Mathematical Biology is Good for Mathematics

Michael C. Reed

About ten years ago I wrote an article, "Why is Mathematical Biology so Hard?" for these *Notices* intending to explain why the applications of mathematics to biology would be very different than the traditional applications to physics and engineering [42]. A lot has happened since then. Mathematical biology has grown from a small field, containing relatively few mathematicians, to a major branch of applied mathematics. The reasons for this growth, which are implicit in the discussion below, are not the point of this article, nor do I want to encourage mathematicians to switch to mathematical biology. Rather, I want to make the case that mathematical biology benefits all mathematicians; it is good for the health of mathematics as a whole.

Most of Science is Biology

That statement seems absurd, almost laughable to many mathematicians who are used to thinking that "science" means physics and chemistry, while biology is just classification, necessary perhaps for training doctors, but not really deep, intellectual, or mathematical. We are in the midst of a biological revolution whose roots lie in the 19th and first half of the twentieth century. In the past twenty-five years the pace of this revolution has accelerated and it has created an enormous biological research community. The American Society for Nephrology has a membership that is comparable in size to the American Mathematical Society, and that's just the kidney. The annual meeting of the Society for Neuroscience attracts around 30,000 attendees,

which is again the approximate number of members of the AMS, and is huge when compared to the size of the Joint Mathematics Meetings (around 7,000). This growth has affected the balance of university research. For example, between September 2012 and May 2013, Duke University awarded 365 PhDs. Of these, 165 (45 percent) were in disciplines of the biological sciences like biochemistry, pharmacology, neurobiology, environmental studies, and so forth. By contrast, seventeen PhDs were awarded in chemistry, six in mathematics, nine in physics, and nine in statistics (for a total of 11 percent) and quite a few of those involved applications to biological problems. Similar numbers also hold at the Ohio State University. In the comparable period, OSU awarded 806 PhDs. Of these, 307 (38.3 percent) were in disciplines of the biological sciences. By contrast, twenty-one PhDs were awarded in chemistry, twenty-one in mathematics, twenty-four in physics, and sixteen in statistics (for a total of 10.2 percent). Today, most of science is biology.

Three factors have spurred the enormous growth of biological research. First, the technological revolution has made it possible to measure biological quantities that were previously out of reach. This drove great leaps in biological understanding and enabled the application of mathematical sciences techniques and physical sciences techniques to biological systems. Second, government and biotech companies are interested in increasing the understanding of biological systems that directly relate to human health and the environment, and therefore have provided new funding. The biotech industry also provides an important new source of employment for PhDs in the biological sciences. Third, universities establish departments and research programs in areas where they will attract students. Almost everyone is interested in human

Michael C. Reed is professor of mathematics at Duke University. His email address is reed@math.duke.edu.

physiology, human health, and the ecology of our planet. To understand the scientific interests of an average educated person, one need only look at the percentage of articles in the Science section of *The New York Times* that are devoted to psychology, biology, ecology, and medicine, compared to mathematics, physics or chemistry.

By many measures, the mathematical sciences community has responded robustly and productively to the biological revolution. For example, in 2007–2008 there were ninety-eight PhDs (9.6 percent of mathematics PhDs) in the subfield of mathematical biology and by 2011–2012 this had risen to 164 PhDs (13.3 percent). Biostatistics has always been a large part of statistics, but even there the PhD numbers have risen from 140 (41.2 percent) in 2007–2008 to 235 (46.3 percent) in 2011–2012. Not only has the community of mathematicians who would self-identify as "mathematical biologists" grown, but many other mathematicians spend part of their research time working on biological problems or working on questions in pure mathematics inspired by biology.

New Questions for Core Mathematics
Modern mathematicians sometimes dismiss physics as "applied mathematics" and thereby ignore the fundamental role that physics played in the development of conceptual ideas in core mathematics. The problem of the planets created the discipline of dynamical systems. The heat, wave, and Maxwell's equations drove the development of partial differential equations. In the modern era, crystallography was partially responsible for the development of group theory, quantum mechanics was an important stimulus for functional analysis, relativity theory for problems in geometry, and quantum field theory for string theory.

