This is to certify that the

dissertation entitled

Developmental and functional constraints on phenotypic covariation during growth and evolution

presented by

Miriam Leah Zelditch

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Zoology

Donald O. Straney
Major professor

Date May 22, 1987

DEVELOPMENTAL AND FUNCTIONAL CONSTRAINTS ON PHENOTYPIC COVARIATION DURING GROWTH AND EVOLUTION

By

Miriam Leah Zelditch

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Zoology

1987

ABSTRACT

DEVELOPMENTAL AND FUNCTIONAL CONSTRAINTS ON PHENOTYPIC COVARIATION DURING GROWTH AND EVOLUTION

By

Miriam Leah Zelditch

The hypothesis that ancestral developmental constraints guide the divergence of descendant taxa asserts that the constraints upon phenotype are developmental in origin and that these intrinsic constraints regulate morphological evolution. In this study, I examine the causal basis of phenotypic integration throughout post-natal ontogeny of the laboratory rat (Rattus norvegicus) and in five species of post-metamorphic Pentremites (Mississippian blastoids). I compare patterns of integration in five age-classes of Rattus and within and between two lineages of Pentremites. The purpose of this analysis is: 1) to test the hypothesis that developmental processes create observed patterns of covariation among characters; and 2) to test the hypothesis that these patterns of covariation are stable over ontogeny and phylogeny.

Hypothetical models derived from developmental theory were evaluated for their ability to reconstruct the observed variance-covariance matrix by confirmatory factor analysis. Confirmatory factor analysis provides a goodness-of-fit value for the fit of the model to the data. Comparisons between patterns of integration in successive age-classes

and related taxa were also made by confirmatory factor analysis, which treats comparative factor analysis as a problem in statistical inference.

Morphogenetic mechanisms, other than general body growth, do not adequately predict the observed variance-covariance among measures. Functional interactions among characters, however, do appear to constrain covariation. Interactions among characters engaged in a common function generate observed phenotypic integration in Pentremites. Furthermore, changes in function throughout growth may explain why patterns of integration vary throughout ontogeny in Rattus.

Patterns of integration change throughout growth in Rattus and during morphological and phylogenetic evolution of Pentremites. These results cast doubt upon the hypothesis that an invariant set of constraints, intrinsic to an ancestral population, guides morphological evolution. Changes in patterns of integration may involve not only changes in the intensity of constraints, or changes in the influence of constraints upon individual characters, but also changes in the identity and nature of the constraints. All of these changes occured during both post-natal growth and evolutionary divergence.

ACKNOWLEDGEMENTS

I thank the members of the Guidance Committee for their assistance. Donald Straney provided continuous support and critical advice. Perhaps most importantly, he asked me numerous questions which I could not answer. The attempts to answer these questions directed much of my research. Richard Houang gave me assistance with LISREL, and was always willing to discuss both methodological and conceptual issues. Robert Anstey gave me not only office space and help with paleontological questions, but also contributed greatly to the development of many of my ideas. Guy Bush kept me concerned with biological mechanisms. James Edwards maintained his interest in my work and my progress and provided criticism even after he took on new responsibilities.

I thank Frank Camillieri for help with factor analysis; particularly for informing me about confirmatory factor analysis. My father also made me aware of the criticisms of traditional factor analysis and maintained a useful skepticism about statistical procedures. I owe a great deal to his concern for methods of theory construction, and for his discussions about meaning and definition.

I also thank my family for the constant emotional,

logistical and intellectual support throughout my whole graduate career. Their commitment to my research enabled me to preserve mine.

This research was supported in part by National Science Foundation Dissertation Improvement Grant (BSR-8313110).

TABLE OF CONTENTS

LIST OF TABLES.....viii

LIST OF FIGURES.....xi

INTRODUCTION.....1

Complex adaptations.....7
Causal analysis of phenotypic constraints..... 10
Restricting the role of developmental constraints....17
Evolutionary dynamics and the adaptive landscape.....22
Morphogenesis and quantitative genetic parameters....34
Analysis of developmental integration.....41
Phenotypic evolution and the evolution of constraints.....44

EVALUATING GENERAL DEVELOPMENTAL MODELS.....49

Data.....50
Factor analysis.....50
Confirmatory factor analysis.....59
Models.....61
Criteria for selecting measures.....82
Discussion.....85

ONTOGENETIC VARIATION IN PATTERNS OF INTEGRATION.....90

Data.....92
Intensity of integration.....94
Causal analysis of patterns of integration.....98
Comparative factor analysis.....110
Discussion.....118

THE EVOLUTION OF PATTERNS OF INTEGRATION.....126

Pentremites.....128
Methods.....143
Analysis of developmental and functional integration.....144
The evolution of functional and developmental constraints.....159
Evolution of constraints and morphological divergence.....169
Discussion.....172

CONCLUSION: CONSTRAINTS AND MORPHOLOGICAL EVOLUTION.....177

BIBLIOGRAPHY.....194

LIST OF TABLES

1. Abbreviations and description of characters represented in Figure 2 (Olson and Miller, 1958).....52

2. Factor-pattern matrices for the exploratory factor analysis of measurements on one-day-old rats.....57

3. Evaluation of models fit to the correlation matrix of measurements on one-day-old rats.....62

4. Evaluation of the static allometry and timing models.....68

5. Hypothetical factor-patterns derived from the path diagrams (Figures 4A, B).....72

6. Evaluation of the tissue-origin and geometric models.....73

7. Sequential evaluation of the components of the composite tissue-origin model, relative to the model of no integration.....77

8. Sequential evaluation of the components of the composite geometric model, relative to the model of no integration.....77

9. Factor-pattern matrices for the two different models common to both geometric and tissue-origin hypotheses.....79

10. Goodness-of-fit values for the size model fit to the limb measures throughout ontogeny.....100

11. Goodness-of-fit values for the size model fit to skull measures throughout ontogeny.....100

12. Goodness-of-fit values for the developmental models fit to the limb measures in five age-classes.....104

13. Goodness-of-fit of the hypothesis that origin in a common embryonic tissue determines observed integration in these five age-classes.....104

14. Goodness-of-fit of the hypothesis that embryonic induction is a source of integration throughout post-natal growth.....105

15. X^2 difference tests for the comparison between the induction model and the conflicting hypothesis that the jaw comprises a mandibular and maxillary unit.....105

16. X^2 difference tests for the evaluation of relative fit of the size and unit models.....107

17. Goodness-of-fit values for the unit model fit to the inclusive skull data.....109

18. X^2 difference tests for the size versus unit models fit to the cranium and jaw data.....109

19. Goodness-of-fit values for the unit models fit to jaw measures in the five age classes.....111

20. Goodness-of-fit values for the hypotheses that factor variance-covariance (Fcova) and the uniqueness of individual characters (U) are invariant in limb measures of sequential age-classes.....115

21.Comparisons among the parameters of integration observed in skull measures of sequential age-classes.....117

22. Description of characters represented in Figure 7....131

23. Formation, species and sample size of the godoniform and pyriform samples analyzed by Olson and Miller.....135

24. Classification of Pentremites species by discriminant function analysis.....137

25. Coefficients for the first canonical variable.....139

26. Goodness-of-fit values for the hypothesis that size alone accounts for observed integration in each sample of Pentremites.....147

27. Goodness-of-fit for the tissue-origin model fit to all five species of Pentremites.....149

28. Probability levels for the simple functional hypotheses fit to each sample of Pentremites.....151

29. X^2 difference tests for the significance of differences between the growth model and the simple functional models 151

30. Probability levels for the models which specify

that growth and interactions between pairs of functional units determine integration.....153

31. Comparison of parameters of integration between pyriform Pentremites.....164

32. Comparison of parameters of integration between P. godoni and P. pyriformis.....164

33. Comparisons in parameters of integration between godoniform and pyriform taxa.....167

34. Relationship between moprhological divergence and stability of intrinsic constraints.....170

LIST OF FIGURES

1. The epigenetic landscape (after Waddington, 1974).....29

2. Schematic representation of skeletal characters analyzed.....51

3. Path diagram for the hypothesis that general body size determines covariation among all characters.....65

4. Path diagrams for the hypotheses that origin in a common embryonic tissue (A) or common geometric orientation (B) determine covariance among characters.....71

5. Schematic representation of the cranium and jaw measures comprising the second set of skull measures.....95

6. Temporal and regional variation in standard scores of overall intensity of integration over postnatal growth in the limb, inclusive skull, cranium and jaw, and within jaw measures.....97

7. Schematic representation of the external morphology of the blastoid and representation of the measures analyzed.....130

8. Hypothesized relationships among the five species of Pentremites (after Waters et al., 1986).....133

9. Plot of the five Pentremites species on the first two canonical variables estimated by discriminant functions.....141

10. Path diagrams for the simplest adequate models fit to the pyriform Pentremites.....157

11. Path diagrams for the simplest adequate models fit to the godoniform Pentremites.....158

INTRODUCTION

A number of morphologists and systematists have recently proposed the hypothesis that developmental mechanisms influence the rate of evolution and determine the likelihood of particular phenotypic changes within a lineage (Alberch et al., 1979; Alberch, 1980, 1982, 1983; Alberch and Alberch, 1981; Alberch and Gale, 1985; Eldredge and Gould, 1972; Frazetta, 1975; Gould, 1977, 1980, 1982, 1984; Hoffmann, 1981; Kurten, 1953; Maderson, 1975; Maderson et al., 1982; Maynard Smith et al., 1985; McNamara, 1986; Shubin and Alberch, 1986; Rachootin and Thomson, 1981; Vavilov, 1922; Waddington, 1976; Wake et al., 1983; Williamson, 1981). Developmental mechanisms are purported to have this power over morphological evolution because they impose constraints upon the distribution of phenotypes within populations and, thereby, constrain the range of potential morphologies available to a lineage. Developmental processes, according to these recent hypotheses, regulate phenotypic evolution because they control the distribution of phenotypic variation.

Apparently, at least as judged by these authors, developmental constraints cause an impressive variety of patterns in morphological evolution. Furthermore, they account for patterns perceived as extremely difficult to

explain by traditional views of morphological evolution. Traditional selectionist theory, whatever its intentions, implies that morphology is almost infinitely labile, capable of the unlimited adaptive change demanded by changing local ecological conditions. The persistence of a phenotype through significant durations of geological time, in a stratigraphic sequence presumably spanning multiple environments, challenges the notion that the phenotype constantly responds to environmental change. Rapid change, concentrated within short intervals of the history of a species, further challenges the conception of phenotypic change as a continuous response to the external environment. A third challenges lies in the absence of transitional phenotypes, intermediate between ancestral and descendant morphologies, either in the fossil record or exhibited in related extant taxa. Evidence from paleontological studies initially suggested that these intermediates might be lacking not because the fossil record is so sparse that intermediates are simply unpreserved, but because morphological change occurs rapidly and in small isolated populations. Apparent stasis, coupled with rapid change, results not from the poverty of the fossil record but from the actual mechanisms of speciation and morphological evolution. As Gould and Eldredge assert (1977), "stasis is data". This evidence of morphological stasis and rapid phenotypic change motivated the search for intrinsic constraints capable of resisting the tendency to adapt to

spatial or temporal variation in ecological conditions.

Developmental constraints provide an explanation for these observations in terms of intrinsic features of the morphogenetic process. Developmental constraints are hypothesized to be responsible for long periods of morphological stability and, despite the apparent contradiction, simultaneously cause rapid phenotypic change (Gould, 1980; Eldredge and Gould, 1972; Hoffmann, 1981; Maderson et al., 1982; Rachootin and Thomson, 1981; Maynard Smith et al., 1985; Waddington, 1976; Wake et al., 1983; Williamson, 1981). Developmental constraints may also induce directional trends in morphological evolution, and explain why particular morphologies recur frequently while others never appear (Alberch, 1980, 1982, 1983; Alberch and Alberch, 1981; Alberch and Gale, 1985; Gould, 1977, 1980, 1982; Maderson et al., 1982, McNamara, 1986; Shubin and Alberch, 1986; Vavilov, 1922; Waddington, 1976). More fundamentally, developmental constraints may be the mechanisms responsible for the coordinated changes in numerous characters, the integrated modifications critical to the origin of complex adaptations (Frazetta, 1975; Gould, 1977, 1980, 1982; Kurten, 1953; Maderson, 1975; Maderson et al. 1982; Olson and Miller, 1958, Maynard Smith et al., 1985).

Such a diverse set of consequences depends upon two features of developmental processes. First, developmental systems may often be inherently constrained, nonlinear,

complex dynamical systems, only rarely capable of being pushed off-course. When modified, they would regularly yield particular results. This kind of system characteristically possesses intrinsic stable states (Alberch, 1980, 1982; Oster and Alberch, 1982; Maynard Smith et al., 1985; Thom, 1975; Waddington, 1976). Barring perturbations that exceed the absorbtive capacity of the system, the stable state will be persistently maintained. Severe perturbations cause switching to an alternative stable state. The transformation between stable states is not a simple linear response to variation, but rather it is discontinuous. Such transitions between stable states, technically referred to as "catastrophes", can be classified, predicted and analyzed and the rules governing the transitions between stable states can be mathematically deduced (Thom, 1975). These rules are supposed to predict the evolution of morphology (Alberch et al., 1979; Maderson et al., 1982; Maynard Smith et al., 1985; Waddington, 1976).

The second property of the developmental system responsible for constraints on morphological evolution derives from the inherently historical character of developmental processes (Gerhart et al., 1982). Not only is development necessarily historical, but also, at the cellular level, it is partially a stochastic process. Each event throughout any developmental process depends not upon some encoded developmental program but upon the specific chain of causal antecedent events. Thus the potential

future states of any cell depend upon the prior states of the system, and perhaps even upon the prior path followed by the system, rather than upon some set of universal instructions. Because development involves progressive restrictions in the fate of cells at each stage of differentiation, it has the characteristic of a branching sequence. The historical branching sequence itself restricts the potential future states of the system (Kauffman, 1983). As a result of the causal linkage between successive states of the system, the branching pattern of developmental history creates a network of interactions between developing morphological traits. Any irregular occurence early in the history of the system can alter all succeeding events. Morphological characters, the outcome of the ontogenetic process, are integrated by the network of interactions so that the cascading effect of changes in early events creates sets of correlated characters jointly dragged along a new developmental pathway.

All of the features of developmentally constrained morphological evolution can be predicted from the historical and dynamic properties of developmental systems. These properties determine the possible distribution of phenotypes. And, according to theory, the distribution of phenotypes within a population necessarily limits the potential phenotypic distribution within a clade. This idea, that developmental constraints underlie patterns of morphological evolution because they circumscribe the range

of potential phenotypic variation, is fundamental to the theory of developmental constraints (Alberch, 1980; Gould, 1984; Maderson et al., 1982; Maynard Smith et al., 1985; Shubin and Alberch, 1986). According to this argument, the pattern and amount of variation at the species level limits the available range of phenotypes in descendant species; the same set of constraints limits the the range of variation throughout the taxonomic hierarchy (Alberch and Alberch, 1981).

The evolutionary patterns apparently so difficult to explain by traditional selectionist theory follow quite easily from the inherent constraints exerted by developmental systems. Morphological stasis, the persistence of a given phenotype despite changing environments and speciation is a consequence of the inherent stability of the developmental system. Even when the environment changes regularly throughout geological time, morphology need not respond because morphology is intrinsically constrained. Thus the average phenotype is unlikely to respond to environmental change.

Apparent punctuations, the rapid changes in average phenotype that may be concentrated in speciation events (Gould and Eldredge, 1977), occur because of transitions between stable states (Maderson et al., 1982). The non-existent phenotypes, the so-called "gaps in morphospace" (Raup, 1966, 1967), are located in the empty region between stable states. The transformation rules that regulate this

switching between stable states impart an inherent tendency to replicate particular states, and so determine directional trends, the continuous modification of some characters within a lineage (Maderson et al., 1982).

Convergence, the recurrence of particular phenotypes in distantly related species, is thus also hypothesized to result from the tendency of an epigenetic system to stabilize at particular points. Related taxa, sharing a common epigenetic system, should stabilize at the same stable states. Furthermore, the transformation rules are also conserved. Thus convergence is a result of inherent stable states and the biased probability of particular transformations, rather than a consequence of adaptation to similar ecological circumstances (Alberch and Alberch, 1981; Gould, 1977; Vavilov, 1922; Wake, 1981; Wake et al., 1983). Intrinsic constraints, imposed by the epigenetic system, produce stasis and rapid change, as well as trends and convergence.

Complex adaptations

Complex adaptations may also be a consequence of the behavior of dynamic systems and the historical nature of developmental processes. Even a less formal analysis of developmental systems, lacking catastrophes and stable states, suggests that processes of development, and their phenotypic products are integrated (Atchley et al., 1981;

Cheverud, 1982; Frazetta, 1975; Olson and Miller, 1958; Van Valen, 1960; Vavilov, 1922). However, a morphogenetic approach to macroevolution explains the two most difficult aspects of complex adaptations: their sudden origin and coordinated modification (Alberch, 1982; Frazetta, 1975; Gould, 1982). Complex adaptations, at their origin, may not be fully realized adaptations because they might still require considerable fine-tuning and adjustments to perform their functions well. However, they could originate with all their necessary features of organization and coordination, because the epigenetic system is integrated (Gould, 1982). Changes in any aspect of the developmental processes would simultaneously affect all developmentally correlated pieces so that a complex as a whole might be fashioned at once, without awaiting modification and assembly of several independent pieces of a mosaic. Natural selection need not progressively modify each piece, and then add it to the complex when it becomes a functional structure; instead, developmental constraints by themselves could create the basic structural framework of complex adaptations and guide their modification.

Developmental integration, resulting from a common response of several features of the phenotype to developmental processes, would cause the coordinated response of the phenotype to natural selection. Selection would have to act upon the whole integrated developmental system. As a result, the evolution of complex adaptations

need not progress at a slow rate, nor by the incremental acquisition of new components. Complex adaptations could originate without macromutations, and the initial stages of a novelty not be supremely well-adapted. But, at their origin, complex adaptations would already possess sufficient organization to be, if not adapted, at least not maladapted (Alberch, 1982; Frazetta, 1975; Gould, 1977, 1982; Maderson et al., 1982).

Accounting for the origin and evolution of complex adaptations by natural selection created serious difficulties for early Darwinian theory. Historically, the difficulty of explaining the origin of complex adaptations by natural selection undermined acceptance of Darwinian theory by morphologists. E. S. Russell (1916) questioned whether Darwin was sufficiently aware of the problem of complex adaptations because natural selection seemed so incapable of creating the necessary complex organic organization. The subject of morphology, throughout the late eighteenth and early nineteenth century, was dominated by The Law of Correlation of Parts (Cuvier, 1769). The Law of Correlation of Parts dictated, on first principles, that all structures within an organism are influenced by the need for harmonious function. This need for harmonious organization forces all characters within the individual to be integrated. Darwin proposed an alternative explanation of biological organization-- organization is a consequence of natural selection. For this organization to be achieved by

natural selection, it must first appear spontaneously, generated by random variation, before it can be favored in competition with more poorly integrated individuals. Over time, parts of the organism would acquire their correlations as new components of organization appeared and were favored by selection. This Darwinian explication of natural selection, however, appeared incapable of building organization and complexity with its piecemeal, gradual tinkering. The origin of complex structures, whose component parts form an integrated functional unit, constituted evidence to many traditional eighteenth and nineteenth century morphologists of a vitalistic evolutionary principle (Owen, 1868; von Baer, 1876a, b; Kolliker, 1864; von Hartmann, 1906; Milne-Edwards, 1867; all cited and discussed in Russell, 1916). Developmental constraints, in contrast, can explain the origin of novelties without invoking vitalistic principles or presuming piecemeal tinkering.

Causal analysis of phenotypic evolution

The ability of developmental constraints to explicate such a wide variety of problematic phenomena may account for their current great appeal. But some of the power attributed to developmental constraints exaggerates their role in morphological evolution. No evolutionary processes require developmental constraints. Nor do current

investigations of particular cases of evolving morphologies support the confident assertions of many authors that developmental constraints predict the distribution and transformations of phenotype within a lineage. The role of developmental constraints is merely inferred from comparisons of repeated occurrence of particular variants in a given taxon (Alberch, 1983; Garcia-Bellido, 1983; Maynard Smith et al., 1985), comparisons among developmental processes among related taxa (Alberch and Alberch, 1981; Alberch and Gale, 1985; Maderson et al., 1982), from examination of heterochronic changes (Alberch et al., 1979; Gould, 1977, 1982; Maderson et al., 1982; Wake, 1980) or by analysis of rates of phenotypic change within a lineage (Eldredge and Gould, 1972; Hoffman, 1981; Maderson et al., 1982; Wake et al., 1983; Williamson, 1981). The constraints are not themselves identified and assessed for their power to control phenotypic distributions. And only rarely (e.g. Maynard Smith et al., 1985) are alternative explanations for phenotypic distributions evaluated as competing or supplementary hypotheses. Yet natural selection and genetic drift can explain the same patterns, and adopting natural selection as the explanation of particular morphological changes may be justified by analysis, not merely by tradition. Natural selection is not a principle to invoke simply because a developmental explanation fails, such as when integrated characters undergo apparent disassociations (Alberch and Alberch, 1981). Developmental constraints are

not the universal null hypothesis. Treating developmental constraints as a null hypothesis errs in the same way, although in the opposite direction, as routinely invoking selection until confronted with self-evidently maladapted traits.

Natural selection, random genetic drift and developmental constraints no doubt cooperate to guide morphological change. They may all be causes of morphological evolution, but different kinds of causes. The idea that natural selection and developmental constraints represent competing causal hypotheses presupposes that they are logically equivalent and mutually exclusive. But, if developmental constraints create the biased distribution of phenotypes, subsequently filtered by natural selection, then developmental constraints supply, at some level, the raw material for evolution. Regarded in this way, the outcome of the developmental constraints (the biased distribution of phenotypes) is one material cause of evolution. Natural selection, on the other hand, is the efficient cause exploiting the materials supplied by development. In terms of Aristotles' classic example of the relationship among causes, the phenotypes generated by developmental processes and the genotypes responsible for the developmental processes, are the bricks and mortar of evolution, the matter employed in the construction of adaptation; natural selection is an efficient cause, the mason responsible for transforming the materials into adaptations. Just as the

particular materials available to the architect suggest, or conversely, restrict his techniques and procedures, development may constrain the potential range of phenotypic variation allotted to natural selection.

However, the material and efficient causes are fused implicitly, and sometimes even explicitly in several arguments elaborated by macroevolutionists: developmental constraints create patterns of morphology, and these patterns of morphology do not merely provide the framework for evolutionary patterns, they are themselves the evolutionary patterns that must be explained (e.g. Gould, 1982). Natural selection is interpreted as responsible specifically (and solely) for adaptation, not for morphology nor for its changes, while the potential effects of random genetic drift tend to be ignored. The causes of morphological change are perceived to lie in the processes that occur within an individual: differential gene activity, interactions between genes and their products, interactions between gene products of different genes, the spontaneous self-assembly of complex macromolecules and tissues, and the interactions between cells and tissues (e.g. Alberch, 1982). According to this view, developmental constraints are the agents of the phenotype and its change, and phenotypic change is evolution.

Developmental constraints, according to some proponents of this morphogenetic approach to macroevolutionary theory, particularly Alberch (1980, 1982), thus supply both material

and efficient causes of phenotypic evolution. Developmental constraints bias the distribution of phenotypes available to natural selection, and, in the absence of selection, create the observed range of phenotypes. Material causes, in effect, are treated as equivalent to efficient causes. Yet, this confounding of different categories of proximate causes is not a simple logical error. Rather, it is an intentional assertion about the causal relationship between the distribution of phenotypes within a population (determined by developmental constraints) and the observed trends in phenotypic evolution (biased by these same constraints). Developmental mechanisms are regarded as both material and efficient causes of phenotypic evolution because, according to this view, the cause of phenotypic change is the cause of evolution.

Such patterns in morphological evolution as stasis, discontinuous change and morphological trends follow naturally from the analysis of the assembly rules and local conditions that govern the development of phenotype. Assembly rules spontaneously generate the range of phenotypes and are intrinsic to the materials of phenotype (the cytoskeleton, cytoplasmic determinants, junctions between cells, protein structure of morphogenetically active macromolecules, quantities and distribution of cell-specific mRNA, etc.). While selection ultimately filters these changes, developmental mechanisms initially determine the potential range of variation. Changes in the materials may

necessarily change the phenotype when assembly rules respond to the changes in local chemical, cytoskeletal or tissue conditions. According to this view, the potential range of phenotypes allowed by development defines the universe of potential morphological change. Only when these variants differ in fitness will selection be responsible for the evolutionary dynamics of morphology.

Morphological evolution is thus equated to phenotypic change, and so the cause of a new morphology is the cause of its evolution. Natural selection is left out of this causal explanation because selection is the cause of something else. It is not the cause of phenotypic evolution per se, but of adaptation. "Adaptation" refers specifically to the outcome of natural selection, which molds a structure to perform a particular function. When a structure currently performs that function, but lacks a history of responding to selection, then it may be mistaken to consider that structure as an adaptation (Gould and Lewontin, 1979). Rather, the current usage of a structure may reflect less the history of selection than the consequences of mutation or the effects of selection acting upon other, developmentally correlated characters. Current usage may be merely accidental. And at least some phenotypic change may be unrelated to adaptation. Phenotypic evolution, in the absence of selection, could occur when intrinsic mechanisms cause correlated changes among integrated characters. This argument, developed largely by Gould (1977, 1980, 1982,

1984; Gould and Lewontin, 1979), diminishes the role of selection in the history of morphology and emphasizes that changes in developmental mechanisms modify the phenotype.

According to traditional microevolutionary theory, mutation is ultimately the source of variation. Mutation, however, is never, by itself, the cause of evolution. Change in the average phenotype is accomplished by natural selection and random genetic drift. Selection, acting upon variation in fitness, results in change in the average phenotype within a population. The phenotypic variation present in a population does not initiate phenotypic change, rather it is passive; it is presented to selection, the active agent of phenotypic change. Thus the traditional microevolutionary theory of morphological evolution differs in several important ways from that of Alberch and Gould, and from the others who emphasize the role of developmental constraints in regulating phenotypic change. The origin of phenotypic variation is not the ultimate cause of phenotypic evolution.

When developmental constraints are regarded as both material and efficient causes of phenotype, and (by extension) of morphological evolution, they are perceived as capable of exerting a considerably greater influence over evolution than natural selection and random drift. According to the hypothesis that the proximate causes of morphological variation are also proximate causes of morphological evolution, developmental constraints do not simply account

for the origin of a structure. They do not merely respond, providing a source of variation. The changes in phenotype that constitute macroevolution do not depend upon the filter manipulated by natural selection; rather, they depend upon the changes in the materials and assembly rules of development.

Restricting the role of developmental constraints

Stasis is one of the evolutionary patterns that initially challenged selectionist theory. However, Lande (1985, 1986) has shown analytically that stasis is a predictable consequence of traditional microevolutionary theory. Lande invokes no intrinsic features of the individual to explain the long duration of any average phenotype. To calculate the probability that a population will depart an adaptive peak, he applied a stationary diffusion model and calculated the expected duration (T) of the interval of stasis for a population at an adaptive peak ($W_{(a)}$), subject to the genetic variance-covariance structure of the characters (G) and extrinsic constraints ($c_{(a)}$, $c_{(v)}$):

$$T = G\,(c_a c_v)^{-1/2}\,\left| W_{(a)}/W_{(v)} \right|\,2Ne \qquad (1)$$

where N_e is the effective population size and $W_{(v)}$ is the depth of the adaptive valley surrounding the original peak. T is therefore largely determined by two extrinsic factors:

by population size and by the relative loss of fitness experienced when passing from an adaptive peak through an adaptive valley. Since the probability of change depends upon the ratio of fitness of the population occupying the adaptive peak to the fitness of the population when in the valley (raised to the power of $2N_e$) and only linearly upon the genetic variance-covariance, the amount of genetic variance-covariance will only exert an appreciable influence upon the duration of a given morphology when the ratio of the two fitnesses is close to 1 (the adaptive landscape is relatively flat) or when the population size is very small. Even when the adaptive peak is only 1.5 times as high as the valley, and the effective population size is on the order of 100, the expected time until a shift between adaptive peaks occurs will be on the order of 10^6 generations (Lande, 1985). Although the adaptive peak may be relatively shallow, large populations will still tend to exhibit stasis. The intrinsic constraints (estimated by genetic variance-covariance in Lande, 1986), on the other hand, have little influence upon the probability of morphological change.

It is during the transition between morphologies that the amount of heritable variation and covariation can exert an influence. Once the population has left an adaptive peak, the next peak colonized is largely a function of both genetic variance-covariance and the local geometry of the adaptive landscape. During the transition between adaptive peaks, a population does not move randomly around the

adjacent valley. Instead, it will progress to the next most accessible adaptive peak. Lande (1986) applied a conditional diffusion model to the transition between adaptive peaks and determined the duration of the transition (T*):

$$T^* = G^{-1} \{c_v^{-1} \ln|N_e c_v (b-v) (v-a)/2| - 1/2\, c_a^{-1} \ln|-N_e c_a (v-a)^2/2| - 1/2\, c_b^{-1} \ln|-N_e c_b (b-v)^2/2|\} \quad (2)$$

where a refers to the original peak and b refers to the peak to which the population evolves. This transition may be rapid. Certainly, when compared to the duration of the ancestral phenotype, the transition between phenotypes can occupy a short time in the history of the lineage.

The local geometry of the adaptive landscape influences the path between peaks, but the probability of following a particular path depends directly upon the heritable variation in a set of characters. In contrast, it depends only logarithmically upon a function of the height of the adaptive peaks, local curvature and population size. The influence of heritable variance-covariance upon following one or another available path between peaks can be represented using Lande's conditional diffusion model for the duration of a transition. The duration of a transition can be interpreted as the inverse of the probability of the transition. Improbable transitions take a relatively long time to occur. Thus high values of T* reflect a low

probability of occurrence. Given two different populations of the same size, each starting at a different peak,(W_{a1}, W_{a2}), the relative probability of one or another of two populations arriving at peak W_b (T^*_1/T^*_2) can be computed as the ratio of the durations of the two transitions:

$$T^*_1/T^*_2 = G_2/G_1 \ln (N_e)\left\{ cv_2/cv_1 \quad \frac{\ln\left|\frac{cv_1}{cv_2}\frac{(b-v_1)}{(b-v_2)}\frac{(v_1-a_1)}{(v_2-a_2)}\right|}{} - \right.$$

$$ca_2/ca_1 \quad \frac{\ln\left|\frac{(-ca_1}{(-ca_2}\frac{(v_1-a_1))}{(v_2-a_2))}\right|}{} -$$

$$\left. \frac{\ln\left|\frac{(b-v_1)^2}{(b-v_2)^2}\right|}{} \right\} \qquad (3)$$

where c_{a1} and c_{a2} refer to the local geometry of the landscape at the initial peaks, c_{v1} and c_{v2} refer to the geometry of the landscape at the two valleys and G_1 and G_2 are the genetic variance-covariance matrices for the two populations. The relative probability that population 1 or 2 will transit to peak W_b is thus proportional to G_1/G_2. Intrinsic constraints (interpreted as the genetic variance-covariance matrix), when unbreakable and resistant to particular changes in phenotype, can influence the direction of the transition. In particular, intrinsic constraints might guide the population towards the foothills of one peak and thereby determine which available adaptive peaks is most accessible.

Lande formalizes the analysis of evolutionary rates and recasts the debate over rates of phenotypic change. He rejects a unified explanation simultaneously accounting for

both stasis and punctuational events. He thus undermines one of the fundamental arguments pressed by the macroevolutionists: the common basis of stasis and punctuational change. However, he does not entirely eliminate the role developmental constraints play as agents of morphological change. Certainly developmental constraints no longer serve to explain the absence of change. Nor do they explain the limits upon the distribution of a given phenotype within a population. But if these constraints are rigid and unbreakable they may retain their significance as the mechanism guiding particular morphological changes (Lande, 1986) and thus may act as potential efficient causes of directional change in evolution.

