
CHARACTER IDENTIFICATION: THE ROLE OF THE ORGANISM*

GÜNTER P. WAGNER¹ AND MANFRED D. LAUBICHLER²

¹*Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520*

²*Program in History of Science, Princeton University, Princeton, NJ 08544*

INTRODUCTION

Some areas of organismic biology are among the mathematically most sophisticated parts of biology. Examples include population ecology, behavioral biology (game theory, ESS), life history theory, and population and quantitative genetics (e.g., Crow and Kimura, 1970; Bulmer, 1980; Berryman, 1981; Hofbauer and Sigmund, 1988; Stearns, 1992; Falconer and Mackay, 1996; Hanski and Gilpin, 1997; Hartl, 1997; Hofbauer and Sigmund, 1998; Maynard-Smith, 1998). Nevertheless, many of these theories do not have the degree of rigor that is typical of other mathematicized sciences such as physics and chemistry. Here we argue that this lack of rigor is due to a shared structural weakness of these biological theories, and that to overcome this deficiency requires in many instances an organismic perspective in character identification.

* Reprinted from *Theory in Biosciences*, Vol. 119 (2000) pp. 20–40

The structural weakness of many mathematical theories in biology is due to the lack of an analytical theory about their range of application, i.e., an analytical method that would allow us to identify and measure those objects in nature that correspond to the ones postulated by the theory. In the case of Newtonian mechanics it is quite clear that the mathematical models only apply to objects that can be characterized by their mass, location in space, and velocity. Physical theories are as powerful as they are, in part because they implicitly contain a definition of the conditions of their validity. When applying physical theories to material phenomena, the required ontological commitments on part of the researcher are rather limited. One only has to assume that the objects described by the theory in fact exist, or at least that these theoretical objects are close enough abstractions of some observed natural phenomenon. In addition, one has to establish an equivalence relation between the mathematical operations of the theory and the natural processes they describe (Rosen, 1962, 1978).

In contrast, the application of mathematical models in biology has been largely ad hoc. In our biological models we postulate the existence of certain theoretical objects, such as populations, genes, or life history characters, but we usually do not reflect too much on the question which properties a biological object must have in order to be a legitimate instance of the model. As a result, few, if any, biological models can be rigorously tested, since it is often not clear whether any empirical result is actually a relevant instance of the model to be tested. In testing biological theories empirically we are thus prone to commit both, Type I and Type II errors (Sokal and Rohlf, 1981). In other words, we might either falsely reject or accept a model.

A particularly striking example of this problem is illustrated by the work of William Murdoch on population regulation in red scale mites in Californian orchards (Murdoch *et al.*, 1996). Two different kinds of models have been proposed to explain population regulation in such situations: the refuge model (Bailey *et al.*, 1962) and a model based on metapopulation dynamics (Gilpin and Hanski, 1991). Murdoch and his collaborators set out to test population regulation empirically. When they started their experiments, it was initially not clear whether the population size variable refers to the number of individuals in an orchard, as suggested by the metapopulation model, or to the number of individuals on a tree, as would be predicted by the refuge model. The experimental results showed that neither is the case, but that the population size variable, i.e., the functional unit that exhibits the phenomenon of population regulation (the dynamical unit), is the population of red scales in the outer branches of a tree. Neither the removal of the refuge population nor the isolation of an individual tree in a cage (to test for metapopulation effects) yielded the effects predicted by either model. Obviously, the problem with both models is

that they do not contain a prescription how to find the limits of the unit to which they actually apply, a problem that has so far been overlooked because both models have rarely been tested in the field. However, as Murdoch points out, the fact that they did not find evidence for refuge or metapopulation effects in their experimental setting does not rule out that these dynamics can exist at another scale (which would be a Type I error). We just do not have any straightforward criteria how to identify the appropriate dynamical units in a given experimental setting.

Theories about the adaptive evolution of organismic properties, such as life history theory, are another example for the ad hoc nature of many models in biology. These theories represent a sophisticated mathematical framework developed to explain the origin and maintenance of life cycle characters in plants and animals (humans included). An ongoing problem with many life history models, however, is that model predictions and experimental results frequently disagree. In most cases these discrepancies are caused by an inappropriate identification of the relevant organismal traits (Stearns, 1992). As a result, the existing models are frequently patched up with ad hoc assumptions about "interactions" among the characters and "constraints" on variation. However, all that "interactions" and "constraints" indicate is that the way the objects were measured in the experiment did not correspond to the units considered in the model. A rigorous theory of life histories would therefore require an operational definition of the biological (i.e., dynamical) properties that are expected of the characters in the model. In addition, it would also have to include methods how to identify these characters in the real world, or at least a procedure of testing whether or not the conditions for the validity of the model have been met experimentally.

The only partial exception to this structural weakness of mathematical models in biology is the theory of natural selection. In the extensive literature about the units of selection one finds explicit definitions of the dynamical properties a class of objects must have in order to be a legitimate instance of the theory of natural selection. Despite continuing controversy about the relative importance of different units of selection (gene, individual, group, etc.), this is one area of biology where mathematical and empirical research have contributed equally to our understanding of the subject (e.g., Fisher, 1930; Dawkins, 1982; Sober, 1984; Endler, 1986; Brandon, 1990; Nagylaki, 1992; Stearns, 1992; Mitton, 1997).

In general, all productive theories should contain at least implicit definitions of the set of objects to which they apply. In biology, examples of such objects are genes, cells, characters, organisms, and populations. Unfortunately, in our empirical studies we are often guided more by a colloquial

and less so by a rigorous technical interpretation of these concepts. Consequently, the problems we described earlier are endemic in biology. In this paper we want to investigate a specific question: how do we define the "right kind of object" in the light of a specific process that we want to understand and given that we have the nucleus of a theory to explain this phenomenon.

