# Phenotypic Variability: Its Components, Measurement and Underlying Developmental Processes

**Article** *in* Evolutionary Biology · November 2007

DOI: 10.1007/s11692-007-9008-1



**3 authors:**

Katherine E Willmore
The University of Western Ontario
**23** PUBLICATIONS **890** CITATIONS
SEE PROFILE

Nathan M. Young
University of California, San Francisco
**62** PUBLICATIONS **2,447** CITATIONS
SEE PROFILE

Joan T. Richtsmeier
Pennsylvania State University
**228** PUBLICATIONS **6,546** CITATIONS
SEE PROFILE

Evol Biol
DOI 10.1007/s11692-007-9008-1

SYNTHESIS

# Phenotypic Variability: Its Components, Measurement and Underlying Developmental Processes

**Katherine Elizabeth Willmore · Nathan M. Young · Joan T. Richtsmeier**

Received: 12 January 2007 / Accepted: 9 July 2007


**Abstract** Variability contrasts with variation in that variability describes the potential for variation, not simply the expressed variation. The power of studying variability lies in creating a conceptual framework around which the relationship between the genotype and phenotype can be understood. Here, we attempt to demonstrate the importance of phenotypic variability, how it structures variation, and how fundamental developmental processes structure variability. Given the broad scope of this topic, we focus on three widely studied properties of variability: canalization, developmental stability and morphological integration. We have organized the paper to emphasize the importance of differentiating between the theory surrounding these components of phenotypic variability, their measurement and the biological factors surrounding their expression. First, we define these properties of variability, how they relate to each other and to variability as a whole. Second, we summarize the common methods of measurement for canalization, developmental stability and morphological integration and the reasoning behind these methods. Finally, we focus on jaw development as an example of how the basic processes of development affect variability and the resultant variation, with emphasis on how processes at all levels of the organismal hierarchy interact with one another and contribute to phenotypic variability.

**Keywords** Variability · Variation · Development · Canalization · Morphological integration · Developmental stability · Phenotype

K. E. Willmore (✉) · J. T. Richtsmeier
Department of Anthropology, Pennsylvania State University, 409 Carpenter Building, University Park, PA 16802, USA
e-mail: kew20@psu.edu

J. T. Richtsmeier
e-mail: jta10@psu.edu

N. M. Young
Department of Orthopaedic Surgery, University of California at San Francisco, 1001 Portrero Avenue, San Francisco, CA 94110, USA
e-mail: nyoung@post.harvard.edu

## Introduction

Phenotypic variation results from both genetic and environmental factors, and a central goal of biology is to understand the complex interactions that mediate the translation from genotype to phenotype. Closely related to phenotypic variation is the concept of variability. Phenotypic variability is defined as the tendency or potential of an organism to vary (Wagner and Altenberg 1996) and corresponds to the range or distribution of potential variation. While variation can be observed and documented, variability cannot be directly quantified. Variation is normally measured as a series of static observations within a sample, each observation representing a single instance of the many phenotypic expressions of the interaction of genetic and environmental factors. While variation is a static observation that represents one of many (although not infinite) possible outcomes, variability, is comprised of all possible outcomes, realized or not, and is an abstraction rather than an observation. Importantly, like variation, variability has limits and is subject to change. What components structure variability, and what are the underlying causes of its change? We argue in this paper that the basic processes of development and their interaction with the environment simultaneously provide the raw material for variability as well as limit it.

A way to conceptualize the relationship between developmental processes, variability and variation is outlined in Fig. 1. Here we show that the basic developmental processes involved in the ontogeny of the standard phenotype (the composite of which represent variation) are the same processes that modulate variability. That is, variability represents a range of potential outcomes (realized or not), and the available developmental processes and their interactions limit that variability. Because the full range of variability is never realized due to lethality and chance, phenotypic variation is a *subset* of all possible phenotypes. Expressed phenotypic variation in a population is what is available to natural selection and this biases the developmental processes available for subsequent generations. Importantly, the developmental processes themselves are evolvable, maintaining enough complex interactions to sustain and produce an ever-changing landscape of variability (Jernvall and Jung 2000; Weiss and Fullerton 2000; Salazar-Ciudad et al. 2001a, b). Developmental processes are dynamic and changing in ways that are both fairly well understood (e.g., through mutation, recombination and tissue interactions) and in ways that we are just beginning to explore (e.g., somatic mosaicism and variable expressivity). The dynamic nature of developmental processes ensures continued variability and allows the relationship depicted in Fig. 1 to continue indefinitely. Thus, variability is an emergent by-product of the developmental processes (including their interactions) involved in the ontogeny of the standard phenotype.

While our depiction of variability and its relationship with developmental processes in Fig. 1 includes all levels

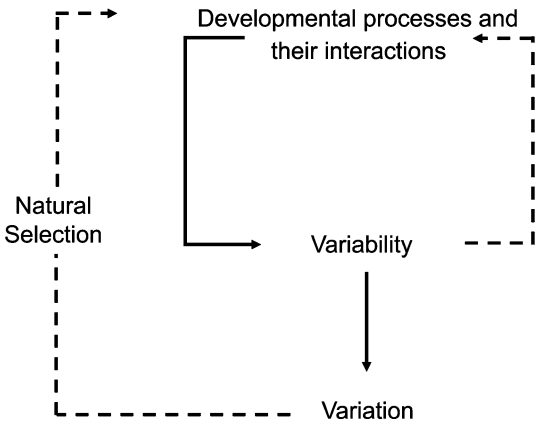


**Fig. 1** Schematic of the relationship and interactions between developmental processes, variability and variation. Starting at the top of this depiction, developmental processes structure variability, which in turn modulates the ontogeny of the organism determining the resultant phenotypic variation. The *solid arrows* indicate a direct effect, whereas the dashed arrows indicate an effect created through a feedback loop. Developmental processes include all elements involved in organismal development including molecular, genetic, cellular, individual, population and environmental factors

of organismal variability, our focus in this paper is phenotypic variability leaving the discussion of related topics such as genetic variability to others (for an excellent discussion see Houle 1998). However, we approach this topic from the perspective of hierarchical systems theory. That is, the phenotype cannot be understood in isolation without knowledge of the parts that make it up and their interactions. It is part of an open system that is comprised of several hierarchies and interactions within and among these hierarchies, from the molecular and cellular level, to the tissue and organ level, to the population and environmental levels. While it is necessary to understand the parts that comprise the phenotype to understand the mechanistic underpinnings of phenotypic variability, the whole is more than just the sum of its parts (Weiss 1971; Jacob 1977). The expression of a phenotype is the result of the complexity of individual factors, their arrangement within hierarchies, and their interactions, and consequently cannot be reduced to the function of a specific mechanism at the genetic level. In fact, the complexity of the potential interactions responsible for the production of the phenotype requires that the search for potential causal mechanisms include a top (phenotype)-down as well as a bottom (molecular)-up approach.

In order to study phenotypic variability effectively, it is necessary to differentiate between theory, measurement and underlying processes. Here we attempt to organize current understanding of variability accordingly. We have limited our discussion of phenotypic variability to three components: canalization, developmental stability and morphological integration (Hallgrímsson et al. 2002). We recognize that there are several other components of variability (e.g., phenotypic plasticity, heterochrony, and heterotopy) but we focus on these three as they have received the greatest attention in the recent literature. Importantly, they are rarely discussed as mutually interacting.

We first discuss the concept or *property* of variability in terms of canalization, developmental stability and morphological integration. We then focus on *patterns* of observable phenotypic variation that are quantified and used as surrogates for direct measurements of variability. Such patterns are thought to carry signatures of those processes that influence phenotypic variability and can be used to pinpoint times or locations that may be more or less susceptible to variation. The patterns of variation used to characterize each of the components of variability are described, and the assumptions and potential pitfalls of using patterns as a statistical measurement of variability are discussed. Finally, we explore sources of variability by examining some of the underlying *processes* that both create and constrain variation at different hierarchical levels. We have chosen to use jaw development as an example to demonstrate the importance of variability and the complexity of its generation. Our treatment of

variability is limited by what is known and we look forward to future discoveries that add to our assessment.

The aim of the current paper is fourfold: (1) to emphasize the importance of the phenotype, and particularly phenotypic variability in modern biology; (2) to provide a review of canalization, developmental stability and morphological integration highlighting the importance of differentiating between theory, measurement and causation; (3) to demonstrate that phenotypic variability is governed by the same developmental processes that create the standard phenotype and; (4) to advocate a re-organization of what is known about variability and how it is studied in terms of a systems-biology perspective of hierarchies and interactions. The irony of advocating a systems view of variability while constraining our discussion to phenotypic variability (and only three of its components at that!) is not lost on the authors. We provide this treatment of phenotypic variability as a template for other studies of variability at the phenotypic or other levels of analysis.

## Properties of Variability

Phenotypic variability is partly composed of three interrelated components: canalization, developmental stability and morphological integration (Hallgrímsson et al. 2002). Each of these components contributes to variability by either limiting phenotypic variation when exposed to perturbations or by biasing the direction of variation. Variability is described as an organismal property (Wagner and Altenberg 1996) and similarly we consider canalization, developmental stability and morphological integration as properties of an organism. Use of these terms is inconsistent in the literature, creating confusion of the underlying concepts for anyone new to the field, and can often create confusion in the interpretation of empirical results. This confusion is partly due to the fact that these components are often defined according to how they are measured and not as theoretical phenomena. Clarity of the common definitions of these properties and their conceptual underpinnings is of extreme importance for ensuring that the questions being asked, and the interpretation of results, are properly understood and conveyed. When these properties are viewed from a theoretical perspective, their boundaries become blurred revealing the interrelationships of these concepts and their role in the greater concept of phenotypic variability.

### Canalization

Canalization refers to the property of an organism that ensures similar phenotypic expression within a group by buffering development against both environmental and genetic perturbations (Zakharov 1992). Canalization does not eliminate phenotypic variation; rather it maintains a certain degree of phenotypic variation under differing genetic and environmental conditions. The concept of canalization was originally developed by both Waddington (1942, 1957) and Schmalhausen (Published in Russian 1938, translated to English 1949). Waddington and Schmalhausen arrived at the idea independently but there are slight differences in their descriptions of canalization.

Waddington approached the topic of canalization from a developmental perspective and described it as the buffering of development against environmental and genetic perturbations. His concept was based on two main observations: (1) the cells and tissues that comprise an organism do not develop in a gradation of forms, rather they form discrete types and (2) processes of development often recover from perturbations and return to their preferred endpoint resulting in normal development. Schmalhausen, referred to canalization as autonomization, which is based on how the norm of reaction interacts with stabilizing selection. The norm of reaction is the range of expected phenotypes under a given set of environmental and genetic conditions. In his view, regulating mechanisms that modulate the norm of reaction under different genetic and environmental conditions are developed under the slow process of stabilizing selection. He argued that organisms with the ability to withstand environmental perturbations while simultaneously responding adaptively to changes in the environment would be favoured by natural selection. That is, there is a range of optimal phenotypes that arise as a balance between buffering development against perturbations, and adaptive change (within the norm of reaction) in response to such perturbations or fluctuations. In this sense, Schmalhausen's view of canalization differs quite drastically from that of Waddington. He included phenotypic plasticity (adaptive change of the phenotype in reaction to different environmental conditions) in his concept of canalization, rather than considering it an opposing process.

Some researchers have differentiated between genetic and environmental canalization (Wagner et al. 1997). Genetic canalization refers to the robustness of the phenotype to the effects of mutation, whereas environmental canalization refers to the robustness of the phenotype in response to environmental perturbations (such as extreme temperatures or exposure to toxins). In fact, Wagner et al. (1997) define canalization exclusively in terms of genetic canalization, and consider what has been defined here as environmental canalization, as the related property developmental stability (see Table 1). While the division between the environmental factors associated with environmental canalization and those associated with developmental stability is somewhat arbitrary, the methods of detecting these two properties differ greatly, and

**Table 1** List of the common definitions used for canalization and developmental stability noting the differences in the types of perturbations buffered by each

| Phenomenon | Properties | References |
|---|---|---|
| Canalization | Buffers development against macro-environmental and genetic perturbations | Waddington (1942, 1957), Clarke (1998) |
| Developmental stability | Buffers development against micro-environmental perturbations | Waddington (1942, 1957), Clarke (1998) |
| Genetic canalization | Buffers development against genetic perturbations | Wagner et al. (1997) |
| Environmental canalization | Buffers development against both micro- and macro-environmental perturbations | Wagner et al. (1997) |

therefore, we feel it is best to acknowledge the difference between environmental canalization and developmental stability here.

Canalization is generally considered a property of an organism that limits phenotypic variation. Essentially, canalization describes the suite of processes that buffer the phenotypic expression of variation created by differing genetic and environmental conditions. In this way, canalization can be viewed as a constraint on variability by limiting the scope of resultant variation expressed in the phenotype. However, while canalization buffers the phenotypic expression of variation, it can also allow genetic variation to accumulate undetected. Therefore, canalization can also facilitate variability in situations where a system is exposed to a perturbation severe enough to surpass a given threshold, thereby revealing 'cryptic' genetic variation that has not previously been exposed to selection (Waddington 1957; Rutherford 2000). This apparent contradiction highlights how the definition of canalization as a property of an organism that limits phenotypic variation, is dependent on a consistent genetic and environmental background.

