From -omics to personalized medicine in nephrology: integration is the key

Francesco Pesce¹, Saad Pathan² and Francesco Paolo Schena³

¹Department of Genomics of Common Disease, School of Public Health, Imperial College London, London, UK, ²Department of Immunology, University College London (UCL), Royal Free Hospital, London, UK and ³Centro Addestramento Ricerca Scientifica in Oncologia (C.A.R.S.O.) Consortium, Valenzano, Bari, Italy

Abstract

Large-scale gene, protein and metabolite measurements (‘omics’) have driven the resolution of biology to an unprecedented high definition. Passing from reductionism to a system-oriented perspective, medical research will take advantage of these high-throughput technologies unveiling their full potential. Integration is the key to decoding the underlying principles that govern the complex functions of living systems. Extensive computational support and statistical modelling is needed to manage and connect the -omic data sets but this, in turn, is speeding up the hypothesis generation in biology enormously and yielding a deep insight into the pathophysiology. This systems biology approach will transform diagnostic and therapeutic strategies with the discovery of novel biomarkers; nephrology; omics; personalized medicine; systems biology.

Keywords: biomarkers; nephrology; omics; personalized medicine; systems biology

Introduction

The complexity and adaptiveness of living systems can be read through self-organized highly interconnected networks whose interacting components, dynamically coordinated in hierarchical patterns, bring to life novel, unexpected emergent properties. Indeed, the system as a whole is unique in its features and, as the flow of life unfolds from DNA to the human being, countless distinct elements contribute to building its layered structure. Each contributes with its own role, yet precisely working seamlessly together.

Systems biology aims to consider the interactions between all the components of biological systems within a manageable format that is amenable to analysis [1]. This approach has been made possible by the derivation of mathematical models, computational tools and the integration of experimental data from genomics, transcriptomics, proteomics and metabolomics. These sciences have paved the way for systems biology as they yield, from each subsystem, a large amount of data capturing all the constituents considered collectively. The -omic cascade, from the potentiality of ‘what can happen’ (genome) through ‘what appears to happen’ (transcriptome) and ‘what makes it happen’ (proteome) to ‘what has happened’ (metabolome; Figure 1), embodies the paradigm of what needs to be modelled. Integration will unveil the full potential of these high-throughput technologies leading to a comprehensive decoding of the upper emergent level, the phenotype. Systems biology is a novel field pitched at decoding -omic dynamic interactions and adding an additional dimension to that of a classical homeostatic model of physiology. Given the high-throughput nature of its experimental techniques, data managing and statistical analysis are computationally intensive. Moreover, while research in traditional biology is driven by successive cycles of hypothesis and testing, systems biology advances, as an extension of -omic techniques, in an unbiased fashion without prior hypothesis but first gathering data and then generating hypothesis after analysis and modelling. Two different ways to proceed have been simplified as the ‘bottom-up’ approach, through -omics, taking advantage of the global measurement of a subsystem component or the ‘top-down’ approach defining integrative models (across scales) from human physiology and disease [2]. The effort to ‘put together’ is fostered by appreciating the holistic and composite characteristics of a problem to catch its emergent properties, rather than dividing a complex problem into its smallest component parts. Systems biology is the suitable complement for reductionism, which has been celebrated as the highest achievement of medical science and, though at its climax, now has shown its limits [3].

Omic experimental data are managed to a great extent by bioinformatics tools and databases enabling classification and categorization into known biological processes,
In silico simulations and these, in turn, revalidate by wet-lab functional experiments. These massive in silico analyses need to be part of a ‘validation loop’ as pointed out by Molina et al. [4], where they are validated by wet-lab functional experiments and these, in turn, refine and strengthen the in silico model.

Application in nephrology

The regulation of the kidney can be compartmentalized into distinct subsystems with interconnections at the different levels of the gene, protein, cell and tissue. Although genomics has revealed novel mutations responsible for kidney diseases, the mechanisms of dysregulation still remain to be understood. Next generation sequencing (NGS) is the latest technology in genomics but previously we have witnessed the dramatic increase in genome-wide association studies (GWASs) based on microarrays to scan hundreds of thousands of single-nucleotide polymorphisms across the whole genome for the association with a disease or quantitative traits. The belief that the field of genomics with simple GWAS or even the more comprehensive NGS could unveil all the genetic determinants of complex diseases upholds the enormous emphasis in reductionism. GWASs have yielded significant genetic associations for complex genetic diseases, such as IgA nephropathy (IgAN) [5–7] and focal segmental glomerulosclerosis (FSGS) [8], but they are hampered by the two major issues of missing heritability and population stratification. In chronic kidney disease (CKD), less than 2% of the estimated heritability of the estimated glomerular filtration rate has been explained [9, 10], whereas the difference in allele frequencies between subpopulations (probably due to different ancestry) is a tricky matter when sampling IgAN for large association studies.

