

EDITORIAL

Evolutionary and ecological functional genomics

Heredity (2008) 100, 101–102; doi:10.1038/sj.hdy.6801015

Genomics, proteomics and related technologies are discovering remarkable variation in genes and their expression, but these discoveries are only a beginning. The next step is discovering how this variation affects the function and fitness of whole organisms in populations. Alongside the well-established communities in biomedical and agricultural functional genomics, a new community has formed: evolutionary and ecological functional genomics (EEFG), which seeks to understand the functional basis of evolutionary forces shaping ecologically important traits in natural biological communities. This requires elucidating the biotic and abiotic challenges to function in natural environments, the molecular and cellular mechanisms used to meet these challenges and the evolutionary mechanisms and processes that influence these mechanisms through time.

Earlier efforts in EEFG faced a fundamental conundrum: 'should EEFG emphasize the focal species of the model organism communities at the potential expense of ecological and evolutionary realism? Or should it ... instead establish new research communities around more ecologically interesting models, even if these models may be less tractable and the process involves delay?' (Feder and Mitchell-Olds, 2003). With changing technology and decreasing costs (Hall, 2007), the time is approaching when genomic tools can be applied to diverse non-model species, characterizing new levels of complexity in natural systems and enabling tests of fundamental hypotheses in ecology and evolution. When DNA sequencing is no longer a limiting factor, then other factors will limit our research. Advances in genomics will return our focus to phenotypes, biological mechanisms and environmental context.

Emerging models

Daphnia has been widely used in ecological studies and whole genome shotgun sequencing now provides the first draft assembly of a crustacean genome (<https://dgc.cgb.indiana.edu/display/daphnia/Introduction>). Eads and colleagues review ecological genomics in *Daphnia*, as well as efforts to integrate genetic, physiological and ecosystem-level responses in this ecological model. Responses of *Daphnia* to stressors such as pollutants, predators, hypoxia and parasites will increasingly be investigated using the genomic resources that have been developed over the last few years. Another ecological model whose genome is now being sequenced is Monkey flower (*Mimulus*), which has a long history in ecological and evolutionary research and encompasses enormous phenotypic, ecological and genomic diversity. The paper by Wu *et al.* provides an overview of the genomic resources currently available for *Mimulus*, and discusses future directions for research. In addition,

Beldade and colleagues discuss community efforts on genomics of butterflies, which provide diverse opportunities for multidisciplinary research on evolutionary processes influencing ecologically important traits. Genomic resources for these studies continues to expand, as part of community-wide efforts to extend ongoing functional genomic research.

Emerging approaches

Shiu and Borevitz review recent advances in microarray technology relevant to studies of EEFG. These tools will enable analyses of genome dynamics, novel functional elements in genomes, increasing understanding of regulatory networks, as well as characterization of microbial community complexity in environmental samples. In addition, Marden examines how alternative splicing of transcripts affects development and organismal phenotypes. Several new approaches incorporating quantitative analyses of alternative splicing will improve functional understanding of trait development, and also may identify functionally important nucleotide polymorphisms. Over the past few years, proteomics has begun to synthesize protein biochemistry, genome biology and bioinformatics, providing a systems-level integration approach that can enhance ecological studies. The paper by Karr introduces the field of proteomics, and considers advantages and disadvantages in comparison to other genomic approaches such as microarray and EST analyses. Among other applications, proteomics will provide new information on metapopulation biology and adaptive processes in nature.

At the level of microbial communities, Ward and colleagues discuss environmental genomics and the issue of microbial species and examine how genomic and metagenomic methods can identify ecologically distinct populations that occupy unique niches, respond uniquely to changing environmental conditions, and how this approach may clarify linkages between genomic variation and community function. In addition, Stinchcombe and Hoekstra examine recent advances and possible future directions combining population genomics and quantitative genetics in model and non-model systems. Taken together, these two approaches offer great potential for uncovering the molecular mechanisms responsible for adaptation.

