Genes Nutr (2010) 5:275–283
DOI 10.1007/s12263-010-0186-6

COMMENTARY

Connecting the Human Variome Project to nutrigenomics

Jim Kaput · Chris T. Evelo · Giuditta Perozzi · Ben van Ommen · Richard Cotton

Received: 17 September 2010 / Accepted: 24 September 2010 / Published online: 15 October 2010

Abstract Nutrigenomics is the science of analyzing and understanding gene–nutrient interactions, which because of the genetic heterogeneity, varying degrees of interaction among gene products, and the environmental diversity is a complex science. Although much knowledge of human diversity has been accumulated, estimates suggest that ~90% of genetic variation has not yet been characterized. Identification of the DNA sequence variants that contribute to nutrition-related disease risk is essential for developing a better understanding of the complex causes of disease in humans, including nutrition-related disease. The Human Variome Project (HVP; http://www.humanvariomeproject.org/) is an international effort to systematically identify genes, their mutations, and their variants associated with phenotypic variability and indications of human disease or phenotype. Since nutrigenomic research uses genetic information in the design and analysis of experiments, the HVP is an essential collaborator for ongoing studies of gene–nutrient interactions. With the advent of next generation sequencing methodologies and the understanding of the undiscovered variation in human genomes, the nutrigenomic community will be generating novel sequence data and results. The guidelines and practices of the HVP can guide and harmonize these efforts.

Keywords Nutrigenomics · Human Variome Project · Harmonization

Note: The views in this article do not necessarily represent those of the U.S. FDA.

J. Kaput (✉)
Division of Personalised Nutrition and Medicine,
FDA/National Center for Toxicological Research,
3900 NCTR Road, Jefferson, AR 72079, USA
e-mail: James.kaput@fda.hhs.gov

C. T. Evelo
Department of Bioinformatics (BiGCaT),
Maastricht University, Maastricht, The Netherlands

G. Perozzi
INRAN—National Research Institute on Food & Nutrition,
Via Ardeatina 546, 00178 Rome, Italy

B. van Ommen
TNO Quality of Life, P.O. box 360,
3700 AJ Zeist, The Netherlands

R. Cotton
Genomic Disorders Research Centre, Level 2, Alan Gilbert
Building, 161 Barry Street, Melbourne, VIC 3053, Australia

Introduction

Nutrigenomics has been called a "post-genome" field of research (e.g., [35]) because it could only develop in a meaningful way after the completion of the sequencing of the human genome. Ideally, nutrigenomics experiments are to be designed, conducted, and analyzed with specific knowledge of genes involved in nutrient metabolism and physiological processes. Data from the Human Genome project and subsequent haplotype mapping projects [14, 59–61, 70] demonstrated that any two humans will differ by 3–5 million bases. However, ongoing re-sequencing of individual genomes has now identified ~18 million single nucleotide polymorphisms [4, 29, 47]. More recent studies have also shown structural variation in the human genome indicating that approximately 13% of the variation between humans may be due to copy number [1, 31, 41, 50, 68, 71]. The ongoing discovery of new SNPs and structural variants shows that the genomic era is far from ending, particularly since the next generation sequencing technologies [42] will

soon allow for complete analyses of all coding sequences (i.e., the exome) of all individuals in a research study (e.g., [9]) or even for the complete sequencing of individual genes, including the regulatory regions.

The catalog of genetic variation is being done by the HapMap and 1,000 Genomes [37, 55] projects. While these projects are providing fundamental information about human genetic variation, they are by design limited by (i) the populations analyzed, (ii) the goal of characterizing variants that are greater than 1% in the population, and (iii) they are unlinked to phenotype. The HapMap project [14, 61] analyzed 45 Han Chinese in China, 45 Japanese in Tokyo, 90 Africans in the Yoruba tribe of Nigeria, and 90 Europeans of Northern European origin. The recently released HapMap3 data [2] extended the analyses to almost 700 individuals but found that 77% of the SNPs analyzed in their study had not been previously characterized. The goal of the 1,000 Genomes project is to analyze ~2,000 individuals from 23 populations (Table 1). Due to these sample sizes, the diversity projects will only be able to detect variants that are greater than ~1% in the populations tested; some rare variants [39] are found during targeted re-sequencing that occurs in these projects yet variation in the many populations not analyzed will not be found. Rare variants are increasingly suspected of being involved in healthy and disease phenotypes [12]. Finally, the goal of these projects is simply to catalog the variants for future use in research linking genotype to phenotype, although some phenotype data are available to research in the consortium. Both HVP and NuGO (Nutrigenomics Organization) aim to document variation causing phenotypic changes such as disease and good health.

The Human Variome Project

The Human Variome Project (HVP; http://www.humanvariomeproject.org/) is an international effort to systematically identify genes, their mutations, and their variants associated with phenotypic variability and indications of human disease or phenotype [15, 16, 33, 51]. The goal of the HVP is to link clinical, medical, and research laboratories for developing knowledge. This knowledge will be accessible to the research and medical communities to improve research strategies and clinical medical practice. One of the key deliverables of the HVP is the creation of a cyber infrastructure to link locus-specific databases (LSDBs [27]). These databases have similar architecture, ontologies, and data elements allowing for interoperability and are curated by local experts. Over 700 LSDBs are

Table 1 1,000 Genomes populations[a]