## Developmental Stability

Developmental stability is the property of an organism that buffers variation of micro-environmental origin, or environmental factors within the organism (Clarke 1998). That is, developmental stability ensures consistent phenotypic expression within individuals given a specific genotype and macro-environment (environment outside the organism). In this way, canalization and developmental stability are similar properties in that they both act to limit phenotypic variation; their differences lie in the types of variation they buffer. Waddington (1957) originally described the variation buffered by developmental stability as developmental noise, which was vaguely described as thermodynamic noise at some unknown process level (Reeve and Robertson 1953). Imprecision in molecular and cellular level processes lead to stochastic noise-like variation that can sometimes be detected at the phenotypic level. This definition of developmental stability can be problematic as it assumes a clear distinction between micro-environmental perturbations such as that caused by developmental noise, and macro-environmental perturbations such as those thought to affect canalization. Environmental perturbations assumed to affect canalization are thought to arise outside of an individual, whereas developmental noise arises within an individual. When one starts to parse out the potential factors contributing to phenotypic variation, both within and among individuals, the distinction between micro and macro environmental influences becomes vague. While the precise differentiation between the variation that is buffered by developmental stability and canalization is controversial, the concept that developmental stability buffers random variation within the individual and canalization buffers genetic and macro-environmental variation is a core distinction between the two. The importance of this distinction will become apparent in the next section when we discuss how to quantify developmental stability as it relies on a random, noise-like pattern of resultant phenotypic variation (Van Valen 1962).

Developmental stability is a property of an organism that reflects the result of two opposing forces: the noise-like variation that perturbs the system, and the buffering mechanisms that limit the phenotypic effects of this variation (Lens et al. 2002, Willmore and Hallgrímsson 2005; Breuker et al. 2006; Van Dongen 2006). Therefore, developmental stability is a relative property. If two populations experience the same degree of developmental noise, then the population with the least amount of within-individual variation is considered to be the more developmentally stable. Issues arise when the level of developmental noise that a population experiences is unknown; lower within-individual variation could result from more efficient buffering mechanisms and hence represent developmental stability, or it could result from low levels of developmental noise.

## Morphological Integration

Morphological integration refers to the phenotypic interdependence of two or more structures that reflects common

function and/or development. This inter-dependence between structures is thought to enhance the overall stability of the organism. The theoretical aspect of integration was well known and incorporated into biology by holists such as van der Klaauw and Schmalhausen. However, the concept was more widely recognized after Olson and Miller's (1958) publication, where they argued that traits that share functional and developmental influences would evolve as integrated units. As pointed out by Chernoff and Magwene (1999), Olson and Miller never formally defined morphological integration. In fact, the theoretical concept of integration (as introduced by Olson and Miller and as defined by most modern biologists) was born out of its statistical measurement. The basic premise of morphological integration describes how different structures that share some factor (e.g., function, developmental precursor and position) will show some sort of relationship in their phenotypic expression. Morphological integration remained relatively unexplored until Cheverud (e.g., 1982, 1984, 1995) conducted a series of studies on the patterns of integration in primate crania, strengthening the work of Olson and Miller (1958) and effectively reviving the concept.

Cheverud (1996) has also expanded the idea of morphological integration beyond the level of the individual. He described integration as occurring at three distinct levels: individual, genetic and evolutionary. The integration proposed by Olson and Miller (1958), or integration between functionally and developmentally related structures occurs at the individual level. Integration created by pleiotropy or linkage disequilibrium constitutes genetic integration, while the coordinated evolution of traits is considered evolutionary integration. These levels of integration do not act in isolation; in fact Cheverud (1996) argued that integration at the individual level causes genetic integration, and that these genetically integrated structures will be inherited together, and therefore, are selected together and result in evolutionary integration.

Related to the concept of morphological integration is modularity. A module is a character, or a set of characters that are more tightly integrated internally than they are with other characters (Wagner 1996). For example, the skull likely forms one module and the limbs form a separate module in a quadruped. The structures that comprise the cranium will show a higher degree of inter-relatedness with other structures of the head than they will with limb traits, and the reverse is true for the limbs. Cranial and limb modules may form due to the physical adjacency of the structures, shared developmental origins, or shared functional demands (Young and Hallgrímsson 2005).

Essential to the concepts of modularity and integration is the hierarchical structure of organisms (Wagner and Altenberg 1996; Chernoff and Magwene 1999; Gass and Bolker 2002). Several levels of integration can occur; the entire organism is a functional whole and is therefore integrated, but within that organism subunits may be integrated through functional, developmental and genetic relationships. Figure 2 demonstrates this concept using a mouse model. In Fig. 2a the entire mouse is integrated by the allometric effect of size; however, the limbs of the mouse also form a module, and within that module the fore- and hind limbs form separate modules. This division of traits is somewhat intuitive; however, Fig. 2b illustrates how integration can be structured by less phenotypically obvious factors. Here, we show how the mouse skull can be divided into modules based on developmental precursors of particular cranial bones. Cranial bones of neural crest origin form one module, whereas cranial bones derived from mesoderm form a separate module and both of these

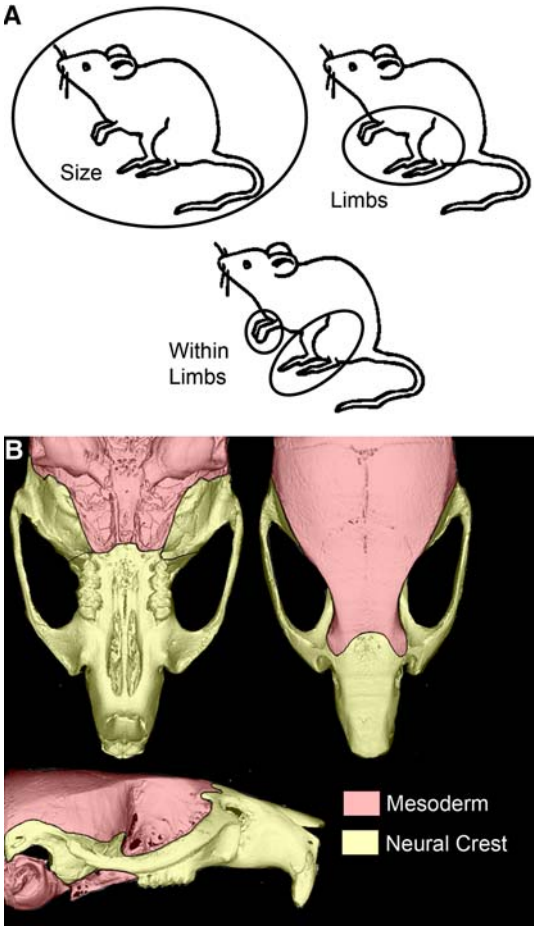


**Fig. 2** Illustration of the modularity concept. (**A**) Modularity as seen in the whole mouse. Modules occur at different hierarchical levels with overall integration influenced by size and then subsets of structures forming the limbs module and finally the within-limb module. (**B**) Modularity within the mouse skull. The cranial bones can be divided into modules based on their developmental precursors. Bones of neural crest origin form one module while bones derived from mesoderm form a separate module

modules join to form the functioning skull. Also implicit in the concept of integration is that the same traits can be part of several different modules; that is, modules are intersecting. For instance, a trait could belong to one module due to shared functional demands, and could be part of a separate module as a result of pleiotropy. We argue that integration and modularity are two sides of the same coin. Modularity is simply nested integration.

Strong integration of many traits is argued to constrain variability because a change in one part of a highly integrated structure will rarely be advantageous to the other parts of that structure or the structure as a whole. Therefore, variability in individual parts will be selected against (Wagner and Altenberg 1996). However, the dissociability of parts into more discrete modules should increase the likelihood that favourable changes in one module will not have an effect on a separate module and will therefore, be more likely to persist within a population. Therefore, large-scale integration constrains variability whereas integration of subsets of organismal structures (modules) can facilitate variability.

Canalization, developmental stability and morphological integration have unique and overlapping relationships with variability. While it is important to have a clear understanding of what these properties represent, their use is strengthened if they can be measured correctly. Below, we outline the more common measurable surrogates of canalization, developmental stability and morphological integration.

## Measuring Variability via Patterns of Variation

Due to the difficulty of direct measurement, most variability studies have employed the indirect approach of measuring the observable patterns of phenotypic variation. As mentioned earlier, phenotypic variation only captures one potential outcome and does not accurately represent the range of possible variation (i.e., variability). This method of measurement also does not allow us to determine the underlying cause of the variation. Because of these limitations, most studies of phenotypic variation and variability have involved perturbing development either through mutation or environmental influence (such as exposure to a toxin or extreme temperature). Using a variety of experimental protocols, researchers are often able to determine specific genotypes that are most susceptible to particular perturbations, specific traits that are most affected, and if using an ontogenetic sample, specific developmental stages most susceptible to the perturbation. In this way, the phenotype and the patterns of phenotypic variation can be used to elucidate potential developmental processes that might be responsible for phenotypic variability under natural conditions.

### Canalization-Measurement Through Patterns of Variation

Corresponding traits in different individuals develop under different environmental and genetic conditions, and therefore, differences in the amount of among-individual variation for a trait will indicate differences in ability to canalize development against genetic and environmental stresses. Thus, among-individual phenotypic variation is the most common measure of canalization. Canalization of a trait is generally inferred by a reduction of the observable phenotypic variance of that trait. Figure 3 illustrates the variance of four populations for two traits. The points within each population represent individual values for mass ($x$-axis) and height ($y$-axis). Each of the populations shows various ranges of values along both axes. A smaller range indicates a smaller variance and the population with the smaller range is considered more highly canalized under the assumption of similar environmental variation. Population A has a narrower range of values along the x-axis than either population B or D. Thus, population A would be considered more canalized for mass than population B or D. Canalization is measured by phenotypic variance; the greater the variance, the less stringent the canalization.

A major drawback of this method of measurement is that canalization can only be determined by comparison with some reference state (Rutherford 2000). The ancestral state of canalization for a structure is unknown; therefore, we rely on comparisons between populations based on hypotheses of causal factors thought to affect canalization. This approach was used by Waddington to demonstrate the existence of canalization based on the hypothesis that animals with aberrant phenotypes resulting from exposure to some environmental stress would show greater variation

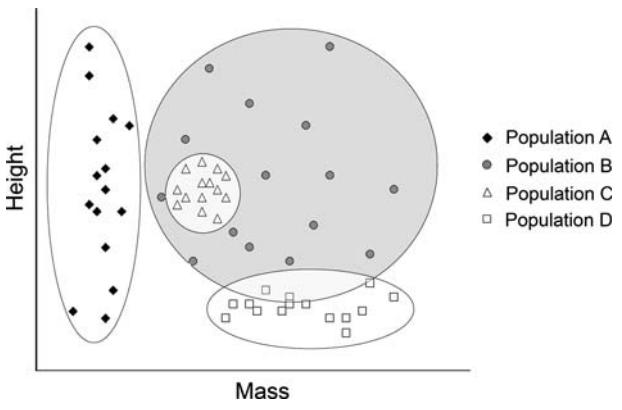


**Fig. 3** Hypothetical scatterplot of phenotypic variances of individuals for both mass and height for four separate populations. The larger the scatter of points, the greater the variance indicating reduced canalization. Note that some populations have low variance for one trait while expressing high variance for the other

than wild-type animals. That is, aberrant individuals would be less canalized than wild-types. Specifically, he compared *Drosophila* phenocopies with wild-type flies. A phenocopy is an aberrant phenotype that resembles the phenotype of a known mutation but is caused by exposure to environmental stress or change rather than mutation. Waddington subjected flies to a heat stress during pupal development that resulted in flies that resembled *cross-veinless* mutants. Of interest to Waddington (1957) was not the mean vein phenotypic change in the flies subjected to the heat stress, but the variable phenotypic expression in response to this stress in the treatment group. Waddington (1957) suggested that the relatively invariant vein morphology in the wild-type flies was evidence that the wild-type is canalized, and only when a stress is severe enough to surpass some threshold will this canalization break down and result in increased phenotypic variation.

Other early studies used a similar approach, comparing wild-type animals with mutants, again under the assumption that the wild-type is more canalized. Rendel (1959) compared variance in scutellar bristle number between wild-type *Drosophila melanogaster* with the sex-linked mutant *scute*; a mutation known to decrease bristle number, particularly on the scutellum. In wild-type flies, the number and pattern of scutellar bristles is effectively invariant. However, marked variance in scutellar bristle number was observed in *scute* mutants (Rendel 1959). Similar studies were conducted by Dunn and Fraser (1958, 1959) using the sex-linked mutation *Tabby* (*Ta*) in mice. The number of secondary vibrissae was shown to decrease in *Tabby* mutants. As with the bristle number in *scute* fly mutants, the interesting trend found in this study was the increased variance around this mean vibrissa number in the mutant mice (Dunn and Fraser 1958, 1959). The results of these early studies were interpreted as demonstrating the canalization of the wild-type and how this canalization can be disrupted if a perturbation (either environmental or genetic) is great enough to surpass some threshold resulting in the phenotypic expression of previously hidden variation. This interpretation has become deeply entrenched in the literature, and canalization is inferred if a trait displays greater variance in the mutant than the wild-type background (Rutherford 2000).

