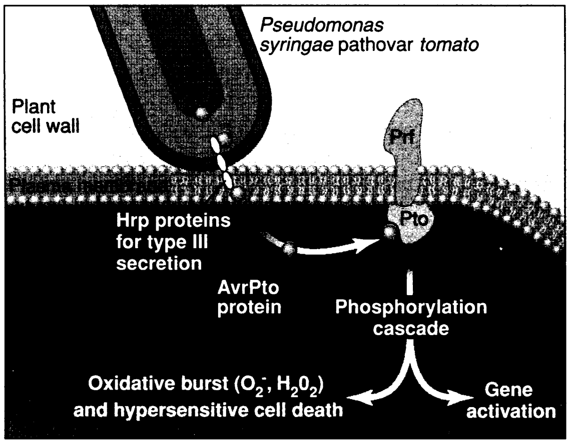

**Molecular specificity.** A direct interaction between products of a plant resistance gene (AvrPto) and an avirulence gene (Pto) in a pathogen accounts for the gene-for-gene specificity between plant and pest.

resembling components of the type III secretion system used by bacteria—such as *Salmonella*, *Shigella*, and *Yersinia*—for the secretion of virulence proteins into mammalian host cells (*5*, *6*). Operation of such a mechanism (see figure) would allow *avr* gene products to interact with cytoplasmic plant proteins, and in fact *avrB* from *P. syringae glycinea* induces an *R* gene–dependent resistance response when expressed as a transgene in plant cells (*7*).

Scofield *et al.* (*2*) and Tang *et al.* (*3*) report a similar effect when *avrPto* from *P. syringae tomato* is expressed in plant cells carrying *Pto*, and then take the story a decisive step further by demonstrating a direct physical interaction between the gene products with yeast two-hybrid genetic selection as an assay system (*2*, *3*). A second protein kinase, Fen, with greater than 80% sequence similarity to Pto, does not interact with the AvrPto protein. *Fen*, which is tightly linked to *Pto*, confers sensitivity to the insecticide fenthion but does not mediate resistance in response to *avrPto*. Creation of chimeric Pto-Fen proteins shows that the Fen gene product can be made competent to interact with the AvrPto ligand by substitution of a small segment of Pto involved in substrate binding, and mutations in either *avrPto* or *Pto* that render the plant-pathogen interaction biologically compatible likewise disrupt the physical interaction between the gene products. Thus, the exquisite specificity that characterizes gene-for-gene biological incompatibility in plant-pathogen interactions can also be discerned at the molecular level.

Protein kinase activity is required for the function of Pto in disease resistance, and mutations at conserved catalytic residues also block the physical interaction with the AvrPto protein. AvrPto binding might directly stimulate Pto kinase activity to trigger the phosphorylation cascade involved in activating the resistance response (see figure) (*8*). Alternatively, AvrPto might bring together two Pto molecules for cross phosphorylation. Pathway activation may also cause phosphorylation of the ligand itself. Genetic evidence indicates that *Pto* action requires a closely linked gene, *Prf*, which encodes a leucine-rich repeat protein (*9*). Although Prf does not appear to be involved in specificity, it may function as an anchor to localize the kinase (see figure). Physical interaction between Prf and Pto might create a two-component receptor system closely resembling the receptor-like protein kinase encoded by *Xa21*. It will be interesting to see which region, if any, of the *Xa21* product interacts with the corresponding AvrXa21 protein.

The isolation of bacterial genes is highly tractable, with over 30 cloned so far, and hence the use of two-hybrid selection may prove to be a powerful approach for the isolation of new *R* genes, assuming at least some function as receptors for the corresponding *avr* gene product. Two-hybrid screening should also give a quick test of the ligand-receptor mechanism in the recognition of viral and fungal pathogens for which *R* genes have been isolated.

AvrPto and Pto are both relatively small, compact proteins and hence should be amenable to structural analysis by x-ray crystallography. It would be immensely satisfying if the series of brilliant studies starting with Flor's work 50 years ago culminate in the elucidation of plant-pathogen recognition systems at the angstrom level. Such information would allow the rational design of *R* genes encoding receptors with novel recognitional specificities—powerful tools for engineering enhanced crop protection in the continuing struggle to secure food production.