The intellectual traffic on the mathematics-biology interface, from mathematics to biology and from biology to mathematics, goes in both directions. We mathematical biologists saw clearly that the enormous development of core mathematics and applied techniques in the twentieth century would find fruitful and important applications to biological systems. The development of these applications began slowly partially because of resistance by biologists and partially because the collection of revealing data, especially on the microscopic level, was difficult. However, in the past fifteen years, resistance has decreased and the development of new measurement tools (often by physicists) has driven an explosion of applications of diverse branches of mathematics and statistics to biological problems. Not only have the traditional tools of ordinary and partial differential equations been used, but graph theory and random graphs have been applied to epidemiology and gene networks [12]. Topology has been used to understand heart fibrillation [28] and algebraic topology to understand neuroscience [8], [5] and image classification [6]. Fractal geometry is used to analyze dendritic arborization [36], and geometric concepts are fundamental for understanding protein folding [29] and the shape of docked proteins. Homological methods are used to characterize the behavior of large data sets and dynamical systems [25]. Combinatorics is used to understand RNA secondary structure [20]. Probability theory [3], [33], stochastic processes [13], and branching processes [14] are becoming central to biological applications and algebraic tools have been used in neuroscience [9] and network analysis [38], [21].

But is there any new mathematics? The answer is "yes, lots." In the twentieth century there were three main influences of biology on mathematics. The theory of evolution and genetics stimulated the fields of statistics, probability, and stochastic processes [41], [40], [26]. The Hodgkin-Huxley equations [22] and Turing's paper on morphogenesis [37] inspired research in reaction-diffusion equations, pattern formation, and traveling waves [17], [16]. Sequencing and reconstruction of the human genome created new questions in probability, statistics and combinatorics [27], [39]. All three of these major influences continue today [11], [34], [32]. In this century, the development of new core mathematics stimulated or inspired by biology has been increasing rapidly as more core mathematicians have gotten acquainted with and involved in biological problems. Biology has created fundamentally new questions in statistics [24] and stimulated the field of algebraic statistics [10]. The issue of how to compare teeth in paleontology led to new questions in conformal geometry [30]. The transport of materials in axons led to new phenomena and theorems in partial differential equations [18]. The theory of biochemical reactions stimulated new theorems in dynamical systems [2], [15] and in queueing theory [31]. The problem of how to compare different proposed phylogenetic trees led to the development of geometric central limit theorems on nonsmooth spaces [23], [43]. Since biological dynamics is very complicated and often parameters are known only approximately or not at all, one needs new coarse-grained methods for the classification of dynamical systems [35], [4]. The issue of how to detect the shape features of proteins stimulated new methods for the shape analysis of surfaces [1]. The effort to understand central pattern generators in the nervous system led to new work exploiting groups of diffeomorphisms to characterize symmetries in the solutions of dynamical systems [19]. The problem of providing low dimensional approximations for very large data sets has led to new questions in harmonic analysis [7].

These brief lists are certainly not comprehensive and are biased by my own knowledge. But they show clearly the breadth of mathematical techniques

being applied to biological problems and that biology is stimulating new mathematics in many disciplines of core mathematics. Two such areas are described in the companion articles by my colleagues Mauro Maggioni and Ezra Miller. It is certain that biology will have as profound an effect on the future of mathematics as physics did in the past.