Code	Population	Description
CHB	Han Chinese	Han Chinese in Beijing, China
CHS	Southern Han Chinese	Han Chinese south
CDX	Dai Chinese	Chinese Dai in Xishuangbanna, China
CHD	Denver Chinese	Chinese in Denver, Colorado (pilot 3 only)
JPT	Japanese	Japanese in Tokyo, Japan
KHV	Kinh Vietnamese	Kinh in Ho Chi Minh City, Vietnam
CEU	CEPH	Utah residents (CEPH) with Northern and Western European ancestry
TSI	Tuscan	Toscani in Italia
GBR	British	British in England and Scotland
FIN	Finnish	Finnish in Finland
IBS	Spanish	Iberian populations in Spain
YRI	Yoruba	Yoruba in Ibadan, Nigeria
LWK	Luhya	Luhya in Webuye, Kenya
GWD	Gambian	Gambian in Western Division, The Gambia
GHN	Ghanaian	Ghanaian in Navrongo, Ghana
MAB	Malawian	Malawian in Blantyre, Malawi
ASW	African–American SW	African Ancestry in Southwest US
AJM	African–American MS	African American in Jackson, Mississippi
ACB	African–Caribbean	African Caribbean in Barbados
MXL	Mexican–American	Mexican Ancestry in Los Angeles, California
CLM	Colombian	Colombian in Medellin, Colombia
PEL	Peruvian	Peruvian in Lima, Peru
PUR	Puerto Rican	Puerto Rican in Puerto Rico

[a] From ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/README.populations

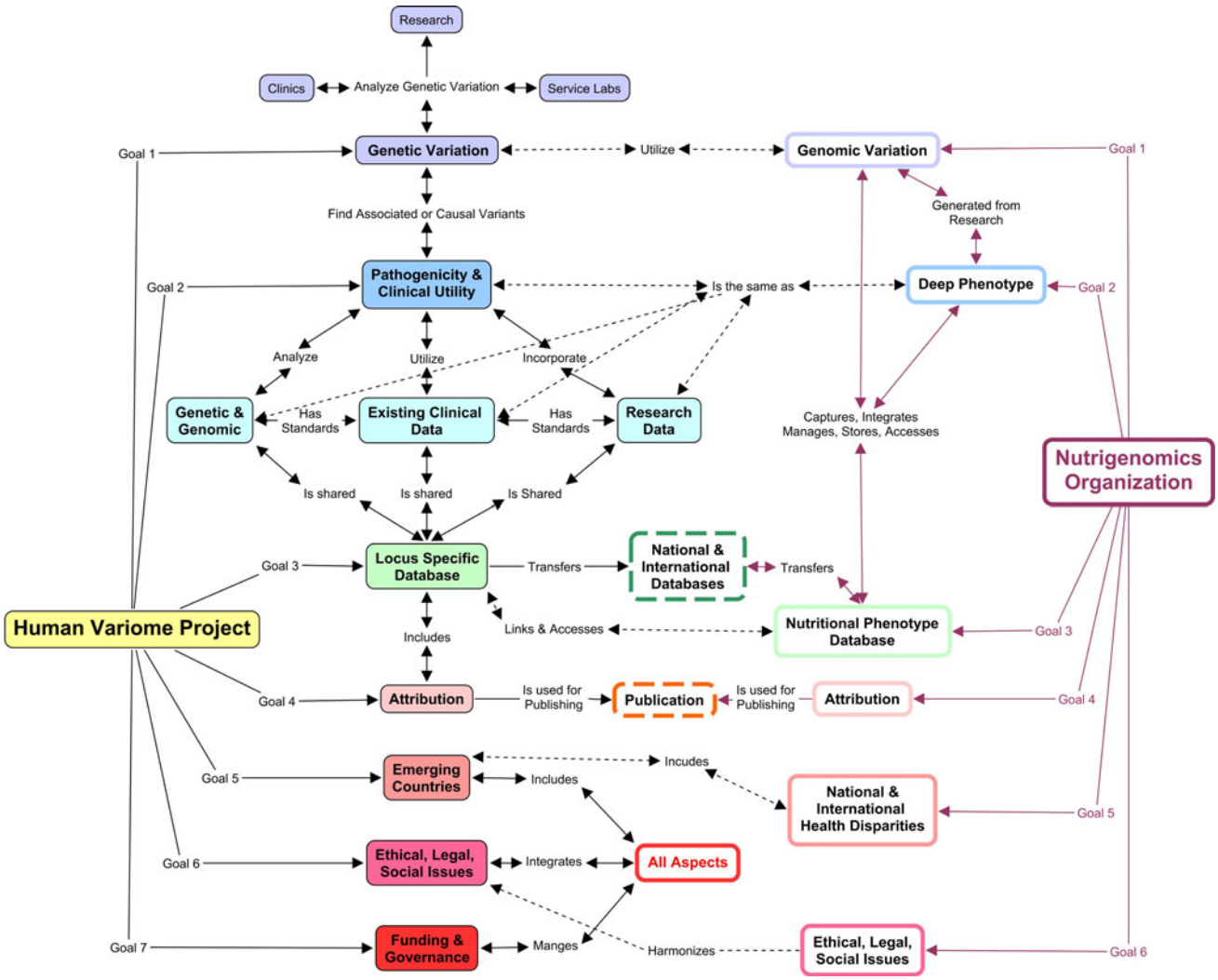

Fig. 1 The key goals of the Human Variome Project and Nutrigenomic Organization (NuGO). Data from nutrigenomics' experiments are from research, population, and clinical studies. See text for details

maintained across the world and accessible at the Human Genome Variation Society website (http://www.hgvs.org/dblist/glsdb.html). Some, but not all, of the information in LSDBs are consolidated in national and international databases such as at the National Center for Biotechnology Information (NCBI—http://www.ncbi.nlm.nih.gov/) and the European Bioinformatics Institute (EBI—http://www.ebi.ac.uk/).