However, while the majority of studies have demonstrated this association between mutation and environmental stress with phenotypic variation, exceptions exist. Not all mutations result in increased phenotypic variance, which is consistent with the premise that canalization is causally heterogeneous (Scharloo 1991). For example, increased haltere variation was *not* found in *Ultrabithorax* (*Ubx*) *Drosophila* mutants (Gibson and van Helden 1997). This result was particularly surprising, as *Ubx* is a homeotic regulatory gene and is expected to have a major impact on development and its effect on haltere development is therefore, assumed to be highly canalized. Exceptions such as this one underscore another challenge to measuring canalization; canalization appears to be trait specific; however, this observation may be due to our trait-by-trait measurement of canalization. If our suggestion that variability, and therefore, canalization, are due to the interaction of developmental processes, than we would expect to see different levels of among-individual variance for different traits depending on the developmental processes associated with each trait. For example, if two traits share a developmental process that is compromised, then we expect to see increased variance in both traits. If the compromised developmental process was only involved in the development of one of the traits, we would expect to see an increase in variance in that trait only. Due to the complexity of developmental interactions we often do not know what traits are affected by a perturbation and therefore, it is difficult to determine traits most suitable for measurement in a study of canalization. In general, most researchers deal with this obstacle by measuring multiple traits and often multiple structures within each trait.

A major aspect of canalization theory developed by Waddington is that development can recover from perturbations through epigenetic compensatory mechanisms (Waddington 1957, 1975). That is, the developmental trajectories themselves are canalized. Waddington referred to these canalized trajectories as chreods (Waddington 1957, 1975). While most studies of canalization have compared a wild-type population with a population of either mutants or phenocopies, Waddington originally argued that canalization should be measured throughout ontogeny. There are a handful of studies that have measured canalization using an ontogenetic approach (Foote and Cowie 1988; Zelditch et al. 1993, 2004; Willmore et al. 2006; Young 2006); however, it is an area of study needing more work. Experimental designs that track canalization through development are an obvious starting point for trying to elucidate potential specific developmental mechanisms that contribute to canalization.

## Developmental Stability—Measurement Through Patterns of Variation

Developmental stability ensures consistent phenotypic expression under the same genetic and environmental conditions, whereas developmental instability represents inconsistencies in phenotypic expression. Therefore, within-individual variation or fluctuating asymmetry (FA) is used to measure developmental instability. Quantitative differences between repeated structures such as vibrissae on mice, or bristles on *Drosophila* within the organism can

be used to determine the within-individual variance. More commonly, the difference in variance between left and right sides of bilaterally symmetric traits is used to measure the variance within individuals. The distribution of the random deviations from perfect symmetry of bilaterally symmetric traits or FA was first defined by Ludwig (1932 as cited by Van Valen 1962). Developmental instability is thought to arise from developmental noise, and the random yet normal nature of left–right differences around a mean of zero found in FA distributions is what would be expected of a noise-like pattern (Palmer and Strobeck 1992). However, it has recently been suggested that a Gaussian distribution may not be the most appropriate (Graham et al. 2003; Klingenberg 2003, Babbitt et al. 2006). The developmental errors associated with FA are assumed to be additive under models that use a normal distribution. However, recent work indicates that non-linear developmental processes might be involved in the generation of FA (Klingenberg 2003). Graham and colleagues (2003) suggest that a log-normal or gamma distribution might be a more accurate representation of these non-linear developmental errors thought to be associated with FA. Currently, there is no consensus on this issue and it is therefore important to exercise care when using a Gaussian distribution to represent FA (Van Dongen 2006).

The association between FA, developmental noise and developmental instability assumes that symmetry is optimal and that any deviation from symmetry constitutes some degree of developmental instability (Palmer and Strobeck 1986). The use of variance between bilaterally symmetric traits has been considered an appropriate measure of developmental stability, as both sides presumably share the same genetic background and environmental effects, and therefore, differences between sides are created through perturbations within the organism (Reeve 1960; Palmer and Strobeck 1986; Clarke 1998). The relationship between FA and developmental stability is similar to that between among-individual variance and canalization. High levels of FA indicate low levels of developmental stability (Van Valen 1962; Palmer and Strobeck 1986).

It is important to differentiate between different types of asymmetry that commonly occur within an organism, as most researchers posit that only FA can be used to measure developmental instability (Van Valen 1962; Palmer and Strobeck 1986, 1992, 2003; Palmer 1996; but see Graham et al. 1993). Directional asymmetry (DA) is normally distributed, but the mean is biased either to the left or right of zero, and antisymmetry is distinguished by a bimodal distribution of deviations, or more subtly by a platykurtotic distribution (See Fig. 4). DA is quite common, the asymmetry of the human heart and brain are examples of this phenomenon. Antisymmetry is much less common; an extreme example is demonstrated by the male fiddler crab.

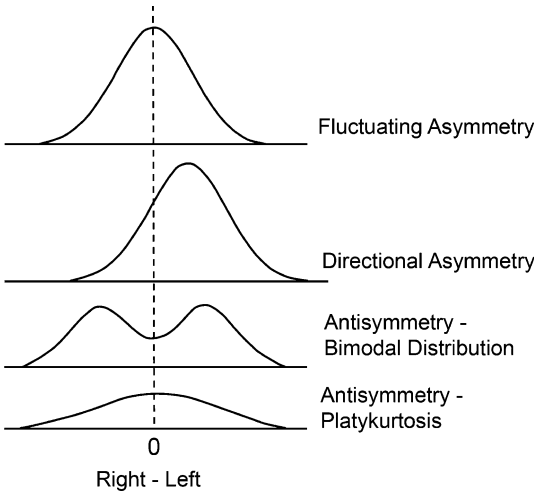


**Fig. 4** Statistical distributions that describe different types of asymmetry as defined by Van Valen (1962). Modified from Palmer (1996)

The right and left claws are initially the same size. However, the first claw to be broken re-develops to become conspicuously smaller than that of the other side. Therefore, fiddler crabs always have one claw larger than the other, but the side that is enlarged is completely random at the population level (Palmer and Strobeck 1992; Palmer 1996). Palmer and Strobeck (1986, 2003) advocate that any study of FA should account for potential DA and antisymmetry as they argue that their underlying mechanisms are genetically controlled and do not represent developmental instability. However, it is not always possible to rule out antisymmetry and DA, partially because the statistical power of techniques used to separate these different types of asymmetries is low, and partly due to what Nijhout and Davidowitz (2003) call environmentally induced asymmetry. Handedness in humans is an example of environmentally induced asymmetry as is any immobile period during development whereby static orientation may create subtle environmental differences between right and left sides that do not reflect developmental instability (Van Dongen 2006).

Organisms exposed to some environmental stress such as a toxin have been found to have elevated levels of FA, which has led to the controversial suggestion that FA is a reliable indicator of stress in natural populations (Graham et al. 1993; Manning and Chamberlain 1994; Leung et al. 2000). The use of FA as an indicator of stress was favoured over more traditional measures, such as fitness components, because FA would be cheaper and simpler to measure. However, as is strongly argued by Palmer and Strobeck (1986, 2003), measuring FA properly is far from simple and the large sample sizes required for its measurement can make the use of FA expensive and labour intensive. Due to the challenges associated with measuring

FA, its use as a bioindicator has been ardently criticized (Bjorksten et al. 2000; Lens et al. 2002).

One such challenge is that the signal-to-noise ratio in measures of FA is extremely low, with FA values typically on the order of 1% of trait size (Lens et al. 2002). This subtlety of FA makes it difficult to detect, and therefore, the degree of FA is probably often underestimated. Additionally, measurement error can have a profound effect on FA (Palmer and Strobeck 2003). Deviations from symmetry are of about the same magnitude as differences created by measurement error (Palmer 1996). Moreover, like FA, measurement error is random and normally distributed about a mean of zero, thereby making it indistinguishable from FA (Palmer and Strobeck 1986; Palmer 1996). Palmer and Strobeck (1986, 2003) outlined an index of FA that can partition out the effects of measurement error, but it requires repeat measures of FA for at least a subset of the sample and can become labour intensive. Finally, the correlation between FA and stress has been found to be highly stress and trait specific (Clarke 1998; Bjorksten et al. 2000; Lens et al. 2002; Rasmuson 2002). Therefore, FA would only be a reliable indicator of stress for certain traits, and those traits are not known in advance. Likewise, not all stressors would have the same magnitude of effect on FA.

An issue of active study and controversy is the relationship between developmental stability and canalization. At the theoretical level their close association is clear; both limit the expression of phenotypic variation. Therefore, the only currently useful distinction between these two phenomena arises due to how they are measured; we use among-individual variance to estimate the inverse of canalization and FA to estimate developmental instability. The general idea has been that different underlying processes govern developmental stability and canalization. However, the results from empirical tests (studied using patterns of phenotypic variation) have been contradictory, some studies indicating that developmental stability and canalization are separate phenomena governed by separate mechanisms (Waddington 1957; Rutherford and Lindquist 1998; Debat et al. 2000; Milton et al. 2003; Réale and Roff 2003) while others have indicated that they represent similar phenomena that at least partly share underlying developmental processes (Clarke 1993, 1998; Klingenberg and McIntyre 1998; Woods et al. 1999; Hallgrímsson et al. 2002; Santos et al. 2005; Willmore et al. 2005; Breuker et al. 2006; Debat et al. 2006). A recent study by Breuker and colleagues (2006) compares among-individual variation with FA in the wings of 115 different strains of *Drosophila melanogaster*. Each of these strains differs by only one deficiency on an otherwise isogenic background, thereby, controlling the genetic variation between populations. The authors found a strong relationship between FA and among-individual variation for wing shape, but not for wing size. They suggest that the relationship between canalization and developmental stability is not universal and will potentially differ between traits and populations. We suggest that the relationship between canalization and developmental stability is based on how the basic processes of development are distributed among the traits of interest and to what degree (and when in development) these processes are exposed to perturbations, be they genetic, micro- or macro-environmental in origin. Clearly, more empirical studies are needed to adequately address this issue. The study by Breuker et al. (2006) is an elegant example of a design that enabled identification of specific underlying developmental mechanisms for the association between canalization and developmental stability. Studies of similar controlled design should prove to be fruitful.

## Morphological Integration—Measurement Through Patterns of Variation

Morphological integration is estimated by the level of covariation or correlation among structures. Strong covariation between traits indicates a high degree of integration and is thought to be caused by shared developmental, functional or genetic influences. Olson and Miller (1951) were the first to suggest using correlation coefficients to measure the level of integration, and later developed a theoretical framework for detecting covariation patterns among morphological traits (Olson and Miller 1958). Most recent studies of morphological integration have followed Olson and Miller's later approach of developing biological hypotheses that suggest how suites of traits might covary using both a priori and a posteriori hypotheses.

Berg (1959) conducted one of the first studies outside of the work by Olson and Miller to use correlation coefficients to determine morphological integration. His study looked at the integration of the reproductive and vegetative organs of several plant species. Olson and Miller (1958) are not cited in his paper, and instead it appears that his work arises from a concept similar to morphological integration that was independently advanced in Russia called correlation pleiades. The concept of correlation pleiades is based on the presence of correlation between some quantitative characters and no such correlation between other quantitative characters (Berg 1959). Gould and Garwood (1969) conducted another early landmark study of morphological integration on the teeth of the rodent *Oryzomys* and the insectivore *Nesophontes*. Following Olson and Miller, they used correlation coefficients to determine levels of morphological integration, but they also introduced factor analysis as a method of estimating morphological integration.

While other researchers had applied Olson and Miller's concept of morphological integration and their method of measurement, it is a series of studies conducted by Cheverud (Cheverud 1982, 1988, 1995; Marroig and Cheverud 2001) that is generally considered the most influential in reviving the concept of morphological integration. These studies tested Moss' (Moss and Young 1960; Moss 1971) functional matrix hypothesis using primate crania (Cheverud 1982, 1995; Marroig and Cheverud 2001). In these studies, cranial traits that shared functional demands or developmental histories were predicted to show strong covariation. The crania of Rhesus macaques (Cheverud 1982), saddle-back tamarins (Cheverud 1995) and New World monkeys (Marroig and Cheverud 2001) all demonstrate correlation patterns that follow the predictions of the functional matrix model (Moss and Young 1960; Moss 1971). These predictions were tested by comparing the correlation matrix derived from the empirical cranial data with a theoretical matrix based on the hypothesis of integration. Matrix correlations are commonly used to compare the correspondence of a priori biological hypotheses with empirical patterns of covariation among traits. Permutation tests are used to evaluate a null hypothesis that the association between two matrices is not different than what would be expected by random chance. A high correlation between matrices indicates that the patterns of correlation found in the skull data closely match the hypothetical matrix, and therefore, implies that the variation is structured by predicted functional and developmental influences. All of the preceding studies used an a priori approach to determine morphological integration. A priori methods allow for specific tests of biological hypotheses and generally yield interpretable results. In these approaches, if the null hypothesis investigated (that the empirical and the hypothetical matrices are similar) is rejected, we know the matrices are dissimilar, but we do not have any information regarding the specifics of how the two matrices differ. If the null hypothesis investigated is not rejected, the hypothetical reasons for the structured variation are accepted. In short, alternate patterns of covariance that could also describe the structural organization are not considered.