New Mathematics Majors

Most undergraduate math majors really like mathematical biology. The reason is that they are excited to see the techniques that they are learning in their mathematics courses, whether calculus, graph theory, ODEs or probability, applied to questions that they are interested in. There's nothing wrong with the traditional applications to physics and engineering, but most of the applications that you can get at with undergraduate mathematics are nineteenth (or eighteenth) century applications, and the students can see that. But in biology they can see mathematics being used to gain understanding in situations where the full scientific picture is not understood by anyone. So, even though they are armed only with some undergraduate courses, they feel they are at the cutting edge, and, in some sense they are. Biological systems are so varied and diverse and mathematical biology is so recent that it is relatively easy to find interesting, exciting "research" projects in which undergraduate mathematics can make a real contribution. And one doesn't need to be a mathematical biologist to create and mentor such projects. If you choose any biology topic (like "prostate cancer," "evolution of the flu," "fetal development") and add the words "mathematical model," a Google search will produce many papers that are accessible to undergraduates. Reading a paper or two, learning the background biology, and working through the mathematics will usually be the core of a good project. One does have to accept that one is not "expert" at the biology in the way that we mathematicians are used to feeling expert compared to our students (in fact the student may know more biology than the faculty member). It helps to realize that even biologists are expert at only 10 percent of biology. There's simply too much biology, it's too difficult, and often the underlying mechanisms are unknown.

The excitement of undergraduates for mathematical biology is a real opportunity for mathematics departments to attract new majors, minors, and joint majors. However, to attract new majors one must design courses that are interesting to freshmen and sophomores even though they have not yet completed the traditional prerequisites for the math major, first year calculus, multivariable calculus, and linear algebra. For twenty-five years I've taught a freshman seminar entitled, "Applications of mathematics to physiology and medicine," that has an interesting clientele. The students are usually quite strong in mathematics and science; often they've arrived at Duke with a five on the BC calculus test. However, they planned to take no mathematics at all their freshman year because they've been told that they already have all the mathematics they need to be biology majors and/or pre-med. In addition to lectures and problems they also do a project and they are amazed to discover that mathematics is useful and can be applied to biological or medical topics that they are interested in. Some of these students become math minors or majors.

Offering freshman and sophomore topics courses is a departure from the traditional mathematics curriculum that had only one way into the major and emphasized the development of technique. The students who've always known they wanted to be mathematicians (like most of the readers of this article) will take the traditional path. It's the others we need to attract: the future doctors, lawyers, economists, who could be math majors (or minors or double majors) instead of biology or economics majors. We know that the mathematical techniques and logical reasoning that we teach will make them better professionals. We just need to interest them by offering attractive courses that change the attitude instilled by the culture and their high school courses that mathematics consists of boring methods that aren't really useful. Of course, mathematical biology is only one of the possible topics for such courses.

Many four-year colleges have discovered that mathematical biology is a good way to attract math majors and this is one of the reasons that the job market for mathematical biologists is excellent. Equally important is the fact that medical research groups have learned the advantage of having mathematicians on staff. In addition, government labs and the large and growing bio-tech industry hire many PhD mathematicians with training in mathematical biology.

A New Public Face for Mathematics

We mathematicians often lament that we are underappreciated. We know that mathematics is beautiful, that it is fundamental, and that it is important for applications. How come all those other people, the general public, the House Science Committee, don't seem to know that? The fact is that we will have a hard time convincing the general public that mathematics is beautiful and fundamental. We have a much better shot at convincing them that mathematics is important and useful. And, here, mathematical biology is making and will continue to make real contributions.

Mathematics, and sometimes new and difficult mathematics, is important for applications. The general public can't understand the mathematics, but they certainly can understand the applications.

The traditional important applications of mathematics have been to physics, from the motions of the planets to quantum mechanics, nuclear fission and the bomb, and fluid flow over airplane wings. Unfortunately, most people just aren't very interested in physics (Voltaire had it right), so they acknowledge the importance but aren't that moved. How about the applications of number theory to cryptography? Again, everyone sees that it is important to have secure communications, but they're not very interested in how it gets done. Leave it to the geeks! But biology is a different story. Everyone is interested in his or her own body and how it works. Everyone wants to be free of disease and live a long time. Everyone (well almost everyone) knows that we'd better be good stewards of our ecosystems or we and our children are doomed. So, when you tell them how mathematics is contributing, they are really interested. And this has the potential, in the long run, to greatly improve the public perception of mathematics.