The HVP has 10 key objectives (Fig. 1):

1. Capture and archive all human gene variation associated with human disease in a central location with mirror sites in other countries. Data governance will ensure security and integrity through the use of auditing and security technologies but nevertheless allow searching across all genes using a common interface.
2. Provide a standardized system of gene variation nomenclature, reference sequences, and support systems that will enable diagnostic laboratories to use and contribute to total human variation knowledge.
3. Establish systems that ensure adequate curation of human variation knowledge from gene-specific (locus-specific), country-specific, or disease-specific database perspective to improve accuracy, reduce errors, and develop a comprehensive data set comprising all human genes.
4. Facilitate the development of software to collect and exchange human variation data in a federation of gene-specific (locus-specific), country-specific, disease-specific, and general databases.
5. Establish a structured and tiered mechanism that clinicians can use to determine the health outcomes associated with genetic variation. This will work as a dialog between those who use human variation data and those who provide them. Clinicians will be

encouraged to provide data and will have open access to complete variation data.
6. Create a support system for research laboratories that provides for the collection of genotypic and phenotypic data together using the defined reference sequence in a free, unrestricted, and open access system and create a simple mechanism for logging discoveries.
7. Develop ethical standards to ensure open access to all human variation data that are to be used for global public good and address the needs of "indigenous" communities under the threat of dilution in emerging countries.
8. Provide support to developing countries to build capacity and to fully participate in the collection, analysis, and sharing of genetic variation information.
9. Establish a communication and education program to collect and spread knowledge related to human variation knowledge to all countries of the world.
10. Continue to carry out research within the opportunities presented by the investigation of human genetic variation and to present these findings to users of this information for the benefit of all.

The Nutrigenomics Organization (NuGO)

NuGO was established as an association of 23 universities and research institutes focusing on jointly developing the exiting research area of nutrigenomics and nutritional systems biology. NuGO evolved from an EU Sixth Framework Network of Excellence and has now transitioned into a global association encompassing individuals and institutions around the globe.

NuGO has two major objectives:

1. stimulating developments in nutrigenomics, nutrigenetics, and nutritional systems biology and incorporating these aspects in nutrition and health research, by joint research projects, conferences, workshops, and training. As a legal, nonprofit entity, NuGO can join as partner in research projects anywhere in the world.
2. shaping the nutrition bioinformatics infrastructure, by initiating, coordinating, facilitating projects in this area and by hosting and disseminating all data, results and information in this area.

The common goals of nutrigenomic community and the Human Variome Project

The HVP established committees to develop action plans to meet the 10 key objectives, which overlap or parallel the goals described in a consensus statement authored by 89 international scientists in the nutrigenomics community [34], which was co-authored by the members of NuGO. The parallel objectives of the HVP and NuGO are the common grounds upon which the two scientific communities base their interaction, as schematically described in Fig. 1.

The HVP captures and classifies genetic variation from voluntary contributions from unlinked clinical, research, diagnostic, and service laboratories. The nutrigenomic community had not specifically designated the identification and characterization of variation as a goal, even though many researchers interested in nutrient–phenotype associations use that knowledge in their experiments (e.g., [45, 56]). The emerging consensus that rare polymorphisms (i.e., <1% in the population) and copy number variants [46] may influence health, disease, and nutrient–gene interactions makes it imperative that nutrigenomic researchers adopt and use sequence technologies and methods as a part of their experimental design and procedures.

Assessment of pathogenicity—the phenotype

Deleterious gene–phenotype associations are described as pathogenicity in clinical settings and are described by genetic, clinical phenotype, and pathology. Basic researchers use the term deep phenotyping [63] when analyzing large numbers of genes, metabolites, proteins, or transcripts, or combinations of omic technologies. These methodologies are being adopted to the clinic for assessments (e.g., [24, 57]) and have led to the application of interdisciplinary, primary care, community-based, and translational research [25, 40, 69] to health and disease studies, blurring the distinction between basic and clinical research [9].

Databases

Data transfer, integration, and access are among the major challenges facing the biomedical researcher in the 21st century. The abstracts for over 19 million are available in PubMed in the U.S. National Library of Medicine's database of publications. Inclusion of quantitative measures of genetic and environmental variations would likely influence the results of many of these studies. However, only a small fraction of full texts and additional data for these publications is electronically accessible. Moreover, much of these data are discipline specific, making it a challenge to mine public-domain results for linked data or knowledge. The HVP is developing standards with international databases (NCBI and EBI) and projects (e.g., European Union Gen2Phen initiative—http://www.gen2phen.org). The nutrigenomic community faces a greater challenge

since few tools or databases are available for nutrition-related research [58]. Two related initiatives are underway to address these limitations. The Nutrigenomics Organization has begun the development of a nutritional phenotype database (dbNP—[64]), a research and collaboration tool, and knowledge base which will allow access to publically available data. A separate yet key component of dbNP is WikiPathways [36, 48], an open, collaborative platform for the curation of biological pathways (http://www.wikipathways.org) based on all data available to the curator.

Emerging countries

About 90% of known SNPs are shared between Asians, Europeans, and Africans and the remaining polymorphisms, called private SNPs, are distributed among these populations [26, 30]. The recent sequencing of the genomes of several individuals (e.g., [5, 38, 66, 67]), along with gene-specific re-sequencing efforts suggest that a larger number of SNPs than previously determined, as well as other sequence variation, exist in the human population. Estimates from African genetic diversity and the Pan Asian SNP initiative indicate that 80 to 90% of human genomic variation resides in the world's emerging countries. The Population Reference Sample (POPRES [44]) targets populations not previously included in the HapMap project, similar to the 1,000 Genomes initiative. The main focus of the HVP effort is the inclusion and analyses of clinical samples from diverse ethnic groups. One of the advantages of including some ethnic populations is the opportunity to study genetic diseases due to consanguinity, large family size, and potential founder effects (e.g., [10, 11, 52]). Although nutrigenomic researchers do not necessarily include such populations, unique genetic groups may yield valuable insights into understanding the distribution of gene–nutrient interactions in health and disease processes [32].

Biomedical research has not typically been the focus of resource for poor countries, even though such activities are likely to produce economic and health benefits for all [20, 54]. The recently announced global alliance for chronic disease (GACD—[19]) addresses the increasing consensus that emerging economies face not only malnutrition but also the development of chronic diseases, a double economic and health burden. Education of health care providers, the public, and government officials is needed for demonstrating the universal nature of the HVP's and NuGO's research efforts, the need to include populations in developing countries, and the benefits from cooperating in biomedical research ([8, 13, 53, 62].

