Feature Book Review

Learning the Lingo

***A Cell Biologist's Guide to Modeling and Bioinformatics*, by Raquell M. Holmes; 2007; 203 pp.; John Wiley and Sons (Hoboken, NJ); ISBN-10:0471164208**

Reviewed by Laurie J. Heyer

Davidson College, Davidson, NC 28035-7118

One of my greatest challenges over the last six years of teaching an undergraduate course in bioinformatics has been finding an appropriate textbook. Although the field of bioinformatics has experienced explosive growth over the last decade, the concomitant increase in available books in this field has failed to produce the right textbook for my course and my students. Each book targets a different audience and applies a different approach to the subject, but the combinatorial diversity results in near misses. To date, bioinformatics books have been aimed at mathematicians, statisticians, computer scientists, molecular biologists, or pharmaceutical and medical researchers. They are pitched at the beginning, middle, or advanced undergraduate level, or to technicians, graduate students, or practicing scientists. Authors take theoretical, abstract, practical, hands-on, and case-based approaches. Topics are focused or comprehensive according to the authors' own experiences and individual opinion of what constitutes bioinformatics. Each book, with its unique combination of target audience, level of difficulty, pedagogical approach and coverage, fills a small niche of the broad computational biology landscape.

In *A Cell Biologist's Guide to Modeling and Bioinformatics*, Holmes attempts the admirable but nearly impossible task of introducing not only bioinformatics, but also computational cell biology, in a slim volume aimed at both practicing biologists and undergraduate students. The result is a whirlwind tour of mathematical and computational approaches to biology. The reader is exposed to an incredible range of ideas, some in sufficient depth to be put directly into practice, others just skimming the surface of available databases and tools.

After outlining the purpose and value of computational and mathematical approaches to biology in a six-page introductory chapter, Holmes hits the highlights of bioinformatics in the next two chapters. Chapter 2 describes methods and databases for finding sequences that are similar to a

query sequence, with emphasis on BLAST at the National Center for Biotechnology Information. The algorithm and parameters are described in enough detail to demystify BLAST; the careful reader will know quite a bit about how to obtain and interpret desired results. Figures containing screen shots are particularly helpful in following along with the description, though they are somewhat difficult to read. Unfortunately, a few misstatements about E-values may confuse the novice BLAST user. An E-value is misinterpreted as a P-value, when in fact these quantities are distinct. A typographical error compounds the confusion between expectation and probability by describing an E-value of 0.02 as implying a 20% chance of obtaining the corresponding alignment score by chance (rather than $1-e^{-0.02}$, or approximately 2% chance). The author's choice to explain global and local alignment algorithms is surprising but welcome—too many biologists never know about these more exact methods for sequence comparison.

Chapter 3 focuses on protein sequence analysis, in particular, the characterization of protein domains and families, and lists several important methods and databases. Holmes, however, did not cover other bioinformatics topics such as:

- algorithms and tools for predicting RNA and protein structure
- finding genes, binding sites, and other motifs in newly sequenced organisms
- comparing whole genomes
- inferring relationships among species in phylogenetic trees
- analyzing genome-wide expression data
- inferring genetic regulatory relationships in gene networks

Chapter 4 discusses computational cell biology models in abstract terms to provide a foundation for subsequent chapters, but the foundation it provides is insufficient. For example, the statement “we know that the plot of the velocity versus substrate concentration of a Michaelis-Menten model

DOI: 10.1187/cbe.08-06-0034

Address correspondence to: Laurie J. Heyer (lahey@davidson.edu).

will create a hyperbolic curve, even without plugging in specific numbers” does not give the reader a feel for the shape of the curve or how it changes with different parameter values. I’m not sure what a hyperbolic curve is, but Holmes clearly misses the opportunity here to graphically illustrate Michaelis-Menten and Hill type kinetics. Furthermore, the section “ODE Essentials” (ODE, a term that is not clearly defined, stands for ordinary differential equations, i.e., equations containing derivatives, but no partial derivatives) is full of jargon and does not clearly convey the key ideas.

The classical computational cell biology modeling concepts described in Chapters 5 through 7—metabolism, cell cycle and calcium dynamics—are the heart and soul of this book. Although the models in these chapters rely on techniques described in Chapter 4, the context provided by the

models, combined with additional explanations of some techniques, should allow the cell biologist to follow the process. Holmes takes the unusual approach of using a different software program for each of these models, making each chapter independent of the others, and providing a broader exposure to modeling tools.

The greatest benefit of this book is for the researcher in cell biology who hopes to begin a collaboration with a mathematician or computer scientist, and needs a working vocabulary of modeling and computational techniques. The book does not claim to train the reader to become a computational biologist for good reason. However, this book delivers on the promise to increase the reader’s confidence in conversations with computationally trained colleagues. With this unusual combination of topics and approaches, yet another niche is filled in the computational biology landscape.