Original Communication

Discovery-Based Nutritional Systems Biology: Developing N-of-1 Nutrigenomic Research

Jim Kaput[1] and Melissa Morine[2]

[1] Clinical Translation Unit, Nestle Institute of Health Sciences, Lausanne, Switzerland
[2] Department of Mathematics, University of Trento, Italy and The Microsoft Research-University of Trento Centre for Computational, Rovereto, Italy

Abstract: The progress in and success of biomedical research over the past century was built on the foundation outlined in R.A. Fisher's *The Design of Experiments* (1935), which described the theory and methodological approach to designing research studies. A key tenet of Fisher's treatise, widely adopted by the research community, is randomization, the process of assigning individuals to random groups or treatments. Comparing outcomes or responses between these groups yields "risk factors" called population attributable risks (PAR), which are statistical estimates of the percentage reduction in disease if the risk were avoided or in the case of genetic associations, if the gene variant were not present *in the population*. High throughput metabolomics, proteomic and genomic technologies provide 21st century data that humans cannot be randomized into groups: individuals are genetically and biochemically distinct. Gene – environment interactions caused by unique dietary and lifestyle factors contribute to heterogeneity in physiologies observed in human studies. The risk factors determined for populations (i.e., PAR) cannot be applied to the individual. Developing individual risk or benefit factors in light of the genetic diversity of human populations, the complexity of foods, culture and lifestyle, and the variety of metabolic processes that lead to health or disease are significant challenges for personalizing dietary advice for healthy or medical treatments for individuals with chronic disease.

Key words: group level analysis, experimental design, systems nutrition research

Biological and biomedical research, like many established human activities, is based on conventions and practices that become standardized over time. Many of these conventions are decades old and do not account for the great advances in knowledge resulting from modern research. Some of the main practices of biomedical research are discussed herein, in an attempt to initiate a dialogue in the nutrition community about new concepts for personalized nutrition and health.

The Design of Experiments

The design of biomedical research experiments rests on standards developed almost 100 years ago. The methodology was codified in RA Fisher's *The Design of Experiments* [12] a 1935 treatise that described the key elements of good research protocols. These benchmarks are a valued requisite for determining the quality of experimental results and Fisher's experimental design is now considered the gold standard for human

DOI 10.1024/0300-9831/a000128

Table I: Characteristics of Experimental Designs.

Randomization	Assign individuals at random to groups or to different groups in an experiment
Replication	Identify the sources of variation, better estimate the true effects of treatments, strengthen the experiment's reliability and validity
Blocking	Reduces known but irrelevant sources of variation between units for estimating the source of variation under study
Orthogonality	The forms of comparison (contrasts) that can be legitimately and efficiently carried out. Comparisons are uncorrelated and independently distributed if the data are normal.
Factorial	Evaluate the effects and possible interactions of several factors
Experiments	(independent variables)

biomedical research studies. Experiments designed according to Fisher's principles (see Table I), when applied to the appropriate scientific question, produce reliable evidence for improving public and personal health. More humans survive childbirth, infancy, childhood, and through advanced age than any time in our species history [44].

One of the key tenets of designing experiments is the requirement for randomization of subjects or other outbred animals to case and control groups. Randomization distributes unknown or unmeasurable characteristics between groups in an attempt to isolate and identify measurable variables that differ between or distinguish those groups. While this step was necessary in the 20th century, the underlying rationale for its use is no longer valid in the post-genomic era, especially for variables that produce small biological effects. For example, one of the key variables that could not be analyzed in sufficient detail was an individual's genetic make-up. However, genome technologies now allow for the characterization at the DNA sequence level (rev in [27, 34]) and soon at the epigenome (i. e., DNA methylation) level (e. g., [10, 11]). Genetic variation is no longer unmeasurable or unknowable. The necessity to analyze genetic makeup has been demonstrated from results of numerous genome sequencing projects: the latest data [8, 43] suggests that individuals differ from each other and reference genomes by about 3.5 million single nucleotide polymorphisms (SNPs), almost 1000 large copy number variants (CNVs), and large numbers of insertions and deletions (indels). An individual's genetic make-up differs from others and has to be measured and incorporated into biomedical research studies at the risk of producing unreliable and irreproducible results.

