PREFACE

This book is dedicated to the premise that nothing is more fundamental to Life than the ability to reproduce. That ability begins with the simple fact that all living organisms store their blueprint for life (the genome) in deoxyribonucleic acid (DNA). DNA is the longest, thinnest biological molecule and the machinery that replicates it must complete the job quickly and accurately if the organism is to flourish. Over the past 60 years, science has elucidated the structure and organization of the genome (Chapter 1) and the molecular mechanisms that reproduce it in a variety of biological systems. Accordingly, the goal of *Genome Duplication* is to explain how genomes are duplicated in living organisms, from the simplest cell to the most complex multicellular organism, and how genome duplication impacts upon and is regulated by cellular activity. *Genome Duplication* is not intended to be an encyclopedia, but rather to identify the underlying concepts and principles that govern genome duplication in a world rich in biodiversity.

The development of concepts and principles that apply to all living organisms through the study of a few well-chosen paradigms is possible, because all living organisms can be classified into three domains: bacteria, archaea, and eukarya (Chapter 2). This book focuses on the intestinal bacterium *Escherichia coli* and the soil bacterium *Bacillus subtilis*, the archaeon *Sulfolobus solfataricus*, the single-cell eukaryotes *Saccharomyces cerevisiae* (budding yeast) and *Schizosaccharomyces pombe* (fission yeast), and multi-cellular eukaryotes *Drosophila melanogaster* (fruit fly), *Xenopus laevis* (African clawed frog), and the mammals *Mus musculus* (mouse) and, of course, *Homo sapiens*. Many examples from other organisms, as well as from bacterial and animal viruses and episomes, are included to illustrate specific points and provide historical context.

Not only do all living organisms use the same molecule to encode their genome, they all use the same mechanism to replicate it—semi-conservative DNA replication by the replication-fork mechanism (Chapters 3 through 6). Remarkably, the sequence of events in the assembly and activation of replication forks is essentially the same throughout the three domains of life. To paraphrase Jacques Monod's famous dictum, what is true for replication forks in bacteria is also true for replication forks in elephants. Thus, genome duplication is arguably the most highly conserved of biological processes and, as such, it is a window into biological evolution. Not surprisingly, the subject of evolution appears throughout the book, but particularly in Chapters 2 and 15.

Genomes do not exist as simple DNA molecules in bacteria, archaea, or eukarya, because DNA is frequently modified chemically and invariably organized into a DNA–protein complex (Chapter 7). These modifications protect the DNA from damage, compact it into a tiny space, and regulate expression of its genetic information. Thus, each time a cell divides, it must dismantle these modifications in order to duplicate its genome, and then restore them as DNA replication proceeds. Moreover, as genome duplication occurs in eukarya, the two sibling molecules (sister chromatids) are embraced by a unique ring structure that remains until they are separated during cell division.

During the course of evolution, the nature of replication origins changed from unique DNA sequences that are essential for assembly and activation of their cognate replication machinery to ill-defined sequences that are determined more by epigenetic than genetic parameters (Chapters 8 through 10). This transition from rigidly defined replication origins to conditionally defined replication origins allowed the evolution of complex, multicellular organisms. The machinery that converts DNA replication origins into two replication forks that travel in opposite directions is impressive (Chapter 11), involving 20 different proteins in bacterial cells and 66 in human cells.

To be productive, all organisms must link cell division tightly to genome duplication. In eukarya, duplication of the nuclear portion of their genome is generally (but not always) restricted to once and only once per cell division (Chapter 12). When a cell has committed itself to dividing, it either completes the process or dies. However, plants and invertebrates frequently employ multiple rounds of genome duplication in the absence of cell division to facilitate post-mitotic growth. Only bacteria and probably some archaea can successfully launch a second set of replication forks from the same replication origin before the first set of forks have terminated.

Eukarya contain a series of feedback loops driving the cell-division cycle in one direction. Checkpoints (Chapter 13) ensure that one phase is completed before the next phase begins. Bacteria contain at least two checkpoints, one to ensure that a cell has produced sufficient materials to support cell division (the restriction checkpoint) and one that senses DNA damage during DNA replication. Eukarya sense DNA damage before, during, and after the DNA replication has taken place. Eukaryotic cells also detect stalled replication forks, prevent entry into mitosis before DNA replication is completed, and prevent cell division before mitosis is completed. If a cell fails to correct the problem, then it is programmed to die.

All of this leads to the understanding and treatment of human disease (Chapter 14). Remarkably, with the exception of diseases caused by prions and RNA viruses, all of the remaining infectious diseases are caused by either microorganisms, DNA viruses, or RNA viruses with a DNA intermediate in their life cycle. This means that differences in the mechanisms by which DNA is duplicated in microbes and viruses compared with mammals offer

therapeutic targets. In fact, current antiviral therapies target the viral genome duplication almost exclusively. Mistakes in genome duplication as well as infection by DNA viruses can also result in noninfectious diseases, such as cancer, a leading cause of death worldwide. DNA replication proteins not only provide useful biomarkers for the early detection of cancer, but most of the current cancer therapies target critical proteins that function at replication forks or that regulate genome duplication. Mistakes in genome duplication also are associated directly with a number of genetic disorders, including those associated with neurodegenerative diseases, such as Huntington's disease, spinocerebellar ataxia, and fragile X syndrome. Thus, it is not hyperbole to state that understanding how genomes are duplicated, how the process of genome duplication is regulated, and how genome duplication differs among the various forms of life is critical to the understanding and treatment of human disease.

Perhaps now one can appreciate how the subject of this book—genome duplication—links together all living organisms, their origins, their evolution, and their future. Unfortunately, the literature on this subject can be intimidating, and students often become lost in the labyrinth of mechanisms created by a diverse group of biosystems and described with a Byzantine nomenclature that challenges professionals as well as novices. *Genome Duplication* attempts to correct these problems, and with a basic, college-level knowledge of biology and chemistry, the reader will find this book both comprehensive and comprehensible. That said, the authors welcome any corrections or suggestions from the reader that would improve this book in a future edition.