

Developmental regulation of skull morphology II: ontogenetic dynamics of covariance

Miriam Leah Zelditch,^{a,*} Jason Mezey,^b H. David Sheets,^c Barbara L. Lundrigan,^d and Theodore Garland Jr.^e

^aMuseum of Paleontology, University of Michigan, Ann Arbor, MI 48109, USA

^bCenter for Population Biology, University of California, Davis, CA 95616, USA

^cDepartment of Physics, Canisius College, Buffalo, NY 14208, USA

^dMichigan State University Museum and Department of Zoology, Michigan State University, East Lansing, MI 48824, USA

^eDepartment of Biology, University of California, Riverside, CA 92521, USA

*Author for correspondence (email: zelditch@umich.edu)

SUMMARY Canalization may play a critical role in molding patterns of integration when variability is regulated by the balance between processes that generate and remove variation. Under these conditions, the interaction among those processes may produce a dynamic structure of integration even when the level of variability is constant. To determine whether the constancy of variance in skull shape throughout most of postnatal growth results from a balance between processes generating and removing variation, we compare covariance structures from age to age in two rodent species, cotton rats (*Sigmodon fulviventer*) and house mice (*Mus musculus domesticus*). We assess the overall similarity of covariance matrices by the matrix correlation, and compare the structures of covariance matrices using common

subspace analysis, a method related to common principal components (PCs) analysis but suited to cases in which variation is so nearly spherical that PCs are ambiguous. We find significant differences from age to age in covariance structure and the more effectively canalized ones tend to be least stable in covariance structure. We find no evidence that canalization gradually and preferentially removes deviations arising early in development as we might expect if canalization results from compensatory differential growth. Our results suggest that (co)variation patterns are continually restructured by processes that equilibrate variance, and thus that canalization plays a critical role in molding patterns of integration.

INTRODUCTION

Canalized developmental processes yield consistent phenotypes regardless of variation in genotypes or environmental conditions encountered over the course of development (Waddington 1942, 1952). Canalization is usually regarded as a process-limiting variation, but it also influences the structure of variation for two related reasons. First, the structure is determined by the balance between processes generating and reducing variance; were variance generated but not also removed, it would accumulate over developmental time. However, at least in the case of rodent skull shape, variance diminishes by 50% early in postnatal growth and then remains nearly constant (Zelditch et al. 2004; Zelditch 2005). Second, the structure of variance is determined by the spatial distribution of the processes that produce and remove variance; as long as the processes removing variance differ in spatial structure from those generating it, the interaction between the two classes of processes, rather than either of them

separately, determines the spatial structure of variation. The structure of variation is usually analyzed in terms of morphological integration or modularity. The first connotes the interdependences among parts because of developmental and/or functional interactions among them (e.g., Olson and Miller 1958; Cheverud 1982; Zelditch 1988); the second refers to the independence among internally cohesive, integrated units (Shlosser and Wagner 2004). To the extent that processes canalizing form interact continually with those generating variation, canalization not only reduces variance, it also plays an ongoing role in structuring integration.

The developmental basis of integration (and modularity) has long been of interest to evolutionary biologists because complex morphologies such as the mammalian skull comprise multiple functionally interdependent units. Should functionally interdependent parts be disproportionate, their ability to perform tasks essential for survival might be compromised, as would be the case if a very long upper jaw is associated with a very short lower jaw. Not surprisingly, canalization and mor-

phological integration (and/or modularity) are thought to evolve by stabilizing selection. In theory, the pattern of (co)variation is expected to evolve to match the fitness surface (e.g., Cheverud 1984, 1996), and alleles whose pleiotropic effects cause interdependent parts to attain appropriate proportions will generally be favored by natural selection. At genetic equilibrium, we therefore expect that individuals with atypically long lower jaws would have atypically long upper jaws because of adaptive pleiotropy. Because pleiotropy would ensure that jaw proportions are normal, jaw shape would not vary, so shape is canalized. Stabilizing selection, however, might not be necessary for the evolution of canalization; in theory, it can evolve even in the absence of stabilizing selection (Siegal and Bergman 2002) and epigenetic interactions, by themselves, could explain the covariation among phenotypic traits. Such covariation would be expected when two or more skeletal traits develop in response to the same soft tissues (e.g., brain, muscles). Epigenetic interactions may also reduce the variance of skeletal shape, the hypothesis we advanced to explain the early postnatal reduction of variance in skull shape (Zelditch et al. 2004). That reduction would require compensatory differential growth, which corrects for prior disproportionate growth. According to this hypothesis, rather than growing at coordinated rates, bones attain normal proportions by compensatory differential growth.