Randomization has serious consequences on the ability to interpret data from experiments where the response of a group exposed to an intervention is compared to a group not exposed (control). Genetic diversity will usually result in phenotypic diversity. Physiological variability has been recognized for millennia and summarized in the modern pre-genomic era by Williams in a 1956 book entitled *Biochemical Individuality* [46]. The ability to account for and measure individual phenotypic variation, however, has only become possible over the past two decades: high throughput tools and methods such as mass spectroscopy are now available to quantitatively measure human physiology, including in response to food (e. g., [5, 47]). A major challenge remains assessing environmental, cultural, and psycho-social variation but modern technology in the form of smart phones and tablets is increasingly used to measure an individual's activities, although linking these data to a common research database remains a challenge [42].

The edifice of statistical methods for group level research must also be re-evaluated:

- Comparing the average measure or response of one group (cases) against those measures in a second group (controls) yields the population attributable risk (PAR) [20, 26, 40]. Hence, PARs are population based – the average of a group – and these values do not apply to individuals because every individual is genetically distinct. PARs are applicable for large effect sizes (e. g., malaria, cholera) but have decreasing reliability as the difference in measures between groups decreases. No statistic is available to assess when PARs may provide knowledge that can be applied to all individuals in the population. The era of personalization requires the development of more personalized risk and benefit factor analysis.
- Power calculations (e. g., [13, 23–25]) do not consider genetic, phenotypic, and environmental diversity. Power analysis is used to calculate the minimum sample size required for detecting an effect of a given size. Alternatively, power calculations also allow one to determine the effect size that may be observed given the number of samples. Regardless of the approach, power analysis for studies of outbred species do not account for genetic, phenotypic, and environmental (such as diet) variation

between individuals. An unintended consequence of increasing the sample size to adequately power an experiment is that increasing the sample size of human studies results in increasing the "noise" while decreasing the signal within each group since each individual added to the study introduces differences in genetic, environmental, lifestyle, and cultural factors. When these factors are not measured in the study, this additional variability can only be considered as "noise" in the analysis.

Analyzing Variables One-By-One

Beadle and Tatum [1] published a seminal research concept and method in 1941 that is summarized as the one gene – one enzyme hypothesis. Although their experimental paradigm revolutionized biomedical research by demonstrating that a mutation in a single gene could alter enzyme activity, their complete hypothesis better reflects biological processes. Their treatise stated that although there would be simple one-to-one relationships, they also predicted that "…relations of great complexity" would also exist. However, their experimental results focused almost exclusively on auxotrophic mutations in Neurospora, which could be rescued by supplementing the media with the appropriate compound. The general applicability of their method contributed to the analysis of single compounds or genes. Even many genome wide association studies (GWAS) rely upon mathematical tools and criteria that analyze the association of multiple but independently-tested SNPs with some complex phenotype [33]. That is, each SNP is considered independently of interactions with other SNPs or genetic variants.

Most biological processes however, are relationships of great complexity and cannot often be characterized or understood by analyzing the effect or activity of a single gene, RNA, protein, enzyme, or metabolite. Organisms consist of systems consisting of interactions with the environment, interactions between organs, interactions between cells, and within cells of interconnected pathways, networks, and relationships. The complex organism is often studied by removing a component (e.g., metabolite or gene) from the cell system. Depending upon the importance of component, removing it from the system may force the system to channel flux through alternate pathways and networks. These concepts were also extended by Hartwell and coworkers [16–18] to the "normal" state in that flux through pathways is constrained by the requirements of life. If one reaction of the system is fast, another reaction may be slow, so that the overall flux is held within a tolerable range.

With this inherent interactivity of biochemical systems, it is difficult to obtain a clear understanding of the system by analyzing each constituent in isolation. Network-based approaches are well adapted for this purpose, as they consider these interactions as the basic framework for the analysis. The variety of approaches for network analysis considers the available data and the goal of the analysis. If high throughput measurements of system elements are available (such as transcriptomics or proteomics), network analysis can be used to identify the most active regions of the system with respect to a given intervention or phenotype. Alternatively, if kinetic information is available for the interactions in the system, dynamic simulation can provide predictions for system behavior following a given perturbation of interest.