To improve the amount of information that can be garnered by comparing matrices, Cole and Lele, (2002) proposed an alternative approach based on the statistical analysis of the differences between the elements of two correlation (or covariance) matrices. In this approach, correlation matrices estimated for each of the two samples are used to estimate a correlation-difference matrix. A correlation-difference matrix is calculated by subtracting the elements of one matrix from the corresponding elements of the other matrix. If the matrices are the same (the null hypothesis), all of these differences are expected to be zero. When one or more elements of the matrix does not equal zero, the correlation differences can be explored to learn more about these differences. A bootstrap approach for testing the null hypothesis of similarity in correlation values was also developed (Cole 2002). The bootstrap approach enables localization of the differences in matrices to particular variables aiding in the biological interpretation of the matrix comparison (see also Richtsmeier et al. 2006).

Another example of an a posteriori approach for detecting morphological integration is cluster analysis (Hallgrímsson et al. 2002; Willmore et al. 2006). This method is used to parse out groups of traits that show a stronger association with each other than with traits from other groups, and these associations are determined without input from the researcher. For example, Willmore et al. (2006) used Ward's cluster analysis to determine if certain craniofacial traits in mice formed modules and to see if these modules changed with age. The clustering is based on statistical association only; none of the traits were coded to indicate any potential relationships based on some biological factor such as function or developmental origin. The resultant modules were consistent between age groups, but did not reveal any obvious biological interpretation. A posteriori methods reveal all possible covariance patterns but these patterns can be extremely difficult to interpret in biological terms. While the premise that patterns of covariance are an appropriate measure of morphological integration is widely accepted, no such consensus exists for the specific methods of determining covariance patterns or the biological interpretation of these patterns.

This brief review of canalization, developmental stability and morphological integration as properties of an organism and the patterns of phenotypic variation used to measure each, reveals two important points. First, our definitions of these components of variability are heavily biased and limited by our methods of measurement and what we are measuring is variation not variability. In fact, our measurement systems impose the artificial separation of these components, which leads us to our second point. That is, the importance of recognizing that each of these components represents overlapping aspects of the same phenomena, namely variability. All three properties limit variability. Developmental stability reflects localized buffering of developmental processes and canalization reflects more global buffering of developmental processes both of which limit variability. Morphological integration ties different structures together including their respective canalization and developmental stability properties. Therefore, morphological integration reflects the organization (often hierarchical in nature) of these canalized and developmentally stable developmental processes. Figure 5 depicts our view of how these properties relate to

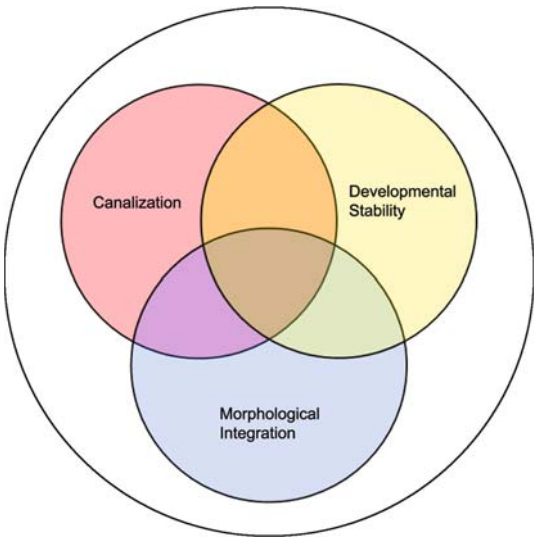


**Fig. 5** Depiction of how the properties of canalization, developmental stability and morphological integration overlap with one another. The area within the outer *black circle*, in which the three components are found, represents phenotypic variability. The placement of these three properties within area outlined for variability demonstrates that they partially but do not completely describe phenotypic variability. Other components of variability such as heterotopy and heterochrony would likewise be placed within the outer circle and show patterns of overlap with the three components depicted here

variability and how they interact with each other. Each of these properties can work in isolation or in combination with one or both of the other two properties. The blurring of boundaries that define these properties suggests that their individual measurement and description is an exercise in futility. However, an understanding of the relationship of canalization, developmental stability and morphological integration with each other and with variability provides us with an opportunity to pinpoint developmental processes affecting variability as demonstrated by Klingenberg (Klingenberg and Zaklan 2000; Klingenberg 2003).

Klingenberg uses the relationship among these properties to differentiate between what he refers to as direct and parallel covariation (Klingenberg and Zaklan 2000; Klingenberg 2003). Direct covariation occurs when variation is transmitted to different traits within a developmental pathway jointly, thereby, potentially integrating parts within a module (Klingenberg and Zaklan 2000; Klingenberg 2003). Direct covariation can occur when a stimulus affects different traits in a similar manner because multiple traits develop from the same precursor or because one component within a system induces another component directly through cell signalling (Klingenberg 2003). Parallel covariation occurs when an environmental or genetic source of stress that arises outside of the developmental pathway affects multiple traits simultaneously (Klingenberg and Zaklan 2000; Klingenberg 2003). Klingenberg argues that the effects of direct and parallel covariation on integration can be separated if genetic and environmental influences are controlled. As noted earlier, because bilateral sides are subjected to the same environmental and genetic influences, covariance patterns of FA (essentially morphological integration of developmental instability) can be used to detect direct developmental interactions. Parallel covariation can be detected through covariance patterns of among-individual variance (morphological integration of the inverse of canalization). Klingenberg (2003) therefore uses comparisons of the patterns and organization of FA with the patterns and organization of among-individual variance to differentiate developmental processes affecting variability that are intrinsic to the developmental system and those that are external to it. This approach combines measures of integration, developmental stability and canalization and emphasized their interrelationships. However, this approach is still subject to all of the potential problems discussed previously for the independent measures of these properties.

## Processes Affecting Phenotypic Variability: A Mammalian Jaw Example

Our thesis is that there are no processes specific to the creation or modulation of phenotypic variability, but that phenotypic variability arises through those same developmental processes that produce the standard phenotype. Our definition of developmental processes is broad, including any factors involved in the development of the phenotype and therefore considers factors at the molecular, genetic, cellular, individual, environmental and population levels.

Using the mammalian jaw as an example we show how the fundamental processes of jaw development also function to structure phenotypic variability of the jaw. We have classified these processes into a hierarchy to clarify the extent of factors affecting phenotypic variability. However, this hierarchy is not tiered in terms of importance or order of influences. The significance of effects from each level is situation specific. Figure 6 depicts the multiple interactions that can occur between levels that influence phenotypic variability. While we describe them in terms of being hierarchical in relation to how we study them, their interactions are not structured hierarchically. In this section we have chosen examples of known developmental factors that affect the jaw, to demonstrate how basic processes of development are responsible for both generating and buffering phenotypic variability. Our examples are necessarily brief and we note that we are not attempting to give a full description of mammalian jaw development.

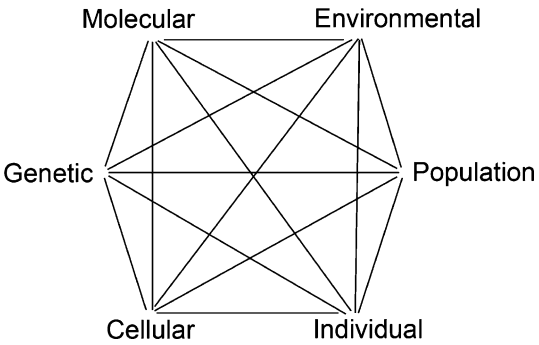


**Fig. 6** A schematic of the interactions that can occur between different hierarchical levels of organismal development. Note that while we describe these levels as hierarchical, their interactions are not structured hierarchically. All levels are equally free to interact with each other

## Processes at the Molecular Level

In this section we use the Dlx gene family to demonstrate how events occurring at the molecular level can both create and buffer variability and that it is the interplay of these effects that determines the resultant variability. The evolution of jaws appears to depend in part on the nested expression domains of Dlx gene family members along the proximodistal axis of the first branchial arch (Graham 2002; Carroll et al. 2005; Depew et al. 2005; Kuratani 2005; Stock 2005). In mammals there are six known members of the Dlx family. These homeobox genes arose as a single pair followed by cluster duplication and currently exist as three sets of linked pairs: Dlx1/2, Dlx5/6 and Dlx3/4 (Dlx-4 formerly Dlx-7) (Depew et al. 2002, 2005; Graham 2002; Weiss and Buchanan 2004).

The two major skeletal units that comprise the jaw are both first branchial arch (BA1) derivates. The maxilla is derived from the proximal end whereas the mandible is derived from the distal end of BA1 (Depew et al. 2002, 2005). Dlx1 and 2 are expressed throughout most of BA1 proximodistally, whereas the linked pairs Dlx5 and 6 and Dlx3 and 4 are more restrictively expressed distally (Depew et al. 2005). These linked pairs exhibit some redundancy, sharing regulatory elements and similar expression patterns (Depew et al. 2002, 2005). The effects of a disruption of one gene can be compensated by the similar function of a redundant gene, thus, increasing the fidelity of the system and potentially widening the range of conditions under which normal development can occur (Thomas 1993; Wagner 1999, 2005). For example, using loss-of-function experiments in mice it was found that several proximal BA1 structures were phenotypically abnormal in Dlx1 −/− mutants (Qiu et al. 1997; Depew et al. 2005). Similar phenotypic effects were found in homozygous Dlx2 −/− mutants however, the effects were not exactly the same (Qiu et al. 1997; Depew et al. 2005). These results suggest that both genes share similar expression patterns in BA1 derived structures but that they also have unique functions. Evidence for redundancy between the genes stems from an exacerbated phenotype for double loss-of-function mutants, indicating that each gene can at least partly compensate for the other (Qiu et al. 1997; Depew et al. 2005). In this way, redundancy between these linked pairs of Dlx genes can help to limit variability in jaw development.

While redundancy between Dlx1 and Dlx2 might help to limit variability of jaw development, the complexity of the code can potentially increase variability. Heterozygous mutant Dlx1 +/− and Dlx2 +/− mice show similar but slightly different phenotypes that are less severe than either homozygote mutant condition indicating that each allele contributes to the code (Depew et al. 2005). Furthermore, compound heterozygous mutants Dlx1/2 +/− had more severe phenotypic effects than either single heterozygous mutant (Dlx1 +/− or Dlx2 +/−) suggesting some synergistic effect between each allele of Dlx1 and Dlx2 (Depew et al. 2005). While the compound heterozygote phenotype was more severe than either single heterozygote, its effects were less severe than either homozygote loss-of-function mutation. These results illustrate the complexity of interactions between these linked genes and their effects on jaw development. Gene–gene interactions can simultaneously create a robust developmental environment by providing alternative developmental routes and can introduce variability through this same complexity. The interactions among the Dlx1 and Dlx2 alleles appear to be non-additive where one Dlx1 allele and one Dlx2 allele is not the functional equivalent of two Dlx1 alleles or two Dlx2 alleles (Depew et al. 2005). Non-additivity of allele function allows for a wide variety of effects caused by reduced gene expression. Loss-of-function experiments typically uncover more detectable differences between phenotypes but we suggest that more subtle differences in gene expression will also affect the phenotype in an equally complex manner.

Equally important to this complex system of interactions that can create and buffer variability at the molecular level are the many factors that can affect gene transcription, translation and expression. The molecular machinery of an organism is vulnerable to the effects of external environmental mutagens as well as metabolites in the internal environment. These environmental effects can damage DNA and disrupt the production of any genes associated with that DNA. Several errors can occur during DNA replication that might potentially lead to variability of development such as double-strand breaks, base-pair mismatches, heterologies, random excision of bases and nucleotides and large structural defects in the normal

Watson-Crick base pairing (Mills et al. 2003; Mohrenweiser et al. 2003). While these defects can arise through normal development, several repair mechanisms are also intrinsic to normal development. For instance, mismatch repair can correct for base-pair mismatches and heterologies, base excision can detect and then repair random excisions of bases and nucleotides as well as large structural defects (Mills et al. 2003; Mohrenweiser et al. 2003). Furthermore, a mechanism called damage recognition cell cycle delay response provides the system with a series of checkpoints that ensures that all the necessary events in the cell cycle are complete before continuing to the next cycle allowing the system enough time to recognize and repair damaged DNA before it is replicated (Kaufmann and Paules 1996; Mohrenweiser et al. 2003). While our treatment of these processes is only superficial, it is clear that the complexity of interactions between factors that can disrupt and factors that ensure proper DNA replication can both create and limit variability of the developmental system, including DNA associated with Dlx genes and jaw development.