According to Lande, developmental constraints can only influence the direction of morphological evolution when they are unchanged (1986). The stable developmental constraints determine the patterns of variation and covariation, which, in turn, restrict the possible directions in which selection and drift can move the population. Modifications in morphogenesis diminish the ability of developmental constraints to regulate morphological evolution (Lande, 1986) because these modifications would alter the patterns of variation and covariation. Thus selection would not be constrained by the ancestral patterns of covariation among characters. Developmental processes exert their constraints because they cause selection to act simultaneously upon

those suites of characters comprising heritable units of variation.

Lande differs from those macroevolutionists (e.g. Alberch, 1982; Gould, 1979, 1982; Maderson et al., 1982) who believe that changes in these processes ultimately cause rapid phenotypic change. For Lande, the mechanisms of morphological evolution lie in those processes which act upon the distribution of phenotypes within a population, not in those processes which precipitate changes in the phenotype. While the direction of evolution might define the potential directions of change, the source of variation is ultimately not the principal mechanism of evolution. Lande details no hypotheses at all to account for novelties. Phenotypic change is not ultimately referred to those mechanisms which create the observed morphology but to the mechanisms which move the population across the adaptive landscape.

Evolutionary dynamics and the adaptive landscape

The metaphor of the adaptive landscape (Wright, 1932a) has organized the traditional rhetoric of phenotypic evolution. Provine (1986) goes so far as to claim that Wright's concept of the adaptive landscape was one of his single most influential, if also most confusing, contributions to evolutionary theory. There has been great heuristic value in the graphical representation of the

relationship between organisms, mechanisms of adaptation, and adapted phenotypes. However, Wright developed at least two versions of the adaptive landscape, later supplemented by a third (Simpson, 1953), which are not easily rendered consistent. In his original version of the landscape, Wright (1932a) envisioned it in terms of the fitness values of gene combinations. However, genotypes do not vary continuously, each combination is discrete. Therefore there is no continuous surface because the axes lack gradations.

The second version of the adaptive landscape developed by Wright (1935) differs from the first in that the axes are defined by gene frequencies rather than by gene combinations. Now each point on the surface represents a whole population (determined by the gene frequencies for each allele in a population) rather than a single individual (determined by its own unique genotype). However, sets of gene frequencies, unlike gene combinations, have no particular adaptive values. Particular gene frequences might have a high adaptive value within one array of gene combinations and a low one in another array. The advantage of this version lay in the possibility of representing the surface of the average population fitness. Thus it contributed, in a major way, to the explication of the quantitative theory of population genetics.

Simpson characterized the landscape in terms of phenotypic, rather than genetic, axes. This is the view adopted later by Wright and developed further by Lande

(1976, 1979, 1985, 1986; Lande and Arnold, 1983). The phenotypic interpretation of the adaptive landscape has the obvious advantage of having being defined by continuous axes of easily estimated variables. Unfortunately, the phenotypic measures are not related, in any obvious way, to the gene frequencies or genetic variance-covariances that are fundamental to the evolutionary theory. If it were only the heuristic value of the graphs that were in doubt, this ambiguity would lead to no serious conceptual confusion. But the confusion is not trivial.

Evolutionary mechanisms entail changes in genetic parameters. This version of the landscape is defined in purely phenotypic terms. Selection acts upon the phenotype, but the effect of selection is to regulate the distribution of genes within the population and the distribution of genes is not represented on this landscape. A path along which certain phenotypic measures increase monotonically and linearly may hint at an underlying continuous genetic path. Lande, at least occasionally (e.g. 1978), behaves as though continuous genetic variation maps quite directly upon continuous phenotypic variation, and that phenotypic clines correspond to phylogenetic trends. And, to some extent, this is obviously true. Quantitative genetics is based upon the idea that sets of modifier genes increment or decrement the expression of a trait: more "+" modifiers produce more of the trait, and the degree to which the trait is expressed is a simple function of the

accumulation of the modifiers. But this is certainly an abstraction, so simplified that it departs radically from the reality of genetic function. This deviation from reality is particularly severe when reality is multivariate.

In a multivariate universe, increasing or decreasing the value of a phenotypic character depends not only upon one set of modifier genes, but also upon interactions among genes. Some genes, referred to as pleiotropic, influence many characters at once. Thus the relevant parameters are not the number of modifiers but the net effect of the whole set of genes, and their interactions, upon the phenotype. The genetic variance-covariance matrix represents the net effect of genetic interactions upon the phenotype, but the genetic variance-covariance matrix, which gives the underlying trajectory between paths on the adaptive landscape, does not uniquely specify a particular phenotype nor adaptive value. Two phenotypes, of similar adaptive value and apparently similar morphologically, may be very dissimilar genetically. The path between these two phenotypes may not be accessible on the underlying genetic map. In particular, if moving between the two phenotypes requires disrupting correlated characters that are mutually regulated by the same pleiotropic gene, then the apparently simple phenotypic change may require major genetic changes. An intermediate phenotype, which appears to lie between two others along the continuous axes of the phenotypic landscape, may actually lie far away from both of the

endpoints in genotype space.

There is no necessary incompatibility between this phenotypic landscape and an underlying landscape derived from quantitative genetic theory. The phenotypic adaptive landscape could be reformulated by drawing a graph comprising two disjunct planes, one phenotypic and one genotypic, along with the mapping function that takes genotypic variables into phenotypic variables. The genotypic axes can be defined by the genetic breeding values in a population rather than particular gene combinations or gene frequencies. The angles between the axes would then give the genetic correlations. This, albeit complex, genetic surface maps onto the phenotypic surface by a function which may be labeled as "development". Such a graphical analysis completes, in a theoretically consistent fashion, the evolutionary landscape metaphor. It may not provide the simplest or most fundamental representation of the relationship between the basic causal agents of phenotype, the genes, and the aspects of phenotype regulated by genes. But it defines the relationship between genotype, phenotype and development.

However, there is a serious limitation on the utility of such a revised landscape. Unfortunately for its heuristic value, the information necessary to predicting phenotypic change lies in the "Development" mapping function. Phenotypic evolution cannot be analyzed without describing the relationship between the genotypic and

phenotypic surfaces. A purely genetic analysis of the evolutionary dynamics says nothing about the actual transformations in phenotype. Without analysis of development there is limited potential for explaining phenotypic change. But this is precisely what quantitative genetic theory treats as irrelevant to the landscape.

Traditional quantitative genetics supplies a road map of genetic change, and a phenotypic adaptive surface. Unfortunately, it does not suggest how to take this road map and use it to define pathways located on the the phenotypic surface. The phenotype is not completely specified by the genotype because the additive genetic contribution is not the exclusive source of phenotypic patterns of variation and covariation. There is no explicit, well-articulated relationship between the additive genetic variance and the observed phenotype. Traditional quantitative genetics does not present any model to represent the mapping function between genotype and phenotype. Lacking a function that links phenotype to genotype, the adaptive landscape is deprived of critical information and restricted in its ability to explain the movement of a population over the adaptive landsape. Phenotypes can be placed on peaks or valleys corresponding to their relative fitness, but they do not evolve on this phenotypic surface. Instead, the evolutionary mechanisms lie in changes in the underlying genetic structure of the population. The dynamic description requires a whole set of additional mechanisms left out of

the representation of the phenotypic landscape. As it stands, the phenotypic adaptive landscape is static.

A different heuristic model supplies different other terms for understanding phenotypic evolution. The metaphor of the "epigenetic landscape" details the pathway taken by a phenotype as it progresses through ontogeny (Waddington, 1957, 1976). The epigenetic landscape refers to the changing phenotype as it moves through ontogenetic time along the potential pathways of phenotypic change (Figure 1). Recently, the geometry of this landscape has increased in complexity with the introduction of catastrophe theory (Thom, 1975; Waddington, 1974). Regions of the landscape are separated from each other by various cusps and more complex catastrophes because epigenesis involves hysteresis. For example, patterns of chondrogenic condensations in the vertebrate limb can be predicted by forces acting between cells and the density of chrondrogenic cells (Oster et al., 1985). Using stress, motion, and conservation equations, the model predicts particular, discrete spatial patterns generated during chondrogenesis. According to the model, early limb bud chondrogenic condensation patterns depend upon osmotic deswelling of the extracellular matrix, which brings the cells closer together, generating intercellular traction forces as a function of increased cell density. This model predicts bifurcations in the pattern of chondrogenic condensations.

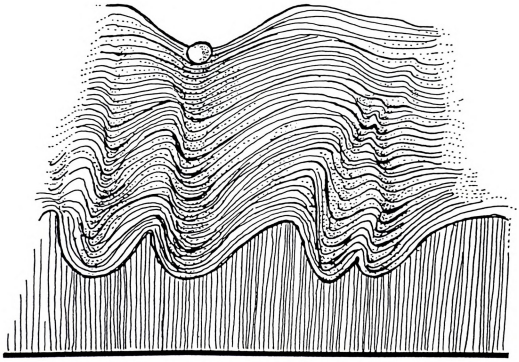


Figure 1.-- The epigenetic landscape (after Waddington, 1976). The phenotype (represented as a ball) proceeds along the valleys of the epigenetic landscape. The pathway of phenotypic change is envisioned as homeostatic. However, the phenotype, when displaced, returns not to its initial position but to another, later poiot along the pathway of change from which it was diverted.

This metaphor of the epigenetic landscape might seem to capture the necessary element of development lacking in the traditional quantitative genetic approach to phenotypic evolution: phenotypes, which occupy different regions of the landscape at various stages of development, are connected to each other through the mapping function of development, with its stable states, transition rules and discontinuities. The paths of phenotypic change are given in the dynamic equations, not simply by genotypes. Genotypes are far removed from the landscape because the genes are non-specific in action.

Unfortunately such models of the developmental process say little about the potential evolutionary modification of morphogenesis. These models may illuminate the mechanisms responsible for chondrogenic condensations, and explain why chondroblast populations of different origins and density produce different cartilage patterns. However, it does not illuminate the evolutionary mechanisms that cause changes in the structure and proportions of the limb. Although these models are used to explain evolutionary change in limb morphology (Shubin and Alberch, 1986), the agents of evolutionary change, the genes and their interactions, play no role in this model. Genes and local environment determine the state of the field variables. In effect, alterations in genes or environment are potential parameter perturbations in the dynamic system. But the mapping function that takes field variables to phenotype cannot be reduced to a genetic

description of morphogenesis.

Alberch (1982) argues that developmental explanations of morphological change need only assume that a genetic basis for the change exists. Specific information about actual genetic changes are unnecessary. Many different alleles, at many different loci, may perturb the same parameters; thus an epigenetic system does not specify a genetic one. Various genotypes may be put into the chrondrogenic development function and they will end up at the same phenotype, as long as they do not produce sufficiently lower levels of hyaluronate or hyaluronidase, or have cytoskeletons that do not respond to stress, or have some tendency to migrate at critical densities. Many different genotypic combinations map to the same phenotype, and genotypes that differ in only one allele may map to phenotypes separated by cusps.

Evolution, according to this epigenetic landscape, cannot be depicted as a continuous process of changes in either genes or gene frequencies. There is, in effect, no genetic dimension to this landscape. But genetic parameters are essential to an evolutionary theory because the evolutionary response of phenotype to selection, random genetic drift or mutation depends upon the genetic basis of the trait (Lande and Arnold, 1983). The epigenetic landscape gives no more than a purely phenomenological description of the evolution of morphology because phenotypic change is not referred to evolutionary mechanisms. Yet, evolutionary

mechanisms must be explicable in terms of changes in the genetic structure of related populations. The epigenetic landscape is visually intriguing, but it captures no more information about evolutionary dynamics than the phenotypic adaptive landscape.

It should be possible to supplement the epigenetic landscape with an additional genetic plane. However, the theory as it currently stands is fundamentally vague on the relationship between genotype and phenotype. While the relationship between genes and epigenesis is characterized as "hierarchical" (e.g. Alberch, 1982), the critical aspect--the structure of the genetic dimension of the relationship--has not been explicitly detailed. The role of the genes is clear: they code for proteins. The proteins influence biochemical interactions within cells, cell properties, and inductive ability of tissues. Individual variation in protein structure, in rate of synthesis, in reaction kinetics, etc. creates phenotypic variation within a population. Thus the epigenetic landscape appears to overlie a genetic axis defined by individual gene combinations.

An alternative, continuous, genetic axis could be constructed to fit the epigenetic landscape. However, there is no theory incorporated in this view of morphogenesis which describes how the genetic basis of the dynamic systems can be defined or how they can evolve. The epigenetic description of phenotypic change permits modification of the

phenotype, but does not implicate any evolutionary mechanisms. In essence, populations do not evolve upon the epigenetic landscape, because they do not alter their genetic structure. They can do no more than change phenotypes. While these morphogenetic explanations of phenotypic change explain how slight changes in genetic structure can generate predictable discontinuities in phenotypes, it does not relate these changes to evolutionary mechanisms: selection and random genetic drift.

Despite the controversy, both traditional quantitative genetic theory and this version of macroevolutionary theory submit essentially static theories of evolution. Traditional quantitative genetic theory cannot describe how the population moves through the phenotypic adaptive landscape because it lacks a trajectory along which the phenotype can move through the landscape. The trajectory lies in the genetic landscape, but there is no well-defined function that relates this genetic trajectory to paths available on the phenotypic landscape. I have defined the mapping function between them to be development, but this definition merely names the missing function; it does not characterize it. Incorporating an underlying genetic landscape to describe the path between phenotypes emphasizes that the function mapping from genotype to phenotype is left out of the analysis. The alternative approach to morphological evolution directly concentrates upon the developmental process itself, but lacks the genetic

trajectory necessary to any analysis of evolutionary mechanisms. While this epigenetic approach recognizes the importance of the genetic basis of evolutionary change, there is no genetic dimension to the landscape in which genetic changes within the population are related to resultant change in the phenotypic trajectory. There is no way to incorporate population dynamics into this description of the developmental process.

Morphogenesis and quantitative genetic parameters

Since the traditional quantitative genetic approach lacks exactly what morphogenetic analysis emphasizes, and the morphogenetic approach lacks what the traditional quantitative genetic theory provides, a synthesis of the two might offer a dynamic approach to phenotypic evolution. Unfortunately the rhetoric, particularly the opposition to quantitative and population genetic models expressed by the adherents of the morphogenetic approach, indicates a distaste for the assumptions and methods involved in such a synthesis, and suggests that a substantial chasm lies between them. Morphologists and geneticists are presumed to ask entirely different kinds of questions and, most importantly, "... In themselves, these genetic approaches offer little insight as to mechanisms of transformation of morphology. Morphogenesis and its modification in evolution do not enter into the equations in any way..." (Maderson et

al., 1982). The equations of quantitative genetics and microevolutionary theory are inadequate because 1) they construct merely genetic models; 2) they employ simple linear models; and 3) statistical descriptions of genotype and phenotype distributions fail to capture the process description of morphogenesis.

These three objections appear to undermine seriously the utility of quantitative genetic models. If models consider only genes, then they may be irrelevant to a study of changing phenotypes. If the linear models are inherently inapplicable then no analysis of quantitative genetic parameters will yield any insight into phenotypic evolution. And if the statistical approach is fundamentally incapable of capturing critical information about developmental processes, then the quantitative genetic approach can never hope to comprehend how particular transformations occur in phenotypic evolution. If these arguments were justified, then the incorporation of traditional quantitative genetics into evolutionary theory would, at best, be useless and, perhaps, even misleading. However, none of these objections is sufficiently well-founded so as to discredit the application of quantitative genetic theory to morphological evolution.

The quantitative genetic models are not merely genetic. It is not as though the genetic parameters of the model are somehow unrelated to phenotype, or incapable of interpretation in terms of phenotypic variation. The genetic

parameters of the model refer to net effects of genes upon phenotype. The description of the phenotype is reduced to a genetic description. But this reduction does not result in some catalog of all genes in the population, nor in a description of the allelic differences at particular loci in different populations. It is not as though quantitative genetic models concentrate solely upon genetic evolution. The genetic description is not deduced for its own sake. Instead, the reduction is performed to permit analysis of evolutionary mechanisms which act upon the heritable portion of phenotypic variation. The purpose of the quantitative genetic analysis is to discern those phenotypic features responding to evolutionary mechanisms. On first principles, the only phenotypic variation that is relevant to evolutionary theory is the heritable variation. Thus the quantitative genetic models extract that part of phenotypic variation which determines the evolutionary response to selection.

Of course, if the bulk of phenotypic evolution is in fact due to changes in the non-additive component of genetic variance, then quantitative genetics, as it currently stands, offers little insight into macroevolution. And there is evidence that the additive genetic variance does not accurately predict observed patterns of phenotypic variation and covariation. Non-additive genetic variance may be responsible for the difference between genetic and phenotypic variance-covariance (Cheverud, 1982 and

references cited therein). Epistasis as well as environmental sources of variance may be a significant source of phenotypic variance neglected by quantitative genetic analysis. Because of non-additive (and non-genetic) sources of variation, analysis of genetic variation might not adequately describe the distribution of phenotypes. Because of the discrepancy between genetic and phenotypic patterns of variation and covariation, the study of the mechanisms which determine the phenotype should not neglect the analysis of phenotypes in favor of genotypes.

However, there is no evidence that the net effect of morphogenetic processes is non-additive. Certainly the processes may be non-linear, but the net effect of these many non-linear processes may be additive. Evidence of the importance of non-linear effects lies in the absence of a satisfactory account of discontinuous changes in phenotype. Despite the long commitment to quantitative genetic analysis, the best explanation of discontinuous phenotypic change is the model of threshold characters. Yet, the argument over the generality of non-linear effects may confuse the ubiquity of non-linear developmental processes with the frequency of non-linear effects. After all, the net effect of many non-linear processes may be additive. A linear model does not presuppose that the causal process is linear, but it does assume that the consequences are.

No one argues that this assumption is not difficult to justify, but the assumption is not unique to quantitative

genetic models. Arguments against linear models should not be directed against quantitative genetic theory because the assumption of linearity is inherent in the statistical techniques employed by both quantitative geneticists and morphogenetic macroevolutionists. The analysis of heterochrony and disassociation, and all multivariate analyses of phenotypic change make the same assumption. Arguments about the adequacy of linear models should motivate a search for more sophisticated approaches to defining the quantitative parameters rather than a rejection of the statistical framework.

Perhaps some of the objections to the linear models as paradigms of macroevolution arise from the belief that evolutionary novelties originate by discontinuous changes in discrete variables. Given the interest in stable states of dynamical systems and hysteresis, change in the distribution of continuous characters might appear to be of little macroevolutionary importance. Yet, changes on the order of 100 standard deviations in the average limb proportions within skinks have been estimated (Lande, 1976) and this should surely qualify as a macroevolutionary change. Abrupt changes in continuous characters even provide most of the examples of punctuated equilibria (Eldredge and Gould, 1972; Gould and Eldredge, 1977). And changes in the patterns of covariance among phenotypic characters might define one kind of macroevolutionary event responsible for creating novelties (Bookstein et al., 1985).

Certainly as quantitative genetic theory now stands, morphogenesis and its modification do not enter into the equations in any articulated, explicit way. Developmental constraints do enter into the theory (Cheverud, 1984; Lande, 1985, 1986), but not in such a way that the particular developmental mechanisms that impose constraints can be extracted from the equations. The absence of any articulated approach to developmental constraints, however, does not reflect a contradiction between quantitative genetic theory and process descriptions of morphogenesis. Such a contradiction could only arise if the process descriptions of morphogenesis were antithetical to quantitative genetic analysis. But the relationship between quantitative genetic analysis and morphogenetic process models is not one of thesis and antithesis. In essence, the morphogenetic process models specify causes of phenotype, while quantitative genetic analysis measures the net effects of these causes.

Quantitative geneticists seem inclined to neglect the mechanisms of development, but the parameters estimated by quantitative geneticists (breeding values, genetic and phenotypic variances and covariances) are not somehow divorced from the developmental mechanisms. Developmental mechanisms, especially the epigenetic interactions among the measured phenotypic characters, determine the parametric values of the measures of phenotype and genotype estimated by quantitative genetic analysis. None of the quantitative

genetic parameters point explicitly to specific developmental processes such as differential gene activity, induction, epithelial-mesenchymal interactions, cell interactions, morphogenetic gradients, mechanical interactions, etc. But these specific developmental mechanisms are the causes of phenotype, and thus of the patterns of phenotypic variation and covariation. They generate the quantitative genetic parameters. Despite the lack of any articulated theory of development, quantitative genetic analysis estimates the net effect, upon phenotype, of morphogenetic processes.

Before any developmental processes can be implicated as causes of phenotypic covariance, the causal basis of covariance must be extracted from the observed patterns of covariance. The studies presented here describe a method for explicitly detecting the developmental basis of phenotypic covariance from traditional parameters estimated by the study of patterns of variation and covariation. These studies, however, differ from traditional quantitative genetic theory in one significant way--they focus upon phenotypic variance-covariation. They thus seek to analyze the constraints acting upon the phenotype. The goals of these analyses of developmental models are: 1) to extract information about processes which cause constraints upon phenotype; and 2) to examine the ontogenetic and evolutionary behavior of the constraints.

Analysis of developmental integration

Olson and Miller (1958) described morphological integration as the interdependence between, and coordination among, the parts of the morphology of an organism. This biological association among morphological characters can be recognized in the patterns of correlations among measures of them. Developmental integration refers specifically to those patterns of covariation created by interactions among characters during ontogeny. Developmental influences that might lead to covariation include growth, both general size increase and local growth gradients; rates of development; timing of developmental events; and tissue interactions. Each of these developmental factors has been implicated as a developmental constraint, although none has yet been demonstrated to determine the patterns of variance and covariance among characters within a population. The study of developmental integration is designed to detect the developmental processes responsible for patterns of covariance among measures. It thus seeks to identify the developmental causes of integration.

Unfortunately, these causes cannot be abstracted by a simple inspection of the variance-covariance matrix. A variance-covariance matrix is complex, responding to numerous causes of covariation. Nor are all the details of specific correlations determined by those developmental processes on which any general theory depends. However, the

associations between the individual characters which result from specific developmental sources of covariation can be detected in the factor-pattern of the variance-covariance matrix.

The additive genetic contribution to each character (X_i) can be simply and traditionally represented as a linear combination of the influence (l_{ik}) of the particular genetic factors (F_k):

$$x_i = l_{i1}F_1 + l_{i2}F_2 + \ldots + l_{ik}F_k + s \qquad (4)$$

where s refers to the component which is specific and unique to that character. When the environment makes no contribution to the character, then the phenotypic and genetic means of that character are the same. Otherwise, the phenotypic mean of the character in the population is the sum of genetic, evironmental and unique components.

The variance-covariance matrix as a whole (V) is specified by :

$$V = lpl^t + E + S \qquad (5)$$

where p refers to the covariance matrix of the factors and l refers to the non-zero loadings of the characters on a factor (Joreskog and Sorbom, 1984). Thus the pattern of zero and non-zero associations between the characters and the factors reflects the joint responses of the characters to the biological sources of covariation. Because these sources of covariation are not directly measured, they are

known as "latent factors".

Many of the recent developments in the analysis of latent variables come from attempts of social theorists to define abstract terms such as "alienation" or "parental values" (see Joreskog and Wold, 1982 for examples of the use of latent variables to measure abstract concepts). The latent variables in these analysis are inherently unmeasurable. Biological theorists, on the other hand, use latent variables for both abstract concepts (e.g. size, shape) and for potentially observable entities (e.g. local morphogenetic gradients, pleiotropic genes). The purpose of analyzing latent variables in biological theory is sometimes be to operationalize latent variables such as size, which can be defined by its measurement model (Bookstein et al., 1985). However, the analysis of latent variables is often designed to determine the effect of specific, but unmeasured, causal agents upon the observed measures. Thus it is the structure of the associations among characters responding to identifiable biological agents that provide causal information about developmental processes. The process models and statistical description of patterns of integration can be united by using process models to predict the statistical parameters of integration. Analysis of specific developmental models which make detailed predictions about the possible developmental factors responsible for observed variance-covariance can suggest the identity of these latent factors.

Phenotypic evolution and the evolution of constraints

The covariance between specific characters that jointly respond to latent factors are potential constraints upon morphological evolution. However, it is not sufficient to demonstrate that developmental mechanisms create patterns of observed integration. These mechanisms might have no effect upon the evolutionary potential of components of integration. Patterns of covariance among characters cannot constrain morphological evolution if they themselves evolve along with morphology. Simple phenotypic evolution, the kind of phenotypic change described by quantitative genetic theory (Lande, 1976; Lande, 1979, Lande and Arnold, 1983, Price and Grant, 1985) occurs when genetic variance-covariance is stable throughout phenotypic change. When these parameters are invariant, then the phenotypic change can be determined by:

$$\triangle X = GP^{-1}s \qquad (6)$$

(Lande, 1976) where P is the phenotypic variance-covariance matrix, G is the genetic variance-covariance matrix and s is the vector of selection differentials. The derived phenotype (X^*) is then a function of the primitive phenotype (X_1) and the incremental change supplied by selection, constrained by the primitive genetic covariance :

$$X^* = X_1 + GP^{-1}s \qquad (7)$$

It is the invariance of genetic integration which makes the microevolutionary process constrained.

The patterns of genetic integration observed in any one population are therefore constraints upon phenotypic evolution only when they are, themselves, historically constrained. However, as is traditional in conventional microevolutionary theory, this formal representation of constraints concentrates solely upon genetic integration. But the constraints upon the phenotype are not exhaustively represented in the additive genetic variance-covariance. To detect the constraints upon phenotype it is necessary to analyze the causes of phenotypic covariance. It is the joint response of phenotypic characters to not only pleiotropy and linkage, but also to epistasis and the environment, which determines phenotypic integration. Developmental constraints upon phenotype are not merely the consequence of the additive genetic contribution to integration. Rather, the patterns of covariance among phenotypic characters reflect the action of developmental constraints. Thus it is the pattern of covariance among characters observed in the phenotype which exhibits the constrained response. And it is the stability of these patterns which must be historically constrained.

Only when the patterns of integration are invariant during phylogenetic and morphological evolution do they

reveal a history of stable developmental constraints. The evolutionary constraint lies in the stability of the pattern. Currently there is no way to evaluate either the frequency or importance of changes in patterns of phenotypic integration. There is no general theory of constraints that covers both stable and dynamic constraints. Thus there is no way to ask how developmental constraints influence phenotypic evolution. Constraints can change, and different kinds of changes may yield different kinds of consequences. Some of the changes in constraints may result from changes in genetic integration. On the other hand, some of these changes may arise from changes in the development function which maps from genotype to phenotype. Whether it is the additive or non-additive component of the genotype which changes, or even if the change lies in the ability of characters to respond to the environment, the changes in constraints are evident in the novel pattern of phenotypic integration. These changes might alter any aspect of the structure of correlations: factor pattern, the covariance among factors, the ability of individual characters to vary uniquely. In the absence of a typology for the changes, the dynamic behavior of constraints cannot even be described, much less probed and examined for the evolutionary consequences. The methods developed in these studies permit direct analysis of the constraints upon phenotype and their behavior throughout ontogeny and phylogeny.

The causal analysis of constraints is designed to

identify the biological factors responsible for observed patterns of phenotypic integration. The first study presented here describes how confirmatory factor analysis (Joreskog, 1969, 1975; Joreskog and Sorbom, 1984) can be used to identify the developmental factors of integration. The second study uses confirmatory factor analysis to analyze and compare the causes of integration in skeletal measures during post-natal growth in laboratory rats (Rattus norvegicus), using the classic data on the ontogeny of integration published by Olson and Miller (1958). The third study employs the comparative study of patterns of integration to examine the stability of these patterns during the evolution of two related lineages of Pentremites, a Missippian blastoid.

There are two different kinds of hypotheses presented in these studies: 1) causal hypotheses which predict the patterns of phenotypic integration in a single population; and 2) hypotheses about the stability of phenotypic integration throughout growth and evolution. The causal analysis identifies those biological processes (e.g. general body growth and local growth gradients, embryonic tissue origin of the characters) which might determine the patterns of integration. Thus these developmental processes are potential constraints upon phenotypic integration because they create covariation among phenotypic characters. While the causal analysis of integration seeks to identify the specific developmental mechanisms responsible for

covariation among characters, the study of variation in patterns of integration directly examines the stability of patterns of integration. The comparative studies test the hypothesis that covariance among characters constrains morphological change. When these developmental processes not only regulate covariation among phenotypic characters, but also limit the ability of individual characters to respond independently to selection, then they generate constraints upon phenotypic evolution.

EVALUATING GENERAL DEVELOPMENTAL MODELS

The goal of the study of developmental integration is to make detailed statements about the particular developmental influences that generate covariance among characters. Although advances in multivariate techniques have superceded the statistical approach to morphological integration described by Olson and Miller (1958), the study of integration continues to be motivated by an interest in the pattern of covariation among biologically associated characters (Andrews et al., 1974; Atchley, 1984; Atchley and Rutledge, 1980; Atchely et al., 1981; Cheverud, 1982; Eldredge, 1973; Gould, 1984; Gould and Garwood, 1969; Leamy, 1975, 1977).

This study describes a procedure for generating developmental models and demonstrates methods for evaluating competing hypotheses. I constructed developmental models based upon: 1) interpretation of exploratory factor analysis, and 2) developmental theory, to evaluate how well mechanistic explanations for developmental integration predict observed variance-covariance among characters. The intent of this study is to suggest those procedures most likely to generate acceptable models. Techniques for constructing and evaluating causal models are presented and applied to a sample data set (covariance and correlation

matrices of a single sample of skeletal characters from one-day-old laboratory rats, Rattus norvegicus) to illustrate methods for testing hypothesized causes of developmental integration.

Data

Developmental models were fit to log transformed osteometric measures taken on a sample of 20 one-day-old laboratory rats, published by Olson and Miller (1958). These laboratory rats are the population studied by Olson and Miller in their classic analysis of the developmental basis of morphological integration.

I selected eighteen of their measures for this analysis (limb and skull measures are represented schematically in Figure 2, described in Table 1) to provide a relatively even coverage of the cranial and post-cranial axial and appendicular skeleton. Olson and Miller took all length measures of axial characters along the longitudinal body axis while the length of appendicular characters was measured along the long axis of the bones; widths were measured parallel to the transverse plane. The one depth measure was taken from the ventral to dorsal surface of the mandible.

Factor analysis

Traditional exploratory factor analysis is designed to extract unobserved, latent causes of covariance among

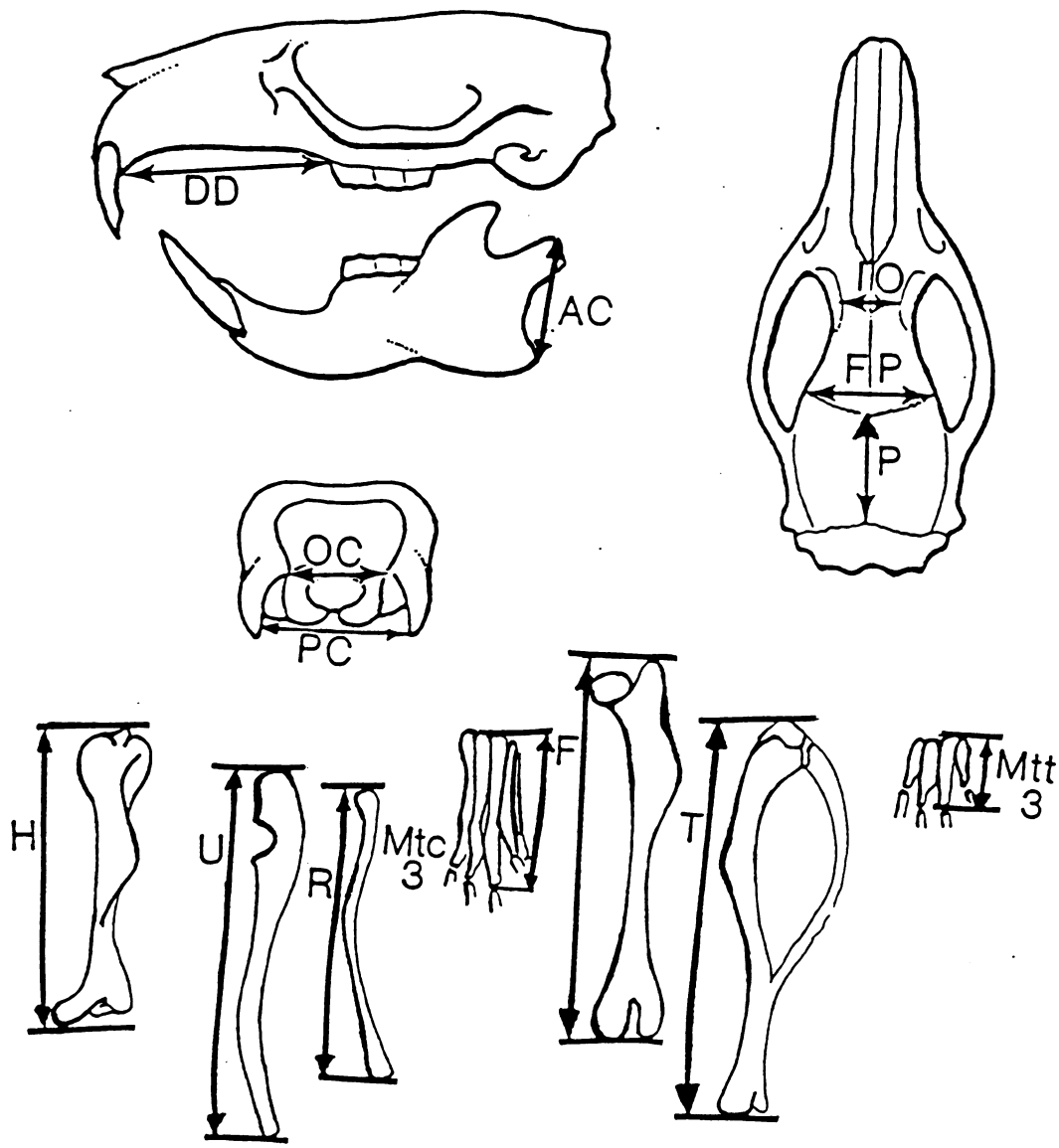


Figure 2.-- Schematic representation of the skeletal characters analyzed. Abbreviations and descriptions of the measures are presented in Table 1.