System biology

Aristotle wrote "…the whole is greater than the sum of its parts" (rev. in [47]) thousands of years ago. The general concepts from such thinking were further developed in von Bertalanffy in General Systems Theory [2], the foundation of modern systems biology concepts [3]. Three key concepts of systems theory are emergence, openness, and concurrency. Emergence is an ability of the system to have capabilities greater than the sum or the parts – an automobile moves only because of the combination and integration of the parts, much like Aristotle's dictum. Machines have utility as metaphors but also do not capture other aspects of system thinking. For example, general systems theory also holds that biological systems are not "closed" like a machine but open [2]. That is, a biological system interacts with and adapts to its environment. A third related concept is concurrency, which describes interacting systems within machines or systems (e.g., a subsystem) that have internal distributed systems/embedded systems/sensor networks. The behavior of the subsystems is determined by the interaction of their internal state with the environment. Similarly, organisms interact with and respond to their environments.

Nutrigenomics describes this type of open, concurrent system by the term gene – environment interactions, classically defined as the statistical main effect of the interaction term [37]. Dominant mutations that cause catastrophic breaks in the system (e.g., inborn errors of metabolism [7]) and environments that over-

ride biological processes (such as cholera) are the exceptions to this rule. Biological systems respond to the signals from the environment such as nutrients, toxins, light exposure, but the system also adapts metabolism to respond to the environmental conditions by changes in transcriptional regulation or intracellular signaling that ultimately affect phenotype [21, 32, 36]. Embedded in this concept is that open systems are dynamic, changing due to changing interactions between environment and host.

The challenge of system thinking is defining the system. An analogy is Google Map: the ability to zoom from an earth view to a street view. Metabolites and proteins, that is reactions, may be considered the street view. The whole system or phenotype is the planet view. In between are layers of organization from the neighborhood of pathways to communities, counties, states, and nations in modules, networks, organ systems. Each layer provides information that can be used to produce some understanding of the biological processes. The current state of system research and analysis focuses on components and interactions such as between pathways or network-network interactions. The results are (usually) maps that attempt to distill knowledge of emergent properties of the biological systems that cannot be directly inferred from the properties of the constituent parts. High-throughput data generation provides the input for systems analysis, permitting a more complete assessment of functional intermediates (i. e., transcripts/proteins/metabolites) between genotype/environment and phenotype. Environmental factors such as diet and lifestyle and cultural and psychosocial dynamics have been notably absent from most biomedical research programs.

