A comment on:

A brief survey of major discoveries that impact on eukaryotic cellular pathway modelling (V.Parisi, V. De Fonzo, F. Aluffi-Pentini)

In their work the authors do a very precious work by driving the attention of the readers on recent achievements of molecular biology that open completely new perspectives on eukaryotic regulation. They actually blur an apparently very straightforward picture turning biological regulation into a complex and intermingled superposition of many and contrasting elements.

The authors message can be summarized as ' the new results show us that the classical differential equation-based modeling-style of metabolic pathways is presently unfeasible, because enzyme concentrations depend upon too many factors to be modeled .

Nevertheless, the great deal of work on these themes will in the near future enable us to take explicitly into consideration all these intricacies and perform realistic life-like simulations'.

In my opinion, it is worth noting that in the history of science (as well as in the empirical practice of medicine) considering the most basic layer as the only possible level for modeling , very rarely was a successful strategy. Since XVII Century, a strict Newtonian scientist like Robert Boyle, was fully aware of the impossibility to model gases at the molecular level: he actually shifted to a completely different level, looking for collective macro-parameters as pressure and volume. This strategy resulted in the very efficient and powerful science of Thermodynamics.

The same holds true in the day-to-day medical practice, where patient-level diagnosis is much more efficient and repeatable than molecular-level analysis of pathological conditions.

In statistical methodology we are used to define the most convenient scale for modeling as the level that maximizes 'non-trivial determinism' (Rand DA, Wilson HB 1995), and this layer, in the case of biological systems in which different scales of space and time interact, only very rarely is the most fundamental one. The maximization of non-trivial determinism has to do with the so called Akaike criterion of the maximization of information coming from an experiment (Stone M. 1977). I'll try and explain such criterion by the following example: any protein is made of the same 20 aminoacid species, this is a very strong deterministic statement but is trivial for modeling since it does not provide any possibility to discriminate different protein functions and structures, or the effect of mutations. On the other hand, considering protein structures as adjacency graphs in which the nodes are the aminoacid residues and the edges the between residues contacts in 3D structure, is a very convenient choice (Bagler G, Somdatta S 2007), given it allows for both a universal and consistent analysis of any protein structure and gives precious indications about specific functional roles at both entire protein and single residue levels (Del Sol et al. 2006).

We are now collecting many indications on the 'most promising level' where to focus our modeling lens for understanding principles of biological regulation and this is not the single cell but the tissue (or in any case organized population of cells) level. Many results point to a largely stochastic behavior of different metabolites (proteins, RNA species) at the single cell level (Elowitz M et al. 2002) coupled with a strongly repeatable and largely invariant metabolic, protein and transcription profile at the tissue level (Giuliani A 2010). Two cell coltures coming from the same kind of tissue of two different individuals share a near to unity correlation coefficient in RNA species abundances, spanning around 30000 gene products each

potentially ranging over four order of magnitude of abundance (Giuliani A 2010). This allowed to sketch very interesting attractor-like dynamics of development and differentiation allowing for a drastic simplification of modeling efforts (Wang et al. 2010, Huang 2009, Chang et al. 2008). As a matter of fact only 200 tissues are present in nature out of the transfinite number of the possible combinations of activation levels of 30000 genes. This tells us of a still unknown 'epigenetic-energetic' landscape driving the feasible solutions we can observe in nature so calling for a 'statistical mechanics – like' modeling effort, more than a molecular *ab-initio* one .

In this realm, contrary to the prediction of the authors, the study of bacteria can still give us very important insights, especially considering large scale organization phenomena like biofilms, huge structures made by the coalescence of billions of bacterial cells each one performing a different role depending on its location in the biofilm (Tolker-Nielsen T, Molin S 2000). This is a very effective (even if primitive) form of cell –cell interaction that can give us very important clues about tissue organization and evolution.

All in all molecular studies are important but by no means they can be considered as the only avenue for a quantitative modeling of biological regulation to be approached.

Alessandro Giuliani

Dept. Ambiente e connessa Prevenzione Primaria
Istituto Superiore di Sanità, Roma, Italy.
Alessandro.giuliani@iss.it

References

Bagler G, Somdatta S 2007. Assortative mixing in Protein Contact Networks and protein folding kinetics. Bioinformatics **23**: 1760-1767.

Chang H, Hemberg M, Barahona M, Ingber DE, Huang S **2008** Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature* **453**: 544-548.

Del Sol A, Fujihashi H, Amoros D, Nussinov R **2006**. Residues crucial for maintaining short paths in network communication mediate signaling in proteins. *Molecular Systems Biology*. Doi: 10-1038/msb4100063.

Elowitz M, Levine A, Siggia ED, Swain P **2002**. Stochastic Gene Expression in a Single Cell. *Science* **297**: 1183-1186.

Giuliani A **2010** Collective motions and specific effectors: a statistical mechanics perspective on biological regulation. *BMC Genomics* **11 (suppl. 1)**: S2.

Huang S **2009** Reprogramming Cell fates: Reconciling Rarity with Robustness. *BioEssays* **31**: 546-560

Rand DA, Wilson HB **1995**. Using spatio-temporal chaos and intermediate-scale determinism to quantify spatially extended ecosystems. *Royal Soc. Proc. Biol. Sci*. **259**: 111-117.

Stone M. **1977**. An asymptotic equivalence of choice of model by cross-validation and Akaike's criterion. *J. of the Royal Statistical Soc. B* **191**: 211-223

Tolker-Nielsen T, Molin S **2000.** Spatial organization of microbial biofilm communities. *Microbial Ecology* **40**: 75-84.

Wang J, Xu L, Wang E, Huang S **2010.** The Potential Landscape of Genetic Circuits Imposes the Arrow of Time in Stem cell Differentiation. *Biophys. J.* **99**: 29-39.