We have shown how jaw development is dependent on the highly coordinated expression of Dlx genes in both space and time. The strength and timing of Dlx expression is dependent on its transcription. Gene transcription is often described as 'bursty' or stochastic (McAdams and Arkin 1999; Fiering et al. 2000; Blake et al. 2003). That is, transcription is regulated by a binary 'on' versus 'off' switching mechanism that relies on thresholds. This threshold mechanism of transcription can both exacerbate and limit variability of gene expression. A threshold system allows for a range of acceptable activity for normal transcription effectively hiding variability at this level. However, if this activity falls below the threshold, transcription will not occur and the protein will not be made increasing variability at this level and potentially affecting any developmental pathway dependent on that protein. Mechanisms exist that help to increase the number of transcripts available to help ensure that they surpass the threshold for normal protein production. These mechanisms work to increase the rate of transcription, lengthen the life of RNA transcripts through reduced activity of RNA degradation machinery, and increase the probability of gene transcription through enhancers (Fiering et al. 2000).

A major determinant of variability at this level is gene networking. The Dlx family does not work in isolation to produce functioning jaws. Rather these genes are part of a much larger network of genes whose interactions result in proper jaw development. In addition to the Dlx gene family, the Bmp, Fgf, Shh, Wnt, Retinoic Acid, Alx, Msx, Otx, Pax, Prx, Fox, Tbx, Gsc, Hox and Endothelin gene families are known to be involved in the development and patterning of the first branchial arch and jaws (Depew et al. 2005). Such a large network of interactions can both limit variability as well as create a cascade of variation that can potentially increase variability. Variability could be increased if there is a disruption of gene expression for a gene whose products largely control the expression of several other genes in major developmental pathways. A relatively minor malfunction of one gene could create a major developmental disruption due to its importance in the network. Conversely, genetic networks can limit the effects of single gene disruptions through what Wagner refers to as distributed robustness (Wagner 2005). Distributed robustness is difficult to measure, but describes how several different parts of a system contribute to a functioning whole, whereby changes or failure of one part can be compensated by the system (Wagner 2005). The robust nature of complex genetic networks has been demonstrated theoretically using several computer models (Wagner 1996; von Dassow et al. 2000; Siegal and Bergman 2002; Hermisson et al. 2003, Hermisson and Wagner 2004; Huerta-Sanchez and Durrett 2007). Of course, empirical evidence that supports these findings is necessary.

## Processes at the Genetic Level

The genotype-to-phenotype map is complicated by interactions at the genetic level that influence the development of phenotypic structures including the jaw. Below we show that basic properties of genetic architecture such as dominance, epistasis and pleiotropy influence mammalian jaw development and how they can potentially affect variability of the jaw.

Dominance refers to the interaction of two alleles at the same locus, and this interaction adds a non-additive component to the genetic variation (Falconer and Mackay 1996). Dominance is believed to reduce the detrimental effects of mutations, and therefore, influence variability (Bagheri and Wagner 2004). Mutational robustness is thought to result in part from dominance, as the system is able to respond more competently to far less gene product, and can therefore, compensate for large reductions in gene product caused by mutation. Using a quantitative trait loci (QTL) approach, dominance was found to affect size and shape variation in mouse mandibles (Klingenberg et al. 2001). While no specific genes were associated with this dominance effect, we can hypothesize how this dominance might affect mandibular variability. Dominance interactions may allow the system to tolerate reduced gene dosage and subsequent gene products required for normal mandibular development.

Another basic form of interaction at the genetic level is epistasis. Epistasis is the interaction between different loci,

whereby an allele at one locus can influence the phenotypic effects of an allele at a separate locus (Wagner et al. 1998). Therefore, the genetic context influences the affect of alleles creating a non-linear interaction between genes and the phenotype (Routman and Cheverud 1997). Schmalhausen (1949) and Waddington (1957) suggested that epistasis suppresses additive genetic and phenotypic variance by creating co-adapted gene complexes. Increased variance found in hybrids has often been attributed to a breakdown of these co-adapted gene complexes. However, epistatic gene interactions have been found to affect FA of centroid size in inbred mouse mandibles (Leamy et al. 2002). Therefore, while epistasis may potentially reduce phenotypic variability, it has also been associated with increased variability in mouse mandibles.

Pleiotropy refers to the interaction of a single genetic locus with more than one phenotypic trait. Pleiotropy has been found to contribute to mandibular morphology of inbred mice particularly between traits that are functionally or developmentally related (Cheverud et al. 1997). In their study, pleiotropic effects of QTLs were mainly restricted to portions of the muscular processes of the ascending ramus and to the alveolar processes of the mandibular corpus. Therefore, a single genetic mutation could potentially disrupt several mandibular components thereby potentially increasing variability. Conversely, pleiotropy provides a way to coordinate development between structures with similar functions potentially limiting phenotypic variability.

Mixtures of these interactions have also been detected, specifically, differential dominance and differential epistasis (or epistatic pleiotropy). Differential dominance is the interaction of dominance and pleiotropy (Ehrich et al. 2003). When the dominance effects of a pleiotropic locus differ from trait to trait, the resulting pleiotropic effects display differential dominance (Ehrich et al. 2003). In a QTL study of mouse mandibles, differential dominance was prevalent and this trait-to-trait variation in dominance results in multivariate overdominance even when overdominance does not exist for any individual trait (Ehrich et al. 2003). Therefore, differential dominance alters the genetic architecture and hence the variability.

When epistatic interactions between pairs of loci vary between different traits, differential epistasis or epistatic pleiotropy occurs (Cheverud et al. 2004). Differential epistasis was demonstrated in a study using maize. The mutant maize *opaque-2* has high levels of the essential amino acid lysine in the endosperm. However, the *opaque-2* mutant also has pleiotropic effects on the kernel, making it soft and susceptible to damage (Moro et al. 1996). When this mutant was crossed into other maize stocks followed by selection, loci capable of modifying the pleiotropic effects of *opaque-2* were revealed (Burnett and Larkins 1999; Geetha et al. 1991). These modifying loci removed the negative pleiotropic effects on kernel integrity while maintaining the high lysine level. Cheverud et al. (2004) found epistatic pleiotropy between QTLs associated with inbred mouse mandibular traits (specifically between individual mandibular traits and mandibular length). Therefore, there is genetic variance in pleiotropic effects of these mouse mandibles that will affect variability of mandibular development when exposed to different selection regimes.

## Processes at the Cellular Level

The contribution of neural crest cells to mandibular formation has been actively studied (Francis-West et al. 2003 and references therein). While our treatment of neural crest involvement in mandibular development is necessarily superficial, we try to demonstrate the complexities of development at the cellular level and how these interactions can both limit and amplify phenotypic variability.

Neural crest cells form the embryonic mesenchyme for the skeletal, dental and connective tissues of the mandible (Atchley and Hall 1991). A series of coordinated events must occur in order for mandibular morphogenesis to proceed, including neural crest induction, differentiation, migration and condensation. Neural crest induction is a two-step process involving first, the dorsal mesodermal induction of neural ectoderm, and second, the epidermal ectodermal induction of neural crest at the epidermal/neural ectodermal border (Hall 1999). Differentiation is dependent on a variety of transcription factors, the expression or lack of expression of Hox genes as well as the region of origin in the hindbrain (Francis-West et al. 2003; Gilbert 2003). Neural crest migration relies on a cell-free pathway within the extracellular matrix (Hall 1999). Neural crest cells produce different proteases such as plasminogen activators (Hall 1999) and ADAM-13 (Francis-West et al. 2003) to degrade the extracellular matrix clearing a pathway for migration. To control the duration and direction of neural crest cell migration there must be some degree of cell-to-cell and cell-to-matrix adhesion which are either inhibited or promoted by glycosaminoglycans such as fibronectin within the extracellular matrix (Hall 1999). Both osteogenesis and chondrogenesis within the mandible are dependent on the formation of neural crest cell condensations (Cottrill et al. 1987; Hurle et al. 1989). If condensations are too small, osteogenesis is not initiated and if condensations are abnormally large, abnormally large bones can develop (Hall 1978; Ede 1983; Thorogood 1983; Johnson 1986). Neural crest cell condensation is dependent on the number of cells that are induced, the rate of their migration, the

number of these migrating cells that divide, the timing of these cellular divisions and the rate of cell death (Hall 1988; Atchley and Hall 1991; Hall 1999).

Interaction between neural crest induction, differentiation, migration and condensation can either increase or limit variability of jaw development. For example, neural crest condensation relies on the preceding events of induction, differentiation and migration. Therefore, there is still the potential to have a normal sized neural crest condensation even if fewer cells have left the neural tube (reduced neural crest induction) through the 'tweaking' of these other events such as: (1) increasing the rate of migration, (2) increasing the proportion of migrating cells that divide, (3) increasing the rate of cell divisions, (4) decreasing rate and proportion of cell death and (5) a combination of the above. However, it is also possible that these other neural crest processes will not be able to compensate for the reduced cellular induction and may exacerbate the effects resulting in even greater variability at the cellular level and potentially at the level of the functioning mandible. The interdependence of these complex cellular events can buffer variability as well as amplify errors potentially increasing variability at the cellular level and beyond.

## Processes at the Individual Level

It is at the individual level where we can see how the parts that make up the whole come together in an orchestrated way to form a functioning structure. Following from our mammalian jaw example, the functioning jaw at the individual level reflects the cumulative effects of the developmental processes at the molecular, genetic and cellular levels as well as their interactions. Furthermore, the functioning jaw is the interface between the phenotype and the external environment, and is therefore, subject to selection. The phenotype and phenotypic variability at the individual level are of great importance due to their direct effect on function and potential success of the individual. The main function of the mammalian jaw is food ingestion necessary for survival. At this level, function is the most important factor, and phenotypic variability can be measured in terms of function.

Epigenetics plays a major role in the development of the functioning jaw. Epigenetics are 'heritable changes in gene expression that are not due to changes in DNA sequence' (Eccleston et al. 2007, p. 395). In addition to the role epigenetics plays in the development of the standard phenotype, it can also strongly influence phenotypic variability. It is easier to see how the dependency of the functioning jaw on coordinated interactions between several component parts could lead to increased variability. Essentially, the variability of the functioning whole represents the cumulative variability of the interacting components. However, in terms of the standard phenotype, the variability of component parts can also cancel out the effects of each other, creating a more stable phenotype at the individual level. Following with our jaw example, the resultant bony structures of the mammalian jaw and their phenotypic variability are influenced by epigenetic interactions with other cranial components such as: other skeletal characters of the cranium and jaw, muscles of mastication, and components of the nervous system.

Effective mastication relies on proper occlusion of the jaws and therefore mandibular and maxillary development must be coordinated. Growth of either the mandible or maxilla must be mirrored by the other in order to maintain proper occlusion. Likewise, the cranium and jaws affect each other's growth and development through their connection at the jaw joint. Herring and colleagues (2002) found that loads transmitted through jaw occlusion and through the jaw joint created variable strain patterns throughout the cranium in miniature pigs. These variable strain patterns could lead to variable cranial morphology, which in turn could affect jaw morphology.

The relationship between the muscles of mastication and the skeletal components of the jaw and cranium demonstrate epigenetic influences. The strength and direction of pull created by the muscles can affect cranial morphology as well as phenotypic variation in the skull. Patterns of muscle contraction during chewing directly relate to strains imposed upon the cranium (Herring and Teng 2000; Rafferty et al. 2000). The strains imposed upon the cranium affect skull shape by creating differential growth at cranial sutures and by adding bone at sites of muscle attachments. Cell proliferation in bone at the cranial sutures is influenced by the polarity of the mechanical loading produced by the muscles with compression inhibiting cell proliferation and tension promoting proliferation (Herring 1993; Herring et al. 2002). Studies have shown the effects of masticatory muscle pull on cranial growth by comparing animals fed diets of different consistencies. Animals fed a hard diet tend to have larger neurocranium dimensions than animals fed a soft diet. These results have been demonstrated in miniature pigs (Ciochon et al. 1997), rats (Katsaros et al. 2002) and ferrets (He and Kiliardis 2003). Additionally, the effects of diet consistency and muscle strain have also been found to alter mandibular alveolar dimensions in rats (Yamada and Kimmel 1991; Bresin et al. 1994). Rats fed a hard diet had thicker mandibular alveolar dimensions demonstrating that differential muscle strain not only affects bone apposition in the vicinity of the insertion of muscles, but also in areas that experience increased bending forces due to muscular action (Yamada and Kimmel 1991; Bresin et al. 1994).

The muscles that impart their effects on the skeletal jaws are controlled by their innervation. A brainstem central pattern generator that is mediated by sensory feedback and conscious control produces the rhythmical jaw movements associated with mammalian mastication (Lund and Kolta 2006). Rhythmical activity of the trigeminal muscle is found in rat fetuses at E19 (Turman 2007). Therefore, rhythmical movements of the jaw activated by a central pattern generator are creating strains on the bony jaw structures of these rats potentially altering the bony phenotype.