## References

1. B. J. Staskawicz, F. M. Ausubel, B. J. Baker, J. G. Ellis, J. D. G. Jones, *Science* **268**, 661 (1995).
2. S. R. Scofield *et al.*, *ibid.* **274**, 2063 (1996).
3. X. Tang *et al.*, *ibid.*, p. 2060.
4. W.-Y. Song *et al.*, *ibid.* **270**, 1804 (1995).
5. H.-C. Huang, S. Y. He, D. W. Bauer, A. Collmer, *J. Bacteriol.* **174**, 6878 (1992).
6. M. Barinaga, *Science* **272**, 1261 (1996).
7. S. Gopalan *et al.*, *Plant Cell* **8**, 1095 (1996).
8. J. Zhou, Y.-T. Loh, R. A. Bressan, G. B. Martin, *Cell* **83**, 925 (1995).
9. J. M. Salmeron *et al.*, *ibid.* **86**, 123 (1996).

---

THEORETICAL BIOLOGY

# Biologists Put on Mathematical Glasses

Torbjörn Fagerström, Peter Jagers, Peter Schuster, Eörs Szathmary

Nothing has shaped the development of thought in physics more than Galileo's statement: "The book of Nature is written in the language of mathematics" (*1*). Physicists take it for granted that the important questions have answers that can be cast into mathematical formulas. Ever since Newton, physics and mathematics have lived in a fruitful symbiosis, with a great deal of cross-fertilization to the benefit of both disciplines. Physics, as we phrase it nowadays, is concerned mainly with the development of a comprehensive and unifying mathematical theory of Nature. The physicists are approaching this goal by experimentation, abstraction, and generalization.

Biology differs from physics in that it has an indispensable historical component. This was stressed already by Ernst Haeckel and most properly phrased by Theodosius Dobzhansky in his statement that "nothing in biology makes sense except in the light of evolution" (*2*). For this and other reasons, much of biology has traditionally described the overwhelming diversity and unique variability of the living world and has used a scientific methodology based on observation, description, and classification. Although generalization and abstraction, where feasible, was always aimed for, mathematics was usually not used in this process. For example, last century's greatest naturalist, Charles Darwin, laid down his great theory of evolution and the origin of species without making use of a single equation.

---

T. Fagerström is in the Department of Theoretical Ecology, S-223 62 Lund, Sweden. E-mail: torbjorn.fagerstrom@teorekol.lu.se. P. Jagers is in the Department of Mathematics, Chalmers and Gothenburg Universities, S-412 96 Gothenburg, Sweden. E-mail: jagers@math.chalmers.se. P. Schuster is at the Institut fur Theoretische Chemie und Strahlenchemie, Universität Wien, A-1090 Wien, Austria. E-mail: peter.schuster@tbi.univie.ac.at. E. Szathmary is at the Collegium Budapest, 1014 Budapest, Hungary. E-mail: szathmary@colbud.hu

This century has witnessed an increased use of mathematics in many fields of biology, prominent examples being cell kinetics, population ecology, epidemiology, and population genetics. In the case of population genetics, the formal approach was introduced already last century by Gregor Mendel, who applied his knowledge of probability theory to the problem of inheritance, and the mathematics of heredity was then taken up and successfully developed by Fisher, Haldane, and Wright. Indeed, the neo-Darwinian synthesis that emerged by the middle of this century owes much of its success to the abstractions, generalizations, and formalizations of theoretical population ecology and theoretical population genetics, and it may well be considered to be the first case of a reasonably mature, proper theoretical biology.

What, then, is theoretical biology about? Can problems be identified in biology that would benefit from current mathematical approaches? Is there an obvious need for the development of new mathematics? These questions were addressed at a meeting in Sweden organized recently by the European Science Foundation.

As to the first question, a mere heap of mathematical models in biology does not constitute theoretical biology. It is rather a discipline, aiming at a coherent body of concepts and a family of models, in which passage from one concept to another, or from one model to another, must follow a regularized pathway. Adding to theoretical biology can occasionally mean little or no mathematical work, at least in terms of "numbers" and "solutions," but the conceptual contribution must then be significant. Waddington's epigenetic landscape is such an example. But these concepts must ultimately turn out to be mathematizable. Mathematical biology, on the other hand, deals with questions where a solid conceptual framework already exists: it constitutes an extension, refinement, and elaboration of established, simple models. It also incorporates the rigorous analysis of mathematical structures applied in biology.

