

## Can ecology help genomics: the genome as ecosystem?

Rodney Mauricio

*Department of Genetics, Davison Life sciences complex, University of Georgia, Athens, GA 30602-7223, USA (Phone: (706)-542-1417; Fax: (706)-542-3910; E-mail: mauricio@uga.edu)*

Received 30 January 2002 Accepted 1 January 2004

*Key words:* autoecology, ecology, emergent properties, epigenetics, epistasis, genomics, microarrays, natural history, systems biology

### Abstract

Ecologists study the rules that govern processes influencing the distribution and abundance of organisms, particularly with respect to the interactions of organisms with their biotic and abiotic environments. Over the past decades, using a combination of sophisticated mathematical models and rigorous experiments, ecologists have made considerable progress in understanding the complex web of interactions that constitute an ecosystem. The field of genomics runs on a path parallel to ecology. Like ecology, genomicists seek to understand how each gene in the genome interacts with every other gene and how each gene interacts with multiple, environmental factors. Gene networks connect genes as complex as the ‘webs’ that connect the species in an ecosystem. In fact, genes exist in an ecosystem we call the genome. The genome as ecosystem is more than a metaphor – it serves as the conceptual foundation for an interdisciplinary approach to the study of complex systems characteristic of both genomics and ecology. Through the infusion of genomics into ecology and ecology into genomics both fields will gain fresh insight into the outstanding major questions of their disciplines.

### Introduction

Genomics has been described as the ultimate integrative discipline, crossing the full spectrum of the biological sciences. Without doubt, genomics is a multidisciplinary pursuit, combining primarily molecular biology and computer science. The genomics era has also brought a renewed interest in systems biology, conceptually a broader multidisciplinary endeavor, and said to bring together biology, chemistry, computer science, engineering, mathematics, and physics (Ideker et al., 2001; Kitano, 2002; Hood & Galas, 2003). Absent in these lists of the 21st century’s new biology is a mention of the field of ecology, the scientific study of the processes influencing the distribution and abundance of organisms, particularly with respect to the interactions of organisms with their biotic and abiotic environments.

This absence is surprising – surprising because both ecologists and genomicists ask similar questions, their respective disciplines have developed along similar intellectual trajectories and share basic epistemological approaches. In many ways, the genome and the ecosystem are parallel constructs and can be studied using similar approaches. The thesis of this paper is that including the field of ecology as part of the study of genomics will lead to advances in both disciplines.

### *A metaphor*

Imagine the Serengeti plain of east Africa: grasses, shrubs, and trees extend over the landscape; giraffe, elephants, and antelope graze over the grasslands; lions, leopards, and hyena hunt and scavenge; vultures, flies, and fungi linger over carrion. Over the past millennium, natural historians

have discovered and described these, and many other, individual species of plants, animals and microbes. Ecologists stepped in over a century ago to study what an individual species *does* in its environment, its 'autoecology'. In other words, we now know how a giraffe manages to live in the Serengeti. In the past century, through a combination of manipulative experiments and mathematical theory, ecologists have made great strides in understanding interactions between individual species (e.g., Wilbur, 1987; Morin, 1999). As a result, to a large degree, we now know how giraffes interact with trees, with other giraffes, with other herbivores, with predators, and even with dung beetles (Jankielsohn et al., 2000): a fairly complex network of interactions.

However, the challenge of ecology is not to understand only the giraffe's role in the Serengeti ecosystem: a complete ecological understanding of the Serengeti would require that we understand the rules regulating how *each* and *every* species in the ecosystem, from bacteria to lions, interacts with every other species and how each species interacts with multiple environmental factors. Needless to say, this is a complicated problem. It is made more complicated by the fact that complex systems are rarely the sum of their parts: emergent properties lead to nonlinearities. Considering the complexity of the problem, ecologists have made astonishing inroads into understanding the natural world, although some remain skeptical (e.g., O'Connor, 2000). Keep the metaphor of the giraffe in the Serengeti in mind as we consider how examination of another 'species' – the gene in its genomic ecosystem – may further accelerate breakthroughs in ecology and genomics.