Phenotypic variability is influenced by the cumulative effects of all of these epigenetic factors involved in jaw development. The functioning jaw reflects the collective interaction of its component parts, and therefore variability in the bony structures of the jaw, in the muscles of mastication and the innervation of these muscles could accumulate to create a highly variable jaw apparatus. However, given the multiple interactions between these components, the cumulative effect of the component variability is most likely non-linear. It is through these interactions where variability at the individual level might be amplified through a cascading effect of developmental errors, or conversely, variability associated with each component part could in fact decrease variability at the individual level. For example, if there was a wide possibility of variant development associated with the development of the mandible and the resultant mandible was much smaller than the standard mandibular phenotype, variability in maxillary development could either amplify the variability of the jaw as a whole or potentially reduce jaw variability. At the functioning jaw level, malocclusion is likely to occur if the maxillary portion of the jaw develops to the standard phenotypic size for the upper jaw, given the small mandible. Through interactions between the upper and lower jaws, the maxilla might also develop to be abnormally small, reflecting development variability of the maxilla, but if the smaller maxilla improves occlusion this same increased developmental variability translates into decreased variability at the individual level (Fig. 7). This example demonstrates that variability is dependent on the level at which it is measured. This level-dependent variability further complicates the genotype-to-phenotype map, and underscores the importance of studying variability using a synthetic systems approach.

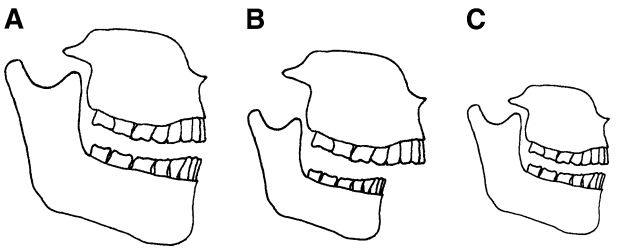


**Fig. 7** A depiction of how the interaction between the maxillary (*upper*) and mandibular (lower) jaw components influence variability of the functioning whole. (**A**) Both maxillary and mandibular components have developed to the standard phenotypic size reflecting relatively low variability in the development of each component, and relatively low variability of the functioning whole due to proper occlusion. (**B**) The maxillary component has developed to the standard phenotypic size but the mandible is much smaller than the standard phenotype, reflecting relatively low variability in maxillary development and higher variability in mandibular development. Malocclusion between the jaws indicates higher variability of the functioning whole. (**C**) Both mandibular and maxillary components are smaller than the standard phenotype reflecting higher variability in the development of both components; however, their integrated (though variable) development has led to proper occlusion and therefore relatively low variability of the functioning jaw

## Processes at the Population and Environmental Levels

We have advocated a systems approach to the study of phenotypic variability including population and environmental level factors. While we emphasize that the factors acting on phenotypic variability at the population and environmental levels are as influential as the factors acting at the molecular, genetic, cellular and individual levels, our (the authors) breadth of understanding of population and environmental factors is limited. Therefore, we limit our discussion of these topics highlighting some of the major factors at work at both of these levels. Readers interested in the topics listed below are encouraged to consult the references provided for a more thorough discussion.

The basic components of the genetic architecture of any individual are filtered from the population from which the individual developed. In order for a developmental process, such as one involved in jaw development to exist within an individual, its genetic underpinnings must exist within the population. A major determinant of variability at the population level is allele frequencies. Allele frequency can change due to genetic drift, chance of survival, mate acquisition and fertility of individuals carrying the alleles (Weiss and Buchanan 2004). Stabilizing selection is another major factor that influences phenotypic variability at the population level (Schmalhausen 1949; Wagner et al. 1997). Stabilizing selection favours the standard phenotype for a given genetic background and set of environmental conditions, and extreme phenotypes (and more importantly their genotypes) on either side of this optimum are removed from the population (Gibson and Wagner 2000). Therefore, alleles associated with extreme phenotypes tend to be selected against and will likely be lost from the population.

Although environmental variables are often ignored, the relationship between the genotype and phenotype can only be defined under a specified set of environmental conditions (Bradshaw 1965; Schlichting 1986; West-Eberhard 1986; Moran 1992; Scheiner 1993; Nijhout

1999). The norm of reaction has been described as the mechanism that transforms environmental variation into a distribution of phenotypic variation (Suzuki et al. 1986). The norm of reaction is the capacity of a genotype to respond to the environment and determines the variety of phenotypes that can be produced by a single genotype in response to environmental variation (Woltereck 1909; Schmalhausen 1949; Bradshaw 1965; Stearns 1989; Schlichting and Pigliucci 1998). Phenotypic plasticity is another factor studied at the environmental level that has a strong influence on phenotypic variability. Phenotypic plasticity refers to phenotypic variation that is induced by environmental factors (Nilsson-Ehle 1914; Schmalhausen 1949; Bradshaw 1965; Stearns 1989). In some circumstances phenotypic variation arising through phenotypic plasticity has been found to be taken over by the genotype through the selection process, so that the novel phenotype(s) develop without the environmental influence. Waddington named this phenomenon genetic assimilation (1942, 1953, 1956, 1957, 1959, 1961). Genetic assimilation stems from Waddington's work on canalization and the idea that development of the phenotype is consistent under a variety of environmental conditions that fall under a particular threshold. If an environmental stimulus surpasses this threshold, genetic variation that has been cryptic due to canalizing processes is uncovered and exposed to selection. Alleles that increase fitness in this new environment will be passed on and become so frequent that the expression of the phenotype will occur in the absence of the environmental stimulus (Waddington 1942, 1953, 1956, 1957, 1959, 1961). A similar phenomenon called genetic accommodation is described by West-Eberhard (2003, 2005). As with genetic assimilation, an environmentally induced trait becomes genetically controlled through exposure of cryptic genetic variation. However, genetic accommodation allows the now genetically controlled trait to remain phenotypically plastic (West-Eberhard 2003; 2005; Braendle and Flatt 2006). Both genetic accommodation and genetic assimilation demonstrate how the environment can have a lasting influence on phenotypic variation and variability beyond the generation influenced by the environment. Unlike the Lamarkian theory of heritable environmental influences, these phenomena act indirectly through exisiting variation to change the genetic architecture.

## A Synthetic Approach

Due to the success of the reductionist approach, we have gained a great understanding of the function of specific genes in isolation. However, our understanding of how these processes interact to form both the phenotype and phenotypic variability is far from complete. A systems approach advocates taking this understanding a step further and looking at how these developmental processes interact within and among levels of organismal hierarchy. The difficulties of conducting studies using this approach are not negligible. Until quite recently we had neither sufficient understanding of particular developmental pathways, nor the computing power and technology to combine information from such pathways. Recent work has taken advantage of such advances using computer modeling to try to determine potential pathways from the genotype to the complex phenotype.

A notable example of this approach is the work of Salazar-Ciudad and Jernvall (2002, 2004, 2005). They argue that it is not possible to predict phenotypic variation based on information of molecular variation alone as phenotypes are produced through development. Therefore, development must be included in studies of phenotypic variation. They further suggest that it is not merely development that needs to be included in theoretical approaches to the genotype–phenotype map; such studies also need to include the dynamics of development (Salazar-Ciudad and Jernvall 2004). That is, one must account for dynamic interactions among developmental components. Salazar-Ciudad and Jernvall (2002) created a mathematical model to predict the development of mammalian teeth. Their model integrates developmental processes from both the genetic and cellular levels producing a more realistic genotype-to-phenotype map. Included in their model are four cell behaviours: (1) cells can secrete signalling molecules, (2) cells can receive signalling molecules and alter their behaviour in response, (3) cells can divide and (4) cells can differentiate. A network of gene products responsible for these cellular behaviours and their interactions is also included in the model (Salazar-Ciudad and Jernvall 2002). Their model is particularly useful as it allows for two-way communication between top (phenotype)-down and bottom (genotype)-up approaches. Using this model, Salazar-Ciudad and Jernvall have determined previously unknown developmental interactions involved in tooth development by altering variational properties at different hierarchical levels within the model (Salazar-Ciudad and Jernvall 2002). This model was also used to determine that developmental processes are themselves evolvable (Salazar-Ciudad et al. 2001a, b; Salazar-Ciudad and Jernvall 2004) and that the evolutionary mechanisms responsible for complex and simple phenotypes are likely different (Salazar-Ciudad and Jernvall 2005). Their model and its applications are excellent examples of how basic developmental processes can create and maintain both the standard phenotype and the variability associated with that phenotype. It also clearly demonstrates how these processes exist and interact at different hierarchical levels

within the organism. A further understanding of how developmental processes affect variability could be gained by extending Salazar-Ciudad and Jernvall's model to include additional processes from hierarchical levels not included in their model. The theoretical findings from such models could then be used to direct further empirical studies of the relationship between phenotypic variability and developmental processes.

In this paper, our focus has been on phenotypic variability, emphasizing the importance of differentiating between its theory, measurement and developmental underpinnings. While our treatment has been based at the phenotypic level, we advocate a systems approach to studying variability. The major obstacle barring our understanding of how the genotype is translated into the phenotype is that the pathways are rarely linear. This nonlinearity arises from the basic developmental processes that form the phenotype, their complex interactions and the variability associated with each process. If we are ever to unravel the mystery of the genotype–phenotype map, we will have to place our current and future understanding of organismal development into a whole systems context.

**Acknowledgements** We thank Miriam Zelditch, David Polly and an anonymous reviewer for their insightful comments on a previous draft of this paper. We also thank Benedikt Hallgrímsson and the members of the Richtsmeier laboratory for their willingness to discuss these issues at great length and for their many helpful suggestions. This work was supported by the National Science Foundation grant BCS-0523637, as well as the William H. Davies Student Fellowship, Medical Science Graduate Research Fellowship, University of Calgary Graduate Studies (to KW).

## Glossary

| | |
|---|---|
| Buffering | any mechanism that limits the phenotypic effects of a perturbation whether it is adaptive or not. |
| Canalization | refers to the property of an organism that ensures similarity of phenotypic expression by buffering development against both environmental and genetic perturbations. |
| Developmental instability | the inverse of developmental stability. |
| Developmental noise | minor environmentally induced developmental accidents. |
| Developmental stability | the property of an organism that buffers variation of micro-environmental origin. That is, developmental stability ensures consistent phenotypic expression within individuals, or within a given genotype and environment. |
| Environmental canalization | refers to the limitation of phenotypic effect caused by exposure to environmental change. |
| Epigenetics | heritable changes in gene expression that are not due to changes in DNA sequence. |
| Fluctuating asymmetry (FA) | random, non-directional deviations from symmetry of bilaterally symmetric traits such as limbs, wings or antennae. FA measures the balance between developmental noise and developmental stability. |
| Genetic architecture | refers to the number of genes and the interactions between them whose products govern a particular trait. |
| Genetic canalization | refers to the robustness of the phenotype to the effects of mutation. |
| Module | a character, or a set of characters that are more tightly integrated internally than they are with other characters. |
| Morphological integration | refers to the inter-relatedness between morphological structures often due to common developmental origins or functional demands. |
| Norm of reaction | the range of expected phenotypes under a given set of environmental and genetic conditions. |
| Perturbation | used here as any disruption to development. Could be genetic, developmental, environmental, or adverse interactions between these factors. |
| Phenotypic plasticity | phenotypic variation induced by the environement. |
| Pleiotropy | the control of one gene on multiple traits. |
| Stabilizing selection | most common form of natural selection, selects for the mean phenotype. |
| Variability | a dispositional term like solubility describing the tendency or propensity to vary. |
| Variation | observed phenotypic differences. |

## References

Atchley, W. R., & Hall, B. K. (1991). A model for development and evolution of complex morphological structures. *Biological Reviews, 66*, 101–157.

Babbitt, G. A., Kiltie, R., & Bolker, B. (2006) Are fluctuating asymmetry studies adequately sampled? Implications of a new model for size distribution. *American Naturalist, 167*, 230–245.

Bagheri, H. C., & Wagner, G. P. (2004). Evolution of dominance in metabolic pathways. *Genetics, 168*, 1713–1735.

Berg, R. L. (1959). The ecological significance of correlational pleiades. *Evolution, 14*, 171–180.

Bjorksten, T. A., Fowler, K., & Pomiankowski, A. (2000). What does sexual trait FA tell us about stress? *TREE, 15*, 163–166.

Blake, W. J., Kaern, M., Cantor, C. R., & Collins, J. J. (2003). Noise in eukaryotic gene expression. *Nature, 422*, 633–637.

Bradshaw, A. D. (1965). Evolutionary significance of phenotypic plasticity in plants. *Advances in Genetics, 13*, 115–155.

Braendle, C., & Flatt, T. (2006). A role for genetic accommodation in evolution? *BioEssays, 28*, 868–873.

Bresin, A., Johansson, C. B., & Kiliaridis, S. (1994). Effects of occlusal strain on the development of the dentoalveolar process in the growing rat. A morphometric study. *European Journal of Experimental Musculoskeletal Research, 3*, 112–122.

Breuker, C. J., Patterson, J. S., & Klingenberg, C. P. (2006). A single basis for developmental buffering of *Drosophila* wing shape. *PloS ONE* 1:e7.

