

Testing hypotheses regarding the genetics of adaptation

Patrick C. Phillips

*Center for Ecology and Evolutionary Biology, University of Oregon, Eugene, OR 97403-5289 USA
(Phone: (541) 346-0916; Fax: (541) 346-2364; E-mail: pphil@uoregon.edu)*

Received 20 December 2003 Accepted 15 February 2004

Key words: association mapping, evolution, genomics, microarrays, QTL mapping, quantitative genetics, statistical genetics

Abstract

Many of the hypotheses regarding the genetics of adaptation require that one know specific details about the genetic basis of complex traits, such as the number and effects of the loci involved. Developments in molecular biology have made it possible to create relatively dense maps of markers that can potentially be used to map genes underlying specific traits. However, there are a number of reasons to doubt that such mapping will provide the level of resolution necessary to specifically address many evolutionary questions. Moreover, evolutionary change is built upon the substitution of individual mutations, many of which may now be cosegregating in the same allele. In order for this developing area not to become a mirage that traps the efforts of an entire field, the genetic dissection of adaptive traits should be conducted within a strict hypothesis-testing framework and within systems that promise a reasonable chance of identifying the specific genetic changes of interest. Continuing advances in molecular technology may lead the way here, but some form of genetic testing is likely to be forever required.

Introduction

How should we view historical developments in evolutionary genetics through the particular lens of the genetics of adaptation? Although it is perhaps a bit premature for such pronouncements, one could argue that we are entering a new era of modern evolutionary genetics. The first era, roughly from 1918–1968, was characterized by the theoretical developments in population and quantitative genetics that have laid the foundation for nearly all other work in evolutionary biology (Figure 1, see also Provine, 1971). This period began with the theoretical reconciliation of quantitative and Mendelian genetics by R.A. Fisher (1918) and rapidly expanded into the codification of population genetics theory in the 1920's and 1930's through the work of Fisher, Sewall Wright and J.B.S. Haldane. It runs on through the beginnings of ecological genetics by the likes of E.B. Ford and others and the application of population genetic

principles to natural populations led by Theodosius Dobzhansky. It ends with a formalization of earlier models by Gustave Malécot and Motoo Kimura into a framework that set the stage for the utilization of the truly genetic data that was soon to follow (Lewontin, 1974). This period could be classified as theory rich and data poor. Most of the theory that we still utilize today was established before we had any knowledge of the nature of the genetic material, and in this sense these approaches are essentially purely genetic and largely devoid of functional context. Fundamental concepts of genetic entities like loci and alleles have hardly changed in population genetics theory, despite tremendous advances in our knowledge of the physical and molecular properties of genes and genomes.

The second era, from 1968 to 1998, was dominated by an explosion of data, frequently collected in the absence of a compelling theoretical context (Lewontin, 1991). In population genetics, the development of protein electrophoresis

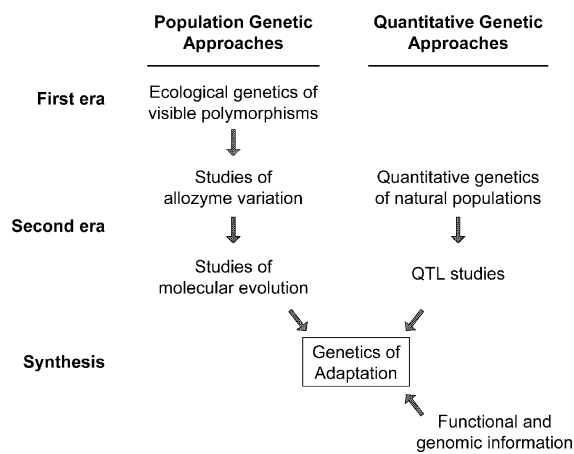


Figure 1. Transitions during the history of population and quantitative genetic approaches to studying the genetics of adaptation. Movement toward a new era of study incorporates these approaches with functional genomic information.

allowed researchers to assess levels of genetic variation in a wide variety of organisms rather than being limited to special cases of known genetic markers (e.g., *Drosophila* chromosomes) or obviously Mendelizing phenotypes (e.g., snail shell polymorphisms). On the quantitative genetic side of things, the theories originally developed by Fisher and greatly expanded by Wright were finally migrated from agricultural systems into a more formal theory of evolutionary quantitative genetics (e.g., Slatkin, 1970; Lande, 1976; Felsenstein, 1977). Here again, researchers could venture into natural populations to ask questions about levels of genetic variation for ecologically important traits. It seemed that no study of the evolutionary ecology of quantitative traits could be complete without an analysis of underlying genetic variation, because evolutionary change is predicated on its existence. To some extent, both the population and quantitative genetic approaches were victims of their own success. Electrophoretic studies revealed ample levels of genetic variation at most loci, while quantitative genetic studies found significant heritability for most traits. Finding genetic variation for its own sake became a hypothesis-free endeavor. Enough studies of this type have now been performed that one need not actually conduct the studies to know their probable outcome. For the most part, average heterozygosity will vary between 0.05 and 0.2 and heritability will fall somewhere between 0.2 and 0.5. Even if a particular estimate were off by

a factor of two or three, would the discussion sections of these particular studies be very different? It is unlikely that they would, which is a testament both to a general lack of precision in these estimates and the lack of a broader hypothesis-testing framework for this work.

Studies of variation per se have developed on one side into much more sophisticated treatments of DNA sequence variation from a molecular evolution viewpoint and on the other side into a formal theory of evolutionary quantitative genetics that treats the entire organism as an integrated whole (Figure 1). Using sequence data, we can address very specific hypotheses regarding historical patterns of selection and rates of evolution of genes of interest, but are frequently far removed from the how, why, what, and where of the adaptive context of that selection. In contrast, in multivariate views of quantitative inheritance, we can measure how selection operates on suites of traits and how trade-offs among traits might structure and constrain the response to selection (Lande, 1988), but are limited to some extent by complexities introduced by the total dimensionality of the system (Charlesworth, 1990) and by the fact that, in order to understand how summary parameters like genetic correlations themselves evolve, we need to have much greater knowledge of the genetic systems underlying these traits (Barton & Turelli, 1989). We are caught between molecular knowledge in the absence of adaptive context and ecological context in the absence of molecular details. One view of the modern challenge to understanding the genetics of adaptation is the need to span this chasm – to be able to move freely from sequence to phenotype to ecological context and, more importantly, to be able to test specific hypotheses at each of these levels.

Are we, then, at the beginning of a self-proclaimed new era? If so, then it is an era that is sure to be dominated by genomic analysis (the 1998 date was chosen because of the publication of the first metazoan genome during this year, The *C. elegans* Sequencing Consortium, 1998). The hope is to use our new abilities to look at genome-wide patterns of genetic variation and gene function to investigate the genetics of adaptation from multiple perspectives. The fear is that we instead will repeat the mistakes of previous technological transitions and collect information in the absence of definitive hypothesis tests; or worse,