#### *The metaphor extended: the genome as ecosystem*

Although the pace of intellectual development has been much more rapid in genomics, the parallels to the development of ecology are unmistakable. Like those legions of systematists identifying the individual species in the ecosystem, geneticists made a cottage industry of identifying single genes until the advent of whole-genome sequencing (and bench geneticists continue to make remarkable progress in carefully reconciling predicted genes with actual ones). In many ways, genomicists reintroduced natural history to biology, albeit a molecular natural history, eschewing hypothesis-driven research

and proclaiming a new phase of 'discovery-based' inquiry (Ideker et al., 2001) with the argument that the field needed to accumulate the basic information upon which hypotheses could later be based.

Like ecologists in the Serengeti, the mainstay of many modern molecular geneticists is attempting to understand the function, the autoecology, of each gene. For many pathways, we know how genes interact with other genes, like we know how giraffes interact with other giraffes or other animals. Molecular geneticists have long understood how genes interact with the environment. Genes live in an ecosystem like animals live in their ecosystem, and although the tools used to study genes and giraffes are clearly different, the broad intellectual approaches to understanding genes and giraffes are not so different.

However, like ecology, the ultimate challenge of genomics is to understand how each gene in the genome interacts with every other gene (epistasis) and how each gene interacts with multiple, environmental factors. Gene networks are just as complex as the 'web' that connects all the species in an ecosystem (Tong et al., 2004). Again, understanding that degree of complexity is a complicated, multidimensional problem. What emergent properties will arise from the complexities of the genome? Will understanding the function of every gene ever allow us to predict complex phenotypes? How pervasive are epigenetic effects (e.g., Waddington, 1942)?

If we see the genome as an ecosystem where genes live, how much more progress will genomicists make in understanding that ecosystem than ecologists have made in understanding their ecosystems? Regardless of the answer to that question, ecology and genomics do have enough to offer one another that the two disciplines may reach their common goal with a healthy interchange of ideas.

#### *What can ecology and genomics offer each other?*

Certainly molecular geneticists have offered ecologists a myriad of tools to understand ecology and in many ways those tools have revolutionized ecology. However, what does ecology offer genomics? The most important thing ecology can offer genomics is experience in simply thinking about, and being trained in thinking about, complex interactions. Most often, this training is

manifested in being able to design experiments that test for complex interactions with both the environment and other individuals or species (Hairston, 1989; Resetarits & Bernardo, 1998).

For example, both geneticists and ecologists use manipulative 'field' experiments. Molecular geneticists use knockout experiments (experimentally excluding genes from a pathway with, for example, targeted mutagenesis or RNAi) to understand how genes interact within the genome and ecologists often experimentally exclude a species from an ecosystem (e.g., with a fence or pesticide) in order to understand the role of that species in the ecosystem. Since ecologists often manipulate multiple species in a factorial fashion, statistical and experimental approaches have been developed that allow for the analysis and interpretation of these data. Most molecular geneticists have tested single mutant, double mutants, and even triple mutants, but it gets exceedingly difficult to examine the factorial effects of every possible combination of four or more independent mutations. Genomics allows the investigator the opportunity to examine the global effects of mutants, but the statistical interpretation of such experiments often clouds the results. The ecologists' experience in designing experiments with an eye towards managing complexity will be directly applicable to the analysis of complex genomic datasets.

For example, many microarray experiments suffer from simple but significant flaws in design that make the data difficult to interpret (Tilstone, 2003). Technical problems arise that could be addressed simply by borrowing concepts from ecology. For example, the slides used for microarrays can sag, causing an attenuation of signal for those spots in the middle. Engineers have worked to improve the physical properties of the slides and computer scientists have worked to account for the signal attenuation. However, ecologists must always account for heterogeneity in their field sites and use a variety of experimental techniques to do so (Cochran & Cox, 1992; Scheiner & Gurevitch, 2001). The simplest field technique, 'spatial blocking,' is easily applied to a microarray (although at a cost of throughput). Rather than apply 10,000 unique spots on a chip, one could spot four replicates of each oligonucleotide or mRNA in distinct spatial blocks on a chip. A simple analysis of variance could

account for the variation due to physical heterogeneity on the slide, whatever the underlying cause.