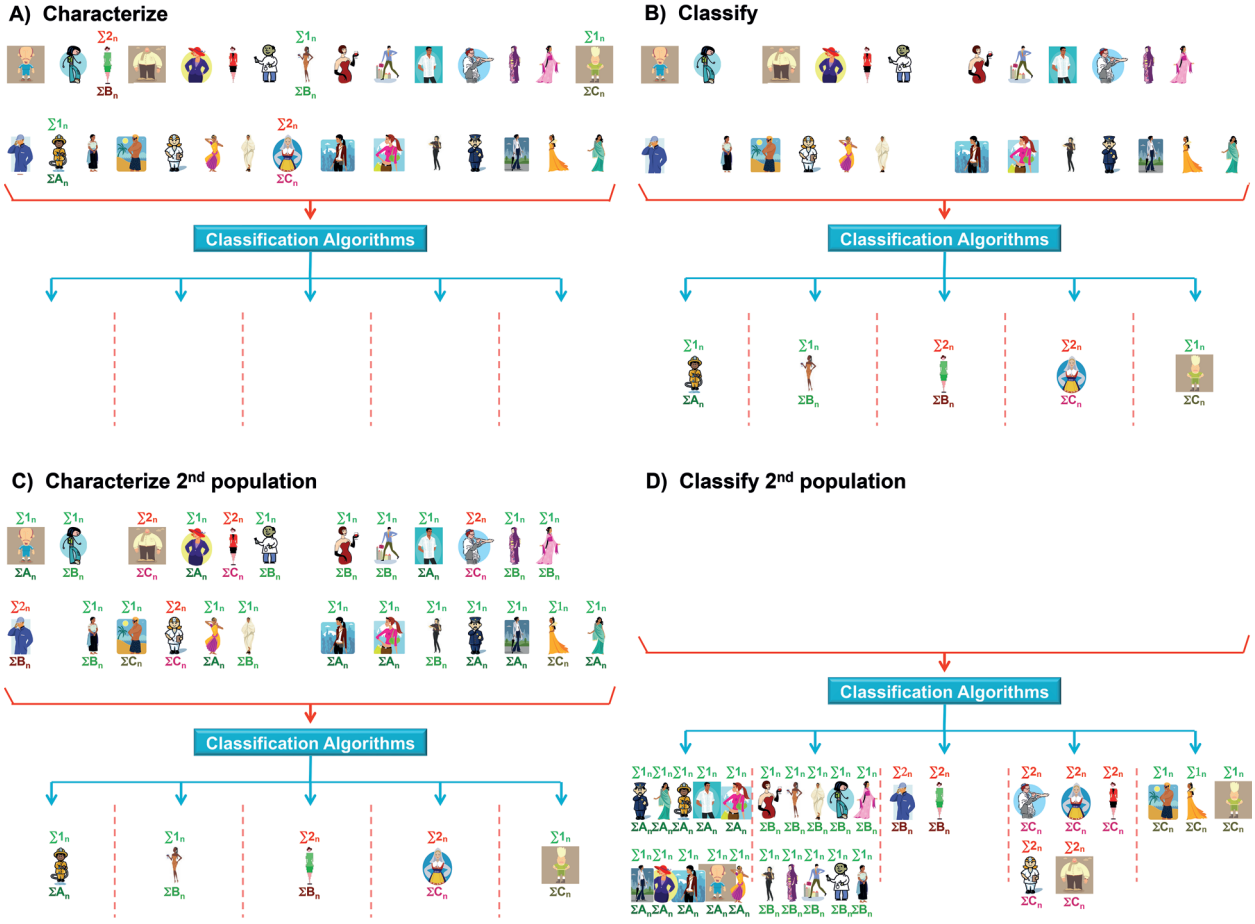

Figure 1: Conceptual strategy for n-of-1 studies. (A) Individuals in a population are "completely" characterized genetically and physiologically and environmentally. The genetic characterization is the sum [Σ of all genetic variations A, B, C and the physiological and environmental characterization is the sum [Σ of metabolomics, proteomic, lifestyle and other designated by numbers. After characterization, (B) classification algorithms are used to sort individuals into response groups that are empirically defined. To test these assignments, (C) individuals in a second population are analyzed as in (A) and then (D) classified as in (B). Each round of analysis and classification improves the probability that an individual is placed within the appropriate response or metabolic group.

While system thinking is essential for understanding biological processes, in most experimental cases, the data for the analyses are limited to metabolite, RNA, and protein in the blood, mRNA or microRNA from selected tissues, and genomic data. Only recently have attempts been made to integrate high dimensional data with measured environmental factors such as nutrient intake and physical activity, as well as multivariate (i.e., well-defined) phenotypes. That is, all of the components of the system are typically not analyzed, and perhaps, may never be analyzed. The final "systems" networks from such analyses can only be representations of the data acquired in the experiment, and as such, cannot be used to fully describe the final phenotype. Nevertheless, such representations better represent the data than two-dimensional, one-to-one relationships depicted in linear maps.

N-of-1 System Biology

These system biology concepts and technologies are beginning to alter experimental design and analyses (e.g., [15, 29–31]). However, most research conducted in the 21st century continues to be reductionistic, focusing on the effects of an individual nutrient or predicted consequences of an individual polymorphism in a gene to explain complex processes.

Strategies for analyses of individual data (often called n-of-1 data [35]) for identifying homeostatic groups and metabolic response groups (e.g., [19, 38]) may circumvent the limitations of case-control experimental designs. In n-of-1 designs, individuals are analyzed completely, or as completely as possible, and then analytical methods are used to find those of like response or characteristics (Figure 1). Such study designs are well suited to the use of unsupervised multivariate statistical approaches for class discovery from high throughput data, paired with network analysis for assessment of the functional context of observed patterns of variation. An interesting possibility from such analyses is that the grouping patterns of individuals at baseline may be different from the grouping patterns following a given intervention. From a translational perspective, these response groups will most likely help to inform optimal dietary interventions on an individual-specific basis. However there is more work to be done in terms of applying appropriate study designs and analytical strategies in order to understand nutritional systems.

Our approaches uses a cyclic method of analyzing a subset of individuals (Figure 1A) and classifying them (Figure 1B), This approach allows the same data sets to be aggregated and used for analysis of population level results [35]. A distinct challenge for this experimental design is the level of data necessary to characterize the system. For example, while it has been long known that humans (and other animals) are hosts to a large variety of microorganisms, the ability to analyze microbial niches in and on the human body became possible only with the advent of high throughput analytical technologies. Mammals are "holobiont" – defined as a multicellular eukaryote plus its colonies of persistent symbionts [14, 41]. Organisms present in gastrointestinal, oral mucosa, urogenital, skin, and airways are being counted, but their functions are challenging to study [22]. The view of n-of-1 has to be expanded to include the microorganisms who live with us.

