Evol Biol (2009) 36:100–117
DOI 10.1007/s11692-009-9049-8

RESEARCH ARTICLE

Integration and Evolvability in Primate Hands and Feet

Campbell Rolian

Received: 15 December 2008 / Accepted: 16 January 2009 / Published online: 17 March 2009

Abstract Morphological integration theory predicts that sets of phenotypic traits that covary strongly due to developmental and/or functional connections between them eventually co-evolve because of a coordinated response to selection, and accordingly become less independently evolvable. This process is not irreversible, however, and phenotypic traits can become less integrated, and hence more independently evolvable, in the context of selection for divergent functions and morphologies. This study examines the reciprocal relationship between shared function, integration and evolvability by comparing integration patterns among serially homologous skeletal elements in the hands and feet of a functionally diverse sample of catarrhine primates. Two hypotheses are tested: (1) species in which the autopods are functionally more similar (e.g. quadrupedal monkeys) have more strongly integrated autopods than species in which the autopods are functionally divergent (e.g. gibbons, humans) and (2) the latter have autopods that are more evolvable, collectively and independently. Morphometric data from selected hand and foot digital rays were used to derive phenotypic variance/covariance matrices. The strength of integration among autopods was quantified using eigenanalysis and a measure of trait variational autonomy. Evolvability was estimated by subjecting phenotypic variance/covariance matrices to simulated random selection gradients, and comparing evolutionary responses among species. Results indicate that integration decreases as hands and feet become functionally divergent, and that the strongly integrated hand and foot skeletons of quadrupedal monkeys respond to selection in a highly collinear manner, even when simulated selective pressures acting on each autopod are in opposite directions in phenotypic space. Results confirm that the evolvability of morphological traits depends largely on how strongly they covary with other traits, but also with body size. The role of pleiotropy as a developmental mechanism underlying integration and evolvability is also discussed.

Keywords Morphological integration · Evolvability · Autopods · Primates · Pleiotropy

C. Rolian (✉)
Department of Cell Biology and Anatomy, University of Calgary, G503 - 3330 Hospital Drive NW, Calgary, AB T2N4N1, Canada
e-mail: cprolian@ucalgary.ca; campbellrolian@gmail.com

Introduction

Morphological integration refers to the tendency of some phenotypic traits to covary more strongly than others because of functional or developmental connections between them. Darwin (1859) was aware of this property of complex organisms, but the study of phenotypic covariation was not formalized until the mid-twentieth century, when Everett Olson and Robert Miller published a series of studies outlining their theory of morphological integration (1951, 1958). Olson and Miller's approach was largely statistical and pattern-based: phenotypic correlations were first quantified, strongly correlated traits were then grouped into subsets, and inferences were made on developmental and functional reasons for increased covariation among these subsets. Integration theory received renewed attention among evolutionary biologists, following a series of landmark studies by James Cheverud on integration in the primate skull (Cheverud 1982, 1995, 1996). Cheverud's insight was to place the study of integration within the theoretical framework of quantitative genetics (Lande

1979, 1980), and to provide a set of processes for the evolution of phenotypic covariation.

Cheverud (1996) identified four kinds of integration organized hierarchically between individuals and populations. Within individuals, integration is driven largely by *functional* and *developmental* factors, as originally formulated by Darwin, and Olson and Miller. Functional integration occurs when different phenotypic traits participate in a common function, and the interactions of the traits affect an organism's performance of that function. Developmental integration occurs when traits interact during organismal development, or are directed by common developmental processes. As Cheverud suggests, these two forms of integration are related, because "development can be viewed as dynamic function, and functional integration in the adult is likely to be achieved through developmental integration" (1996, p. 45). At the level of populations, morphological traits become *genetically* integrated when sets of phenotypic traits are inherited together, more or less independently of other sets of traits. This co-inheritance occurs mainly via pleiotropy and linkage disequilibrium. Finally, Cheverud proposed that genetic integration leads to *evolutionary* integration, when morphological traits co-evolve because of a coordinated response to selection. Importantly, Cheverud argued that pleiotropy is a key mechanism linking these four types of integration (see also Wagner 1996; Wagner and Altenberg 1996).

This hierarchical view of integration implies that as phenotypic traits become integrated at the population level, their responses to selection become increasingly coordinated, while their ability to respond to independent selective pressures is reduced. In other words, as traits become more strongly integrated, their independent evolvability may be reduced. This reduction in evolvability is not irreversible, or else increasing integration among traits would eventually lead to the complete loss of evolvability in complex organisms. Instead, Cheverud's model predicts that integration among traits can also be selected against, and can be achieved through a reduction in pleiotropic interactions among these traits ('parcellation', Wagner and Altenberg 1996, 'dissociation', Hansen 2003). Parcellation may occur, for example, when morphological structures that once shared common functions, or share a common developmental basis, are selected to be increasingly functionally specialized (Hallgrimsson et al. 2002). Parcellation further predicts that traits that become less integrated as a result of selection against functional and/or developmental integration should also become more independently evolvable.

The complex reciprocal relationships between individual and population levels of integration, and especially between integration and evolvability, have received comparatively little attention in integration studies (reviewed in Magwene and Chernoff 1999, see also Hansen 2003). This may be due in part to theoretical and practical difficulties in teasing apart functional and developmental causes of integration within individuals, and until recently to a lack of tools for examining the relationship between integration and evolvability within the framework of quantitative genetics. Here I address some of these relationships by quantifying and comparing integration—and its impact on evolvability—in the hand and foot skeletons of Old World monkeys, apes and humans (catarrhine primates).

Primate hands and feet (autopods) provide a useful tool to study the relationship between hierarchical levels of integration. Autopods are serially homologous structures that evolved from an ancient genetic duplication event (Hall 1995; Ruvinsky and Gibson-Brown 2000). By definition, serially homologous structures share large portions of their genetic and developmental architecture, a fact which has been empirically validated in vertebrate autopods (Margulies et al. 2001; Shou et al. 2005). Primate hands and feet are thus already developmentally integrated and prone to pleiotropic effects. At the same time, however, catarrhine autopods have evolved to perform a diversity of functions in postural, locomotor and manipulative behaviors. Importantly, catarrhines vary substantially in the extent to which such functions are shared *between* the hands and feet: from arboreal and terrestrial quadrupedal monkeys (Cercopithecoids) in which the hands and feet share many functions in grasping and propulsion, to forelimb-dominated Hominoids (e.g. gibbons) in which the arms and hands have become specialized for vertical climbing and suspensory behaviors, and humans, in which the functional divergence of the autopods is extreme (Stern and Oxnard 1973; Fleagle 1999).

Catarrhine autopods thus vary in the extent to which they are functionally integrated, but presumably less so in how strongly they are developmentally integrated. This dual property can be used to test hypotheses about the unique role of functional convergence (and divergence) in driving population-level integration. In addition, Hansen and Houle (2008) have developed tools for quantifying integration and evolvability in multivariate data. Like Cheverud's model of integration, these tools are based in a quantitative genetics framework, and can be used to test hypotheses regarding how functional integration within catarrhine autopods ultimately affects the ability of hands and feet to evolve independently in response to different selective challenges.

This study tests two related hypotheses. The first is that the strength of morphological integration, measured as phenotypic covariation among hand and foot skeletal elements, increases in relation to increasing functional convergence between autopods (see also Young and Hallgrimsson 2005). Put differently, primates in which the hands and feet have evolved divergent functions, such as

gibbons and humans, will show weaker patterns of integration than species in which the autopods share common functions, such as quadrupedal monkeys. As a corollary, the second hypothesis to be tested is that species whose autopods are more strongly integrated have autopods that are less independently evolvable. Specifically, I test the prediction that as a consequence of greater functional integration, the autopods of quadrupedal monkeys show more strongly coordinated responses to simulated natural selection, and are less able to evolve independently in response to these selective pressures.

Materials and Methods

Sample

Eleven catarrhine primates are included in this study (Table 1). The sample is equally divided among Cercopithecoids and Hominoids, and covers a spectrum of functional divergence of the hands and feet, including arboreal and terrestrial quadrupedal monkeys in which the hand and foot share similar functions in grasping and propulsion, forelimb-dominated suspensory and knuckle-walking apes in which the hand and foot have relatively less functional overlap, and humans, in which the functional dissociation of the hand and foot is extreme. All samples are adults of known sex with no pathology in the appendicular skeleton. Specimens were considered adult when their long bone epiphyses were fused or in the process of fusing, i.e. when a gap formerly occupied by the cartilaginous growth plate was visible but the epiphyses were attached to the shaft by bony struts. A few older non-human primate individuals suffering from mild cases of osteoarthritis were included as the pathology did not visibly deform the articular surface or diaphyses of the affected bones.

Data

Morphological data were derived from the appendicular skeleton from a single side in each individual. In most cases at least one side was available in disarticulated form in each limb. In a small number of individuals ($\sim$80) where a critical element was missing, broken or showed evidence of healed fractures, the element was replaced with its homolog from the contralateral side. The metapodials, proximal and middle phalanges of all digits in the hand and foot were placed ventrally in anatomical position on a flatbed scanner (Microtek i320 ScanMaker, Carson, CA) and scanned in TIFF format at 300 pixels per inch (ppi). The scanner method has been validated in previous integration studies (e.g. Hallgrimsson et al. 2002; Young and Hallgrimsson 2005). Distal phalanges were not included as they are rarely preserved in skeletal preparations. The identity of the phalanges was determined by size and morphological differences, or when available by comparison with the articulated side of the skeleton.