Given that the level of variation is constant over much of postnatal skull development in all rodents examined to date (Zelditch et al. 2004; Zelditch 2005), it might appear that compensatory differential growth plays an important role only early in the developmental process, that is, during the interval in which variance decreases by 50%. Additionally, previous studies have found that the structure of integration changes significantly only during that same interval (Zelditch 1988; Zelditch and Carmichael 1989). Thus, it would appear that compensatory differential growth has a minor role in structuring adult patterns of integration beyond that initial phase of postnatal growth. However, compensatory differential growth reduces variance only if the rate at which new variance is generated is lower than the rate at which variance diminishes. If new variance is continually generated, at the same rate as variance is compensated, the level could be constant. That balance between processes generating and removing variance could maintain variance in dynamic equilibrium but nevertheless alter its structure. The primary question addressed herein is whether the constancy in variance exhibited by rodent skulls following the initial reduction in variance results from coordinated growth that forestalls new variation from being generated, or instead from a balance between processes generating and removing variation.

To distinguish between these two hypotheses for the constancy of variance, we examine the ontogenetic dynamics of (co)variances, testing two hypotheses: (1) the structure of

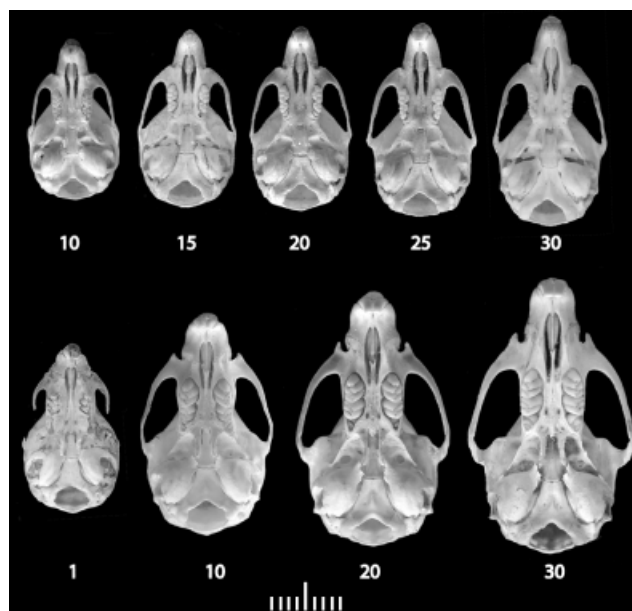


Fig. 1. Skulls of *Mus musculus domesticus* (top) and *Sigmodon fulviventer* (bottom) in ventral view at selected ages; the age of each individual, in days postnatal, is shown below the skull.

variation is constant throughout postnatal development; (2) disproportions arising early in development are gradually corrected, leading to a progressive departure from either the structure of variation characteristic of the youngest sample (prior to the reduction of variance) or from the next youngest (the youngest age at which variance has attained its minimum level).

MATERIALS AND METHODS

Samples

We examined samples of cotton rats (*Sigmodon fulviventer*) and house mice (*Mus musculus domesticus*) from the age at which skulls are well enough ossified to measure reliably to the age of sexual maturity (Fig. 1). More complete information about these samples, including the comparability of developmental ages between species and the number of litters contributing to each sample, are given in Zelditch et al. (2004). Briefly, our samples of cotton rats comprised offspring of wild-caught parents reared in the Michigan State University Museum and sacrificed at the day of birth ($N = 18$), and 10 ($N = 18$), 20 ($N = 17$), 30 ($N = 18$), 40 ($N = 11$), and 50 ($N = 12$) days after birth. Our samples of house mice comprised offspring of laboratory-reared parents of the out-bred Hsd/ICR strain, obtained from Harlan Sprague Dawley; mice were bred, reared and sacrificed at the University of Wisconsin-Madison under the supervision of T. Garland. Sampling was carried out at 5-day intervals during the phase of high growth rates and rapid changes in form, and at 10-day intervals thereafter; the samples comprised 10 ($N = 25$)-, 15 ($N = 21$)-, 20 ($N = 15$)-, 25 ($N = 15$)-, 30 ($N = 13$)-, 40 ($N = 25$)-, and 50 ($N = 29$)-day-old mice. Based on previous analyses, we found 1-day-old cotton rats to be comparable with 10-day-

old house mice in degree of maturity; 10-day-old cotton rats are approximately comparable with 15-day-old house mice. From 20 days on, the two species are nearly equal in degree of maturity at any given age (Zelditch et al. 2003).

Morphometric methods

To examine the ontogenetic dynamics of covariances of skull shape, we used landmark-based geometric shape analysis. Landmarks for skulls of cotton rats and house mice are shown in Fig. 2, A and B; the landmarks differ between species because some could not be reliably located in both. The subset of landmarks common to both includes all those sampled on house mice skulls, with the exception of the zygomatic landmark (ZA). Even though the subset of landmarks common to both species enables comparing them, it lacks some of the landmarks needed to capture important features of skull shape, most notably the most lateral braincase landmarks (GL and ZA) and the several anterior palatal landmarks and occipital landmarks available only for cotton rats. For that reason, our analyses were carried out using both the complete set of skull landmarks available for each species and the reduced set of landmarks common to both.