The NIH National Center for Minority Health and Health Disparities (NCMHD) recently called for a paradigm shift to include minority, low socioeconomic and rural populations, and individuals in biomedical research [21, 22]. From the perspective of the science underlying both the HVP and nutrigenomic research efforts, understanding the full spectrum of genetic or metabolic spectra will not be possible without the involvement of all ancestral groups. Including these populations and individuals may allow for a more rapid translation of basic science to society. Community-based participatory research collaborations may provide forums for addressing cultural and ethical concerns of biomedical research [40]. The genomic sovereignty/equality for all countries to be involved in their research efforts is an accepted norm of the HVP and nutrigenomic research communities. The value of 'human capital' within all populations is acknowledged and treasured. Real and tangible benefits of biomedical research to improve health will be generated for participating populations; the voluntary participation of the greatest number of countries would ensure targeted interventions to improve personal and public health.

Ethics

HVP and nutrigenomic researchers are committed to adhering to the highest ethical principles governing research, data sharing, and ultimately enabling this new knowledge to benefit all of the humanity. Ethical guidelines specifically for LSDBs were previously published [17], and new guidelines are in process [49]. NuGO has also published bioethical guidelines [6, 7]. The ethics of international health research continue to evolve with a greater emphasis on development and social justice [28], values which are consistent with the stated goals of the HVP and nutrigenomics researchers [34, 51].

A key ethical concern for the Human Variome Project and eventually for data generated from individuals in nutrigenomics research is the accessibility of research data on public websites. For HVP, rare mutations in a population open the possibility of identifying the research participant. For both fields of research, polygenic analyses (e.g., whole genome scans) have yielded the identification of single participants in research studies. Such polygenic analyses generate data that could be used for re-identifying individual patients [18] but is less likely in single gene diseases.

The HVP is developing an ethics review committee with a subcommittee focused on issues related to LSDB for (i) providing counsel when dilemmas arise, (ii) overseeing guidelines, (iii) identifying best practices, (iv) determining how best to ensure privacy in all situations, (v) formulating how to handle data for which explicit consent does not exist or is not possible to achieve, and (vi) developing a consent form that is consistent for all LSDBs but which can

be adapted to the requirements of individual countries. NuGO has an ethics committee, and several members are also members of the HVP ethics group.

Funding and governance

International collaborations of the scope of the HVP and nutrigenomics efforts are rarely funded. Rather, both organizations believe that distributed financing for distributed science is more likely to occur. To make a distributive model feasible, best practices and harmonized protocols are necessary. The HVP is developing the standards for LSDB's and reporting, and NuGO has a rich tradition of developing standard operating procedures for methodologies (http://www.nugo.org/sops). Several international efforts are currently being designed for type 2 diabetes and micronutrient genomics [65]. Guidelines and policies of the HVP will be adopted for both of these protocols since both require analyses of phenotype based on genetic differences in study participants. The HVP is managed by the Genomic Disorders Research Center (Melbourne, Australia; http://www.genomic.unimelb.edu.au), and NuGO has a coordinating council consisting of European scientists from member Institutions.

Attribution and publication

Large scale science requires novel approaches for attributing contributions. High throughput technologies generate a large number of raw data whose availability represents a great advantage to all scientists working in the field. These data sets are often unpublished and made available to all through deposition in public databases. This creates the need to define new criteria for the attribution of unpublished data and requires a policy of credits and incentives to be coordinated between database curators and the editors of scientific journals. For the nutrigenomics projects, harmonized, distributive funding allows for the publication of local studies, and the possibility of comprehensive analyses of all data by a team consisting of individuals from multiple disciplines and regions. The HapMap project and other international projects have previously developed procedures for publication credit. The HVP faces additional challenges since clinics and service laboratories describe pathogenicity and genetic variation, but academic incentives to publish are not paramount for these professions. Novel approaches for attribution have been proposed [23, 33, 43] and are starting to be applied by some databases and continue to be developed by the HVP. A public debate is also in progress among the editors of genetics journals to coordinate efforts toward the attribution of data deposition and database linking (http://www.gen2phen.org/wiki/hvp-publication-credit-and-incentives-recommendations).

Connecting the projects

Members of the HVP and nutrigenomics community have attended conferences and workshops, including the 3rd Asia Pacific Nutrigenomics conference (Melbourne, 2008), the Fifth China Health Annual for the announcement of the HVP node in Beijing (2008), the Symposium of Princess Al-Jawhara Center of Excellence in Research of Hereditary Disorders (King Abdulaziz University. Jeddah, Saudi Arabia), the Third HVP planning meeting (Costa Brava, Spain) and at the Fourth HVP meeting (Roadmap: Project structure) at UNESCO (Paris, France). NuGO organized a micronutrient genomics workshop at the UNESCO meeting. Future meetings include co-hosting workshops at the International Conference on Nutrigenomics (ICON) in Guaruja', Brazil (September 2010). The HVP and NuGO are exploring the best means to formally recognize their ongoing interactions. The HVP has a similar ongoing relationship with the International Society for Gatrointestinal Hereditary Tumors (http://www.insight-group.org).

Organisms are complex systems that are regulated by interactions between multiple environmental factors and multiple genes. Analyzing complex systems requires a multidisciplinary and multi-technological approach (i.e., omics). Since each person or team has a limited level of complexity, the ideal, and indeed, the only, approach to understand biological systems is to distribute the task among many teams (see [3]). The interaction between the HVP and nutrigenomics researchers leverages the expertise of both groups and fosters a more complete analysis of the human organism. These articles will assist the process of populating the human genome sequence with variations which are biologically and medically relevant and eventually provide the complete "human variome."