Burnett, R. J., & Larkins, B. A. (1999). *Opaque2* modifiers alter transcription of the 27-kDA $\gamma$-zein genes in maize. *Molecular & General Genetics, 261*, 908–916.

Carroll, S. B., Grenier, J. K., & Weatherbee, S. D. (2005). *From DNA to diversity: Molecular genetics and the evolution of animal design* (2nd ed.). Malden MA: Blackwell Publishing Ltd.

Chernoff, B., & Magwene, P. M. (1999). Morphological integration: Forty years later. In: E. C. Olson, & R. L. Miller (Eds.), Morphological integration (pp. 319–353). Chicago: University of Chicago Press.

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. *Journal of Theoretical Biology, 110*, 155–171.

Cheverud, J. M. (1988). A comparison of genetic and phenotypic correlations. *Evolution, 42*, 958–968.

Cheverud, J. M. (1995). Morphological integration in the saddle-back tamarin (*Saguinus fuscicollis*) cranium. *American Naturalist, 145*, 63–89.

Cheverud, J. M. (1996). Developmental integration and the evolution of pleiotropy. *American Zoologist, 36*, 44–50.

Cheverud, J. M., Routman, E. J., & Irschick, D. J. (1997). Pleiotropic effects of individual gene loci on mandibular morphology. *Evolution, 51*, 2006–2016.

Cheverud, J. M., Ehrich, T. H., Vaughn, T. T., Koreishi, S. F., Linsey, R. B., & Pletscher, L. S. (2004). Pleiotropic effects on mandibular morphology II: Differential epistasis and genetic variation in morphological integration. *Journal of Experimental Zoology (Molecular and Development Evolution)*, *302B*, 424–435.

Ciochon, R. L., Nisbett, R. A., & Corruccini, R. S. (1997). Dietary consistency and craniofacial development related to masticatory function in minipigs. *Journal of Craniofacial Genetics and Developmental Biology, 17*, 96–102.

Clarke, G. M. (1993). The genetic basis of developmental stability. I. Relationships between stability, heterozygosity and genomic coadaptation. *Genetica, 89*, 15–23.

Clarke, G. M. (1998). The genetic basis of developmental stability, V. Inter- and intra-individual character variation. *Heredity, 80*, 562–567.

Cole, T. M. III. (2002). MI Boot Windows-based software for bootstrap comparison of morphological integration. University of Missouri–Kansas School of Medicine, Kansas, MO.

Cole, T. M., III, & Lele, S. (2002). Bootstrap-based methods for comparing morphological integration patterns. *American Journal of Physical Anthopology, Supplement, 34*, 55.

Cottrill, C. P., Archer, C. W., & Wolpert, L. (1987). Cell sorting and chongrogenic aggregate formation in micromass culture. *Developmental Biology, 122*, 503–515.

Debat, V., Alibert, P., David, P., Paradis, E., & Auffray, J. -C. (2000). Independence between developmental stability and canalization in the skull of the house mouse. *Proceedings of Royal Society of London, Series B, 267*, 423–430.

Debat, V., Milton, C. C., Rutherford, S., Klingenberg, C. P., & Hoffmann, A. A. (2006). HSP90 and the quantitative variation of wing shape in *Drosophila melanogaster*. *Evolution, 60*, 2529–2538.

Depew, M. J., Lufkin, T., & Rubenstein, J. L. R. (2002). Specification of jaw subdivisions by Dlx genes. *Science, 298*, 381–385.

Depew, M. J., Simpson, C. A., Morasso, M., & Rubenstein, J. L. R. (2005). Reassessing the Dlx code: The genetic regulation of branchial arch skeletal pattern and development. *Journal of Anatomy, 207*, 501–561.

Dunn, R. B., & Fraser, A. S. (1958). Selection for an invariant character—"vibrissae number"—in the house mouse. *Nature, 181*, 1018–1019.

Dunn, R. B., & Fraser, A. S. (1959). Selection for an invariant character, vibrissae number in the house mouse. *Australian Journal of Biological Sciences, 12*, 506–523.

Eccleston, A, DeWitt, N., Gunter, C., Marte B., & Nath, D. (2007). Epigenetics. *Nature, 447*, 395.

Ede, D. A. (1983). Cellular condensations and chondrogenesis. In: B. K. Hall (Ed.), Cartilage, development, differentiation and growth, Vol. 2. (pp. 143–186). New York: Academic Press.

Ehrich, T. H., Vaughn, T. T., Koreishi, S. F., Linsey, R. B., Pletscher, L. S., & Cheverud, J. M. (2003). Pleiotropic effects on mandibular morphology I. Developmental morphological integration and differential dominance. *Journal of Experimental Zoology (Molecular and Development Evolution), 296B*, 58–79.

Falconer, D. S., & Mackay, T. F. C. (1996). *Introduction to quantitative genetics*. New York: Longman Press.

Fiering, S., Whitelaw, E., & Martin, D. I. K. (2000). To be or not to be active: the stochastic nature of enhancer action. *Bioessays, 22*, 381–387.

Foote, M., & Cowie, R. H. (1988). Developmental buffering as a mechanism for stasis: Evidence from the pulmonate *Theba pisana*. *Evolution, 42*, 396–399.

Francis-West, P. H., Robson, L., & Evans, D. J. R. (2003). Craniofacial development: The tissue and molecular interactions that control development of the head. In: F. Beck, B. Christ, W. Kriz, W. Kummer, E. Marani, R. Putz, Y. Sano, T. H. Schiebler, G. L. Schoenwolf, & K. Zilles (Eds.), Advances in anatomy and cell biology. Vol. 169. Berlin: Springer-Verlag.

Gass, G. L., & Bolker, J. A. (2002). Modularity. In: W. Olson (Ed.), Keywords and concepts in evolutionary developmental biology. Cambridge: Harvard University Press.

Geetha, K. B., Lending, C. R., Lopes, M. A., Wallace, J. C., & Larkins, B. A. (1991). *Opaque-2* modifiers increase $\gamma$-zein synthesis and alter its spatial distribution in maize endosperm. *Plant Cell, 3*, 1207–1219.

Gibson, G., & van Helden, S. (1997). Is function of the Drosophila homeotic gene Ultrabithorax canalized? *Genetics, 147*, 1155–1168.

Gibson, G., & Wagner, G. (2000). Canalization in evolutionary genetics: a stabilizing theory? *BioEssays, 22*, 372–380.

Gilbert, S. F. (2003). *Developmental biology* (7th ed.) Massachusetts: Sinauer Associated Inc.

Gould, S. J., & Garwood, R. A. (1969). Levels of integration in mammalian dentitions: An analysis of correlations in

*Nesophontes micrus* (Insectivora) and *Oryzomys couesi* (Rodentia). *Evolution, 23*, 276–300.

Graham, A. (2002). Jaw development: Chinless wonders. *Current Biology, 12*, 810–812.

Graham, J. H., Freeman, D. C., & Emlen, J. M. (1993). Developmental stability: A sensitive indicator of populations under stress. In: M. A. Lewis (Ed.), Environmental toxicology and risk assessment (pp. 136–158). Philadelphia: American Society for Testing and Materials.

Graham, J. H., Shimizu, K., Emlen, J. M., Freeman, D. C., & Merkel, J. (2003). Growth model and the expected distribution of fluctuating asymmetry. *Biological Journal of the Linnean Society, 80*, 57–65.

Hall, B. K. (1978). *Developmental and cellular skeletal biology*. New York: Academic Press.

Hall, B. K. (1988). The embryonic development of bone. *American Scientist, 76*, 174–181.

Hall, B. K. (1999). *The neural crest in development and evolution*. New York: Springer-Verlag.

Hallgrímsson, B., Willmore, K., & Hall, B. K. (2002). Canalization, developmental stability, and morphological integration in primate limbs. *American Journal of Physical Anthropology Supplement, 35*, 131–158.

He, T., & Kiliaridis, S. (2003). Effects of masticatory muscle function on craniofacial morphology in growing ferrets (*Mustela putorius furo*). *European Journal of Oral Sciences, 111*, 510–517.

Hermisson, J., Hansen, T. F., & Wagner, G. P. (2003). Epistasis in polygenic traits and the evolution of genetic architecture under stabilizing selection. *American Naturalist, 161*, 708–734.

Hermisson, J., & Wagner, G. P. (2004). Canalization in evolutionary genetics: a stabilizing theory? *BioEssays, 22*, 372–380.

Herring, S. W. (1993). Epigenetic and functional influences on skull growth. In: J. Hanken, & B. K. Hall (Eds.), The skull, Vol. 1 (pp. 153–206). Chicago: University of Chicago Press.

Herring, S. W., & Teng, S. (2000). Strain in the braincase and its sutures during function. *American Journal of Physical Anthropology, 112*, 575–593.

Herring, S. W., Decker, J. D., Liu, Z.-J., & Ma, T. (2002). Temporomandibular joint in miniature pigs: anatomy, cell replication, and relation to loading. *Anatomical Record, 266*, 152–166.

Houle, D. (1998). How should we explain variation in genetic variance of traits? *Genetica, 102/103*, 241–253.

Huerta-Sanchez, E., & Durrett, R. (2007). Wagner's canalization model. *Theoretical Population Biology, 71*, 121–130.

Hurle, J. M., Gana, Y., & Marcias, D. (1989). Experimental analysis of the in-vivo chondrogenic potential of the interdigital mesenchyme of the chick leg bud subjected to local ectodermal removal. *Developmental Biology, 132*, 368–374.

Jacob, F. (1977). Evolution and tinkering. *Science, 196*, 1161–1166.

Jernvall, J., & Jung, H. S. (2000). Genotype, phenotype, and developmental biology of molar tooth characters. *American Journal of Physical Anthropology, 43*, 171–190.

Johnson, D. R. (1986). *The genetics of the skeleton*. Oxford: Clarendon Press.

Katsaros, C., Berg, R., & Kiliaridis, S. (2002). Influence of masticatory muscle function on transverse skull dimensions in the growing rat. *Journal of Orofacial Orthopedics, 63*, 5–13.

Kaufmann, W. K., & Paules, R. S. (1996). DNA damage and cell cycle checkpoints. *The FASEB Journal, 10*, 238–247.

Klingenberg, C. P. (2003). A developmental perspective on developmental instability: Theory, models and mechanisms. In: M. Polak (Ed.), Developmental Instability (DI): Causes and Consequences (pp. 14–34). Oxford: Oxford University Press.

Klingenberg, C. P., & McIntyre, G. S. (1998). Geometric Morphometrics of developmental instability: analyzing patterns of fluctuating asymmetry with Procrustes methods. *Evolution, 52*, 1363–1375.

Klingenberg, C. P., & Zaklan, S. D. (2000). Morphological integration between developmental compartments in the Drosophila wing. *Evolution, 54*, 1273–1285.

Klingenberg, C. P., Leamy, L. J., Routman, E. J., & Cheverud, J. M. (2001). Genetic architecture of mandible shape in mice: Effects of quantitative trait loci analyzed by geometric morphometrics. *Genetics, 157*, 785–802.

Kuratani, S. (2005). Developmental studies of the lamprey and hierarchical evolutionary steps towards the acquisition of the jaw. *Journal of Anatomy, 207*, 489–499.

Leamy, L. J., Routman, E. J., & Cheverud, J. M. (2002). An epistatic genetic basis for fluctuating asymmetry of mandible size in mice. *Evolution, 56*, 642–653.

Lens, L., VanDongen, S., Kark, S., & Matthysen, E. (2002). Fluctuating asymmetry as an indicator of fitness: Can we bridge the gap between studies? *Biological Reviews, 77*, 27–38.

Leung, B., Forbes, M. R., & Houle, D. (2000). Fluctuating asymmetry as an indicator of stress: comparing efficacy of analyses involving multiple traits. *American Naturalist, 155*, 101–115.

Ludwig, W. (1932). *Das Rechts-links problem im Teirreich und beim Menschen*. Berlin (Germany): Springer.

Lund, J. P., & Kolta, A. (2006). Generation of the central masticatory pattern and its modification by sensory feedback. *Dysphagia, 21*, 167–174.

Manning, J. T., & Chamberlain, A. T. (1994). Fluctuating asymmetry in gorilla canines: a sensitive indicator of environmental stress. *Proceedings of Royal Society of London, Series B, 255*, 189–193.

Marriog, G., & Cheverud, J. M. (2001). A comparison of phenotypic variation and covariation patterns and the role of phylogeny, ecology, and ontogeny during cranial evolution of new world monkeys. *Evolution, 55*, 2576–2600.

McAdams, H. H., & Arkin, A. (1999). It's a noisy business! Genetic regulation at the nanomolar scale. *Trends in Genetics, 15*, 65–69.

Mills, K. D., Ferguson, D. O., & Alt, F. W. (2003). The role of DNA breaks in genomic instability and tumorigenesis. *Immunological Reviews, 194*, 77–94.

