incorrectly predicted orientation of 11 genes [25]. Numerous N-termini of proteins were also corrected with such protein-centric evidence [25]. Proteogenomics thus offers another mechanism for effective linkage of the two new sister fields in personalized medicine – nutrigenomics and nutripoteomics – so that high-throughput data can be integrated and reconciled not only from the microbiome, its host and from food but also between their genomic and proteomic complements. In the case of human-associated microbes, such proteomic data are particularly valuable, since they offer crucial evidence of the viability of these microbes, which cannot be deduced from genomic data.

Alliance of genomics and proteomics presents a unique opportunity to weave nature and nurture in understanding the complex nutrition related outcomes such as obesity and healthy eating. Despite intensive studies on food choice and consumption, food over-consumption and obesity remain as a global threat, and an epidemic impacting both developed and low and middle income countries (LMICs). New approaches to combat the obesity epidemic include, for example, the Brain-to-Society (BtS) model of motivated adaptive behaviours (e.g., food choice); these aim to address the nested challenges emerging from cognitive and biological susceptibility as well as their interaction with dynamic societal, economic and policy systems [26]. Proteomics biomarkers could aid as an integrated measure of environmental influences (including of social systems) on the host, together with the constitutive genomic markers of obesity, food consumption or food choices (e.g., see the work on chocolate consumption and craving) [27, 28]. Proteomics expands our toolbox to enable personalized medicine and better delineate complex phenotypes and motivated adaptive behaviors such as food preference. Still, when confronted with pervasive uncertainty in the governance of human-environment interactions and the linked social-ecological systems that underlie healthy eating and obesity, caution is necessary to avoid panaceas or universal remedies that claim to heal all diseases [29].

A keyword search of the Pubmed database on October 14, 2010 returned five (nutripoteomics) and 41 (proteogenomics) records for these two novel hybrid fields in personalized medicine. This, however, is likely to change. Biotechnologies related to nutrition have implications for the global society at large. The current proposal for an

international federated effort to map the human proteome, and the examples of a new productive alliance between genomics and proteomics outlined above, might further accelerate these emerging subfields of personalized medicine from a linear to an exponential growth phase in the next few years.

5. POPULATION BIOBANKS: READY FOR NUTRIPROTEOMICS AND PROTEOGENOMICS DATA?

The HPP proposes to characterize the “normal-ome” (non-pathological state) with respect to protein variation in different cell types. The first challenge in such a substantive project concerns the following question and dilemma: on which basis samples should be chosen and regarded as “normal”? Some biobanks are already collecting samples for plasma, urine, or other fluids and tissues for proteome analysis. Another challenge is the marked scale up necessary for sample collection in a manner that is suitable for both genomic and proteomic analyses [30]. Population biobanks would be well served by investing in large informatics capabilities in order to optimally utilize the new proteome catalogue to be generated by the HPP for the study of their prospective population cohorts. Repositories of proteomics data should be more stringent for the quality of data deposited by proteomic platforms. Harmonization and integration of data from the proteomics datasets will add further complexity to future population based omics research.

Large scale data intensive science projects (e.g., HGP, the Biomarker Consortium) require tool building, e.g., population biobanks and a data commons, which can be creatively mined, further driving applied and conceptual innovations. Precompetitive collaboration, defined as competitors sharing early stages of research that benefit all [31], is one possible approach to tool building in this early stage of nutripoteomics and proteogenomics science on the critical path to personalized medicine. In these efforts, it is essential, however, to avoid the creation of a false hierarchical dichotomy between infrastructure tool-building science and subsequent discovery oriented science; both are inseparable and rely on each other to materialize. Finally, innovation in ethics and governance frameworks are needed for proteomics, as with the genomics data intensive science.

CONCLUSIONS

Social science critique of technology has been effective in identifying the complex linkages between science and society [12]. On the other hand, such linkages were often discerned *post-facto*, after a technology future is locked into a certain trajectory, either due to proliferation of expectations or after firm beliefs and strong opinions are formed on benefits/risks of technology applications. An important corollary is that such *post-facto* social science engagement does not allow policy interventions to maximize benefits from technologies or avoid undesirable impacts [7, 32]. In the current complex and data intensive 21st century innovation ecosystems [21], fostering sustainable personalized medicine innovations calls for real-time technology assessment as well as prospective and participatory science policy.

