Math. Struct. in Comp. Science (2014), vol. 24, iss. 3, e240307, 3 pages. © Cambridge University Press 2014
doi:10.1017/S0960129512000771

Toward a probabilistic biology

JEAN-JACQUES KUPIEC

Centre Cavaillés, ENS Paris,
29 rue dUlm, 75005 Paris, France
Email: jean-jacques.kupiec@ens.fr

Received 1 February 2011, revised 23 June 2012

Stochastic gene expression (SGE) is now considered to be an established fact and has become an important subject of research (Viñuelas *et al.* 2012). During the last decade, the availability of new techniques has made possible the production of more precise and more spectacular data showing the extensive variability in gene expression occurring between individual cells. However, evidence supporting probabilistic models of cell behaviour and gene expression has been available for quite a long time – for example, see the reviews in Laforge *et al.* (2005) and Golubev (2010). It should be noted that the first model of stochastic cell differentiation was proposed in 1964 (Till *et al.* 1964), which is almost at the same time as the genetic programming theory (Jacob and Monod 1961), which is deterministic in nature. One can thus wonder why the determinist view of biology has remained dominant for such a long time, and what makes a probabilistic view more acceptable nowadays? In this short paper, I will briefly argue that there is a strong epistemological obstacle to the acceptance of probabilism in biology, and that even today SGE is not integrated into a fully probabilistic approach of cellular processes, but rather into a concept that I call 'determinism with noise'. I will argue that a new fully probabilistic theoretical framework is needed to truly integrate the stochastic aspects of cell physiology.

1. The epistemological obstacle

Throughout its history, biology has been a theatre of intense debate. Issues concerning mechanism and vitalism, structure and function, and fixism and evolutionism are just a few examples. In contrast, the belief that physiological processes occurring inside the organism must be deterministic has never been questioned. From Aristotle's finalism in Antiquity, Descartes's *animal-machine* in the seventeenth century, Claude Bernard's *milieu intrieur* in the nineteenth century to Jacob and Monod's genetic programming theory in the twentieth century, with some very rare exceptions, all biological theories have been deterministic. Indeed, biology faces a particular difficulty that is less apparent in physics. When we deal with living beings, we also deal with Man, and, as 'human' beings, we have a strong reluctance to accept the idea that we could be governed by probabilistic processes since it seems to imply that life is meaningless. Darwin's theory of evolution has faced strong opposition because it is based on purposeless chance events, but when it comes to our own individual ontogenesis, the idea that it could be the result of probabilities and of disorder within us is even more unacceptable. Etymologically, the word organism has the same root as organisation and, indeed, the organism has always been thought of as a sort of kingdom of order, while chance events and organisation seem to be antagonistic.

https://doi.org/10.1017/S0960129512000771 Published online by Cambridge University Press

2. Determinism with noise

Since the emergence of molecular biology, biological order has been explained by genetic information. In the framework of molecular biology, an organism is, indeed, the expression of the information encoded in DNA, and the mechanism by which information is transformed into an actual being during ontogenesis is thought to be governed by the so-called stereospecificity of proteins. According to this theory, proteins are thought to interact in a specific one-to-one manner, like pieces of a puzzle, in order to avoid the possibility of multiple combinations between them, which could be a source of randomness. On this basis, specific networks of protein interactions (signalling, metabolic or structural) are meant to build and control cells. In the case of signal transduction and gene regulation, each interaction has been conceived as a cybernetic switch, and whole networks as cybernetic circuits or programs (Jacob and Monod 1961).

In order to explain the variability of gene expression, it has been proposed that disturbances caused by thermal agitation affect the way networks (programs) function. Indeed, for interactions to occur, proteins have to diffuse in the cell space in order to encounter their molecular partner. For this reason, there are always fluctuations in the operations of a cell network, and two cells are never exactly in the same state, even if they belong to the same population of cells contained in a homogeneous medium. This is also true for gene expression, which is controlled by transcriptional regulators interacting with the regulatory sequences of genes. Because of thermal fluctuations, one gene can be active in a cell while the same gene is not activated in another cell, explaining the observed expression variability of this gene. It is clear that this introduces stochastic fluctuations in cell behaviour, but does not challenge the idea of molecular interaction networks. The cell is still globally determined by an underlying molecular order corresponding to these networks.

3. Constrained randomness

Determinism with noise is, in fact, already an out-of-date concept. Protein networks have been precisely analysed, and it has been demonstrated that proteins are not specific in their interactions. One particular protein can always interact with a number of molecular partners, and at least 10% of an organism's network (several thousand proteins) corresponds to proteins that can interact with at least 100 molecular partners, leading to immense combinatorial possibilities (Kupiec 2009). Taking these data into account completely demolishes the idea that there can be an underlying molecular order controlling cell functions and implies that molecular interactions are intrinsically stochastic. For this reason, biology needs a new theoretical framework in place of the genetic programming theory, and I have proposed a chance-selection mechanism to explain cell behaviour (Kupiec 1983). In this model, selective constraints arising from the cell's environment (internal and external to the organisms) limit the intrinsic randomness of protein interactions, which allows for cell activity.

References

Golubev, A. (2010) Random discrete competing events vs. dynamic bistable switches in cell proliferation in differentiation. *Journal of Theoretical Biology* **267** (3) 341–54.

https://doi.org/10.1017/S0960129512000771 Published online by Cambridge University Press

Jacob, F. and Monod, J. (1961) Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology* **3** 318–356.

Kupiec, J.-J. (1983) A probabilist theory for cell differentiation, embryonic mortality and DNA C-value paradox. *Speculation in Science and Technology* **6** 471–478.

Kupiec, J.-J. (2009) *The Origin of Individuals*, World Scientific.

Laforge, B., Guez, D., Martinez, M. and Kupiec, J.-J. (2005) Modeling embryogenesis and cancer: an approach based on an equilibrium between the autostabilization of stochastic gene expression and the interdependence of cells for proliferation. *Progress in Biophysics and Molecular Biology* **89** 93–120.

Till, J. E., McCulloch, E. A. and Siminovitch, L. (1964) A stochastic model of stem cell proliferation, based on the growth of spleen colony-forming cells. *Proceedings of the National Academy of Sciences* **51** 29–36.

Viñuelas, J., Kaneko, G., Coulon, A., Beslon, G. and Gandrillon, O. (2012) Towards experimental manipulation of stochasticity in gene expression. *Progress in Biophysics and Molecular Biology* **110** 44–53.

https://doi.org/10.1017/S0960129512000771 Published online by Cambridge University Press