Two-dimensional digital landmarks were placed on homologous morphological features in each taxon using TPSDig2 (Rohlf 2005). Landmarks were used to derive lengths of the metapodial and phalangeal elements, and the

Table 1 Sample composition

Species	Males	Females	Total	% Wild	Provenance
Homo sapiens (Hom)	102	102	204	–	CMNH, NMNH
Pan troglodytes (Pan)	37	61	98	95	PCM, AMNH, NMNH, MCZ
Gorilla gorilla (Gor)	53	44	97	95	PCM, AMNH, NMNH, MCZ
Pongo pygmaeus (Pon)	21	30	51	70	NMNH, BSM, MCZ
Hylobates lar (Hyl)	34	33	67	100	MCZ
Macaca mulatta (Mac)	36	68	104	60	NMNH, UPR
Papio hamadryas (Pap)	39	43	82	40	UT, UIUC
Cercopithecus (Cer)	31	32	63	80	NMNH, AMNH, UM
Erythrocebus patas (Ery)	24	34	58	0	UPR
Trachypithecus cristatus (Tra)	21	38	59	100	BSM, MCZ
Total	–	–	883	–	–

Abbreviations: *AMNH* American Museum of Natural History, New York; *BSM* Bayerische Staatssammlung, Munich; *CMNH* Cleveland Museum of Natural History; *MCZ* Harvard Museum of Comparative Zoology, Cambridge, *NMNH* National Museum of Natural History, Washington; *PCM* Powell-Cotton Museum, Birchington; *UM* Department of Anthropology, University of Minnesota, Minneapolis; *UIUC* Department of Anthropology, University of Illinois—Urbana-Champaign; *UPR* Laboratory for Primate Morphology and Genetics, University of Puerto Rico, San Juan; *UT* Department of Anthropology, University of Texas—Austin. The *Cercopithecus* sample comprises *C. mitis* and *C. ascanius* (see text). The three-letter species abbreviations are used in all following tables and figures

width of the distal end of the proximal phalanges in the first, third and fifth rays of the autopods, using the intrinsic scaling factor determined by the scanning resolution (Fig. 1). The first, third and fifth rays were chosen because they are easily distinguished on the basis of size and morphology, and avoid potential sex effects known to affect lengths of the second and fourth digital rays (2D:4D ratios, e.g. McFadden and Bracht 2003, 2005). To minimize redundancy, the lengths of the phalangeal elements were summed within a digit, providing three variables per digit (metapodial length, phalangeal length and distal width), for a total of 18 variables. Emphasis was placed on element lengths as these have been shown to vary substantially within and among taxa in relation to autopod functional specializations (e.g. Jouffroy et al. 1991; Susman 1979; Strasser 1994).

Estimation of Phenotypic VCV Matrices

The hypotheses outlined above were tested by comparing the strength of integration and potential for evolvability embodied in the autopod phenotypic variance/covariance (VCV) matrices of the catarrhine sample (Table 1). Testing these hypotheses within a quantitative genetics framework ideally requires knowledge of additive genetic variances and covariances among autopod traits (Lande 1979; Hansen and Houle 2008), but such data are difficult to obtain. Phenotypic and genetic VCV matrices are proportional, however, and several studies have shown that the phenotypic VCV matrix derived from morphometric traits can serve as a proxy for its genetic counterpart, particularly if heritabilities do not differ significantly between traits and are comparable across species, which is likely the case for serially homologous structures (Cheverud 1988; Roff 1995, 1996; Ackermann and Cheverud 2000, 2004; Porto et al. 2009, but see Willis et al. 1991 for a contrasting view). This study therefore uses phenotypic VCV matrices in lieu of their genetic counterparts, under the assumption that the latter are directly proportional to the former, and have similar scaling factors across autopod traits and across species.

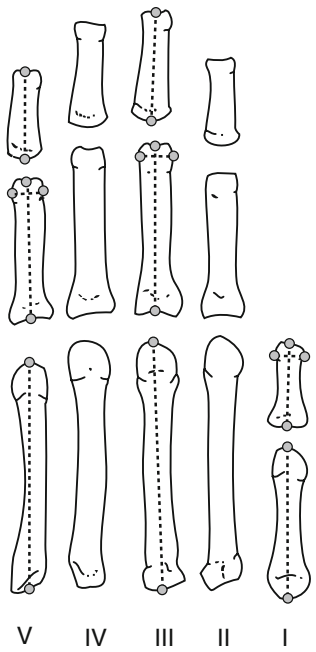

Fig. 1 Illustration of the flatbed scanner method for data collection. Right hand elements from a *Cercopithecus mitis* specimen are laid out ventrally in anatomical position, numbered from right to left (I = thumb, V = fifth digit). Digital landmarks (*grey circles*) were placed as shown, and at homologous locations in the foot scan, to obtain element length and width data from both autopods (*dashed lines*). Note that the lengths of the phalangeal elements in digits III and V were summed prior to analysis

The digital ray data were used to estimate autopod phenotypic VCV matrices for each taxon. Raw data were first adjusted to reduce the effects of confounding sources of phenotypic (co)variance that are not relevant to this study. All species were corrected for sexual dimorphism by adding the difference between sex means to the individuals in the group with the smaller mean (Marroig and Cheverud 2004). Whenever necessary, the same approach was also used to first reduce the effects of other sources of phenotypic (co)variance within sex, such as size dimorphism related to population substructuring. For example, the *Macaca* sample consists of 60% free-ranging individuals from the Cayo Santiago colony (Laboratory for Primate Morphology and Genetics, University of Puerto Rico), and 40% captive individuals from the Smithsonian Institute. Sex-specific ANOVAs revealed significant size differences between the groups, with the LPMG individuals being on average 6% larger. Data from the wild and captive groups were pooled within sex by adding the mean differences between the wild and captive subpopulations to each captive individual. This approach was also used to pool the *Cercopithecus* sample, which comprises blue monkeys (*C. mitis*) and the smaller redtail monkeys (*C. ascanius*). Separately, these two samples are not sufficiently large for statistical purposes. Pooling of the species was deemed reasonable as they are closely related and among the most arboreal of the guenons (Gebo and Sargis 1994; Tosi et al. 2004). A total of 10 taxa was used for analyses.

Taxonomic differences in body size and in the range of size-related variation can also affect patterns and magnitudes of integration and evolvability in the vertebrate skeleton (Hallgrimsson and Maiorana 2000; Marroig and Cheverud 2005). Therefore, all analyses were carried out on VCV matrices estimated from both raw and size-corrected datasets, after the sex and subsample adjustments described above. The scaling method employed scales all

individuals in a group to the same size, and adjusts autopod element size and shape to account for allometric effects (Lleonart et al. 2000). Lleonart et al.'s method was adapted by using the logged geometric mean (GM) of the raw morphological data as the overall size measure, and regressing each logged variable in turn onto this measure (Jungers et al. 1995). The slopes of the regressions represent allometric growth coefficients, and are used to adjust the variables in each individual using the following equation:

$$Y_i^* = Y_i[X_0/X_i]^b \quad (1)$$

where Y_i and X_i are the values of the specific trait and overall size (GM) for individual "i", respectively, Y_i* is the theoretical value for the trait at the group average size, X_0 is the group average size (average GM) and b is the allometric regression slope for each trait (Marroig and Cheverud 2004, p. 269). Adjusted data scale all individuals within a group to the same average body size, but preserve individual element shape covariances. To compare taxa, each individual's adjusted data is divided by its sample's average GM, scaling all individuals in all groups to a theoretical body size of 1. Raw and scaled data were then used to derive phenotypic VCV matrices. Unless noted otherwise, all analyses were carried out using custom-written routines in Matlab v 7.1 (Mathworks, Natick, MA).

Analyses and Hypothesis Testing

Quantifying Autopod Functional Convergence

Testing the hypotheses outlined above requires the ability to quantify the degree of shared function between hands and feet in each taxon. To do so, a set of 20 postural and locomotor behavioral categories was first selected based on the primate behavioral ecology literature (e.g. Ripley 1967; Cant 1987; Cannon and Leighton 1994; Gebo and Chapman 1995; Remis 1995; Hunt et al. 1996; Doran 1997; Isbell et al. 1998; Wells and Turnquist 2001). This dataset captures the diversity of recorded postural and locomotor behaviors in this sample of catarrhines. The 20 behavioral categories were further subdivided into main functions for the hand and foot, producing a dataset comprising a total of 33 autopod functional categories. Note that the total number of autopod functions is less than double the number of behaviors, because hand and foot function in some behaviors is identical. For example, for the posture behavioral category "quadrupedal standing", the function of both autopods is to provide weight support, whereas in the locomotor category "vertical climbing", the hands are used mainly for tensile support (i.e. above-head grasping), while the feet primarily provide propulsion.

Next, the hand and foot in each taxon were scored for each of the 33 functional categories. Methodological differences between studies in the ways behavioral data are reported preclude the use of scores that directly reflect the relative amount of time that each autopod is used in each function. Each functional category for each autopod was instead assigned an integer score between 0 and 4, where 0 means a given autopod function and/or behavior is not observed for that species, and 4 means that autopod function and postural/locomotor behavior are used almost exclusively in the repertoire of a taxon. To obtain an index of functional convergence between the hands and feet, a vector correlation was then computed between the normalized 33-element vectors for the hand and foot. This vector correlation is the cosine of the angle between the normalized hand and foot vectors, and thus measures on a scale of 0 to 1 the similarity of the hand and foot in a 33-dimensional functional multivariate space. This autopod functional convergence index (FCI) will be 1 if the scores of the hand and foot in each category are identical, and will be zero when and only when the hand and foot vectors are orthogonal in functional multivariate space, reflecting hands and feet that do not share any common functions.

Quantifying Integration and Evolvability

Hansen and Houle (2008) have developed a set of tools for quantifying integration and evolvability of multivariate characters, based on Lande's equation characterizing the expected evolutionary response of a multivariate dataset to selection:

$$\Delta z = \mathbf{G}\beta \quad (2)$$

where Δz is a vector of change in the means of the characters between generations, $\mathbf{G}$ is the additive genetic variance/covariance matrix, and β is a selection gradient, a vector of partial regression coefficients of relative fitness on the characters (Lande 1979; Lande and Arnold 1983; Hansen and Houle 2008). As explained above, this study uses the phenotypic VCV matrix in lieu of its genetic counterpart. Two indices of integration and two indices of evolvability were designed by adapting some of Hansen and Houle's measures of evolvability and other measures of integration (Wagner 1984, 1990).

