Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Genetics & Development

Genomes and evolution: multidimensional approaches to understanding diversity

Editorial overview

Sarah Teichmann and Nipam H Patel

Current Opinion in Genetics & Development 2008, **18**:469–471

Available online 26th December 2008

0959-437X/$ – see front matter

DOI 10.1016/j.gde.2008.11.003

Sarah Teichmann

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 0QH, UK
e-mail: sat@mrc-lmb.cam.ac.uk

Sarah Teichmann works on the evolution of protein–protein and protein–nucleic acid interactions using data from genome sequences, protein structures and functional genomics. A particular focus is on the prediction, evolution and dynamics of transcriptional regulators and their interactions. The classification and analysis of protein complexes is a further area of interest.

Nipam H Patel[1,2]
[1] Department of Molecular Cell Biology, 519A LSA #3200, Berkeley, CA 94720-3200, USA
[2] Department of Integrative Biology, 3060 VLSB #3140, Berkeley, CA 94720-3140, USA
e-mail: nipam@uclink.berkeley.edu

Nipam Patel works on the evolution of development, primarily in arthropods. Areas of interest include the evolution of segmentation and the developmental basis for appendage diversification and early cell fate specification. He is particularly interested in developing new model systems that can be experimentally manipulated to address macroevolutionary questions.

The availability of complete genome sequences has transformed developmental biology by allowing comprehensive comparisons of gene repertoires across organisms. A fundamental question in developmental biology is how development evolved in the first place, in the sense of multicellular organisms consisting of distinct cell types. Rokas reasons that this occurred through the expansion of genetic toolkit components that are often found in both unicellular and multicellular lineages. The recent sequencing of the unicellular choanoflagellate *Monosiga brevicollis* highlights this for protein families in signalling, while transcription factor families also follow this principle. For instance, the famous homeodomain transcription factors are found across plants, fungi, animals and other eukaryotes, but they are expanded and contribute to developmental body-building most significantly in animals. Using transcription factor annotations in completely sequenced genomes from the transcription factor database DBD [1], Rokas highlights the many lineage-specific expansions of transcription factor families across animals, plants and fungi.

Oakley and Rivera continue the discussion by describing a general approach to understanding the evolution of complexity, and then illustrate the approach with specific examples derived from studies that reveal how neural complexity may have come about. They define three general mechanisms for evolving complexity that function at many biological levels (genes, gene networks, cell types, tissues and organs), first, copying and differential divergence; second, fission and differential divergence and third, fusion of copied parts. The fact that these mechanisms can be applied at so many levels allows data from many different disciplines to be integrated together into a way that advances our ability to connect genotype with phenotype.

Complete genome sequences also allow comparisons of sex chromosomes in animals, in order to learn about the evolution of sex determination. The sequencing of a large number of *Drosophilid* and mammalian species sheds light on X chromosome evolution in these species. Gurbich and Bachtrog argue that genes with male-biased expression tend to be either lost or retroposed off the X chromosome. In *Drosophila miranda*, a neo-X chromosome was formed relatively recently in evolution, which shows extremely high levels of adaptation. This chromosome as well as sex chromosomes overall may exhibit evolutionary traits that shed light on the principles of genome evolution in general.

Barske and Capel also discuss the evolution of sex determination systems, but focus on the relationship between genetically and environmentally based signals that control sexual differentiation. While most everyone is

familiar with genetically based systems for sex determination, some fish and reptiles have a temperature-dependent mode of sex determination. While the mechanisms behind these two different modes of sex determination were thought to be quite distinct, more recent analyses suggest that there is actually a continuum of modes of sex determination between these two seemingly discrete categories, and highlight the need to expand our understanding of the molecular mechanisms of sex determination to a wider array of non-mammalian species.

Genome evolution occurs not just through changes in gene content and chromosomal location of genes, but also through changes in gene expression. Gene expression can now be measured on a genome-wide scale not only by array experiments, but also by large-scale RNA *in situ* hybridisation projects. Lecuyer and Tomancak describe a variety of different techniques that are being pioneered in *Drosophila*, in terms of both the generation of images and also their processing. Techniques are advancing in terms of spatial and temporal resolution, as well as annotation of the resulting images, which is crucial for their use for the broader community.

Large-scale RNA *in situ* hybridisation projects illustrate the tremendous variety of gene expression patterns in an organism like *Drosophila*. The intricate regulation of genes is achieved to a large degree by an interplay between the repertoire of transcription factors in the organism and their binding sites. Bonn and Furlong review *cis*-regulatory networks in Drosophila development, where considerable numbers of transcriptional regulatory interactions are known for the segmentation, dorsoventral patterning and mesoderm networks. These networks are significantly augmented by recent global ChIp-chip data, which reveal new principles regarding the organisation and dynamics of *cis*-regulatory modules. Furlong and colleagues have analysed the binding of a small number of transcription factors at several developmental timepoints, and find that their occupancy varies. This must depend on cooperativity with other factors, potentially including chromatin-level regulation.