Landmarks were digitized on both sides of the skull but the data for one side were then reflected across the midline and coordinates of bilaterally homologous landmarks were averaged. This procedure precludes analyzing both directional and fluctuating asymmetry but allowed us to include individuals that were damaged or visibly deformed on one side. The configurations of half-skulls were geometrically scaled (to unit centroid size (CS)) and superimposed using Generalized Least Squares Procrustes superimposition, which preserves all information about shape, removing only that related to scale, position, and orientation (Rohlf and Slice 1990). To ease interpretation of the graphical results, skulls are depicted after reflecting them back over the midline.

Although geometric scaling removes variation in size, it does not eliminate the impact of that variation in size on shape. We removed that allometric variation so that samples would not be judged different solely because they differ in static allometry, although comparisons were also carried out without removing that variation to ensure that samples would not be judged different if they are similar with that component included. To remove the within-age variation correlated with within-age variation in size, we estimated the expected shape at the average size for each age by multivariate regression of shape on size. The dependent variable (shape) comprised the full set of partial warp scores (including scores on the uniform component); the independent variable (size) was measured by CS, which is the square root of the squared distance between each landmark and the centroid of the landmark configuration, summed over all landmarks. To the expected shape at the average size for each age-class, we added the residuals from the regression.

Informal comparisons of variation patterns were performed by visual inspection of the principal components (PCs) of variation within each age-class; these components represent statistically independent dimensions of variation. Formal comparisons were performed by statistically comparing covariance matrices, testing the hypothesis that these are either equal or proportional from age to age, and by comparing the dimensions in which variance is con-

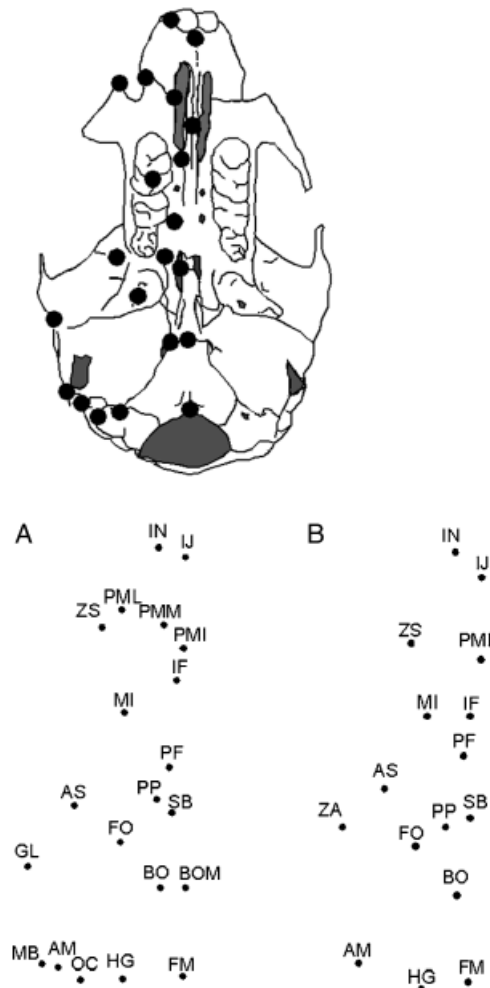


Fig. 2. Landmarks sampled on skulls of both species: (A) *Sigmodon fulviventer*; (B) *Mus musculus domesticus*. *S. fulviventer*: Junctionure between incisors on premaxillary bone (IJ); premaxilla–maxilla suture where it intersects outline of the skull in photographic plane (PML); lateral margin of incisive alveolus where it intersects outline of the skull in photographic plane (IN); anteriormost point on the zygomatic spine (ZS); suture between premaxillary and maxillary portions of palatine process (PMI); premaxilla–maxilla suture lateral to incisive foramen (PMM); posteriormost point of incisive foramen (IF); medium mure of first molar (MI); posterior palatine foramen (PF); posterolateral palatine pit (PP); junction between squamosal, alisphenoid and frontal on squamosal–alisphenoid side of suture (AS); midpoint along posterior margin of glenoid fossa (GL); anteriormost point of foramen ovale (FO); lateralmost point on presphenoid–basisphenoid suture where it intersects the sphenopalatine vacuity in the photographic plane (SB); the most lateral point on basisphenoid–basioccipital suture (BO); midpoint of basisphenoid–basioccipital suture (BOM); hypoglossal foramen (HG); junction between paroccipital process and mastoid portion of temporal (OC); midpoint of foramen magnum (FM); juncture of mastoid, squamosal and bullae (MB); juncture between mastoid and medial end of auditory tube (AM). *M. m. domesticus*: a subset of the landmarks described above, with the interior corner formed by intersection of zygomatic arch with braincase (ZA). The set of landmarks common to both species include all those visible on *M. m. domesticus* with the exception of ZA.

centrated. To determine whether covariance matrices differ, we estimated the matrix correlation (R_M) between successive ages; R_M is a conventional metric of similarity between covariance matrices (e.g., Mantel 1967; Cheverud et al. 1989; Klingenberg et al. 2003). It is calculated as a Pearson product-moment correlation across equivalent entries of two covariance matrices. For these landmark coordinate data, an entry is the covariance between a pair of coordinates, such as the x -coordinate of landmark BO and the y -coordinate of landmark FM in one sample; elements along the diagonal of the matrix are the variances of each coordinate, which are omitted from the analysis. The correlation is thus calculated over the covariances between homologous coordinates in two samples. R_M can be interpreted like a standard correlation coefficient; it ranges from -1.0 to 1.0 . When two samples have equal or proportional covariance matrices, $R_M = 1.0$, if the two matrices are unrelated, $R_M = 0.0$, and if they are maximally dissimilar, $R_M = -1.0$.