Systems Nutrition Research

Nutrition research, defined here as the entire environment that a holobiont is exposed to, is more challenging to study. Clinical research methods can be used – that is, experiments conducted in highly controlled settings. However, translation of results from such studies is challenging because real life has a wider range of variables than laboratory or clinical settings. One approach to conduct translational n-of-1 research is through community-based participatory research study [28]. CBPR's central focus is developing a partnership among researchers and individuals in a community that allows for more in depth lifestyle analyses but also translational research that simultaneously helps improve the health of individuals and communities. As importantly, the unmeasureable factors or influences are incorporated into the overall physiological measures. Describing the setting of the research is metadata, which puts the research results into context even if not a part of the formal analytical process. While CBPR has been gaining much interest in the social and nutritional sciences fields [6, 39], relatively few studies have used this method for biomedical research [4, 9, 45]. CBPR is a process whereby the participants provide information and biological samples on an ongoing basis, and the biomedical researcher provides existing knowledge as well as results from the ongoing study. Research is "personalized" since one individual is assessed and informed even in the community setting. The applications are more immediate than population based methods and are targeted to the community and individual. Since genetic

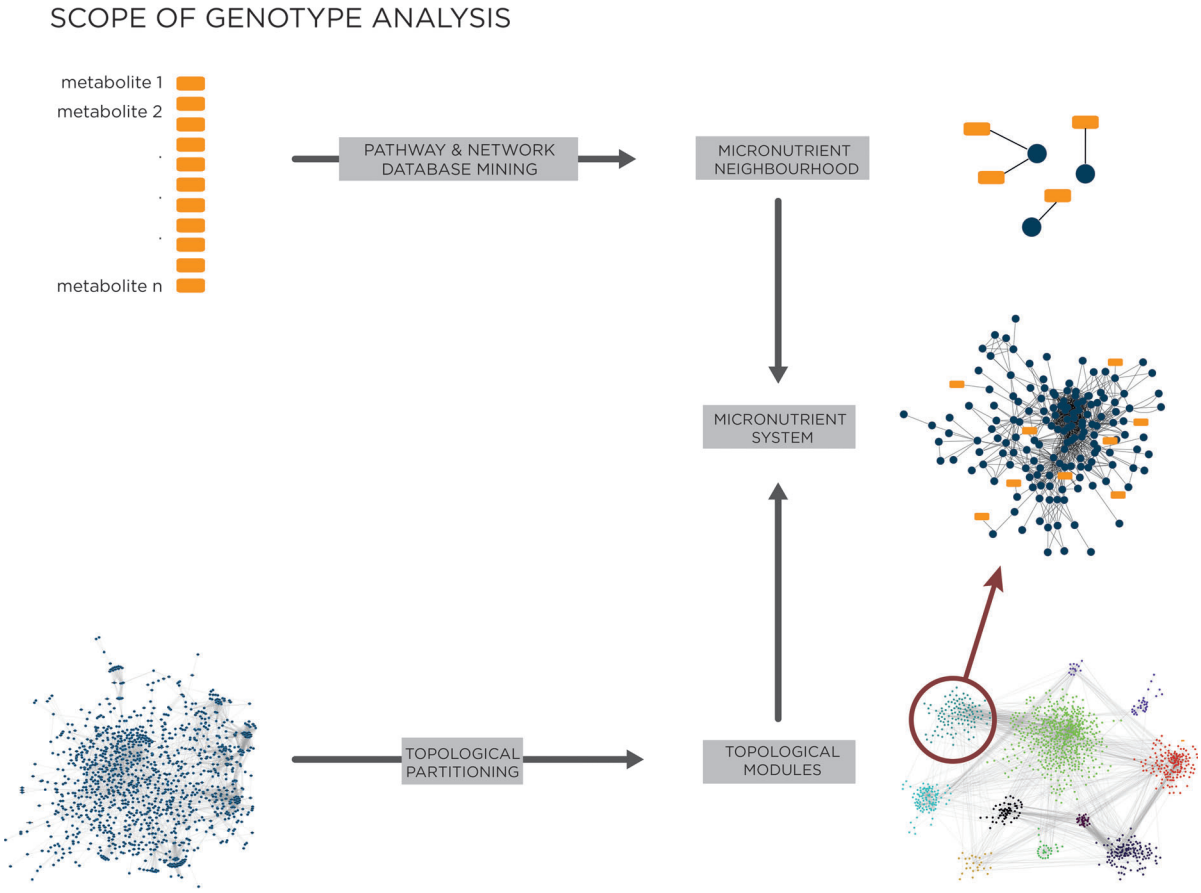

Figure 2: Middle Out Genotype Analyses. Two approaches were developed for the analyses of genotypes associated with metabolite levels or patterns. The top left shows metabolites that may be analyzed from subjects. An example may be vitamins in the blood samples. The pathways and networks that interact with, metabolize, or regulated by these metabolites are identified from pathway databases. This set of pathways and genes is called a neighborhood, and in this case, a micronutrient neighborhood. However, the pathways and networks may be unconnected due to the nature of pathway databases. A second approach uses a global interaction network derived from protein-protein interaction data. After topological partitioning, the function of the modules (subnetworks) are determined by pathways and genes within them (lower right). Modules which are enriched for micronutrient genes identify those systems most affected by changes in vitamin (in this example) status. In some cases, the neighborhood genes and one or more of the global network modules overlap as is shown in the micronutrient system in the middle right of this figure. Modules and subnetworks differ from pathways in that related functions are clustered. These representations provide a different approach to representing biological knowledge.