Autopod Integration The variance of the eigenvalues (VE) of the VCV matrices was computed to assess the overall magnitude of integration embodied in each taxon's autopod VCV matrix (Wagner 1984, 1990). VE is a measure of multivariate tightness at the population level: a higher VE indicates that most variance in the data can be accounted for by a small number of principal components with large eigenvalues, corresponding to stronger

covariation between traits (stronger integration). The VE method was designed for correlation matrices in which (co)variances are standardized (Wagner 1984). However, VE can be computed from VCV matrices and made commensurate across species with unequal variances by first dividing the eigenvalues by the trace of the VCV matrix. The sum of the eigenvalues of a VCV matrix, by definition, equals the total variance in the data (the trace). Thus, by standardizing eigenvalues by the trace, each eigenvalue is expressed as a portion of the total variance, and the sum of the adjusted eigenvalues is 1 in all taxa.

VE is predicted to be lowest in taxa whose hands and feet are functionally divergent, reflecting more weakly functionally integrated autopods, and will increase in relation to increasing functional convergence of the hands and feet. In other words, VE will show a positive relationship with the functional convergence index described above. VEs were estimated using both raw and scaled data. Statistical significance of differences in VE between species was determined by bootstrapping the original data, and recomputing VE for each iteration of the bootstrapped sample (Manly 1991). The *P*-value for a pairwise comparison corresponds to the number of times VE in the species with the smaller VE exceeds the bootstrapped values in the species with the larger VE, divided by the number of iterations (1000).

Trait Autonomy As a second measure of integration in catarrhine autopods, I computed the autonomy of each trait with respect to all others in the dataset. Hansen and Houle define 'autonomy' as the fraction of the total genetic (or phenotypic) variance of a trait or set of traits that is independent of potentially constraining traits (i.e. traits with which they covary). The autonomy of a single trait, z_j, with respect to all others is given by:

$$\text{Trait autonomy} = \left(\left[\mathbf{P}^{-1} \right]_{jj} \left[\mathbf{P} \right]_{jj} \right)^{-1} \quad (3)$$

where $\mathbf{P^{-1}}$ is the inverse of the phenotypic VCV matrix, and jj identifies the jth trait along the diagonal of either matrix (Hansen and Houle 2008, p. 1207). Trait autonomy is the fraction of phenotypic variance in each trait that is free of constraints imposed by covariance with other traits. As such, the average autonomy across all traits provides an estimate of how parcellated the dataset is (Wagner and Altenberg 1996): traits that are more integrated will have a greater fraction of their variation potential 'locked up' as covariance with other traits, hence a lower trait autonomy. For each taxon, a trait autonomy index was obtained as the average of the 18 trait autonomies computed using Eq. 3. Trait autonomy is predicted to be lower in species in which the hands and feet are functionally convergent. In other words, the trait autonomy index will show a negative relationship with the functional convergence index. Average trait autonomy was estimated using both raw and scaled datasets. Statistical significance of differences in average trait autonomy between species was evaluated using the same resampling strategy as above.

Directional Evolvability Broadly defined, evolvability is the ability of a population to respond to a selective challenge (Hansen 2003; Hansen and Houle 2008). In the context of Lande's multivariate response to selection equation (Eq. 2), evolvability is the capacity of a population multivariate mean to respond in the direction that selection pushes it, given population-specific constraints on patterns of variation described by its VCV matrix. Mathematically, measures of evolvability capture the extent to which the population mean response vector (Δz) coincides with the selection gradient vector (β) in multivariate space, both in magnitude and direction. Hansen and Houle (2008) define multivariate evolvability, $E(\beta)$, as the length of the projection of the evolutionary response vector (Δz) on the selection gradient (β), divided by the strength of selection (the norm of β):

$$E(\beta) = |\Delta z(\beta)| \cos\theta / |\beta| \quad (4)$$

where $\cos\theta$ is the cosine of the angle between the selection gradient and the response vector in multivariate space. This measure of evolvability is unique to each direction in multivariate phenotypic space, and evolvability is thus different for each β. For example, multivariate evolvability is highest when β points in the direction of the eigenvectors of the VCV matrix, (Schluter 1996; Marroig and Cheverud 2005). The evolutionary potential of a population can be assessed by subjecting its VCV matrix to a number of random selection gradients, and computing the average evolvability over all selection gradients. If the simulated selection gradients are symmetrically distributed in p-dimensional space and the mean gradient length is 1, Hansen and Houle (2008, p. 1206) have determined that the average evolvability of a population in p-dimensional space is equivalent to the average eigenvalue of its VCV matrix, in other words to the average character variance.

Because it is equal to the average trait variance, Hansen and Houle's measure of average evolvability in p-dimensional space correlates strongly with the total amount of variance in a population. As such, this property precludes comparing average evolvabilities among species with unequal total variance. Moreover, unlike VE, average evolvability (i.e. the average eigenvalue) cannot be expressed as a proportion of the trace, as this standardized value would be 1/p in all species. Instead, I used a related measure of evolvability derived from Eq. 4, directional evolvability, which is simply the cosine of the angle between the response vector and the selection gradient in p-space:

$$\text{Directional evolvability} = \cos\theta = |\beta| \cdot |\Delta z(\beta)| \quad (5)$$

This index of evolvability is the portion of Hansen and Houle's evolvability that measures the extent to which the population mean responds in the *direction* of the selection gradient, while ignoring the *magnitude* of the response in this direction (see also Marroig et al., this volume). The index ranges from 0 to 1, where 0 implies that the direction of the evolutionary response vector is orthogonal to a given selection gradient in phenotypic space, and 1 implies that the response is parallel to the direction of selection. Like Eq. 4, the index described by Eq. 5 is unique for different selection gradients in p-dimensional space. The average directional evolvability in response to a set of symmetrically distributed random selection gradients thus gives an estimate of the evolvability of p traits in phenotypic space. The average directional evolvability of the raw and scaled autopod datasets was obtained as follows: the autopod phenotypic VCV matrix in each taxon was subjected to a set of 1000 symmetrically distributed random selection gradients with coefficients between −1 and 1, normalized to a length of 1. For each selection gradient, directional evolvability was computed as the cosine of the angle between the normalized selection gradient ($|\beta|$) and evolutionary response vectors ($|\Delta z(\beta)|$).

The average directional evolvability of a species' autopods—the directional evolvability index—is the average of the 1000 cosines. The statistical significance of this index can be evaluated against a null hypothesis of no common orientation between selection and response vectors, by comparing it against the distribution of vector correlations among 10000 pairs of random p-element vectors (Cheverud and Marroig 2007). If the computed index exceeds 95% of these random correlations, there is significantly more common orientation among the selection and response vectors than expected by chance alone. For 18 elements, 95% of the random vector correlations are 0.39 or less.

A directional evolvability index close to 1 implies that the population mean can respond in any and all directions that selection pushes it in phenotypic space. In contrast, directional evolvability will decrease as the variance embodied in the VCV matrix becomes constrained along fewer principal axes of variation, i.e. as the traits become more strongly integrated. Thus, traits that are more highly integrated should be—as a whole—less evolvable over the entire p-dimensional phenotypic space. With regard to the second hypothesis, this leads to the prediction that evolvability is inversely related to the functional convergence index. Statistical significance of the difference in the evolvability index among species for raw and scaled data was assessed using the same resampling strategy as in the VE method.

Autopod Evolutionary Collinearity The collinearity of the evolutionary responses of the hands and feet to each of the 1000 random selection vectors described above was also evaluated. The dataset comprises two subsets of nine homologous variables measured in the hand and foot (Fig. 1). Thus, each 18-element evolutionary response vector (Δz) can divided into two homologous 9-element hand and foot response vectors, Δz_{hand} and Δz_{foot}, respectively, and these can be normalized and correlated. The resulting vector correlation is the cosine of the angle between the normalized hand and foot response vectors in a phenotypic space comprising nine serially homologous dimensions (e.g. where the first dimension is 'first metapodial length', etc.). This cosine measures the extent to which the evolutionary responses of the hand and foot to random selection gradients are collinear. As such, it is also a measure of how much the magnitude and patterns of covariation between hands and feet constrains their ability to evolve in different directions in multivariate space. Note that although both are based on the cosine between two vectors, unlike the evolvability index, this index is a measure of evolutionary constraint: the higher the cosine, the more the hand and foot respond in a parallel direction, regardless of whether the selection coefficients acting on homologous elements are similar.

The autopod evolutionary constraint index was computed as the cosine of the angle between the normalized hand and foot evolutionary response vectors for each random selection gradient, averaged over the 1000 random selection gradients. The statistical significance of this index was evaluated against the expected distribution of 10,000 pairs of random vector correlations, as above. For 9-element vectors, 95% of the vector correlations are 0.55 or less. The prediction is that species with functionally convergent hands and feet will have a significantly higher constraint index, reflecting the fact that their autopods are more strongly integrated, and hence less independently evolvable. The statistical significance of the predicted differences among species in the evolutionary constraint index—based on raw and scaled data—was assessed using the same resampling strategy as above (1,000 iterations).

Results

Functional Convergence Index

The autopod functional convergence index (FCI) for each taxon is presented in Fig. 2 in relation to the phylogenetic relationships of the sample. FCI ranges from 0 in humans to 0.93 in *Cercopithecus*, and is generally lower in Hominoids than in Cercopithecoids (Fig. 2). An interesting exception is *Pongo*, in which FCI is almost as high as the

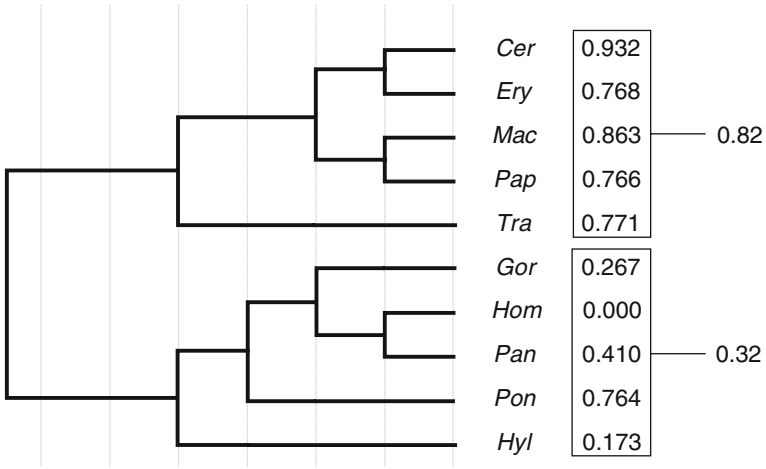

Fig. 2 Autopod functional convergence indices and phylogenetic relationships among the 10 catarrhine taxa used in this study. The phylogenetic tree is derived from Page and Goodman (2001). Branch lengths are not scaled. Numbers on the right are the average functional convergence indices for Cercopithecoids (top) and Hominoids (bottom). Species abbreviations as in Table 1

Cercopithecoids. *Pongo*'s elevated FCI is not surprising when considering that orangutans locomote using a combination of hands and feet for grasping (clambering), and are considered to be the most quadrumanous of the apes (Cant 1987; Rose 1988). In contrast, humans have an index of 0, reflecting an evolutionary history in which the hands and feet presumably underwent selection for functional specializations in the context of manipulative behavior and bipedal locomotion, respectively (Darwin 1871).

