Is biological randomness a statistical physics concept?

Ovidiu Radulescu,[1] Adrian Devenyi[2]

One indication of the relevance of stochastic laws in biology is the fact that genetically similar cells can behave in very different ways. This suggests an analogy with statistical physics, in which stochastic laws govern the behavior of large ensembles of particles. Biological randomness has important theoretical and practical consequences in medicine and may be responsible for the resistance of pathogens or cancer to drug treatment. Modeling biological randomness is now a major field in theoretical biology and has been approached with methods adapted from physics. Like in physics, biological systems are governed by hierarchical processes involving variables with different timescales. Contrary to most physical processes, biological systems can exhibit inversions in the relationship between timescale and hierarchical rank. As a consequence, microscopic fluctuations can be transmitted to and even amplified by the phenotype. Because of these peculiarities, new approaches are needed for the study of biological randomness.

Randomness in physics and biology

Randomness, as opposed to determinism, means that the result of an experiment is not unique, but belongs to a set of different outcomes. The outcomes are not necessarily equivalent and one can define a distribution associating a probability to each outcome. Randomness does not imply absence of laws. On the contrary, mathematical models can compute probability distributions and make statistical predictions. Randomness does not exclude causality, that can be assessed by conditional probabilities. Randomness and statistical laws are well established concepts in physics, economics and social sciences. They prevail also in biology, though in a subtly different way. In this essay we compare random phenomena in physics and biology.

Randomness in Physics: Statistical Physics vs. Quantum Mechanics

The idea that matter is built from constituent parts (atoms) dates from antiquity (Leucippus, Democritus, Epicurus). Reconciling the atomistic hypothesis with the macroscopic bulk properties had been a challenge that was settled by the kinetic theory of gases, owing to the contributions of Clausius, Boltzmann, Maxwell, Gibbs and Einstein, among others. The atoms in this theory obeyed the laws of Newtonian mechanics. If we knew the exact positions and velocities of all the atoms in a container, we could calculate exactly both the future and past behaviors of the system. Given that a macroscopic quantity of gas contains on the order of 10^{23} atoms, it is practically impossible to exactly calculate the dynamics of such a system. From a practical point of view, the motions of the individual atoms are random. This type of randomness is called apparent and it stems from our imperfect knowledge of the system.[3] However, an approximation was found by averaging over the random microscopic motion of the constituent particles. The Central Limit Theorem (CLT) ensured that for a sufficiently large ensemble of particles, exact knowledge of the microscopic details of the particles' motion was unnecessary.

By contrast, the development of Quantum Mechanics in the first half of the 20th century, brought about a different kind of randomness in Physics: intrinsic. For a quantum system, even having exact knowledge about the initial state only allows us to predict the dynamics in terms of probabilities. The phil-

osophical consequences of this intrinsic randomness are still subject to debate, but as a physical theory, quantum mechanics is well established and completely agrees with the experimental evidence.

Variation and stochastic fluctuations in biology

Variation is fundamental in biology. Darwin describes his theory as "descent with modification", suggesting that organisms are submitted to variations and that these variations are transmitted to offspring. This heritable variation has been related to changes of the DNA sequence, generically named mutations. Mutations are subject to the effect of natural selection that keeps or eliminates them according to their effect on reproductive success (fitness) in a given environment. An essential feature of Darwin's thinking is that mutations do not have a purpose, but rather randomly exhaust all the possibilities. Although genetic variability is well accepted in biology, the idea that variability and random changes can occur somewhere else than in genes was only recently accepted. In fact, genes do not lead immediately and unconditionally to protein production. The protein synthesis is a multi-step process, in which all of the steps are dependent on the presence or absence of regulatory proteins (activators or repressors).[4] Furthermore, regulatory proteins can access the genes only if DNA chromatin unfolds which – similarly to polymer dynamics in statistical physics –, depends on multiple stochastic events.[5] These changes, called "epigenetic" to distinguish them from mutations consisting in the modification of the DNA, are triggered by a small number of molecules or supramolecular complexes which are submitted to "atomistic" randomness of their individual motion. It was thus experimentally shown that, at the scale of the individual cell, protein production is intermittent, combining quick and slow random steps.[6] The randomness of the protein levels leads to random choice of the phenotype and interaction among genes can render this choice bi- or multi-modal. For instance, the compound eye of the fruit fly is made of multiple units named ommatidia; during the development independent random choices are made by photoreceptor cells of ommatidia to become one of two possible cell types.[7] Stem cells, undifferentiated cells capable of generating part or all the cell types of a multicellular organism, are also able to take random bimodal decisions.[8]