In many cases, and in particular in the most problematic ones, the theories we are concerned with refer to suborganismal objects, such as genes or cellular and organismal characters (traits). The relationship of these objects to the individual organism can be of one of two kinds: (i) the organism can be thought of as a composite entity "made up" of its traits and characters or (ii) the traits can be thought of as (conceptual) abstractions of the organism. These two scenarios differ as to which object—part or whole—is ontologically prior. In the first case the characters or parts are ontologically prior to the higher level object or the organism. The relationship between atoms and molecules is of this kind. Molecules are composed of atoms and many of their molecular properties can be derived from the properties of their atomic parts. In the second case the higher level unit is ontologically prior. In this instance the suborganismal objects (characters) are defined as conceptual abstractions of a higher level integrated whole and thus ontologically secondary. Here we argue that most biological objects at the suborganismal scale are of the second kind. In other words, we assume the ontological primacy of organisms and derive the objects relevant to the theory, i.e., the biological characters, by means of a conceptual decomposition of the organism. We further argue that if we define biological characters as conceptual abstractions, we also need an appropriately defined organism concept within biological theory.

BIOLOGICAL EXAMPLES FOR THE ONTOLOGICAL PRIMACY OF HIGHER LEVEL OBJECTS

In the previous section we stated that the organism as a whole is ontologically prior to its composite parts. Here we will give additional examples of the ontological primacy of higher level entities from other areas of biology. We thus argue that the ontological primacy of higher level entities is a general principle that characterizes most (if not all) of biology. For a systematic discussion of these issues and the related question of emergence we refer to the book by Maynard-Smith and Szathmáry (1995).

Arguably the cell is the most fundamental biological unit. The cell is also the unit of ontological primacy in molecular biology. The subject matter of molecular biology is the investigation of the molecular mechanisms that underlie the fundamental processes of life, such as DNA replication, protein

synthesis, regulation of gene expression, cross-membrane transport, metabolic pathways, and intracellular communication. All these processes take place within or between cells. Furthermore, they are both enabled and constrained by this cellular milieu. The cellular context not only guarantees the specific physicochemical and spatial conditions that are required by these highly specific chemical interactions, it is also the level at which the *functional* roles of these processes can be assigned. The specific nonmolecular, i.e., *functional*, characterizations—*messenger* RNA, *metabolic* enzymes, *transcription* factors, *transport* molecules, etc.—of all objects in molecular biology represent exactly the kind of abstractions that are derived by means of a theory-guided conceptual decomposition of a higher level unit, the cell, that is ontologically prior to its component parts. Molecular biology therefore differs from biochemistry in that it does not study the interactions between these molecules *qua* molecules, but investigates them within the functional context of a cell. In this sense the level of the cell can be said to be ontologically prior to its component parts.

Similarly, a population is a biological object that is ontologically prior to its functionally characterized traits. Any functional concept in population biology, such as refuge, migration balance, predator-prey relation, and demographic pattern, is a conceptual abstraction that presupposes the reality of a population and therefore its ontological primacy. The same is also true for the functional concepts in quantitative and population genetics. Notions, such as the average effect of an allele, used to identify theoretically relevant properties of lower level objects, such as alleles, are all population specific.

In our view, these examples, although generally known to biologists, deserve further attention in that they point to a common feature of biological objects, one that, in our opinion, should also be reflected by the structure of biological concepts and theories. The objects and functional characters that are at the center of every biological theory are not just *found objects* waiting to be collected in nature, rather they are conceptual abstractions determined by the parameters of a specific theoretical interest (Griffiths, 1997; Laubichler, 1997b; Wagner, 1997; Griesemer, 2000; Laubichler and Wagner, 2000a,b; Wimsatt, 2000). As such, they can only be identified within the context of the object that is ontologically prior to them. Therefore, a precise formulation of these objects of ontological primacy has to be part of every theory as well.

The organism concept plays a central role in many biological theories, even though it is rarely defined explicitly. The reason for the centrality of the organism concept in biology lies in the fact that a large number of biological phenomena only occur in an organismal context. Biological characters or traits simply do not exist independent of the organism in the way atoms can (and do) exist outside of molecules. It is therefore rather surprising that we currently do

not have an analytical concept of the organism that would reflect its importance. The following section discusses the role of the organism with regard to the problem of character identification in evolutionary biology but first we will briefly outline the significance of the organism as a unit of integration in the context of biological theories.

THE ANALYTICAL ROLE OF THE ORGANISM CONCEPT WITHIN A THEORY OF BIOLOGICAL OBJECTS

The organism is a privileged biological object that deserves special attention. Living objects face two fundamental challenges: (i) to maintain their integrity in the face of a changing internal and external environment and (ii) to assure their propagation in the face of inevitable death. Self-maintenance and reproduction are thus the two most crucial biological processes. Among present-day biological objects, the cell is the minimum level of complexity that is capable of both reproduction (cell division) and self-maintenance (the maintenance of an internal cellular milieu.) Accordingly, the cell is a privileged biological object with respect to self-maintenance and reproduction. Here we restrict our treatment to present-day biological objects and do not investigate any objects that might have been important during the early stages of life's history on earth (Miller and Orgel, 1974; Gesteland *et al.*, 1999).

Organisms are also capable of self-maintenance and reproduction. Furthermore, they are composed of cells. We have seen before that the cell is the object of ontological primacy for many of the processes studied by molecular biology. The crucial next question is, when do we need to refer to the organism in order to explain a biological phenomenon because it cannot be explained on any other level of complexity?

A situation where this question becomes relevant is when one asks what is the difference between a multicellular animal, a so-called metazoan, and a colony of cells? What is added that makes metazoans so distinctly organismic compared to a mere aggregation of cells? Both the colony as well as the metazoan organism are multicellular and in both there is a certain degree of cooperation between the cells. The metazoan animal, however, is considered an organism, while a colony is put in a separate category. A possible answer to this question can be found when one asks what organizational features are necessary in a metazoan animal in order to realize, at the multicellular level, the functions that are carried out by the cell in single celled organisms. A necessary function of cellular organization is the maintenance of the chemical gradient between the inside and the outside of the cell. In the cell a necessary component of this function is the cell membrane, which separates the cytosole from the outside and

regulates the traffic of molecules between inside and outside with the aid of receptor and channel molecules. We propose that the epithelial surface, which completely surrounds the body cavity of all metazoans and actively regulates the exchange of molecules between the body cavity and the outside, fulfills the same function in metazoan organisms. In contrast, cell colonies and mesozoan organisms like sponges are unable to regulate the chemical composition of their inside. In fact, there is no well-defined inside or body cavity in these animals. We thus propose to define metazoans as animals that have the potential to regulate the chemical composition of their body cavity through the activity of an epithelial body surface. Note that the gut and internal genital organs like the uterus are topologically part of the body surface and that the gut lumen and the uterine cavity are spaces external to the body cavity. Regulation of the chemistry of intercellular spaces through epithelial covers is an example of an emergent biological phenomenon that cannot be explained at any level of complexity other than the organism. In addition, this view helps identify the stages of evolution at which a new organismal level of organization has been established, with all its consequences for the further differentiation of this group of organisms.