Mathematics inspired by physics is largely based on symmetry considerations, but symmetry never played the same fundamental role in biology as it does in physics: there are no quantum numbers of life. Nevertheless, theoretical biology has so far largely borrowed its methods from physics; in particular, the use of differential equations—this brilliant product of the Newtonian worldview—and the techniques for analyzing their dynamical properties have dominated much of theoretical biology. One must acknowledge, however, that biology is only partly Newtonian. It is individuals who make up populations, at all levels from molecules and tumors to whole species, and we must not forget that variation occurs at all these levels in the most conspicuous manner. It is not without reason that Boltzmann remarked that Darwin's work was an intuitive "statistical mechanics of populations."

Modern probability theory offers possibilities to describe individual behavior, even in situations that exhibit much individual variation, and even when these variations do not follow the standard distributions of elementary statistics. From these individual properties, characteristics of the whole, like rates of growth, evolution, or extinction, might then be deduced. A more rigorous application of probability theory to biology, in terms of individuals with varying behavior, rather than in the streams or fluxes of classical physics, should be close to biological thinking. This may even inspire developments in mathematics and eventually turn out to be relevant for modern physics, which long ago left the tradition of thinking in terms of continua in favor of systems of interacting particles, as discrete as are the individuals of biological populations.

In addition, biology should also stimulate the development of new branches of mathematics, tailored to the specific needs of theoretical biology. Most urgent is the current explosion of data produced in the various branches of biology. Leading the pack in this regard are genome sequencing and the molecular genetics of development. Similarly, as taxonomists and ecologists gather more information on Earth's biodiversity, databases that are already large will continue to grow. This data explosion requires theory to be advanced because "no new principle will declare itself from below a heap of facts," as Sir Peter Medawar stated precisely. To this field of urgent need for theory one can add a list of unique biological phenomena, like reproduction, selection, adaptation, symbiosis and "arms races," that await to be cast in a rigorous theoretical framework. Advances by mathematicians on these topics should be welcomed, but in a critical spirit, by the theoretical biology community.

In trying to meet the increasing need for conceptualization, formalization, and abstraction, biology should borrow some of the virtues of physics. At the same time biologists must neither deny nor forget their heritage, nor fall into the traps of shortlived fads. Several latecomers among the concepts of theoretical physics are key concepts in biology as well. This applies to network dynamics, which is relevant to the analysis of metabolism or the immune system; to self-organization, which is what developmental biology is partly about; and to the emergence of new properties from synergy of interactions, which is presumably why multicellularity arose eons ago.

All these topics were discussed at the recent meeting in Sweden, together with an outlook to the open questions of biology. Why and how do information and complexity increase in evolution? What causes Nature to build more complex things in a hierarchical manner with principles that are recurrent at all levels? Genes, for example, are integrated into genomes, cells into multicellular organisms, and individuals into societies. Recurrence of the same principle gives rise to ever higher forms of complex life in an apparently open-ended evolutionary process. "Organization"—the operation of control systems in specific canalizing structures—has been an integrative concept in many areas of biology, but theoretical biology has so far had limited success in describing it in formal terms. Related to this problem is the ongoing occupation of evolutionary biologists with equilibrium situations and microevolution: usually, the applied dynamic (if explicit at all) is a closed one. Major transitions in evolution—such as the origin of life, the emergence of eukaryotic cells, and the origin of the human capacity for language, to name but a few—could not be farther away from an equilibrium. Also, they cannot be described satisfactorily by established models of microevolution. What is needed is an open-ended model, in which evolutionary novelties (or, rather, representations thereof) can continue to arise indefinitely. This model must be related to a currently nonexisting theory of variation, which in turn must be related to the theory of organization of objects, the evolution of which we would like to describe. A theory of development (ontogenesis) is still missing.

The origin of language is an even more formidable problem. Attempts to solve it must be based on ingredients in the theory of evolution, neurobiology, and formal linguistics. Because important insights have been mathematized in all these fields already, their joint application to this problem will have a strong mathematical element as well. But how much of this can be achieved by the standard methods is an open question.

Our ultimate goal must be a unifying theory of biology emerging from the forthcoming synthesis of three great disciplines: molecular, developmental, and evolutionary biology. Such a concept will provide the theoretical basis for a biology of the future with its own tools and methods, some coming from mathematics, some from computer science, and others, perhaps, from somewhere else. We suspect that an enormously exciting period lies ahead.

### References

1. S. Drake, *Discoveries and Opinions of Galileo* (Doubleday-Anchor, New York, 1957).
2. T. Dobzhansky, *Am. Biol. Teach.* **35**, 125 (1973).