Why is randomness in biology different from physics?

Fluctuations in biology result from small numbers

Biological systems are complex systems with numerous actors and internal degrees of freedom. Like atoms in a statistical physics system, the ensemble is submitted to random internal interaction and behaves stochastically. In physics, the central limit theorem CLT guarantees that microscopic fluctuations are averaged out and have limited impact at the macroscopic level. The applicability of similar ideas to biology was first questioned by Erwin Schrödinger, who noticed that biology deals essentially with small numbers N of particles.[9] Direct application of CLT implies that relative noise scales with $N^{-1/2}$, which means that fluctuations are relatively larger at smaller N. As a matter of fact, biology needs error-correction systems to decrease noise to the $N^{-1/2}$ limit and below.[10] For biological systems, N is in the range 1-10,000, much smaller than for physical systems, where N is on the order of 10^{23}. In the absence of error-correction, fluctuations show up at many scales and exceed the CLT law. This property can have profound consequences: the defect of expression of a single or of a few genes can dramatically change the phenotype and the uncorrected effect of mutations can lead to diseases such as cancer.

Heterogeneity of biological systems and failure of limit theorems

In physics, a general theory called "hydrodynamic fluctuations" was proposed for explaining macroscopic fluctuations of classical and quantum mechanical systems,

independently of the microscopic details of the system.[11] Recently, the hydrodynamic theory was successfully extended to define the behavior of "active matter" used as a paradigm for many biological systems: bird flocks, fish schools, bacterial films, muscle filaments, etc. A characteristic of active matter is the consumption and dissipation of energy at all times. This fundamental fact leads to intriguing properties such as the possibility of giant macroscopic fluctuations.[12]

However, the applicability of hydrodynamic theory requires a certain amount of homogeneity in the system. There are numerous examples in biology where this is not possible, the best example being the heterogeneity of cancer. Tumor cell populations are strongly heterogeneous, meaning that cells are distributed among genetically and/or epigenetically different states.[13] Under application of drugs, the dynamic equilibrium is broken and cells in resistant, previously less competitive states can develop, which explains why heterogeneity leads to drug resistance. The same principle explains the development of resistance to antibiotics in bacteria populations.

One explanation for the strong heterogeneity of these biological systems is what we call the "inverted hierarchy of timescales": in biology, contrary to physics, the microscopic fluctuations are at least as slow as the macroscopic ones.[14] Indeed, atoms or molecules of non-biological matter have simple shapes and fast dynamics, whereas in biology sophisticated molecular machines work slowly at molecular level. Protein synthesis has several stages, some of them being very slow.[15] The fluctuations generated by the slow molecular processes are not averaged out and generate wide distributions of cell states. These microscopic distributions have their own slow dynamics that impact the macroscopic dynamics of a tissue, organ or organism.

Conclusion

Given the failure of traditional physics approaches to explain all the aspects of biological randomness, a new synthesis is needed to bring these phenomena together. The new theory must take into account the impossibility of separating the organization scales, the wide distribution of cell properties and the slow dynamics of these distributions. A number of multi-scale mathematical models exist already, based on coupling spatio-temporal and structural dynamics. In these models, cells are distributed in space, but also in a "structure" space of intrinsic properties.[16] Alternatively, hybrid stochastic models use discrete stochastic modeling at the scale of molecular slow processes and deterministic continuous modeling at phenotype (macroscopic) level.[17] These models are effective in capturing properties of specific systems, but are still unable to reveal universal laws such as those of statistical physics. It may be that in biological systems, particularities are more determinant than in physics, and therefore less amenable to a general model. However, physicists' definition and understanding of life is continuously evolving and the last 20 years have brought new theories such as hydrodynamics of active matter, theory of stochastic biochemical networks, various types of multiscale modeling. Our understanding of biological heterogeneity and drug resistance is improving, and hopefully will soon find application in targeted and personalized therapies.