In the case of multicellular organisms reproduction and self-maintenance are also more elaborate processes; reproduction frequently involves highly specialized organs and behavioral patterns, and self-maintenance is achieved through a variety of different cellular, physiological, and behavioral mechanisms that utilize highly complex structures, such as the nervous and endocrine systems. In addition, multicellularity requires a whole new kind of biological process—development.

Higher organisms have to be rebuilt in every generation. During development the potential conflict between the competing interests of cells and the organism as a whole becomes apparent. The functional and structural integration of the organism requires the specialization of cells, which in turn leads to the loss of their potential for independent cellular reproduction, a phenomenon already recognized by Weismann in his distinction between the germ plasm and the somatoplasm (Weismann, 1892). As a result of this complex integration, the organism itself becomes the privileged biological object with respect to the processes of reproduction and self-maintenance. Leo Buss described this process as the "evolution of individuality" and suggested that many intricate features of development can be understood as direct consequences of this conflict of interest between individual cells and integrated multicellular organisms (Buss, 1987). In cancer we see the effects of a reversal of the organism's control over the reproductive interests of its cells (see, e.g., Weinberg, 1998).

The principle of functional integration is not confined to higher organisms. It is by now standard wisdom in biology that the eukaryotic cell is the result of a symbiotic integration between two or more formerly independent prokaryotic organisms (Margulis, 1970, 1981). In addition, Maynard-Smith and Szathmáry (1995) suggest that life's history on earth is marked by a series of "major transitions in evolution" that have led to the emergence of newly integrated levels of complexity (see also Margulis and Fester, 1991). These phenomena raise important technical and conceptual problems, such as the question under what conditions does the integrated "whole" gain functional precedence over its composite parts, i.e., when can we say that the integrated "whole" acts as a biological individual in a specific biological process? Within the context of an integrated higher organism certain parts, such as organs or organ systems, acquire the role of functional biological characters. They are involved in specific biological processes, such as reproduction and self-maintenance. However, their very existence depends on the ontologically prior object, the organism. We identify these biological characters by means of a theory-guided conceptual abstraction. We also want to note that defining biological character through conceptual abstraction is formally equivalent to the mathematical problem of subsystem decomposition (Klir, 1985).

In theoretical biology the organism concept functions as a unit of integration. Units of integration, such as the cell and the organism, are ontologically prior objects in the context of specific biological processes. The qualifier "in the context of specific biological processes" is important because not every biological phenomenon can be understood at the same level of complexity. Finding the appropriate level of description is ultimately an empirical question. Privileging any particular level of description *a priori* is bad metaphysics and even worse science (Brandon, 1995). In the biological hierarchy the role of any particular object can shift relative to the specific process under investigation. The cell can be the liminal reference frame, i.e., the object of ontological primacy, for the process of DNA replication or gene transcription, but it can be a functionally characterized part (character) in the context of an organismal process, such as sexual reproduction. Similarly, the organism can act either as an integrated whole or it can be a part within a larger object, such as a population. Our success in developing consistent biological theories that explain central biological processes ultimately depends on our ability to establish the appropriate level of description and to define the range of objects that are legitimate instances of the theory.

The following section discusses two concrete examples—the questions of quantitative inheritance and adaptive evolution—of how functionally relevant

characters can be identified in the context of a mathematical theory representing a specific biological process.

CONTEXT-DEPENDENT UNITS OF SELECTION AS AN EXAMPLE OF FUNCTIONAL CHARACTER DECOMPOSITION

So far we have argued for the need of an analytical method that would allow us to identify the range of objects a given biological theory applies to, but we have not given any concrete examples of such a method. In this section we discuss how recent work on the unit of selection question can help to illustrate the decomposition problem and the related question of the individualization of specific characters (Laubichler, 1997a; Wagner *et al.*, 1998; Laubichler and Wagner, 2000b). The unit of selection question is particularly well suited for this purpose because the process of natural selection is among the best understood biological phenomena (Darwin, 1859; Fisher, 1930; Williams, 1966; Dawkins, 1976; Sober, 1984; Endler, 1986; Grant, 1989).

Traditionally two different units of selection have been recognized: the replicator and the interactor (Dawkins, 1976; Hull, 1980). Both concepts are defined to capture important aspects of the process of natural selection, accurate replication of an entity type and selective interaction with the environment, respectively. In addition, Brandon (1982) introduced the conceptually important distinction between units and levels of selection. Here we analyze the role of a different kind of unit of selection, one that can be defined as a *unit of inheritance* (Laubichler, 1997a; Wagner, Laubichler *et al.*, 1998; Laubichler and Wagner, 2000b) for technical details, but see also Wimsatt (1980) and Lloyd (1988) for further discussions of these issues.

The concept of a unit of inheritance refers to objects that contribute to the additive genetic variance of a character in a population. Consequently, these objects play an important role in the process of natural selection. The additive genetic variance is defined as the fraction of the total phenotypic variance that determines the heritability of a phenotypic (quantitative) character and therefore its ability to respond to natural selection. It can be expressed as the variance of the breeding values in a population which in turn are determined by the average effects of all the alleles in question (Falconer and Mackay, 1996; Lynch and Walsh, 1998). The technical reason why the additive genetic variance of a population has to be expressed in the form of the average effects of alleles lies in the fact that, in the case of sexual reproduction, genotypes do not cross the generational boundary intact, only genes do. In each generation, the gametes combine to reassemble the genotypes. Therefore, the breeding value of any

particular genotype depends on the values of the individual genes it carries, in other words their average effects.

This is the reason why genes are privileged units of inheritance, but are they the only ones that matter? This is ultimately an empirical question. In order to address this problem we have to develop a conceptual framework that would enable us to ask this question in the first place. Over the last decades quantitative genetics has been biased toward the assumption that most problems can be approximated by a theory based on additive effects of genes (Bulmer, 1980; Falconer and Mackay, 1996). One aspect of this view is the idea that the single locus paradigm for defining average effects and breeding values can easily be extended to a multilocus case. It is simply assumed that the average effects of all alleles on a gamete add up to the average effect of each gamete and the multilocus genotype, respectively. Such a view supports the idea that genes are the relevant units of selection (Dawkins, 1976, 1982). Here we report how, by reconceptualizing the problem of defining average effects, we were able to address the question of the unit of inheritance operationally.