References

1. Al-Sukhni W, Gallinger S (2008) Germline copy number variation in control populations. Cytogenet Genome Res 123(1–4): 211–223
2. Altshuler DM, Gibbs RA, Peltonen L, Dermitzakis E, Schaffner SF, Yu F, Bonnen PE, de Bakker PI, Deloukas P, Gabriel SB, Gwilliam R, Hunt S, Inouye M, Jia X, Palotie A, Parkin M, Whittaker P, Chang K, Hawes A, Lewis LR, Ren Y, Wheeler D, Muzny DM, Barnes C, Darvishi K, Hurles M, Korn JM, Kristiansson K, Lee C, McCarrol SA, Nemesh J, Keinan A, Montgomery SB, Pollack S, Price AL, Soranzo N, Gonzaga-Jauregui C, Anttila V, Brodeur W, Daly MJ, Leslie S, McVean G, Moutsianas L, Nguyen H, Zhang Q, Ghori MJ, McGinnis R, McLaren W, Takeuchi F, Grossman SR, Shlyakhter I, Hostetter EB, Sabeti PC, Adebamowo CA, Foster MW, Gordon DR, Licinio J, Manca MC, Marshall PA, Matsuda I, Ngare D, Wang VO, Reddy D, Rotimi CN, Royal CD, Sharp RR, Zeng C, Brooks LD, McEwen JE (2010) Integrating common and rare genetic variation in diverse human populations. Nature 467(7311):52–58

