

PREFACE

12860

The last decade has seen a veritable explosion in the amount of raw information generated by molecular biologists worldwide. Coming to grips with this cascade of information has now become for many a full-time vocation, in contrast to the not-too-distant past when such analyses were practiced primarily by disgruntled graduate students who preferred handling a pipette to banging at a keyboard. The burgeoning field of bioinformatics, loosely defined at the intersection of molecular and computational biology, has given rise to numerous important discoveries and holds the promise of revealing even more of nature's secrets. For many, the attraction to bioinformatics rests in its newness as an untamed area of biology. For others, the attraction lies elsewhere—between the reductionist's fondness for details at the chemical level and the phylogeneticist's elation at understanding relationships within the parade of kingdoms.

Those with a conversational knowledge of bioinformatics have heralded its virtues broadly, calling it everything from the magic bullet that will cure all infirmities known to man to a brute-force powertool for dismantling sequence data to simply a sexy way to do science. In reality, bioinformatics is a new way to do hard, meaningful work. The field is one of methods that are constantly in flux and not fully developed, not unlike the early, golden years of biochemistry when the methods of choice involved melting and spinning a target biomolecule, in contrast to the much finer and more delicate methods available in the lab today. In the race to push bioinformatics along, however, a number of high-profile movers and shakers have conspired to downgrade bioinformatics from being a science in its own right to merely a function that can be done by buying the right kit. What keeps this field centered is the academic user community itself, whether the members are at private universities or at government-sponsored research centers. The significant advances being made in bioinformatics rest here, ranging from gathering and curating the raw data to developing new and elegant methods for examining the data, all in an open environment where information and expertise are freely shared. The bioinformatics community is unique in that, outside of the commercial sector, the *esprit de corps* is much more open than in many of the competitive arenas in biology. It is from this last thought that this book takes its inspiration, in enabling scientists who wish to know more about how to do sequence analysis to jump aboard and experience the ride that has enthralled so many. Through this book, we hope that the reader will come to appreciate the rigor of these methods and to understand that, not unlike the laboratory, the running of controls and the realization of what the methods can and cannot do are of the utmost importance. Most of all, we encourage the readers to just simply play.

12860

There are many individuals we both thank, without whose efforts this volume would not have become a reality. First and foremost, our thanks go to the all the authors whose indi-

vidual contributions make up this book. The combination of expert and professional viewpoints, coupled with general good-naturedness even under tight time constraints, has made working with these men and women a pleasure.

Inasmuch as a good portion of this book is devoted to the tools and databases developed at the National Center for Biotechnology Information (NCBI) at the NIH, we thank the all the members of the NCBI team for their hard work and dedication to maintaining the public databases under their charge and, more importantly, for making the data in those databases both of the highest quality and easily accessible to the scientific community. The computational biology community in general has been very generous with the fruits of its labor, making freely accessible tools and specialized databases without which elementary sequence analysis could not take place. We thank all those who have undertaken such projects, including those not specifically mentioned in this volume, for making possible many significant advances in this new age of sequence-based biology.

We also thank our editor, Ann Boyle, for her patience, encouragement, and support. This book marked a lot of firsts for us, so in the process of learning the ins and outs, we also developed a very strong friendship with her. We look forward to working with her again in the future.

From BFFO: I would like to acknowledge the constant support, love, and trust provided by my spouse, Nancy Ryder. Her respect for this work, as well as her having given me the space such an effort required, is beyond our vows, and I can only love her more for it. I also thank Mark Boguski, who invited me to the NCBI four years ago. Mark continues to direct my attention to a variety of new and interesting projects, and his continued enthusiasm and interest will always be inspirational to me.

And from ADB: My heartfelt thanks for the unbounded support and encouragement provided by David Landsman. My friendship with David began many years ago, when he was willing to take on as a post-doc a physical biochemist with little knowledge of what a brand-new field called bioinformatics was all about. I don't think either of us would have guessed back in 1992 that our teaching and research collaborations would lead to the writing of a textbook. Without his strong influence on how to think about computational biology, with the emphasis on *biology*, this book would not have been possible.

*Andreas D. Baxevanis
B.F. Francis Ouellette*