Those expected values for R_M presume that sample sizes are large and because ours were small, we could obtain correlations lower than 1.0 even if samples were drawn from the same population. Although our samples were above the size regarded as minimally sufficient for comparative studies of vertebrate skulls (Polly 2005), they were small enough that correlations lower than 1.0 could be expected even if samples were drawn from the same population. To estimate the correlation between two samples drawn from the same population, which is the maximum we would expect in comparisons between samples, we calculated R_M between each matrix and itself, resampling the data from a single age-class 1000 times (resampling is performed with replacement). The average correlation obtained from those 1000 bootstrap samples can be taken as a measure of the repeatability of the matrix. The same procedure is applied to both covariance matrices being compared, yielding the bootstrap distribution of values for the correlations between matrix **A** and itself, $R_{M(AA)}$, and between matrix **B** and itself, $R_{M(BB)}$. When the correlation between two matrices is within the 95th percentile range of the bootstrap distribution, we conclude that two matrices differ no more than expected by chance, that is, by no more than could be explained by sampling. We note that most studies test a different null hypothesis, namely, that the two matrices are more *similar* than expected by chance but we are asking whether samples, drawn from the same laboratory colony and separated by only 5 or 10 days in age, differ by more than expected by chance.

Using the repeatabilities for the two matrices, $R_{M(AA)}$ and $R_{M(BB)}$, we can adjust the observed values of between-age matrix correlations (R_M) to take the impact of sampling into account. Following Cheverud (1995), the adjusted correlation $R_{M(\text{adj})}$ is computed as:

$$R_{M(\text{adj})} = R_{M(\text{obs})} / \sqrt{R_{M(AA)} R_{M(BB)}}$$

where $R_{M(\text{obs})}$ is the matrix correlation between the observed covariance matrices for two age-classes.

Two covariance matrices could differ significantly, as judged by R_M , but still be similar in their dimensions of variation, so a method comparing those dimensions offers an important supplement to comparisons based only on a measure of overall similarity. It is possible that samples differ significantly, but only in the relative amount of variability allocated to the same combinations of

traits. For example, in one sample, the width of the braincase relative to the face might be the most variable feature, with rostral width relative to palatal width the next most, whereas in another sample, it might be that rostral relative to palatal width is the most variable feature, with braincase width relative to facial width the next most. In such cases, variance is reapportioned. A method for comparing structures of variation is particularly necessary when comparisons are based on covariances among coordinates of individual landmarks because the superimposition procedure determines the variance allocated to individual landmarks. To test the hypothesis that samples differ solely in the amount of variability, we examined the structure of variation using a method related to Common Principal Components Analysis (CPCA; Flury 1988). CPCA is now widely used to compare covariance or correlation matrices (Steppan 1997; Marroig and Cheverud 2001; Polly 2005). Using this method, a series of hypotheses, defined by Flury are tested: (1) the two matrices are equal; (2) the two matrices are proportional; (3) the two matrices have the same PCs; (4) the two matrices share some but not all PCs; and (5) the two matrices are unrelated. CPCA assumes that the ordering and identify of PCs is unambiguous (Flury 1987), meaning that PC1 accounts for statistically significantly more variance than PC2, and PC2 for statistically significantly more variance than PC3, etc. When axes instead are statistically equal in length, designating one as PC1 is effectively arbitrary because the sample does not have a unique major axis. Under those conditions, the distribution is more nearly hyperspherical than hyperelliptical, which is evident in the equality of the eigenvalues of the matrix. Under those conditions, sampling from a single hyperspherical distribution multiple times could yield a variety of apparently different PC1s. Thus, apparently large differences between samples could be due simply to chance. To test the null hypothesis that the first eigenvalue is equal to the second, we used Anderson's (1963) test for the distinctness of eigenvalues; that null hypothesis could be rejected only for one sample—the 10-day-old house mice. Consequently, for our comparisons, we used an alternative to CPCA also developed by Flury (1987), based on one approach devised by Krzanowski (1979, 1982).