References

1. P. AGARWAL, H. EDELSBRUNNER, J. HARER, and Y. WANG, Extreme elevation on a 2-manifold, *Discrete Comput. Geom.* **36** (2006), 553–572.

2. D. ANDERSON, A proof of the global attractor conjecture in the single linkage class case, *SIAM J. Appl. Math.* **71** (2011), no. 4, 1487–1508.

3. D. ANDERSON, J. MATTINGLY, H. F. NIJHOUT, and M. REED, Propagation of fluctuations through biochemical systems, *Bull. Math. Biol.* **69** (2007), 1791–1813.

4. Z. ARAI, W. KALIES, H. KOKUBU, K. MISCHAIKOW, H. OKA, and P. PILARCZYK, A database schema for the analysis of global dynamics of multiparameter systems, *SIAM J. Appl. Dyn. Sys.* **8** (2009), 757–789.

5. G. CARLSSON, Y. DABAGHIAN, F. MEMOLI, and L. FRANK, A topological paradigm for hippocampal spatial map formation using persistent homology, *PLoS Comp. Biol.* **8** (2012), no. 8, e1102581.

6. G. CARLSSON, T. ISHKHANOV, V. DESILVA, and A. ZOMORODIAN, On the local behavior of spaces of natural images, *Int. J. Comput. Vision* **76** (2008), 1–12.

7. R. COIFMAN, I. KEVREKIDES, S. LAFON, M. MAGGIONI, and B. NADLER, Diffusion maps, reduction coordinates and low dimensional representation of stochastic systems, *SIAM J. Mult. Mod. Sim.* **7** (2008), 842–864.

8. C. CURTO and V. ITSKOV, Cell groups reveal structure of stimulus space, *PLoS Comp. Biol.* **4** (2008), e1000205.

9. C. CURTO, V. ITSKOV, A VELIZ-CUBA, and N. YOUNGS, The neural ring: an algebraic tool for analyzing the intrinsic structure of neural codes, *Bull. Math. Biol.* **75** (2013), 1571–1611.

10. M. DRTON, B. STURMFELS, and S. SULLIVANT, *Lectures on algebraic statistics*, Springer, New York, 2009.

11. R. DURRETT, *Probability models for DNA sequence evolution*, Springer-Verlag, New York, 2008.

12. _____, *Some features of the spread of epidemics and opinions on a random graph*, PNAS **107** (2010), 4491–4498.

13. _____, Population genetics of neutral mutations in exponentially growing cancer cell populations, *Ann. Appl. Prob.* **23** (2013), 230–250.

14. R. DURRETT, J. FOO, K. LEDER, J. MAYBERRY, and F. MICHOR, Intratumor heterogeneity in evolutionary models of tumor progression, *Genetics* **188** (2011), 461–477.

15. B. A. EARNSHAW and J. P. KEENER, Global asymptotic stability for nonautonomous master equations, *SIAM J. Appl. Dyn. Syst.* **9** (2010), 220–237.

16. P. C. FIFE, *Mathematical aspects of reacting and diffusing systems*, Springer, New York, 1979.

17. R. FITZHUGH, *Biological engineering*, ch. Mathematical models of excitation and propagation in nerve, 1–85, McGraw-Hill, New York, 1969.

18. A. FRIEDMAN and B. HU, Uniform convergence for approximate traveling waves in linear reaction-hyperbolic systems, *Arch. Rational Mech. Anal.* **186** (2007), 251–274.

19. M. GOLUBITSKY and I. STEWART, Nonlinear dynamics of networks: the groupoid formalism, *Bull. Amer. Math. Soc.* **43** (2006), 305–354.

20. C. HEITSCH, A. CONDON, and H. HOOS, *DNA computing*, ch. From RNA Secondary Structure to Coding Theory: A Combinatorial Approach, pp. 215–228, Springer, New York, 2003.

21. F. HINKELMAN, D. MURRUGARRA, A. JARRAH, and R. LAUBENBACHER, A mathematical framework for agent-based models of complex biological networks, *Bull. Math. Biol.* (2011), 1583–1603.