Table 1.-- Description of characters (Olson and Miller, 1958). The abbreviations specific to Figure 2 are enclosed in parentheses.

No.	Abbreviation	Description
		Skull measures
1.	Fr-Par(w)	Width of the fronto-parietal suture across the vault of the skull, measured as a segment of arc, from ends of the sutures as seen in dorsal aspect.(FP)
2.	Io(w)	Minimum interorbital width, measured normal to skull length.(IO)
3.	Par(l)	Distance along the mid-line from the junction of the interparietal and parietal to the junction of frontal and parietal.(P)
4.	Dias(l)	Length of the dental diastema on the upper jaw, from base of first cheek tooth (or its position in young specimens) to base of incisor.(DD)
5.	Occ(w)	Maximum distance between outer margins of occipital condyles.(OC)
6.	Para(w)	Distance between basal tips of paraoccipital processes.(PC)
7.	Ang-cor(d)	Height of jaw from base of angular process to the top of coronoid condyle.(AC)
		Post-cranial axial measures
8.	Atlas(w)	Maximum width of the atlas.
9.	L9v(w)	Maximum width of the postzygopophysis of the 9th vertebra.
10.	L22v(w)	Maximum width of the postzygopophysis of the 22nd vertebra.
11.	Sacral(w)	Maximum width of the sacral rib.

Table 1 (cont'd).

Limb measures

12.	Hum(l)	Maximum length of the humerus.(H)
13.	Ulna(l)	Maximum length of the ulna.(U)
14.	Rad(l)	Maximum length of the radius.(R)
15.	Mtc3(l)	Maximum length of the 3rd metacarpal.(Mtc3)
16.	Fem(l)	Maximum length of the femur.(F)
17.	Tib(l)	Maximum length of the tibia.(T)
18.	Mtt3(l)	Maximum length of the 3rd metatarsal.(Mtt3)

ıres. The factor variates associated with extracted
›r axes may be regarded as mathematical constructs with
le predictive value, or, in contrast, the latent
ıbles may be viewed as real inferred variables that
:ct the behavior of observed variables (see Joreskog and
1982, particularly Bookstein, 1982b, for a discussion
lternative concepts of inferred factor variates). If
ıt variables correspond to real but unobserved
ctors, then the interpretation of the factor analysis
an identification of these causal variables. I
ed the latent variables as biologically real and used
to suggest a causal explanation for the covariance
en the observed variables.

The ease with which factor analysis discriminates among
endent sources of covariance constitutes a principal
tage of factor analysis. However, the notorious
uity of factor analysis, and its ability to detect
ently meaningful structure in random data (Armstrong,
, set serious limitations upon the use of factor
sis for theory construction or evaluation of hypotheses
ik, 1972). The infinite number of equally valid
ions is only one source of ambiguity. An even more
us limitation lies in the ad hoc causal interpretation
from the pattern of variation and covariation.

I performed exploratory factor and principal component
ses on the rat measures to suggest potential
›pmental causes of integration. The principal

nents approach differs substantially from that of r analysis. In principal components analysis, all nce is explained by the common components, there are no e components. The study of integration should not e common responses among all characters *a priori*. er, principal components analysis is a classic ution of the problem of ambiguity in factor analysis; fore I included it in this analysis for comparison.

Principal components and initial factors were extracted the correlation matrix. All factors with eigenvalues er than 1 were retained for analysis. The initial rs were rotated to find a simple structure in which variable loads highly on only one axis (Harman, 1967; son, 1976; Thorndike, 1978; Thurstone, 1933 and 1947 ss procedures for selecting the rotated factor ion). No rotation successfully achieves ideal simple ture, and all are equally valid. I used both varimax, maximizes the number of very high and near-zero ngs for the columns, and the quartimax rotation, which izes the variance in factor loadings across rows of the r matrix (Thorndike, 1978).

A model derived from interpretation of an exploratory r analysis might seem guaranteed to fit the data from it was derived. However, the hypothetical factor x based upon interpretation of particular exploratory r analytic solutions could fit poorly if the pretation is weak, or if the interpretation ignored as

.al some factors that are actually important sources of
ıtion in the data. Thus I tested the goodness-of-fit of
ıterpretation of the factor analysis, not the fit of the
›ratory analysis itself.

sults of the exploratory factor analysis.-- In none of
:our solutions (Table 2) is any developmental cause of
ration apparent. The principal components, initial
›r solution and the quartimax factor solutions for the
lation matrix all produce a similar structure. In all
, a general factor accounts for most of the variance.
characters, except for the fronto-parietal suture
h, have high positive loadings on the general factor.
second axis comprises a unit of several skull and limb
res. The third axis reveals a group comprising
nts of both skull and post-cranial axial skeleton,
ps reflecting a contrast between proximal and distal
nts of the different skeletal systems.

The varimax factor solution differs considerably from
thers. Both the first and second axes are characterized
›sitive loadings for most characters, but on neither
are the loadings generally high. The fronto-parietal
e width is the only character with a significant
ive loading on any of the axes. The second factor
iates the interorbital width, jaw depth, width of the
brae and all limb length measures. This might imply
ration among two gradients: proximo-distal and medial-
al. The third axis comprises the sacrum, distal

: 2.-- Factor-pattern matrices for the exploratory
)r analysis of measurements on one-day-old rats.

Principal Components			Initial Factors		
FACTOR1	FACTOR2	FACTOR3	FACTOR1	FACTOR2	FACTOR3
r(w) -.27	.69	.34	-.25	.23	.40
.79	.03	-.32	.77	.15	-.29
) .64	.00	.35	.61	-.12	.20
l) .72	.30	.29	.71	.08	.35
) .75	.23	.46	.75	-.03	.59
w) .59	-.63	.23	.59	-.63	-.15
or(d) .79	.03	.02	.77	-.03	-.01
(w) .77	.17	-.10	.75	.16	.00
) .60	-.12	-.48	.57	.12	-.30
w) .58	.21	-.65	.57	.47	-.44
l(w) .68	-.40	.36	.66	-.47	.08
) .70	.46	.05	.68	.34	-.28
) .83	.12	.08	.81	-.13	.00
l) .92	.04	-.16	.92	.13	-.12
l) .81	.32	-.09	.81	.29	.12
) .93	-.22	.07	.94	-.25	-.09
) .89	-.05	-.10	.89	.02	-.13
l) .75	-.02	.06	.73	-.03	-.07

imax Rotation			Quartimax Rotation		
FACTOR1	FACTOR2	FACTOR3	FACTOR1	FACTOR2	FACTOR3
r(w) .15	-.23	-.47	-.20	-.49	-.09
.33	.70	.32	.76	.14	.32
.55	.17	.32	.62	-.02	-.22
.) .74	.24	.14	.73	-.26	.18
.93	.06	.18	.78	-.34	-.43
) .23	.08	.84	.56	.59	-.33
or(d) .51	.43	.38	.77	.08	-.01
w) .53	.52	.21	.75	-.07	.12
.17	.57	.26	.56	.17	.30
) .11	.86	.01	.56	.01	.65
(w) .46	.08	.67	.65	.32	-.38
.70	.42	-.05	.71	-.31	.05
.54	.40	.48	.81	.15	-.08
) .55	.67	.35	.92	.05	.19
) .66	.55	.09	.82	-.24	.13
.55	.46	.66	.92	.30	-.09
.51	.60	.43	.88	.13	.12
) .48	.41	.36	.72	.08	.00

mb long bones and all hindlimb bones, possibly
ting a relationship among serially homologous
ters of the limb.

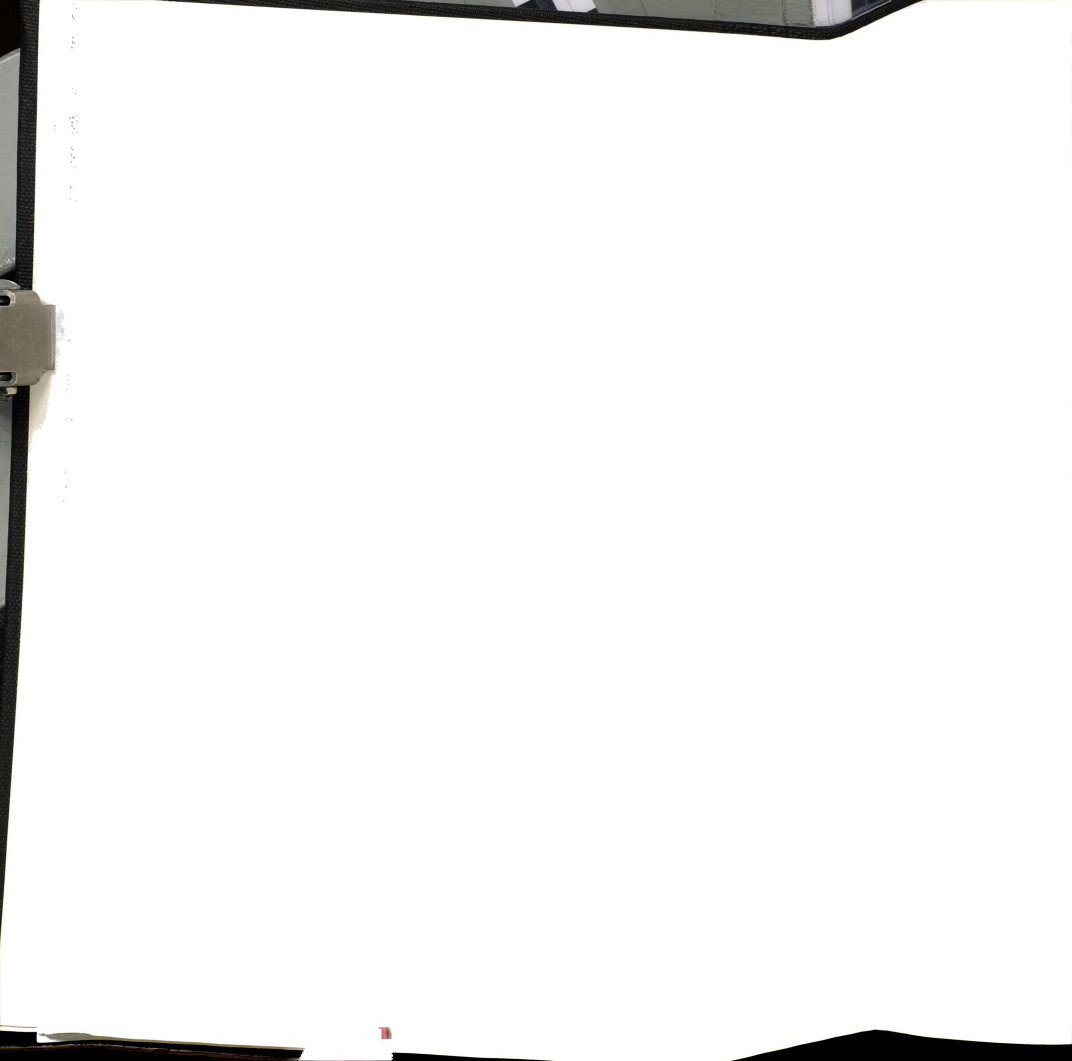
he consensus of the four solutions indicates the
ce of a general factor along which most characters
positively and highly, the standard criteria for
izing a size factor. An additional relationship among
homologs or among characters lying along a proximo-
gradient is also indicated by these solutions. Thus
l body growth and an independent local axial and
o-distal limb growth gradients may be the latent
of developmental integration. However, not all
gs are consistent with this interpretation. Neither
ed gradient affects all characters which lie along the
nt. And some of these gradients, such as the one
unifies the proximo-distal and medial-lateral
ters, lack theoretical justification.

dditional rotations, particularly oblique rotations,
be necessary to discern other potential mechanisms
sible for developmental integration or to corroborate
ypothetical causes of integration inferred from this
is. Perhaps it is implausible to assume that
pmental factors are independent, an assumption
it in the choice of orthogonal rotations. The failure
scern an interpretable underlying structure in these
may possibly result from the small sample size; much
nt covariance may be spurious. However, even multiple

tent solutions, derived from both orthogonal and
e rotations extracted from a large sample of measures,
rely suggest hypotheses. These hypotheses might be
orated by replication of the studies in other samples.
tent structure detected in these other studies might
a similar set of causes acting to determine
ation. Yet, the causal theory remains merely an
nce derived by interpretation of the perceived latent
ure. This causal theory should be subsequently
ted to rigorous testing.

matory factor analysis

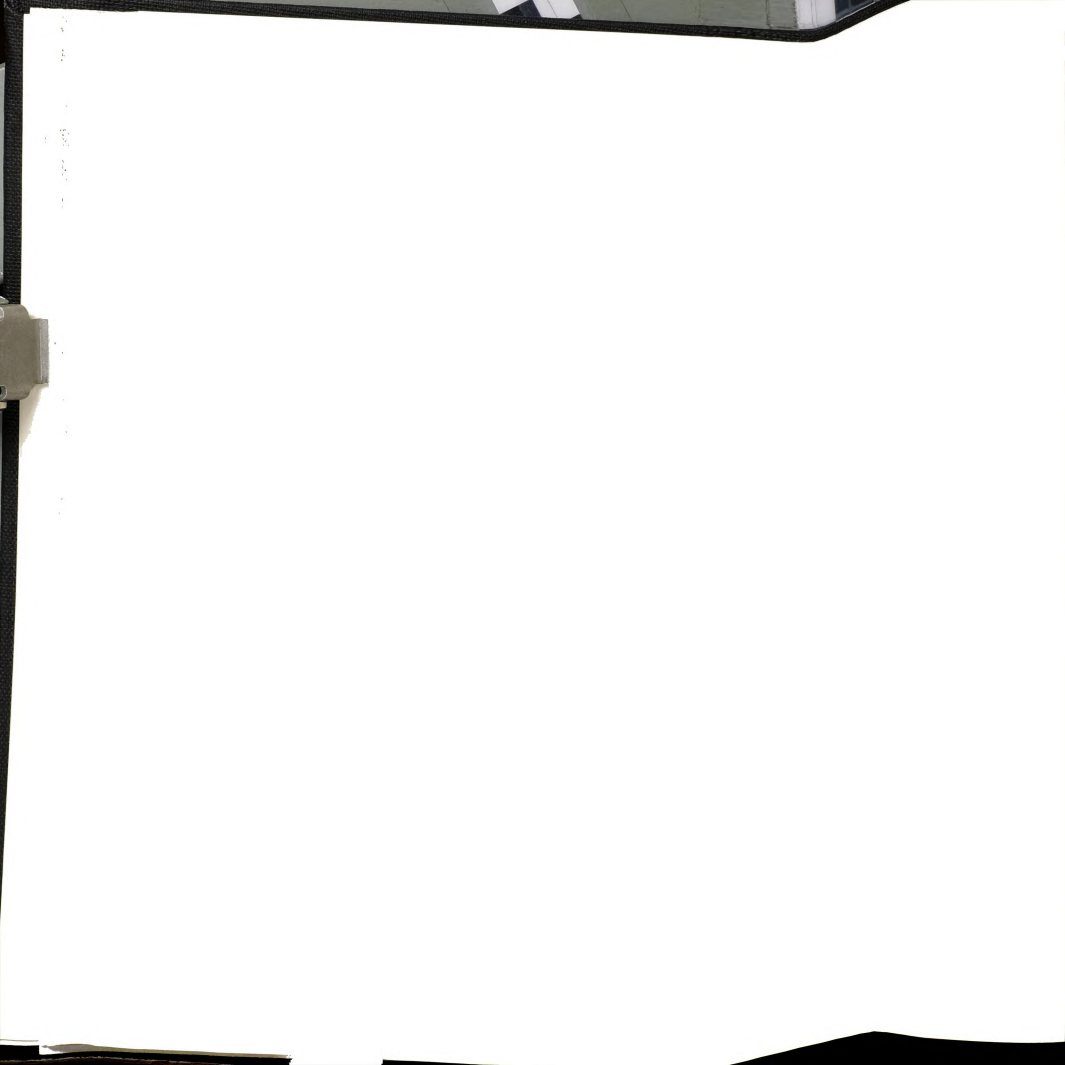
nlike exploratory factor analysis, confirmatory factor
is directly evaluates how well theoretical models fit
ed patterns of variation. The theoretical determinants
pecified a priori, a model is constructed from the
esized causes of covariance, and this model is then
o the observed measures. The association between
factors and observed variables, the regression
cients for the variables on the common factors, the
ance between factors, and unique variances for the
dual variables (or all parameters simultaneously) may
cified in the target matrix and tested.
used the LISREL program (Joreskog and Sorbom, 1984;
ble for IBM-compatible microcomputers) to fit
etical developmental models to the observed measures.
imposes strict constraints upon the hypothetical



: matrix: coefficients must be set equal to a
:ular number or matched to some other coefficients, or
totally unrestricted. Individual coefficients cannot
: specified as greater or less than others, but must be
) some predicted value.

Vhile LISREL can estimate unspecified parameters,
s are rarely identified unless estimates for some of
oefficients are provided. Identification of a model
res that the same value is estimated for any given
eter within the model in all factor-structures
ating the same variance-covariance matrix. Maximum
ihood estimates are not available when the model is not
ciently identified. When identified, LISREL provides a
oodness-of-fit statistic for the hypothesis that the
rved covariance is constrained by the model, against the
rnative that the covariance is unconstrained (Joreskog
Sorbom, 1984). The goal of these confirmatory factor
yses is to not reject the null hypothesis, thus
pting that the model derived from a causal hypothesis
uately reconstructs all observed variance-covariance.

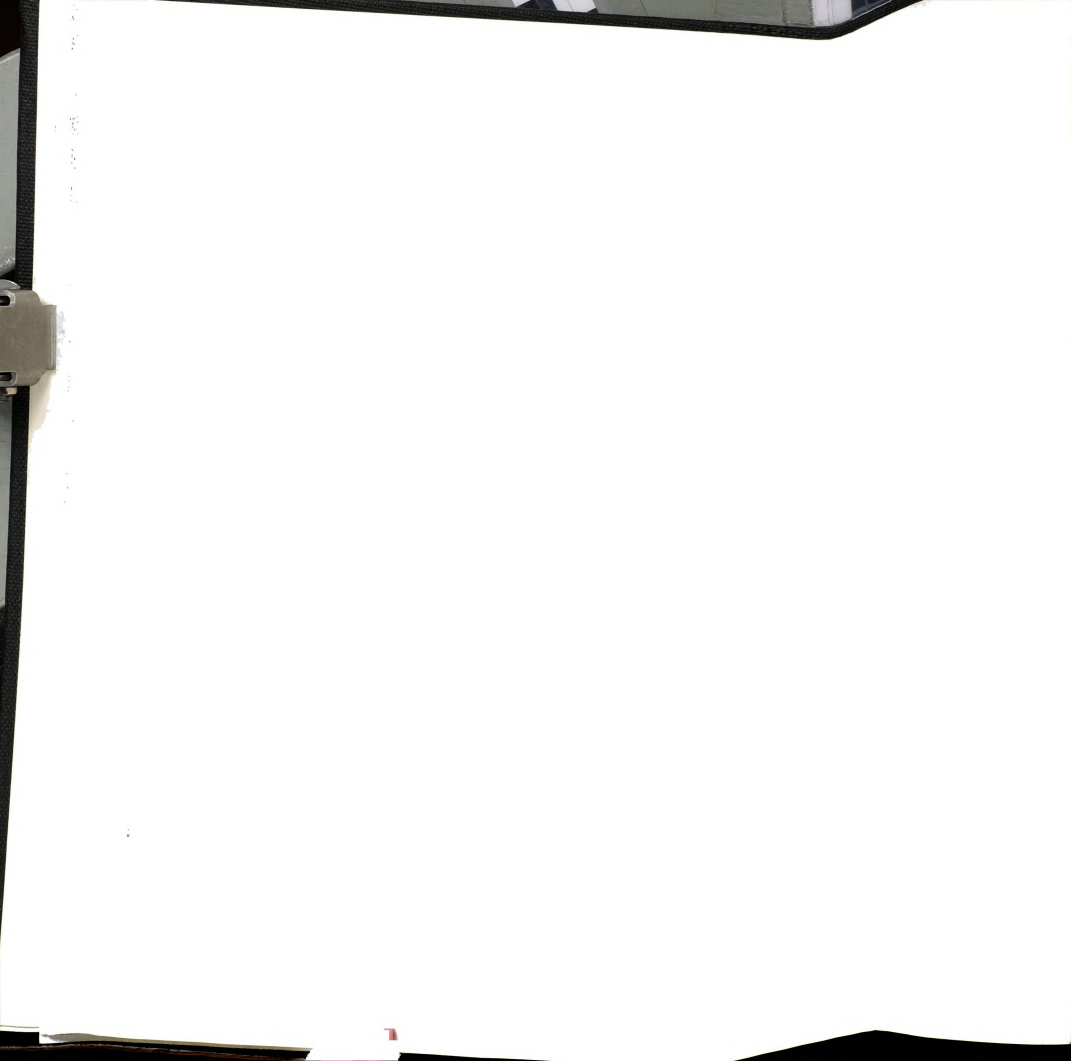
In the following application of confirmatory factor
ysis, I analyze ad hoc models developed from my
rpretation of the results of exploratory factor analysis
a priori models constructed to test biological
theses about the causes of observed integration.



els derived from exploratory factor analysis.-- I fit odels derived from the interpretation of exploratory sis to the correlation matrices, the same matrix zed to produce the exploratory factor solutions. This dure is not inherently circular because it is the pretation of the latent structure, rather than the t structure itself, which is tested.

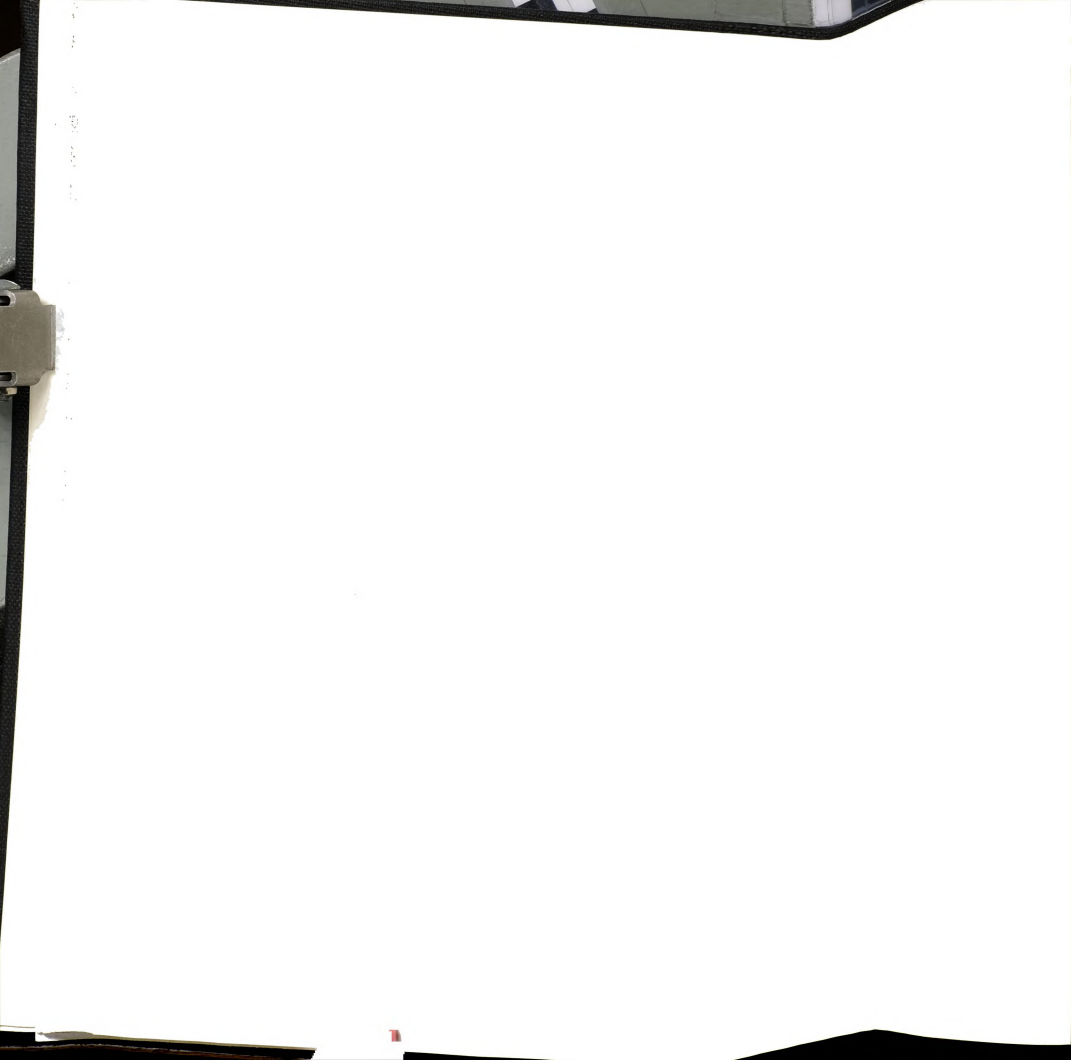
I interpret the exploratory factor analysis to imply the appendicular skeleton lies along a proximo-distal ent uniting serially homologous limb characters. This ent does not covary with the axial skeletal measures, lie along a polarized axial gradient, forming a ar group of cranial and post-cranial axial characters. I also fit both a simple size model, derived by fitting first general axis to all measures, and a more complex thesis which specifies that the cranial characters form unit, the post-cranial characters comprise a second the appendicular characters constitute a third unit the length and width characters form a fourth bipolar

valuation of models derived from exploratory factor ysis fit to the correlation matrix.-- None of these ls fit well (Table 3). Unfortunately, since all models fit to the correlation matrix, the X^2 value has able meaning, at best. The models can be compared, but by ranking them according to their X^2 values relative



3.-- Evaluation of models fit to the correlation of measurements on one-day-old rats.

›del	No. of factors	X^2/d.f.
	1	1.63
- bipolar axial + appen.	3	1.65
cranial+ post-cranial axial ›n.+ bipolar length/width	5	1.71



.e degrees of freedom, not by rigorous statistical

Based upon the relative rank of the alternative models, imple size model fits at least as well as the more ex models derived from the interpretations of the ratory factor analysis. The simple size model even rs to fit somewhat better than the most complex model ising five factors. Thus the secondary factors, which nt for the covariance among characters other than that o general body size, do not add any information above eyond that supplied by the general size model.

This analysis is limited in two ways: 1) because the s are fit to the correlation, rather than covariance, x; and 2) because the models are not devised a priori biological hypotheses but are ad hoc explanations of red latent variables. The only conclusion that can be is that causal interpretations of the biological es of integration, derived from exploratory factor sis, are not guaranteed to fit the data.

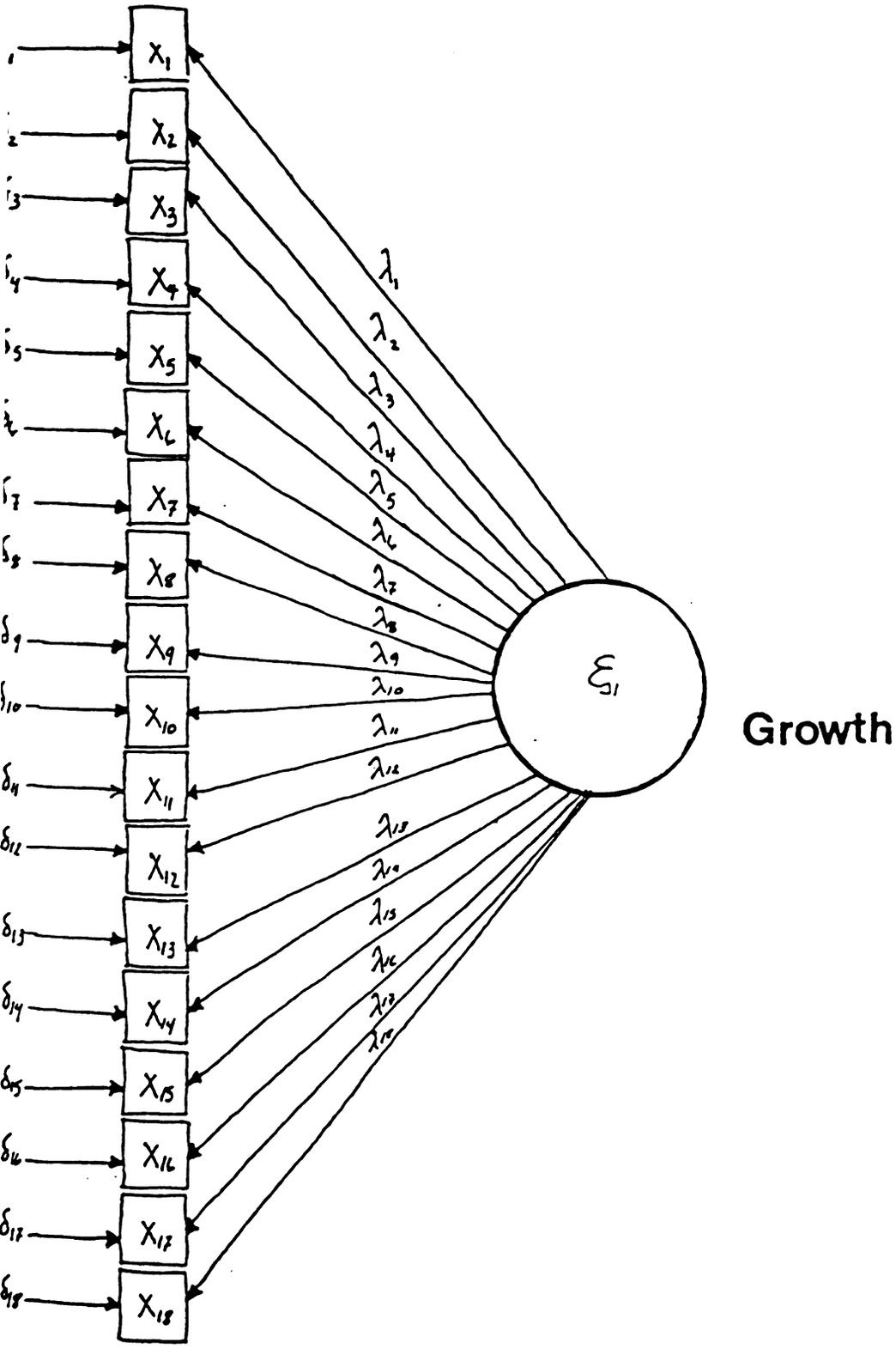
For the remainder of this study, all models will be fit he covariance matrix, and will be developed on the basis heory.

owth models.-- The following example details the edure for generating a hypothetical factor matrix for a le growth model. Sewall Wright proposed that the system orrelations among measures could be viewed in terms of s of causation (1921, 1968). He devised the path

am to graphically represent the interactions between bserved measures and the hypothesized causal latent oles. Path analysis unites the causal interpretation ne statistical description of systems of correlations. ath diagram (Fig. 3) details the interactions among oles expected by the growth model: all measurements of tal characters are influenced by growth, the only influence upon the measures. According to this nesis, the tendency of all characters to increase with asing body size causes the individual characters to .