Beyond providing guidance in experimental design, ecologists can contribute a nuanced approach to studying the interactions of genes with the environment that goes beyond simple microarray gene expression studies done in a few different environments. For example, an investigation of mutant phenotypes performed under realistic ecological conditions could be valuable in shedding light on the 'genetic uncertainty principle' where a reverse genetics approach has not yielded an informative mutant phenotype (Tautz, 2000). The failure of a gene knockout to produce a visible phenotype could be due to genetic redundancy, but it could also be masked by the permissive environments in which most mutants are screened (Gilliland et al., 1998; Meagher et al., 2000).

In addition to being an experimental science, ecology is also a highly mathematical discipline. While some cell and molecular biologists have employed complex mathematics in their work, there remains an enormous potential in the synergy between the kind of datasets genomicists generate and the mathematical approaches that ecologists have refined over the last century. Very simple mathematical models were derived early in the history of ecology to predict population growth (logistic equation) and to study interactions among species (Lotka-Volterra equation). Today, ecology has developed a firm mathematical foundation (May 1976; Dieckmann et al., 2000; May, 2001; Okubo & Levin, 2001; Cushing et al., 2002). Mathematics is an essential tool to understanding complex systems. Models are used to generate hypotheses that can be experimentally tested. For example, a model of a complex network can be generated, along with a predicted response to a perturbation. Perturbation experiments can be performed and the observed results compared with the model. Mathematics will be essential to guide the course of experimentation in genomics as the complexity of systems increases. When applied to genomics, these models will focus in detail on the specific molecular mechanisms of individual genes and proteins and their interactions. Further models could explicitly incorporate deterministic environmental parameters as well as environmental stochasticity.

This approach has been recently advocated by systems biologists who favor an applied mathematics and computational approach to biology (Hood & Galas, 2003). Further evidence of the common path taken by ecology and genomics lies in the recent establishment of systems biology as an intellectual discipline. Systems biology has an antecedent in systems ecology. Systems ecology is a branch of ecology that attempts to understand the structure and function of ecosystems by concentrating on energy inputs and outputs of the system (Odum, 1983; Patten & Jørgensen, 1995). Systems ecology was developed partly as a way to confront the complexity of systems. The system itself is a black box and the approach trades off the ability to understand the details of the components of the system for understanding the system as a whole. Whether systems biologists embrace a deep systems approach or if they simply apply mathematics to molecular biology at a global scale (Ideker et al., 2001), the path of modern biology will be paved with mathematics; and ecologists have been strolling that way for decades (May, 1976).

Ecologists clearly have something to offer to genomics, but genomics will continue to be critical to advances in ecology. Certainly, techniques created for genomics have found application in ecology. Craig Venter's attempt to use sequencing to identify every microbe in the Sargasso Sea is an example of the power of genomics to identify all the players in a complex ecosystem. And ecologists have started using some of the tools of genomics in their own work (Jackson et al., 2002). Nevertheless, genomics could have an even more profound intellectual contribution to ecology. As physics infused ecology in the 1970s, a focused interest on the ecology of the genome may give great insight into biological systems at higher levels of organization. For example, perhaps gene networks are, at some level, fundamentally different from food webs. The present research interest in genetic networks could have substantial application to ecologists' work on species interactions (e.g., Barkai & Leibler 1997; Bergman & Siegal 2003). Genetic systems, like ecological systems, seem to be more stable the more connected they are. Although this result makes some intuitive sense in a genetic system, it is unclear why it seems to be the case in ecological systems. For many questions, modeling the genome as an ecosystem will have direct

applications to understanding any complex system, including ecosystems.

### *Final thoughts*

In this paper, I have attempted to outline some of the common approaches that genomics and ecology have taken to addressing the outstanding questions in their disciplines. I see unmistakable similarities in these two seemingly disparate fields. It strikes me that both ecology and genomics have much to offer each other. And since genomics is still in many ways establishing its paradigms, now seems the appropriate time for each field to take full advantage of the others' strengths. Will the infusion of ecological ideas into genomics help to make more sense of genomes than we presently have of ecosystems? Will a new synthesis of ecology and genomics lead us into this new century of biology? I do not know. But if I were a beginning graduate student in genetics, I would look at the course offerings in math. If I were a beginning ecology or math graduate student, I would look over at what the geneticists were doing. And if I were hiring systems biologists, I would take a careful look at ecologists.

### **Acknowledgments**

I greatly appreciate my friend (and molecular geneticist) Greg Copenhaver's indulgence over the years: his comments on earlier versions of the manuscript were welcome. He has always been a welcome sounding board. Amy Bouck also made thoughtful comments on an earlier draft of this manuscript.