[1] LPHI, UMR CNRS 5235 – University of Montpellier, Montpellier, France.
[2] Jermain Capital Management, NY, USA.

[3] M. N. BERA, A. ACIN, M. KUS *et al.*, "Randomness in quantum mechanics: philosophy, physics and technology", *Rep. Prog. Phys.* 80(12), 2017, 124001.
[4] F. JACOB, J. MONOD, "On the Regulation of Gene Activity", *Cold Spring Harbor Symposia on Quantitative Biology* 26, 1961, 193-211.
[5] T. CREMER, C. CREMER, "Chromosome territories, nuclear architecture and gene regulation in mammalian cells", *Nat. rev. gen.* 2(4), 2001, 292-301.
[6] E. M. OZBUDAK, M. THATTAI, I. KURTSER *et al.*, "Regulation of noise in the expression of a single gene", *Nature Genetics* 31, 2002, 69-73; L. CAI, N. FRIEDMAN, X. S. XIE, "Stochastic protein expression in individual cells at the single molecule level", *Nature* 440(7082), 2006, 358-362; A. RAJ, C. PESKIN, D. TRANCHINA *et al.*, "Stochastic mRNA Synthesis in Mammalian Cells", *PLoS biology* 4(10), 2006, e309.
[7] A. HODGKINSON, G. UZ´E, O. RADULESCU *et al.*,

"Signal propagation in sensing and reciprocating cellular systems with spatial and structural heterogeneity", *Bull. Math. Biol.* 80 (7), 2018, 1900-1936.

[8] J. TILL, E. MCCULLOCH, L. SIMINOVITCH, "A stochastic model of stem cell proliferation, based on the growth of spleen colony-forming cells", *PNAS* 51(1), 1964, 29-36.

[9] E. SCHRÖDINGER, *What is Life?: With Mind and Matter and Autobiographical Sketches*, Cambridge, Cambridge University Press, 1992.

[10] A. N. GORBAN, O. RADULESCU, "Dynamical robustness of biological networks with hierarchical distribution of time scales", *IET Systems Biology* 1(4), 2007, 238-246.

[11] D. FORSTER, *Hydrodynamic fluctuations, broken symmetry, and correlation functions*, Boca Raton, CRC Press, 2018.

[12] M. C. MARCHETTi, J-F. JOANNY, S. RAMASWAMY *et al.*, "Hydrodynamics of soft active matter", *Rev Mod Phys*, 85(3), 2013, 1143.

[13] A. MARUSYK, V. ALMENDRO, K. POLYAK, "Intra-tumour heterogeneity: a looking glass for cancer?", *Nature Rev. Cancer* 12(5), 2012, 323-334.

[14] O. RADULESCU, G. INNOCENTINI, J. HORNOS, "Relating network rigidity, time scale hierarchies, and expression noise in gene networks", *Phys. Rev. E* 85(4), 2012, 041919.

[15] J. M. RASER, E. K. O'SHEA, "Control of stochasticity in eukaryotic gene expression", *Science* 304(5678), 2004, 1811-1814.

[16] A. HODGKINSON, G. UZ´E, O. RADULESCU *et al.*, "Signal propagation in sensing and reciprocating cellular systems with spatial and structural heterogeneity", *op. cit.*

[17] G. INNOCENTINI, A. HODGKINSON, O. RADULESCU, "Time Dependent Stochastic mRNA and Protein Synthesis in Piecewise-deterministic Models of Gene Networks", *Frontiers in Physics* 6, 2018, 46.