Genome **Medicine**

Predictive, preventive, personalized and participatory medicine: back to the future

Charles Auffray[1]*, Dominique Charron[2] and Leroy Hood[3]

Abstract

The pioneering work of Jean Dausset on the HLA system established several principles that were later reflected in the Human Genome Project and contributed to the foundations of predictive, preventive, personalized and participatory (P4) medicine. To effectively develop systems medicine, we should take advantage of the lessons of the HLA saga, emphasizing the importance of exploring a fascinating but mysterious biology, now using systems principles, pioneering new technology developments and creating shared biological and information resources.

HLA and the Human Genome Project

Shortly before he passed away last year, the late Jean Dausset said he wanted to be remembered for four disciplines to which he had made significant contributions: anthropology, immune response, transplantation and predictive medicine [1]. All of these were based on his seminal discovery of the human leukocyte antigens (HLA) in the 1950s. In 1980 he received the Nobel Prize in Physiology or Medicine jointly with Baruj Benacerraf and George Snell 'for their discoveries concerning genetically determined structures on the cell surface that regulate immunological reactions'.

Through the international and highly collaborative HLA workshops that he pioneered with colleagues in the 1960s and which have been running successfully ever since, Jean Dausset and his colleagues established the important role of the histocompatibility antigens in the control of the immune response and in the susceptibility or resistance to a wide variety of diseases, including those with an auto-immune component, and they realized that these defined a key element of the biologically unique

identities of individuals. The major histocompatibility complex (MHC), in which the genes encoding the HLA antigens are located, turned out to be a high gene density, high diversity and low recombination genetic system that has evolved to become a master coordinator of individual specificity in both innate and cognate immunity [2]. The highly polymorphic nature of the HLA antigens, initially detected through serological and cellular assays with the active participation of blood and organ transplantation donors, was exploited early on to analyze the genetics and migration history of human populations. Understanding the susceptibility or resistance to diseases and infections thus laid an early cornerstone for the creation of predictive medicine through personalized disease prevention.

These contributions expanded in scope and depth with the biochemical and functional characterization of the MHC antigens expressed on a variety of cell types, particularly within the immune system. With the cloning and sequencing of the human and murine histocompatibility genes in the early 1980s, the characterization of the HLA and H-2 loci became test cases that contributed to demonstrating the feasibility and value of genome projects, as they were the first multi-megabase genetic loci to be sequenced entirely, defining for the first time all of the genes in this gene-rich region as well as also providing a wealth of polymorphic markers for genetic linkage studies. This definition of all the genes in an important region through discovery science by sequencing a chromosomal region was the forerunner of the pioneering contribution the genome project made in essentially defining all of the genes of the human genome (and making them available to all scientists). We believe that it is important to take advantage of the lessons of the HLA/H-2 saga to enable effective and efficient deployment of the systems medicine roadmap [3,4].

Technological development is essential for high throughput, high precision and low cost data generation [5]. The initially lengthy and cumbersome serological, cellular, biochemical and genetic HLA typing assays have greatly benefited from laboratory automation and are conducted under standard operating procedures in centralized facilities in compliance with regulatory

*Correspondence: charles.auffray@vjf.cnrs.fr
[1]Functional Genomics and Systems Biology for Health, CNRS Institute of Biological Sciences, 7 rue Guy Moquet, BP8, 94801 Villejuif cedex, France
Full list of author information is available at the end of the article

authorities. Quality control exercises, standards and professional accreditation schemes were developed very early on by the HLA community to become the professional platform of histocompatibility medicine worldwide. This has been instrumental in setting the framework for the Human Genome Project, in which many MHC scientists were involved. Rapid and striking recent advances in high throughput and low cost sequencing technology are now providing access to the exponential characterization of individual genomes and transcriptomes, enabling, for example, systems genetics approaches for the identification of disease-related genes and polymorphisms [6]. Moreover, there are now exciting possibilities for analyzing autoimmune diseases, allergies and transplantation biology in powerful new assays, such as sequencing the MHC locus from family members to delineate MHC genes and their linkage relationships, using protein chips that contain more than half of the human proteins to search for autoantibodies in sera, or the ascertainment of the complete repertoires of functional antibody molecules in B cells and T-cell receptors in T cells. Together these new dimensions of exploring patient data open up new and unique opportunities for attacking immune-related diseases. In order to unravel with a similar coverage and depth the role of other key biomolecules, such as proteins, lipids, sugars and other metabolites, groundbreaking innovative technological developments will be necessary.

