Theory in Biology

Theory is available light

Martin A. Nowak

Many people praise Gregor Mendel as the founding father of genetics. Few, however, know that as a student at the University of Vienna he excelled in Mathematics, but failed Botany twice, which prevented him from pursuing an academic career. He entered a monastery and continued to do science. He found the rules of genetic inheritance because he had a clear mathematical hypothesis even to the extent of ignoring ambiguous results that did not fit his mathematical expectation.

Ronald Fisher considered Mendel as a mathematician with an interest in biology. Fisher, together with J.B.S. Haldane and Sewall Wright, unified Mendelian genetics with Darwinian evolution by formulating a precise mathematical description of evolutionary dynamics. This work was done in the 1930s and is the very foundation of our current understanding of evolution, without which, as we know, nothing makes sense in biology.

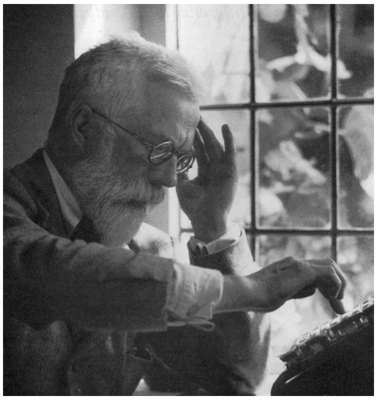
It is no coincidence that Gregor Mendel, the founder of genetics, trained as a mathematician. A portrait painted after Mendel's death, preserved in the Augustinian Abbey of St Thomas in Brno and printed in *Gregor Mendel: the First Geneticist* (1996), Oxford University Press.

In the 1960s, the Japanese mathematical biologist Motoo Kimura conceived the neutral theory of evolution, which assumes that the overwhelming majority of molecular mutations do not affect the fitness of an individual. Today most methods for reconstructing phylogenies from genetic data assume the neutral theory is correct. In a brilliant PhD thesis submitted in 1964, Bill Hamilton formulated the theory behind 'selfish genes'. In the 1970s, John Maynard Smith brought game theory to biology: evolutionary game theory is a generic description of evolutionary dynamics whenever fitness is not constant, but depends on the composition of the population. At about the same time, Manfred Eigen and Peter Schuster pioneered an information theoretic approach to evolution with quasi-species theory, while Josef Hofbauer and Karl Sigmund studied rigorously the replicator equation, one of the most important descriptions of evolutionary dynamics.

Contributions such as these — and many others — have made it clear that any discourse of evolutionary dynamics ultimately requires exact mathematical formulations. All evolutionary biologists I have ever met appreciate mathematics and understand its role in scientific research.

Evolution does not stand alone as a biological discipline that would be unthinkable without theory. The work of Robert May, for example, has extended earlier studies in ecology and epidemiology to ensure that these subjects are also grounded in precise mathematical formalism. Clearly, the spread of epidemics or the dynamic interaction of species in an ecosystem cannot be understood without the help of mathematical methods. Notably, the most cited ecologist of our time, David Tilman, is a theoretician.

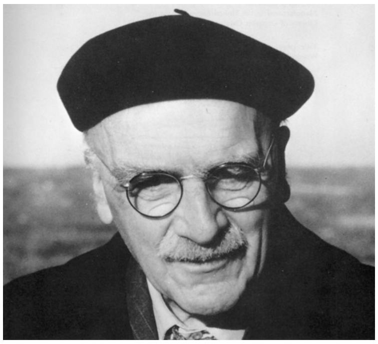
In immunology, molecular and cellular experiments have generated fascinating insights, but nobody can argue about the population dynamics of infectious agents and immune cells without the aid of mathematical models. Crucial questions, such as what determines the virus load in HIV infection, how fast is the turnover rate of HIV-infected cells, or what explains the long and variable asymptomatic period between HIV infection and development of AIDS, cannot be answered by a purely verbal analysis of experimental observations.



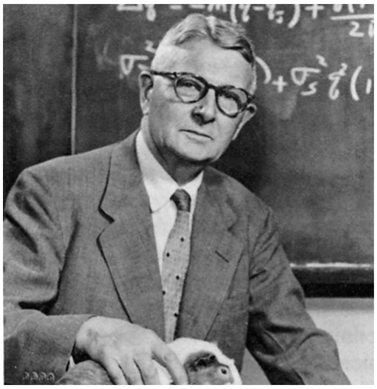
Sir Ronald Fisher, with Haldane and Wright one of the founders of population genetics. (Courtesy of Antony Barrington-Brown and the Science Photo Library.)

It is sometimes stated that biology is too complicated for mathematical investigation, but this argument is more logical if reversed: surely only simple mechanisms can be understood without the use of mathematical analysis.

In cancer genetics, one of the most crucial questions concerns the role of genetic instability in tumor progression. Mutations in oncogenes or tumor suppressor



J.B.S. Haldane ended his book '*The causes of evolution*' (1932) with the words "The permeation of biology by mathematics is only beginning". (Courtesy of Klaus Patau.)