The growth models incorporate precise estimates for the coefficients. Estimated factor coefficients for the allometry model derive from Jolicoeur's (1963) stration that the first principal component of a iance matrix of log transformed measures is a general variable. I used the regression coefficients from the principal component to estimate the responses of the cters to the growth factor. Since my data are not tudinal, the response of these variables to a general axis reflects only static allometry.

I fit an alternative timing model. This model tests dea that the timing of ossification should predict the coefficients on a general growth axis. I chose s Gompertz model (Barton and Laird, 1969; Laird, 1966; et al., 1965, 1968) and used her explicit relationship en initiation of development and predicted allometric



gure 3.-- Path diagram for the hypothesis that general dy size determines covariation among all characters. By nvention, observed variables (x), the skeletal measures, e enclosed in rectangles. The latent variable (ξ) growth enclosed in a circle. Arrows represent the direction of e causal influence. The factor coefficients (λ) are timated by the predicted intensity of response of the served variables to the latent variables. The unique riance associated with each variable (δ) is uncorrelated th all other factors.

cients:

$$dt = (\ln k)\ (1/a) \qquad (8)$$

dt is the difference in the time of initiation of pment, k is the allometric coefficient for the le on the general size axis and a is the decay rate ter. I used normal tables for rat development dson, 1924) to determine the timing of initiation of cation, and computed the allometric coefficients based the differences in timing. I computed timing ences relative to the characters ossifying earliest, caled the subsequent characters by this value. Thus loadings provide differential growth estimates rather absolute growth rate estimates. I used an estimate of n Laird (1966).

luation of the simple growth models.-- When the model t to the covariance matrix, the X^2 statistic is the ihood ratio for the hypothesis that the population iance is constrained according to the model, under the ption that the variables are multinormally distributed. ^{2} goodness-of-fit value is a global estimate of the fit e model to the structure of the population covariance x. Fit is thus a function of the model as a whole, of bility to reconstruct the observed variance-covariance x. Instead of treating fit as a function of specific idual correlations, LISREL evaluates the net fit of a to the data.

The small sample size creates a complication in using

obability level of the χ^2 statistic as a test of the
tween the model and the data. Although the X^2 is a
on of sample size and increases as sample size
ses, Boomsma (1982) has shown that in samples where N
there are considerable deviations from the expected X^2
, with a tendency for the calculated X^2 statistic to
large. The X^2 generally stabilizes in samples
r than 400 for models varying in the number of factors
variables. Since morphometric analysis frequently
s a large number of variables and relatively small
sizes, the X^2 value by itself should not
tically lead to rejection of biologically plausible
.

Neither of the growth models reconstructs observed
iance (Table 4). However, despite their failure to
and their conflicting predictions, they both account
a substantial proportion of integration. The Fit
ment (F. I.) between a model of total independence
variables and a substantive model measures the
mation contained within the model of dependence. The
of the theoretically interesting models against the
of total independence (t.i.) among measures given by
er and Bonett (1980):

$$F.I. = (X^2/d.f._{(model\ 1)} - X^2/d.f._{(model\ 2)}) / (X^2/d.f._{(t.i.)} - 1) \quad (9)$$

res the increase in information supplied by the more
ex model. The models, however, must be fit to the

able 4. Evaluation of the static allometry and timing odels.

odel	X^2	df	p
tatic allometry	249.16	152	.000
iming	329.84	152	.000

nce-covariance matrix since the F. I. employs the X2
s.

Judged by the Bentler-Bonett F. I., the static
etry model improves 69% over the hypothesis that no
ration is present in these data, while the timing model
nts for merely 44% more integration than this model of
ntegration. Growth, if not developmental timing, is
bly an important developmental influence upon
iance. However, growth is probably not the only cause
served integration.

nplex factor models.-- The complex models are derived
hypotheses which predict the patterns of covariation
characters forming groups on the secondary factors of
ration (the terminology for these group factors comes
Wright, 1932b and is used by Bookstein et al., 1985).
secondary factors comprise discrete sets of
cters.

The tissue-origin hypothesis predicts that all
cters derived from the same embryonic tissue covary.
skeletal characters, except those of ectodermal neural
origin, are derived from mesoderm. However, the
cular source of mesoderm is critical to morphogenesis.
roblasts differentiating from sclerotome appear to
r substantially from those derived from lateral plate
erm (Kosher, 1983; Zwilling, 1961, 1968). The specific
es: lateral plate, sclerotome and head mesenchyme are
latent determinants of integration. Each source of

letal tissue is thus one of the latent variables, a rce of integration. This hypothesis produces the path gram of Figure 4A.

The path diagram for the alternative geometric model gure 4B) isolates the length measures as a single egrated unit; width measures form another component of egration; the single depth measure does not associate h any other variable.

These two path diagrams specify the target factor ttern matrices represented in Table 5A and B. A target rix translates the path diagram into a factor pattern rix that can be fit to the observed variance-covariance rix. I specified coefficients of the target matrix to her be zero or non-zero. The regression coefficient of variable on each factor was set to a value of 1 to serve a scale for estimating the free coefficients (Joreskog Sorbom, 1984). All other non-zero weights were estimated LISREL. The tissue origin and geometry models were equately identified by specifying the target pattern rices given in Table 5 and by specifying the uniqueness efficients of the scaling variables.

Evaluation of the tissue-origin and geometric models.-- ither the tissue-origin nor geometric model fits the served variance-covariance matrix (Table 6). Furthermore, dged by the F. I., both complex models account for less tegration than the static allometry model. Lacking the owth factor, little observed integration can be

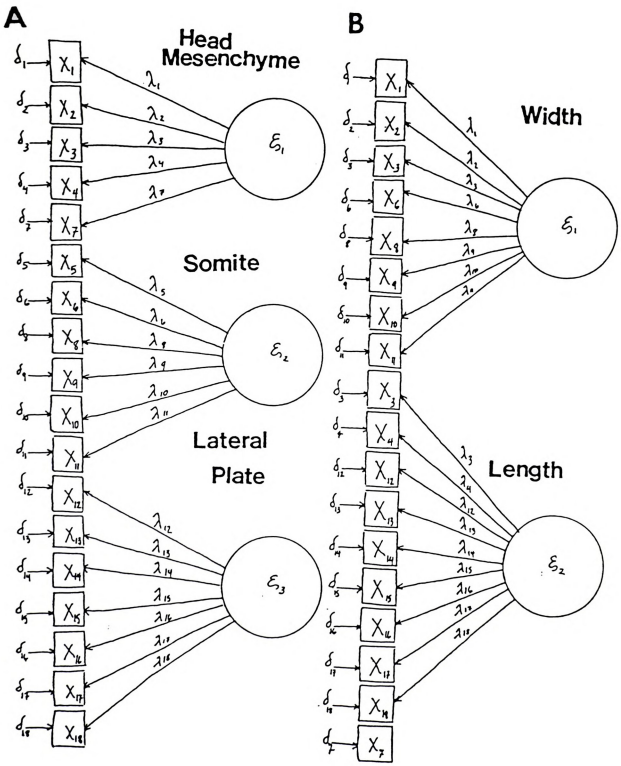


Figure 4. Path diagrams for the hypothesis that origin in a common embryonic tissue (A) or common geometric orientation (B) determine covariance among characters.

Table 5.--Hypothetical factor-patterns derived from path diagrams (Fig. 4A, B). Each variable associated with each latent variable has a non-zero factor coefficient, estimated by LISREL and arbitrarily represented here as 1.

	A. Tissue-Origin			B. Geometric	
Character	Head Mesenchyme	Somite	Lateral Plate	Length	Width
Fr-Par(w)	1	0	0	0	1
Io(w)	1	0	0	0	1
Par(l)	1	0	0	1	0
Dias(l)	1	0	0	1	0
Occ(w)	0	1	0	0	1
Para(w)	0	1	0	0	1
Ang-cor(d)	1	0	0	0	0
Atlas(w)	0	1	0	0	1
L9v(w)	0	1	0	0	1
L22v(w)	0	1	0	0	1
Sacral(w)	0	1	0	0	1
Hum(l)	0	0	1	1	0
Rad(l)	0	0	1	1	0
Ulna(l)	0	0	1	1	0
Mtc3(l)	0	0	1	1	0
Fem(l)	0	0	1	1	0
Tib(l)	0	0	1	1	0
Mtt3(l)	0	0	1	1	0

Table 6. Evaluation of the tissue-origin and geometric models.

Model	X^2	df	p	F.I.
Tissue-origin	244.22	132	.000	.591
Geometric	251.81	131	.000	.584

reconstructed by these complex models. The poor fit of these two models suggests that a size factor must be incorporated into the analysis of the local, secondary factors of integration. However, either tissue origin or geometric orientation of the measures might still explain some integration.

Composite models.-- I combined the geometric and tissue-origin models with the static allometry factor. These models have a more complex factor structure than the other models since they predict that each character is affected by at least two latent variables: 1) general growth, and 2) a local secondary factor, either tissue-origin or geometric orientation.

The relative fit of the tissue-origin and geometric group models does not predict which composite model will fit better. A group model might explain some of the covariance accounted for by the general axis, but little else, thus overlapping the general factor without adding additional information. A poorly fitting group model, when combined with the general axis, might better reconstruct the covariance among all measures if it provides a better hypothesis for the structure of the residual covariance.

Evaluation of composite multiple factor models.-- Neither the static allometry + tissue-origin (Table 7) nor the static allometry + geometric models fit the observed variance-covariance matrix (Table 8). Yet, each might still represent an improvement upon the individual growth,

embryological or geometric components of the model. The improved fit obtained from adding geometric or tissue-origin factors to simple growth models can be estimated by the X2 difference test because the simple models nest within the more complex models. When the models nest and differ in the degree of complexity they can be compared according to their X^2 difference value. A large X^2, relative to the degrees of freedom, indicates that more information might be contained within the data than predicted by the model. A large X^2 difference between models, relative to the difference in degrees of freedom, indicates a significant difference in goodness-of-fit between the two models. Despite relatively poor absolute fit of any individual model, the relative fit of the different models identifies those that can be rejected and suggests which models deserve further examination. Thus we can reject models which capture relatively little of the information contained within the data.

The difference between the X^2 values, relative to the difference in degrees of freedom, is not a statistically rigorous test. The probability level associated with a particular X^2 difference value, relative to the difference in degrees of freedom of the models, only asymptotically approaches the X^2 value for the hypothesis that the model predicts the structure of the data. Furthermore, the small sample size violates the assumptions of this large-sample test. Therefore, I used very stringent values of alpha

< .005 to evaluate statistical significance of X2 differences between models to minimize the risk of concluding that unnecessary components of integration add further information to the model.

The static allometry model fits significantly better than the model of no integration, and the composite static allometry + tissue-origin improves significantly upon the tissue-origin model (Table 7). However, the simple static allometry model accounts for as much integration as the more complex composite static allometry + tissue-origin model. Despite a considerable increase in complexity, no information is contained the composite model that is not already present in the simple static allometry model. The composite geometric model follows the same pattern (Table 8).

Although static allometry is not an adequate explanation of observed integration, it accounts for as much covariance as more complex models. This analysis suggests that neither tissue-origin nor geometric orientation of measures influences observed integration in these data to any significant extent. But there is no basis here for discriminating between the competing tissue-origin and geometric models. Whenever competing models both fit poorly or well, it may be necessary to distinguish between the alternative causal models before eliminating either.

Table 7.-- Sequential evaluation of components of the composite tissue-origin model, relative to the model of no integration. Differences that are significant at the .005 level are indicated by *. ("Allometry" refers to the static allometry model).

Model	Number of common factors	Tests X^2	d.f.	Comparison Models	ΔX^2	Δd.f.
1.No integration	0	471.97	153			
2.Allometry	1	249.16	152	M_1-M_2	222.54	1 *
3.Tissue-origin	3	244.22	132	M_2-M_3	4.94	20
4.Allometry+ tissue-origin	4	220.59	128	M_3-M_4	23.63	4 *
				M_2-M_4	28.57	24

Table 8.-- Sequential evaluation of components of the composite geometry model, relative to the model of no integration. Differences that are significant at the .005 level are indicated by *. ("Allometry" refers to the static allometry model).

Model	Number of common factors	Tests X^2	d.f.	Comparisons Models	ΔX^2	Δd.f.
1.No integration	0	471.97	153			
2.Allometry	1	249.16	152	M_1-M_2	222.54	1*
3.Geometric	3	251.81	135	M_2-M_3	-5.65	17
4.Allometry + geometric	4	221.85	131	M_3-M_4	29.96	4*
				M_2-M_4	27.31	21

Comparisons between the tissue-origin and geometric models.-- Unfortunately, there is no general procedure for directly comparing conflicting causal models since non-nested models cannot be compared rigorously by the X^2 difference test. To compare the conflicting tissue-origin and geometric models, I derived each of them from the set of predictions common to both models. Specific predictions derived from either the tissue-origin or geometric model were subsequently incorporated into the models consistent with both hypotheses. The improved fit of these models comprising specific predictions, over the models devised from predictions common to both hypotheses, could then be calculated by the X^2 difference test.

There were two different target matrices consistent simultaneously with both tissue-origin and geometric models (Table 9). The geometric and tissue-origin models make identical predictions regarding the behavior of the limb measures: all limb measures form a single group according to both hypotheses since all limb bones differentiate from the lateral plate and all limb measures were taken along the proximo-distal limb axis. Furthermore, both models agree that skull widths form a group distinct from appendicular lengths. Skull length measures and the post-cranial axial width measures, however, must be excluded from this shared model. While these cranial and post-cranial width measures covary amongst themselves according to their geometric orientation, they belong to separate groups according to

Table 9.-- Factor pattern matrices for the two different models common to both geometric and tissue-origin hypotheses.

	Model 1		Model 2	
Fr-Par(w)	1	0	0	0
Io(w)	1	0	0	0
Par(l)	0	0	0	0
Dias(l)	0	0	0	0
Occ(w)	0	0	1	0
Para(w)	0	0	1	0
Ang-cor(d)	0	0	0	0
Atlas(w)	0	0	1	0
L9v(w)	0	0	1	0
L22v(w)	0	0	1	0
Sacral(w)	0	0	1	0
Hum(l)	0	1	0	1
Rad(l)	0	1	0	1
Ulna(l)	0	1	0	1
Mtc3(l)	0	1	0	1
Fem(l)	0	1	0	1
Tib(l)	0	1	0	1
Mtt3(l)	0	1	0	1

tissue origin. Thus no model consistent with both geometric and tissue-origin models can include these sets of characters as either a single factor or integration or on two distinct factors of integration. One model consistent with both hypotheses predicts that limb lengths form one unit of integration and that head-mesenchyme derived skull widths form another unit (Model 1 of Table 9). The alternative common model specifies two factors of integration: 1) the occipital and post-cranial axial widths, and 2) the limb lengths (Model 2 of Table 9). Because the occipital and post-cranial axial characters are all derived from somites and follow a common geometric orientation, they form one integrated unit. Similarly, the limb measures jointly originate in the lateral plate and follow a common geometric orientation. This common model excludes skull length measures.

Evaluation of the tissue-origin versus geometric model.-- The two models which incorporate those predictions common to both the geometry and tissue-origin models do not improve upon the static allometry model ($\Delta X^2 = 22.32, \Delta df = 16, .25 > p > .10$; $\Delta X^2 = 10.63, \Delta df = 12, .75 > p > .50$ respectively).

Predictions specific to the tissue-origin model do not improve upon either the static allometry or common models. The addition of skull length to the skull width factor does not account for any more integration than the common model comprising simply skull widths and appendicular lengths ($\Delta X^2 = 13.16, \Delta df = 15, .75 > p > .50$). The only hypothesis both

consistent with the tissue origin model, and improving upon the size model, collapses the head mesenchyme and somite factors into a single unit. This model, similar to that derived from the exploratory factor analysis, does improve significantly upon both the common models (ΔX^2 = 40.94, Δdf = 5, $p < .005$ for the first common model; ΔX^2 = 52.74, Δdf = 9, $p < .005$ for the second) and upon the size model (ΔX^2 = 63.37, Δdf = 21, $p < .005$). The F. I. for this model, compared to the model of no integration, is .80; thus it improves upon the model of no integration by 80%. However, despite its relatively good fit, this model has weak ties to theoretical determinants.

A specifically geometric model merges the skull, occipital and post-cranial widths into a width factor. This model is not consistent with a tissue-origin model since it combines dermatocranial (head mesenchyme) characters with the post-cranial and occipital (somite) group. The skull length measures are not incorporated into this model, but it improves significantly upon both the common models (ΔX^2 = 21.34, Δdf = 2, $p < .005$ for the first common model and ΔX^2 = 33.03, Δdf = 6, $p < .005$ for the second common model). Furthermore, it improves significantly upon the static allometry model (ΔX^2 = 43.66, Δdf = 18, $p < .005$). This model although it ignores all skull length characters, improves upon the model of no integration by 75%.

Adding the skull lengths to the factor comprising appendicular lengths creates a second specifically geometric

model. According to this model, the occipital and post-cranial widths constitute a group distinct from the length measures. Only the second of the two common models is consistent with this specifically geometric prediction. This second geometric model improves upon the common model (ΔX^2 = 30.34, Δdf = 5, p < .005) and the static allometry model (ΔX^2 = 52.41, Δdf=22, p < .005). It improves upon the model of no integration by 76%, despite neglecting all width characters.

The tissue-origin and geometric models do not appear to differ substantially in their relative ability to reconstruct the observed structure of covariation among measures. These laborious comparisons between nested components of competing models are necessary because only nested models can be compared by statistical tests. These two theoretically different models, while not nested, do make several similar predictions about patterns of covariation. When models overlap in their predictions to the extent evident here, but are not nested, no critical tests are possible. Relative fit of competing models might be better evaluated by constructing critical tests a priori and designing measurements to distinguish between the models.

Criteria for selecting measures

These data do not permit any discrimination between covariance due to tissue origin, measurement scheme or local

longitudinal and transverse morphogenetic gradients. A good fit of the geometric model would argue that alternative sets of measures must be employed to discriminate between the hypotheses of covariance due to measurement, a non-biological explanation, and covariance due to local growth fields. A truss measurement scheme (Bookstein, 1982a; Bookstein et al., 1985; Humphries et al., 1981; Strauss and Bookstein, 1982)) and tensor analysis of the truss data (Bookstein, 1984) might more effectively distinguish growth axes oblique to the body axes.

Taking measurements by alternative measurement schemes, such as the truss measurement procedure, and fitting hypotheses to both sets of measures may permit discrimination between those factors responding to the measurement scheme and those biological factors responsible for integration. Fitting the tissue-origin model and geometric models to alternative measurement schemes could increase the ability to discriminate between them. By measuring widths of the long bones, which grow along the long axis, and measuring lengths of the axial characters that grow along the width axis, patterns of covariance due to growth can be distinguished from the covariance due to measurement. Stronger support for the tissue-origin model requires: 1) greater discrimination between the geometry and tissue origin hypotheses, 2) less overlap between the simple tissue-origin model and size, and 3) greater sampling of characters within a skeletal region that are differentiated

from different tissues, such as the skull. The truss measurement scheme, if indeed it does reduce redundant sampling of global size dimensions, should also discriminate between covariance due to general size, or to morphogenetic gradients parallel to measurement axes, or even covariance resulting from common behaviour of characters derived from a common population of differentiating cells.

The kinds of models amenable to analysis obviously depend upon the data they address. Complex developmental models, particularly those that hypothesize inductive or morphogenetic interactions, require highly detailed coverage of the skeleton. Cheverud's (1982) analysis of neurocranial and orofacial clusters exemplifies the kind of refined models that can be tested when the measures ensure full coverage of a single complex system. A tissue-origin model requires less detailed coverage of local components of the skeleton, but needs samples of measures derived from neural crest, skull mesodermal mesenchyme, sclerotome and somatic mesoderm. In effect, the models must be constructed prior to data collection so that the characters which permit discrimination between alternative hypotheses can be identified before collection of data. This discrimination may require analysis of measures chosen by several measurement schemes. Certainly it will always require choosing characters that behave differently according to the competing hypotheses.

Discussion

Despite the frequent use of exploratory factor analysis to suggest sources of covariance in observed data, my results cast doubt upon the value of causal inferences drawn from exploratory analysis. Any exploratory factor analysis will fit the data better than theoretical models because exploratory factor analysis provides a maximally fitting structure for the given number of factors. However, the good fit provided by exploratory analysis may force a sacrifice of a meaningful picture of the causes of integration. Even when hypotheses may seem premature or obscure, and an exploratory analysis most justified, the exploratory analysis cannot reliably suggest a biological explanation for the patterns of covariance. Not only did the interpretation of the causes of covariance suggested by exploratory factor analysis generate several poor models for these data, but none of the exploratory factor analyses support either of the theoretical hypotheses.

The advantages of a confirmatory approach to the analysis of developmental integration are both statistical and conceptual. It permits rigorous evaluation of causal hypotheses that predict the structure of morphometric data. Poor fit is likely when fitting any general model by confirmatory factor analysis because the general model lacks factors unique to each population, such as the nutritional history of a population, which may influence phenotypic integration. The purpose of this analysis is to examine the

patterns of covariance among characters responding to identified developmental influences. Thus it requires a method for detecting the influence of specific developmental factors upon particular sets of characters. These general developmental models do not exhaustively account for all covariance in the data, so they do not fit as well as an exploratory factor solution. However, the ability to reject poorly fitting general developmental models in favor of those better able to reconstruct observed covariance is perhaps the principal advantage of this procedure.

The small sample size employed in this study may hinder identification of the sources of developmental integration, other than growth, in this population. Because of the small sample size, spurious covariance might be indistinguishable from covariance due to the biological factors incorporated in the models. While this analysis demonstrates procedures for evaluating relative fit of alternative models, the apparent failure of all developmental models may well result from small sample size rather than from a weak response to developmental interactions. However, the failure of the developmental models, particularly the failure to distinguish between competing models, suggests strategies for improving model construction and selection of measures.

The ablity to reject poorly fitting models in favor of those better able to reconstruct observed covariance is perhaps the principal advantage of this procedure. This advantage is limited when competing hypotheses fit equally

well or equally poorly. Ambiguous results, such as the virtually identical fit of the geometric and tissue-origin models, motivate an approach to measurement and model construction that concentrate upon the predicted differences between the models. Neither comparisons of fit increments nor estimates of the improvement over some shared set of predictions can provide the necessary critical test. Under these conditions, any attempt to choose between them requires laborious and statistically suspect procedures. The geometric and tissue-origin hypotheses cannot be directly contrasted because the measures were not selected to discriminate between the two hypotheses. Only when the data are chosen according to the a priori hypotheses can critical tests between competing hypotheses be constructed. When one model fits, and the alternative does not, confirmatory factor analysis can supply a method for distinguishing between conflicting explanatory schemes.

Size, by itself, accounts for a substantial portion of the observed integration in these one-day-old rats. Although the apparent influence of size may reflect a bias in the measurement scheme, the joint response of all characters to general body growth may be the dominant developmental constraint. The failure of developmental timing model to fit these data might imply that the simple timing parameters (e.g. age at onset of ossification) and the Gompertz model are inadequate to explain the allometric relations among characters. Fitting growth models more

sophisticated than the Gompertz model (Ebert, 1980; Rickert, 1979; Schnute, 1981) may improve the fit of a timing model. Another source of poor fit may lie in the procedure employed by Olson and Miller for measurement of skeletal characters in one-day-old rats. Since they included unossified cartilaginous models as well as bony elements in their measures, these measures may confound chondrification and ossification rate. Furthermore, the measures were chosen without reference to the location of the centers of ossification. Thus my estimate of the time at which ossification occured for the actual measures is quite rough since the normal tables report timing of ossification for the bone itself, not necessarily for the portion of bone sampled by these measures. Extrapolations from the timing of ossification from one bone to another cannot give precise estimates of timing because proximity of the bones does not adequately predict relative timing of development; for example, the sphenoid initiates ossification twenty-two hours after the onset of ossification in the basisphenoid and forty-eight hours after the onset of ossification of the pterygoid process (Donaldson, 1924).

While size is the dominant source of integration, other factors might be discerned by reducing the spurious covariance and by fitting alternative developmental models. The failure of all developmental models to reconstruct observed variance-covariance adequately indicates that the factors incoporated in the model do not exhaustively account

for observed integration. The results of this analysis emphasize the complexity of the factors that generate phenotypic covariance. Current controversies over the role of development in morphological evolution have concentrated upon few potential sources of developmental constraints, and have neglected to demonstrate that these constraints exert much influence over the developing phenotype. The apparent complexity of developmental integration warrants both more sophisticated attempts to test hypotheses and more caution in suggesting the developmental basis of morphological evolution.

ONTOGENETIC VARIATION IN PATTERNS OF PHENOTYPIC INTEGRATION

Changes in the structure of integration during ontogeny may present a serious challenge to the notion that a particular set of developmental constraints guides phenotypic change. If patterns of integration vary throughout the course of ontogeny, then selection would not be constrained by some set of unbreakable and irresistable constraints. Instead, selection could act upon any adaptive age-specific pattern of integration, however unique to that age. Variable constraints upon morphology would then perhaps influence the time at which selection could act effectively, rather than constrain the evolving characters.

If patterns of integration reflect functional interactions among characters during growth, rather than reflecting only embryological interactions, they might be expected to change considerably within the early stages of life. Laboratory rats develop from hairless, blind neonates to sexually mature adults within six weeks. At approximately ten days after birth the rats open their eyes and also undergo considerable hormonal changes. For example, they end the critical period for responding to testosterone (Swanson and van der Weff ten Bosch, 1963) and attain normal adult levels of somatomedin-like activity (Olsen et al., 1980). Not only eye-opening and changes in hormonal levels

but also changes in function might cause patterns of integration to vary throughout ontogeny. Twenty-one days is approximately the time of weaning of laboratory rats. Changes in patterns of integration might occur at weaning, resulting from changing interactions among skeletal characters engaged in mastication. Furthermore, puberty might influence patterns of integration. Puberty is reached at variable ages in different strains of laboratory rats; forty-one-day-old rats are approaching or are in early stages of puberty (Parker and Mahesh, 1976).

Phenotypic and genetic covariance do appear to change during the course of postnatal growth (Atchley, 1984; Atchley and Rutledge, 1980). Perhaps these differences in covariance structure reflect a change in developmental constraints. But they may be nothing more than minor modifications of a constrained pattern. Before concluding that developmental constraints change along with changes in covariance, we need to identify those developmental processes which constrain covariance and ask if those processes vary over ontogeny. Thus we need both a causal analysis of constraints and a comparative analysis of these constraints over ontogeny.

This study examines the influence of developmental interactions upon observed phenotypic integration in five age-classes of a single population of laboratory rats (Rattus norvegicus). I evaluate developmental models which make specific predictions about the pattern of covariance

among developmentally associated characters. The dynamics of these patterns are subsequently examined by confirmatory factor analysis through a comparison of patterns of integration between the age-classes. Confirmatory factor analysis not only permits rigorous statistical tests of causal hypotheses but also allows comparative factor analysis to be treated as a problem in statistical inference (Joreskog, 1969; Joreskog and Sorbom, 1984; Sorbom, 1974).

The purpose of this study is to identify developmental constraints upon phenotype and to ask if these constraints persist in their influence throughout postnatal growth. I address two questions: 1) do morphogenetic processes constrain observed covariance among phenotypic characters? and 2) are the constraints on integration invariant throughout postnatal growth?

Data

All measures are taken from data published by Olson and Miller (1958). They comprise log transformed osteological characters from five cross-sectional samples of a single population (N=20 for each sample) of laboratory rats at five ages: 1-day, 11-days, 21-days, 41-days and 250-days (adult).

I analyzed limb and skull measures separately. The sample size limits the number of measures which can be analyzed at one time. The small sample size may be at least

partially responsible the failure to fit any developmental models to the data set which comprised skull, post-cranial axial and limb measures (Chapter 2). When samples sizes are so small, it may be difficult to distinguish between biologically meaningful and random patterns of variation. To reduce the random covariance among measures belonging to separate anatomical units of the skeleton, I partitioned the measures into skeletal subsets.

I used the set of limb measures analyzed in the previous study (Chapter 2; measures represented in Figure 2, abbreviations and descriptions given in Table 1). The limb characters comprise measures of the length of the humerus, radius, ulna, femur, tibia, third metatarsal and third metacarpal along the proximo-distal axis. I also analyzed three sets of skull measures. The first set, the inclusive skull set, comprises measures analyzed in the previous study. These measures were chosen to sample characters of the facial, neurocranial, occipital, and jaw components of the skull. I selected measures for another set of skull measures, the cranium and jaw set, in order to improve the coverage of the mandible, maxilla and neurocranium. I removed the occipital characters, and excluded the interorbital width (a character lacking clearly defined anatomical landmarks). In their stead, added a measure of the posterior neurocranium, the parietal-interparietal suture length, to the measures of the fronto-parietal suture length and parietal bone length, and sampled several

measures on both maxilla and mandible (Figure 5). The third set of skull measures uses the jaw characters of the second data set to permit a more fine-scaled study of a single structural unit.

Intensity of integration

The partitioning of skeletal characters into separate sets: 1) inclusive skull; 2) cranium and jaw; 3) jaw; and 4) limb measures is not motivated merely by the small sample size. It also permits estimation of the intensity of integration specific to regions of the skeleton. While the intensity of integration may be estimated as a property of a population, it can also be regarded as a property of a set of characters. In this analysis I estimate the intensity of integration specific to the whole skull, to the jaw, and to the limb measures, in each age-class.

To estimate overall integration, I used the standardized X^2 (Lindgren, 1968) of the model of no integration (n.i.) which specifies complete independence among all measures.

$$\text{Overall Integration} = (X^2 - df_{(n.i.)}) \; / \; (2df^{1/2}) \qquad (10)$$

Asymptotically, the standardized X^2 follows the z-distribution and permits comparisons between standard scores of integration in samples of measures that differ in the number of variables and degrees of freedom. Unlike other

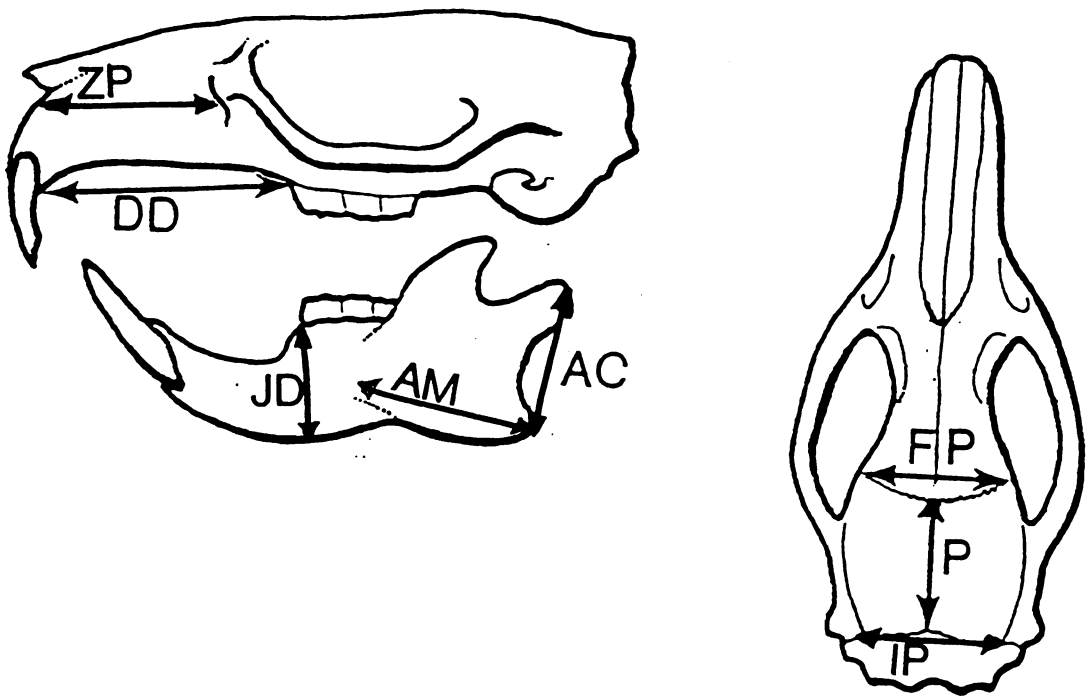


Figure 5.-- Schematic representation of the cranium and jaw measures comprising the second set of skull measures. Abbrevations not defined in Table 1: IP = width of the parietal-interparietal suture; AM = distance between the angular process and the most anterior point of insertion to the masseter; AC = distance between the angular and coronoid processes; ZP = distance from the most anterior point of the zygomatic to the most anterior extension of the premaxilla).

measures of the intensity of integration (Cheverud et al., 1983; Olson and Miller, 1958; Van Valen, 1960), this measure is not scaled to range between 0 and 1.

