Commentary
A need for a 'whole-istic functional genomics' approach in complex human diseases: arthritis

Ashok R Amin

Hospital for Joint Diseases/NYU School of Medicine, New York, USA

Correspondence: Ashok R Amin (e-mail: ashok.amin@msnyuhealth.org)

Abstract

'Genomic tools', such as gene/protein chips, single nucleotide polymorphism and haplotype analyses, are empowering us to generate staggering amounts of correlative data, from human/animal genetics and from normal and disease-affected tissues obtained from complex diseases such as arthritis. These tools are transforming molecular biology into a 'data rich' science, with subjects with an '-omic' suffix. These disciplines have to converge and integrate at a systemic level to examine the structure and dynamics of cellular and organismal function ('functionomics') simultaneously, using a multi-dimensional approach for cells, tissues, organs, rodents and Zebra fish models, which intertwines various approaches and readouts to study the development and homeostasis of a system. In summary, the postgenomic era of functionomics will facilitate narrowing the bridge between correlative data and causative data, thus integrating 'intercoms' of interacting and interdependent disciplines and forming a unified whole.

Keywords: arthritis, genomics, inflammation, proteomics

Introduction

Arthritis is a complex disease with an unknown etiology. Some of the common underlining symptoms include inflammation, dysfunction of joints due to destruction of cartilage and soft tissue. Based on the clinical symptoms, arthritis can be classified as osteoarthritis (OA), rheumatoid arthritis, synovial lipomatosis, avascular necrosis, crystal deposition disease, Goud and other diseases [1].

A major challenge we face in the postgenomic era is the characterization of genes involved in oligogenic and polygenic disorders such as arthritis. This is because, unlike monogenic diseases, pedigrees from complex diseases reveal no Mendelian inheritance patterns, and gene mutations are neither sufficient nor necessary to explain the disease phenotypes.

Arthritis is a disease with complex traits influenced by various risk factors. Multiple genetic, environmental and epistatic determinants represent the greatest challenge for genetic analysis, largely due to the difficulty of isolating the phenotype of one gene amid the noise of other genetic and environmental influences. It may be recognized that the complexity is hidden in idealized laboratory settings and in normal operations, but this complexity becomes conspicuous when one notices a rare cascading failure, primarily due to paradoxical features that keep together the robustness, modularity, feedback, repair and fragility of the complex biological system in arthritis.

The knowledge of new genomic information and the tools to decipher it obviates the necessity to reassess our working hypothesis. The 'genomic tools' will, for the first time, allow us to analyze small amounts of surgical samples (such as needle biopsies) and to analyze clinical samples or cells (yielding 10–100 pg nucleic acids) in the context of the whole genome.

Preliminary genomic analysis in OA has already resurrected the debate on OA or osteoarthrosis based on the

IL = interleukin; OA = osteoarthritis; PCR = polymerase chain reaction.

semantic issues in the definition of inflammation in cartilage in the postgenomic era of molecular medicine [2,3]. This has challenged a 20-century-old definition of inflammation proposed by Cornelius Celsius. He defined inflammation (redness and swelling with heat and pain [*rubor et tumor cum calore et dolor*]) as an entity using a singular rather than a plural noun, implying that it might be a single process or a type of process. The avascular, alymphatic and aneural human OA-affected articular cartilage harboring chondrocytes (like activated macrophages, but not normal chondrocytes) shows superinduction of inflammatory mediators as observed by gene chip analysis, but fails to show the cardinal signs of inflammation [3]. These types of analyses will not only facilitate development of unbiased hypotheses at the molecular level, but will also assist us in following the scent to the identification and characterization of novel targets and disease markers for pharmacological intervention, gene therapy and diagnosis.

A system approach to arthritis

'General System Theory', proposed in 1940, has pervaded all fields of science and has penetrated into popular thinking in psychology, economics and social sciences. The postgenomic revolution has redefined 'System Biology' or 'Whole-istic Biology' [4,5]. Unraveling the genetics of human diseases such as arthritis will require moving beyond the focus on one gene at a time to exploring pleiotropism, epistasis and environmental dependency of genetic effects by integrating various technologies and datasets forming a unified whole. There is consensus among various investigators that a single genetic approach is not sufficient to give a comprehensive analysis of a complex disease, but rather would require an entire arsenal of approaches as recently described by Amin and coworkers [5,6].

A strategy for genomic analysis in arthritis

Reliable analysis of complex human diseases such as arthritis will require graspable knowledge of the functional interactions between key components of cells (such as T cells, macrophages, neutrophils, osteoclasts, chondrocytes and synovial cells), tissues (synovium, bone and cartilage) and systems (mobile joints in animal models such as rodents and Zebra fish), as well as the interactions that change in the disease state (clinical material and diagnosis) (Fig. 1). This information resides neither in the genome nor in individual gene(s)/protein(s), but it seems to lie at the level of protein interactions within the context of subcellular, cellular, tissue, organ and system structure.