In our analysis we did not *a priori* assume that the average effects of the gametes can be approximated by the sum of the average effects of their alleles, i.e., we did not privilege the gene as a unit of inheritance. We were thus required to identify the average effects of objects other than individual alleles as well. Accordingly, we defined average effects for whole gametes, for functionally relevant associations of alleles on these gametes, and for all the individual alleles that make up these gametes (Laubichler, 1997a; Wagner *et al.*, 1998). Within our framework of analysis the identification of individual contributions to the total additive genetic variance thus becomes a decomposition problem. What fraction of the total additive genetic variance of the genotypes can be assigned to each of those individual entities that cross the generational boundary, such as, gametes, chromosomes, or single alleles?

To frame the question this way represents a conceptual shift in the way we analyze the unit or selection problem. Rather than defining a specific set of potential units of selection *a priori*, we now begin our analysis by identifying the most inclusive functionally relevant object. We then have to decompose the total additive genetic variance into the irreducible contributions of all objects in question.

However, neither the gametes nor the single alleles are the objects of ontological primacy in this case. Heritability is defined at the level of the phenotype, i.e., the organism. It is a phenotypic measurement, a regression of the midparent to the midoffspring phenotype (Falconer and Mackay, 1996). Organisms represented by their phenotypes are thus the ontologically prior objects in this case. However, phenotypes do not cross the generational

boundary and therefore we need to identify functionally defined objects, i.e., characters, that can account for the observed phenotypic correlation between parents and their offspring. In other words, we need to identify units of inheritance, which in turn explain the observed heritability.

We have a well-established theory of quantitative genetics that describes this process (Bulmer, 1980; Falconer and Mackay, 1996; Lynch and Walsh, 1998). The next crucial step was to develop an analytical method that allows us to identify the range of objects our theory applies to. We have developed such an analytical method for the case of a two locus two allele system with epistasis and linkage disequilibrium (Laubichler, 1997a; Wagner *et al.*, 1998; Laubichler and Wagner, 2000b). To determine the relative importance of all possible units of inheritance—in the case of a two locus two allele model gametes and alleles—we calculated the average effects of both and then defined (i) the additive gametic (Agam), (ii) the additive genic (Agen), and (iii) the irreducible additive gametic value (AEgam) of a genotype. The irreducible additive gametic value of a genotype represents the fraction of the additive gametic value that cannot be accounted for by an additive combination of the average effects of the alleles. The total additive genetic variance can then be decomposed into its components:

$$VA(total)=VAgen+VEAgam+Cov(Agen \times EAgam). \quad (1)$$

The question which biological object (gamete, allele, or both) is a functionally relevant unit of inheritance can now be addressed operationally. It is now no longer an either/or problem, but rather a question of determining the relative importance of all elements of analysis. We found that the relative importance of different units of inheritance is population specific. Depending on the amounts of linkage disequilibrium and the values for the epistasis coefficient (a measure for the strength of nonadditive genetic effects), as well as the allele frequencies in a population, the magnitude of the irreducible gametic component varies. In general, we can refute the assumption that the additive genic variance is always a good approximation of the total additive genetic variance. In Fig. 1 we give an example that illustrates the change in the relative importance of different units of inheritance in the course of evolution. The model in question is a so-called corridor model, used to study the role of pleiotropic effects on the fitness of complex adaptations (Wagner, 1988, 1989; Baatz and Wagner, 1997). Figure 1 shows a simulated evolutionary trajectory of a population from a point far away from its equilibrium until it reaches its fixed point. During the course of evolution the relative importance of different components of the additive genetic variance in fitness varies dramatically. While initially most of the additive

genetic variance can be attributed to the irreducible gametic component, by the time the population nears its equilibrium the genic component dominates.

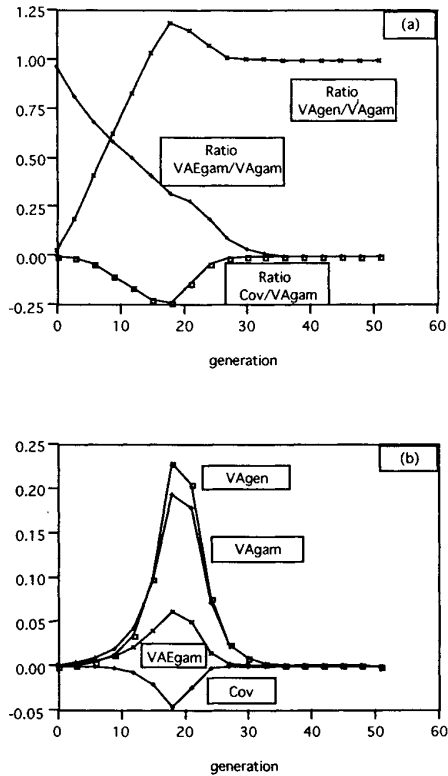


FIGURE 1 Additive genetic variance of the genotypic values for fitness in a corridor model. Shown are the fractions of the total additive genetic variance (VAgam) that can be attributed to the genic (VAgam), the irreducible gametic (VAEgam) and the covariance components (Cov) (a) and the absolute values for these variance components. The relative importance of different units of inheritance varies in the course of evolution. Initially, when the population is far away from its equilibrium point, the irreducible gametic component dominates, while closer to the equilibrium point the total additive genetic variance can be approximated by the additive contributions of individual alleles. Details can be found in Laubichler (1997a).

This type of analysis shows the kind of results we obtained when we turned the unit of selection question into a decomposition problem. We identified the functionally relevant characters (in our case units of inheritance) by their contributions to the total additive genetic variance. Units of inheritance are thus defined as abstractions of an ontologically primary object, the organism. In the next section we will generalize this theory of character decomposition. We will present the outline of an analytical theory that allows us to identify functionally relevant objects based on the notion of quasi-independence of characters.