Transcriptomics is the high-throughput expression profile of RNAs in a specific cell or tissue, in a particular moment or condition, to capture its specific and complete gene expression set. This has been the most widely used -omic technique so far [11]. Many examples are available in IgAN [12, 13], diabetic nephropathy (DN) [14], FSGS [15], lupus nephritis (LN) [16] and CKD [17]. A European cooperative multicentre project, the European Renal cDNA Bank—Kroener-Fresenius Biopsy Bank (ERCB—KFB), was initiated in 1998 with the purpose of studying gene expression following a common protocol for the high-throughput application of gene expression analysis of renal biopsies. Transcriptomics has been mainly based on microarrays so far, while measuring RNA by NGS is an unbiased method for genome-wide transcription factor binding-site profiling and transcription characterization that is called RNA-Seq [18]. The new approach for whole-transcriptome sequencing with RNA-Seq also yields quantitative results with greater accuracy than microarrays. This technique can measure the absolute quantity and capture transcriptome dynamics across distinct systems without extensive normalization of data sets. Protein–DNA interactions by chromatin immunoprecipitation (ChIP) combined with the quantitative measurements of mRNA is also being used to better investigate regulatory mechanisms influencing the transcriptome. After the microarray era, it is time for ChIP-Seq, applying high-throughput sequencing to this technique as well, in order to achieve higher specificity, sensitivity and genome-wide comprehensiveness [19, 20]. Moreover, it has to be taken into account that transcription factor binding and promoter expression can be regulated by methylation, chromatin rearrangement and epigenetic mechanisms, which are also influenced by an environmental factor. Other avenues for predicting kidney disease include the measurement of micro-RNAs (miRNAs) [21]. These are 21–23 nucleotides in length that are evolutionarily conserved and are able to regulate the expression of protein-coding genes. It is worth noting that miRNAs are stable in tissues and biological fluids; therefore, they may be used as potential diagnostic biomarkers [22]. They are implicated, in podocyte damage [23, 24], ischaemia-reperfusion injury [25], renal fibrosis [26, 27] and glomerulonephritis pathogenesis [28–32]. A systems biology approach has been recently used to describe the role of miRNAs in the global transcriptional reprogramming during the progression of autosomal dominant polycystic kidney disease (ADPKD) [33].

Now, we might consider that, if GWASs consist of linking a genetic variance to a disease or a quantitative phenotypic trait, by extension, gene expression levels represent a quantitative trait to be used for this purpose. This example of integration between genomics and transcriptomics, which discover genetic loci influencing gene expression levels, is called expression quantitative trait locus mapping and has been applied by Papeta et al. [34] in the murine model of HIV-associated nephropathy.

Proteomics is the large-scale profile of proteins, and data suggest that there is the potential to validate novel...
bimarkers in renal diseases as reviewed by Konvalinka et al. [35]. Given its non-invasive nature and easy collection, urine is an ideal biofluid for biomarker discovery and exploitation in renal diseases through proteomics as shown by the application in DN [36], LN [37], IgAN [38] and renal transplant rejection [39]. The design of proteomic studies is critically important as the proteome may be affected by several confounding factors related to both the sample characteristics and the technique itself. Different methods used in proteomics include two-dimensional polyacrylamide gel electrophoresis, liquid chromatography-mass spectrometry (LC-MS), matrix-assisted laser desorption/ionization-time-of-flight MS and capillary electrophoresis coupled with electrospray ionization MS. The efforts now need to be focused on the standardization of procedures for clinical applications. The Human Kidney and Urine Proteome Project has been set up to provide the standardized protocol for urine collection and storage (http://www.hkupp.org), and the EuroKUP (‘Urine and Kidney Proteomics’) is an action supported by COST (‘European Cooperation in the Field of Scientific and Technical Research’) promoting translational proteomic research in kidney diseases by bringing together and facilitating interactions between basic scientists and clinicians. Mischak et al. [40] have recently analysed the issues in translating proteomic biomarkers into clinical practice proposing an implementation roadmap that strongly relies on a multidisciplinary approach to make it work.

Metabolomics can be defined as the non-targeted measurement of all the low-molecular-weight compounds, particularly endogenous but also exogenous, present in a fluid or tissue of a living organism. Metabolites represent the downstream execution of genome, transcriptome and proteome information. All metabolic profiling is based on nuclear magnetic resonance, gas chromatography or LC-MS methods, which are less expensive compared with other -omic technologies such as proteomics and gene arrays [41]. This technique has a wide range of applications in kidney diseases. From early diagnostic biomarkers in acute kidney injury to the analysis of the biochemical basis of the uremic syndrome, glomerulonephritis, ADPKD and kidney cancer, several papers describe the potential and show the pitfalls of this -omic science as extensively reviewed in Weiss and Kim [42].