The intensity of overall integration estimates the intercorrelations among measures. Poor fit of the hypothesis of no integration results in a high X^2 value for the fit of the model. Increasing interdependence among measures increases the X^2 value, and thus increases the estimate of overall integration. This index does not depend upon the absolute amount of covariance in the data, nor upon the number of variables. It differs from other measurres of integration because it does not depend upon average correlations between characters nor upon the average eigenvalues of the components. Rather, it depends strictly upon the inability of the model of no integration to reconstruct the structure of the observed covariance.

The overall interdependence among characters within the population does not exhaustively estimate the parameters of integration. The estimate of overall integration collapses the different aspects of integration into a single summary, reflecting the amount of covariance among all characters.

Estimates of the intensity of integration.-- There is both temporal and regional variation in standardized overall intensity of integration (Fig. 6). While there are no confidence intervals surrounding the standard scores, limb measures appear more integrated than skull measures. Different skull data sets differ, but less strikingly, in

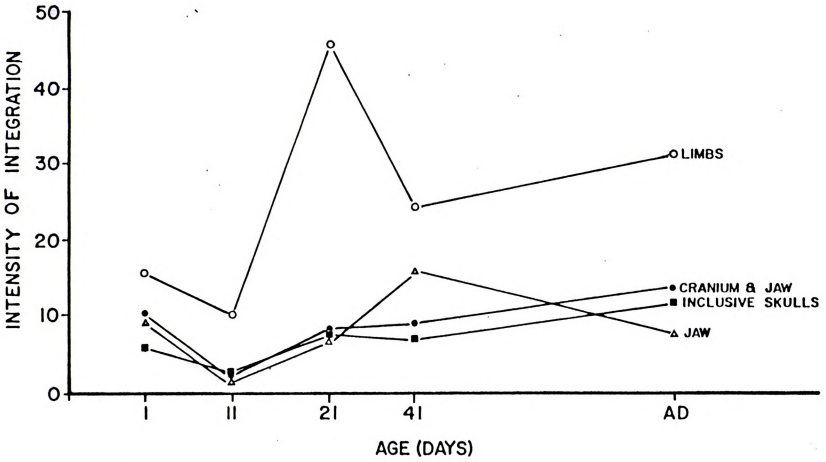


Figure 6.-- Temporal and regional variation in standard scores of overall intensity of integration over post-natal growth in the limb; inclusive skull; cranium and jaw; and within jaw measures.

intensity of integration and in their patterns of temporal variation. The jaws achieve their highest degree of integration at puberty, and subsequently this intensity declines; in contrast, the more inclusive sets of skull measures both increase after maturity. Increasing integration in the post-pubescent skull coupled with decreasing integration in the jaw may reflect a decrease in regionalization of integration.

Causal analysis of patterns of integration

I fit several models, both explicitly developmental models and models of structural integration which depend upon neither developmental nor functional hypotheses, to each set of measures (see Chapter 2 for details of the procedure used for analyzing causal models by confirmatory factor analysis). The goal of this causal analysis is to identify the sources of integration. I fit each model to each age-class; however, not all models are identified in all five age classes (see Chapter 2 for a discussion of model identification).

Size.-- I fit the size model to each data set, hypothesizing that a single, general axis predicts all observed covariance. I first evaluated the fit of this size model, and then included the size factor as a component in all other models. Thus the size model predicts not only that size is a cause of phenotypic integration, but also that

size is the only cause of phenotypic integration.

Evaluation of the size model.-- Size is the predominant source of integration observed among the measures of the limbs (Table 10). The hypothesis that size alone accounts for observed integration adequately reconstructs observed covariance among limb measures at all ages except at adulthood. Thus, only in the adult population is there any secondary factor of integration.

Integration among skull measures is also strongly determined by size (Table 11). Size, however, is not an adequate explanation of the patterns of observed integration among characters in the inclusive skull data at eleven days of age nor in the adult population. Nor is size adequate to account for observed integration among the measures of the jaw and cranium between birth and puberty. Yet, at no age is there evidence that a secondary factor determines observed integration in the jaws. Indeed, integration within the jaw complex is quite weak at eleven days, evident both in the low overall intensity of integration (Figure 6) and in the failure of a general size factor to add any information to the model of no integration ($X^2 = 3.82$, df = 1, $.025 < p < .05$).

Evidently, size is a major source of integration in the skeleton, but secondary factors of integration may be present as well. All other models predict that size is one source of integration; these alternative models differ from each other in their hypotheses about causes of secondary

Table 10.-- Goodness-of-fit values for the size model fit to the limb measures throughout ontogeny.

Age (days)	X^2	df	p
1	17.08	20	.648
11	19.48	20	.491
21	17.63	20	.612
41	23.87	20	.248
Ad	32.44	20	.039

Table 11.-- Goodness-of-fit values for the size model fit to skull measures throughout ontogeny.

Age (days)	Inclusive Skull			Jaw + Cranium			Jaw		
	X^2	df	p	X^2	df	p	X^2	df	p
1	19.10	20	.515	28.71	27	.375	7.39	9	.596
11	35.36	20	.018	37.28	27	.090	12.31	9	.196
21	14.44	20	.808	59.74	27	.000	7.55	9	.580
41	22.22	20	.326	16.52	27	.942	5.25	9	.813
Ad	30.27	20	.066	36.96	27	.096	12.86	9	.169

factors of integration.

Morphogenetic processes.-- I fit two developmental models to the limb measures. The first hypothesis, the limb-bud model, predicts that the forelimb and hindlimb measures each constitute units of integration. According to this hypothesis, the anterior and posterior limb buds are developmentally distinct, a hypothesis suggested by their independent response to some mutations and teratogens (Grueneberg, 1963). I fit an alternative developmental model which recognizes the serial homology of fore- and hindlimb structure. According to this serial homology model, all long bones should form a single unit, while the metacarpal and metatarsal form another unit. By hypothesis, the shared mechanism of distal outgrowth unites the serially homologous fore- and hindlimb bones, while the (mammalian) digits form an integrated unit of characters due to their common dependence upon local cell-death.

To the set of inclusive skull measures, I fit the tissue-origin by hypothesizing that all structures derived from paraxial mesoderm, whether somitic or somitomeric, covary, while neural crest derivatives covary among themselves. Currently, hypotheses about the exact tissue origin of skull bones in mammals are tentative. I presumed that mammal skull bones originate from the same tissue precursors as the skull bones of chicks. Thus the mammalian homologs of the chick parietal and the posterior portion of the frontal bone derive from paraxial somitomeric mesoderm;

the occipital characters derive from somitic mesoderm (sclerotome), while the anterior portion of the frontal bone and the first branchial arch derivatives are neural crest in origin (Noden, 1982).

I fit a hypothesis of embryonic induction to the second skull data set. I asked if the cranial characters, responding to the inductive influence of the neural tube (and perhaps also notochord), form an integrated unit distinct from the branchial arch characters. The cranial vault is highly influenced by the quantity and pressure of the cranial contents, while the facial skull is largely unaffected by cranial contents (Young, 1959). Furthermore, branchial arch derivatives differ in their inductive stimuli, relying upon permissive inductive interactions with mandibular epithelium (Hall, 1982). A good fit of this hypothesis might demonstrate that a shared inductive influence generates covariance. However, not only must the induction model fit the data, but also an alternative, theoretically inconsistent model, should fail. This second model (the cranium + maxilla + mandible model), forces derivatives of the maxillary and mandibular process to form separate units of integration. It conflicts with both the tissue origin and induction models; thus if it fits as well as both, it casts doubt upon tissue origin and induction as mechanisms of integration.

Evaluation of the morphogenetic models of integration.--At each age at least one of the developmental models fit the

limb measures (Table 12). Yet, these developmental models fit no better than the simple size model.

The hypothesis that common origin in neural crest, or somitic or paraxial mesoderm determines integration generally fails to fit the characters of the inclusive skull measure (Table 13). And even when the model does fit, it captures no more information than already explained by either size or the model of no integration at all. The induction model fits in the 11-day-old population, but perhaps the relatively good fit of the tissue-origin model in this age-class is an artifact of its complexity. Since the 11-day-old population is so poorly integrated overall, a complex model, requiring little covariance among the individual factors, may capture this relative independence of the individual characters.

The hypothesis that a shared response to embryonic inductive stimuli determines integration fits the cranium + jaw measures in the 21-day-old population (Table 14). At 21 days, the skull apparently comprises three units of integration: 1) size, 2) cranium, and 3) jaw. However, the equal fit of the conflicting cranium + maxilla + mandible model at this age (Table 15), which separates characters responding to a common inductive stimulus, suggests that no inductive factor forces all jaw characters to covary.

Perhaps there are inductive influences limited to specific regions of the jaw, but it is unlikely that this decomposition of the skull into cranial and jaw measures is

Table 12. Goodness-of-fit values for the developmental models fit to the limb measures in five age-classes. None show significant improvement over the size model as judged by the X^2 difference test.

Model	Age (in days)	X^2	df	p
Limb bud	1	4.94	10	.890
	11[a]	---		
	21	3.65	10	.962
	41	184.30	10	.000
	Adult	31.08	9	.000
Serial homology	1	7.87	11	.895
	11[a]	---		
	21	4.47	11	.954
	41	9.60	11	.566
	Adult	282.95	11	.000

[a] Maximum likelihood estimates not available. The models were not sufficently identified in this population.

Table 13.-- Goodness-of-fit of the hypothesis that a origin in a common embryonic tissue determines observed integration in five age-classes. Significant improvement over the size model, as judged by the X^2 difference test is indicated by *.

Age (in days)	X^2	df	p
1	40.39	10	.000
11	15.83	10	.104 *
21	141.32	10	.000
41	86.76	10	.000
Adult	12.58	10	.248

Table 14.-- Goodness-of-fit of the hypothesis that embryonic induction is a source of integration throughout postnatal growth. Significant improvement over size is indicated by *.

Age (in days)	X^2	df	p
1	19.76	20	.492
11	18.21	16	.312
21	15.74	16	.471*
41	11.13	16	.801
Ad	15.21	16	.509

Table 15.-- X^2 difference tests for the comparison between the induction model and the conflicting hypothesis that the jaw comprises mandibular and maxillary units.

Age (in days)	ΔX^2	Δ df	p	
1	14.06	7	>.050	a
11	4.44	3	>.250	
21	5.40	3	>.100	
41	-- b			
Ad	-.13	3	<.995	c

a The induction model fit to measures 1 day-old rats

b needed fewer restrictions to identify the model. Maximum likelihood estimates not available.

c These negative X^2 values were treated as zeroes. They probably reflect unstable X^2 values resulting from the small sample size.

a delayed, transient response to embryonic induction. In the adult population, the cranium and jaw cohere to form a single unit of integration.

Structural integration.-- I fit models of structural integration to each data set. The unit models predict that all characters within a skeletal subset are integrated. As well as responding to a general size factor, the unit model predicts that all the characters within each skeletal subset constitute a structural unit. These unit models depend upon no explicit or articulated biological hypothesis; rather, they conflict with all developmental hypotheses because they do not distinguish among characters that differentiate from different tissues nor by different morphogenetic processes. A relatively good fit of these models constitutes grounds for suspecting that hypothesized developmental mechanisms are not causes of integration.

Evaluation of the models of structural integration.-- The unit model of integration, which depends upon no causal hypothesis, fits the limb measures as well or better than the size model (Table 16).

However, not all limb measures invariably belong to the limb unit. The third metatarsal cannot be forced to covary with the other limb measures except in the neonatal and 41-day-old samples. In the 1-day-old population, both models which associate limb measures into a single unit fit (X^2 = 6.48, df = 12, p = .890; X^2 = 6.48, df = 13, p = .927). In the 11-day-old population the third metatarsal definitely

Table 16. X^2 difference tests for the evaluation of relative fit of the size and unit models. The 41-day old population is fit to the model of integration which associates the third metatarsal with the other limb measures, other populations are fit to the model which exludes the third metatarsal. Significant differences between the fit of competing models are indicated by *.

Age	ΔX^2	Δdf	p
1	10.06	7	>.100
11	12.49	7	>.050
21	8.99	7	>.100
41	13.33	8	>.025
Adult	26.06	7	<.005*

does not associate with the other limb measures (X^2 = 125.40, df = 12, p <.0001 for the hypothesis that the third metatarsal is united with the other limb measures). This same pattern is repeated in the 21-day-old population when the model excluding the third metatarsal fits significantly better than the alternative unit model (ΔX^2= 277.41, Δdf = 1, p <.005 for the comparison between the two models of integration). The pattern changes in the 41-day-old population. While the third metatarsal is neither necessarily dissociated from, nor associated with, other limb measures in the neonate, removing the third metatarsal from the limb unit renders the model inconsistent with the variance-covariance structure of the limb measures in the 41-day-old population (X^2 = 604.73 df = 13, p < .0001).

However, even when the third metatarsal is incorporated into the secondary limb factor, the model of structural integration fits no better than the simple size model. Only in the adult population does the hypothesis of a secondary factor improve upon the simpler growth model, and the only hypothetical secondary factor which effectively reconstructs more of the observed covariance structure than the size model is this structural model of integration.

In the set of inclusive skull measures, the unit model never improves upon the fit of the size model (Table 17). At weaning, the unit model fits cranium and jaw measures better than does the size model, but signficantly more poorly than the induction model (Table 18). The unit model

Table 17. Goodness-of-fit values for the unit model fit to the inclusive skull data. None show significant improvement over the size model as judged by X^2 difference tests.

Age	X^2	df	p
1	8.11	12	.777
11	80.59	12	.000
21	5.91	12	.920
41	7.69	12	.908
Adult	13.44	12	.222

Table 18. X^2 difference tests for the size versus unit models fit to the cranium and jaw data. Significant differences are indicated by *.

	Size versus Unit			Unit versus Induction		
Age	ΔX^2	Δdf	p	ΔX^2	Δdf	p
1	16.56	9	>.050	7.61	2	>.010[a]
11	17.46	9	>.050	1.61	2	>.250
21	25.47	9	<.005*	18.53	2	<.005*
41	8.07	9	>.500	-2.69	2	<.995[b]
Adult	18.75	9	>.025	3.00	2	<.100

[a] The induction model fit to measures of 1-day-old rats needed fewer restrictions to identify the model.

[b] These negative X^2 values were treated as zeroes. They probably reflect unstable X^2 values resulting from the small sample size.

never improves upon the size model as an explanation of integration in the jaw (Table 19).

Based upon the relative fit of the alternative, competing models, size is an adequate explanation of observed integration in the pre-adult limb. A secondary factor is necessary to explain the covariance among the measures of the adult limb; the best model for this secondary factor is the non-developmental unit factor.

Size is also an adequate explanation of observed covariance in inclusive skull measures in the pre-adult populations, except at eye-opening, when there is no evidence of any general component of integration. The tissue-origin model fits best in this age group, but this good fit is probably due to the complexity of the model. The adult population shows some indication of a secondary unit factor, although the unit model does not account for signicant amount of integration.

Size accounts for observed integration among the cranium and jaw measures except at weaning, when at least three factor are present. At weaning, the neurocranium and jaw form separate units of integration. The jaw retains its integration throughout ontogeny.

Comparative factor analysis

Comparisons between populations subjected to independent exploratory factor analyses suffer from the

Table 19. Goodness-of-fit values for the unit models fit to jaw measures in the five age classes. None show significant improvement over the size model.

Age (in days)	X^2	df	p
1	118.56	3	.000
11	1.82	3	.610
21	8.56	3	.036
41	.35	4	.986
Adult	83.69	3	.000

difficulties of exploratory factor analysis and from additional problems induced by the often improper approach to comparisons of factor structure (Mulaik, 1972). When used for comparative studies, exploratory factor analysis may provide inadequate, and even misleading, representations of the similarity of two populations. A statistically valid comparison of the factor structures of any two populations must not only 1) be based upon analysis of covariance (not correlation) matrices, and 2) compare factor-pattern (not factor-structure) matrices but also 3) rotate the independent factor-pattern matrices to make them as similar as possible (Mulaik, 1972). Without this attempt to make the factors as similar as possible, and to assess the homology of the rotated factors, there is inadequate evidence that apparent differences between the populations reflect different biological causes of covariance.

The procrustean approach to comparing factor-pattern matrices (e.g. Meredith, 1964) finds the hypothetical population factor-pattern matrix from which all the sample factor-pattern matrices differ least. Confirmatory factor analysis extends the procrustean approach and treats the comparison between factor-pattern matrices as a problem in statistical inference (Sorbom, 1974), providing a X^2 goodness-of-fit value for the null hypothesis that the differences between samples is due to chance.

I used LISREL to perform the multi-sample confirmatory factor analysis. I fit all populations simultaneously to a

common model and I also compared successive age-classes to ask if the patterns of integration are disrupted between particular ages during growth. While each population might independently fit a common model, they could still differ among themselves if they deviate from this common model in different ways. To test the invariance of the developmental constraints, I forced the populations to share a common factor-pattern matrix. Populations might still differ in particular factor coefficients, factor variance-covariance matrix and uniqueness coefficients. For example, the populations might each have a size factor, differing only in static allometric coefficients. Or the same two factors might be present, but be independent in one population while correlated in another. Finally, the same factors of integration might be present, but in one population these factors might explain all of the variance in the characters, while in another the characters might each be less strongly influenced by the common factors. I fit increasingly restricted models to the data, comparing first the factor-pattern, then asking if the factor variance-covariance or uniqueness coefficients are the same, and finally forcing factor-pattern, factor variance-covariance, and uniqueness coefficients to be equal.

Comparative analysis of patterns of integration.-- There is no major change in patterns of integration in the limb. All five populations can be fit to the same hypothetical factor-pattern, the model of structural integration (X2 =

41.47, df = 65, p= .990 for the hypothesis that the factor-pattern matrix is invariant). However, there are changes in the details of integration, evident in the poor fit of the more restrictive hypothesis which forces the factor variance-covariance matrix to be the same across populations and specifies that unique variance is invariant across age (X^2 = 236.28, df = 93, p = .000).

Comparisons between successive age-classes reveals that the only interval of change occurs between 1 and 11 days (Table 20). During this interval the proportion of variance specific to the individual characters and the factor variance-covariance both change. However, all five populations can be simultaneously fit to a common restricted model simply by relaxing the restriction that the unique variances, during the interval between 1 and 11 days, are invariant (X^2 = 97.31, df = 98, p = .501). All of the differences in interactions among limb measures are restricted to the interval between birth and eye opening.

The patterns of integration within the inclusive skull data set undergo similar changes in the details of integration during postnatal growth. When all samples are fit to a common model, the unit model of structural integration, the model fails (X^2 = 93.65, df = 60, p = .004). Thus there may be repatterning in the skull. However, comparisons between successive age-classes show that the factor pattern matrix is stable (X^2 = 22.69, df = 24, p = .538) even during the interval when neither the uniqueness

Table 20. Goodness-of-fit for the hypotheses that factor variance-covariance (Fcova) and the uniqueness of individual characters (U) are invariant in limb measures of sequential age-classes.

Age	1--11			11--21		
X^2	df	p	restricted parameters	X^2	df	p	restricted parameters
			None	25.27	34	.861	Fcova
				33.35	38	.684	U
				48.94	41	.185	U+Fcova

Age	21-41			41--Adult			
X^2	df	p	restricted parameters	X^2	df	p	restricted parameters
28.29	34	.743	Fcova	22.51	34	.934	Fcova
				50.98	38	.078	U
				51.21	41	.132	U+Fcova

coefficients or the factor variance-covariance can be forced to be equal (Table 21).

The only significant change during this interval between eye-opening and puberty involves a change in the unique variance of a single character: the length of the dental diastema. Relaxing the requirement that the unique variance in this one character is invariant between eye opening and puberty generates an acceptable fit of the most restrictive model (X^2 = 85.03, df = 67, p = .068).

The set of skull measures that encompass relatively more information about the jaw and its integration with the cranial measures (Skull 2) undergoes actual repatterning. This repatterning occurs around the time of weaning. The most restrictive model fits adequately for the interval between birth and eye opening, but no model at all, not even the hypothesis that the factor pattern is stable, fits between eye opening and weaning (X^2 = 248.32, df = 32, p < .0001). Furthermore, no restrictions at all can be added to the model when the 21-day-old population is analyzed simultaneously with the 41-day-old and adult populations, although the 41-day-old and adult populations are virtually identical in factor structure. Between weaning and puberty, the specific variance associated with the individual characters also undergoes further change, until the details of integration are stabilized at puberty.

The jaw measures can all be fit simultaneously to a common factor-pattern, the structural unit of integration

Table 21. Comparisons among the parameters of integration observed in skull measures of sequential age-classes. (Skull 1 = inclusive skull set; Skull2 = the expanded cranium and jaw measures); and Jaws.

Age	1--11				11--21			
Measures	X^2	df	p	restricted parameters	X^2	df	p	restricted parameters
Skull 1	38.47	33	.236	Fcova				None
Skull 2	47.38	46	.416	Fcova				None
	62.80	51	.124	U				
Jaws	9.29	13	.751	Fcova	9.28	13	.751	Fcova
	19.93	15	.175	U	13.54	15	.561	U

Age	21--41				41--Adult			
Measures	X^2	df	p	restricted parameters	X^2	df	p	restricted parameters
Skull 1	23.32	33	.894	Fcova	33.23	33	.668	Fcova
	40.22	37	.330	U				
	39.78	40	.480	U+Fcova				
Skull 2	47.38	46	.416	Fcova	40.36	46	.707	Fcova
	62.80	51	.124	U	49.15	51	.547	U
	71.51	54	.056	U+Fcova	58.73	54	.306	U+Fcova
Jaws	3.77	9	.926	Fcova	23.27	17	.141	Fcova
					14.21	15	.510	U

model, (X^2 = 6.98, df = 15, p = .958 for the invariant factor-pattern hypothesis), but they still show dynamic patterns of integration during ontogeny (X^2 = 369.69, df = 47, p < .0001 for the hypothesis that factor-pattern, factor variance-covariance and unique variance are invariant in all samples). Although the interactions between skull and jaw are readjusted at the time of weaning, among themselves the jaw measures retain their original pattern of associations. Between eye opening and weaning the only change involves the unique variance of the measure of the jaw depth at the anterior margin of M_1. When this one uniqueness coefficient is not constrained during this interval, the most restrictive model fits well (X^2 = 14.39, df = 17, p = .649). Between weaning and puberty the pattern matrix does not change; however, no other restrictions can be incorporated into the model. Between puberty and adulthood the uniqueness of the individual characters constitutes an additional invariant parameter.

Discussion

General body size is the dominant, and often the only, cause of integration in these five populations of laboratory rats. It is the one source of integration which affects all regions of the skeleton, at virtually all ages. Only the inclusive skull measures sampled in the 11-day-old rat do not exhibit a significant response to size. And size is the

only cause of integration in the limb until adulthood. More complex developmental models apparently specify causal factors that exert no detectable influence upon observed patterns of integration. Perhaps alternative developmental models could capture more information than contained within a simple size model. But as long as the size factor adequately predicts so much of the observed covariance, alternative models will rarely improve upon it.

The overwhelming influence of size, however, could be an artifact of the analysis. The variability observed in each age class might include variation in both gestational age (Hughes and Tanner, 1970) and growth rate. And the choice of measures may also exaggerate the relative influence of size upon patterns of integration. When so many measures are defined by a small set of measurement axes, principally by the longitudinal and proximo-distal axes, the measures may provide a redundant sample of a limited set of dimensions. Alternative measurement schemes, such as the truss measurement scheme (Bookstein et al., 1985; Strauss and Bookstein, 1982), may contain more information about oblique axes. However, the importance of this size factor may be real, not just an artifact of the measures. Only by fitting both size and other more complex models to several sets of measures, defined by alternative measurement schemes, can the biological importance of size be distinguished from the influence of size imposed upon the data.

Even when general body size does not adequately predict observed covariation among measures, the secondary factors of integration do not evince an unambiguously developmental cause. The theoretically more conservative models of integration, which depend upon no explicitly developmental causation, suffice to predict observed patterns of phenotypic integration. While these models of integration may seem less useful than the developmental models, they impose no causal interpretation where none is clearly required by the data.

The secondary factors of integration, such as the cranium and jaw components at weaning, may reflect transient changes in functional interactions among phenotypic characters. During the first six weeks of life, these rats undergo changes in overall intensity and structure of integration. Unfortunately, these age classes provide only a gross sample of ontogeny. Thus, a correlation between patterns of integration and specific functional stages is merely suggested by these results. However, the timing of these changes within the skull and jaw suggests that functional interactions might create age-specific patterns of integration and underlie the transformation in observed phenotypic integration.

Skeletal morphology is well known to respond to diet (Beecher et al., 1983; Bouvier and Hylander, 1981; Moore, 1965). The transition from suckling to grinding occlusion may be responsible for the decrease in unique variance of

the length of the dental diastema. Within three weeks after weaning, the length of the dental diastema loses its ability to vary independently and becomes increasingly responsive to the common factors influencing the rest of the skull. This change in the degree to which the length of the dental diastema is associated with the other skull and jaw characters may follow from a direct environmental influence upon phenotypic integration.

However, preparation for adult function, not just current or past function, also may be responsible for the observed pattern of integration. The neonatal pattern of integration is repatterned before the effects of grinding should be evident. Thus it is not the actual transition in diet which forces the cranium and jaw to undergo repatterning; nor is the behavioral transition from suckling to weaning a direct response to change in diet. The neuromuscular transition follows an orderly sequence from stereotyped jaw opening and closing movements, through coordinated phases of jaw opening and closing followed by coordinated asymmetrical contractions. The gradual maturation of neuromuscular control is well developed before weaning is complete (Herring, 1985). Prenatal muscle loading is also necessary for normal cartilage development (Atchley et al., 1981).

These data do not permit a more subtle analysis of the epigenetic functional interactions responsible for postnatal skeletal integration. Testing hypotheses about the

influence of functional interactions on developmental integration requires more intensive sampling of ontogeny. Many events, not just preparation for weaning, may influence the repatterning of skull integration between eye opening and weaning. Furthermore, the measures analyzed in this study do not provide an adequate coverage of skeletal structures associated with the musculature of jaw opening, closing and lateral movements. However, the preliminary correspondence between variation in integration and changes in behavior suggest that such functional considerations should be incorporated into hypotheses of developmental constraints. The ontogeny of anatomical structure is, in part, a consequence of the ontogeny of function. Embryonic and infant muscle-firing patterns may be important determinants of integration. Even before masticatory behavior begins, muscle loading affects the anatomical structures; developmental models cannot neglect these embryonic preparations for future functional interactions.

Certainly, patterns of phenotypic skeletal integration may reflect cellular and tissue interactions. Chondrogenic cell density of different mesenchymal populations differs, with consequent differences in the shape of chondrogenic condensations (Kosher, 1983; Zwilling, 1961, 1968). Furthermore, mechanical interactions resulting from common responses to the growth of soft tissues may influence skeletal organization. Cranial morphology, unlike the jaws and teeth, may depend upon brain growth as well as inductive

stimuli from neural tube tissue (Young, 1959). But developmental interactions must be broadly characterized to include neurological and muscular influences upon chondrogenesis. The concept of developmental constraints should be expanded to include not only the cellular mechanisms of morphogenesis but also the intrinsic functional mechanisms which influence form.

The skull is consistently poorly integrated. However, this weak integration is not because the skull exhibits complex patterns of integration. At weaning and in the adult the acquisition of secondary factors does not depress the intensity of integration. Nor does the presence of a secondary limb factor in the adult preclude an increase in integration between puberty and adulthood. While integration and complexity have been perceived as in conflict (Olson and Miller, 1958), increasing complexity does not occur at the expense of overall integration in these rats.

The notion that complexity and integration might be contradictory presumes that integration is a simple feature of a population. Integration, however, is complex. It is not merely a function of the number of integrated suites of characters; it is also a function of the interactions between these suites. Perhaps even more importantly, it is also a function of the influence of the factors of integration upon the individual characters. Overall integration in these rats is lowest when the common factors exert little influence upon the ability of characters to

vary independently. It is relatively high both when size determines the joint behavior of all characters and when size is only one of the effective causes of integration. Secondary factors, rather than decreasing overall integration, may even increase overall integration. These secondary factors may emerge by an increasing response of individual characters to common factors as the individual characters lose their ability to vary uniquely.

The constraints upon integration are stable throughout ontogeny, only temporarily disrupted between eye opening and weaning in the skull. Otherwise, the changes in the patterns and intensity of integration primarily reflect changes in the degree to which the individual characters respond to common factors, and the interactions between the common factors. Yet, because of temporal variation in these aspects of integration, the age at which selection acts might strongly influence the possible evolutionary change in morphology. For example, if selection were to act upon rats as they undergo weaning, the jaw and cranium could be affected separately. In contrast, selection upon the adult could not so precisely modify the jaw without concomitant effects upon the neurocranium. Similarly, if selection were to act upon pubescent rats, the third metatarsal might be forced to respond with all limb characters despite its prior, and subsequent, independence. Even when only a single factor generates all integration, the individual characters can only be modified according to how intensely this factor

constrains their ability to act independently, an aspect of integration which varies throughout postnatal growth. Thus the potential evolutionary consequences of integration may depend upon the schedule of integration and the particular age at which selection acts. The schedule of integration, arising from regular transitions in the effects of cellular, physiological and behavioral causes of integration upon individual characters, may itself represent a developmental constraint on morphological evolution.

Functional causes of integration and functional causes of repatterning may be purely phenotypic responses to extrinsic environmental factors. If so, then these functional factors may be critical to the origin of phenotypic integration, but of little use in predicting the response to selection. The failure of the complex developmental models to predict significant aspects of observed phenotypic covariation among measures does not deny to developmental mechanisms their evolutionary role. Yet developmental constraints guiding evolutionary modification of the phenotype may be equally responsive to embryonic functional interactions. Both preparation for function and actual function might create the patterns of integration represented in genetic covariance.

THE EVOLUTION OF PATTERNS OF INTEGRATION

The constraints evident in patterns of phenotypic integration limit the ability of individual characters to behave randomly. These constraints determine patterns of covariation among characters. Characters might covary because of their common response to biological factors such as growth, and perhaps to functional interactions as well. Yet, developmental and functional covariation need not coordinate the evolution of integrated characters. .To regulate the direction of morphological evolution within a lineage, developmental and functional constraints within the ancestral population must define the potential paths of morphological change.

Historical variation in the structure of integration would undermine the idea that morphological change is constrained by intrinsic factors possessed by all taxa within a lineage. If the patterns of covariation change during phylogenetic and morphological differentiation, then morphological transformation is not controlled by some set of constraints intrinsic to the ancestor. Rather, the constraints would evolve as the morphology evolves.

Historical variation in these constraints can be examined by testing the hypothesis that related taxa share common patterns of integration. In this study, I use Olson

and Miller's (1958) classic historical study of functional integration in five samples of Pentremites, a Mississippian blastoid, to examine the evolution of constraints on complex functional characters.