DECOMPOSING ORGANISMS INTO A SET OF CHARACTERS

Adaptation is the result of spontaneous variation and natural selection. Since competitive success in natural selection depends on the overall lifetime performance of the organism, or even a group of organisms, adaptation is an organism level property. Nevertheless, it is often assumed that it is possible to study the adaptation of individual parts more or less independently of the rest of the organism. This assumption is at the heart of the character concept in evolutionary biology. The postulate of the independence of parts also underlies many research programs in biology (occasional lip service to the wholeness of the organism not withstanding). It raises the question under what conditions certain parts can be conceptually isolated from the rest of the organism and still fulfill the conditions of validity of the theory that we apply to them, for instance the theory of natural selection. Lewontin (1978) gave a heuristic answer to this question with the notion of *quasi-independence*. Quasi-independence refers to the assumption that it is possible to selectively adapt a specific character without simultaneously disturbing the remaining characteristics of the organism. Here we want to outline a mathematical approach that gives a precise meaning to this intuitive notion of quasi-independence. We also note that the problem of character identification in evolutionary biology is structurally the same as the so-called natural subsystem decomposition in general systems theory (Klir, 1985; Zwick, this volume).

In the context of the adaptationist research program in evolutionary biology the problem of character identification can be interpreted as a generalization of the unit of selection problem (earlier discussion). However, in contrast to much of the unit of selection debate, we do not assume *a priori* that certain objects are relevant characters, i.e., adaptations. Rather we start our investigations with the organism as the primary unit of analysis and construct characters as abstractions

from the ontologically prior object. We use the following criterion for constructing characters: *given the equations that describe the dynamics of natural selection among individuals (see later), how can we lump the organism "types" (be it genotypes or phenotypes) into equivalence classes (i.e., character states), such that the dynamics of these abstract types is still predicted by the same equations without any loss of generality.* The answer to this question reveals that the symmetries of the selection equations determine the conditions under which abstraction (i.e., identification of a character) is theoretically justified.

Let $T = \{t_1, \dots\}$ be a set of "types," i.e., the set of possible states of a complex adaptive system (CAS), such as genotypes or phenotypes. Associated with the set shall be a Malthusian fitness function $m: T \rightarrow R$ with $m_i = m(t_i)$, and a frequency distribution $x_i = x(t_i) \in [0, 1]$. Ignoring the transmission function (i.e., the influence of mutation and recombination) the frequency distribution shall obey the Crow-Kimura differential equation

$$\dot{x}_i = x_i(m_i - \bar{m}) \quad (2)$$

with $\bar{m} = \sum_{i \in T} x_i m_i$ (Crow and Kimura, 1970). For the theory developed here it is important to note the invariance properties of this equation. It is easy to show that the equation is invariant to a translation of the fitness values, i.e., the dynamics of the type frequencies does not change, if each fitness value is changed by the same additive constant.

A partitioning of T will be denoted as $\Pi = \{\pi_1, \dots, \pi_n\}$, where each π_j is an equivalence class of T . Each equivalence class π_j can be thought of as an allele in classical population genetics. The frequency of an equivalence class is $p_j = \sum_{t_i \in \pi_j} x_i$.

Following Rosen it is clear that any natural subsystem decomposition requires that the set T can be represented as a Cartesian product of at least two sets: $T = T_1 \times T_2$ (Rosen, 1978; see also Kim and Kim this volume). The same conclusion can also be reached from considerations about "independent variational characters" (Kim and Kim, this volume). In our context we want to represent T by the product of the focal partitioning and its "orthogonal complementary partitioning": $T = \Pi \times \bar{\Pi}$. Later we briefly discuss how one can obtain an *oc*-partitioning from a given one.

It is easy to see that an *oc*-partitioning can only exist if each equivalence class of the original partitioning has the same size M or if the number of elements of T is NM . An *oc*-partitioning can be constructed from a family of invertible 1-1 functions $F = \{f_{ij} | f_{ij}: \pi_j \rightarrow \pi_i\}$. Each function in this family maps each element of one equivalence class to its corresponding member in another equivalence class. A possible interpretation of such a function is that we seek two genotypes that are identical at a second locus. If F is also transitive, i.e., if $s = f_{ij}(t)$ and $t = f_{ji}(u)$ then $s = f_{ii}(u)$, then F defines a complementary partitioning $\bar{\Pi} = \{\bar{\pi}_1, \dots, \bar{\pi}_M\}$. Each class in this partitioning is $\bar{\pi}_k = \{s \equiv t \text{ if there is a } f_{ij} \in F | s = f_{ij}(t)\}$. This definition leads to an equivalence class since the functions in F are invertible and transitive. It is also clear that for each t in T there is an $i \leq N$ and $j \leq M$ such that $\{t\} = \pi_i \cap \bar{\pi}_j$ and the type set T can be represented as $\Pi \times \bar{\Pi}$.

Using the concept of a *oc*-partitioning, we can define two important terms: Pi-additivity and Pi-LE (linkage equilibrium), i.e., additivity of fitness and linkage equilibrium with respect to a given *oc*-partitioning.

Definition 1: a fitness function is additive with respect to a partitioning Π (Pi-additive) if there is an associated *oc*-partitioning and $m(f_{ij}(t)) = m(t) + c_{ij}$ for all $t \in \pi_j$.

Definition 2: a frequency distribution over T is in linkage equilibrium with respect to a *oc*-partitioning (Pi-LE) if for all $t \in \pi_j$ and all k , $x_k(f_{ij}(t)) = x_j(t)$ with $x_j(t) = \frac{x(t)}{p_j}$.

Note that Pi-LE implies that $x(t \in \pi_j \cap \bar{\pi}_k) = p_j \bar{p}_k$, the same criterion as in the classical definition of linkage equilibrium.

Let us now consider the conditions under which the dynamics of the frequency distribution over a partitioning is independent of the information lost because of the partitioning. In other words we want to understand under what conditions the equations governing the frequency distribution over a partitioning are dynamically sufficient. Each partitioning can be seen as an abstraction. We concentrate on the differences among the members of separate equivalence classes, but ignore the variation that exists among the members of the same equivalence class. It will be shown that the functional independence of a

subsystem is closely tied to the invariance properties of the dynamical equations that govern the system. In our case the relevant equations are the Crow-Kimura differential equations, introduced earlier. These equations are invariant under a translation of the scale of Malthusian fitness values. From this property we obtain the following Lemma:

Lemma: the dynamics of the frequency distribution over a partitioning of T is independent of the rest of the system if each equivalence class has the same variance in fitness.