HLA and systems biology

The generation and management of massive amounts of sequence information are becoming dematerialized commodities based on community-agreed standards and reagents. This follows a trend that was established early on in the HLA community with the nomenclature committees and the provision of sets of standardized reference reagents for common use in the HLA workshops. This has the merit of refocusing attention on the definition of precise biological and medical questions and the identification of appropriate biological resources, which are the necessary starting points of any successful systems biology approaches. In this context, public-private partnerships based on well characterized and openly shared biological resources and data repositories, taking into account issues of intellectual property rights and active patient participation, individually and through patient organizations, are vitally important [7]. The pre-competitive Innovative Medicines Initiative, a joint undertaking of the European Union and the European Federation of Pharmaceutical Industry Associations [8], and the Program on Public-Private Partnerships of the United States National Institutes of Health Roadmap [9] are initial examples of these developments. They should be encouraged to expand and network on a worldwide

basis in the same spirit as the HLA workshops, with active participation of technology providers, patient organizations and regulatory authorities.

Analysis of the sequence, polymorphism, transcription and protein-expression profiles of various MHCs and of the three-dimensional structure of the histocompatibility antigens have required and triggered the development of a wide variety of computational tools for data collection, storage and analysis. They enabled elucidation of structure-function relationships of the encoded cell surface antigens, illuminating their ability to bind and present foreign peptides to immune effector cells. They also provided the computational framework for the formation of electronic public repositories, which have been greatly expanded with the advent of genomics. The development in the recent years of increasingly powerful computational resources is proving essential for maintaining the ability to access and analyze the exponentially increasing flow of diverse data types on biomolecules, cells, organs and the associated normal and pathological phenotypes. Their implementation in computer grids is providing computing-on-demand cloud-based resources that can be managed externally, enabling researchers to focus their attention on the development and use of mathematical and computational tools. The tools required for the successful application of these computational resources to biomedical problems are complex because systems biology approaches typically involve iterative cycles of experiment and modeling [10,11].

The challenge of the 21st century for all scientific and engineering disciplines is understanding and managing complexity. In this context, biology is unique in having initiated the development of powerful tools to attack biological complexity, such as systems approaches to the control of physiology, development and complex disease [12-14]. Emerging high throughput biological tools are driven by the emergence of microfluidics and nanotechnology and by the biology-driven creation of powerful and novel computational and mathematical tools. This raises the possibility that biology will dominate 21st century science in part because it holds the technological capability to attack some of the most fundamental problems of humankind, such as global healthcare and well-being, energy and environmental management, sustainable agriculture and nutrition.

HLA and systems medicine

Hematopoietic stem cell and organ transplantation represent a unique opportunity to study the complex and systemic effects of introducing a new genome in an individual. The drastic biological perturbations caused depend on both the genetic make-up of the donor and that of the recipient. Donor-recipient histocompatibility

can be assessed by correlating biological and physio-logical effects, thus identifying the genetic contribution of MHC alleles, haplotypes and their combinations. This opens up, for example, the possibility of neutralizing MHC genetic differences by performing hematopoietic stem cell transplantation between identical siblings. Such a clinical setting makes it possible to highlight the effects of other immunogenetic systems, such as the minor histocompatibility loci, which have a more limited yet significant influence on the outcome of transplantation [15]. Once again the genome project has transformed the search for the genes encoding minor histocompatibility loci through its complete parts list of all human genes. There is a strong incentive to develop and validate cell therapeutics using allogeneic cells and/or tissues to treat chronic diseases and aging. The experience gained in transplantation medicine will directly benefit regenerative medicine by fostering an integrated approach of biology, medicine, logistics, and regulatory and social issues.

Among the major bottlenecks in translating systems biology into systems medicine are the limited number of clinical cases that can be included in randomized trials and the number of genetic, environmental and nutritional variables that cannot be easily accounted for. Trans-plantation medicine has been confronted with this issue from the beginning of its development. Validation of the role of each individual HLA locus was established sequentially by comparing populations of patients of different but known genetic composition (monozygotic twins, siblings, haplotype-matched family members and HLA-matched unrelated individuals). This has been instrumental in deciphering the role of the different HLA loci, alleles or haplotypes. With the increasing number of genetic systems and alleles that have to be taken into account, the availability of real patients is limited and insufficient. This triggers the need to develop virtual patient models that mimic the patients' main charac-teristics, from which testable hypotheses can be generated and validated on the small number of actual patients available.

In addition to transplantation and regenerative medi-cine, other important diseases, such as auto-immune diseases for which HLA has been confirmed as the major genetic factor, could benefit from an HLA-based systems biology approach. Insulin-dependent diabetes mellitus, rheumatoid arthritis, ankylosing spondylitis and multiple sclerosis are among the first obvious candidates. Moreover, given that recent genome-wide association studies have revealed a rather unexpected immune component in psychiatric disorders such as schizophrenia [16], these could also be amenable to the same approach. The complexity of the HLA system has led us right into the complexities of predictive, preventive, personalized and participatory (P4) medicine.