While moving beyond to broaden the snapshot taken from a biological system, spanning the genome to the metabolome level, it is noteworthy that the scope to integrate data sets generated from different -omic techniques has to face the straits of lacking common identifiers (Table 1). Software environments like ‘Bioinformatics Resource Manager’ [43] have been developed to aid in managing, merging and integrating these data with functional annotation and interaction data from public sources also providing connectivity to visual analytic tools based on networks such as Cytoscape [44]. Ghosh et al. [45] provided a comprehensive review of these tools along the whole computational workflow of systems biology. Such a complex integrative process is not straightforward but has clear advantages. Even in related domains like transcriptomics and proteomics, there are potential pitfalls to be aware of, as Perco et al. [46] pinpointed when they compared and analysed the transcriptomic and proteomic profiles in the context of CKD, thus proposing the mapping of heterogeneous -omic profiles on the level of protein interaction networks as a method. A combination of transcriptomics and proteomics with conventional toxicological parameters has been proven to add an extra value in the identification of novel, specific candidate biomarkers for induced kidney proximal tubule damage [47] and using a similar methodology, based on ingenuity pathway analysis, but applied on urine metabolome, Lv et al. [48] revealed a time-dependent biochemical perturbation induced by gentamicin toxicity.

Online resources to take the advantage of the systems biology approach to research, specifically designed for nephrology, are now becoming available as well. Platforms like ‘Nephromine’ (http://www.nephromine.org/resource/login.html) and KUPCKB (The Kidney & Urinary Pathway Knowledge Base) [49] are valuable tools for data mining, exploration and network visualization. Moreover, given the challenging interdisciplinary nature of systems biology, the collaborative initiatives born to tackle this issue, such as SysKid ‘Systems Biology towards Novel Chronic Kidney Disease Diagnosis and Treatment’ (http://www.syskid.eu/), are remarkable.

### Conclusions

The need of biomarkers in nephrology for early diagnosis, prediction of progression and evaluation of treatment response is a priority because so far, diagnosis and treatment decisions have been based on histology, limited serologic markers and clinical parameters [50]. In the near future, -omics will improve the classification schemes in renal disease and transplantation on a refined molecular level with the advent of renal histogenomics in clinical practice [51]. As Trusheim et al. [52] underline, moving from empirical to stratified and then individualized medicine in the therapeutic continuum requires three key factors in order to define clinically specific subclasses: refined molecular fingerprints, multiple therapeutic options and clinical biomarkers.

Set to be a highly effective discovery tool, completed by functional studies, systems biology promises to identify disease pathways and biomarkers putting a significant dent in the last step towards personalized medicine. The ‘P4’ medicine of the future according to Leeroy Hood, the co-founder and president of the ‘Institute for Systems

<table>
<thead>
<tr>
<th>Table 1. Techniques used in -omics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Genomics</strong></td>
</tr>
<tr>
<td><strong>Transcriptomics</strong></td>
</tr>
<tr>
<td><strong>Proteomics</strong></td>
</tr>
<tr>
<td><strong>Metabolomics</strong></td>
</tr>
</tbody>
</table>

NGS, next generation sequencing; 2D-PAGE, two-dimensional polyacrylamide gel electrophoresis; MALDI-TOF-MS, matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry; CE-ESI-MS, capillary electrophoresis coupled with electrospray ionization mass spectrometry; NMR, nuclear magnetic resonance; GC-MS, gas chromatography-mass spectrometry; LC-MS, liquid chromatography-mass spectrometry.
**Omic integration in nephrology**

Biology’ in Seattle, will be personalized, predictive, preventive and participatory medicine [53]. The extended biomarker panel obtained at disease onset, or as an early individual predictor, might also be further exploited to evaluate the time-dependent trend of such extended and dynamic multi-level -omic profiling as actually demonstrated by Chen et al. [54].

Indeed, a probabilistic health plan will be based on individual genetic and molecular fingerprints, managing to prevent diseases in at-risk subjects and to tailor custom therapeutic interventions, with the patient always actively involved in the choice at each stage.

**Acknowledgements.** F.P. is supported by long-term fellowships from ERA-EDTA (European Renal Association-European Dialysis and Transplant Association) (ERA-EDTA ALTF 72-2010 and ERA-EDTA ALTF 89-2011).

Conflict of interest statement. None declared.

**References**


Received for publication: 22.5.12; Accepted in revised form: 10.9.12