Autopod Integration

The index of overall integration among autopod traits, given by the variance of the trace-adjusted eigenvalues of the VCV matrix, is reported for the raw and scaled datasets in Table 2. There are significant differences among species in the overall magnitude of integration. Specifically, species in which the hands and feet are functionally more divergent have less strongly integrated autopods, even after removing the effects of size-related variation (Table 2). Among Hominoids, the orangutan has a relatively high index of integration not significantly different from any of the Cercopithecoids, consistent with its higher FCI (Table 2). In fact, as predicted, there is a significant positive relationship between FCI and VE (Fig. 3a, $r^2 = 0.44$, $P = 0.04$), indicating that the magnitude of covariation within and between autopods decreases as they become increasingly specialized for non-overlapping functions and behaviors. This relationship remains significant after removing the effect of body size-related variation (Fig. 3b, $r^2 = 0.43$, $P = 0.04$).

Trait Autonomy

The average autopod trait autonomy for the raw datasets ranges from 0.12 in *Papio* to 0.28 in *Gorilla* (Table 3). In

Table 2 Pairwise comparisons of raw (below diagonal) and scaled (above diagonal) variance of eigenvalues (VE)

		Hom	*Pan*	*Gor*	*Pon*	*Hyl*	*Mac*	*Pap*	*Cer*	*Ery*	*Tra*
		0.0063	**0.0060**	**0.0089**	**0.0120**	**0.0085**	**0.0091**	**0.0092**	**0.0096**	**0.0110**	**0.0088**
Hom	**0.028**	*Hom*	-	<0.001	<0.001	<0.001	<0.001	0.001	<0.001	<0.001	<0.001
Pan	**0.032**	-	*Pan*	<0.001	<0.001	<0.001	<0.001	0.001	<0.001	<0.001	<0.001
Gor	**0.021**	<0.001	<0.001	*Gor*	0.007	-	-	-	-	0.004	-
Pon	**0.035**	0.025	-	<0.001	*Pon*	0.003	0.01	0.013	0.031	-	0.007
Hyl	**0.029**	-	-	0.001	0.037	*Hyl*	-	-	-	0.002	-
Mac	**0.034**	0.017	-	<0.001	-	0.038	*Mac*	-	-	0.006	-
Pap	**0.041**	<0.001	0.003	<0.001	0.026	<0.001	0.015	*Pap*	-	0.006	-
Cer	**0.032**	-	-	<0.001	-	-	-	0.003	*Cer*	0.018	-
Ery	**0.034**	0.047	-	<0.001	-	-	-	0.012	-	*Ery*	0.003
Tra	**0.037**	0.003	0.048	<0.001	-	0.007	-	-	0.041	-	*Tra*

Species VEs are shown in columns and rows next to their respective taxa. The eigenvalues in each species have been standardized by the trace of that species' VCV matrix prior to estimating VE. The probability shown is the number of times that the species with the smaller VE has a resampled estimate which is larger than the species with the larger VE, divided by the number of replicates (1000)

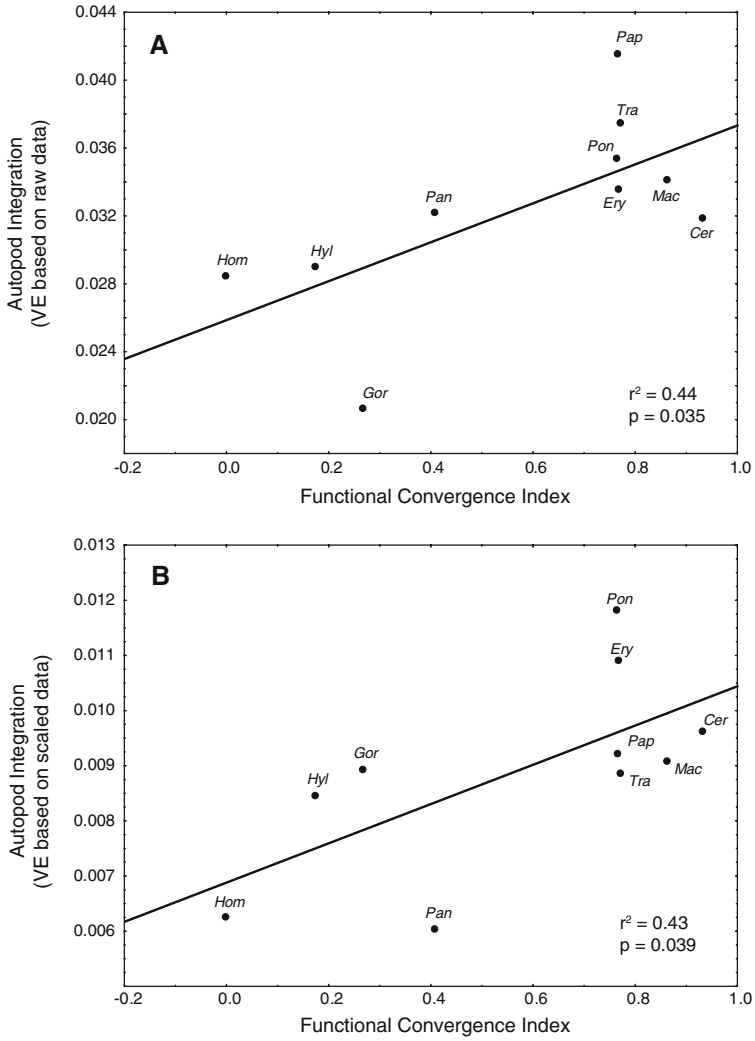

Fig. 3 Catarrhine autopod integration, given by the variance of eigenvalues of the phenotypic VCV matrix, regressed against FCI, based on raw (**a**) and scaled datasets (**b**) (see also Table 2). Species abbreviations as in Table 1

Table 3 Pairwise comparisons of average trait autonomies based on raw (below diagonal) and scaled data (above diagonal)

		Hom	*Pan*	*Gor*	*Pon*	*Hyl*	*Mac*	*Pap*	*Cer*	*Ery*	*Tra*
		0.075	**0.045**	**0.090**	**0.061**	**0.066**	**0.017**	**0.030**	**0.022**	**0.025**	**0.041**
Hom	**0.261**	*Hom*	-	<0.001	<0.001	0.003	-	0.012	-	0.035	-
Pan	**0.252**	-	*Pan*	<0.001	<0.001	0.006	-	0.034	-	-	-
Gor	**0.289**	-	-	*Gor*	-	-	<0.001	-	<0.001	-	0.037
Pon	**0.173**	<0.001	0.019	0.002	*Pon*	-	<0.001	0.013	<0.001	-	0.005
Hyl	**0.219**	-	-	-	-	*Hyl*	<0.001	-	0.018	-	-
Mac	**0.182**	0.001	-	0.005	-	-	*Mac*	0.001	-	0.012	-
Pap	**0.125**	<0.001	<0.001	<0.001	-	0.041	0.049	*Pap*	-	-	-
Cer	**0.165**	<0.001	0.011	0.002	-	-	-	-	*Cer*	-	-
Ery	**0.186**	0.002	-	0.01	-	-	-	-	-	*Ery*	-
Tra	**0.191**	0.002	-	0.018	-	-	-	-	-	-	*Tra*

Average trait autonomies are shown in columns and rows next to their respective taxa. Trait autonomy measures the proportion of phenotypic variance in a trait that is independent of covariance with all other traits in the dataset. The probability shown is the number of times that the species with the smaller average trait autonomy has a resampled estimate which is larger than the species with the larger autonomy, divided by the number of bootstrapped replicates (1000)

other words, in catarrhine autopods, only 12–28% of the average variation of the traits is independent of covariation with all other traits, both within and between autopods. Once the effect of body size-related variation is removed, the average trait autonomy across all taxa decreases to less than 0.05 (range 0.017–0.09), indicating that—on average—less than 5% of the variation in these autopod traits is free of variational constraints imposed by other traits, and by covariation with body size. There are significant differences in the average trait autonomy across catarrhines, however, and species with functionally divergent autopods tend to have higher average trait autonomies (Table 3). Here again the quadrumanous *Pongo* falls within the range of the quadrupedal Cercopithecoids, with an average trait autonomy that is significantly lower than other Hominoids in which the autopods share fewer functions. As predicted, there is a strong negative relationship between FCI and the average trait autonomy (Fig. 4a, $r^2 = 0.65$, $P < 0.005$), indicating that functionally convergent autopods tend to have a smaller fraction of trait phenotypic variation that is independent of covariation with other traits. This negative relationship remains highly significant after removing the effect of body size-related variation across taxa (Fig. 4b, $r^2 = 0.68$, $P < 0.005$).