and omic data developed from population studies can not yet be reliably associated with health outcomes in individuals, the initial information flows between researcher and community collaborator focused on nutritional assessments and dietary advice. As more gene – nutrient or omic – nutrient associations are proven, the information flow will include biomedical data and results.

We applied this strategy to children and teens (ages 6 to 14), offering improved nutritious breakfast, lunch, and two healthy snacks during a 5-week summer day camp. Data aggregation (i. e., population level) results for nutrient intakes, healthy eating index scores, metabolite levels in plasma, and population genetic results were analyzed (Monteiro et al, in preparation). The n-of-1 analyses identified metabolic groups consisting of patterns of plasma vitamin levels and red blood cell metabolites. We also developed a middle-out (as opposed to top-down or bottom-up) procedure (Figure 2) to analyze genotype data by defining and testing associations of single nucleotide polymorphisms of genes involved in micronutrient metabolism, related gene networks, and their protein-protein interactions (Morine, Monteiro, et al, in preparation). Topological partitioning and enrichment analysis identified modules (subnetworks) enriched in genes encoding proteins that participate or regulate pathways associated with micronutrients. One of these micronutrient

modules was significantly enriched with SNPs that differed statistically between groups defined by patterns of metabolite levels.

Our n-of-1 experimental design accounts for individual genetic and lifestyle (including dietary) differences and produces data that can be analyzed at the population level, metabolic group level, and individual level. It has not escaped our attention that this experimental design can be applied to a variety of biomedical research questions.

Summary

The advent of technology permitting the more complete characterization of biological systems has provided data demonstrating the genetic, physiological, and biochemical diversity in the human (and other animal) specie. Analyzing complex processes cannot reasonably be done using approaches of measuring the effects of a single gene, protein, RNA, or enzyme or how a single nutrient or chemical alters the system. New experimental designs that account for the diversity of genotype X environment interactions need to be developed in order to improve personal and public health.

References

1. Beadle G and Tatum E. 1941. Genetic Control of Biochemical Reactions in Neurospora. *Proceedings of the National Academy of Sciences of the United States of America* 27: 499–506.

2. Von Bertalanffy L. 1950. The Theory of Open Systems in Physics and Biology. *Science* 111(2872):23–29.

3. Von Bertalanffy L. 1951. General system theory – A new approach to unity of science. *Human Biology* 23: 203–361.

4. Boyer Mohatt, GV, Pasker, RL, Drew, EM, and McGlone, KK BB. 2007. Sharing Results from Complex Disease Genetic Studies: A Community Based Participatory Research Approach. *International Jounal of Circumpolar Health* 66: 11–22.

5. Ceglarek U, Leichtle A, Brugel M, Kortz L, Brauer R, Bresler K, Thiery J, Fiedler GM, and Brügel M. 2009. Challenges and developments in tandem mass spectrometry based clinical metabolomics. *Mol Cell Endocrinol* 301 (1–2): 266–271.