Milton, C. C., Huynh, B., Batterham, P., Rutherford, S. L., & Hoffmann, A. A. (2003). Quantitative trait symmetry independent of Hsp90 buffering: distinct modes of canalization and developmental stability. *Proceedings of National Academy of Sciences of the United States of America, 100*, 13396–13401.

Mohrenweiser, H. W., Wilson, D. M., & Jones, I. M. (2003). Challenges and complexities in estimating both the functional impact and the disease risk associated with the extensive genetic variation in human DNA repair genes. *Mutation Research, 526*, 93–125.

Moran, P. A. P. (1992). The evolutionary maintenance of alternative phenotypes. *American Naturalist, 139*, 971–989.

Moro, G. L., Habben, J. E., Hamaker, B. R., & Larkins, B. A. (1996). Characterization of the variability in lysine content for normal and opaque2 maize endosperm. *Crop Science, 36*, 1651–1659.

Moss, M. (1971). Functional cranial analysis and the functional matrix. *American Speech Hearing Association Report, 6*, 5–18.

Moss, M., & Young, R. (1960). A functional approach to craniology. *American Journal of Physical Anthropology, 18*, 281–292.

Nijhout, F. H. (1999). Control mechanisms of polyphonic development in insects. *BioScience, 49*, 181–192.

Nijhout, H. F., & Davidowitz, G. (2003). Developmental perspectives on phenotypic variation: canalization, and fluctuating asymmetry. In: M. Polak (Ed.), Developmental instability (DI): Causes and consequences (pp. 3–13). Oxford: Oxford University Press.

Nilsson-Ehle, H. (1914). Vilka erfarenheter hava hittills vunnits rörande möjligheten av växters acklimatisering? *Kungl Landtbruksakad Hand Tidskr, 53*, 537–572.

Olson, E. C., & Miller, R. L. (1951). A mathematical model applied to a study of the evolution of species. *Evolution, 5*, 256–338.

Olson, E. C., & Miller, R. L. (1958). *Morphological integration*. Chicago: University of Chicago Press.

Palmer, A. R. (1996). Waltzing with asymmetry. *BioScience, 46*, 518–532.

Palmer, A. R., & Strobeck, C. (1986). Fluctuating Asymmetry: Measurement, Analysis, Patterns. *Annual Review Ecology Systematics, 17*, 391–421.

Palmer, A. R., & Strobeck, C. (1992). Fluctuating asymmetry as a measure of developmental stability: Implications of non-normal distributions and power of statistical tests. *Acta Zoologica Fennica, 191*, 57–72.

Palmer, R., & Strobeck, C. (2003). Fluctuating asymmetry analysis unplugged. In: M. Polak (Ed.), Developmental instability (DI): Causes and consequences (pp. 279–319). Oxford: Oxford University Press.

Qiu, M., Bulfone, A., Ghattas, I., Meneses, J. J., Christensen, L., Sharpe, P. T., Presley, R., Pederson, R. A., & Rubenstein, J. L. R. (1997). Role of the Dlx homeobox genes in proximodistal patterning of the branchial arches: Mutations of Dlx-1, Dlx-2 and -2 alter morphogenesis of proximal skeletal and soft tissue structures derived from first and second arches. *Developmental Biology, 185*, 165–184.

Rafferty, K. L., Herring, S. W., & Artese, F. (2000). Three-dimensional loading and growth of the zygomatic arch. *Journal of Experimental Biology, 203*, 2093–3004.

Rasmuson, M. (2002). Fluctuating asymmetry—indicator of what? *Hereditas, 136*, 177–183.

Réale, D., & Roff, D. A. (2003). Inbreeding, developmental stability, and canalization in the sand cricket *Gryllus firmus*. *Evolution, 57*, 597–605.

Reeve, E. C. R. 1960. Some genetic tests on asymmetry of sternopleural chaetae number in *Drosophila*. *Genetical Research, 1*, 151–172.

Reeve, E. C. R., & Robertson, F. W. (1953). Analysis of environmental variability in quantitative inheritance. *Nature, 171*, 874–875.

Rendel, J. M. (1959). Canalization of the scute phenotype of Drosophila. *Evolution, 13*, 425–439.

Richtsmeier, J. T., Aldridge, K. A., DeLeon, V. B., Panchal, J., Kane, A. A., Marsh, J. L., Yan, P., Cole, T. M., III. 2006. Phenotypic integration of neurocranium and brain. *Journal of Experimental Zoology (Molecular and Development Evolution), 306B*, 1–19.

Routman, E. J., & Cheverud, J. M. (1997). Gene effects on a quantitative trait: two-locus epistatic effects measured at microsatellite markers and at estimated QTL. *Evolution, 51*, 1654–1662.

Rutherford, S. L. (2000). From genotype to phenotype: buffering mechanisms and the storage of genetic information. *Bioessays, 22*, 1095–1105.

Rutherford, S. L., & Lindquist, S. (1998). Hsp90 as a capacitor for morphological evolution. *Nature, 396*, 336–342.

Salazar-Ciudad, I., Newman, S. A., Solé RV. 2001a. Phenotypic and dynamical transitions in model genetic networks I. Emergence of patterns and genotype–phenotype relations. *Evolution & Development, 3*, 84–94.

Salazar-Ciudad, I., Solé RV, Newman SA. 2001b. Phenotypic and dynamical transitions in model genetic networks II. Application to the evolution of segmentation mechanisms. *Evolution & Development, 3*, 95–103.

Salazar-Ciudad, I., & Jernvall, J. (2002). A gene network model accounting for development and evolution of mammalian teeth. *Proceedings of National Academy of Sciences of the United States of America, 99*, 8116–8120.

Salazar-Ciudad, I., & Jernvall, J. (2004). How different types of pattern formation mechanisms affect the evolution of form and development. *Evolution & Development, 6*, 6–16.

Salazar-Ciudad, I., & Jernvall, J. (2005). Graduality and innovation in the evolution of complex phenotypes: Insights from development. *Journal of Experimental Zoology (Molecular and Development Evolution), 304B*, 619–631.

Santos, M., Fernández Iriarte, P., & Céspedes, W. (2005). Genetics and geometry of canalization and developmental stability in *Drosophila subobscura*. *BMC Evolutionary Biology, 5*, 7.

Scharloo, W. (1991). Canalization: Genetic and developmental aspects. *Annual Review of Ecology Systematics, 22*, 65–93.

Scheiner, S. M. (1993). Genetics and evolution of phenotypic plasticity. *Annual Review of Ecology Systematics, 24*, 35–68.

Schlichting, C. D. (1986). The evolution of phenotypic plasticity in plants. *Annual Review of Ecology Systematics, 17*, 667–693.

Schlichting, C. D., & Pigliucci, M. (1998). *Phenotypic evolution: A reaction norm perspective*. Sunderland MA: Sinauer.

Schmalhausen, I. I. (1949). *Factors of evolution*. Chicago: University of Chicago Press.

Siegal, M. L., & Bergman, A. (2002). Waddington's canalization revisited: developmental stability and evolution. *Proceedings of National Academy of Sciences of the United States of America, 99*, 10528–10532.

Stearns, S. C. (1989). The evolutionary significance of phenotypic plasticity. *BioScience, 39*, 436–445.

Stock, D. W. (2005). The Dlx complement of the leopard shark, *Triakis semifasciata*, resembles that of mammals: Implications for genomic and morphological evolution of jawed vertebrates. *Genetics, 169*, 807–817.

Suzuki, D. T., Griffiths, A. J. F., Miller, J. H., & Lewontin, R. C. (1986). *An introduction to genetic analysis*. New York: WH Freeman.

Thomas, J. H. (1993). Thinking about genetic redundancy. *Trends in Genetics, 9*, 305–309.

Thorogood, P. (1983). Morphogenesis of cartilage. In: B. K. Hall (Ed.), Cartilage, development, differentiation and growth. Vol. 2. (pp. 223–255). New York: Academic Press.

Turman, J. E., Jr. 2007. The development of mastication in rodents: From neurons to behaviors. *Archives of Oral Biology, 52*, 313–316.

Van Dongen, S. (2006). Fluctuating asymmetry and developmental instability in evolutionary biology: past, present and future. *Journal of Evolutionary Biology, 19*, 1727–1743.

Van Valen, L. (1962). A study of fluctuating asymmetry. *Evolution, 16*, 125–142.

von Dassow, G., Meir, E., Munro, E., & Odell, G. (2000). The segment polarity network is a robust development module. *Nature, 406*, 188–192.

Waddington, C. H. (1942). The canalisation of development and the inheritance of acquired characters. *Nature, 150*, 563.

Waddington, C. H. (1953). Genetic assimilation of an acquired character. *Evolution, 7*, 118–126.

Waddington, C. H. (1956). Genetic assimilation of the bithorax phenotype. *Evolution, 10*, 1–13.

Waddington, C. H. (1957). *Strategy of the genes*. New York: MacMillan.

Waddington, C. H. (1959). Canalization of development and genetic assimilation of acquired characters. *Nature, 183*, 1654–1655.

Waddington, C. H. (1961). Genetic assimilation. *Advances in Genetics, 10*, 257–293.

Waddington, C. H. (1975). *The Evolution of an evolutionist*. Ithaca, New York: Cornell University Press.

Wagner, A. (1999). Redundant gene functions and natural selection. *Journal of Evolutionary Biology, 12*, 1–16.

Wagner, A. (2005). Distributed robustness versus redundancy as causes of mutational robustness. *BioEssays, 27*, 176–188.

Wagner, G. P. (1996). Homologues, natural kinds and the evolution of modularity. *American Zoologist, 36*, 36–43.

Wagner, G. P., & Altenberg, L. (1996). Complex adaptations and the evolution of evolvability. *Evolution, 50*, 967–976.

Wagner, G. P., Booth, G., Bagheri-Chaichian, H. 1997. A population genetic theory of canalization. *Evolution, 51*, 329–347.

Wagner, G. P., Laubichle, MD, Bagheri-Chaichian, H. (1998). Genetic measurement theory of epistatic effects. *Genetica, 102–103*, 569–580.

Weiss, K. M., & Fullerton, S. M. (2000). Phenogenetic drift and the evolution of genotype–phenotype relations. *Theoretical Population Biology, 57*, 187–195.

Weiss, K. M., & Buchanan, A. V. (2004). *Genetics and The logic of Evolution*. New York: John Wiley.

Weiss, P. A. (1971). The basic concept of hierarchic systems. In: P. A. Weiss (Ed.), Hierarchically organized systems in theory and practice (pp. 1–43.). New York: Hafner Publishing Company.

West-Eberhard, M. J. (1986). Alternative adaptations, speciation, and phylogeny. *Proceedings of National Academy of Sciences of the United States of America, 83*, 1388–1392.

West-Eberhard, M. J. (2003). *Developmental plasticity and evolution*. New York: Oxford University Press.

West-Eberhard, M. J. (2005). Phenotypic accommodation: Adaptive innovation due to developmental plasticity. *Journal of Experimental Zoology (Molecular and Development Evolution), 304B*, 610–618.

Willmore, K. E., Klingenberg, C. P., Hallgrímsson, B. 2005. The relationship between fluctuating asymmetry and environmental variance in Rhesus macaque skulls. *Evolution, 59*, 898–909.

Willmore, K. E., Hallgrímsson, B. (2005). Within individual variation: Developmental noise versus developmental stability. In: B. Hallgrímsson, & B. K. Hall (Eds.), Variation: A central concept in biology (pp. 191–218). New York: Elsevier Academic Press.

Willmore, K. E., Leamy, L., Hallgrímsson, B. (2006). Effects of developmental and functional interactions on mouse cranial variability through late ontogeny. *Evolution & Development, 8*, 550–567.

Woltereck, R. (1909). Weitere experimentelle untersuchungen über artveränderung, speziell über das wesen quantitaver artunterschiede bei Daphnien. *Verh Deutsch Zool Gesellsch, 19*, 110–173.

Woods, R. E., Sgrò CM, Hercus, M. J., & Hoffmann, A. A. (1999). The association between fluctuating asymmetry, trait variability, trait heritability, and stress: A multiply replicated experiment on combined stresses in *Drosophila melanogaster*. *Evolution, 53*, 493–505.

Yamada, K., & Kimmel, D. B. (1991). The effect of dietary consistency on bone mass and turnover in the growing rat mandible. *Archives of Oral Biology, 36*, 129–138.

Young, N. M. (2006). Function, ontogeny and canalization of shape vairance in the primate scapula. *Journal of Anatomy, 209*, 623–636.

Young, N. M., Hallgrímsson, B. (2005). Serial homology and the evolution of mammalian limb covariation structure. *Evolution, 59*, 2691–2704.

Zakharaov, V. M. (1992). Population phenotgenetics: analysis of developmental stability in natural populations. *Acta Zoologica Fennica, 191*, 7–30.

Zelditch, M. L., Bookstein, F. L., & Lundrigan, B. L. (1993). The ontogenetic complexity of developmental constraints. *Journal of Evolutionary Biology, 6*, 621–641.

Zelditch, M. L., Lundrigan, B. L., & Garland, T. (2004). Developmental regulation of skull morphology. I. Ontogenetic dynamics of variance. *Evolution & Development, 6*, 194–206.