A system biology approach to functional genomics in arthritis is shown in Fig. 1. The scheme shows the role and involvement of various cell types, tissues and organs, and the use of animal models to understand the pathophysiology of arthritis. Understanding expression and functions of 'uncharacterized genes' in target cells and various (normal and disease) tissues requires the use of different cell types in the complex interaction and interplay. The synovium can be classified and analyzed as normal and hypertrophic, and the latter can be subdivided as cartilage invasive and noninvasive in different forms of arthritis [7]. The subchondral bone has been impacted significantly in these diseases, as observed by the remodeling and thickening in OA. The combined role of all five cell types (T cells, macrophages, neutrophils, osteoclasts and chondrocytes) is important to understand the pathogenesis of arthritis [8]. They may be acting as complex traits fine tuning the disease process.

Mouse and Zebra fish models (knockin/knockout) have been proven to mimic symptoms observed in man, as shown for type II collagen and endothelin, respectively [9,10]. For example, endothelin and its receptor were found to be differentially expressed in normal and human OA-affected cartilage (Amin, Attur and Dave, unpublished data, 2003). A mutation of *sucker* that encodes a Zebra fish endothelin 1 showed distortion of the ventral cartilage, the pharyngeal segments and craniofacial development in endothelin receptor-deficient mice [10,11]. Functional genomics requires an integrated team of experts including biochemists, cell biologists, structural biologists, physiologists and geneticists to create a unified whole due to the unknown nature of genes to be analyzed and the type of expertise regained. The structure–function relationship of differentially expressed genes in normal and diseased tissue can be analyzed in cells to organ cultures, as recently described for a type II IL-1β decoy receptor [12].

At least four technologies have been extensively used for gene mining and functional genomics. Figure 1 also shows various approaches that can be applied selectively or simultaneously to various cell types, organs, and animal models and human subjects to understand the structure–function relationship of genes in arthritis. These include gene expression arrays, real-time PCR, proteomics, high-throughput DNA sequencing, single nucleotide polymorphism and haplotyping analysis, and 2D-matrix assisted laser desorption ionization-time of flight (2D MALD-TOF) [13,14].

Gene and protein mining technologies such as gene expression array, proteomics, single nucleotide polymorphism, haplotyping and linkage disequilibrium, and microsatellites generate a significant amount of correlative data that requires annotating using various bioinformatic platforms. Although computer-intensive disciplines and bioinformatics allow clustering analysis for gene expression arrays and provide insight into the 'correlation' among genes and biological phenomena, they have limitations in revealing the 'causality' of regulatory relationships and/or predicting *ab initio* gene structure, gene function and protein folds from the raw sequence data.