6. Chen DT, Jones L, and Gelberg L. 2006. Ethics of clinical research within a community-academic partnered participatory framework. *Ethn Dis* 16 (1 Suppl 1): S118–35.

7. Copeland S. 2008. A review of newborn screening in the era of tandem mass spectrometry: what's new for the pediatric neurologist? *Semin Pediatr Neurol* 15 (3): 110–116.

8. Durbin RM, Abecasis GR, Altshuler DL, Auton A, Brooks LD, Gibbs RA, Hurles ME, and McVean GA. 2010. A map of human genome variation from population-scale sequencing. *Nature* 467 (7319): 1061–1073.

9. Ebbesson SOE. 2011. The legacy of the Alaska Siberia Medical Research Program: a historical perspective. *International journal of circumpolar health* 70 (5): 584–93.

10. Fazzari MJ and Greally JM. 2010. Introduction to epigenomics and epigenome-wide analysis. *Methods Mol Biol* 620: 243–265.

11. Feinberg AP. 2010. Genome-scale approaches to the epigenetics of common human disease. *Virchows Arch* 456 (1): 13–21.

12. Fisher R. 1935. The Design of Experiments. In *The Design of Experiments.* London: Oliver and Boyd.

13. Foppa I and D. S. 1997. Power and Sample Size Calculations For Case-Control Studies of Gene-Environment Interactions With a Polytomous Exposure Variable. *American Journal of Epidemiology* 146 (7): 596–604.

14. Gilbert SF, Sapp J, and Tauber AI. 2012. A Symbiotic View of Life: We Have Never Been Individuals. *The Quarterly Review of Biology* 87 (4): 325–341.

15. De Graaf AA, Freidig AP, De Roos B, Jamshidi N, Heinemann M, Rullmann JAC, Hall KD, Adiels M, and Van Ommen B. 2009. Nutritional systems biology modeling: from molecular mechanisms to physiology. *PLoS computational biology* 5 (11): e1000554.

16. Hartman IV JL. 2006. Genetic and Molecular Buffering of Phenotypes. In *Nutritional Genomics. Discovering the Path to Personalized Nutrition*, ed. J Kaput Rodriguez, R.L., pp. 105–134. Hoboken, NJ: John Wiley and Sons.

17. Hartman JL, Garvik B, and Hartwell L. 2001. Principles for the buffering of genetic variation. *Science* 291 (5506): 1001–1004.

18. Hartman JL, Iv JLH, Garvik B, and Hartwell L. 2009. Principles for the Buffering of Genetic Variation. 1001 (2001):

19. Van Hemert JL and Dickerson J a. 2012. Discriminating response groups in metabolic and regulatory

pathway networks. *Bioinformatics (Oxford, England)* 28 (7): 947–54.

20. Kaput J. 2008. Nutrigenomics research for personalized nutrition and medicine. *Curr Opin Biotechnol* 19 (2): 110–120.

21. Kaput J and Rodriguez RL. 2004. Nutritional genomics: the next frontier in the postgenomic era. *Physiol Genomics* 16 (2): 166–177.

22. Koren O, Knights D, Gonzalez A, Waldron L, Segata N, Knight R, Huttenhower C, and Ley RE. 2013. A Guide to Enterotypes across the Human Body: Meta-Analysis of Microbial Community Structures in Human Microbiome Datasets. *PLoS Computational Biology* 9 (1):, ed. JA Eisene1002863.

23. Kraft P and Cox DG. 2008. Study designs for genome-wide association studies. *Adv Genet* 60: 465–504.

24. De La Vega FM, Gordon D, Su X, Scafe C, Isaac H, Gilbert DA, and Spier EG. 2005. Power and sample size calculations for genetic case/control studies using gene-centric SNP maps: application to human chromosomes 6, 21, and 22 in three populations. *Hum Hered* 60 (1): 43–60.

25. Lee ML and Whitmore GA. 2002. Power and sample size for DNA microarray studies. *Stat Med* 21 (23): 3543–3570.

26. Levine B. 2007. What does the population attributable fraction mean? *Prev Chronic Dis* 4 (1): A14.

27. Mardis ER. 2011. A decade's perspective on DNA sequencing technology. *Nature* 470 (7333): 198–203. Nature Publishing Group.

28. McCabe-Sellers B, Lovera D, Nuss H, Wise C, Ning B, Teitel C, Clark BS, et al. 2008. Personalizing nutrigenomics research through community based participatory research and omics technologies. *OMICS* 12 (4): 263–272.