Directional Evolvability

The directional evolvability index, given by the average of the cosine of the angle between 1,000 random selection gradients and their corresponding evolutionary response vectors, is shown in Table 4. The evolvability index of all taxa—whether based on raw or scaled data—is greater than 0.39, which corresponds to the 95th percentile of the expected distribution of cosines among 10,000 random 18-element vectors. The selection gradients and their evolutionary responses are thus significantly more parallel in

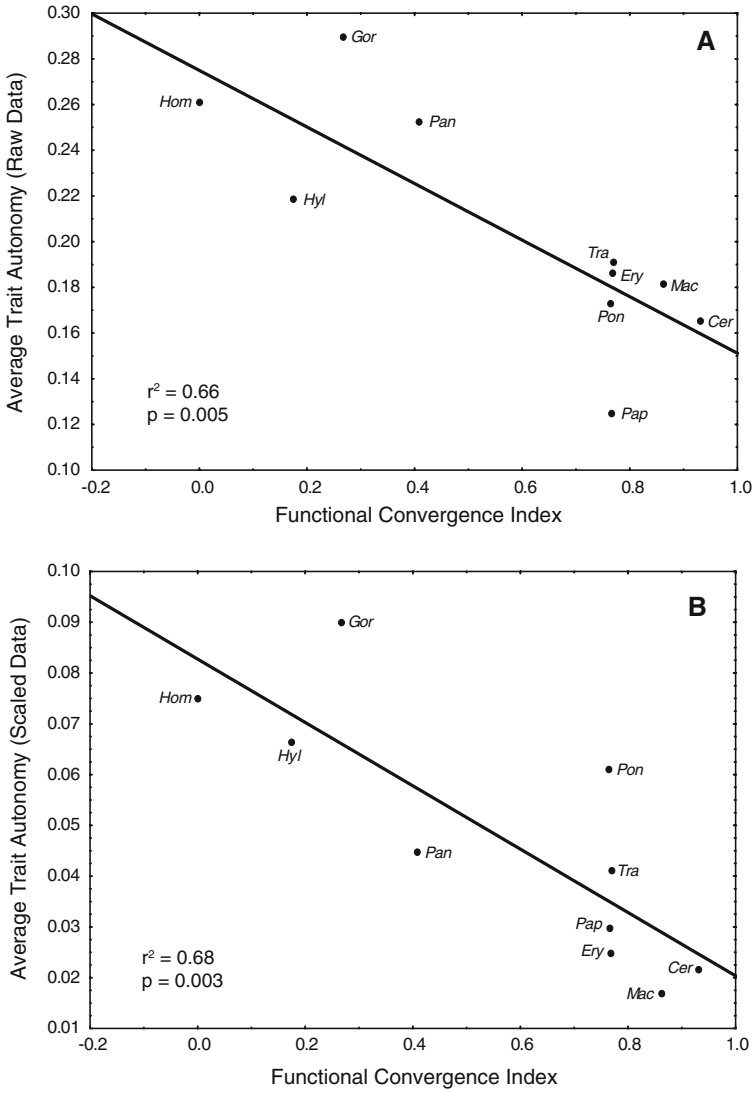

Fig. 4 Catarrhine average autopod trait autonomy, regressed against FCI, based on raw (**a**) and scaled datasets (**b**) (see also Table 3). Trait autonomy is the proportion of variance in a trait that is free of covariance with other traits in the dataset. Species abbreviations as in Table 1

Table 4 Pairwise comparisons of autopod evolvability indices based on raw (below diagonal) and scaled data (above diagonal)

		Hom	*Pan*	*Gor*	*Pon*	*Hyl*	*Mac*	*Pap*	*Cer*	*Ery*	*Tra*
		0.657	**0.675**	**0.638**	**0.587**	**0.602**	**0.654**	**0.634**	**0.637**	**0.616**	**0.557**
Hom	**0.478**	*Hom*	-	-	<0.001	<0.001	-	-	-	0.032	<0.001
Pan	**0.466**	-	*Pan*	-	0.004	0.017	-	-	-	-	<0.001
Gor	**0.556**	<0.001	<0.001	*Gor*	-	-	-	-	-	-	0.005
Pon	**0.466**	-	-	<0.001	*Pon*	-	0.041	-	-	-	-
Hyl	**0.486**	-	-	0.001	-	*Hyl*	-	-	-	-	-
Mac	**0.417**	0.001	0.011	<0.001	-	0.008	*Mac*	-	-	-	0.001
Pap	**0.355**	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	*Pap*	-	-	0.025
Cer	**0.461**	-	-	<0.001	-	-	-	<0.001	*Cer*	-	-
Ery	**0.464**	-	-	<0.001	-	-	-	<0.001	-	*Ery*	-
Tra	**0.415**	0.001	0.01	<0.001	-	0.008	-	0.008	-	-	*Tra*

Species evolvability indices are shown in columns and rows next to their respective taxa. The evolvability index is obtained by computing the cosine of the angle between a random selection gradient vector and its evolutionary response in multivariate space, averaged over the number of random selection gradients applied to a given VCV matrix (1000). The probability shown is the number of times that the species with the smaller evolvability index has a resampled estimate which is larger than the species with the larger index, divided by the number of bootstrapped replicates (1000)

multivariate space than expected by chance alone, indicating that the autopods in all taxa are able to respond to some degree in the direction that selection pushes the population mean in 18-dimensional phenotypic space.

There are, however, differences among taxa in the directional evolvability index based on the raw data, with species like *Gorilla*, *Homo* and *Hylobates* showing significantly higher evolvabilities than the majority of the Cercopithecoids (Table 4). In fact, there is a weak negative relationship between FCI and the evolvability index, although it is not significant (Fig. 5a, $r^2 = 0.34$, $P = 0.08$). This trend suggests that the autopods of species in which the hand and foot have diverged in function are better able to respond to selection pressures in any direction of phenotypic space than species in which the hands and feet are more strongly integrated. Note that this negative trend between FCI and evolvability disappears after controlling for body-size related variation (Fig. 5b, $r^2 = 0.07$, $P = 0.44$), and there are fewer statistically significant differences in evolvability among taxa based on scaled data (Table 4).

Autopod Evolutionary Collinearity

The index of autopod evolutionary constraints is reported in Table 5. This index measures the common directionality in the evolutionary responses of the hands and feet to the same 1,000 random selection vectors used to assess directional evolvability (Table 4). Note that vector correlations among the 9-element hand and foot portions of the random selection gradients are normally distributed with a mean of 0. A little over half of these correlations are negative, reflecting simulated selection pressures that push hand and foot means in opposite directions in phenotypic space. Despite this, indices based on the raw data are high and range from 0.81 to 0.97, significantly greater than the level of 0.55 expected by chance alone. Such high average cosines show that even though selection pressures acting on hand and foot traits may be in opposite directions in phenotypic space, their responses to selection are highly parallel.

There are significant differences among species autopod constraints index, with species such as *Gorilla* and *Homo* having the lowest index, and quadrupedal species like *Cercopithecus, Macaca* and *Papio* having the highest. Interspecific differences in the autopod constraints index based on raw data correspond broadly to differences in the degree of functional convergence among autopods, as shown by the significant positive relationship between FCI and the collinearity index (Fig. 6a, $r^2 = 0.42$, $P = 0.04$). This relationship indicates that—as predicted—species with functionally convergent autopods have hands and feet that show a more strongly coordinated response to selection and are less independently evolvable.

The positive relationship between FCI and the autopod constraints index remains even after removing the effects of body size-related variation (Fig. 6b, $r^2 = 0.43$, $P = 0.04$). In the scaled data, however, the collinearity of the evolutionary responses of the hands and feet is substantially reduced, from an average of 0.9 to 0.6 (Table 5). The index for *Gorilla, Homo* and *Trachypithecus* is less than 0.55, implying that the average cosine of the size-free hand-foot responses in these species is not significantly greater than that expected by chance alone.

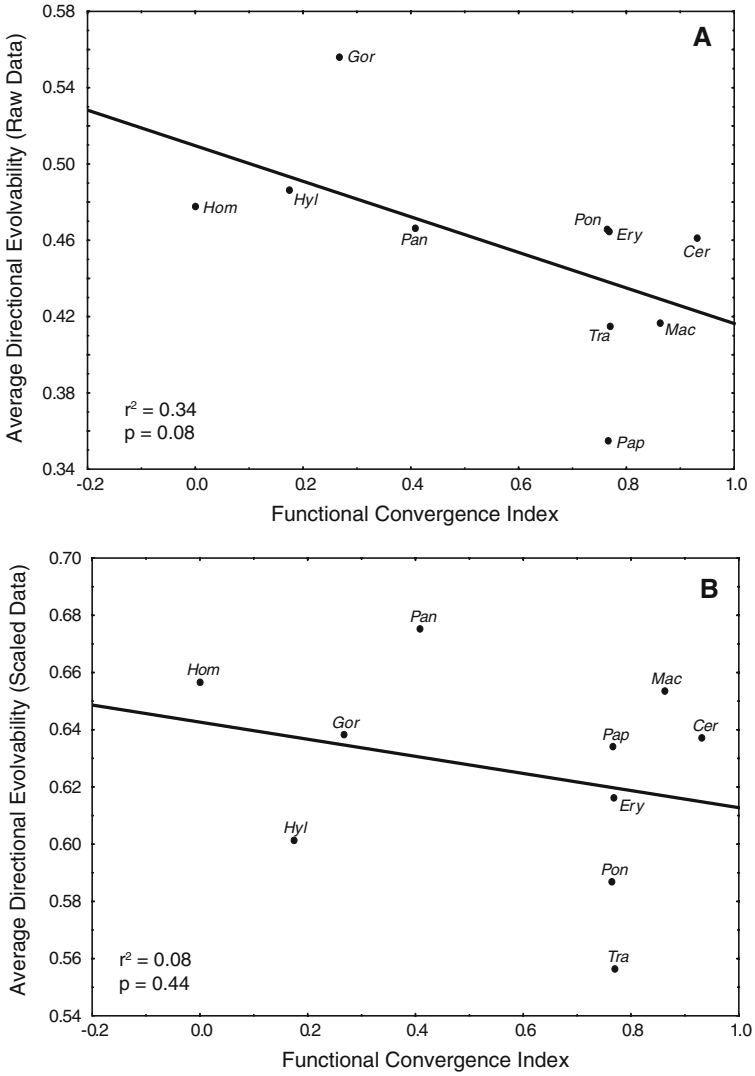

Fig. 5 Catarrhine autopod directional evolvability, regressed against FCI, based on raw (**a**) and scaled datasets (**b**) (see also Table 4). Directional evolvability is the cosine of the angle between a selection gradient and its evolutionary response, and measures the ability of the autopod traits to respond in the direction that selection pushes the population multivariate mean in phenotypic space. Species abbreviations as in Table 1

Discussion

Cheverud's hierarchical model for the evolution of integration predicts that traits that share a common developmental basis or partake in common functions within individuals eventually become genetically integrated at the level of populations, primarily through increasing pleiotropic effects (Cheverud 1996; Wagner 1996). At the population level, the model further predicts that traits that are integrated due to shared development or function co-evolve as a result of a coordinated response to selection, and become less independently evolvable. Importantly, however, integration theory predicts that neither of these relationships is irreversible: traits that exhibit high levels of phenotypic covariation can evolve to be less integrated and more independently evolvable, for example in the context of selection for divergent functions.