3. Bar-Yam Y (2004) Making things work: solving complex problems in a complex world. NECSI- Knolwedge Press, Cambridge, p 306
4. Baye TM, Tiwari HK, Allison DB, Go RC (2009) Database mining for selection of SNP markers useful in admixture mapping. BioData Min 2(1):1
5. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR, Boutell JM, Bryant J, Carter RJ, Keira Cheetham R, Cox AJ, Ellis DJ, Flatbush MR, Gormley NA, Humphray SJ, Irving LJ, Karbelashvili MS, Kirk SM, Li H, Liu X, Maisinger KS, Murray LJ, Obradovic B, Ost T, Parkinson ML, Pratt MR, Rasolonjatovo IM, Reed MT, Rigatti R, Rodighiero C, Ross MT, Sabot A, Sankar SV, Scally A, Schroth GP, Smith ME, Smith VP, Spiridou A, Torrance PE, Tzonev SS, Vermaas EH, Walter K, Wu X, Zhang L, Alam MD, Anastasi C, Aniebo IC, Bailey DM, Bancarz IR, Banerjee S, Barbour SG, Baybayan PA, Benoit VA, Benson KF, Bevis C, Black PJ, Boodhun A, Brennan JS, Bridgham JA, Brown RC, Brown AA, Buermann DH, Bundu AA, Burrows JC, Carter NP, Castillo N, Chiara ECM, Chang S, Neil Cooley R, Crake NR, Dada OO, Diakoumakos KD, Dominguez-Fernandez B, Earnshaw DJ, Egbujor UC, Elmore DW, Etchin SS, Ewan MR, Fedurco M, Fraser LJ, Fuentes Fajardo KV, Scott Furey W, George D, Gietzen KJ, Goddard CP, Golda GS, Granieri PA, Green DE, Gustafson DL, Hansen NF, Harnish K, Haudenschild CD, Heyer NI, Hims MM, Ho JT, Horgan AM, Hoschler K, Hurwitz S, Ivanov DV, Johnson MQ, James T, Huw Jones TA, Kang GD, Kerelska TH, Kersey AD, Khrebtukova I, Kindwall AP, Kingsbury Z, Kokko-Gonzales PI, Kumar A, Laurent MA, Lawley CT, Lee SE, Lee X, Liao AK, Loch JA, Lok M, Luo S, Mammen RM, Martin JW, McCauley PG, McNitt P, Mehta P, Moon KW, Mullens JW, Newington T, Ning Z, Ling Ng B, Novo SM, O'Neill MJ, Osborne MA, Osnowski A, Ostadan O, Paraschos LL, Pickering L, Pike AC, Chris Pinkard D, Pliskin DP, Podhasky J, Quijano VJ, Raczy C, Rae VH, Rawlings SR, Chiva Rodriguez A, Roe PM, Rogers J, Rogert Bacigalupo MC, Romanov N, Romieu A, Roth RK, Rourke NJ, Ruediger ST, Rusman E, Sanches-Kuiper RM, Schenker MR, Seoane JM, Shaw RJ, Shiver MK, Short SW, Sizto NL, Sluis JP, Smith MA, Ernest Sohna Sohna J, Spence EJ, Stevens K, Sutton N, Szajkowski L, Tregidgo CL, Turcatti G, Vandevondele S, Verhovsky Y, Virk SM, Wakelin S, Walcott GC, Wang J, Worsley GJ, Yan J, Yau L, Zuerlein M, Mullikin JC, Hurles ME, McCooke NJ, West JS, Oaks FL, Lundberg PL, Klenerman D, Durbin R, Smith AJ (2008) Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456(7218):53–59
6. Bergmann MM, Bodzioch M, Bonet ML, Defoort C, Lietz G, Mathers JC (2006) Bioethics in human nutrigenomics research: European nutrigenomics organisation workshop report. Br J Nutr 95(5):1024–1027
7. Bergmann MM, Gorman U, Mathers JC (2008) Bioethical considerations for human nutrigenomics. Annu Rev Nutr 28:447–467
8. Bhan A, Singh JA, Upshur RE, Singer PA, Daar AS (2007) Grand challenges in global health: engaging civil society organizations in biomedical research in developing countries. PLoS Med 4(9):e272
9. Biesecker LG, Mullikin JC, Facio FM, Turner C, Cherukuri PF, Blakesley RW, Bouffard GG, Chines PS, Cruz P, Hansen NF, Teer JK, Maskeri B, Young AC, Manolio TA, Wilson AF, Finkel T, Hwang P, Arai A, Remaley AT, Sachdev V, Shamburek R, Cannon RO, Green ED (2009) The clinseq project: piloting large-scale genome sequencing for research in genomic medicine. Genome Res 19(9):1665–1674
10. Bittles A (2001) Consanguinity and its relevance to clinical genetics. Clin Genet 60(2):89–98
11. Bittles AH (2002) Endogamy, consanguinity and community genetics. J Genet 81(3):91–98
12. Cirulli ET, Goldstein DB (2010) Uncovering the roles of rare variants in common disease through whole-genome sequencing. Nat Rev Genet 11(6):415–425
13. Cohen ER, Masum H, Berndtson K, Saunders V, Hadfield T, Panjwani D, Persad DL, Minhas GS, Daar AS, Singh JA, Singer PA (2008) Public engagement on global health challenges. BMC Public Health 8:168
14. Consortium IHGS (2003) The international HapMap project. Nature 426(6968):789–796
15. Cotton RG (2009) Collection of variation causing disease—the Human Variome Project. Hum Genomics 3(4):301–303
16. Cotton RG, Phillips K, Horaitis O (2007) A survey of locus-specific database curation. Human genome variation society. J Med Genet 44(4):e72
17. Cotton RG, Sallee C, Knoppers BM (2005) Locus-specific databases: from ethical principles to practice. Hum Mutat 26(5):489–493
18. Craig DW, Pearson JV, Szelinger S, Sekar A, Redman M, Corneveaux JJ, Pawlowski TL, Laub T, Nunn G, Stephan DA, Homer N, Huentelman MJ (2008) Identification of genetic variants using bar-coded multiplexed sequencing. Nat Methods 5:887–893
19. Daar AS, Nabel EG, Pramming SK, Anderson W, Beaudet A, Liu D, Katoch VM, Borysiewicz LK, Glass RI, Bell J (2009) The global alliance for chronic diseases. Science 324(5935):1642
20. Daar AS, Thorsteinsdottir H, Martin DK, Smith AC, Nast S, Singer PA (2002) Top ten biotechnologies for improving health in developing countries. Nat Genet 32(2):229–232
21. Dankwa-Mullan I, Rhee KB, Stoff DM, Pohlhaus JR, Sy FS, Stinson N Jr, Ruffin J (2010) Moving toward paradigm-shifting research in health disparities through translational, transformational, and transdisciplinary approaches. Am J Public Health 100 Suppl 1:S19–S24
22. Dankwa-Mullan I, Rhee KB, Williams K, Sanchez I, Sy FS, Stinson N Jr, Ruffin J (2010) The science of eliminating health disparities: summary and analysis of the NIH summit recommendations. Am J Public Health 100 Suppl 1:S12–S18
23. Editorial (2008) Human variome microattribution reviews. Nat Genet 40 (1)
24. Gowda GA, Zhang S, Gu H, Asiago V, Shanaiah N, Raftery D (2008) Metabolomics-based methods for early disease diagnostics. Expert Rev Mol Diagn 8(5):617–633
25. Hebert JR, Brandt HM, Armstead CA, Adams SA, Steck SE (2009) Interdisciplinary, translational, and community-based participatory research: finding a common language to improve cancer research. Cancer Epidemiol Biomarkers Prev 18(4):1213–1217
26. Hinds DA, Stuve LL, Nilsen GB, Halperin E, Eskin E, Ballinger DG, Frazer KA, Cox DR (2005) Whole-genome patterns of common DNA variation in three human populations. Science 307(5712):1072–1079
27. Horaitis O, Talbot CC Jr, Phommarinh M, Phillips KM, Cotton RG (2007) A database of locus-specific databases. Nat Genet 39(4):425
28. Ijsselmuiden CB, Kass NE, Sewankambo KN, Lavery JV (2010) Evolving values in ethics and global health research. Glob Public Health 5(2):154–163
29. Johnson AD (2009) Single-nucleotide polymorphism bioinformatics: a comprehensive review of resources. Circ Cardiovasc Genet 2(5):530–536
30. Jorde LB, Wooding SP (2004) Genetic variation, classification and 'race'. Nat Genet 36 Suppl 1:S28–S33
31. Kang TW, Jeon YJ, Jang E, Kim HJ, Kim JH, Park JL, Lee S, Kim YS, Kim JY, Kim SY (2008) Copy number variations (CNVs) identified in Korean individuals. BMC Genomics 9:492
32. Kaput J (2008) Nutrigenomics research for personalized nutrition and medicine. Curr Opin Biotechnol 19(2):110–120