Sewell Wright studied genetic variation in subdivided populations and invented the ‘fitness landscape’. (Courtesy of The University of Chicago Press.)

genes enhance the net growth rate of the cell, whereas mutations in genetic instability genes increase its mutation rate. Chromosomal instability refers to the phenomenon that most cancer cells have increased rates of gaining or losing whole chromosomes. Do mutations that trigger genetic instability accelerate cancer progression? How long does it take to activate oncogenes or inactivate tumor suppressor genes in populations of dividing cells with and without genetic instability? Both questions require mathematical analysis.

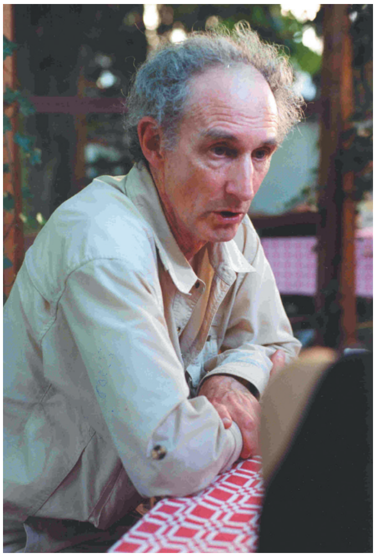
The very concept of a tumor suppressor gene was first proposed by Al Knudson in 1971, based on a mathematical analysis of incidence patterns of retinoblastoma in children. Much earlier, at a time before the nature of DNA mutation had been revealed, Richard Doll used mathematics and statistics to prove that smoking causes cancer.

The Hodgkin–Huxley equation is central to neuro-physiology and has inspired much experimental and theoretical work. Moreover, cognitive functions of the brain cannot be understood quantitatively without the formalism of computer science, learning theory or neural networks. Human language, for example, was studied for thousands of years, but the tremendous progress of the last 50 years coincides with the insistence on mathematical rigor.

Understanding biophysical principles of molecular motors requires the mathematics of chemical kinetics and thermodynamics. Protein folding and predictions of the secondary or tertiary structures of RNA is another vast and exciting field which demonstrates the fruitful interaction of mathematics and biology.

Although mathematical biology has a long and impressive history, as I have argued here, the present time is particularly exciting. Advances in cellular and molecular biology, mostly in the context of the human genome project, have yielded extensive amounts of information that not only require mathematical and computational analysis, but represent invaluable resources for testing new theories. In genomics, mathematical techniques are needed to align sequences, search for genes or characterize the interactions of genetic control networks. Notably, Eric Lander, one of the leaders of the genomic revolution, came top of his class at Princeton University in mathematics. Evolutionary biology can ask questions concerning the organization of whole genomes, the selection pressure acting on individual genes, the evolution of chromosomes or the relationship between the complexity of organisms, the size of the genome and the number of genes. The data for evaluating exciting new theories may already exist. Somebody only has to ask the right questions.

Students feel this excitement. Harvard undergraduates in mathematics, physics or computer science, for example, are keen to participate in biological research projects. Similarly, biology students now feel the need to take advanced classes in mathematics. The choice of the students will define the direction of biology.



Lord May of Oxford, President of the Royal Society, brought a clear mathematical line into ecology and epidemiology.

All these examples show that biology is already permeated by essential theory in every corner. This is not surprising, because mathematics is the most powerful language of science that we have. Mathematical models provide a logical link between assumption and conclusion. Verbal interpretations do not have the same amount of rigor. Insisting on purely non-mathematical interpretation of empirical observations limits the range of questions that can be asked and conclusions that can be reached.

For any scientific investigation, theory is available light.

Background reading

Anderson, R.M. and May, R.M. (1991). Infectious Diseases of Humans, Oxford Univ Press.
Fisher, R.A. (1930). The Genetical Theory of Natural Selection, Oxford Univ Press.
Ewens, W.J. (2004). Mathematical Population Genetics, Springer.
Gillespie, J.H. (1991). The causes of molecular evolution, Oxford Univ Press.
Hamilton, W. D. (1996). Narrow roads of gene land. NY: Freeman.
Hofbauer, J. and Sigmund, K. (1988). Evolutionary Games and Replicator Dynamics, Cambridge University Press
Kimura, M (1983). The neutral theory of molecular evolution, Cambridge University Press.
May, R. M. (1972). Stability and complexity of model ecosystems, Princeton University Press.
Maynard Smith, J. (1982). Evolution and the theory of games. Cambridge University Press.
Nowak, M. A. and May, R. M. (2000). Virus Dynamics, Oxford University Press.

Program for Evolutionary Dynamics,
Department of Organismic and
Evolutionary Biology, Department of
Mathematics, Harvard University,
Cambridge, Massachusetts 02138, USA.
E-mail: martin_nowak@harvard.edu