At this early phase of nutriproteomics, proteogenomics and the current definition of what could be the HPP, there is ample room for real-time engagement of proteomics and genomics data intensive sciences with social science and policy research (Box 1). Past lessons from, and the progress made in the fields of nanotechnology [33, 34] and climate change [35] are valuable in this context. Furthermore, a renewed alliance of genomics and proteomics offers a viable strategy to weave nature and nurture in personalized medicine research and practice. The recent call for a gene-centric HPP and the emerging field of proteogenomics already attest to the promise of such an alliance. Through upstream engagement from the outset, and together with real-time analyses of novel biotechnology applications, we can increase the likelihood of cultivating innovations with added societal value [7, 8, 12].

Box 1

Foresight on Points to Consider to Cultivate Nutriproteomics and Proteogenomics Innovations with Added-Societal Value

Science

- Tool-building and infrastructure science are essential first steps to enable future discoveries in novel data intensive sciences such as nutriproteomics and proteogenomics.
- Pre-competitive collaboration warrants further careful attention as a potential enabler of early phase tool-building in these emerging fields of personalized medicine.
- Dynamic, user-friendly, web-based servers that allow the personalized medicine community to continuously revise and re-annotate genes are timely.
- International federated efforts are needed to scale up biobanking and data intensive sciences towards a productive alliance between genomics and proteomics so that both systems biology and personalized medicine are enabled.
- A population and global public health focus needs to be cultivated for nutriproteomics and data intensive sciences, given their broad and growing importance in both developed and low and middle-income countries (LMICs).

Society

- Measures to prevent a false hierarchy or dichotomy between infrastructure science and discovery science are needed. These two domains of 21st century data intensive science are inseparable and firmly depend on each other. Hence, a new culture of collaboration beyond individual based entrepreneurship is timely.
- Parallel innovations in research ethics and science governance are essential to enable large-scale data intensive science.

- Measures to bring together the complex systems in both biology and society in understanding nutrition, healthy eating and obesity should be adopted.
- A culture of ongoing, real-time, evidence-based vigilance is necessary to reflect on how best to foster personalized medicine innovations with added societal value.

Acknowledgments

We thank Ahmed El-Sohehy (University of Toronto) for a critical review of the editorial manuscript and Don Husereau (CADTH, Ottawa) for discussions on constructive technology assessment. This manuscript was supported in part by an operating research grant from the Canadian Institutes of Health Research (#84620) and a career investigator salary for science and society research in personalized medicine from the Fonds de la recherche en santé du Québec to Ozdemir.

ABBREVIATIONS

HGP	Human Genome Project
HPP	Human Proteome Project
HTA	Health technology assessment
HUPO	Human Proteome Organization
LMICs	Low and middle-income countries

References

1. Patel SA, King CC, Lim PK, et al. Personalizing stem cell research and therapy: the arduous road ahead or missed opportunity? *Curr Pharmacogenomics Person Med*. 2010; 8(1):25–36. [PubMed: 20563265]
2. Kaput J, Astley S, Renkema M, et al. Harnessing nutrigenomics: development of web-based communication, databases, resources, and tools. *Genes Nutr*. 2006; 1(1):5–11. [PubMed: 18850216]
3. Castle, D., Cline, C., Daar, AS., et al. Science, society and the supermarket: the opportunities and challenges for nutrigenomics. Hoboken, NJ: Wiley-Interscience; 2006.
4. Garcia-Bailo B, Toguri C, Eny KM, et al. Genetic variation in taste and its influence on food selection. *OMICS*. 2009; 13(1):69–80. [PubMed: 18687042]
5. Poland GA, Ovsyannikova IG, Jacobson RM. Application of pharmacogenomics to vaccines. *Pharmacogenomics*. 2009; 10(5):837–52. [PubMed: 19450131]
6. Joly Y, McClellan KA, Knoppers BM. Personalized vaccines and public health genomics: anticipating and monitoring the ELSIs. *Curr Pharmacogenomics Person Med*. 2010; 8(1):4–6.
7. Ozdemir V, Husereau D, Hyland S, et al. Personalized medicine beyond genomics: New technologies, global health diplomacy and anticipatory governance. *Curr Pharmacogenomics Person Med*. 2009; 7(4):225–30. [PubMed: 20613883]
8. Ozdemir V, Suarez-Kurtz G, Stenne R, et al. Risk assessment and communication tools for genotype associations with multifactorial phenotypes: the concept of “edge effect” and cultivating an ethical bridge between omics innovations and society. *OMICS*. 2009; 13(1):43–61. [PubMed: 19290811]
9. Kussmann M. Nutriproteomics – Linking proteomics variation with personalized nutrition. *Curr Pharmacogenomics Person Med*. 2010; 8:245–56.
10. Rizkallah MR, Saad R, Aziz RK. The human microbiome project, personalized medicine and the birth of pharmacomicrobiomics. *Curr Pharmacogenomics Person Med*. 2010; 8(3):182–93.

