Preface

Expressed sequence tags (ESTs) are single-pass reads derived from randomly selected cDNA clones. As such, they provide a highly cost-effective route for the purposes of gene discovery. Over the past couple of decades, we have seen a remarkable rise in their use. To date ESTs have been generated for over 1,500 different eukaryotes with new species datasets being produced almost daily. These rich sources of sequence data provide a wealth of information that has been used to explore issues of eukaryotic diversity. Typically, EST sequencing has been used to identify novel and/or "interesting" genes associated with a species for which full genome sequencing is not currently economically viable; however, ESTs have also proved valuable for purposes of genome annotation and also as alternatives to microarray analyses, for the assessment of relative levels of gene expression.

With the advent of "Next Generation" sequencing technologies such as the Illumina Solexa, ABI SOLID, and 454 Life Sciences platforms, EST-based resources are now becoming readily accessible for labs with even limited budgets. This volume aims to introduce the reader to many of the fundamental concepts underlying the generation and analysis of ESTs. Targeted mainly at life-science researchers (clinician, zoologist, botanist, etc.) interested in incorporating ESTs into their research programs, chapters are provided that give detailed descriptions of the various methods used to generate and analyze EST datasets. While primarily aimed at scientists interested in exploiting EST technology as a means of surveying the genetic diversity of an organism of interest, this book will also be of interest to those researchers who wish to use EST technology for other purposes such as expression profiling, analysis of alternative transcripts, and phylogenomics.

Following a brief overview of ESTs and their previous and current uses (**Chapter 1**), the first section of this volume (**Chapters 2–7**) focuses on general strategies surrounding EST projects and provides details on the methods employed to prepare cDNA libraries from a range of organisms (protists and fungi, plants, and animals). Within these chapters, the reader is presented with a number of protocols that outline the construction of a variety of different types of cDNA libraries including libraries that are enriched for full-length cDNA sequences, libraries that are normalized with respect to relative abundance, and libraries that make use of splice-leaders that some eukaryotes employ particularly for splicing of polycistronic units. **Chapter 8** then discusses available sequencing resources and provides protocols for the generation of sequence information from these libraries. **Chapters 9–11** present state-of-the-art software tools (which can operate on a single desktop workstation) used for processing the raw sequence information generated by the sequencing platforms and extracting biologically meaningful sequence data in the form of ESTs. Protocols are given to remove low-quality sequence or vector contaminants and process the ESTs into a format that can be submitted to the public EST database resource – dbEST. Additional bioinformatic pipelines are described that (1) cluster ESTs into nonredundant sets of related

gene objects; (2) annotate these clusters with functional associations; and (3) store the results within a centralized Web-accessible database resource. **Chapters 12–13** outline statistical methods to analyze these data, both to examine differential expression and also to explore their evolutionary properties. Finally, the last two chapters (**14 and 15**) discuss how ESTs can be used to construct microarray resources and, in more global terms, how they can be usefully applied within the context of human health.

This book assumes that the reader has some knowledge of molecular biology techniques and is also familiar with basic bioinformatics concepts.

John Parkinson

http://www.springer.com/978-1-58829-759-4

Expressed Sequence Tags (ESTs)
Generation and Analysis
Parkinson, J. (Ed.)
2009, X, 332 p. 29 illus., Hardcover
ISBN: 978-1-58829-759-4
A product of Humana Press