Emerging issues

Felix and Wagner examine concepts of redundancy and robustness, taking examples from development in the nematode *Caenorhabditis elegans*. This discussion considers two prominent mechanistic causes of robustness, namely redundancy and distributed robustness. Fay and Wittkopp examine the role of natural selection in the evolution of gene regulation, and discuss the advantages and disadvantages of methods that have been applied to this question, with an emphasis on caveats imposed by the properties of regulatory networks.

Roy *et al.* discuss the evolution of allelic dimorphism in malarial surface antigens. In this system, the malaria parasite is under constant onslaught from the human

immune system, and parasite antigens show unmistakable evidence of strong selection. Studies show that the alleles encoding parasite antigens often form two allelic classes, with much more similarity within each class than between classes. Their discussion examines several mechanisms that may influence this unexpected pattern of variation.

Kramer and Kalisz discuss how causes of phenotypic variation and genetic constraints can be identified by applying genomic techniques including fine-mapped QTLs, reciprocal heterologous transformation and ecological genetic studies of model systems. The future challenge for plant evolutionary and developmental biologists is to construct a comprehensive evolutionary framework combining processes that simultaneously constrain and generate phenotypic and genetic variation. In addition, Ungerer *et al.* examine how applications of genomic approaches in ecology are beginning to provide interdisciplinary insights into biological phenomena that previously lacked a genetic context. They also consider various approaches and methodologies that will be useful for ecological genomic studies, and consider areas for future research.

Finally, Bangert *et al.* examine the genetic structure of a foundation species, and consider scaling of community phenotypes from an individual to regional scale. On the basis of long-term studies of poplars (*Populus*), they consider genetic variation in phytochemistry and tree architecture, which affect diverse organisms and ecosystem processes. In addition, they consider how community attributes may be influenced by an interaction between genetic polymorphism and environmental variation across geographic scales.

Bioinformatics and community databases

Progress in EEFG also requires continuing emphasis on bioinformatics. A growing number of evolutionary model organisms are becoming targets for large-scale mapping, EST, and genome sequencing projects, and building an informatics infrastructure for handling these data is an expensive, time-consuming and daunting prospect. The Generic Model Organism Database Project

(GMOD: <http://www.gmod.org>) (Stein *et al.*, 2002) is an organized effort among developers from Flybase, Grapeme, MGD (for mouse), SGD (for yeast), WormBase and other established database groups to provide a ready-made 'construction kit' for a genomic database. GMOD consists of a variety of matching open-source software components (a database schema, web-based graphical user interfaces, annotation software, and so on) and building upon best practices in database design and genomic data standards. GMOD provides a mature off-the-shelf solution for managing standard genomic data. The National Evolutionary Synthesis Center (NESCent, www.nescent.org) is committed to facilitating the adoption of GMOD within the evolutionary community through training and support, and also to further development of the GMOD toolkit for management of evolutionary data types (for example, information about geographic, phenotypic, and genetic variation and information about phylogenetic relationships). The considerable effort that would otherwise be spent reinventing the wheel with each new sequencing project can instead be devoted to addressing the scientific questions that make these models so compelling.

Acknowledgements

We thank Todd Vision for contributing to the discussion of model organisms databases.

T Mitchell-Olds¹, M Feder² and G Wray¹

¹Department of Biology, Duke University, Durham, NC, USA

and ²Department of Organismal Biology and Anatomy,

The University of Chicago, Chicago, IL, USA

References

- Feder M, Mitchell-Olds T (2003). Evolutionary and ecological functional genomics. *Nat Rev Genet* **4**: 651–657.
- Hall N (2007). Advanced sequencing technologies and their wider impact in microbiology. *J Exp Biol* **210**: 1518–1525.
- Stein LD, Mungall C, Shu S, Caudy M, Mangone M, Day A *et al.* (2002). The generic genome browser: a building block for a model organism system database. *Genome Res* **12**: 1599–1610.