### **References**

- Barkai, N. & S. Leibler, 1997. Robustness in simple biochemical networks. *Nature* 387: 913–917.
- Bergman, A. & M.L. Siegal, 2003. Evolutionary capacitance as a general feature of complex gene networks. *Nature* 424: 549–552.
- Cochran, W.G. & G.M. Cox, 1992. *Experimental Designs*. Wiley & Sons, New York, 611 pp.
- Cushing, J.M.R., Costantino, B., Dennis, R., Desharnais, & S. Henson (eds.), 2002. *Chaos in Ecology: Experimental Nonlinear Dynamics*. Academic Press, New York. 225 pp.
- Dieckmann, U., R. Law & J.A.J. Metz (eds.), 2000. *The Geometry of Ecological Interactions: Simplifying Spatial*

- Complexity. Cambridge University Press, Cambridge. 580 pp.
- Gilliland, L.U., E.C. McKinney, M.A. Asmussen & R.B. Meagher, 1998. Detection of deleterious genotypes in multigenerational studies. I. disruptions in individual *Ara-bidopsis* actin genes. *Genetics* 149: 717–725.
- Hairston, N.G., Sr., 1989. *Ecological Experiments: Purpose, Design and Execution*. Cambridge University Press, Cambridge. 384 pp.
- Hood, L. & D. Galas, 2003. The digital code of DNA. *Nature* 421: 444–448.
- Ideker, T., T. Galitski & L. Hood, 2001. A new approach to decoding life: Systems biology. *Annu. Rev. Genomics Hum. Genet.* 2: 343–372.
- Jackson, R.B., C.R. Linder, M. Lynch, M. Purugganan, S. Somerville & S.S. Thayer, 2002. Linking molecular insight and ecological research. *Trends Ecol. Evol.* 17: 409–414.
- Jankielsohn, A., C.H. Scholtz & S.V. Louw. 2000. Effect of habitat transformation on dung beetle assemblages: a comparison between a South African nature reserve and neighboring farms. *Environ. Entomol.* 30: 474–483.
- Kitano, H., 2002. Systems biology: a brief overview. *Science* 295: 1662–1664.
- May, R.M., 1976. *Theoretical Ecology: Principles and Applications*. W. B. Saunders, New York. 317 pp.
- May, R.M., 2001. *Stability and Complexity in Model Ecosystems*. Princeton University Press, Princeton, New Jersey. 292 pp.
- Meagher, S., D.J. Penn & W.K. Potts, 2000. Male–male competition magnifies inbreeding depression in wild house mice. *Proc. Natl. Acad. Sci. USA* 97: 3324–3329.
- Morin, P.J., 1999. *Community Ecology*. Blackwell Science, Malden, MA. 424 pp.
- O'Connor, R.J., 2000. Why ecology lags behind biology. *The Scientist* 14(20): 35.
- Odum, H.T., 1983. *Systems Ecology: An Introduction*. Wiley and Sons, New York, 644 pp.
- Okubo, A. & S.A. Levin (eds.), 2001. *Diffusion and Ecological Problems*. Springer Verlag, New York. 488 pp.
- Patten, B.C. & S.E. Jørgensen (eds.), 1995. *Complex Ecology. The Part-Whole Relation in Ecosystems*. Prentice Hall, Englewood Cliffs, New Jersey. 705 pp.
- Resetarits, W.J. & J. Bernardo (eds.), 1998. *Experimental Ecology: Issues and Perspectives*. Oxford University Press, New York. 488 pp.
- Scheiner, S.M. & J. Gurevitch (eds.), 2001. *Design and Analysis of Ecological Experiments*. Oxford University Press, New York. 415 pp.
- Tautz, D., 2000. A genetic uncertainty problem. *Trends Genet.* 16: 475–477.
- Tilstone, C., 2003. Vital statistics. *Nature* 424: 610–612.
- Tong, A. H. Y. et al., 2004. Global mapping of the yeast genetic interaction network. *Science* 303: 808–813.
- Waddington, C.H., 1942. The epigenotype. *Endeavour* 1: 18–20.
- Wilbur, H.M., 1987. Regulation of structure in complex systems: experimental temporary pond communities. *Ecology* 68:1437–1452.