Preface

What constitutes long-range control is, by its nature, difficult to define. In Drosophila, *cis*-acting long-range elements located at a distance of a kilobase or so from the target gene's promoter are considered to be operating at a long distance. On the other hand, in mammals the classic long-range element is the locus control region (LCR) of the beta-globin locus, sitting 50kb from the globin gene cluster. We now know that *cis*-regulatory elements can lie at a distance of at least a megabase, but the absolute distance that regulators are able to span to control gene expression is unknown. The properties that dictate these regulatory limits are still open to investigation. But no matter the distance, the goal is the same; that is, to convey regulatory information across an intervening stretch of DNA to control target gene expression. Although some aspects of regulatory control have been known for three quarters of a century, the recent availability of multiple genomic sequences and high-throughput technologies has generated an emerging view of regulatory mechanisms. In this volume, we cover some of the most recent advances in this fast-moving area of genomic mechanisms involved in gene regulation.

Control of gene expression is the major key to regulating cellular function in developing and adult organisms. Transcriptional control may be imposed by different mechanisms, ranging from regional chromatin organization to very specific transcription factor binding at gene-specific sites. One of the earliest observed phenomena was position effect variegation in Drosophila, which is explored from the current vantage point here. Drosophila studies were frequently at the forefront in subsequent discoveries, such as the identification of more widely acting regulators like the polycomb group of proteins and some of the earliest analyses of *cis*-regulatory sequences. One advantage of Drosophila is the availability of multiple species sequences for comparison, so that divergence as well as conservation of homologous regulatory sites can be assessed at the bioinformatic and functional level. It has been possible to explore how alterations in regulatory elements mirror the evolutionary changes in detailed expression pattern and changing morphology.

Switching from Drosophila to mammals, and utilizing initial information from human disease variants, one of the best explored gene expression systems has been the beta-globin complex where the exquisitely controlled developmental switching from closely linked embryonic to fetal and adult gene copies has been studied in great detail. Many aspects of now more widely used technology for the study of expression control were developed for this system.

It is a salutary lesson then to see that the alpha-globin complex, which maps to a different chromosome, reveals significant differences, as well as some similarities in the regulatory mechanisms utilized to control expression of the other subunit for this crucial tetrameric protein where levels of expression of the two different subunits are very tightly coordinated. The four clustered Hox complexes in mammals constitute another classical set of multigene assemblies, which have been elegantly dissected to reveal additional layers of complexity.

Imprinting, or parent-of-origin-specific gene expression at a number of important growth regulatory loci, began to be explored systematically in mice a few decades ago. Here, methylation differences were identified and the role of noncoding RNAs and insulator elements was gradually uncovered, revealing a complex system of regulation that is re-set every generation. As with the globin genes, human disease-associated abnormalities have contributed significantly to our understanding of this esoteric level of gene expression control. Insights from some of the implicated genes suggested that altered methylation and other epigenetic changes may play a key role in the development of malignancies.

With the advent of multiple genome databases, sophisticated bioinformatics method have been developed for the identification of conserved regulatory elements across different evolutionary distances. Methodologies have also been developed to couple bioinformatic identification to functional analysis. A major emerging theme from this has been that regulatory changes play a key role in the evolution of different vertebrate species, much as has already been observed for Drosophila. Interestingly, the possibility of broader sequence comparison analysis has revealed that the existence of long-range control, sometimes over regions of more than a megabase, has helped to shape genome evolution significantly. The largest regulatory distances are associated with major developmental control genes which fulfill multiple roles during early development and often in later maintenance of gene function. Some of these gene domains are so large that they overlap with flanking gene territories so that neighboring genes have to maintain their links with each other through evolution, leading to regions of conserved synteny.

Not surprisingly, mutations in regulatory regions are increasingly shown to be associated with human disease. However, currently, we are only observing the tip of the iceberg. It is becoming clear that many common disease-association studies are identifying noncoding region variants as the underlying cause of these later onset disorders. It will be exciting, and potentially useful for disease management and treatment, to see what aspects of fine tuning are altered in different anomalies. Areas for future exploration will include the mechanisms through which physiological and environmental changes are translated into altered gene function through long-range regulatory networks.

We want to thank all the authors who delivered the 13 excellent reviews for this volume and our publishing editors for help and guidance in getting the book out.

Robert E. Hill and Veronica van Heyningen