This study quantified and compared patterns and magnitudes of integration embodied in the phenotypic variance/covariance matrix of a functionally diverse sample of catarrhine hands and feet, in order to (i) examine the reciprocal relationship between functional and evolutionary integration and (ii) evaluate the impact of integration on the evolvability of morphological traits. Primate autopods are a valuable model for addressing these relationships, because like all serially homologous structures they share the majority of their developmental architecture, yet vary across taxa in the extent to which the hand and foot share common functions. Thus, working from the null hypothesis that primate hands and feet are already largely developmentally and genetically integrated, this study focused on the influence of convergence in function on evolutionary integration and evolvability.

Relationships Between Different Levels of Integration

The results of the integration analyses support the idea that the degree of shared function among organismal traits

Table 5 Pairwise comparisons of autopod evolutionary constraints indices based on raw (below diagonal) and scaled data (above diagonal)

		Hom	*Pan*	*Gor*	*Pon*	*Hyl*	*Mac*	*Pap*	*Cer*	*Ery*	*Tra*
		0.246	**0.595**	**0.504**	**0.589**	**0.691**	**0.753**	**0.670**	**0.683**	**0.648**	**0.532**
Hom	**0.827**	*Hom*	<0.001	0.028	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.002
Pan	**0.906**	0.003	*Pan*	-	-	-	<0.001	-	-	-	-
Gor	**0.815**	-	<0.001	*Gor*	-	0.002	<0.001	0.013	0.005	-	-
Pon	**0.879**	-	-	0.045	*Pon*	-	<0.001	-	-	-	-
Hyl	**0.906**	0.003	-	0.001	-	*Hyl*	-	-	-	-	0.002
Mac	**0.964**	<0.001	0.002	<0.001	<0.001	0.002	*Mac*	-	-	0.017	<0.001
Pap	**0.970**	<0.001	<0.001	<0.001	<0.001	<0.001	-	*Pap*	-	-	0.016
Cer	**0.921**	0.003	-	<0.001	-	-	0.004	0.002	*Cer*	-	0.005
Ery	**0.913**	0.033	-	0.015	-	-	0.002	<0.001	-	*Ery*	-
Tra	**0.875**	0.041	-	0.011	-	-	<0.001	<0.001	-	-	*Tra*

The evolutionary constraint indices are shown in columns and rows next to their respective taxa. The autopod evolutionary constraints index is obtained by computing the average cosine of the angle between the vectors of the hand and foot evolutionary responses to the same 1000 random selection gradients that were applied to obtain the evolvability index. The probability shown is the number of times that the species with the smaller index has a resampled estimate which is larger than the species with the larger index, divided by the number of bootstrapped replicates (1000)

within individuals drives evolutionary integration at the population level. The strength of integration in the VCV matrix of each catarrhine taxon was estimated in two complementary ways: the variance of the eigenvalues of the VCV matrices provides a measure of how evenly the variation in each dataset is distributed among principal components, while the average trait autonomy describes the average portion of variance in autopod traits that is independent of covariance with all other traits. Both measures of integration show a significant relationship with the degree of functional convergence of the autopods (Figs. 3 and 4): species in which the autopods were found to be more convergent in their respective functions in postural and locomotor behaviors have the highest VE and the lowest trait autonomy (Tables 2 and 3). Moreover, analyses based on scaled data show that these relationships are not due to interspecific differences in mean body size and size-related variation.

These results show that the skeletal elements of functionally convergent autopods covary more strongly in relation to their means, and have more variation 'locked up' as covariation with other autopod skeletal elements. From a functional perspective, such strong covariation among autopods in quadrupedal and quadrumanous species may have been maintained or selected for because it improves performance in shared functions such as grasping and propulsion on identical substrates. Conversely, strong integration and reduced autonomy were likely selected against in species like *Gorilla* and *Homo*, if the performance of increasingly specialized functions and behaviors (e.g. bipedalism vs manipulation) was adversely affected by such strong covariation in size and shape between homologous autopod elements.

Note that differences in integration and autonomy often parallel the phylogenetic division between Cercopithecoids and Hominoids (Tables 2 and 3). This is not entirely surprising, as functional differences also largely parallel this division (Fig. 2, Fleagle 1999). However, *Pongo*, in which the autopods show the greatest convergence in function among Hominoids (Fig. 2, Cant 1987), also has the highest VE and lowest autonomy among Hominoids, and even among some Cercopithecoids whose autopods are also hypothesized to be functionally integrated. This pattern that transcends phylogenetic relationships suggests that differences in integration and autonomy are driven more by functional effects than by phylogenetic heritage. More importantly, relatively closely related species like *Homo* and *Pongo*, at opposite ends of both the functional similarity and the integration/autonomy spectrums, illustrate the point that morphological integration is evolutionarily reversible, and that patterns and magnitudes of covariation in the hands and feet are relatively labile, changing in accordance with the degree of functional convergence among these structures.

These results are consistent with previous analyses of integration in the vertebrate postcranium (e.g. Hallgrimsson et al. 2002; Young and Hallgrimsson 2005). For example, Young and Hallgrimsson (2005) examined the relationship between serial homology, functional divergence and integration in the limbs of a phylogenetically and functionally diverse sample of mammals. Although the degree of functional convergence between limbs was not

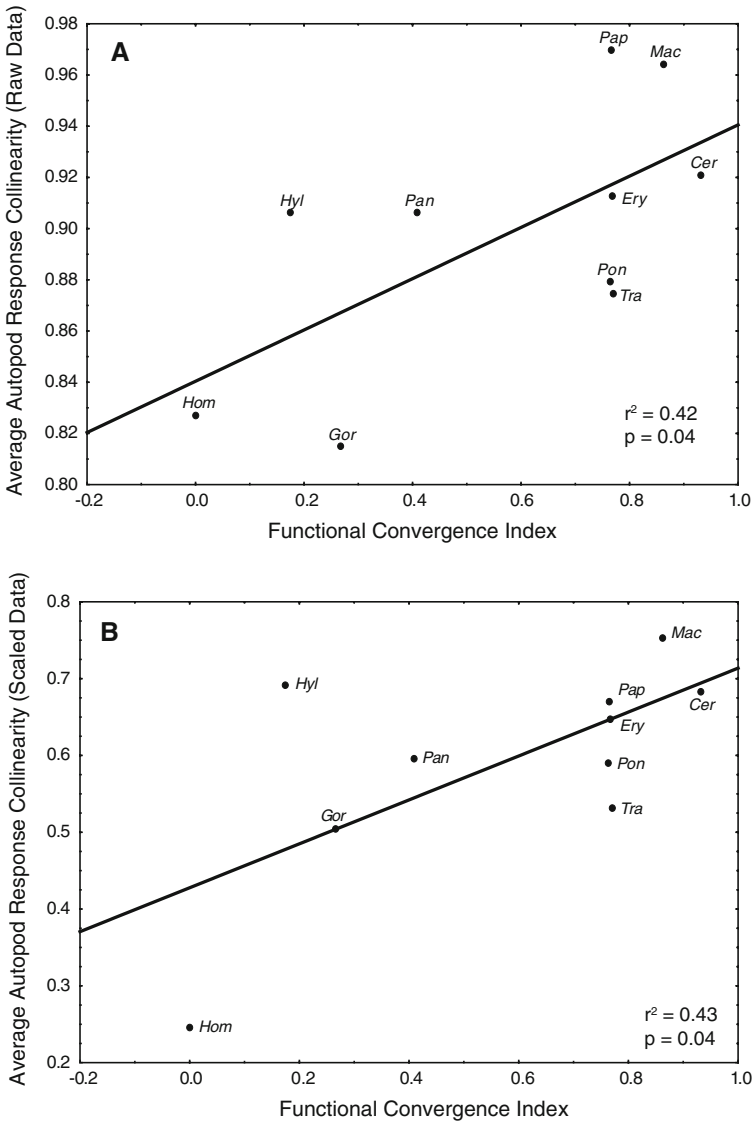

Fig. 6 Catarrhine autopod evolutionary collinearity, regressed against FCI, based on raw (**a**) and scaled datasets (**b**) (see also Table 5). Evolutionary collinearity is the cosine of the angle between the hand and foot evolutionary response vectors subjected to a random selection gradient, and measures the ability of the hands and feet to respond in the direction that selection pushes their respective population multivariate means in phenotypic space. Species abbreviations as in Table 1

quantified, and the dataset contained a single element from each autopod, Young and Hallgrimsson found that covariation among serially homologous limb elements was weaker in species whose limbs were functionally divergent (e.g. bats, gibbons). The results of the present study extend these findings to a larger sample of autopod elements, and further confirm the reversible relationship between functional and evolutionary integration in a sample of ten closely related species. This last point is important, as significant differences between even closely related species like orangutans, African apes and humans (Tables 2 and 3) suggest that patterns and magnitudes of integration evolve rapidly in the context of species' functional diversification.

Relationship between Integration and Evolvability

The second objective of this study was to assess the impact of integration on the evolvability of the autopods. Evolvability captures the ability of a population mean to respond to a selective challenge. Here, the evolvability of catarrhine autopods was assessed by comparing how the multivariate mean of each taxon responds to simulated selection. Thousand random selection gradients—symmetrically distributed in phenotypic space—were applied to each taxon's phenotypic VCV matrix and the responses compared. Evolvability was evaluated in two complementary ways: the average cosine of the angle between the selection vectors and their corresponding evolutionary responses measures the ability of the autopod trait means to respond in the direction that selection pushes them (directional evolvability), while the average cosine of the angle between hand and foot response vectors measures the extent to which mean responses in hand and foot homologous traits are parallel in phenotypic space.