Looking into the future, on the basis of its past history, we can be certain of at least one thing: the HLA system will continue to have a pivotal integrative role in the advent of P4 medicine and should therefore be con-sidered as a hallmark of systems medicine.

Authors' contributions
CA wrote the first draft, and all authors contributed to and approved the final draft.

Author details
[1]Functional Genomics and Systems Biology for Health, CNRS Institute of Biological Sciences, 7 rue Guy Moquet, BP8, 94801 Villejuif cedex, France. [2]Jean Dausset Laboratory, Inserm U940 & Paris Diderot University, Hôpital Saint Louis 1, Avenue Claude Vellefaux, 75475 Paris cedex 10, France. [3]Institute for Systems Biology, 1441 North 34th Street, Seattle, WA 98103-8904, USA.

References
1. Charron DJ: **Obituary: Jean Dausset (1916-2009).** *Nature* 2009, **460**:338.
2. Charron DJ: **Immunogenetics today: HLA, MHC and much more.** *Curr Opin Immunol* 2005, **17**:493-497.
3. Auffray C, Chen Z, Hood L: **Systems medicine: the future of medical genomics and healthcare.** *Genome Med* 2009, **1**:2.
4. Clermont G, Auffray C, Moreau Y, Rocke DM, Dalevi D, Dubhashi D, Marshall DR, Raasch P, Dehne F, Provero P, Tegner J, Aronow BJ, Langston MA, Benson M: **Bridging the gap between systems biology and medicine.** *Genome Med* 2009, **1**:88.
5. Hood L, Heath JR, Phelps ME, Lin B: **Systems biology and new technologies enable predictive and preventative medicine.** *Science* 2004, **306**:640-643.
6. Roach JC, Glusman G, Smit AF, Huff CD, Hubley R, Shannon PT, Rowen L, Pant KP, Goodman N, Bamshad M, Shendure J, Drmanac R, Jorde LB, Hood L, Galas DJ: **Analysis of genetic inheritance in a family quartet by whole-genome sequencing.** *Science* 2010, **328**:636-639.
7. Auffray C: **Sharing knowledge: a new frontier for public-private partnership in medicine.** *Genome Med* 2009, **1**:29.
8. **The Innovative Medicines Initiative** [http://www.imi-europe.org/Pages/default.aspx]
9. **NIH Common Fund** [http://nihroadmap.nih.gov/]
10. Wolkenhauer O, Fell D, De Meyts P, Blüthgen N, Herzel H, Le Novère N, Höfer T, Schürrle K, van Leeuwen I: **SysBioMed report: advancing systems biology for medical applications.** *IET Syst Biol* 2009, **3**:131-136.
11. Stein L: **The case for cloud computing in genome informatics.** *Genome Biol* 2010, **11**:207.
12. Bonneau R, Facciotti MT, Reiss DJ, Schmid AK, Pan M, Kaur A, Thorsson V, Shannon P, Johnson MH, Bare JC, Longabaugh W, Vuthoori M, Whitehead K, Madar A, Suzuki L, Mori T, Chang DE, Diruggiero J, Johnson CH, Hood L, Baliga NS: **A predictive model for transcriptional control of physiology in a free living cell.** *Cell* 2007, **131**:1354-1365.
13. Oliveri P, Tu Q, Davidson EH: **Global regulatory logic for specification of an embryonic cell lineage.** *Proc Natl Acad Sci USA* 2008, **105**:5955-5962.
14. Hwang D, Lee IY, Yoo H, Gehlenborg N, Cho JH, Petritis B, Baxter D, Pitstick R, Young R, Spicer D, Price ND, Hohmann JG, Dearmond SJ, Carlson GA, Hood LE: **A systems approach to prion disease.** *Mol Syst Biol* 2009, **5**:252.
15. Dickinson AM, Charron D: **Non-HLA immunogenetics in hematopoietic stem cell transplantation.** *Curr Opin Immunol* 2005, **17**:517-525.
16. Shi J, Levinson DF, Duan J, sanders AR, Zheng Y, Pe'er I, Dudbridge F, Holmans PA, Whittemore AS, Mowry BJ, Olincy A, Amin F, Cloninger CR, Silverman JM, Buccola NG, Byerley WF, Black DW, Crowe RR, Oksenberg JR, Mirel DB, Kendler KS, Freedman R, Gejman PV: **Common variants on chromosome 6p22.1 are associated with schizophrenia.** *Nature* 2009, **460**:753-757.