This alternative method, common subspace analysis asks whether the variance within the samples is contained within a common subspace. The basic idea is graphically depicted in Fig. 3, where we show two samples compared with respect to two dimensions (defined by PC1 and PC2) within a three-dimensional space. In this example, only the proportion of variance along those two dimensions differs thus the samples occupy the same two-dimensional subspace. More technically, the method asks whether samples differ significantly in the set of eigenvectors spanning a given number of dimensions. The difference between those sets of eigenvectors is measured by the minimum angle through which one subspace must be rotated to align it with the other (the method for making that determination is given in Appendix A). The approach used by Krzanowski (1979) was to use the sum of the squared cosines of the angles between the individual pairs of eigenvectors as a measure of the difference between subspaces, rather than the more intuitive magnitude of the total rotation used here (see Equation (A6)).

To determine whether that angle is larger than expected by chance, we compared it to the range of angles that could be ob-

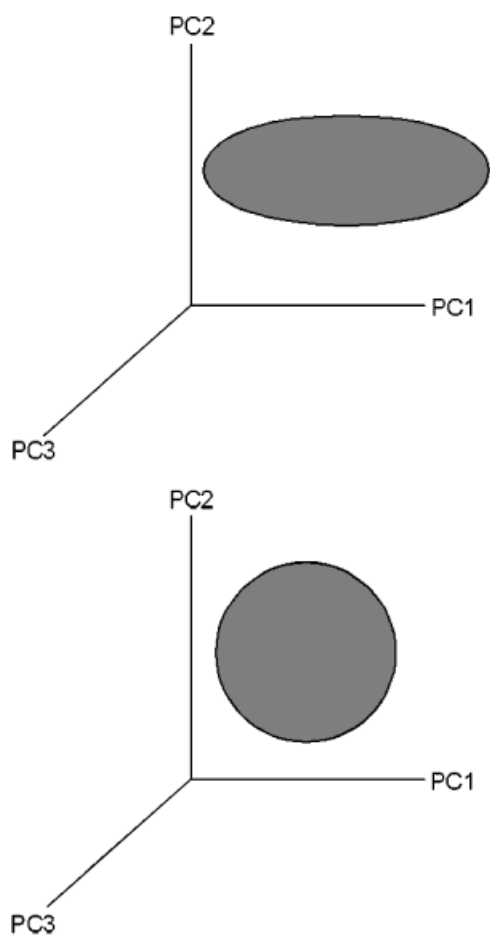


Fig. 3. Two distributions occupying the same two-dimensional subspace of a three-dimensional space. In this case, comparisons between the distributions could not be made by comparing individual principal components (PCs) to each other; the first and second PCs of the circle are not distinct because the axes are equal in length. Instead, comparisons are made between planes, testing the hypothesis that the difference between axes spanning the two-dimensional subspaces is no greater than expected by chance.

tained were the null hypothesis true. More specifically, we resampled within each individual age-class, randomly subdividing each into two groups. The sample sizes for each subdivision are determined by the sample sizes of the two age-classes being compared; when those differed, the larger sample was divided into two unequally sized samples, one having the sample size of the larger sample, the other the sample size of the other subdivision. Both subdivisions of the smaller sample have that smaller sample size to avoid generating bootstrap sets with sample sizes larger than the sample from which it is drawn. The PCA is then carried out for each subdivision within an age-class, and the angle between those two subdivisions is estimated (the same procedure is applied to the other age-class as well). Repeating this 400 times gives the bootstrap distribution of the within-age angles. When the observed angle between the subspaces of two age-classes exceeds the 97.5% range of the within-group angles generated by the bootstrap procedure, we concluded that the observed angle could not have arisen

by a random subdivision of a single group. We use the 97.5% rather than 95% range because we also tested the hypothesis that the two samples are no more similar than expected by chance. To test that other null hypothesis that the samples are no more similar than expected by chance, we compared subspaces of randomly generated data (i.e., variation at landmarks is independent, identical, and uniformly distributed) of the same dimensionality and sample size as the actual data. The minimum angle between randomly generated samples provides an estimate of the angle expected for two unrelated samples. By using the 97.5% range for each test we maintain the Type 1 error rate of 5% over both.

To choose the number of axes to include in the comparison we used two criteria. The first was the minimum number of axes required to reduce the *within-age* angles below that expected for random data. That gives the smallest subspace that can be compared meaningfully (limiting it still further would result in subspaces that can be differentiated only if they are significantly more different than unrelated samples). The second criterion was the number of axes required to account for 80% of the variance within each sample. We consider only 80% because PCs accounting for the last 20% are individually trivial (each accounts for <1% of the variance). Rather than asking if subspaces defined by those trivial axes are the same, we ask whether variance is concentrated in the same dimensions.

Procrustes superimposition and the calculation of CS were performed in CoordGen, removal of the variation related to size was performed in Standard6, PCs analysis and Anderson's test of the distinctness of eigenvalues were performed using PCAGen, and the comparison among subspaces was performed using SpaceAngle; all programs are part of the Integrated Morphometrics Programs (IMP), produced in Matlab6 (Mathworks 2000); compiled standalone versions running in Windows are freely available at <http://www2.canisius.edu/~sheets/morphsoft.html>. The calculation and bootstrapping of matrix correlations were performed using Mace (Marquez 2004).