Busser, Bulyk and Michelson also discuss developmental regulatory networks, including those in fly. They emphasise the hierarchical nature of developmental networks as a series of subnetworks at increasingly fine granularity. The interactions in networks can be reliably identified by integrating computational searches with data on binding motifs and genome-wide expression. There are several high-throughput approaches for the determination of recognition motifs of transcription factors, such as protein-binding microarrays, one-hybrid and SELEX. These approaches can be applied not only to fly development but also to embryonic stem cells, which depend on networks to maintain their undifferentiated state as well as their ability to self-renew and differentiate.

This is true not only of embryonic stem cells but also of haematopoietic stem cells, which differentiate into all the cells types of the lymphoid and myeloid lineages of haematopoiesis. The cell types of haematopoiesis are powerful models of differentiation, as they can be isolated experimentally using cell surface markers. Miranda-Saavedra and Gottgens mention that network motifs of positive feedforward loops are likely to be important in both embryonic and haematopoietic stem cells. They review both the classical 'bottom-up' approaches that identify individual molecules in transcriptional regulatory networks, as well as large-scale 'top-down' approaches that are starting to be applied to haematopoietic cell types.

The formation of neural crest cells represents another well-studied example of cell differentiation. The evolution of neural crest is tightly linked to many aspects of vertebrate development and evolution as these cells form a wide variety of derivatives including craniofacial cartilage, pigment cells and teeth, all of which have contributed to vertebrate diversification. Baker reviews our detailed understanding of the molecular circuitry that underlies neural crest formation and differentiation, and explores where these genetics networks originated and how they have been elaborated to give rise to new structures. These analyses have been greatly facilitated by recent studies in a diversity of taxa, such as lampreys, hagfish, amphioxus and ascidians that help us reconstruct the early evolutionary history of vertebrates.

Ravi and Venkatesh explore the genomic aspects of a remarkable radiation within the vertebrates; the evolution of teleost fish. There are about 27 000 species of teleost fish, and they display a remarkable diversity of morphology and behavioural adaptations. The authors discuss the implications of a number of events, including whole genome duplication, chromosomal rearrangements and the rapid rate of protein sequence divergence and divergence of putative *cis*-regulatory elements that may have contributed in various ways to this diversification. They also discuss how various hypotheses could be further tested by genomic sequencing of additional taxa whose lineages branched off near the base of the teleost radiation.

Within the teleosts exists one of the classic evolutionary models of rapid phenotypic evolution, the cichlid fish of East Africa. In less than 100,000 years, more than 500 species have arisen in Lake Victoria alone. Karaku and Meyer review recent progress in understanding possible mechanisms that underlie cichlid evolution and the current experimental approaches that are being pursued. This system has been particularly useful for understand-

ing the interplay between ecological habitats, adaptation and phenotypic evolution. The authors conclude that the cichlid system is well poised to provide general insights into evolutionary processes that apply to all organisms.

Reed, Papa and Martin review another classic example of evolution described in so many textbooks, namely the diversification of *Heliconius* butterflies and the genetics that underlies the remarkable mimicry systems that exists within these beautiful butterflies. Combining genetic and genomic approaches, researchers are rapidly homing in on the loci responsible for generating colour pattern differences within different *Heliconius* species, but have already come to the remarkable conclusion that homologous loci between different species underlie much of the variation that is seen, even when the colour patterns look quite different from one another. Further work is still needed to narrow down mapped loci to individual genes, but then it should be possible to finally understand how allelic variation can control such complex evolutionary differences.

Flowers and Purugganan provide a population genetics perspective on the evolution of genomes. It is important to keep in mind that the differences we see between species had their origins at the level of variation within a population, and thus are reflective of the genomic processes that go on at the population level. The authors focus on understanding plant genome evolution, although their approach and conclusions are relevant to all genomes. They discuss the extent to which population size, transposable elements and gene family diversification within populations can drive genome evolution. These types of population models lead to both the modification of existing theories, and new theories, as to the origin of genome diversification that we see between species.

It is known that genome variation can be driven by transcription factors and their binding sites, and now there is increasing evidence that chromatin-level and even higher order regulation also plays a role. Babu, Janga, de Santiago and Pombo review work that shows that regulatory elements cannot just be distal to a promoter, but even on a different chromosome. Such inter-chromosomal interactions and larger chromosome territories in the nucleus represent a new way of thinking about gene regulation. This impacts both evolution and development of organisms and genomes, from yeast to human.

References

1. Wilson D, Charoensawan V, Kummerfeld SK, Teichmann SA: **DBD-taxonomically broad transcription factor predictions: new content and functionality**. *Nucleic Acids Res* 2008, **36**:D88-D92.