Comparisons between the average morphology of related species sometimes reveal similar trends in morphological evolution in distantly related taxa. Instances of convergence and parallelism, and the limited number of discrete classes of forms, suggest the operation of constraints (Alberch, 1982; Alberch and Alberch, 1981; Maderson et al., 1982; McGhee, 1982; Raup, 1966, 1967). But the convergence in average morphology between distantly related taxa can neither identify the source of constraints, nor document the importance of these constraints upon evolving morphologies. Too many other variables, such as the interaction between selection and ancestral phenotype, might influence the average morphology. A direct examination of constraints requires that the constraints themselves be extracted, analyzed and compared in related taxa. Only by identifying constraints, and isolating them for comparison, can their historical behavior be directly examined.

The constraints studied here are both developmental and functional. There are two developmental hypotheses: 1) observed integration is a consequence of the common response of all characters to general body growth; and 2) characters derived from a common tissue origin will form discrete units of integration. Functional hypotheses assert the importance

of interactions among characters engaged in a common activity. These functional hypotheses do not suggest that functionally integrated characters need form a complex adaptation. Rather, the functional hypotheses specify that characters covary because they act together during the life of the organism. More complex functional hypotheses claim that several functional components of integration interact to create units of integration comprising several suites of characters engaged in different activities.

The behavior of components of integration during morphological evolution is analyzed by comparing the patterns of integration within and between two related lineages. These comparisons test the hypothesis that morphological evolution occurs within a framework of stable constraints.

Pentremites

Pentremites is a Carboniferous genus of spiraculate blastoids. The external morphology of post-metamorphic blastoids comprises calcitic plates, secreted by mesodermal tissue located between opposing plates (Macurda 1967). Plates are regularly arranged upward, from the basal plates nearest the stem to the apical deltoids (Beaver, 1967). The deltoid plates are bounded by ambulacra, through which water and food particles move. The mouth, surrounded by deltoid plates, and the anus, an opening within a deltoid plate, are located at the summit of the calyx. Spiracles are also

located at the upper extremity of each deltoid plate. In Pentremites, the anus is joined with one spiracle, just above the posteriod deltoid, forming the large anispiracle. Openings into the digestive or respiratory system are evident either as gaps between plates or as excavations within plates.

I selected a subset of measures published by Olson and Miller (1958) (represented schematically on the dissected specimen in Figure 7, described in Table 22). Because of the small sample sizes in several populations, I could not analyze all their measures. I removed colinear measures because the morphometric analysis cannot estimate parameters unless the covariance matrix is postive-definite. I also excluded from analysis those measures intended to represent general body "size" and "shape" and measures encompassing several functional units. All measures selected belong to basal, deltoid, digestive or respiratory complexes.

Little is known about the biology of Pentremites. There are no close modern relatives, nor are there any modern analogs to blastoids. Crinoids and echinoids provide the basis for the interpretation of function in Pentremites. Miller and Chave designated the respiratory and digestive complexes as "functional groups", because they might be expected to respond adaptively to environmental changes (these inferences of biological function are discussed in Olson and Miller, 1958). The deltoid characters were

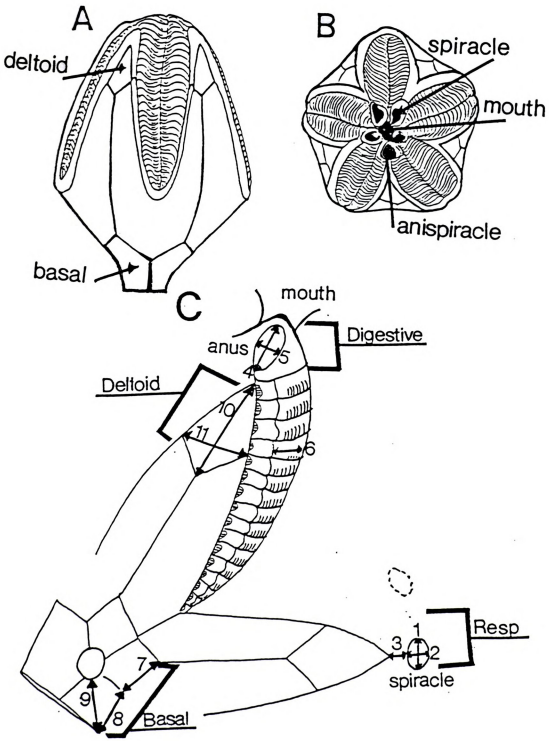


Figure 7.-- Schematic representation of the external morphology of the blastoid and representation of the measures analyzed. (A is a lateral view, B is an view of the summit of the calyx, C is a dissected specimen. The measures are drawn on the dissected specimen). Abbreviations and descriptions of the measures are presented in Table 22.

Table 22.-- Description of characters represented in Figure 7. The numbers of the measures on this list correspond to the numbered measures on the diagram.

Respiratory

1 Outer edge of spiracle to margin of oral opening

2 Width of spiracle

3 Exposed tip of deltoid plate to oral opening along line bisecting spiracle

Digestive

4 Outer edge of anal opening to margin of oral opening

5 Width of anal opening

6 Length of food groove

Basal

7 Length of radial-basal suture adjacent to azygous plate

8 Length of radial-azygous plate suture

9 Length of azygous plate from center of base to distal apex of azygous

Deltoid

10 Length of exposed portion of deltoid

11 Width of deltoid from tips of lateral margins of paired radials

also considered as functional characters, but more because of the close topographic association between deltoid, digestive and respiratory characters than because of any specific function attributed to the deltoid plates. The deltoid and basal characters might be better interpreted as structural characters, because they are features of the external skeleton and serve to support and protect internal organs. The basal plates contribute to the architecture of the calyx associated with attachment of the calyx to the stem; the deltoid plates form the lateral, apical plates of the calcified skeleton. I have termed these structural characters as components of "functional groups" because the basal and deltoid characters are engaged in a common activity. I do not use the term to imply that "functional groups" represent adaptations.

The genus Pentremites comprises two morphological types: elongate pyriform and bullet-shaped godoniform morphs. Pyriform and godoniform taxa are members of separate monophyletic groups (Fig. 8) that share a common ancestor (Waters et al. 1986). The three godoniform species in this study (P. godoni, P. tulipaformis and P. robustus) are closely related. Both Olson and Miller (1958) and Waters et al. (1986) interpret P. godoni as ancestral to P. tulipiformis, and P. tulipiformis as the ancestor of P. robustus. These three godoniform taxa do not comprise a strictly monophyletic group, as P. gutschicki, the sister-group of P. robustus, is excluded. However, even without

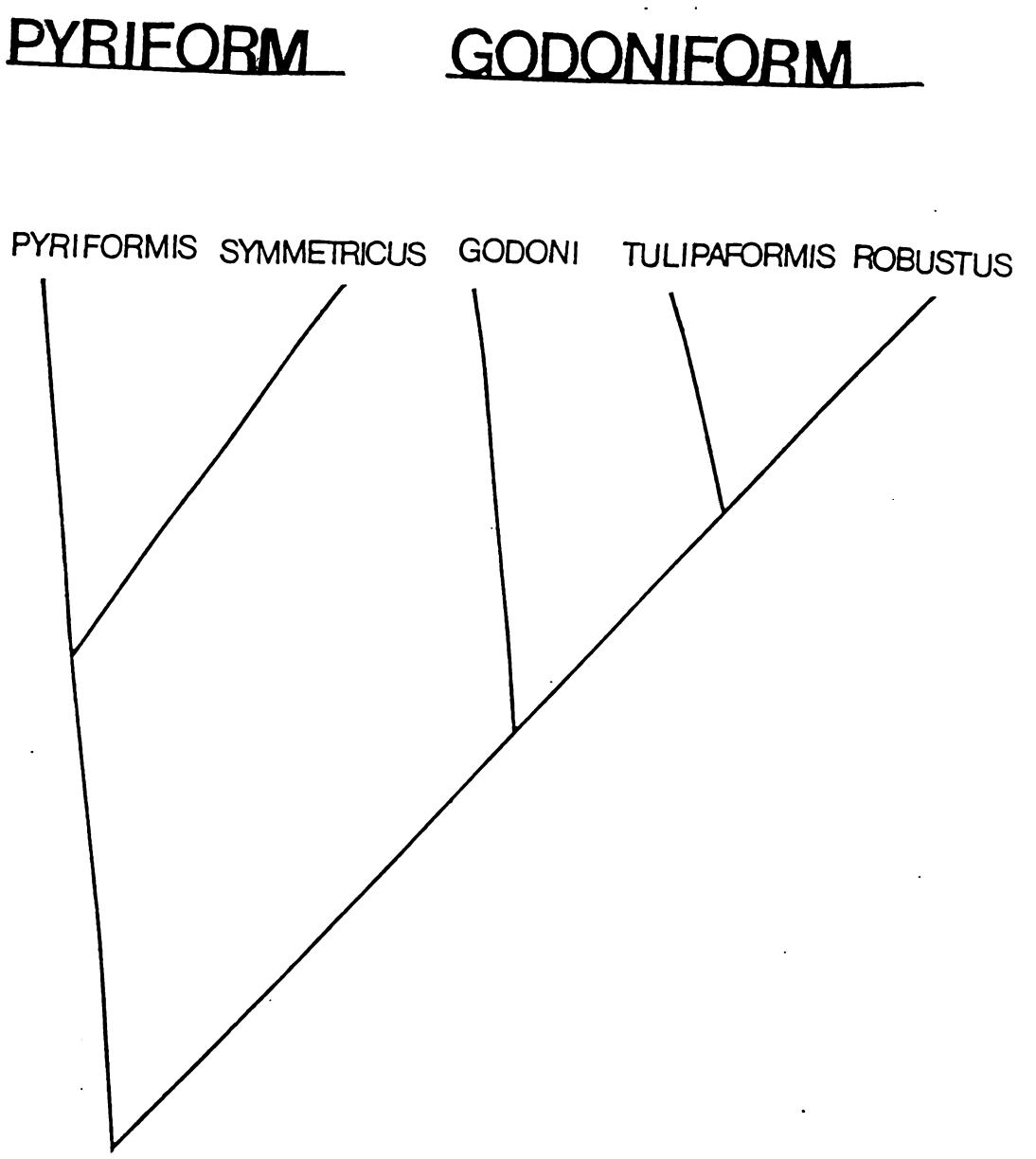


Figure 8.-- Hypothesized relationships among the five species of Pentremites (after Waters et al., 1986). According to Olson and Miller (1958) P. pyriformis is ancestral to P. symmetricus; Waters et al. consider P. symmetricus to be the ancestor of P. pyriformis).

speculating about ancestral-descendant relationships, the three godoniform taxa are closely related.

Relationships between the two pyriform taxa (P. symmetricus and P. pyriformis) are more difficult to interpret. While Olson and Miller interpreted P. pyriformis as ancestral to P. symmetricus, Waters et al. regard P. symmetricus as ancestral to P. pyriformis. The cladogram does not represent any hypothesized ancestral-descendant relationships. However, the historical relationship between these two taxa is critical to any interpretation of transformations of constraints. Therefore, I have interpreted the comparisons between these two taxa in accord with both phylogenetic hypotheses.

Current systematic analysis allocates the two species of Olson and Miller to five different species (Waters et al. 1986). Since the measurement procedures of Olson and Miller destroyed the specimens, their identifications cannot be checked. I have used stratigraphic position and locality to identify species analyzed in this study (summarized in Table 23). More than one morphologically similar population could be represented in each sample. However, both multivariate analysis of variance (BMDP4V; Dixon, 1985) and stepwise discriminant analysis (BMDP7M; Dixon, 1985) clearly discriminate the several godoniform taxa and successfully distinguish the two pyriform taxa from each other. Misclassifications reflect the similarity between the populations in the distribution of the characters

Table 23.-- Formation and species and of the godoniform and pyriform samples analyzed by Olson and Miller. The table is arranged chronologically, with the ealiest formation listed last. The number of specimens measured of each species is enclosed in parentheses.

Formation	Godoniform	Pyriform
Glen Dean	P. robustus (28)	
Golconda	P. tulipaformis (18)	P. symmetricus (17)
Paint Creek	P. godoni (40)	P. pyriformis (31)

contributing to the discriminant function. Only 14% of the individuals were misclassified (Table 24), all involving a failure to discriminate between the externally dissimilar Paint Creek godoniform and pyriform species. Six specimens of P. godoni were classified as P. pyriformis, while nine specimens of P. pyriformis were classified as P. godoni.

Analysis of morphological differentiation.-- The differentiation of average morphology in the evolution of Pentremites was analyzed by multivariate analysis of variance and by stepwise discriminant analysis. Both of these procedures analyze differences between the population means. This analysis of morphological differentiation investigates the degree to which, and the ways in which, the average morphology of the populations change. Multivariate analysis of variance tests the hypothesis that populations do not differ in their means when several characters are analyzed simultaneously. The null hypothesis is that the different samples, in this case the different species, are drawn from a common population. Discriminant analysis identifies the characters which contribute most to the differences between samples, even when the samples do not differ significantly. The result of the discriminant analysis is one or more linear combinations of characters which best discriminate between the samples.

Discriminant analysis was used to examine the patterns of differentiation in Pentremites. Multivariate analysis of variance was used to test the significance of the

Table 24.-- Classification of *Pentremites* species by discriminant function analysis. (Abbreviations: Pyr = *P. pyriformis*; Sym = *P. symmetricus*; God = *P. godoni*; Tulip = *P. tulipiformis*; Rob = *P. robustus*).

Taxon		%Correct	Number of Cases classified by the discriminant function				
			PYRIFORM		GODONIFORM		
			Pyr	Sym	God	Tulip	Rob
PYRIFORM	Pyr	80.6	25	0	6	0	0
	Sym	100.0	0	17	0	0	0
GODONIFORM	God	76.9	9	0	30	0	0
	Tulip	100.0	0	0	0	18	0
	Rob	100.0	0	0	0	0	28

differences among the average morphologies.

Morphological differentiation in Pentremites.-- While godoniform and pyriform taxa can be distinguished by overall shape of the calyx, the functional characters examined in this analysis do not show such clear differentiation between these two groups, as evident in the failure of the discriminant analysis to distinguish between pyriform and godoniform species at a common stratigraphic level. The inability of the discriminant analysis to separate godoniform and pryiform taxa is probably a result of the measures selected for this analysis. Much of the discrimination between species is due to differences in the average length of the radial-basal suture adjacent to the azygous plate (Table 25). Basal characters, and the single respiratory character, are responsible for most of the discrimination among species. However, some of the differences are explained by differences in the average width of the anal opening, length of the food groove and the remaining basal character, the length of the azygous plate. Evidently, characters drawn from all four functional units contribute to the differences among these five species.

However, this discriminant function analysis identifies a canonical variable that simultaneously distinguishes among all five groups. It is probably strongly influenced by P. robustus, which differs so strongly from the other taxa. More subtle differences between taxa, and the particular differences which distinguish between sister-taxa, are not

Table 25.-- Coefficients for the characters associated with the first canonical variable. This variable accounts for 99.86% of the dispersion among species of Pentremites.

Character	Coefficient
Outer edge of spiracle to margin of oral opening	2.162
Width of anal opening	1.429
Length of the food groove	- .324
Length of radial-azygous plate suture	1.594
Length of azygous plate from center of base to distal apex of azygous	-6.783
Length of exposed portion of deltoid	.216
Width of deltoid from tips of lateral margins of paired radials	.109

likely to define the first discriminant function. Some of these more subtle differences are evident in the canonical variables which account for little of the discrimination. The third canonical variable distinguishes between the Paint Creek pyriform and godoniform taxa, and between the Golconda pyriform and godoniform taxa, but accounts for only 0.19% of the total dispersion. Excluding the Glen Dean P. robustus from the discriminant function analysis would magnify the differences between the remaining taxa. However, the purpose of this analysis is not to identify those differences acquired by the derived taxa, but rather to examine the patterns of morphological differentiation in both lineages.

Species of Pentremites at the same stratigraphic level are similar to each other in the characters associated with the first two canonical variables (Figure 9). The Paint Creek species resemble each other more closely than they do their closer relatives. The same pattern is repeated in the Golconda species. P. tulipaformis and P. symmetricus are similar to each other and distinguished from the Paint Creek species.

Apparently, godoniform and pyriform lineages replicate the same average morphology. The primitive godoniform (P. godoni), is almost indistinguishable from P. pyriformis. The relatively derived P. robustus is only slightly different from the primitive pyriform P. symmetricus. The evidence for similar trends in the transformations in each of these two lineages is not altered if P. symmetricus is interpreted as

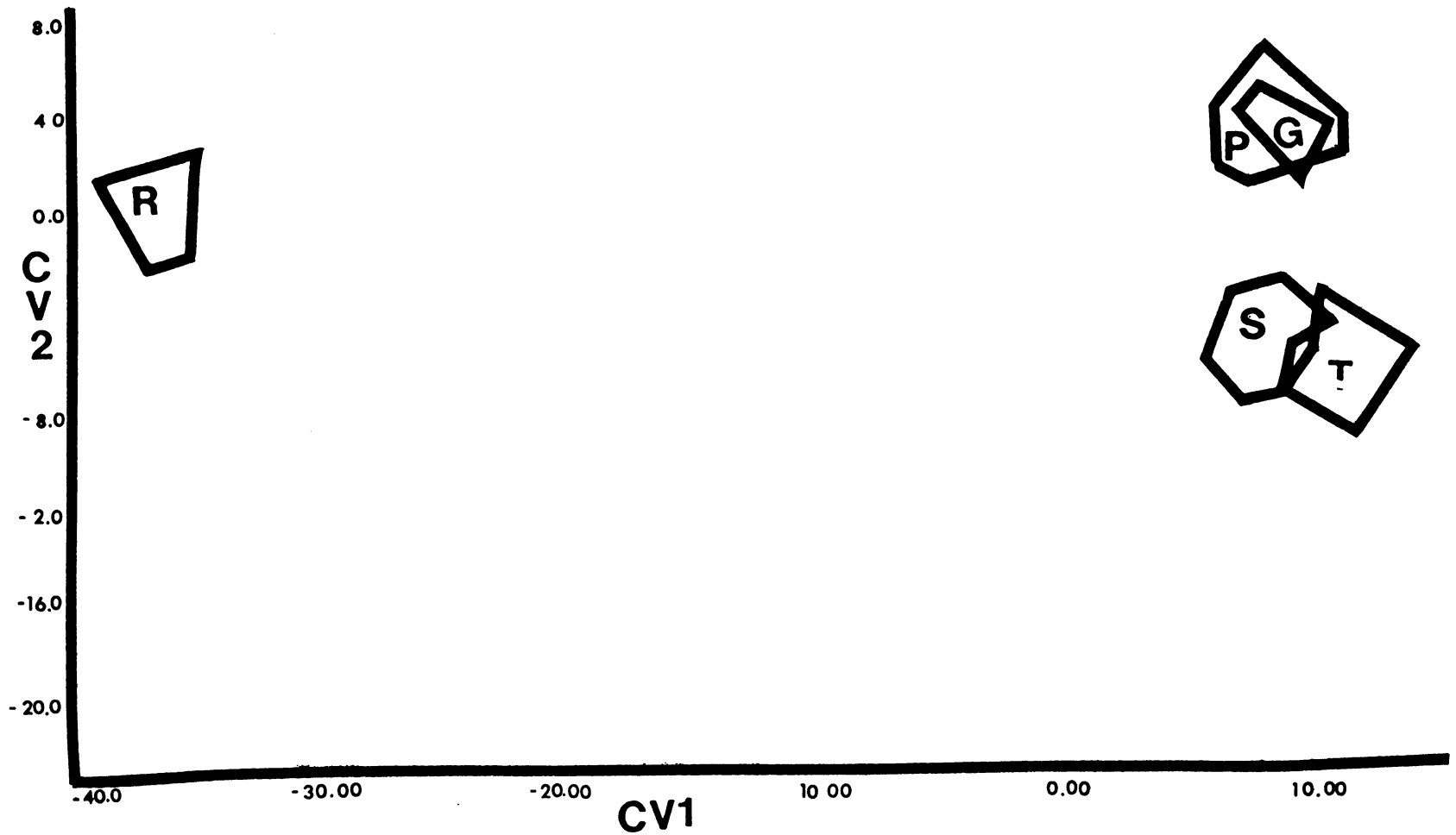


Figure 9.-- Plot of the five Pentremites species on the first two canonical variables estimated by discriminant functions. (Abbreviations: S= P. symmetricus; P= P. pyriformis; G= P. godoni; T= P. tulipaformis; R= P. robustus.

derived rather than primitive. If P. pyriformis is the primitive pyriform then the two lineages both follow the same trend, in the same direction, converging upon a common morphology. The separation between three discrete clusters representing 1) Paint Creek species, 2) Golconda species and 3) P. robustus, coupled with common transformations in morphology, suggests that functional morphology of Pentremites might be constrained. The gap between the discrete clusters cannot be interpreted as evidence of inaccessible morphologies because the discriminant function analysis separates taxa as far as possible, thus these gaps will result from a discriminant function analysis even when average morphologies are fairly similar. Yet, it might appear that there are restrictions upon the number of possible variants upon the Pentremites functional plan, and these variants recur in distantly related taxa.

Analysis of differences in average morphology does not demonstrate that the limited number of morphological variants in Pentremites is a result of intrinsic constraints. Nor does it identify particular constraints upon phenotypic evolution. It does not even argue for constraints upon the potential evolutionary transformations. Patterns of covariation, which are determined by the intrinsic constraints upon phenotype, might evolve. The discriminant analysis concentrates upon divergence in the average morphology, in the means of the characters in the related species, rather than upon the constraints. The

similarity between pyriform and godoniform taxa at a given stratigraphic level might even result from selection upon the average morphology, rather than from any intrinsic constraints. Similarities and differences in average morphology may yield a misleading picture of the patterns of integration in each population and of the historical transformations in integration. The evolution of morphology in these Pentremites may appear constrained, but this analysis merely suggests the hypothesis. Direct analysis of the constraints upon phenotype in Pentremites, and examination of the stability of the phenotypic constraints throughout phylogeny, can test the hypothesis that the evolution of functional morphology in Pentremites occurs within a stable framework of constraints.

Methods

Causal analysis of integration.-- The developmental and functional constraints upon the phenotype of Pentremites were analyzed by fitting competing developmental and functional models to observed variance-covariance matrices in each species (see chapters 2 and 3 for details of the procedure). Comparisons between nested models used the X^2 difference test. These models were fit to each species independently and evaluated, by confirmatory factor analysis, for their ability to reconstruct the observed variance-covariance.

Comparative analysis of integration.-- Comparisons between

the species employed comparative confirmatory factor analysis (details of the procedure for comparative confirmatory factor analysis are provided in Chapter 3). Three hypotheses were tested for each comparison: 1) that the species share a common factor-pattern (F); 2) that the covariance between factors and the variance explained by each factor is invariant across populations (Fcova); or 3) that the degree to which the individual characters vary uniquely is invariant across all species sharing a common factor-pattern (U) .

Analysis of developmental and functional integration

This analysis concentrates upon the functional basis of integration. In addition, two specifically developmental models are evaluated: growth and differentiation from a common tissue origin are hypothesized to be the developmental constraints upon phenotypic covariance. The growth model uses the regression weights of each character on the first principal component (see Chapters 2 and 3) to specify the associations between the individual characters and the growth factor. This hypothesis asserts that a shared response to general body growth determines all observed covariance.

The tissue-origin model predicts that, in addition to a growth factor, mesodermal derivatives (basal and deltoid characters) comprise one unit of integration while the endodermal derivatives (respiratory and digestive

characters) comprise another integrated unit. Other models, also consistent with the hypothesis that origin in a common tissue determines covariation, could be constructed. Little is known about the developmental biology of Pentremites. It is possible that all measures, both of structural plates and of functional characters, respond to the tissue interactions required for synthesis and construction of skeletal plates. This hypothesis asserts that the digestive and respiratory characters covary because the internal structures are derived from endoderm, while the basal and deltoid plates are secreted by mesodermal structures.

Each of the functional models incorporates the general growth factor. In addition, the functional models claim that secondary factors of integration are determined by functional interactions. Inferences about the adaptive nature of the functional groups are not required before defining expected patterns of covariation a priori. In this analysis, I construct functional groups based upon the hypothesis that characters which act as a unit, whether functional or structural, will covary. According to the simplest functional hypotheses, all covariance not explained by growth is due to interactions among characters engaged in a common activity. For example, the basal model predicts that the covariance among measures is due to the covariance of all basal measures with each other, and that all other covariance is a function of general growth. When the functional models improve significantly in fit over the

growth model, the interactions among functionally related characters are a cause of integration. According to these hypotheses, each discrete functional unit constitutes a unit character.

After evaluating the fit of the growth and simple functional models, more complex models were constructed by combining functional units to ask if interactions among functional units constitute a source of integration. For example, the basal-deltoid model forces the basal and deltoid characters to covary as a single factor, a unified component of integration. This hypothesis predicts that characters engaged in different activities create coherent networks of interactions which could evolve as units. When the complex functional models significantly improve upon nested simple functional models, as judged by the X^2 difference test (Chapter 2), then interactions between functional units contribute to observed integration.

Evaluation of causal developmental models.-- Growth is an important, but not the sole explanation of observed covariance (Table 26). In no species does the growth model alone adequately reconstruct observed variance-covariance. However, in P. tulipaformis, the patterns of covariance due to growth almost adequately account for integration ($p = .036$). While not an exhaustive explanation for observed phenotypic integration, a general response to overall body growth in all five samples accounts for a substantial proportion of observed phenotypic variance-covariance. As

Table 26.-- Goodness-of-fit values for the hypothesis that size alone accounts for observed integration in each sample of Pentremites. The null hypothesis is that the growth model reconstructs observed variance-covariance. The % improvement of the growth model over the model of no integration estimates the proportion of total integration accounted for by size.

Taxon	X^2	df	p	% improvement
P. symmetricus	99.44	54	0.000	63
P. pyriformis	121.85	54	0.000	64
P. godoni	148.78	54	0.000	76
P. tulipaformis	74.13	54	0.036	51
P. robustus	140.74	54	0.000	55

measured by the percent improvement of the growth model relative to the model of no integration, the proportion of integration due to growth ranges from 51% in P. tulipaformis, to a high of 76% in P. godoni. Thus general source of covariance, due to the tendency of all characters to increase in size with increasing body size, represents a primary constraint upon phenotypic integration.

The tissue-origin model fits P. godoni and P. tulipaformis well (Table 27). However, the poor ability of this model to reconstruct observed integration in P. symmetricus, and its equivocal fit in both the P. pyriformis and P. robustus populations, casts doubt upon the hypothesis that origin in a common embryological tissue determines covariance among characters. The tissue-origin hypothesis specifies that the developmental basis of covariation is common to all these taxa because the tissue origin of these characters is a universal feature of this lineage. It is unlikely that these five species differ in which embryological tissues gives rise to these characters. Thus the variation in fit of this model is most likely a consequence of the failure of the hypothesis, rather than a result of variation in embryological tissue origin. However, it is possible that origin in a common embryonic tissue is a developmental basis of covariation in these taxa, which fails to fit P. pyriformis and P. robustus because other sources of covariation also influence the observed patterns of covariance.

Table 27.-- Goodness-of-fit for the tissue-origin model fit to all five species of Pentremites. The growth factor is incorporated into the tissue-origin model.

Taxon	X^2	df	p
P. symmetricus	555.96	43	0.000
P. pyriformis	53.90	40	0.070
P. godoni	46.87	40	0.211
P. tulipaformis	45.48	43	0.369
P. robustus	56.72	40	0.042

Evaluation of functional models.-- The simple functional models, which all include the growth factor, account for some observed integration (Table 28). Yet, at least some of the integration explained by these simple functional models may merely be due to the incorporation of the growth factor within each model. As judged by the X^2 difference test (Table 29), the basal characters do contribute significantly, over and above growth, to integration in P. pyriformis, P. godoni, and P. robustus. Similarly, the deltoid characters also contribute significantly to integration in P. godoni, P. robustus, and both pyriform samples. The respiratory characters contribute to integration in both P. tulipaformis and P. pyriformis, but are clearly inconsistent with the observed integration in P. robustus and P. symmetricus. The digestive characters do not form a unit of integration in any of the five samples.

Only in P. tulipaformis is all covariance among phenotypic measures determined by general growth and interactions among characters within a single functional unit (the respiratory). In P. tulipaformis the respiratory unit contributes significantly more than growth to observed integration, and the simple respiratory model adequately reconstructs observed covariance ($p = .275$). In other cases, the X^2 difference tests reveals that functional units do contribute to integration but do not adequately account for observed integration. The functional units, therefore, do contribute to observed integration, but

Table 28.-- Probability levels for the simple functional hypotheses fit to each sample of Pentremites. Each simple functional model predicts that all observed covariance is due to general growth and interactions among characters within the specified functional unit.

	Basal	Deltoid	Respiratory	Digestive
P. symmetricus	0.001	0.012	0.000	0.000
P. pyriformis	0.026	0.000	0.000	0.000
P. godoni	0.005	0.000	0.000	0.000
P. tulipaformis	0.041	0.037	0.275	0.008
P. robustus	0.000	0.000	0.000	0.000

Table 29.-- X^2 difference tests for the significance of differences between the growth model and the simple functional models. The null hypothesis for the test is that the two models fit equally well. When the difference in fit is not significant, the functional model explains no integration not already explained by the growth model. There are no improvements significant at the .01 level, those significant at the .005 level are indicated by **.

	Basal		Deltoid		Resp		Dig	
	ΔX^2	Δdf	ΔX^2	Δdf	ΔX^2	Δdf	ΔX^2	Δdf
P. symmetricus	12.28	4	23.12	3**	-456.52	4	-456.524	4
P. pyriformis	50.58	4**	29.25	3**	20.69	4**	.56	4
P. godoni	69.65	4**	19.43	3**	6.21	4	5.14	4
P. tulipaformis	5.53	4	3.63	3	18.62	4**	-4.48	4
P. robustus	21.38	4**	46.51	3**	-1008.82	4	6.36	4

at least some of the observed covariance probably depends upon interactions among functional units.

In P. pyriformis, the model combining both digestive and respiratory characters fits significantly better than either simple digestive or respiratory models ($\Delta X^2 = 19.93$, $\Delta df = 3$; $p < .005$ for the comparison between digestive + respiratory and simple respiratory models; $\Delta X^2 = 40.06$, $\Delta df = 3$; $p < .005$ for the comparison between digestive + respiratory and simple digestive models). Thus digestive and respiratory characters interact to comprise a weakly integrated unit in P. pyriformis. But this interaction between digestive and respiratory characters still contributes relatively little to observed integration, as judged by the poor fit of this model ($p = .000$). A similar pattern of improvement above the nested simple functional models, coupled with poor fit of the digestive + respiratory model, is repeated in all other taxa except for P. tulipiformis and P. robustus (Table 30).

No other pair of functional units contributes significantly to observed integration in more than two species. The deltoid and digestive complex accounts for significantly more integration than either of the nested deltoid or digestive units in P. symmetricus and P. robustus, but does not improve upon the nested functional components in the other three species. In no species does the combination of basal and deltoid characters account for more integration than the simple basal or deltoid models.

Table 30.-- Probability levels for the models which specify that growth and interactions between pairs of functional units determine observed integration. These models were compared to the simple functional models by X^2 difference tests; significant improvement (.01 level) over the simple nested functional models are indicated by *.

	Sym	Pyr	God	Tulip	Rob
Bas + delt	0.035	0.017	0.006	0.039	0.000
Bas + resp	0.016*	0.015	0.000	0.276	0.000
Bas + digs	0.001	0.000	0.003	0.028	0.000*
Delt + resp	0.000	0.000	0.000	0.072	0.000
Delt + digs	0.045*	0.001	0.000	0.000	0.001*
Resp + digs	0.001*	0.000*	0.000*	0.259	0.000

All other models which comprise a pair of functional units, improve in fit upon both nested functional models in at least one species. Yet, in no case does this improvement result in a model which successfully reconstructs observed variance-covariance. All complex models which fit better than both simpler nested models still fit poorly.

However, achieving successful fit does not require, in all cases, the addition of multiple functional units to a single factor. In P. pyriformis a model which specifies that basal and deltoid characters constitute two factors accounts for observed integration (X^2 = 43.70, df = 46; p = .569). In this sample, the basal and deltoid characters represent two components of integration. This hypothesis that the deltoid and basal characters each constitute a factor of integration, when combined with the growth factor, is the simplest model which can successfully reconstruct observed variance-covariance in this species. In P. symmetricus however, this model fails to fit (X^2 = 555.96, df = 45; p = .000). Specifying that respiratory characters are associated with the basal characters, forming a single unit of integration, improves the fit (X^2 = 56.34, df = 43, p = .085 for this model). Further modifying the model, incorporating the digestive and deltoid characters into a single unit of integration, produces an acceptable fit (p = .234). This model, which interprets the covariance among characters as a result of a single general growth factor and two secondary factors (basal + respiratory and digestive + deltoid) is the

simplest model which reconstructs observed integration well in P. symmetricus..