RESULTS

PCs of variation

For each age-class in both species, half or more of the variance is explained by the first three PCs, although it usually takes five or six to explain 80% of the variance in samples of cotton rat skulls and up to seven in samples of house mice (Figs. 4, 5). Variation appears to be evenly distributed over the first few dimensions, as expected in light of our inability to reject the null hypothesis that the first two eigenvalues are equal. Even when PC1 appears to account for much more of the variance than PC2, as in samples of 10-, 20-, and 25-day-old house mice, that difference is not statistically significant.

At most ages, in both species, the proportions of the braincase relative to the face are highly variable, as shown by the lateral displacements of the landmark at the base of the zygomatic arch (GL or ZA in cotton rats and house mice, respectively) or by a shortening of the braincase relative to the face, which is indicated by a posterior displacement of that

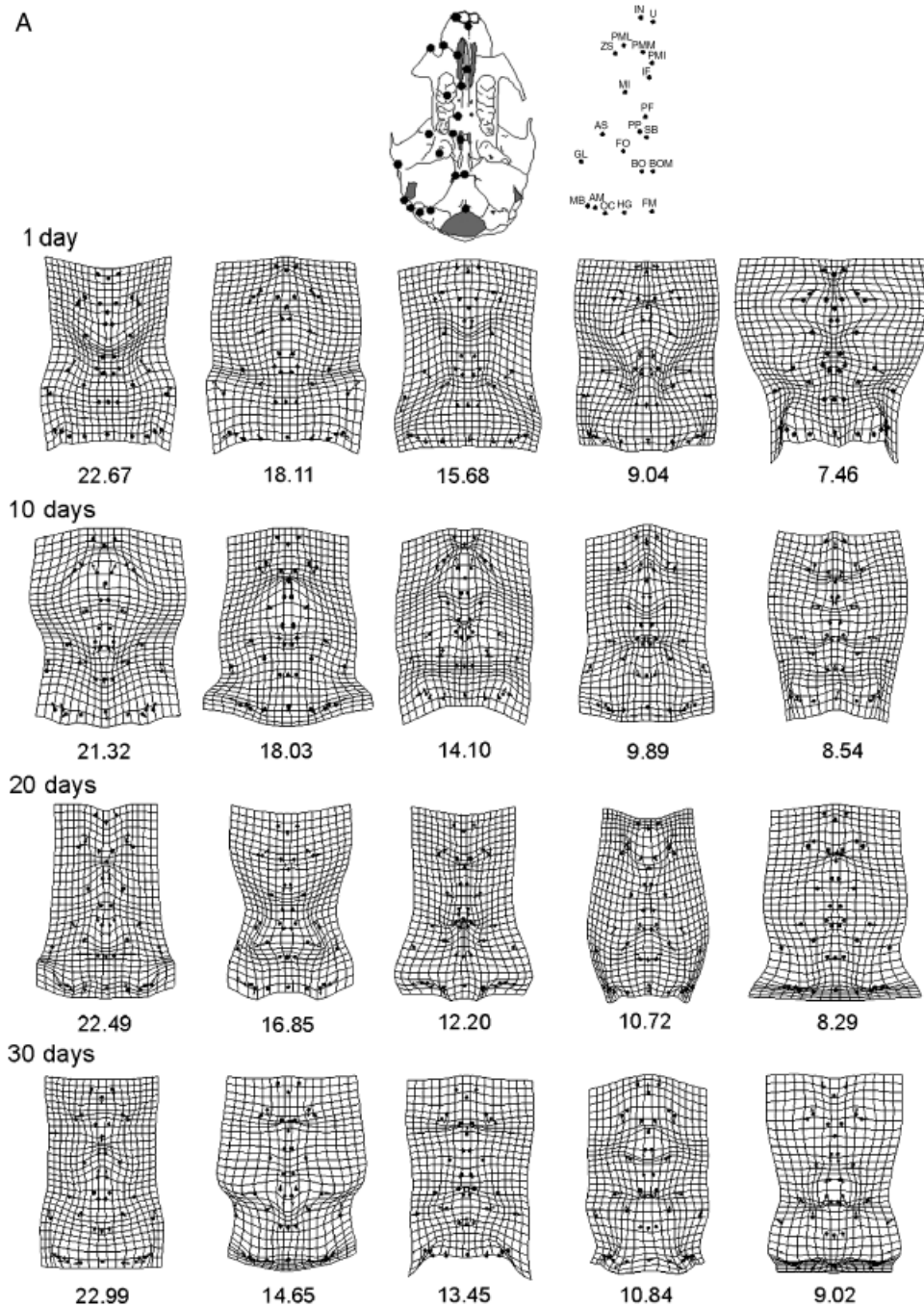


Fig. 4. The structure of variation for each sampled age of cotton rats. Shown are the first five principal components (PCs) of the variation of the complete set of landmarks after removing the effects of size. Each PC is depicted as a deformation by the thin-plate spline. Age of the sample is indicated above each row; the percentage of variation explained by each PC is indicated below.