The directional evolvability index based on raw data shows a weak negative relationship with the functional

convergence index, and several species with functionally divergent autopods had significantly greater average evolvabilities than species hypothesized to be more functionally integrated (Table 4). These data suggest that autopods that are more weakly integrated are better able to respond to selective challenges, no matter what direction selection pushes the mean. On its own, however, this finding only shows that the autopods are collectively more evolvable, and does not prove that the functionally specialized hands and feet of species like *Homo* and *Hylobates* are more independently evolvable. Indeed, the observed reduction in phenotypic covariation and integration in these species could relate as much to within-autopod covariation as to covariation among autopods.

The results of the autopod evolutionary collinearity analysis suggest, however, that covariation among autopods is in fact reduced in taxa with functionally divergent autopods. There is a significant positive relationship between the collinearity index and the functional convergence index based on raw and scaled data (Table 5, Fig. 6), indicating that the hands and feet of species like *Homo*, *Gorilla* and *Pan* are better able to respond in the directions that selection pushes each autopod multivariate mean, regardless of whether these directions coincide. This pattern is more apparent once the effects of size-related variation are removed. Based on scaled data, most Cercopithecoid hands and feet still show more parallel responses than expected by chance alone, while several Hominoids have average cosines that are near or below the 95th percentile (Table 5).

Interestingly, size-related variation significantly impacts the evolvability of the autopods. For example, the directional evolvability of the autopod dataset increased in all taxa once the effects of size-related variation were removed (Table 4). This increase in evolvability suggests that body-size related variation acts as a form of constraint, channeling evolutionary responses along this axis of variation, and making other evolutionary directions in phenotypic space less accessible. The autopod evolutionary collinearity data are consistent with this interpretation. When size-related variation is included, hand and foot responses are highly parallel across taxa, even though over half the selection gradients pushed hand and foot multivariate means in opposite directions. Collinearity in autopod responses is explained by the fact that size-related variation channels evolutionary responses along its main axis of variation in phenotypic space. Once size-related variation is removed, other directions in phenotypic space are opened up, and the hand and foot trait means are better able to respond in the direction that selection pushes them individually. This view is consistent with the idea that—at least for traits that describe morphological size and shape—body size-related variation determines the path of least evolutionary resistance in phenotypic space (Schluter 1996; Marroig and Cheverud 2005).

Notwithstanding the influence of size-related variation on evolvability, the data suggest that skeletal traits that are more strongly integrated are less evolvable, collectively and individually (Tables 4 and 5). It is important to note, however, that average evolvability was measured as the capacity to respond to selection in *any* direction in phenotypic space. Most of these simulated directions probably don't reflect selective pressures that natural populations would experience. Further, even though they are—on average—less evolvable over the entire phenotypic space, more strongly integrated traits may be more evolvable, both in terms of the magnitude and direction of the evolutionary response, along particular directions of this space, notably those that coincide with the orientation of the leading eigenvectors of their VCV matrix (Porto et al. 2009).

The Paradox of Pleiotropy

The analyses presented here suggest that factors that increase covariation among traits increase their ability to respond to selection in a coordinated manner, and limit the ability of each trait to respond individually to potentially different selective pressures. A morphological trait's capacity to evolve in response to selection thus depends in large part on how strongly it covaries with other such traits and with body size. Mechanistically, evolutionary changes in the magnitude of integration are thought to be achieved through natural selection acting on the phenotypic expression of genetic variation in pleiotropy (Cheverud 1996; Pavlicev et al. 2008). For example, pleiotropic interactions are expected to be selected against if the covariation they induce at the level of the phenotype reduces fitness, and vice versa (Wagner 1996; Wagner and Altenberg 1996).

Pleiotropic effects are well documented in the development of serially homologous structures like vertebrate autopods, with over 90% of gene transcripts simultaneously expressed in embryonic murine fore- and hindlimbs (Margulies et al. 2001; Shou et al. 2005). At the same time, however, changes in the extent of pleiotropic effects are thought to be the main mechanism by which both integration and parcellation evolve. At face value, this suggests that pleiotropy plays contradictory roles in the evolution of integration and parcellation in serially homologous structures. For example, if the autopods evolve convergent functions, as in Cercopithecoids and *Pongo*, then pleiotropic effects already present as a result of serial homology may facilitate correlated evolution of the hands and feet. Conversely, in evolutionary contexts where the hands and feet become functionally and morphologically divergent, as in the remaining Hominoids, the existence of such extensive pleiotropy could be seen as an obstacle to their independent evolution.

The diversity of specialized functions and uniquely derived morphologies in this sample of catarrhine autopods reveals, however, that pleiotropic interactions during the development of serially homologous structures aren't a significant obstacle to their independent evolution. How then does morphological and functional diversity evolve among structures that use virtually identical genetic and developmental pathways? The answer to this question is not yet known, but could involve a relative increase in the importance a small number of key genes that are uniquely expressed in each limb, such as members of the T-box family of DNA-binding proteins (Khan et al. 2002; Rallis et al. 2003). Alternatively, or additionally, differences in the patterns and magnitude of covariation among serially homologous structures could be achieved not through different pathways, but through the differential regulation of common developmental pathways, including the differential control of gene expression via tissue-specific enhancers (Shapiro et al. 2004; Menke et al. 2008). Either situation suggests that genetic variation in pleiotropy is not the only developmental mechanism driving the evolution of integration and parcellation. Future research into the evolution and tissue-specific roles of regulatory elements may improve our knowledge of how integration, evolvability and morphological diversity can evolve among serially homologous structures in complex organisms (Rudel and Sommer 2003; Wilkins 2002).

Acknowledgements Thanks to my co-organizer Kat Willmore, who helped put together the AAPA symposium at which some of these results were presented. Thanks also to Benedikt Hallgrimsson, for the opportunity to publish proceedings from the symposium in a special issue of the journal Evolutionary Biology. This research was supported by a National Science Foundation Doctoral Dissertation Improvement Grant (BCS 0647624) and a Canadian Natural Sciences and Engineering Research Council Postgraduate Scholarship. I am grateful to J. Chupasko (Museum of Comparative Zoology), D. Dunbar and T. Kensler (Laboratory for Primate Morphology and Genetics, University of Puerto Rico), L. Jellema (Cleveland Museum of Natural History), L. Gordon and D. Hunt (National Museum of Natural History), E. Westwig (American Museum of Natural History), M. Tappen and J. Soderberg (University of Minnesota), S. Leigh and J. Polk (University of Illinois - Urbana-Champaign), M. Harman (Powell-Cotton Museum), D. Hills (Natural History Museum, London), M. Hiermeier (Bavarian Zoological State Collection) and L. Shapiro (University of Texas—Austin) for providing access to specimens in their care. Finally, thanks to my dissertation committee members, D. Lieberman, D. Pilbeam, G. Lauder and B. Hallgrimsson, for helping me through the design, implementation and analysis of this project, and for reading earlier incarnations of the manuscript.

References

Ackermann, R. R., & Cheverud, J. M. (2000). Phenotypic covariance structure in tamarins (genus *Saguinus*): A comparison of variation patterns using matrix correlation and common principal component analysis. *American Journal of Physical Anthropology, 111*, 489–501. doi:10.1002/(SICI)1096-8644(200004)111:4<489::AID-AJPA5>3.0.CO;2-U.

Ackermann, R. R., & Cheverud, J. M. (2004). Detecting genetic drift versus selection in human evolution. *Proceedings of the National Academy of Sciences of the United States of America, 101*, 17946–17951. doi:10.1073/pnas.0405919102.

Cannon, C. H., & Leighton, M. (1994). Comparative locomotor ecology of gibbons and macaques—selection of canopy elements for crossing gaps. *American Journal of Physical Anthropology, 93*, 505–524. doi:10.1002/ajpa.1330930409.

Cant, J. G. H. (1987). Positional behavior of female Bornean orangutans (*Pongo pygmaeus*). *American Journal of Primatology, 12*, 71–90. doi:10.1002/ajp.1350120104.

Cheverud, J. M. (1982). Phenotypic, genetic, and environmental morphological integration in the cranium. *Evolution: International Journal of Organic Evolution, 36*, 499–516. doi:10.2307/2408096.

Cheverud, J. M. (1988). A comparison of genetic and phenotypic correlations. *Evolution: International Journal of Organic Evolution, 42*, 958–968. doi:10.2307/2408911.

Cheverud, J. M. (1995). Morphological integration in the saddle-back tamarin (*Saguinus fuscicollis*) cranium. *American Naturalist, 145*, 63–89. doi:10.1086/285728.

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

Cheverud, J. M., & Marroig, G. (2007). Comparing covariance matrices: Random skewers method compared to the common principal components model. *Genetics and Molecular Biology, 30*, 461–469. doi:10.1590/S1415-47572007000300027.

Darwin, C. (1859). *On the origin of species by means of natural selection*. London: J. Murray.

Darwin, C. (1871). *The descent of man and selection in relation to sex*. New York: D. Appleton and company.

Doran, D. (1997). Ontogeny of locomotion in mountain gorillas and chimpanzees. *Journal of Human Evolution, 32*, 323–344.

Fleagle, J. G. (1999). *Primate adaptation and evolution*. San Diego: Academic Press.

Gebo, D. L., & Chapman, C. A. (1995). Positional behavior in 5 sympatric old-world monkeys. *American Journal of Physical Anthropology, 97*, 49–76. doi:10.1002/ajpa.1330970105.

Gebo, D. L., & Sargis, E. J. (1994). Terrestrial adaptations in the postcranial skeletons of guenons. *American Journal of Physical Anthropology, 93*, 341–371. doi:10.1002/ajpa.1330930306.

Hall, B. K. (1995). Homology and embryonic development. *In "Evolutionary Biology Vol 28", 28*, 1–37.

Hallgrimsson, B., & Maiorana, V. (2000). Variability and size in mammals and birds. *Biological Journal of the Linnean Society. Linnean Society of London, 70*, 571–595. doi:10.1111/j.1095-8312.2000.tb00218.x.