11. Douma KF, Karsenberg K, Hummel MJ, et al. Methodology of constructive technology assessment in health care. *Int J Technol Assess Health Care*. 2007; 23(2):162–8. [PubMed: 17493301]
12. Guston DH, Sarewitz D. Real-time technology assessment. *Technol Soc*. 2002; 24(1–2):93–109.
13. A gene-centric human proteome project: HUPO--the Human Proteome Organization. *Mol Cell Proteomics*. 2010; 9(2):427–9. [No authors listed]. [PubMed: 20124355]
14. The call of the human proteome. *Nat Methods*. 2010; 7(9):661. [No authors listed]. [PubMed: 20827799]
15. Armengaud J. Proteogenomics and systems biology: quest for the ultimate missing parts. *Expert Rev Proteomics*. 2010; 7(1):65–77. [PubMed: 20121477]
16. Armengaud J. A perfect genome annotation is within reach with the proteomics and genomics alliance. *Curr Opin Microbiol*. 2009; 12(3):292–300. [PubMed: 19410500]
17. Blouin C, Dubé L. Global health diplomacy for obesity prevention: lessons from tobacco control. *J Public Health Policy*. 2010; 31(2):244–55. [PubMed: 20535105]
18. Nilsson T, Mann M, Aebersold R, et al. Mass spectrometry in high-throughput proteomics: ready for the big time. *Nat Methods*. 2010; 7(9):681–5. [PubMed: 20805795]
19. Armengaud J. Systems biology and synthetic biology: Understanding biological complexity on the critical path to personalized medicine. *Curr Pharmacogenomics Person Med*. 2010; 8:257–67.
20. Rabilloud T, Hochstrasser D, Simpson RJ. Is a gene-centric human proteome project the best way for proteomics to serve biology? *Proteomics*. 2010; 10(17):3067–72. [PubMed: 20648483]
21. Hey, T., Tansley, S., Tolle, K. *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Microsoft Research; Redmond, Washington: 2009.
22. Kolker E. A vision for 21st century US Policy to support sustainable advancement of scientific discovery and technological innovation. *OMICS*. 2010; 14(4):333–5. [PubMed: 20673064]
23. de Groot A, Dulerio R, Ortet P, et al. Alliance of proteomics and genomics to unravel the specificities of Sahara bacterium *Deinococcus deserti*. *PLoS Genet*. 2009; 5(3):e1000434. [PubMed: 19370165]
24. Aziz RK, Bartels D, Best AA, et al. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics*. 2008; 9:75. [PubMed: 18261238]
25. Baudet M, Ortet P, Gaillard JC, et al. Proteomics-based refinement of *Deinococcus deserti* genome annotation reveals an unwonted use of non-canonical translation initiation codons. *Mol Cell Proteomics*. 2010; 9(2):415–26. [PubMed: 19875382]
26. Dubé, L., Bechara, A., Dagher, A., et al. *Obesity Prevention: The Role of Brain and Society on Individual Behavior*. Elsevier; New York: 2010.
27. Rezzi S, Ramadan Z, Martin FP, et al. Human metabolic phenotypes link directly to specific dietary preferences in healthy individuals. *J Proteome Res*. 2007; 6(11):4469–77. [PubMed: 17929959]
28. Aziz RK. A hundred-year-old insight into the gut microbiome! *Gut Pathog*. 2009; 1(1):21. [PubMed: 19968883]
29. Ostrom E, Janssen MA, Anderies JM. Going beyond panaceas. *Proc Natl Acad Sci U S A*. 2007; 104(39):15176–8. [PubMed: 17881583]
30. Kraut A, Marcellin M, Adrait A, et al. Peptide storage: are you getting the best return on your investment? Defining optimal storage conditions for proteomics samples. *J Proteome Res*. 2009; 8(7):3778–85. [PubMed: 19397304]
31. Weber, S. *The Success of Open Source*. Harvard University Press; Cambridge, MA: 2004.
32. Knoppers BM. Genomics and policymaking: from static models to complex systems? *Hum Genet*. 2009; 125(4):375–9. [PubMed: 19252928]
33. Guston D. 2007Toward anticipatory governance. Accessed October 23, 2010Available from: <http://nanohub.org/resources/3270>
34. Barben, D., Fisher, E., Selin, C., et al. Anticipatory governance of nanotechnology: Foresight, engagement and integration. In: Hackett, EJ,Amsterdamska, O,Lynch, M., Wajeman, J., editors. *The Handbook of Science and Technology Studies*. Cambridge, MA: MIT Press; 2008. p. 979-1000.

35. Quay R. Anticipatory governance -- A tool for climate change adaptation. *J Am Plann Assoc.* 2010; 76(4):496–511.