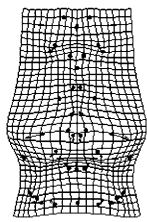
zygomatic landmark (Figs. 4, 5). Another feature seen at several ages in both species looks like a localized constriction of the skull in the region where the alisphenoid, frontal and squamosal intersect (AS), or in the region of the foramen ovale (FO); rather than indicating a change in shape of the skull (making it appear either like an hourglass or bulging in those regions), this is likely because of variability in the locations of those foramina. Similarly, variation at the posterior end of the incisive foramen (IF), which might suggest changes

in proportions of the maxilla, could be because of variation in the length of that foramen within the bone.

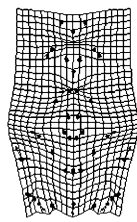
In cotton rats, there is considerable variation in the length of the IF, especially early in postnatal growth; later, the variation does not seem localized to the length of that foramen but rather it lies in the relative lengths of the premaxilla and maxilla (PMI) (Fig. 4). The width of the palate across the tooth-row also varies, as does the occipital region (in multiple directions). At most ages, there is also an interesting sugges-

B

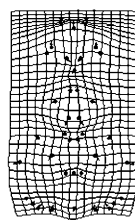
40 days



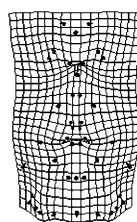
22.37



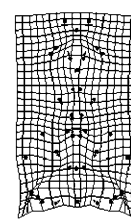
21.20



15.24

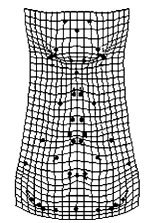


11.62

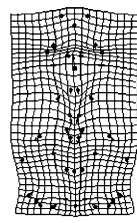


10.52

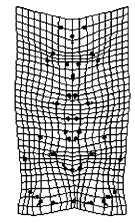
50 days



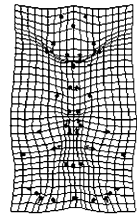
24.50



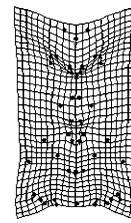
20.64



13.54



11.54



8.92

Fig. 4. (B). Continued

tion of asymmetry in the anterior palate; only asymmetry along the midline could be detected because the coordinates are averaged across both sides and reflected over the midline to yield the bilaterally symmetric skulls shown in Figs. 3 and 4; consequently, the coordinates of the two sides cannot differ. The midline points are not mathematically constrained to lie along the midline of the skull and some evidently do not, suggesting deviations from symmetry substantial enough to be visible, especially in the palatine process at the suture between PMI.

Two or more PCs may describe conflicting patterns of variation within the same region of the skull; for example, PC1 of 15-day-old house mice (Fig. 5) suggests that much of the variation within that age-class lies in the relative width of the cranium and the width of the cranial base relative to the lateral braincase, whereas PC3 and PC4 both describe variation in relative width of the posterior cranial base. Individuals with high scores on PC1 but low scores on PC3 have wide basicrania, especially across the sphenoid anteriorly, whereas individuals with high scores on both PC1 and PC3 have wide basicrania, with a more posterior widening of the cranial base. That several components describe variation within the basicranium indicates the multiplicity of directions in which basicranial shape varies.

Variation in house mice appears to be generally smoother, less often limited to a single landmark (which would look like a large and very local contraction or expansion of the grids). That apparent smoothness could be due partly to the sparser sampling of the house mouse skull, but the consistency across ages, especially in the dominant components of variation, is

striking. At most ages we see variation within the braincase, albeit complicated by what might be local variation of AS within the orbit, and variation in the landmark at the zygomatic spine (ZS), which could indicate either a localized variation in width at the premaxillary–maxillary suture or a more localized variation in the orientation of the spine. Another striking feature that is found repeatedly is the variation in width at the sphenoccipital suture (SB), suggesting that the area occupied by the bullae varies among individuals. The sharp crimping of the grid in this region does not indicate that some individuals are remarkably narrow and others are extremely broad at that suture. Rather, as discussed above, scores on these latter components must be interpreted in conjunction with components describing a more generalized broadening of the cranium.

Comparing covariance matrices

The correlations between each covariance matrix and itself are fairly low, and the confidence intervals around these correlations are quite broad (Table 1). Consequently, even low correlations would not be sufficient statistical grounds for deciding that age-classes differ by more than expected by chance. However, the correlations between age-classes are generally very low, even after adjusting them upwards to take into account the impact of sampling (Table 2). Most correlations are lower than 0.25, and the maximum is only 0.475 (for the comparison between 30- and 40-day-old house mice based on the complete set of landmarks for that species). Even the highest correlation is statistically significantly much lower