Figure 1
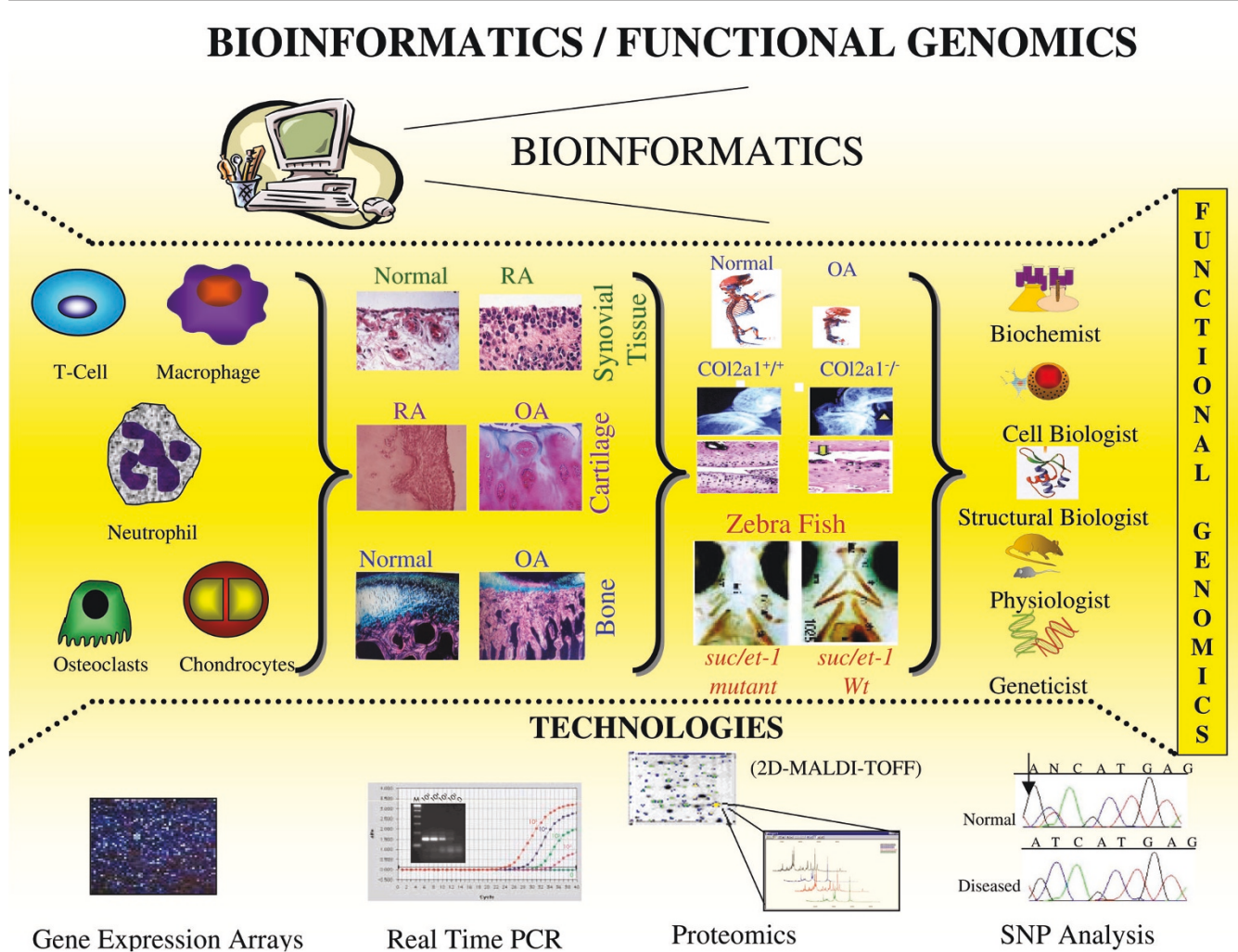
An integrative system biology approach to functional genomics in arthritis. 2D-MALDI-TOFF, 2D-matrix assisted laser desorption ionizartion-time of flight; OA, osteoarthritis; RA, rheumatoid arthritis; PCR, polymerase chain reaction; Wt, Wild type.

The key to bioinformatics is integration, interpretability between various data platforms and the ability to visualize retrieved complex data in a way that aids their interpretation. Integrating various incompatible bioinformatics platforms is essential. Such efforts are currently under way by the Interoperable Informatics Infrastructure Consortium, a computer hardware 14-member organization. In summary, bioinformatics facilitates deriving hypotheses allowing us to enter the network structure, followed by identifying structure–function relationships using other tools.

Functional genomics

Genomics has provided us with a massive 'parts catalog' for the human body in normal and disease states. Proteomics seems to define some of these individual 'parts' and the structures they form in detail. There is no 'user's guide' describing how these parts are put together to allow these interactions that sustain life or cause diseases. However, the new emerging field of functional genomics will provide such information.

Functional genomic analysis involves a systematic effort to understand the function of genes and gene products (transcripts and proteins) and their role in biological systems (cells, tissues and organisms), until now classically performed for single genes (e.g. generation of mutants, analysis of proteins and transcripts), in the context of the whole genome. While an understanding of genes and proteins continues to be important, the focus should be on ascertaining a system's structure and its dynamics.

Inspecting genome databases and expression arrays (of an enzyme, transporter, receptor or ligand) without their integrative functional knowledge with respect to various

forms of arthritis will be a starting point for functional genomics in this area. These include a gene-driven approach and a phenotype-driven approach. Both strategies are complimentary, leading collectively to association of the phenotype with genotypes, as recently reported [5,6].

Conclusions and future directions

Functional genomics will begin to mature in the coming decade into a coherent science (as molecular biology did in the last half of the previous century), and its constituent fields will become clearer. It is likely to give a whole new meaning to clinical and genomic-based translational research and biomarkers of over 35,000 possible data points. The potential for its applications are infinite. The present climate faces several challenges for those attempting to perform genomic research on human subjects, including informed consent, public acceptance, sample collection and storage, and current technological capabilities and cost. Among the several subcategories of genomics, functional genomics is most closely linked to pharmacogenomics. This has generated hype and hope for a continuous metamorphosis of molecular medicine, individualized drug therapy and pharmaceutical drug development.

