



ELSEVIER

# Genomes and evolution From sequence to organism

Editorial overview

Evan E Eichler and Nipam H Patel

Current Opinion in Genetics & Development 2003,  
13:559–561

0959-437X/\$ – see front matter  
© 2003 Elsevier Ltd. All rights reserved.

DOI 10.1016/j.gde.2003.10.014

## Evan E Eichler

Department of Genetics, Center for  
Computational Genomics, Case Western  
Reserve University School of Medicine and  
University Hospitals of Cleveland, Cleveland,  
OH 44106, USA  
e-mail: eee@po.cwru.edu  
URL: <http://eichlerlab.gene.cwru.edu/eichler/>

Evan Eichler is an Associate Professor  
and Director of the Bioinformatics  
Core Facility in the Department of  
Genetics, Case Western Reserve  
University School of Medicine. The  
long-term goal of his research is to  
understand the evolution, pathology  
and mechanisms of recent genome  
duplication in human.

## Nipam H Patel

Depts. of Integrative Biology and Molecular Cell  
Biology, University of California, 3060 VLSB  
#3140, Berkeley, CA 94720-3140, USA  
e-mail: nipam@uclink.berkeley.edu  
URL: <http://www.patellab.org>

Nipam Patel's laboratory studies the  
evolution of developmental  
processes, with a primary interest in  
the evolution of arthropod  
segmentation and body patterning.

## Introduction

The field of evolutionary biology has been revitalized by both the recent availability of whole-genome sequences from a variety of disparate organisms, and by the use of genomic methods in organisms where more limited sequence information is available. Old and new questions regarding the nature of genetic variation, the origin of species, the developmental basis of morphological change, and the molecular basis of genome evolution are being explored with respect to their functional significance. Evolutionary biologists are exploiting genome sequence information in an unprecedented fashion. This issue of *Current Opinion in Genetics & Development* is designed to encapsulate some of the recent excitement and progress in this area of research. Surveying the reviews, two central themes emerge. First, these reviews highlight the use of whole-genome sequence to assess fundamental properties of molecular evolutionary change and its relationship to protein variation and changes in *cis*-regulation that ultimately underlie functional variation. Important lessons from these studies are presented that reinforce some old concepts, dispel others, and simultaneously introduce a few new ones. Second, comparative genomics emerges as being central to inferring patterns of genome evolution as well as their potential impact in terms of the evolution of developmental patterns and mechanisms. Both close and distant phylogenetic relationships are critical in this functional inference. These reviews highlight how recent advances in genomics have significantly enhanced the way in which many disciplines operate.

## Fundamentals of evolutionary change

Since the introduction of neutral theory of molecular evolution by Kimura in 1968 [1], the null hypothesis for the bulk of genomic evolutionary change has been that the underlying events are neutral in their effect and random in their distribution. Although large-scale genomic data have upheld this insightful and, once, controversial theory, some refinements have become apparent. The first set of reviews advances our understanding of the pattern and nature of genetic variation. Various perspectives of mutational change are showcased. They range from single base-pair changes to larger evolutionary changes in genomes. **Ellegren, Smith and Webster** summarize interspecific patterns of single-nucleotide variation among non-coding portions of mammalian DNA. Considerable variation and regional biases in mutation rates are observed. The authors put forward an attractive theory involving both mechanistic biases as well as selection of genome structure to explain the origin of such local variation in neutrally evolving DNA. Two sets of reviews, one by **Tishkoff and Verrelli** and the other by **Maloof**, explore patterns of single-nucleotide change within humans and *Arabidopsis* revealing patterns of discrete blocks of linkage disequilibrium. The potential value of such patterns of variation as tools in understanding recent

population structure and mapping disease and quantitative loci are discussed. The review by **Sankoff** takes a different perspective, examining large-scale patterns of rearrangement in both prokaryotes and eukaryotes. Here, the complexity of interpreting not only inversions and translocations but also other events including transposition, horizontal gene transfer, duplication and large-scale deletion are examined. The recent non-random distribution of such rearrangements, termed inhomogeneities, is noted with some caution. This review emphasizes the importance of developing appropriate computational algorithms to model the complete spectrum of large-scale genome rearrangements.

### Evolution of animal diversity

Over the past two decades, the evolution of developmental variation has received renewed interest in large part because of the advances made in unraveling the genetic basis of development in several model species. More recently, broader-scale genome evolution data has also begun to contribute to our understanding of the evolution of development. **Venkatesh** reviews our understanding of the remarkably 'plastic' properties of fish genomes, a plasticity that has been hypothesized to play a role in the remarkable radiation of teleost fishes. Whereas the existence of whole-genome duplications in teleost lineages remains controversial, the evidence for frequent gene duplications, polyploidization, and speciation are quite evident. **Aboobaker and Blaxter** focus on the evolution, in nematodes, of one particular region of animal genomes thought to have a fundamental role in developmental evolution: the Hox complex. Although the organization of this complex is generally well conserved in most animal lineages, the Hox complex has undergone extensive alterations in the nematode lineage, possibly reflecting an increasing tendency towards deterministic developmental modes in this group of animals. **Brooke and Holland** illustrate the many ways that genomic data can help unravel the evolution of multicellularity and early animal lineages. They review how recent genomic studies have helped clarify our understanding of the basal branches of the animal partition of the 'Tree of Life', and how the comparative analysis of specific genes has helped reveal the evolutionary history of several developmental innovations.