Hallgrimsson, B., Willmore, K., & Hall, B. (2002). Canalization, developmental stability, and morphological integration in primate limbs. *Yearbook of Physical Anthropology, 45*, 131–158. doi:10.1002/ajpa.10182.

Hansen, T. F. (2003). Is modularity necessary for evolvability? Remarks on the relationship between pleiotropy and evolvability. *Bio Systems, 69*, 83–94. doi:10.1016/S0303-2647(02)00132-6.

Hansen, T. F., & Houle, D. (2008). Measuring and comparing evolvability and constraint in multivariate characters. *Journal of Evolutionary Biology, 21*, 1201–1219. doi:10.1111/j.1420-9101.2008.01573.x.

Hunt, K. D., Cant, J. G. H., Gebo, D. L., Rose, M. D., Walker, S. E., & Youlatos, D. (1996). Standardized descriptions of primate locomotor and postural modes. *Primates, 37*, 363–387. doi:10.1007/BF02381373.

Isbell, L. A., Pruetz, J. D., Lewis, M., & Young, T. P. (1998). Locomotor activity differences between sympatric patas

monkeys (*Erythrocebus patas*) and vervet monkeys (*Cercopithecus aethiops*): Implications for the evolution of long hindlimb length in *Homo*. *American Journal of Physical Anthropology, 105*, 199–207. doi:10.1002/(SICI)1096-8644(199802)105:2<199::AID-AJPA7>3.0.CO;2-Q.

Jouffroy, F. K., Godinot, M., & Nakano, Y. (1991). Biometrical characteristics of primate hands. *Human Evolution, 6*, 269–306. doi:10.1007/BF02437254.

Jungers, W. L., Falsetti, A. B., & Wall, C. E. (1995). Shape, relative size, and size adjustments in morphometrics. *Yearbook of Physical Anthropology, 38*, 137–161. doi:10.1002/ajpa.1330380608.

Khan, P., Linkhart, B., & Simon, H. (2002). Different regulation of T-box genes *Tbx4* and *Tbx5* during limb development and limb regeneration. *Developmental Biology, 250*, 383–392.

Lande, R. (1979). Quantitative genetic-analysis of multivariate evolution, applied to brain—body size allometry. *Evolution: International Journal of Organic Evolution, 33*, 402–416. doi:10.2307/2407630.

Lande, R. (1980). The genetic covariance between characters maintained by pleiotropic mutations. *Genetics, 94*, 203–215.

Lande, R., & Arnold, S. J. (1983). The measurement of selection on correlated characters. *Evolution: International Journal of Organic Evolution, 37*, 1210–1226. doi:10.2307/2408842.

Lleonart, J., Salat, J., & Torres, G. J. (2000). Removing allometric effects of body size in morphological analysis. *Journal of Theoretical Biology, 205*, 85–93. doi:10.1006/jtbi.2000.2043.

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

Manly, B. F. J. (1991). *Randomization and Monte Carlo methods in biology*. London: Chapman and Hall.

Margulies, E. H., Kardia, S. L. R., & Innis, J. W. (2001). A comparative molecular analysis of developing mouse forelimbs and hindlimbs using Serial Analysis of Gene Expression (SAGE). *Genome Research, 11*, 1686–1698. doi:10.1101/gr.192601.

Marroig, G., & Cheverud, J. M. (2004). Cranial evolution in sakis (Pithecia, Platyrrhini) I: Interspecific differentiation and allometric patterns. *American Journal of Physical Anthropology, 125*, 266–278.

Marroig, G., & Cheverud, J. M. (2005). Size as a line of least evolutionary resistance: Diet and adaptive morphological radiation in New World monkeys. *Evolution: International Journal of Organic Evolution, 59*, 1128–1142.

McFadden, D., & Bracht, M. S. (2003). The relative lengths and weights of metacarpals and metatarsals in baboons (*Papio hamadryas*). *Hormones and Behavior, 43*, 347–355. doi:10.1016/S0018-506X(02)00048-X.

McFadden, D., & Bracht, M. S. (2005). Sex differences in the relative lengths of metacarpals and metatarsals in gorillas and chimpanzees. *Hormones and Behavior, 47*, 99–111. doi:10.1016/j.yhbeh.2004.08.013.

Menke, D. B., Guenther, C., & Kingsley, D. M. (2008). Dual hindlimb control elements in the *Tbx4* gene and region-specific control of bone size in vertebrate limbs. *Development, 135*, 2543–2553. doi:10.1242/dev.017384.

Olson, E. C., & Miller, R. L. (1951). A mathematical model applied to a study of the evolution of species. *Evolution: International Journal of Organic Evolution, 5*, 325–338. doi:10.2307/2405677.

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

Page, S. L., & Goodman, M. (2001). Catarrhine phylogeny: Noncoding DNA evidence for a diphyletic origin of the mangabeys and for a human-chimpanzee clade. *Molecular Phylogenetics and Evolution, 18*, 14–25. doi:10.1006/mpev.2000.0895.

Pavlicev, M., Kenney-Hunt, J. P., Norgard, E. A., Roseman, C. C., Wolf, J. B., & Cheverud, J. M. (2008). Genetic variation in pleiotropy: Differential epistasis as a source of variation in the allometric relationship between long bone lengths and body weight. *Evolution: International Journal of Organic Evolution, 62*, 199–213.

Porto, A., de Oliveira, F. B., Shirai, L. T., De Conto, V., & Marroig, G. (2009). The evolution of modularity in the mammalian skull I: Morphological integration patterns and magnitudes. *Evolutionary Biology*. doi:10.1007/s11692-008-9038-3.

Rallis, C., Bruneau, B. G., Del Buono, J., Seidman, C. E., Seidman, J. G., Nissim, S., et al. (2003). *Tbx5* is required for forelimb bud formation and continued outgrowth. *Development, 130*, 2741–2751. doi:10.1242/dev.00473.

Remis, M. (1995). Effects of body-size and social-context on the arboreal activities of lowland gorillas in the Central-African-Republic. *American Journal of Physical Anthropology, 97*, 413–433. doi:10.1002/ajpa.1330970408.

Ripley, S. (1967). Leaping of Langurs—a problem in study of locomotor adaptation. *American Journal of Physical Anthropology, 26*, 149. doi:10.1002/ajpa.1330260206.

Roff, D. A. (1995). The estimation of genetic correlations from phenotypic correlations—a test of Cheverud's conjecture. *Heredity, 74*, 481–490. doi:10.1038/hdy.1995.68.

Roff, D. A. (1996). The evolution of genetic correlations: An analysis of patterns. *Evolution: International Journal of Organic Evolution, 50*, 1392–1403. doi:10.2307/2410877.

Rohlf, F. J. (2005). *TPSDig2*. Stony Brook: State University of New York.

Rose, M. D. (1988). Functional anatomy of the cheiridia. In J. H. Schwartz (Ed.), *Orangutan biology* (pp. 299–310). Oxford: Oxford University Press.

Rudel, D., & Sommer, R. J. (2003). The evolution of developmental mechanisms. *Developmental Biology, 264*, 15–37. doi:10.1016/S0012-1606(03)00353-1.

Ruvinsky, I., & Gibson-Brown, J. J. (2000). Genetic and developmental bases of serial homology in vertebrate limb evolution. *Development, 127*, 5233–5244.

Schluter, D. (1996). Adaptive radiation along genetic lines of least resistance. *Evolution: International Journal of Organic Evolution, 50*, 1766–1774. doi:10.2307/2410734.

Shapiro, M. D., Marks, M. E., Peichel, C. L., Blackman, B. K., Nereng, K. S., Jonsson, B., et al. (2004). Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature, 428*, 717–723. doi:10.1038/nature02415.

Shou, S. M., Scott, V., Reed, C., Hitzemann, R., & Stadler, H. S. (2005). Transcriptome analysis of the murine forelimb and hindlimb autopod. *Developmental Dynamics, 234*, 74–89. doi:10.1002/dvdy.20514.

Stern, J. T., & Oxnard, C. (1973). Primate locomotion: Some links with evolution and morphology. *Primatologia, 4*, 1–93.

Strasser, E. (1994). Relative development of the hallux and pedal digit formulas in Cercopithecidae. *Journal of Human Evolution, 26*, 413–440. doi:10.1006/jhev.1994.1026.

Susman, R. L. (1979). Comparative and functional morphology of hominoid fingers. *American Journal of Physical Anthropology, 50*, 215–236. doi:10.1002/ajpa.1330500211.

Tosi, A. J., Melnick, D. J., & Disotell, T. R. (2004). Sex chromosome phylogenetics indicate a single transition to terrestriality in the guenons (tribe Cercopithecini). *Journal of Human Evolution, 46*, 223–237. doi:10.1016/j.jhevol.2003.11.006.

Wagner, G. P. (1984). On the eigenvalue distribution of genetic and phenotypic dispersion matrices—evidence for a nonrandom

organization of quantitative character variation. *Journal of Mathematical Biology, 21*, 77–95.

Wagner, G. P. (1990). A comparative study of morphological integration in *Apis mellifera* (Insecta, Hymenoptera). *Zeitschrift fur Zoologische Systematik und Evolutionsforschung, 28*, 48–61.

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: International Journal of Organic Evolution, 50*, 967–976. doi:10.2307/2410639.

Wells, J. P., & Turnquist, J. E. (2001). Ontogeny of locomotion in rhesus macaques (*Macaca mulatta*): II. Postural and locomotor behavior and habitat use in a free-ranging colony. *American Journal of Physical Anthropology, 115*, 80–94. doi:10.1002/ajpa.1059.

Wilkins, A. S. (2002). *The evolution of developmental pathways*. Sunderland: Sinauer Associates.

Willis, J. H., Coyne, J. A., & Kirkpatrick, M. (1991). Can one predict the evolution of quantitative characters without genetics? *Evolution: International Journal of Organic Evolution, 45*, 441–444. doi:10.2307/2409678.

Young, N. M., & Hallgrimsson, B. (2005). Serial homology and the evolution of mammalian limb covariation structure. *Evolution: International Journal of Organic Evolution, 59*, 2691–2704.