The Golconda godoniform P. tulipaformis is adequately fit by the simple hypothesis that general growth and interactions among the respiratory characters determine integration. No other components of integration contribute significantly to the pattern of integration in this sample. Other functional units can be added to the respiratory unit without a loss of fit. For example, the model which best fit the pyriform P. symmetricus also fits the godoniform P. tulipaformis well ($p = .560$). While it does not fit better than the simpler respiratory model, it does not conflict with the pattern of covariance. This same model, which associates basal and respiratory characters on a single factor, and deltoid and digestive characters on a second factor, also fits P. godoni well ($X^2 = 46.87$, $df = 40$; $p = .211$). Only such a model, comprising all four functional units, as well as the general growth factor, fits the pattern of integration in P. godoni. In contrast, the hypothesis that the basal and deltoid characters each constitute a component of integration fits P. godoni relatively poorly ($p = .062$). Adding the respiratory characters to the basal component does not significantly improve the fit ($\Delta X^2 = 4.18$, $\Delta df = 3$; $p < .250$). In this primitive godoniform species, all functional characters contribute to observed integration.

P. robustus is unique because no model at all can

reconstruct observed integration well. Even the model which fits best in this species (associating the deltoid and basal characters on one factor, and digestive and respiratory characters on the other) fits relatively poorly ($p = .042$). The model comprising the deltoid and digestive factor and the basal and respiratory factor, which fits well in all other populations, clearly conflicts with the patterns of integration in P. robustus ($X^2 = 1149.56$, $df = 40$; $p = .000$). I could not construct any a priori model to fit this sample acceptably; instead, LISREL modified the specifications to produce an acceptable fit for a model which associates the width of the anal opening with the basal unit, and with the remaining digestive, respiratory and deltoid characters.

Growth and interactions among characters engaged in a common function evidently do constitute causes of observed phenotypic integration in these samples of Pentremites, except in P. robustus. Figure 10 summarizes the pattern of integration in the pyriform species, Figure 11 summarizes these patterns in the godoniform species. Evidently, the functional interactions are not restricted to characters within a single functional unit. Only in P. tulipaformis does general growth and interactions within a single functional unit explain all observed covariation among characters. In both pyriform species, and in the godoniform P. godoni, interactions between the basal and respiratory characters or between deltoid and digestive characters

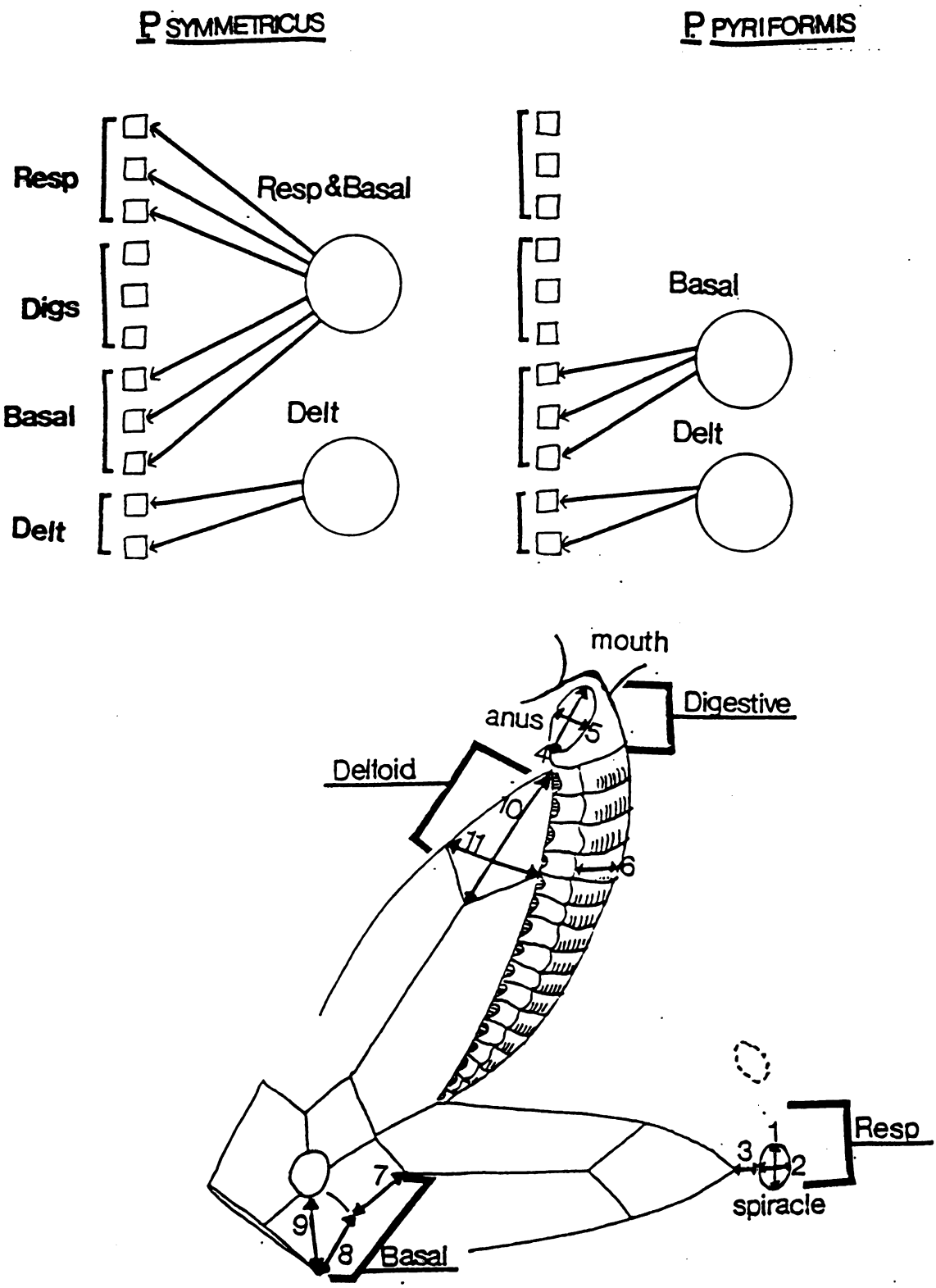


Figure 10.-- Path diagrams for the simplest adequate models fit to the pyriform Pentremites.

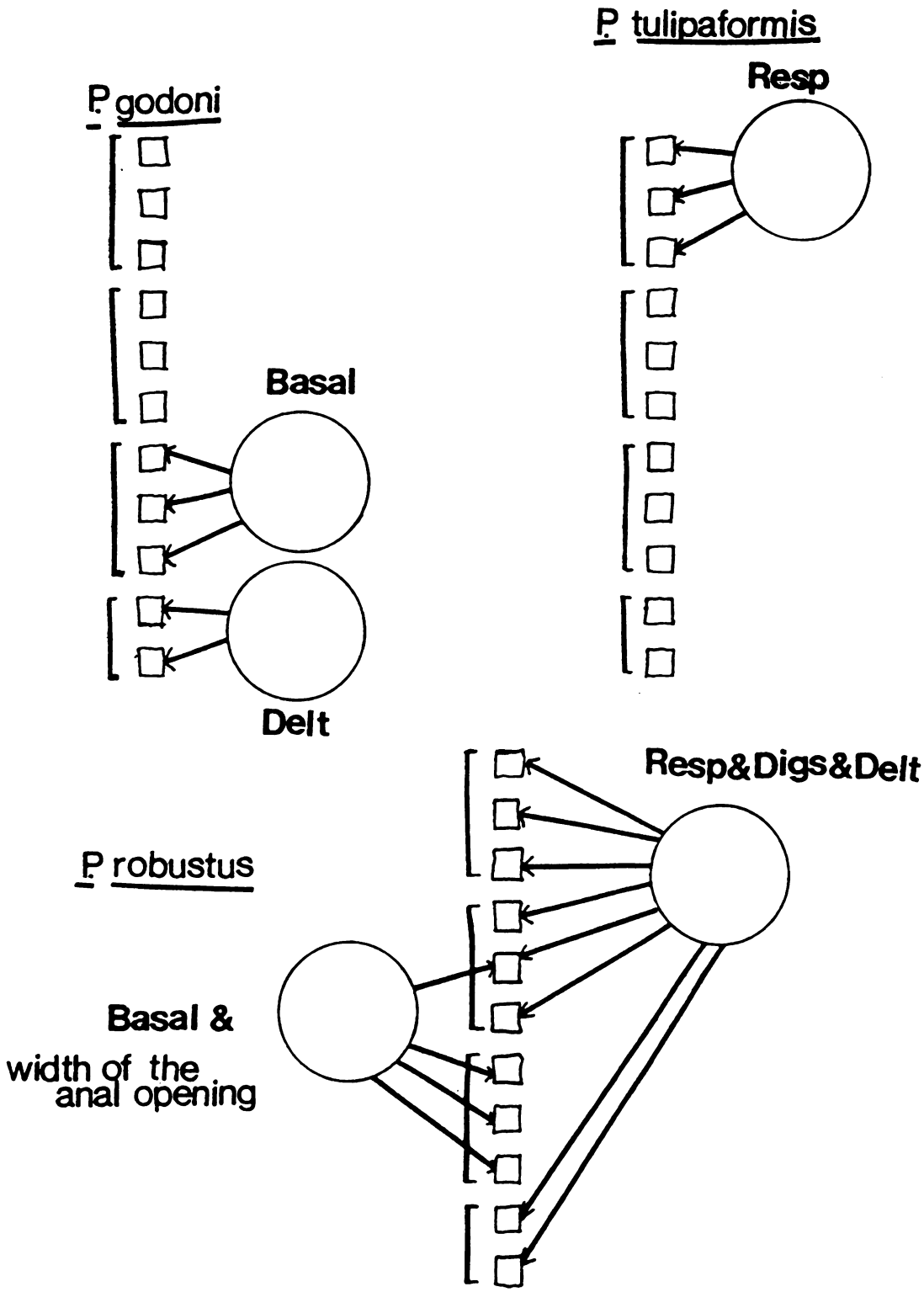


Figure 11.-- Path diagrams for the simplest adequate models fit to the godoniform Pentremites.

account for integration.

Developmental factors other than growth might further control covariation among characters, but they could not be discerned in this analysis. One developmental factor, however, that does not appear to constrain patterns of variation and covariation is the tissue origin of the characters. The fit of the model which associates mesodermal and endodermal derivatives into two discrete units varies across taxa, although the characters, presumably, do not differ in embryological origin. Thus it is unlikely that tissue origin determines covariation. Yet, perhaps the tissue origin of the endodermal characters matters less than the interaction between endodermal derivatives and exoskeletal characters. The spiracular, oral and anal characters are located within gaps and excavations into the plates; thus the position and proportions of the endodermal characters may be determined by growth of the external plates. More sophisticated developmental models might reveal developmental constraints upon integration, but such models would require measurements designed to capture more developmental information and might also require more sophisticated developmental hypotheses.

The evolution of functional and developmental constraints

The study of integration in any set of species uncovers the intrinsic constraints which might make particular

evolutionary changes likely. Yet, because the constraints themselves may evolve, constraints upon phenotype need not act as constraints upon phenotypic evolution. While the constraints upon phenotype may determine observed morphology, the evolutionary transformation in morphology could come about through changes in the constraints. Evolutionary transformations in the constraints can be directly analyzed by comparing the ability of particular models to simultaneously fit related taxa.

Construction of comparative models.-- The hypothesis that the factor-pattern is invariant claims that the same latent variables determine covariation among characters in all five species. It asserts that the constraints are preserved throughout morphological and phylogenetic evolution. Thus a comparison between the factor-pattern of phenotypic integration directly tests the hypothesis that the constraints are stable throughout the history of these Pentremites. If this hypothesis fails, then it is probable that the constraints themselves have evolved. A change in factor-pattern implies that different causes underlie phenotypic covariation within the different species.

Even when the framework of constraints is invariant, as judged by the common factor-pattern, the constraints may behave differently in different species. For example, the populations may differ in the degree to which a particular constraint influences integration. Differences in the variance of the factor of integration or in the covariance

evolutionary changes likely. Yet, because the constraints themselves may evolve, constraints upon phenotype need not act as constraints upon phenotypic evolution. While the constraints upon phenotype may determine observed morphology, the evolutionary transformation in morphology could come about through changes in the constraints. Evolutionary transformations in the constraints can be directly analyzed by comparing the ability of particular models to simultaneously fit related taxa.

Construction of comparative models.-- The hypothesis that the factor-pattern is invariant claims that the same latent variables determine covariation among characters in all five species. It asserts that the constraints are preserved throughout morphological and phylogenetic evolution. Thus a comparison between the factor-pattern of phenotypic integration directly tests the hypothesis that the constraints are stable throughout the history of these Pentremites. If this hypothesis fails, then it is probable that the constraints themselves have evolved. A change in factor-pattern implies that different causes underlie phenotypic covariation within the different species.

Even when the framework of constraints is invariant, as judged by the common factor-pattern, the constraints may behave differently in different species. For example, the populations may differ in the degree to which a particular constraint influences integration. Differences in the variance of the factor of integration or in the covariance

between factors reflects a change in influence of the constraint. Thus the stable set of constraints may still evolve in their influence over the stable integrated units.

Furthermore, individual characters may differ, in the different populations, in how effectively they are restricted by the constraints. A decrease in the degree to which a specific character can vary independently implies that its variation is increasingly constrained by that factor of integration. Changes in the degree to which individual characters are regulated by the constraints can be identified by differences in the uniqueness of the individual characters.

All comparisons were performed by fitting populations simultaneously to the model which best reconstructed variance-covariance structure in most samples. The simple respiratory model, which effectively predicted the structure of integration in P. tulipaformis alone was not used for this comparison. Nor was the model which predicted that basal and deltoid characters each determine a factor of integration employed for comparisons. Neither of these two models fit more than one population. In constrast, the model which specifies that general growth constitutes a general factor of integration; basal and respiratory characters together comprise one secondary factor of integration and the complex of digestive and deltoid characters another secondary factor fit all populations except for P. robustus. While this model is more complex

than required by P. tulipaformis and P. pyriformis, it fits these two populations acceptably. This is the only model which fits all populations other than P. robustus. To compare P. robustus to the other species, I used the model, generated by LISREL, which adequately reconstructed observed variance-covariance matrix in this one species as well as the model which fitted all other species.

Stability of constraints.--The two pyriform species show no divergence in factor-pattern when simultaneously fit to a common model (Table 31). Furthermore, they show no divergence in factor-covariance nor in the proportion of variance unique to each character ($p = .198$; and $p = .175$ for the two hypotheses respectively). Thus I cannot reject the hypothesis that the intrinsic constraints upon phenotype are common to both pyriform species, evident in the invariant factor-pattern. And the effects of these constraints upon particular characters do not differ in these two species. In both species, the constraints influence the patterns of integration to the same degree, and are similar in the degree to which the factors of integration covary. Even the degree to which the constraints regulate the independent variation of the individual characters is the same.

The godoniform lineage, however, shows a greater divergence between the Paint Creek and Golconda species (Table 32). Although there is no change in factor pattern, P. godoni and P. tulipaformis differ in factor-covariance

Table 31.-- Comparison of parameters of integration between pyriform Pentremites. Reported are the goodness-of-fit values for the hypotheses that factor-pattern (F), factor-covariance (Fcova) or proportion of variance unique to each character (U) are invariant. The null hypothesis is that the populations do not differ.

	X^2	df	p
F	91.21	84	.277
Fcova	96.89	86	.198
U	103.47	91	.175

Table 32.-- Comparisons of parameters of integration between P. godoni and P. tulipaformis. Reported are the goodness-of-fit values for the hypotheses that factor-pattern (F), or factor-covariance (Fcova) or proportion of variance unique to each character (U) are invariant.

	X^2	df	p
F	85.65	80	.312
Fcova	112.47	86	.029
U	165.19	97	.000

and in the proportion of unique variance in each character. Apparently the same constraints may determine integration in these two species, but these constraints differ between these species in their influence. In particular, the respiratory characters covary more intensely in P. tulipaformis, while the basal, deltoid and digestive characters account for little observed integration. However, the patterns of integration, over and above the covariation among respiratory characters, exhibit the same structure in P. godoni and P. tulipaformis. It is the proportion of integration due to covariation among the characters which has changed, not the pattern of interactions. Despite a stable framework of constraints, the behavior of the constraints has changed.

When P. robustus, the terminal member of the godoniform lineage, is compared to P. tulipaformis, its putative ancestor, the hypothesis that they share a common factor-patterns fails (X^2 = 114.51, df = 80, p = .007). Apparently even the identity of the constraints has changed. Similarly, P. robustus differs from P. godoni (p = .000 for the hypothesis that the factor pattern is stable). Furthermore, the three godoniform taxa cannot be fit simultaneously to a common model (p = .000), despite the similarity between P. godoni and P. tulipaformis. Thus during the divergence of P. robustus, the godoniform lineage undergoes a divergence in the nature, not just the intensity, of constraints.

While it appears that the constraints themselves are stable within each lineage, changing significantly only in the divergence of P. robustus, comparisons between the pyriform and godoniform lineage reveal a more complex transformation (Table 33). P. pyriformis is similar not only to P. symmetricus, but also to P. godoni and P. tulipaformis. However, P. symmetricus differs from P. godoni in all parameters and only slightly resembles P. tulipaformis in factor-pattern. Despite the similarity between P. symmetricus and P. pyriformis, P. symmetricus differs from other taxa similar to P. pyriformis. Apparently, those patterns of integration common to the two pyriform taxa must differ from those patterns shared by P. pyriformis and the godoniform taxa.

If P. pyriformis is interpreted as primitive to P. symmetricus, then P. symmetricus, P. godoni and P. tulipaformis share some aspects of primitive similarity. Over time, as the two lineages diverge, they retain different aspects of the primitive pattern of integration. If P. symmetricus is interpreted as the primitive pyriform species, then P. pyriformis has converged upon the godoniform set of constraints. Although the primitive pyriform and godoniform taxa differ in factor-pattern, under this view P. pyriformis becomes modified to share common complexes of integrated characters with the godoniform P. godoni and P. tulipaformis.

Whichever phylogenetic hypothesis is adopted, the same

Table 33.-- Comparisons in parameters of integration between godoniform and pyriform taxa. Presented are the p levels for the hypotheses that the factor pattern (F), factor-covariance (Fcova) and uniqueness of the characters (U) are common.

	P. godoni	P. tulipaformis
	F Fcova U	F Fcova U
P. symmetricus	.000/.000/.000	.011/.000/.000
P. pyriformis	.229/.000/.138	.748/.587/.000

conclusion is deduced: the constraints themselves evolve. This process appears to be initially gradual and continuous: there is no change in factor pattern within the early stages of each lineage. The divergence between the lineages occurs by acquiring (or retaining) different aspects of a pattern. For example, in P. tulipaformis, the respiratory unit decreases its primitive association with the basal unit, and the deltoid and digestive units decrease in intensity of integration. In the pyriform lineage, the association between the deltoid and digestive characters is retained, but the relative contribution of the digestive and basal characters to the pattern of integration changes. No change is abrupt until the appearance of Glen Dean P. robustus. In constrast to this gradual transformation in factor-pattern, the repatterning of integration, which occurs only in the evolution of P. robustus, entails a disruption of the ancestral pattern.

The factor-pattern of integration reflects the set of constraints acting in each population. The invariance of this pattern constitutes evidence of shared constraints, preserved through morphological and phylogenetic divergence. Similarities in the factor-covariance and the degree to which each character varies independently reveal not only a stable framework of constraints, but also the stable behavior of these shared constraints. Evidently, all of these features of integration are stable within the pyriform lineage, but all evolve within the godoniform lineage.

If P. symmetricus is interpreted as the primitive pyriform taxon, then the primitive pyriform and godoniform taxa are similar in their factor-covariance. This similarity means that each component of integration is equally under the influence of the constraints, and the constraints interact to the same degree, in these two species. The two species differ, however, in the degree to which the individual characters are free to vary independently. Thus the constraints, while they equally influence components of integration, differ in their ability to regulate the variation in the individual characters.

Evolution of constraints and morphological differentiation

Changes in constraints can yield dramatic differences in morphology. P. robustus has diverged the most from the aspects of integration shared by the other godoniform taxa. It is also significantly different in morphology from all other taxa, both godoniform and pyriform. The difference between P. robustus and the other Pentremites was clearly revealed in the discriminant analysis. The significance of this difference is confirmed by multivariate analysis of variance (Table 34). Repatterning, in this case, certainly does accompany morphological difference.

The multivariate analysis of variance indicates that the two Paint Creek taxa are significantly different in average morphology, as are the two Golconda species. Yet, morphology can evolve even within a stable framework of

Table 34.-- Relationship between morphological divergence and stability of intrinsic constraints. The F-ratios from multivariate analysis of variance are presented on the upper diagonal; the X^2 values for the hypothesis that the factor-pattern matrix is invariant are presented on the lower diagonal. P. robustus was highly significantly different in factor-pattern from all other Pentremites. Differences significant at the .05 level are indicated by *; at the .005 level are indicated by **.

	Sym	Pyr	God	Tulip	Rob
Sym	---	81.58**	107.76**	67.96**	3456.30**
Pyr	91.24	---	5.49**	108.03**	2724.30**
God	176.71**	95.41	---	56.85**	6095.20**
Tulip	115.56*	71.23	85.65	---	2842.36**

constraints. P. pyriformis and P. tulipaformis may share a common set of constraints, as indicated by the stability of factor-pattern, but they are significantly different in average morphology.

The discriminant analysis casts a somewhat different light on this analysis of the relationship between morphological similarity and the evolution of constraints. The characters which serve to discriminate among all five taxa classify P. pyriformis with P. godoni, and P. symmetricus with P. tulipaformis. While the two Paint Creek taxa do share common constraints, the two Golconda species differ significantly in factor-pattern.

They are classified together by the discriminant function because they do not differ in the means of those characters associated with the canonical variables of the discriminant function. Thus the apparent description of constrained evolution in the lineage drawn from the analysis of dispersion gives a misleading picture of the historical transition in constraints. Discriminant analysis forms a cluster of P. tulipaformis and P. symmetricus, although the multivariate analysis of variance reveals that they differ significantly in average morphology. The apparent convergence upon the same average morphology may be merely a consequence of the attempt to distinguish between all five taxa simultaneously by the principal axes of dispersion. When taxa do not differ in the characters which contribute most to the discriminant function, they are not

distinguished in discriminant analysis. The apparently replicated domains of variation exhibited in each of the two lineages appear to be an artifact of the analysis.

Discussion

The patterns of phenotypic covariation in each of these five species of Pentremites are developmentally and functionally constrained. The principal constraint upon phenotypic integration is the response of all characters to general body growth. More than half of the covariation among characters in each population can be explained as a response to overall body growth. In addition to this developmental source of integration, functionally interacting characters covary among themselves to form complex characters. However, functional units rarely represent discrete unit characters. Instead, interactions among functional units, and the organization of these functional units into distinct factors, create complex networks of integrated characters. It is these networks which constitute the complex characters capable of responding in a unified way to evolutionary processes.

When constraints upon phenotype are unbreakable and resistant to change, the common response of all characters to general body growth and interactions among functional characters could have the power to direct phenotypic transformations. However, as long as the intrinsic

constraints are malleable and capable of variation, the ancestral constraints lose their power to control the path adopted by the evolving population. The results of this study suggest that even when the identity of the intrinsic constraints may remain stable, evident in the invariant factor-pattern, the behavior of these constraints can be variable and capable of evolutionary change. Only within the pyriform lineage were constraints invariant. More frequently, some modification of constraints occurred, ranging from alterations in the importance of the particular constraints to destruction and repatterning of constraints.

Transformations in constraints can modify the influence of the constraints upon individual characters, removing particular characters from the complex and thus permitting mosaic evolution of formerly integrated characters. Alternative transformations could decrease or increase the influence of constraints upon the phenotype, affecting the amount of variation available in a whole complex. The most dramatic change entails a disruption and subsequent reconstruction of the constraints. All of these modifications occur within Pentremites and cause the evolutionary potential of the complex characters to differ in the different species.

This particular analysis of evolving constraints is limited by the tentative nature of the identification of species and of the hypothesized relationships among species. Furthermore, the small sample sizes may render some

conclusions suspect. However, alternative phylogenies will not alter the principal conclusion: that constraints change over the course of evolution. The small sample size might hinder discovery of developmental sources of covariation, because spurious covariation might be confounded with covariation due to shared developmental processes. But the small sample size does not affect the conclusion that the differences between species are significant-- indeed, with larger sample sizes even more differences might be discerned.

Even with small sample sizes, the difference in factor-pattern between P. robustus and the other Pentremites is striking. Repatterning may constitute one paradigm for macroevolutionary change (Bookstein et al., 1985). This analysis confirms that significant morphological change can accompany repatterning. The morphological difference, however, is not the key to analysis of evolutionary mechanisms. Even if repatterning and morphological change are independent, as they may be, repatterning of constraints might fundamentally alter the paths of phenotypic change available to a population and thus define a category of macroevolutionary change potentially independent of morphological divergence.

The breakdown and reconstruction of constraints throughout phylogeny does not result in derived taxa with weak constraints upon phenotype. In no population are the characters free to vary independently. Rather, growth and

interactions among characters seem to determine observed patterns of integration. In P. robustus the pattern of integration is complex, and not due to interactions between discrete functional units. However, even in this species, phenotypic characters appear restricted in their patterns of variation and covariation by their response to intrinsic constraints.

These intrinsic factors constraining covariation in each population may differ between ancestor and descendant species. As the descendant species diverges from its ancestor, it may gradually acquire a new set of constraints upon phenotypic covariation. In effect, the path followed by the descendant need not be one available to the ancestor. Instead, a descendant might follow a new, and evolving path. It might even have available a whole new range of potential morphologies available. The patterns of integration among characters in the ancestral population do not necessarily determine historical transformations because the constraints themselves can evolve. The constraints upon phenotypic evolution lie not only in the constraints upon the phenotype, but also in the constraints upon change in patterns of integration throughout the history of a lineage.

This analysis of evolving phenotypic constraints in Pentremites documents that modifications of the pattern of integration do occur. The mechanisms responsible for changing constraints, however, cannot be specified, as no analysis has yet explored the relationship between such

processes as selection, mutation, speciation and patterns of integration. Perhaps intense selection could alter patterns of phenotypic correlations, or a mutation in a gene involved in pattern formation could change the patterns of integration among characters. Evolving systems of pleiotropy might also alter the patterns of phenotypic integration by changing patterns of genetic integration. And perhaps different kinds of changes in integration are caused by different mechanisms. Yet, the absence of any theory postulating the mechanisms responsible for changing patterns of integration does not diminish the importance of these changes.

Changes in patterns of phenotypic integration throughout morphological and phylogenetic evolution undermine the idea that intrinsic constraints, imposed by developmental and functional interactions, create stable paths along which the phenotype evolves. Whenever the constraints upon phenotype lack historical stability, constraints upon the ancestral phenotype cannot act as constraints upon phenotypic evolution. The path of phenotypic change might then be determined by evolving, rather than ancestral, constraints.

CONCLUSION: CONSTRAINTS AND MORPHOLOGICAL EVOLUTION

Of all the developmental processes investigated in these studies, only general body growth appears to influence strongly the observed patterns of integration. No morphogenetic mechanism, other than growth, significantly influences covariation. Nor does any developmental process seem to define paths for evolving morphologies that are stable over either ontogeny or phylogeny. However, the results of these studies do not seriously threaten the hypothesis that developmental constraints, responsible for patterns of phenotypic covariation, guide the direction of morphological evolution. Only a small number of models were tested; thus there may be a variety of morphogenetic mechanisms (not explored in these studies) which determine observed covariation. Samples sizes were small, and the procedures employed in the analysis are designed for large samples. In addition, all individuals analyzed were either post-natal or post-metamorphic, and morphogenetic constraints might exert stronger influences at earlier stages of development. Furthermore, the measures analyzed were not specifically designed to test the models; thus they might not respond to the postulated developmental mechanisms as sensitively as might other measures.

However, the results of these studies do challenge the perhaps naive idea that a stable set of ancestral morphogenetic mechanisms invariantly exert a steady, dominant influence upon the evolving phenotype and regulate its evolution. While there may well be mechanisms of morphogenesis which limit the random behavior of individual characters and cause particular characters to covary, their effects are neither obvious nor universal within a lineage. Even tissue interactions, classically regarded as a potential constraint upon the mosaic evolution of complex adaptations (Hall, 1975; Maderson, 1975; Maderson et al., 1982), seem unable to explain observed integration. The developmental processes which might influence patterns of covariation are not evident merely by inspection of the average morphology within an age-class or lineage. More sophisticated and refined morphogenetic models must be analyzed before we can conclude that morphogenesis does not constrain phenotypic covariation.

Perhaps the failure to detect developmental constraints is a result of the rather naive search for a set of universal mechanisms which regulate cell-cell and tissue interactions within groups of distantly related organisms. Patterns of covariation among characters change, both during ontogeny and during phylogeny, and the mechanisms responsible for these patterns of covariation are probably similarly dynamic. Common tissue origin, shared mechanisms of chondrogenesis, joint response to an inductive stimulus,

etc. seem too universal in scope to act as dynamic constraints upon phenotypic evolution. Indeed, there is no reason to suppose that developmental processes responsible for observed covariation act as universal constraints for an entire lineage. Rather than inherited from the ancestor, and preserved unchanged throughout the history of a lineage, the constraints upon phenotype may themselves evolve and alter the available paths.

The analysis of functional integration in Pentremites, and the temporal variation in patterns of integration in Rattus suggest that characters engaged in a common function covary. Interactions among characters engaged in a common activity during post-natal life also appear to constrain covariation, but no developmental model incorporated these functional considerations. Both the dynamics of integration in the rat, and the patterns of integration in post-metamorphic Pentremites, appear to reflect the functional interactions among characters.

There may well be a developmental basis for this functional integration. On the other hand, the functional aspects of observed integration may, in part, arise from actual use of the morphological structures. If so, then these observed patterns of functional integration may be essentially environmental in origin and incapable of constraining any evolutionary response to selection. It is only the heritable aspects of integration that have any ability to guide morphological evolution. When functional

interactions generate purely phenotypic covariation, and are essentially independent of genetic covariance, then they are potentially irrelevant to evolutionary theory.

However, there is no reason to suppose that the joint heritability of functionally integrated characters should be lower than that of morphogenetically integrated characters. Patterns of pleiotropy may evolve to ensure this genetic coordination among characters whose coordinated activity would increase fitness (Cheverud, 1984). Thus the fitness of an individual might be determined, to some extent, by its genetic integration among functionally related characters. Just because functional interactions among characters might respond to use and disuse of characters, and thus respond to particular environmental cues such as diet, temperature, local water currents, etc., does not mean that they are exclusively environmental in origin.

Despite the attention given to such embryological factors as the mechanisms of chondrogenesis in the vertebrate limb (Shubin and Alberch, 1986), the embryology of function is neglected in studies of the developmental basis of macroevolutionary change in morphology. Yet, the ontogeny of function may establish developmental constraints which override the influence of morphogenetic mechanisms upon anatomical structures. Prenatal and infant muscle loading and neuromuscular interactions may affect not only the shape of the cranium and jaw (Atchely et al., 1984) but also the patterns of integration in the skull. The concept

of developmental constraints should be expanded to include epigenetic functional interactions. A narrow view of developmental constraints which neglects the development of function in favor of cell and tissue interactions appears inadequate to explain observed integration through post-natal growth of the rat and Pentremites.

Mechanisms of chondrogenesis, local growth gradients and tissue interactions may exert an influence not detected in these studies. The inability to detect any effects of these processes may be not only a consequence of the choice the choice of models tested, but also of of measures selected and the populations analyzed. Conventional measures, such as those taken on Rattus, may impede discovery of such biological processes. The developmental hypotheses focus upon local interactions among cells and tissues, but the measures redundantly sample a small set of dimensions of variation. Thus the geometry of the measurement scheme might limit the analysis to the discovery of global factors acting to determine the common response of all characters. In effect, the significance of growth as a developmental factor of covariation might be less a consequence of the actual importance of growth than of the redundancy of the measurement scheme.