A lot clearly needs to be done as more than 40% of the 35,000 genes (and possibly 120,000 different proteins they may code) have not been ascribed any functional attribute [15], neither a biochemical function (e.g. kinase), a cellular function (e.g. a specific signaling pathway) or a function at the tissue/organism level (e.g. synovial hypertrophy, cartilage homeostasis, etc.). There is presently a significant amount of 'data dumping' generated by arrays and automation that does not make much sense. To explore such a vast genome space, new technologies that exploit and link genome and clinical data to ask entirely new kinds of questions about the complex nature of arthritis will be essential. Modern biologists, both accomplished professionals and students, are unfortunately ill-prepared for this changing role because of the understandable bias in their background towards experimental techniques and results. Ultimately, we will have to adapt.

Competing interests

None declared.

References

1. McCarty DJ: **Differential diagnosis of arthritis: analysis of signs and symptoms.** In *Arthritis and Allied Conditions: A Textbook of Rheumatology*. Edited by Koopman WJ. Philadelphia: Lippincott; 1998:39-50.

2. Aigner T, McKenna L: **Molecular pathology and pathobiology of osteoarthritic cartilage.** *Cell Mol Life Sci* 2002, **59**:5-18.
3. Attur MG, Dave M, Akamatsu M, Katoh M, Amin AR: **Osteoarthritis or osteoarthrosis: the definition of inflammation becomes a semantic issue in the genomic era of molecular medicine**. *Osteoarthritis Cartilage* 2002, **10**:1-4.
4. Chong L, Ray LB: **Whole-istic biology [abstract]**. *Science* 2002, **295**:1661.
5. Attur MG, Dave, MN, Tsunoyama K, Akamatsu M, Kobori M, Miki J, Abramson SB, Katoh M, Amin AR: **'A system biology' approach to bioinformatics and functional genomics in complex human diseases: arthritis.** *Curr Issues Mol Biol* 2002, **4**:129-146.
6. Amin AR: **Pharmacogenomics: hype, hope, or metamorphosis of molecular medicine and pharmaceuticals**. *Trends Pharmacol Sci* 2002, **23**:583.
7. Watanabe N, Ando K, Yoshida S, Inuzuka S, Kobayashi M, Matsui N, Okamoto T: **Gene expression profile analysis of rheumatoid synovial fibroblast cultures revealing the overexpression of genes responsible for tumor-like growth of rheumatoid synovium**. *Biochem Biophys Res Commun* 2002, **294**:1121-1129.
8. Gu J, Marker-Hermann E, Baeten D, Tsai WC, Gladman D, Xiong M, Deister H, Kuipers JG, Huang F, Song YW, Maksymowych W, Kalsi J, Bannai M, Seta N, Rihl M, Crofford LJ, Veys E, De Keyser F, Yu DT: **A 588-gene microarray analysis of the peripheral blood mononuclear cells of spondyloarthropathy patients.** *Rheumatology* 2002, **41**:759-766.
9. Aszodi A, Hunziker EB, Olsen BR, Fassler R: **The role of collagen II and cartilage fibril-associated molecules in skeletal development.** *Osteoarthritis Cartilage* 2001, **9(suppl A)**:S150-S159.
10. Clouthier DE, Williams SC, Yanagisawa H, Wieduwilt M, Richardson JA, Yanagisawa M: **Signaling pathways crucial for craniofacial development reveal by endothelin-A receptor-deficient mice**. *Dev Biol* 2000, **217**:10-24.
11. Miller CT, Schilling TF, Lee K, Parker J, Kimmel CB: **Sucker encodes a zebrafish endothelin-1 required for ventral pharyngeal arch development**. *Development* 2000, **127**:3815-3828.
12. Mukundan AG, Dave M, Cipolletta C, Kang P, Goldring MB, Patel IR, Abramson SB, Amin AR: **Reversal of autocrine and paracrine effects of interleukin 1 (IL-1) in human arthritis by type II IL-1 decoy receptor**. *J Biol Chem* 2000, **275**:40307-40315.
13. Bajorath J: **Rational drug discovery revisited: interfacing experimental programs with bio- and chemo-informatics.** *Drug Discov Today* 2001, **6**:989-995.
14. Lai E: **Application of SNP technologies in medicine: lessons learned and future challenges.** *Genome Res* 2001, **11**:927-929.
15. Yaspo, ML: **Taking a functional genomics approach in molecular medicine.** *Trends Mol Med* 2001, **7**:494-501.

Correspondence

Ashok R Amin, PhD, Director, Rheumatology Research and Laboratory for Functional and Pharmacogenomics in Musculoskeletal Diseases, Hospital for Joint Diseases/NYU School of Medicine, 301 East 17th Street, Room 1600, New York, NY 10003, USA. Tel: +1 212 598 6537; fax: +1 212 598 7604; e-mail: ashok.amin@msnyuhealth.org