33. Kaput J, Cotton RG, Hardman L, Watson M, Al Aqeel AI, Al-Aama JY, Al-Mulla F, Alonso S, Aretz S, Auerbach AD, Bapat B, Bernstein IT, Bhak J, Bleoo SL, Blocker H, Brenner SE, Burn J, Bustamante M, Calzone R, Cambon-Thomsen A, Cargill M, Carrera P, Cavedon L, Cho YS, Chung YJ, Claustres M, Cutting G, Dalgleish R, den Dunnen JT, Diaz C, Dobrowolski S, Dos Santos MR, Ekong R, Flanagan SB, Flicek P, Furukawa Y, Genuardi M, Ghang H, Golubenko MV, Greenblatt MS, Hamosh A, Hancock JM, Hardison R, Harrison TM, Hoffmann R, Horaitis R, Howard HJ, Barash CI, Izagirre N, Jung J, Kojima T, Laradi S, Lee YS, Lee JY, Gil-da-Silva-Lopes VL, Macrae FA, Maglott D, Marafie MJ, Marsh SG, Matsubara Y, Messiaen LM, Moslein G, Netea MG, Norton ML, Oefner PJ, Oetting WS, O'Leary JC, de Ramirez AM, Paalman MH, Parboosingh J, Patrinos GP, Perozzi G, Phillips IR, Povey S, Prasad S, Qi M, Quin DJ, Ramesar RS, Richards CS, Savige J, Scheible DG, Scott RJ, Seminara D, Shephard EA, Sijmons RH, Smith TD, Sobrido MJ, Tanaka T, Tavtigian SV, Taylor GR, Teague J, Topel T, Ullman-Cullere M, Utsunomiya J, van Kranen HJ, Vihinen M, Webb E, Weber TK, Yeager M, Yeom YI, Yim SH, Yoo HS (2009) Planning the Human Variome Project: the Spain report. Hum Mutat 30(4):496–510
34. Kaput J, Ordovas JM, Ferguson L, van Ommen B, Rodriguez RL, Allen L, Ames BN, Dawson K, German B, Krauss R, Malyj W, Archer MC, Barnes S, Bartholomew A, Birk R, van Bladeren P, Bradford KJ, Brown KH, Caetano R, Castle D, Chadwick R, Clarke S, Clement K, Cooney CA, Corella D, Manica da Cruz IB, Daniel H, Duster T, Ebbesson SO, Elliott R, Fairweather-Tait S, Felton J, Fenech M, Finley JW, Fogg-Johnson N, Gill-Garrison R, Gibney MJ, Gillies PJ, Gustafsson JA, Hartman Iv JL, He L, Hwang JK, Jais JP, Jang Y, Joost H, Junien C, Kanter M, Kibbe WA, Koletzko B, Korf BR, Kornman K, Krempin DW, Langin D, Lauren DR, Ho Lee J, Leveille GA, Lin SJ, Mathers J, Mayne M, McNabb W, Milner JA, Morgan P, Muller M, Nikolsky Y, van der Ouderaa F, Park T, Pensel N, Perez-Jimenez F, Poutanen K, Roberts M, Saris WH, Schuster G, Shelling AN, Simopoulos AP, Southon S, Tai ES, Towne B, Trayhurn P, Uauy R, Visek WJ, Warden C, Weiss R, Wiencke J, Winkler J, Wolff GL, Zhao-Wilson X, Zucker JD (2005) The case for strategic international alliances to harness nutritional genomics for public and personal health. Br J Nutr 94(5):623–632
35. Kaput J, Rodriguez RL (2004) Nutritional genomics: the next frontier in the post genomic era. Physiol Genomics 16(2): 166–177
36. Kelder T, Pico AR, Hanspers K, van Iersel MP, Evelo C, Conklin BR (2009) Mining biological pathways using wikipathways web services. PLoS ONE 4(7):e6447
37. Kuehn BM (2008) 1,000 genomes project promises closer look at variation in human genome. JAMA 300(23):2715
38. Levy S, Sutton G, Ng PC, Feuk L, Halpern AL, Walenz BP, Axelrod N, Huang J, Kirkness EF, Denisov G, Lin Y, MacDonald JR, Pang AW, Shago M, Stockwell TB, Tsiamouri A, Bafna V, Bansal V, Kravitz SA, Busam DA, Beeson KY, McIntosh TC, Remington KA, Abril JF, Gill J, Borman J, Rogers YH, Frazier ME, Scherer SW, Strausberg RL, Venter JC (2007) The diploid genome sequence of an individual human. PLoS Biol 5(10):e254
39. Li B, Leal SM (2009) Discovery of rare variants via sequencing: implications for the design of complex trait association studies. PLoS Genet 5(5):e1000481
40. McCabe-Sellers B, Lovera D, Nuss H, Wise C, Ning B, Teitel C, Clark BS, Toennessen T, Green B, Bogle ML, Kaput J (2008) Personalizing nutrigenomics research through community based participatory research and omics technologies. OMICS 12(4):263–272
41. McElroy JP, Nelson MR, Caillier SJ, Oksenberg JR (2009) Copy number variation in African Americans. BMC Genet 10:15
42. Metzker ML (2010) Sequencing technologies—the next generation. Nat Rev Genet 11(1):31–46
43. Mons B (2005) Which gene did you mean? BMC Bioinformatics 6:142
44. Nelson MR, Bryc K, King KS, Indap A, Boyko AR, Novembre J, Briley LP, Maruyama Y, Waterworth DM, Waeber G, Vollenweider P, Oksenberg JR, Hauser SL, Stirnadel HA, Kooner JS, Chambers JC, Jones B, Mooser V, Bustamante CD, Roses AD, Burns DK, Ehm MG, Lai EH (2008) The population reference sample, POPRES: a resource for population, disease, and pharmacological genetics research. Am J Hum Genet 83(3):347–358
45. Ordovas JM, Tai ES (2008) Why study gene-environment interactions? Curr Opin Lipidol 19(2):158–167
46. Perry GH, Dominy NJ, Claw KG, Lee AS, Fiegler H, Redon R, Werner J, Villanea FA, Mountain JL, Misra R, Carter NP, Lee C, Stone AC (2007) Diet and the evolution of human amylase gene copy number variation. Nat Genet 39(10):1256–1260
47. Phillips C (2009) SNP databases. Methods Mol Biol 578:43–71
48. Pico AR, Kelder T, van Iersel MP, Hanspers K, Conklin BR, Evelo C (2008) WikiPathways: pathway editing for the people. PLoS Biol 6(7):e184
49. Povey S, Al Aqeel AI, Cambon-Thomsen A, Dalgleish R, den Dunnen JT, Firth HV, Greenblatt MS, Barash CI, Parker M, Patrinos GP, Savige J, Sobrido MJ, Winship I, Cotton RG (2010) Practical guidelines addressing ethical issues pertaining to the curation of human locus-specific variation databases (LSDBs). Hum Mutat [Epub ahead of print]
50. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, Freeman JL, Gonzalez JR, Gratacos M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J, Valsesia A, Woodwark C, Yang F, Zhang J, Zerjal T, Armengol L, Conrad DF, Estivill X, Tyler-Smith C, Carter NP, Aburatani H, Lee C, Jones KW, Scherer SW, Hurles ME (2006) Global variation in copy number in the human genome. Nature 444(7118):444–454
51. Ring HZ, Kwok PY, Cotton RG (2006) Human Variome Project: an international collaboration to catalogue human genetic variation. Pharmacogenomics 7(7):969–972
52. Saadallah AA, Rashed MS (2007) Newborn screening: experiences in the Middle East and North Africa. J Inherit Metab Dis 30(4):482–489
53. Seguin B, Hardy BJ, Singer PA, Daar AS (2008) Genomic medicine and developing countries: creating a room of their own. Nat Rev Genet 9(6):487–493
54. Singer PA, Daar AS (2001) Harnessing genomics and biotechnology to improve global health equity. Science 294(5540):87–89
55. Siva N (2008) 1,000 genomes project. Nat Biotechnol 26(3):256
56. Smith CE, Ordovas JM (2010) Fatty acid interactions with genetic polymorphisms for cardiovascular disease. Curr Opin Clin Nutr Metab Care Current. 13(2):139–144
57. Spratlin JL, Serkova NJ, Eckhardt SG (2009) Clinical applications of metabolomics in oncology: a review. Clin Cancer Res 15(2):431–440
58. Stumbo P, Weiss R, Newman J, Pennington J, Tucker K, Wiesenfeld P, Illner A-K, Klurfeld D, Kaput J (2010) Omics and environmental assessments: survey, needs, and plans of the US nutrition and nutrigenomic communities. J Nutr (in press)
59. Takeuchi F, Serizawa M, Kato N (2008) HapMap coverage for SNPs in the Japanese population. J Hum Genet 53(1):96–99
60. Teo YY, Sim X, Ong RT, Tan AK, Chen J, Tantoso E, Small KS, Ku CS, Lee EJ, Seielstad M, Chia KS (2009) Singapore genome variation project: a haplotype map of three Southeast Asian populations. Genome Res 19(11):2154–2162

