Published in final edited form as:

Curr Opin Genet Dev. 2009 December ; 19(6): 533–534. doi:10.1016/j.gde.2009.11.003.

Genomes and evolution: advances in identifying gene regulatory effects of evolution

Martha L. Bulyk and

Brigham & Women's Hospital and Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115, USA mlbulyk@receptor.med.harvard.edu

Michael Levine

Department of Molecular & Cell Biology, University of California–Berkeley, Berkeley, CA 94720, USA mlevine@berkeley.edu

The focus of this issue of *Current Opinions in Genetics and Development* is the identification and function of *cis* regulatory DNAs underlying the development and evolution of fungi, plants, and animals. There is little doubt that one of the outstanding challenges of the 21st century is the deciphering of the ‘dark matter’ of the genome; namely, the noncoding regulatory DNAs that control where and when every gene is switched on and off during development and in response to the environment. The evolutionary diversity of life on earth, including variations in human populations, relies extensively on changes in noncoding regulatory DNAs.

In this special issue the editors are pleased to present an eclectic mixture of 12 articles reflecting the diverse roles and methods of study of regulatory DNAs. We somewhat arbitrarily consider these papers in three categories: the role of *cis* regulatory DNAs in animal development, the identification of *cis* regulatory DNAs using comparative genome analysis, and the global evolution of regulatory systems.

Cis regulatory DNAs and animal development

The three papers by Davidson, Heintzman and Ren, and Lucchessi describe the central importance of *cis* regulatory DNAs in animal development. Several well-established and emerging models are considered, including endomesoderm specification in the early sea urchin embryo, dosage compensation of the X chromosome in the *Drosophila* embryo, and limb development (and evolution) in mammals.

Studies of the sea urchin have led to tremendous gains in understanding the nature of gene regulatory networks and *cis* regulatory elements. Davidson discusses subcircuits that effectively perform logic operations for the spatial control of gene expression in development. Transcription is regulated further by chromatin remodeling complexes. Heintzman and Ren describe advances in experimental strategies to identify long-range regulatory sequences, focusing on enhancers and insulators, throughout the human genome. Chromatin immunoprecipitation studies of histone modifications and the transcriptional coactivator p300 have led to the identification of ‘marks’ for enhancer elements; similar studies of the CTCF have led to identification of enhancer-blocking insulator elements. Lucchessi focuses on the

© 2009 Elsevier Ltd. All rights reserved.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

MSL complex responsible for dosage compensation in *Drosophila* in their description of experimental advances in studies of complexes that operate at the whole-chromosome level to regulate transcription.

Identification of *cis* regulatory DNAs by comparative genomics

The five papers by Noonan, Meireless-Filho and Stark, Wohlbach *et al.*, Wilson and Odom, and Dimas and Dermitzakis describe the ever-growing arsenal of methods for identifying *cis* regulatory elements. Foremost among these, driven in part by the diminishing costs of high-throughput genome sequencing, is the use of comparative sequence analysis. Such approaches have identified critical regulatory elements conserved among flies and among fungi. Moreover, sequence polymorphisms are beginning to identify putative regulatory DNAs in divergent populations of humans as well. Additional whole-genome methods are also considered, including chromatin immunoprecipitation coupled with high-throughput sequencing ('ChIP-seq') assays for the identification of *in vivo* transcription factor occupancies.

Noonan reviews recent testing of conserved noncoding sequences in *in vivo* reporter assays, which has revealed tissue-specific enhancers active during mammalian development. Noonan also describes recent studies of human-specific sequence changes in conserved noncoding sequences that appear to be "fast-evolving" in humans. Experimental studies of such regions from various species will aid in our understanding how species-specific regulatory changes lead to differences in animal morphology.

While sequence conservation can aid in the discovery of functional elements, Meireless-Filho and Stark describe various phenomena that can result in challenges for computational identification of *cis* regulatory elements. TF binding site divergence has been observed in mammals (and yeast) and suggests divergence in gene regulatory network connections. Functional conservation of regulatory elements despite lack of nucleotide identity can indicate motif degeneracy or binding site turnover with compensatory changes. Wohlbach *et al.* focus on *Ascomycota* fungi in their examination of divergence in gene expression across species. They discuss in particular how redundancy in gene copies, functional elements and molecular interactions may play a major role in the evolution of regulatory divergence. Wilson and Odom describe studies of the interactions of mammalian transcription factors interact with DNA. Recent results suggest that mutations in *cis* regulatory differences contribute to species-specific differences in gene expression. While variation in regulatory elements can reflect species-specific differences in gene regulation, polymorphisms in regulatory elements contribute to diversity within populations. Dimas and Dermitzakis focus in human populations in their discussion of how regulatory variation contributes to gene expression level differences among individuals.

Evolution of regulatory systems

The four papers by Degnan *et al.*, Hufton and Panopoulou, Bourque, and Keeling review the regulatory "toolkit" and the consequences of whole-genome changes on gene regulation. Transcription factors are a major component of the regulatory toolkit. Degnan *et al.* focus on the evolution of transcription factor classes and families during bilaterian development. They describe a number of transcription factor families that appear to be specific to metazoans and thus may have contributed to the evolution of animal multicellularity. Genome diversification can result from whole-genome duplication. Hufton and Panopoulou discuss whole-genome duplication among plants, yeast, vertebrates, and human cancers and the variety of ensuing outcomes including genome restructuring and genome stasis. Repetitive DNA has often been overlooked or ignored in genome analyses. Bourque focuses on transposable elements as a rich source of innovation in genes, regulatory elements and genome structures including structural rearrangements throughout evolution. Finally, horizontal gene transfer, best known for its

effects in prokaryotes, is also at play in eukaryotes. Keeling provides examples of potential adaptive advantage of horizontal gene transfer in eukaryotes, including in parasitism, pathogenesis, and particular metabolically challenging environmental conditions.

Outlook

Altogether, these short reviews provide an impressive overview of the many methods and systems being used in modern biology to address the great challenge of deciphering the genomics of gene regulation in development and evolution. The availability of more genome sequences, combined with the production of various high-throughput datasets on the functions of *cis* regulatory elements and *trans* regulatory factors and their effects on gene expression, will further enhance our understanding of how genomes shape cellular regulatory networks and phenotypes through evolution.

Biographies

Martha Bulyk is professor of Medicine and Pathology (and Health Sciences and Technology) at Harvard Medical School and Brigham & Women's Hospital. She uses computational and experimental approaches to study transcriptional regulatory networks in model organisms and in the human genome. Current research focuses on the development of various computational methods, including methods that employ phylogenetic conservation, to identify likely DNA regulatory elements in the human genome and in other model organisms. They are also applying novel microarray technologies for the high-throughput characterization of the binding specificities and regulatory roles of transcription factors (TFs).

Michael Levine is professor of Genetics, Genomics and Development at UC Berkeley. Mike's lab studies gene networks that control animal development and disease. Research focuses largely on how noncoding regions of the genome function to control the differential patterns of gene expression, both spatial and temporal, that define cell behavior. The developmental stage of multicellular organisms in particular requires very explicit control of gene expression in order to layout different cell tissue types. Using model developmental systems, including the early *Drosophila* embryo, the sea squirt, *Ciona intestinalis*, and the flour beetle, *Tribolium castaneum*, Mike and his colleagues are working to develop a deeper understanding of genetic regulatory codes.