29. Morine MJ and Priami C. 2013. *Analysis of biological systems.* Imperial College Press.

30. Morine MJ, Tierney AC, Van Ommen B, Daniel H, Toomey S, Gjelstad IMF, Gormley IC, et al. 2011. Transcriptomic coordination in the human metabolic network reveals links between n-3 fat intake, adipose tissue gene expression and metabolic health. *PLoS computational biology* 7 (11): e1002223.

31. Morine MJ, Toomey S, McGillicuddy FC, Reynolds CM, Power KA, Browne JA, Loscher C, Mills KHG, and Roche HM. 2012. Network analysis of adipose tissue gene expression highlights altered metabolic and regulatory transcriptomic activity in high-fat-diet-fed IL-1RI knockout mice. *The Journal of nutritional biochemistry:*

32. Muller M and Kersten S. 2003. Opinion: Nutrigenomics: goals and strategies. *Nat Rev Genet* 4 (4): 315–322.

33. Nakagawa S. 2004. A farewell to Bonferroni: the problems of low statistical power and publication bias. *Behavioral Ecology* 15 (6): 1044–1045.

34. Ng PC and Kirkness EF. 2010. Whole genome sequencing. *Methods Mol Biol* 628: 215–226.

35. Nikles J, Mitchell GK, Schluter P, Good P, Hardy J, Rowett D, Shelby-James T, Vohra S, and Currow D. 2011. Aggregating single patient (n-of-1) trials in populations where recruitment and retention was difficult: the case of palliative care. *Journal of clinical epidemiology* 64 (5): 471–80. Elsevier Inc.

36. Ordovas JM and Shen J. 2008. Gene-environment interactions and susceptibility to metabolic syndrome and other chronic diseases. *J Periodontol* 79(8 Suppl): 1508–1513.

37. Ottman R. 1996. Gene-environment interaction: definitions and study designs. *Prev Med* 25 (6): 764–770.

38. O'Sullivan A, Gibney MJ, Connor AO, Mion B, Kaluskar S, Cashman KD, Flynn A, Shanahan F, and Brennan L. 2011. Biochemical and metabolomic phenotyping in the identification of a vitamin D responsive metabotype for markers of the metabolic syndrome. *Molecular nutrition & food research* 55 (5): 679–90.

39. Plumb Collins, N, Cordeiro, JN, Kavanaugh-Lynch, M M. 2008. Assessing Process and Outcomes,: Evaluating Community Based Participatory Research. *Progress in Community Health Partnerships: Research, Education, and Action* 2 (2): 87–97.

40. Rockhill B, Newman B, and Weinberg C. 1998. Use and misuse of population attributable fractions. *Am J Public Health* 88 (1): 15–19.

41. Rosenberg E and Zilber-Rosenberg I. 2011. Symbiosis and development: the hologenome concept. *Birth defects research. Part C, Embryo today: reviews* 93 (1): 56–66.

42. Stumbo PJ, Weiss R, Newman JW, Pennington JA, Tucker KL, Wiesenfeld PL, Illner AK, Klurfeld DM, and Kaput J. 2010. Web-enabled and improved software tools and data are needed to measure nutrient intakes and physical activity for personalized health research. *J Nutr* 140 (12): 2104–2115.

43. Sudmant PH, Kitzman JO, Antonacci F, Alkan C, Malig M, Tsalenko A, Sampas N, Bruhn L, Shendure J, and Eichler EE. 2010. Diversity of human copy number variation and multicopy genes. *Science* 330 (6004): 641–646.

44. Vaupel JW. 2010. Biodemography of human ageing. *Nature* 464 (7288): 536–42.

45. Wells MI. 2000. Beyond cultural competence: A model for individual and institutional cultural development. *Journal of Community Health Nursing* 17 (4): 189–199.

46. Williams RP. 1956. *Biochemical Individuality: The Basis for the Genetotrophic Concept.* New Canaan, CT: Keats Publishing.

47. Zivkovic AM and German JB. 2009. Metabolomics for assessment of nutritional status. *Curr Opin Clin Nutr Metab Care* 12 (5): 501–507.

Jim Kaput

Clinical Translation Unit
Nestlé Institute of Health Sciences
Quartier de l'Innovation
EPFL Campus
1015 Lausanne
Switzerland
James.Kaput@rd.nestle.com