Proof: it is easy to show that, ignoring transmission, the selection equation for the frequency distribution over that partitioning is

$$\dot{p}_j = p_j(\bar{m}_j - \bar{m}) \quad (3)$$

with $\bar{m}_j = \frac{1}{p_j} \sum_{t \in \pi_j} m(t)x(t)$. These equations, however, do not account for the

selection that is going on between members of the same equivalence class. Rather, competition between members of the same class leads to a dynamics of the mean fitness of the equivalence classes. This dynamics, ignored by the partitioning, induces a transformation of the parameters of the system of equations (3). It has already been mentioned earlier that these types of equations are invariant under any transformation of the form $\bar{m}_j \rightarrow \bar{m}'_j = \bar{m}_j + c$. Consequently the equations for the class frequencies remain unaffected by this dynamics, if the parameters all change with the same rate, i.e., the equations remain invariant if $\dot{\bar{m}}_j = \dot{\bar{m}}_k$ for all j and k .

It is well known that the rate of change in mean fitness is equal to the variance in fitness (one version of the fundamental theorem of natural selection)

$$\dot{\bar{m}}_j = \sum_{t \in \pi_j} m(t)^2 x_j(t) - \bar{m}_j^2 = V(m|\pi_j) = V_j. \quad (4)$$

Consequently, the dynamics of the frequencies over the partitioning is dynamically sufficient if $V_j = V_k$.

Next we will consider what conditions need to be met by the frequency distribution over T and the fitness function in order to establish the partitioning

as an independent subsystem. We will show that Pi-additivity and Pi-LE are necessary and sufficient in the sense defined next.

Proposition: Pi-additivity of the fitness function and Pi-LE are necessary and sufficient for the dynamical independence of the partitioning in the following sense:

1. Pi-LE is necessary and sufficient to guarantee $V_j = V_k$ for all Pi-additive fitness functions.
2. Pi-additivity is necessary and sufficient to make $V_j = V_k$ for all Pi-LE frequency distributions and to maintain Pi-LE under selection.

The proof of this proposition is somewhat technical and can be found in the Appendix. It is important to note that this result does not state that Pi-LE and Pi-additivity are the only conditions that fulfill the conditions of the lemma, since it is only stated that if one is given, then the other is necessary and sufficient. There are a number of special cases that formally fulfill the criterion of equal genetic variance, but are highly singular and dynamically unstable. A more general theorem has not yet been proven.

This result explains why additivity and linkage equilibrium play such a prominent role in theories of population and quantitative genetics. First, they are the conditions that assure that the dynamics of allele frequencies at a locus, or the frequency distribution of a quantitative character, are dynamically independent of the rest of the genotype (in the case of genes) or the phenotype (in the case of quantitative characters). In other words, additivity and linkage equilibrium are the conditions under which the theory of natural selection can be applied to sub-organismal characters. Second, these conditions are directly dictated by the invariance properties of the dynamical equations themselves, i.e., the Crow-Kimura equations for overlapping generations with weak selection (Crow and Kimura, 1970). Analogous results can be derived for Wright's selection equation for non-overlapping generations (Carter and Wagner, unpublished). These conditions also guarantee that those properties of the organism that have been ignored by the abstraction of characters have no impact on the dynamics of the system, since the equations that govern the dynamics of these characters are invariant to their influence.

It is interesting to note that the conditions for dynamical independence require a match between the frequency distribution of the types and the fitness function. The fitness function is determined by the physiology of the organism and its interactions with the environment, whereas the frequency distribution is determined largely by the transmission function, i.e., the probability

distribution with which certain types are generated by a certain parental type (Altenberg and Feldmann, 1987). This means that, only if a character defined by the dynamics of selection is simultaneously also a unit of transmission, such as a gene under Mendelian patterns of inheritance, can it be a dynamically independent character. Hence, there is a strong relationship between a unit of inheritance and a unit of selection, not in the material sense, but as defined by their dynamical properties that may, but do not need to, correspond to a physical unit.

Another interpretation of this result is that the conditions of Pi-additivity and Pi-LE guarantee the independence of one character from the other and that this enables these characters to act as independent modules of phenotypic adaptation (Wagner and Altenberg, 1996). Our results thus define the conditions under which complex organisms can maintain their evolvability, despite their physical complexity. Furthermore these conditions imply what kind of evolutionary changes can lead to the origin of a new independent character. The origin of a new character is based on the convergence of a functional property (Pi-additivity) and a property of the genetic/developmental system that maintains Pi-LE. Only if the genetic representation (Wagner and Altenberg, 1996) that leads to Pi-LE matches the pattern of functional constraints (Pi-additivity) can one speak of a new independent entity. This statement has obvious consequences for the definition of evolutionary novelties.

CONCLUSION

We began this paper with the claim that the structural weakness of many mathematical theories in biology is caused by their lack of an analytical theory that identifies the objects the theory applies to. We argued that in order to overcome this deficit, in many cases one needs an organismal perspective. One of the roles that we ascribed to the organism in theoretical biology is to provide the functional context from which biological objects (characters) can be abstracted. We also argued that in these cases the organism is ontological prior to its composite parts and that the individualization (abstraction) of biological characters is in essence a subsystem decomposition problem.

We gave two analytical examples for such a decomposition theory, the unit of inheritance and a generalized theory of biological characters. In both cases, we defined the relevant objects through their role within an organismal process represented by dynamical equations. Classes of objects are defined by their invariance with respect to specific transformations, i.e., by their stability within the process. We argued that this process specificity is characteristic for all suborganismic characters and even for biological objects in general.

Finally, we argued that the cell and the organism are privileged biological objects with respect to the processes of reproduction and self-maintenance. Self-maintenance implies the regulation of an internal milieu. We therefore propose that, in the case of animals, the organismal level of complexity arose with the origin of epithelial surfaces that separate the interior, or the body cavity, from the external environment.