22. A. L. HODGKIN and A. F. HUXLEY, A quantitative description of membrane current and its application to conduction and excitation in nerve, *J. Physiol.* **117** (1952), 500–544.

23. T. HOTZ, S. HUCKEMAN, H. LE, J. S. MARRON, J. C. MATTINGLY, E. MILLER, J. NOLEN, M. OWEN, V. PATRANGENARU, and S. SKWERER, Sticky central limit theorems on open books, *Ann. Appl. Prob.* **23** (2013), 2238–2258.

24. I. M. JOHNSTONE, On the distribution of the largest eigenvalue in principal components analysis, *Ann. Statist.* **29** (2001), 295–327.

25. T. KACZYNSKI, K. MISCHAIKOW, and M. MROZEK, Computational homology, *Applied Mathematical Sciences*, vol. 157, Springer, New York, 2004.

26. S. KARLIN, Evolutionary processes and theory, *Academic Press*, New York, 1986.

27. S. KARLIN and S. F. ALTSCHUL, Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes, *PNAS* **87** (1990), no. 6, 2264–2268.

28. J. P. KEENER, The topology of defibrillation, *J. Theor. Biol.* **230** (2004), 459–473.

29. J. LIANG, H. EDELSBRUNNER, P. FU, P. SUDHAKAR, and S. SUBRAMANIAM, Analytical shape computation of macromolecules: I. molecular area and volume through alpha shape, *Proteins* **33** (1998), 1–17.

30. Y. LIPMAN and I. DAUBECHIES, Conformal Wasserstein distances: Comparing surfaces in polynomial time, *Adv. in Math.* **227** (2011), 1047–1077.

31. W. H. MATHER, J. HASTY, L. S. TSUMRING, and R. J. WILLIAMS, Factorized time-dependent distributions for certain multiclass queueing networks and an application to enzymatic processing networks, *Queueing Systems* **69** (2011), 313–328.

32. R. MIHAESCU and L. PACHTER, Combinatorics of least square trees, *PNAS* **105** (2008), no. 36, 13206–13211.

33. C. MITCHELL and M. REED, Neural timing in highly convergent systems, *SIAM J. App. Math.* **68** (2007), 720–737.

34. W.-M. NI, The mathematics of diffusion, CBMS Conf. Ser. Appl. Math. 82, *SIAM*, Philadelphia, 2011.
35. J. RINZEL, Mathematical topics in population biology, morphogenesis and neurosciences, *Lecture Notes in Biomath*, no. 71, ch. A formal classification of bursting mechanisms in excitable systems, 267–281, Springer.
36. D. RISTANOVIC and N. T. MILOSEVIC, Fractal analysis: methodologies for biomedical researchers, *Theor. Biol. Forum* **105** (2012), 99–118.
37. A. M. TURING, The chemical basis of morphogenesis, *Phil. Transact. Royal Soc. B* **237** (1952), 37–72.
38. A. VELIZ-CUBA, An algebraic approach to reverse engineering finite dynamical systems arising from biology, *SIAM J. Appl. Dynam. Sys.* **11** (2012), 31–48.
39. M. WATERMAN, *Introduction to computational biology*, Chapman Hall/CRC, Boca Raton, FL, 1995.
40. S. WRIGHT, *Evolution and the genetics of populations: Genetics and biometric foundations, vol. 1-4.*, U. Chicago Press, Chicago, 1984.
41. ———, The roles of mutation, inbreeding, crossbreeding and selection in evolution, *Proc. 6th Int. Cong. Genetics* **1** (1984), 356–366.
42. M. REED, Why is mathematical biology so hard? *Notices AMS* **51** (2004), 338–342.
43. S. HUCKEMANN, J. MATTINGLY, E. MILLER, and J. NOLEN, Sticky central limit theorems at isolated hyperbolic planar singularities, *Elec. J. Prob.* **78** (2015), 1–34.