### Purifying and adaptive evolution

Of course, there are important exceptions to the basic neutral model of evolutionary change. One of the most obvious is founded on selection. Based on the premise that most of the functionally important regions of the genome effectively evolve more slowly, **Cooper and Sidow** tackle the issues surrounding the identification of regulatory regions from comparative whole-genome sequence data. A consideration of the scope of organisms useful for such purposes and the importance of coupling phylogenomic detection methods with experimental vali-

ation are emphasized. **Li and Wang** explore other advances in computational methods, such as the identification of clustered transcription factor binding sites, that allow regulatory regions to be identified. Combined with other experimental data, such computational approaches can help to characterize regulatory networks within organisms and will be especially important in identifying potential examples of *cis*-regulatory evolution that are believed to underlie much of developmental evolution.

Providing a counterbalance to the Cooper and Sidow perspective of conservation, **Swanson and Copley, Goodstadt and Ponting** examine the opposite spectrum of functionally important changes: the rapid evolution of genes and gene families that are essential for adaptation of a species to its ecological niche. Swanson surveys the methods for detection of these regions of rapid evolutionary change using genomic and population data. Copley *et al.* suggest that lineage-specific protein domains are the hallmarks of adaptive evolution within genome sequence data. Such expansions along with the formation of more complex proteins within increasing numbers of domains (domain accretion) are the driving forces behind vertebrate proteome evolution. Both adaptive evolution reviews conclude that genes associated with reproduction, immunity and disease resistance are the most common substrates of such bursts of positive selection. The discovery and function of other, less obvious, domains within organisms remains a significant challenge within evolutionary biology.

### Structural dynamics

Other regions of the genome appear to evolve at an accelerated pace due to the structural properties of the underlying sequence or their location within the genome. Unlike most single-nucleotide mutation rates that are relatively constant as a function of generation time, dynamic mutations are those in which the probability of an initial mutational event increases the probability of additional events [2]. Such events are therefore subject to episodic bursts of activity. Homology-driven duplications, whole-genome duplications, microsatellite evolution and retrotransposition may be viewed in this light. **Jackson** explores the bias for segmental duplications to accrue within the pericentromeric regions of primate genomes. Rapid evolutionary genomic turnover, outpacing that of 'normal' neutral sequence, is described leading to the formation of complex mosaic blocks of juxtaposed segmental duplications abutting centromeric DNA. A tantalizing model involving segmental duplication, centromere repositioning and the euchromatization of pericentromeric sequence is proposed as a potential mechanism for the formation of novel 'chimeric' genes during hominoid evolution. **Seoighe and Paterson et al.** examine the evidence both in favor and against ancient whole-genome duplications within the genomes of *Arabidopsis*, *Saccharomyces cerevisiae* and *Caenorhabditis elegans*.

Since Ohno's advocacy of whole-genome duplication in 1970 [3] as a major force in gene evolution, the topic has been the subject of controversy and at times acrimonious debate. Although the debate continues to rage, whole-genome sequence data are beginning to shed some light on the subject. Most of the available genomic analyses, at least for these organisms, support a minimum of one ancient whole-genome duplication during their evolution. The relative importance of such episodes of polyploidization to radically restructure eukaryote genome and proteome content, however, as compared to more gradual processes of segmental duplication has not been satisfactorily resolved.

Finally, **Deininger, Moran, Batzer and Kazazian** — leading experts in the field of retrotransposition — explore the potent mutagenic properties of LINE, SINE and LTR parasites in sculpting eukaryotic genomes. Extreme variation in the rates of retrotransposition among different species and the homology-driven rearrangement that these elements may promote makes these elements unpredictable and multifaceted agents of evolutionary change. By virtue of these properties and the sheer abundance of these elements, the authors argue that mobile elements are the dominant force in reshaping most of the eukaryotic genome.

### Conclusions

Although some of the topics addressed in this issue of *Current Opinion in Genetics & Development* may be per-

ceived as esoteric endeavors of molecular evolution, the results of these areas of research have ramifications that extend far beyond the simple reconstruction of hypothetical historical events. Understanding the historical pattern of genetic variation is the link to understanding the molecular basis of morphological variation. The development of grains that are disease resistant, the mapping of quantitative traits and the discovery of susceptibility factors that associate with complex genetic diseases require that the nature and pattern of genetic variation be understood. Genome sequence and patterns of change can be seen as only a first step in the process. Further theoretical and technological advances will be required to address the functional significance of such change in a high-throughput fashion. Understanding such functional changes in the context of genetic variation is key to this progress. In the era of genomics, the future of evolutionary biology is bright. Its relevance to commerce and biomedical sciences is becoming self-evident. Significant advances are assured as long as developmental biology and molecular evolution continue to cross-fertilize within the field.

### References

1. Kimura M: **Evolutionary rate at the molecular level.** *Nature* 1968, 217:624-626.
2. Richards RI, Sutherland GR: **Heritable unstable DNA sequences.** *Nat Genet* 1992, 1:7-9.
3. Ohno S: *Evolution by Gene Duplication.* London: Allen and Unwin; 1970.