ACKNOWLEDGMENTS

We thank the following individuals for helpful discussions on the subject: Lee Altenberg, Robert Brandon, James Griesemer, Paul Griffiths, and Bill Wimsatt. GPW would thank Junhyong Kim for his intellectual companionship in developing many of the ideas presented in this paper. JK presented his own interpretation of these issues in another paper (Kim and Kim, this volume). He also acknowledges his collaboration with Ashley Carter on extending the results on character decomposition (results not yet published). This is contribution # 62 of the Yale Center for Computational Ecology.

APPENDIX

Remark: this formulation of the proposition recognizes that there are other constellations of fitness functions and frequency distributions which make $V_j = V_k$, but they are all highly singular and do not guarantee this condition under a wide variety of conditions.

PROOF OF PROPOSITION:

Part 1: since the fitness function is assumed to be Pi-additive the variance of the k^{th} equivalence class is

$$\begin{aligned} V_k &= \text{Var}(m|\pi_k) = \text{Var}(m(f_{jk}(s)) + c_{jk} | s \in \pi_k) \\ \text{Var}(m(f_{jk}(s)) + c_{jk} | s \in \pi_k) &= \text{Var}(m(f_{jk}(s)) | s \in \pi_k) \end{aligned}$$

The difference $V_k - V_j$ can be written as

$$\sum_{t \in \pi_j} m(t)^2 [x_k(f_{kj}(t)) - x_j(t)] - 2\bar{m}_j \sum_{t \in \pi_j} m(t) [x_k(f_{kj}(t)) - x_j(t)]$$

The fitness function can be arbitrarily scaled such that $\bar{m}_j = 0$, leading to

$$V_k - V_j = \sum_{t \in \pi_j} m(t)^2 [x_k(f_{kj}(t)) - x_j(t)]$$

Since $m(t)^2 \geq 0$ the only condition which makes $V_k = V_j$ for all Pi-additive fitness functions is $[x_k(f_{kj}(t)) - x_j(t)] = 0$ for all $t \in \pi_j$, i.e., Pi-LE.

Part 2: assume the frequency distribution over T is Pi-LE, then the variances can be written as

$$V_k = \sum_{t \in \pi_j} (m(f_{kj}(t)) - \bar{m}_k)^2 x_k(f_{kj}(t))$$

$$V_j = \sum_{t \in \pi_j} (m(t) - \bar{m}_j)^2 x_j(t)$$

we now replace the fitness values of the types in the class π_k by

$$m(f_{kj}(t)) = m(t) + d_k(f_{kj}(t))$$

then we have

$$V_k = \sum_{t \in \pi_j} (m(t) + d_k(f_{kj}(t)))^2 x_j(t) - \bar{m}_k^2$$

$$\bar{m}_k^2 = \bar{m}_j^2 + 2\bar{m}_j \sum_{t \in \pi_j} d_k(f_{kj}(t))x_j(t) + \left(\sum_{t \in \pi_j} d_k(f_{kj}(t))x_j(t) \right)^2$$

which can be simplified to

$$\begin{aligned} V_k &= V_j + 2E_j(m(t)d_k(f_{kj}(t))) - 2\bar{m}_j E_j(d_k(f_{kj}(t))) + \\ &\quad + E_j(d_k(f_{kj}(t))^2) - E_j(d_k(f_{kj}(t)))^2 \\ V_k &= V_j + 2\text{Cov}_j(m(t), d_k(f_{kj}(t))) + \text{Var}_j(d_k(f_{kj}(t))) \end{aligned}$$

Hence, $V_k = V_j$ if and only if

$$2\text{Cov}(m(t)d_k(f_{kj}(t))) + \text{Var}_j(d_k(f_{kj}(t))) = 0.$$

This equation is trivially fulfilled if there is no variation in any of the equivalence classes because then the variances and covariances are also equal to zero. Another condition is that $d_k(f_{kj}(t)) = \text{const}$, which is the same as Pi-additivity. Note that zero variance also implies zero covariance. Finally, the condition is fulfilled if

$$\frac{\text{Cov}_j(m(t)d_k(f_{kj}(t)))}{\text{Var}_j(d_k(f_{kj}(t)))} = -\frac{1}{2}$$

which is the regression coefficient of d on m equal to $-1/2$, or if the linear regression equation reads

$$m(t) = -\frac{1}{2}d_k(f_{kj}(t)) + \text{const},$$

or if

$$m(f_{kj}(t)) = -m(t) + 2 \text{const}$$

This is an interesting condition that could be called reverse additivity. It is the same as additivity, but the fitness gradient is exactly reverse. Even if this condition fulfills the requirement $V_k = V_j$, it is incompatible with Pi-LE, since it leads to the selection of noncorresponding elements in the classes π_j and π_k . Thus only Pi-additivity is both sufficient for $V_k = V_j$, as well as necessary for arbitrary Pi-LE distributions and compatible with the maintenance of Pi-LE. This proves the proposition.

LITERATURE CITED

- Altenberg, L., and Feldmann, M. W. (1987). Selection, generalized transmission, and the evolution of modifier genes. I. The reduction principle. *Genetics* **117**:559-572.
- Baatz, M. and Wagner, G. P. (1997). Adaptive inertia caused by hidden pleiotropic effects. *Theor. Popul. Biol.* **51**:49-66.
- Bailey, V. A. et. al. (1962). Interaction between hosts and parasites when some host individuals are more difficult to find than others. *J. Theor. Biol.* **3**:1-18.
- Berryman, A. A. (1981). "Population Systems: A General Introduction." Academic Press, New York.