61. The International HapMap C. (2005) A haplotype map of the human genome. Nature 437(7063): 1299-1320
62. Tindana PO, Singh JA, Tracy CS, Upshur RE, Daar AS, Singer PA, Frohlich J, Lavery JV (2007) Grand challenges in global health: community engagement in research in developing countries. PLoS Med 4(9):e273
63. Tracy RP (2008) 'Deep phenotyping': characterizing populations in the era of genomics and systems biology. Curr Opin Lipidol 19(2):151–157
64. van Ommen B, Bouwman J, Dragsted L, Drevon C, Elliott R, de Groot P, Kaput J, Mathers J, Muller M, Pepping F, Saito J, Scalbert A, Radonjic M, Rocca-Serra P, Travis T, Wopereis S, Evelo C (2010) Challenges of molecular nutrition research 6: the nutritional phenotype database to store, share and evaluate nutritional systems biology studies. Genes Nutr (in press)
65. van Ommen B, El-Sohemy A, Hesketh J, Kaput J, Fenech M, Evelo C, McArdle H, Bouwman J, Lietz G, Mathers J, Fairweather-Tait S, van Kranen H, van Kranen H, Elliott R, Wopereis S, Ferguson L, Méplan M, Perozzi G, Allen L, Rivero D, Group aTMGPW (2010) The micronutrient genomics project: creating a community driven knowledge base for micronutrient research. Genes Nutr (in press)
66. Wang J, Wang W, Li R, Li Y, Tian G, Goodman L, Fan W, Zhang J, Li J, Guo Y, Feng B, Li H, Lu Y, Fang X, Liang H, Du Z, Li D, Zhao Y, Hu Y, Yang Z, Zheng H, Hellmann I, Inouye M, Pool J, Yi X, Zhao J, Duan J, Zhou Y, Qin J, Ma L, Li G, Zhang G, Yang B, Yu C, Liang F, Li W, Li S, Ni P, Ruan J, Li Q, Zhu H, Liu D, Lu Z, Li N, Guo G, Ye J, Fang L, Hao Q, Chen Q, Liang Y, Su Y, San A, Ping C, Yang S, Chen F, Li L, Zhou K, Ren Y, Yang L, Gao Y, Yang G, Li Z, Feng X, Kristiansen K, Wong GK, Nielsen R, Durbin R, Bolund L, Zhang X, Yang H (2008) The diploid genome sequence of an Asian individual. Nature 456(7218):60–65
67. Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, He W, Chen YJ, Makhijani V, Roth GT, Gomes X, Tartaro K, Niazi F, Turcotte CL, Irzyk GP, Lupski JR, Chinault C, Song XZ, Liu Y, Yuan Y, Nazareth L, Qin X, Muzny DM, Margulies M, Weinstock GM, Gibbs RA, Rothberg JM (2008) The complete genome of an individual by massively parallel DNA sequencing. Nature 452(7189):872–876
68. Yim SH, Kim TM, Hu HJ, Kim JH, Kim BJ, Lee JY, Han BG, Shin SH, Jung SH, Chung YJ (2010) Copy number variations in East-Asian population and their evolutionary and functional implications. Hum Mol Genet 19(6):1001–1008
69. Zerhouni EA (2007) Translational research: moving discovery to practice. Clin Pharmacol Ther 81(1):126–128
70. Zhang W, Ratain MJ, Dolan ME (2008) The HapMap Resource is providing new insights into ourselves and its application to pharmacogenomics. Bioinform Biol Insights 2:15–23
71. Zogopoulos G, Ha KC, Naqib F, Moore S, Kim H, Montpetit A, Robidoux F, Laflamme P, Cotterchio M, Greenwood C, Scherer SW, Zanke B, Hudson TJ, Bader GD, Gallinger S (2007) Germline DNA copy number variation frequencies in a large North American population. Hum Genet 122(3–4):345–353