Book Review

Review of "Next-generation DNA sequencing informatics" by Stuart M. Brown (Editor)

Jennifer K. Sehn

Department of Pathology, Washington University School of Medicine, MO 63110, USA

E-mail: *Jennifer K. Sehn - jsehn@path.wustl.edu
*Corresponding author
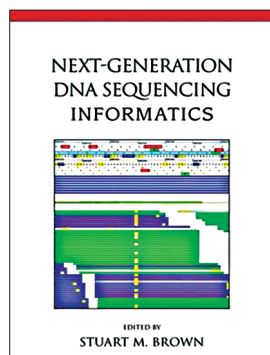
This article may be cited as:
Sehn JK. Review of "Next-generation DNA sequencing informatics" by Stuart M. Brown (Editor). J Pathol Inform 2013;4:25.

Available FREE in open access from: http://www.jpathinformatics.org/text.asp?2013/4/1/25/119003

Stuart M. Brown. (Editor)
New York: Cold Spring Harbor Laboratory Press,
Cold Spring Harbor; 2013
Pages: 256
ISBN-10: 1936113872

With the emergence of massively parallel sequencing methods (also known as next generation sequencing [NGS]) in basic science and clinical testing, there is a growing recognition of the central role of bioinformatics in the analysis of the deoxyribonucleic acid (DNA) sequence reads. The bioinformatics are complex, and familiarity with the nuts and bolts issues associated with the analysis of large datasets has thus far been restricted to experts with a background in computer science or computational mathematics. There is an especially clear need for practical reference materials for clinical NGS applications; the lack of books to guide real-world clinical applications is not only frustrating for pathologists and other physicians, but is also an impediment to widespread adoption of NGS techniques in patient care settings. The book Next-Generation DNA sequencing informatics is one of the early offerings positioned to fill this void.

The book, edited by Stuart M. Brown, with contributions drawn largely from the faculty at New York University, serves as a foundation for understanding the available sequencing platforms. It compiles numerous resources for a broad range of basic science and translational NGS applications. The text is appropriate for a mixed audience, including novice users of NGS who require definitions of common sequencing terminologies, as well as more advanced users with interest in a detailed review of the mathematical theories behind NGS data analysis. And, for the mathematically inclined, there are practical summaries of the algorithms underlying sequence alignment to reference genomes and de novo genome assembly. There is no doubt that this book provides a valuable discussion of the bioinformatics pipelines that are in common use in basic science settings for the analysis of genomic DNA from NGS of various organisms. The usefulness to researchers is enhanced by the presentation of software and visualization tools for the management of data sets from the related RNAseq and CHiP-seq approaches.

Despite these useful features, several aspects of the book limit its utility for those involved in NGS in clinical settings. First and foremost, throughout the book, the discussion is focused almost exclusively on detection of single nucleotide variants. Readers are left to find other sources for discussion of databases and internet resources

Access this article online

Quick Response Code:

Website:
www.jpathinformatics.org

helpful for the annotation and interpretation of this class of sequence variants. In addition, there is virtually no discussion of bioinformatics approaches for identification of the other three classes of sequence variation, namely small insertions and deletions, copy number variants and structural variants, though these three classes of mutations occur in a significant proportion of inherited disorders and account for a large percentage of the sequence variation seen in malignancies. Pertinent alignment issues related to the detection of these classes of sequence variation, such as gapped versus ungapped approaches, are also not covered. Finally, the bioinformatics approaches described in this book are focused on testing performed to identify constitutional variants, without much consideration of issues in identification of somatic mutations.

At a more general level, the book primarily covers the analysis of NGS data sets derived from fully –omic approaches, including genomes, transcriptomes and the like. In contrast, NGS performed to direct patient care is currently focused on gene panels and exomes, whether in the setting of constitutional sequence variants in inherited diseases or somatic mutations in the setting of cancer. Pathologists and geneticists developing NGS assays for the clinical laboratory will need additional discussion of the opportunities provided by targeted sequencing approaches.

For physicians and scientists with little or no background in NGS bioinformatics, this book will be a helpful resource. However, clinicians seeking to deploy NGS in clinical laboratory settings will need additional reference works to guide development of their bioinformatics pipelines.