- Brandon, R. (1982). "The Levels of Selection." Biennial Meeting of the Philosophy of Science Association, Philosophy of Science Association.
- Brandon, R. N. (1990). "Adaptation and Environment." Princeton University Press, Princeton.
- Brandon, R. N. (1995). "Concepts and Methods in Evolutionary Biology." Cambridge University Press, Cambridge.
- Bulmer, M. G. (1980). "The Mathematical Theory of Quantitative Genetics." Calderon Press, Oxford.
- Buss, L. (1987). "The Evolution of Individuality." Princeton University Press, Princeton.
- Crow, J. F., and Kimura, M. (1970). "An Introduction to Population Genetics Theory." Harper and Row, New York.
- Darwin, C. (1859). "The Origin of Species." John Murray, London.
- Dawkins, R. (1976). "The Selfish Gene." Oxford University Press, Oxford.
- Dawkins, R. (1982). "The Extended Phenotype." Oxford University Press, Oxford.
- Endler, J. A. (1986). "Natural Selection in the Wild." Princeton, Princeton University Press.
- Falconer, D. S., and Mackay, T. F. C. (1996). "Introduction to Quantitative Genetics." Longman, Edinburgh.
- Fisher, R. A. (1930). "The Genetical Theory of Natural Selection." Calderon Press, Oxford.
- Gesteland, R. F., Cech, T. R. *et al.*, eds. (1999). "The RNA World: The Nature of Modern RNA Suggests a Prebiotic RNA." Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- Gilpin, M., and Hanski, I., eds. (1991). "Metapopulation Dynamics: Empirical and Theoretical Investigations." Academic Press, London.
- Grant, P. (1989). "Ecology and Evolution of Darwin's finches." Princeton University Press, Princeton.
- Griffiths, P. (1997). "What Emotions Really Are? The Problem of Psychological Categories." University of Chicago Press, Chicago.
- Griesemer, J. (2000). "Reproduction and the Reduction of Genetics in Development" (unpublished manuscript).
- Hanski, I., and Gilpin, M., eds. (1997). "Metapopulation Biology: Ecology, Genetics, and evolution." Academic Press, San Diego.
- Hartl, D. L. (1997). "Principles of Population Genetics." Sinauer, Sunderland, MA.
- Hofbauer, J., and Sigmund, K. (1988). "The Theory of Evolution and Dynamical Systems: Mathematical Aspects of Selection." Cambridge University Press, Cambridge.
- Hofbauer, J., and Sigmund, K. (1998). "Evolutionary Games and Population Dynamics." Cambridge University Press, Cambridge.
- Hull, D. (1980). Individuality and selection. *Annu. Rev. Ecol. Syst.* 1:311-332.
- Kim, J., and Kim, M. (2000). The mathematical structure of characters and modularity. In "The Character Concept in Evolutionary Biology" (G. P. Wagner, ed.). Academic Press.
- Klir, G. (1985). "The Architecture of Systems Problem Solving." Plenum Press, New York.
- Laubichler, M. D. (1997a). Identifying Units of selection: Conceptual and Methodological Issues. *Biology*. Yale University, New Haven:vii+197.
- Laubichler, M. D. (1997b). The nature of biological concepts." *Eur. J.Semiotic Stud.* 9(2):251-276.
- Laubichler, M. D., and Wagner, G. P. (2000a). Levels of selection in a two locus two allele system (in preparation).
- Laubichler, M. D., and Wagner, G. P. (2000b). Organisms and character decomposition: steps towards an integrative theory of biology. *PSA* (in press).
- Lewontin, R. (1978). Adaptation. *Sci. Am.* 239:156-169.

- Lloyd, E. (1988). "The Structure and Confirmation of Evolutionary Theory." Greenwood Press, New York.
- Lynch, M., and Walsh, B. (1998). "Genetics and Analysis of Quantitative Traits." Sinauer, Sunderland, MA.
- Margulis, L. (1970). "The Origin of Eucaryotic Cells." Yale University Press, New Haven.
- Margulis, L. (1981). "Symbiosis in Cell Evolution." W. H. Freeman, New York.
- Margulis, L., and Fester, R., eds. (1991). "Symbiosis as a Source of Evolutionary Innovation." MIT Press, Cambridge, MA.
- Maynard-Smith, J. (1998). "Evolutionary Genetics." Oxford University Press, Oxford.
- Maynard-Smith, J., and Szathmáry, E. (1995). "The Major Transitions in Evolution." Oxford University Press, Oxford.
- Miller, S. L., and Orgel, L. E. (1974). "The Origins of Life on Earth." Prentice-Hall, Englewood Cliffs.
- Mitton, J. (1997). "Selection in Natural Populations." Oxford University Press, Oxford.
- Murdoch, W. W. *et al.* (1996). Refuge dynamics and metapopulation dynamics: an experimental test. *Am. Nat.* **147**(3):424-444.
- Nagylaki, T. (1992). "Introduction to Theoretical Population Genetics." Springer, Berlin.
- Rosen, R. (1962). "Church's thesis and its relation to the concept of realizability in biology and physics." *Bull. Math. Biophys.* **24**:375-393.
- Rosen, R. (1978). "Fundamentals of Measurement and Representation of Natural Systems." North-Holland, New York.
- Sober, E. (1984). "The Nature of Selection." MIT Press, Cambridge, MA.
- Sokal, R. R., and Rohlf, J. F. (1981). "Biometry. The Principles and Practice of Statistics in Biological Research." W. H. Freeman, New York.
- Stearns, S. C. (1992). "The Evolution of Life Histories." Oxford University Press, Oxford.
- Wagner, G. P. (1988). The influence of variation and of developmental constraints on the rate of multivariate phenotypic evolution. *J. Evol. Biol.* **1**:45-66.
- Wagner, G. P. (1989). Multivariate mutation-selection balance with constrained pleiotropic effects. *Genetics* **122**:223-234.
- Wagner, G. P. (1997). The structure of biological concepts and its relation to the dynamics of biological organizations. *Eur. J. for Semiotic Stud.* **9**:299-320.
- Wagner, G. P., and Altenberg, L. (1996). Complex adaptations and the evolution of evolvability. *Evolution* **50**:967-976.
- Wagner, G. P., Laubichler, M. D. *et al.* (1998). Genetic measurement theory of epistatic effects. *Genetica* **102/103**:569-580.
- Weinberg, R. A. (1998). "One Renegade Cell: How Cancer Begins." Basic Books, New York.
- Weismann, A. (1892). "Das Keimplasma: Eine Theorie der Vererbung." Gustav Fischer, Jena.
- Williams, G. C. (1966). "Adaptation and Natural Selection." Princeton University Press, Princeton.
- Wimsatt, W. (1980). "The Unit of Selection and the Structure of the Multi-level Genome." Biennial Meeting of the Philosophy of Science Association, Philosophy of Science Association.
- Wimsatt, W. C. (2000). Emergence as non-aggregativity and the biases of reductionism. In "Natural Contradictions: Perspectives on Ecology and Change" (P. J. Taylor and J. Haila, eds.), (forthcoming).
- Zwicky, M. (2000). Wholes and parts in general systems methodology. In "The Character Concept in Evolutionary Biology" (G. P. Wagner, ed.), Academic Press.