The geometry of the measurement scheme is not the only bias introduced by the selection of measures. These measures were not selected a priori, according to the hypotheses. Thus the hypotheses might effectively predict

the covariation among another set of characters, which do respond to the hypothesized morphogenetic processes. Any attempt to salvage developmental information from phenotype must examine those aspects of phenotype most likely to respond to the hypothesized developmental constraints. Thus I cannot conclude that morphogenesis exerts no influence upon phenotypic covariation. I can only conclude that morphogenesis exerts an indetectable influence, at best, upon these particular characters.

No analysis of the causes of integration can ever exhaust all potential hypotheses, or even construct all possible models specified by these hypotheses. So it is quite possible that the developmental factors which would have adequately reconstructed observed phenotypic variance-covariance were not analyzed. However, the rejection of particular models at least excludes certain hypotheses from further consideration. While this approach to identifying potential morphogenetic causes of integration may never falsify the hypothesis that some morphogenetic agent is responsible for covariation, it can test particular hypotheses about specific agents.

The approach to developmental constraints adopted here employs a specific definition of constraints. Constraints are defined as the biological processes which generate patterns of integration. Developmental constraints are those specifically developmental processes which determine covariation among developmentally associated characters.

Other approaches to developmental constraints emphasize such properties of constrained systems as discrete, bounded distributions of mean phenotypes (e.g. Alberch, 1982; Oster and Alberch, 1982; Waddington, 1976). According to such alternative definitions, constraints are the rules which prevent particular phenotypes from being made, or from surviving. My concept of constraint as a source of phenotypic covariation, however, departs from the emphasis upon the average phenotype. Rather than searching for the causes of limits upon body size, or the for underlying developmental basis for the recurrence of certain proportions among characters, or for the rules according to which specific patterns in morphology are determined (e.g. Alberch and Alberch, 1981; Alberch and Gale, 1985; Raup, 1966, 1967; Shubin and Alberch, 1986) my approach concentrates upon the causal analysis of covariance among characters. The causal analysis of this non-random variation seeks to extract information about the biological mechanisms responsible for covariation from the structure of observed phenotypic covariance. The procedure employs statistical inference to ask if particular biological processes determine observed covariance. When the statistical model, deduced from the biological theory, reconstructs observed variance-covariance structure, the hypothesized biological mechanisms are interpreted as constraints.

The descriptive aspects of morphology are neglected by this focus upon character covariance, but this approach is

justified by its ability to unite morphogenetic and quantitative genetic studies of constrained phenotypic evolution. From morphogenetic theory, predictions about the patterns of covariation among characters influenced by particular developmental processes can be formulated and tested. The quantitative genetic approach to the study of phenotypic evolution makes predictions about the evolutionary consequences of these patterns of covariation. In effect, this definition of developmental constraint as the factors of covariance determined by developmental mechanisms fuses the otherwise disparate modes of analyzing the causes of morphological evolution.

The mechanisms of morphological evolution are selection and random genetic drift. The interplay between the evolutionary mechanisms and the constrained patterns of variation produce the observed patterns of phenotypic evolution. The developmental factors which determine the phenotypic variance-covariance structure are not necessarily constraints upon phenotypic evolution. Rather, these developmental constraints reflect those biological processes which affect the phenotype <u>within</u> a population. Constraints upon phenotypic evolution lie in those processes which limit variation in the structure of covariance during evolution. Thus studying patterns of phenotypic covariation within populations and subsequently comparing these patterns in related populations distinguishes between 1) the factors which constrain covariation among phenotypic characters; and

2) the factors which constrain historical, geographic and temporal variation in these patterns of covariation. The analysis of phenotypic integration within populations addresses the causal basis for constraints upon integration. The comparisons among factor-patterns in successive age-classes, in geographic populations and in related species tests hypotheses about the stability of the patterns of phenotypic integration.

Even if morphogenetic processes were to constrain the phenotype, they would not necessarily define limits upon evolutionary change. Discovery of the morphogenetic basis of phenotypic covariation would merely establish that constraints upon the phenotype within an individual population are developmental in origin. But to play a role in phenotypic evolution, these constraints must determine how one phenotype can be transformed another. When the constraints are unchanged by speciation or selection, they can direct the transformation in average phenotype. However, when the constraints upon the ancestral phenotype vary throughout the course of ontogeny, or even themselves evolve, they cannot so unambiguously define the set of potential paths available to all descendants.

The comparative study of integration in Rattus norvegicus demonstrates that developmental constraints are not stable throughout the course of post-natal growth. Rather, these patterns of integration are dynamic. The factors of covariation change, as does the amount of

integration associated with these factors and the degree to which these factors influence particular characters. Changes in patterns of phenotypic covariation during the course of post-natal growth may be either programmatic, determined by intrinsic genetic factors, or due to immediate responses to the environment. Certainly these dynamics can yield constraints on evolution. Selection might be forced to act at particular ages, or upon particular patterns of integration, because of the sequence and timing of dynamic integration. Selection, acting upon specific phenotypes with specific patterns of integration, would produce different effects depending upon the age at which it acts. The constraints would not lie in any particular pattern of integration, because no one pattern limits the possibility of achieving a different one. The outcome of selection would depend upon the age at which it acts.

Not only are patterns of integration unstable throughout ontogeny, they also change throughout the course of evolution. The results of the analysis of evolving constraints in Pentremites undermines the idea that intrinsic constraints acting within an ancestral population define a set of potential paths in the adaptive landscape. Intrinsic ancestral constraints appear unlikely to limit the direction of phenotypic change to particular paths available to the ancestral population because the paths available to a population at a given time can change.

According to traditional microevolutionary theory, the

genetic covariation (the explicit representation of intrinsic constraints) is constant throughout morphological evolution. Selection produces predictable effects upon the phenotype because the selection differential and the intrinsic constraints interact to create the new phenotype. And much of macroevolutionary theory regards the developmental processes as conservative, a "resilient developmental programme" (Alberch and Alberch, 1981) which prevents the expression of mutation and precludes particular phenotypic changes. Perhaps developmental mechanisms determine phenotypic integration at earlier stages of ontogeny so that comparisons between early stages of ontogeny within a lineage would reveal a stable pattern of integration. However, the variability in the patterns of phenotypic integration in samples of related Pentremites makes the hypothesis that these constraints are transmitted from ancestor to descendant, and guide the morphological transformation within a lineage, appear untenable.

In the evolution of Pentremites, patterns of integration undergo both minor modification and repatterning. Thus the constraints acting in any one population may not be able to guide evolution in all descendant species. In effect, ancestral constraints do not demarcate stable paths along which the evolving population can move. Rather, the paths between phenotypes changes as the phenotype evolves. Unfortunately, changes in these paths cannot be perceived by examination of the changes in average

phenotype. The relationship between changes in the average morphology evolution and the paths defined by intrinsic constraints is complex.

The derived pyriform species, P. pyriformis, shares a common set of constraints with the primitive godoniform, P. godoni, and differs little in morphology, as judged by discriminant analysis. However, the primitive pyriform, P. symmetricus, differs from all godoniform species in constraints, but morphologically resembles P. godoni as closely as does P. pyriformis in the characters associated with the first two canonical variables. Thus, P. pyriformis seems to have preserved its primitive morphology, but converged upon the characteristic godoniform constraints. During its course of evolution, the godoniform P. tulipaformis diverges morphologically not only from the pyriform species in morphology but also from P. godoni. Yet, the primitive godoniform constraints are not lost, merely modified. In contrast to the minor modifications of constraints exhibited in the transition between the two pyriform species and the evolution of P. tulipaformis from its primitive godoniform ancestor, P. robustus not only undergoes substantial morphological change, but also acquires a unique set of constraints.

Neither traditional microevolutionary theory nor macroevolutionary theory can explain how patterns of phenotypic might be reorganized during evolution. Repatterning of integration may be considered as a paradigm

instance of a macroevolutionary event (Bookstein et al., 1985). But there is no reason to suspect that any of the mechanisms which alter the pattern or details of integration lie outside the boundaries of traditional microevolutionary theory. The causes of repatterning cannot yet be specified because they have not yet been investigated. Repatterning might result from any of several different biological processes: mutation, intense selection, perhaps migration and introgression, or even from chromosomal inversions. While current microevolutionary theory lacks both the concept and a term for repatterning, none of the potential causes of repatterning are outside the scope of microevolutionary theory. Thus, repatterning is a macroevolutionary event consistent with microevolutionary theory but so far excluded from it.

Similarly, changes in the patterns of integration which merely modify the details of integration are excluded from current theory. These specific parameters of integration have not been distinguished, explicated and compared in related species. Thus the biological mechanisms which change both factor covariation and the ability of particular characters to vary independently cannot yet be identified. Yet, these parameters of integration are simply aspects of covariation structure, and so amenable to quantitative analysis.

Traditional microevolutionary theory, as it now stands, cannot analyze the dynamics of developmental constraints

throughout evolution. Nor can it isolate the causes of developmental constraints upon phenotype because it does not investigate the relationship between the phenotype and genotype. The causal study of epigenetic interactions and analysis of the mechanisms of evolving patterns of integration should be incorporated into microevolutionary theory. Study of development integration and unstable constraints would expand, rather than challenge, the quantitative genetic approach to morphological evolution.

By emphasizing only those additive aspects of genotype, and presuming that the structure of genetic covariance is transmitted from one generation to another, traditional evolutionary theory neglects critical information about phenotypic evolution. Traditional microevolutionary theory emphasizes the genetic aspects of evolutionary theory because it is the genotype which determines the potential response to selection and drift. Yet, studies of the patterns of phenotypic integration are not merely substitutes for genetic analysis. The actual phenotypic outcome of natural selection, drift and mutation is not always adequately summarized in genetic covariation. Traditional evolutionary theory studies changes in the genetic structure of the population, but the effects upon phenotype of these changes must also be incorporated into theory.

Only when the causes of genetic covariance are also responsible for phenotypic covariation can developmental

constraints be adequately estimated by the genetic factors. And only when the patterns of genetic integration are invariant throughout the course of evolution can the processes which translate genetic covariance into phenotypic covariance be ignored. This invariance is presumed in the current microevolutionary analysis of phenotypic change. Under such conditions, when genetic covariance is invariant, the descendant phenotype can be derived from the ancestral phenotype by a simple process of selection. The differences between the two species would not be a consequence of changes in developmental mechanisms, but of selection acting within a stable framework of constraints. Yet, even under these restrictive and perhaps unrealistic conditions, the descendant phenotype might not be predictable. Perhaps the two populations differ in their response to environmental factors, or in the interactions between genes, or in the interaction between environment and genes. While these environmental and epistatic factors might not be transmitted from one species to another they will certainly affect the phenotype within each population.

Currently, the relationship between genetic and phenotypic covariation is unclear. Yet, the phenotypic consequences of changes in genetic covariance structure may be critical to any understanding of phenotypic evolution. Considering the frequency of changes in patterns of integration throughout ontogeny and phylogeny, evolutionary and developmental processes responsible for these changes

should be explored further. As currently framed, traditional quantitative genetic theory of microevolution cannot support this analysis because it lacks the necessary terms and concepts. Although it is possible to estimate the effects of selection upon the phenotype, and to explain how genetic covariance constrains this response, it is not possible to ask if changes in phenotype are caused by: 1) changes in genetic covariance; 2) changes in the way in which the genetic covariance is expressed through development; or 3) selection acting through these genetic constraints. There are too few variables in the equations.

Some of the observed changes in patterns of phenotypic integration may reflect underlying changes in genotypic integration. Recent analysis of the stability of additive genetic variance suggests that genetic covariance structure is stable only within species, at best, and may differ even between subspecies (Lofsvold, 1986). If this instability in genetic covariance structure is a frequent consequence of geographic differentiation and speciation, the theory of microevolution needs to add a term representing an operator acting upon genetic covariation. At the very least, this potential for change in genetic covariance must be explicitly recognized by microevolutionary theory.

Changes in the developmental constraints which translate genotype into phenotype, perhaps in part produced by changes in genetic covariance, should also be amenable to analysis. Of course any evolutionary change in morphology

must have an underlying genetic basis. However, changes in the non-additive portion of genetic variation will not alter the genetic covariation, but may repattern the developmental constraints. And changes in particular biological mechanisms responsible for covariation may cause changes in genetic covariance but leave developmental constraints undisturbed. Not all aspects of genetic covariance determine the pattern of developmental integration. Some of the genetic covariance among characters reflects a joint response to factors of integration that are not developmental in origin. Thus the theoretical framework of microevolution must distinguish between genetic and developmental constraints upon phenotype and their evolution.

To perform this analysis, the constraints upon phenotype must be regarded as a critical component of the analysis of phenotypic evolution. The patterns of integration evident in analysis of genetic covariance cannot substitute for analysis of phenotypic patterns. And the theory should comprehend dynamics of integration (both phenotypic and genetic). Extending the theoretical framework to incorporate unstable patterns of genetic covariation, and evolving constraints upon phenotype, can place genotype and phenotype into the same dynamic landscape, and provide a role for development in evolution. Furthermore, the analysis of dynamic constraints might reveal how evolving morphology affects the evolutionary behavior of intrinsic constraints.

BIBLIOGRAPHY

BIBLIOGRAPHY

Alberch, P. 1980. Ontogenesis and morphological diversification. Amer. Zool. 20:653-667.

Alberch, P. 1982. Developmental constraints in evolutionary processes. Pages 313-332 in Bonner, J. T. (ed.). Evolution and Development. Springer-Verlag, New York.

Alberch, P. 1983. Morphological variation in the neotropical salamander genus Bolitoglossa. Evolution 34:906-919.

Alberch, P. and J. Alberch. 1981. Heterochronic mechanisms of morphological diversification and evolutionary change in the neotropical salamander Bolitoglossa occidentalis (Amphibia: Plethodontidae). J. Morph. 167:249-264.

Alberch, P. and E.A. Gale. 1985. A developmental analysis of an evolutionary trend: Digital reduction in amphibians. Evolution 39:8-23.

Alberch, P., S.J. Gould, G.F. Oster, D.B. Wake. 1979. Size and shape in ontogeny and phylogeny. Paleobiology. 5:296-317.

Andrews, H. E., J. C. Brower, S. J. Gould, R. A. Reyment. 1974. Growth and variation in Eurypterus remipes Dekay. Bull. geol. Instn. Univ. Upsala: N. S., 4,6:81-114.

Armstrong, J. S. 1967. Derivation of theory by means of factor analysis or Tom Swift and his electric factor analysis machine. The American Statistician 21:17-21.

Atchley, W. R. 1984. Ontogeny, timing of development, and genetic variance-covariance structure. Amer. Natur. 123:519-540.

Atchley, W. R., S. W. Herring, B. Riska, and A. A. Plummer. 1984. Effects of the muscular dysgenesis gene on developmental stability in the mouse mandible. Jour. Craniofac. Gen. Dev. Biol. 4:179-189.

Atchley, W. R. and J. J. Rutledge. 1980. Genetic components of size and shape. I. Dynamics of components

of phenotypic variability and covariability during ontogeny in the laboratory rat. Evolution 34:1161-1173.

Atchley, W. R., J.J. Rutledge, and E. Cowley. 1981. Genetic components of size and shape. II. Multivariate covariance patterns in the rat and mouse skull. Evolution 35: 1037-1055.

Barton, A. D. and A. K. Laird. 1969. Analysis of allometric and non-allometric differential growth. Growth 33:1-16.

Beaver, H. H. 1967. Morphology. Pages 300-350 in Moore, R. C. (ed.) Treatise on Invertebrate Paleontology Part S: Echinodermata. Geol. Soc. Amer. and Univ. of Kansas, Lawrence, Kansas.

Beecher, R. M., R. S. Corruccini, and M. Freeman. 1983. Craniofacial correlates of dietary consistency in a nonhuman primate. J. Craniofac. Genet. Dev. Biol. 3:193-202.

Bentler, P. M. and D. G. Bonett. 1980. Significance tests and goodness of fit in the analysis of covariance structures. Psychological Bulletin 88:588-606.

Bookstein, F. L. 1982a. Foundations of morphometrics. Ann. Rev. of Ecol. and Syst. 13:451-470.

Bookstein, F. L. 1982b. Panel discussion-- Modeling and method. Pages 317-321 in Joreskog, K. G. and H. Wold (eds.) Systems Under Indirect Observation: Causality, Structure, Prediction I. North-Holland, Amsterdam.

Bookstein, F. L. 1984. A statistical method for biological shape comparison. Jour. Theor. Biol. 107:475-520.

Bookstein, F. L., B. Chernoff, R. L. Elder, J. M Humphries, Jr., G. R. Smith, R. E. Strauss. 1985. Morphometrics in Evolutionary Biology. The Academy of Nat. Sci. of Phila., Spec. Pub. 15.

Boomsma, A. 1982. The robustness of LISREL against small sample sizes in factor analysis models. Pages 149-173 in Joreskog, K. G. and H. Wold (eds.). Systems Under Indirect Observation: Causality, Structure, Prediction I. North-Holland, Amsterdam.

Bouvier, M. and W. L. Hylander. 1981. Effect of bone strain on cortical bone structure in macaques (Macaca mulatta). J. Morph. 167:1-12.

Cheverud, J. M. 1982. Phenotypic, genetic and environmental morphological integration in the cranium.

Evolution 36:499-516.

Cheverud, J. M. 1984. Quantitative genetics and developmental constraints on evolution by selection. J. theor. Biol. 110:155-171.

Cheverud, J. M., J. J. Rutledge, and W. R. Atchely. 1983. Quantitative genetics of development: Genetic correlations among age-specific values and the evolution of ontogeny. Evolution 37:895-905.

Cuvier, 1826. Histoire des Progres des Sciences naturelles depuis 1789,i. Paris.

Dixon, W. J. 1985. BMDP Statistical Software. Univ. Calif. Press, Berkeley.

Donaldson, H. H. 1924. The Rat. Wistar Institute Memoirs, 6.

Ebert, T. A. 1980. Estimating parameters in a flexible growth equation, the richards function. Can. J. Fish. Aquat. Sci., 37:687-692.

Eldredge, N. 1973. Systematics of Lower and Lower Middle Devonian species of the tilobite Phacops Emmrich in North America. Bull Amer. Mus. Nat. Hist. 151:285-338.

Eldredge,N. and S.J. Gould. 1972. Punctuated equilibria: an alternative to phyletic gradualism. In Schopf, T.J.M. (ed.). Models in Paleobiology. Freeman, Cooper and Co., San Francisco

Frazetta, T.H. 1975. Complex Adaptations in Evolving Populations. Sinauer Assoc., Sunderland.

Garcia-Bellido, A. 1983. Comparative anatomy of patterns in the genus Drosophila. Pages 227-255 in Goodwin, B. C., N. Holder and C. C. Wylie (eds.). Development and Evolution. Cambridge Univ. Press, Cambridge.

Gerhart,J. C., S. Berking, J. Cooke, G. L. Freeman, A. Hildebrandt, H. Jokusch, P. A. Lawrence, C. Nusslein-Volhard, G.F. Oster, K. Sander, H. W. Sauer, G. S. Stent, N. K. Wessells, L. Wolpert. 1982. The cellular basis of morphogenetic change Group Report. Pages 87-114 in Bonner, J. T. (ed.). Evolution and Development. Springer-Verlag, New York.

Gould, S. J. 1977. Ontogeny and Phylogeny. Belknap Press, Cambridge.

Gould, S.J. 1980. Is a new and general theory of evolution emerging? Paleobiology 6:119-130.

Gould, S.J. 1982. Change in developmental timing as a mechanism of macroevolution. Pages 333-346 in Bonner, J. T. (ed.). Evolution and Development. Springer-Verlag, New York.

Gould, S. J. 1984. Covariance sets and ordered geographic variation in Cerion from Aruba, Bonaire and Curacao: a way of studying nonadaptation. Syst. Zool. 33:217-238.

Gould, S.J. and N. Eldredge. 1977. Punctuated equilibria: the tempo and mode of evolution reconsidered. Paleobiol. 3: 115-151.

Gould, S. J. and R. A. Garwood. 1969. Levels of integration in mammalian dentitions: an analysis of correlations in Nesophontes micrus (Insectivora) and Oryzomys couesi (Rodentia). Evolution 23:276-300.

Gould, S. J. and R. C. Lewontin. 1979. The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. Proc. R. Soc. Lond. B 205:581-598.

Grueneberg, H. 1963. Pathology of development. Wiley Co., New York.

Hall, B. K. 1975. Evolutionary consequences of skeletal differentiation. Amer. Zool., 15:329-350.

Hall, B. K. 1982. Distribution of osteo- and chondrogenic neural crest-derived cells and of osteogenically inductive epithelia in mandibular arches of embryonic chicks. J. Embryol. Exp. Morph. 68:127-136.

Harman, H. 1967. Modern Factor Analysis. Univ. Chicago Press, Chicago.

Herring, S. 1985. The ontogeny of mammalian mastication. Amer. Zool. 25:339-349.

Hoffmann, H. 1981. Biological controls of the punctuated versus gradual mode of species evolution. Pages 57-63 in J. Martinell (ed.). International Symposium of Concept and Method in Paleontology. Dept. of Paleo,University of Barcelona, Barcelona.

Hughes, P. C. R. and J. M. Tanner. 1970. A longitudinal study of the growth of the black-hooded rat: methods of measurement and rates of growth for skull, limbs, pelvis, nose-rump and tail lengths. J. Anat. 106:349-370.

Humphries, J. M., F. L. Bookstein, B. Chernoff, G. R.

Smith, R. L. Elder and S. G. Poss. 1981. Multivariate discrimination by shape in relation to size. Syst. Zool., 30:291-308.

Jolicoeur, P. 1963. The multivariate generalization of the allometry equation. Biometrics 19:497-499.

Joreskog, K. G. 1969. A general approach to confirmatory factor analysis. Psychometrika 34:183-202.

Joreskog, K. G. 1975. Structural equation models in the social sciences: specification, estimation and testing. Pages 265-287 in Krishnaiah, P. R. (ed.). Applications of Statistics. North-Holland, Amsterdam.

Joreskog, K. G. and D. Sorbom. 1984. LISREL VI: Analysis of Linear Structural Relationships by Maximum Likelihood, Instrumental Variables and Least Squares Methods. Scientific Software, Mooresville, Indiana.

Joreskog, K. G. and H. Wold. 1982. Systems Under Indirect Observation: Causality, Structure, Prediction. North-Holland, Amsterdam.

Kauffman, S. 1983. Developmental constraints: Internal factors in evolution. Pages 195-225 in Goodwin, B. C., N. Holder and C.C. Wylie (eds.). Development and Evolution, Cambridge Univ. Press, Cambridge.

Kolliker, A. 1864. Ueber die Darwinische Schoepfungstheorie. Zeit. f. Wiss. Zool. 14:74-86.

Kosher, R. A. 1983. The chondroblast and the chondrocyte. Pages 59-85 in Hall, B. K. (ed.). Cartilage, 1. Academic Press, New York.

Kurten, B. 1953. On the variation and population dynamics of fossil and recent mammal populations. Acta Zool. Fenn. 76: 1-122.

Laird, A. K. 1966. Dynamics of embryonic growth. Growth 30:263-275.

Laird, A. K., A. D. Barton and S. Tyler. 1968. Growth and time: an interpretation of allometry. Growth 32:347-354.

Laird, A. K., S. Tyler and A. D. Barton. 1965. Dynamics of relative growth. Growth 29:249-263.

Lande, R. 1976. Natural selection and random genetic drift in phenotypic evolution. Evolution 30:314-334.

Lande, R. 1978. Evolutionary mechanisms of limb loss in

tetrapods. Evolution 32:73-92.

Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain : body size allometry. Evolution 33:402-416.

Lande, R. 1985. Expected time for random genetic drift of a population between stable phenotypic states. Proc. Natl. Acad. Sci. 82:7641-7645.

Lande, R. 1986. The dynamics of peak shifts and the pattern of morphological evolution. Paleobiology 12:343-354.

Lande, R. and S.J. Arnold. 1983. The measurement of selection on correlated characters. Evolution 37:1210-1226.

Leamy, L. 1975. Component analysis of osteometric traits in randombred house mice. Syst. Zool. 24:176-190.

Leamy, L. 1977. Genetic and environmental correlations of morphometric traits in randombred house mice. Evolution 31:357-369.

Lindgren, B. W. 1968. Statistical Theory. The Macmillan Co., New York.

Lofsvold, D. 1986. Quantitative genetics of morphological differentiation in Peromyscus. I. Tests of the homogeneity of genetic covariance structure among species and subspecies. Evolution 40:559-573.

Macurda, D. B., Jr. 1967. Development and hydrodynamics of blastoids. Pages 356-382 in Moore, R. C. (ed.) Treatise on Invertebrate Paleontology Part S: Echinodermata. Geol. Soc. Amer. and Univ. of Kansas, Lawrence, Kansas.

Maderson, P.F.A. 1975. Embryonic tissue interactions as the basis for morphological change in evolution. Amer. Zool. 15:315-327.

Maderson, P.F.A., P. Alberch, B.C. Goodwin, S. J. Gould, A. Hoffman, J.D. Murray, D.M. Raup, A. de Ricqles, A. Seilacher, G.P. Wagner, D.B. Wake. 1982. The role of development in macroevolutionary change Group Report. Pages 279-312 in Bonner, J. T. (ed.). Evolution and Development. Springer-Verlag, New York.

Maynard Smith, J.M. R. Burian, S. Kauffman, P. Alberch, J. Campbell, B. Goodwin, R. Lande, D. Raup, L. Wolpert. 1986. Developmental constraints and evolution. Quart. Rev. Biol. 60:265-287.

McGhee, G. R., Jr. 1980. Shell form in the biconvex articulate Brachiopoda: A geometric analysis. Paleobiology 6:57-76.

McNamara, K. J. 1986. The role of heterochrony in the evolution of Cambrian trilobites. Biol. Rev. 61: 121-156,

Meredith, W. 1964. Rotation to achieve factorial invariance. Psychometrika 23:187-206.

Milne-Edwards, H. 1867. Rapport sur les Progres recents des Sciences zoologique en France. Paris.

Moore, W. J. 1965. Masticatory function and skull growth. J. Zool. 146:123-131.

Morrison, D. F. 1976. Multivariate Statistical Methods. McGraw-Hill Book Co., New York.

Mulaik, S. A. 1972. The Foundations of Factor Analysis. McGraw-Hill Book Co., New York.

Noden, D. M. 1982. Patterns and organization of craniofacial skeletogenic and myogenic mesenchyme: A persepctive. Pages 167-203 in Dixon, A. D. and B. G. Sarnat (eds.). Factors and Mechanisms Influencing Bone Growth. Alan R. Liss, Inc., New York.

Olsen, R. F., P. J. Wangsness, W. H. Patton, and R. J. Martin. 1980. Relationship of serum somatomedin-like activity and fibroblast proliferative activity with age and growth in the rat. Growth 44:19-28.

Olson, E.C. and R.J. Miller. 1958. Morphological Integration. Univ. Chicago Press, Chicago.

Oster, G.F. and P. Alberch. 1982. Evolution and bifurcation of developmental programs. Evolution 36: 444-459.

Oster, G.F., J.D. Murray and P.K. Maini. 1985. A model for chondrogenic condensations in the developing limb: the role of extracellular matrix and cell tractions. J. Embryol. exp. Morph. 89: 93-112.

Owen, R. 1868. Anatomy of Vertebrates, iii.

Parker, C. R.,Jr. and V. B. Mahesh. 1976. Hormonal events surrounding the natural onset of puberty in female rats. Biol. Repro. 14: 347-353.

Price, T.D. and P. R. Grant. 1985. The evolution of ontogeny in Darwin's finches: a quantitative genetic

approach. Amer. Nat. 125: 169-188.

Provine, W. B. 1986. Sewall Wright and Evolutionary Biology. The Univ. of Chicago Press, Chicago.

Rachootin,S. and K.S. Thomson, 1981. Epigenetics, paleontology and evolution. Evolution Today 2:181-193.

Raup, D. 1966. Geometric analysis of shell coiling: General problems. J. Paleo. 40:1178-1190.

Raup, D. 1967. Geometric analysis of shell coiling: Coiling in ammonoids. J. Paleo. 41:43-65.

Rickert, W. E. 1979. Growth rates and models. Fish Physiology, 8:677-743.

Russell, E. S. 1916. Form and Function. The Univ. Chicago Press, Chicago.

Schnute, J. 1981. A versatile growth model with statistically stable parameters. Can. J. Fish. Aquat. Sci. 38:1128-1140.

Shubin, N. H. and P. Alberch. 1986. A morphogenetic aproach to the origin and basic organization of the tetrapod limb. Evol. Biol. 20:319-388.

Simpson, G.G. 1953. The Major Features of Evolution. Columbia Univ. Press, New York.

Sorbom, D. 1974. A general method for studying differences in factor means and factor structures between groups. Brit. Jour. Math. Stat. Psych. 27:229-239.

Strauss, R. and F. L. Bookstein. 1982. The truss: body form reconstruction in morphometrics. Syst. Zool 31:113-135.

Swanson, H. E. and J. J. van der Weff ten Bosh. 1963. Sex differences in growth of rats, and their modification by a single injection of testosterone priopionate shortly after birth. J. Endocrin. 26:197-207.

Thom, R. 1975. Structural Stability and Morphogenesis. W. A. Benjamin, Inc., Reading, Mass.

Thorndike, R. M. 1978. Correlational Procedures for Research. Gardner Press, Inc., New York.

Thurstone, L. L. 1933. The Theory of Multiple Factors. Edward Bros., Ann Arbor.

Thurstone, L. L. 1947. Multiple Factor Analysis. Univ.

Chicago Press, Chicago.

Van Valen, L. 1960. Morphological integration. Evolution 19:347-349.

Vavilov, N.I. 1922. The law of homologous series in variation. J. Genet. 12: 67-87.

von Baer, K. E. 1876a. Ueber den Zweck in den Vorgangen der Natur. Reden, ii. St. Petersburg.

von Baer, K. E. 1876b. Ueber Darwin's Lehre. Reden, ii. St. Petersburg.

von Hartmann,E. 1906. Das Problem des Lebens. Biologische Studien. Bad Sacha.

Waddington, C.H. 1957. The Strategy of the Genes. Allen and Unwin, Ltd., London.

Waddington, C. H. 1974. A catastrophe theory of evolution. Pages 253-266 in Waddington, C. H. 1976. The Evolution of an Evolutionist. Cornell Univ. Press, Ithaca.

Waddington, C. H. 1976. The Evolution of an Evolutionist. Cornell Univ. Press, Ithaca.

Wake, D. B. 1980. Evidence of heterochronic evolution: a nasal bone in the Olympic salamander, Rhyacotriton olympicus. J. Herpetol. 14:292-295.

Wake, D. B., G. Roth and M. L. Wake. 1983. On the problem of stasis in organismal evolution. J. Theor. Biol. 101:211-224.

Waters, J. A. S. Horowitz and D. B. Macurda, Jr. 1986. Ontogeny and phylogeny of the Carboniferous blastoid Pentremites. J. Paleo. 59: 701-712.

Williamson, P. G. 1981. Palaeontological documentation of speciation in Cenozoic molluscs from Turkana Basin. Nature. 293:437-443.

Wright, S. 1921. Correlation and causation. Jour. Agric. Res. 20:557-585.

Wright, S. 1932a. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. Proc. Sixth Int. Congr. Genetic. 1:356-366.

Wright, S. 1932b. General, group and special factors. Genetics 17:603-619.

Wright, S. 1935. Evolution in populations in approximate

equilibrium. J. Genet. 30:235-266.

Wright, S. 1968. Evolution and the Genetics of Populations. I. Genetic and Biometric Foundations. Univ. of Chicago Press, Chicago.

Young, R. W. 1959. The influence of cranial contents on postnatal growth of the skull in the rat. Am. J. Anat. 105:383-415.

Zwilling, E. 1961. Limb morphogenesis. Adv. Morphogenesis 1:301-330.

Zwilling, E. 1968. Morphogenetic phases in development. Devel. Biol. Suppl. 2:184-207.