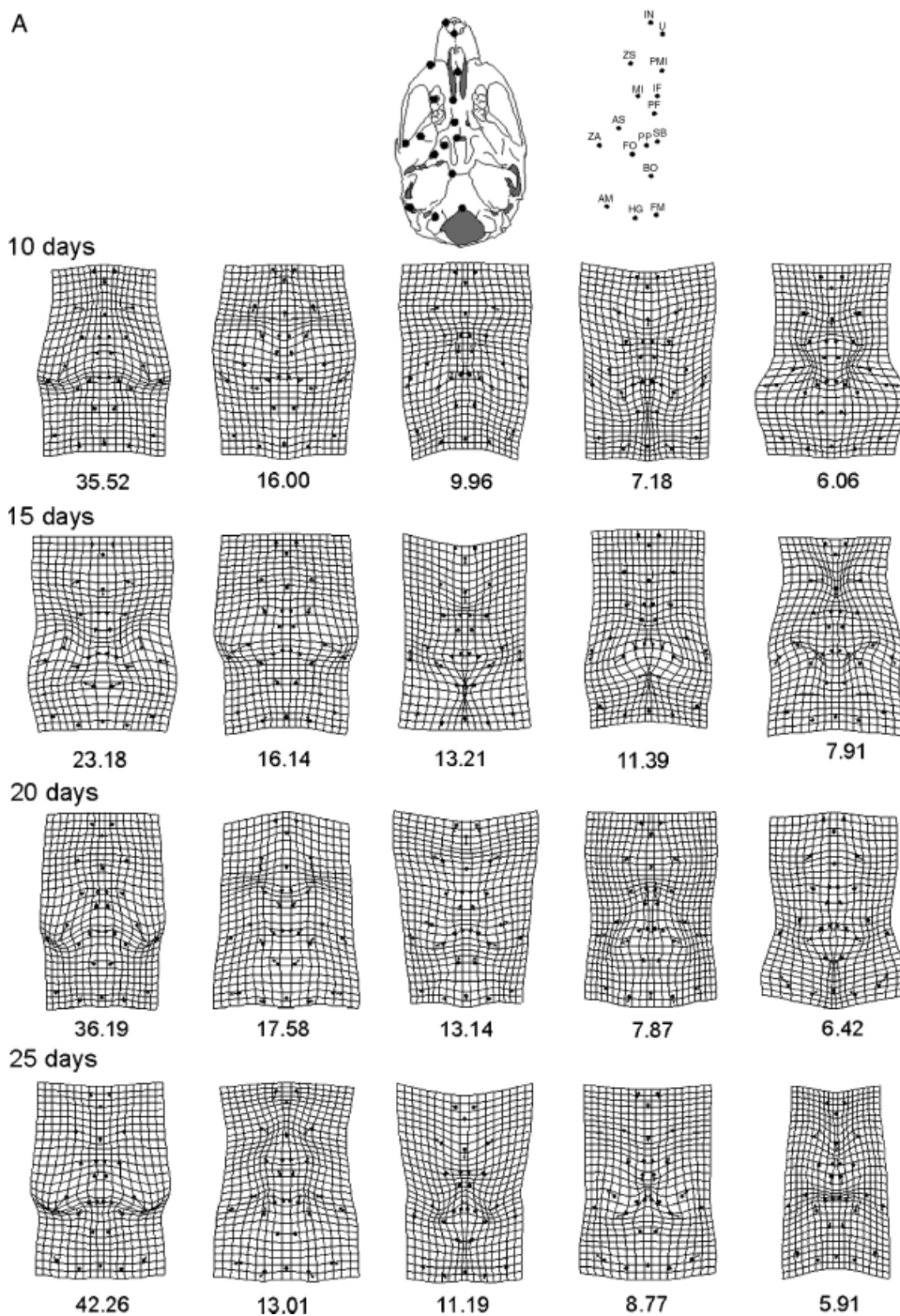


Fig. 5. The structure of variation for each sampled age of house mice. Shown are the first five principal components (PCs) of the variation of the complete set of landmarks after removing the effects of size. Each PC is depicted as a deformation by the thin-plate spline. Age of the sample is indicated above each row; the percentage of variation explained by each PC is indicated below.

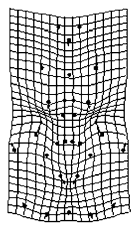
than 1.0. Removing the allometric component of shape has very little impact on the results; for example, when that component is included, the correlation between the 10- and 15-day-old house mice is 0.205 (compared to 0.228 when that component is excluded). We can thus reject the hypothesis that the covariance matrices of successive age-classes are either equal or proportional. Each comparison suggests a sig-

nificant (and large) change in covariance matrices from age to age.

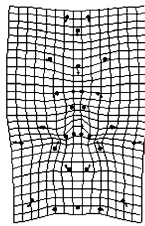
The correlations between the youngest age-class and all older ones are typically low, as are the correlations between the youngest age-class in which variance has reached its equilibrium value and all older samples (Fig. 6). In the case of cotton rats, a progressive decrease would be unexpected

B

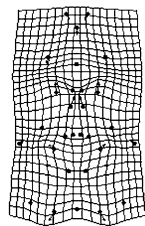
30 days



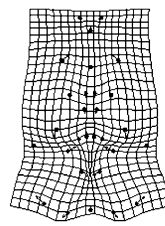
28.00



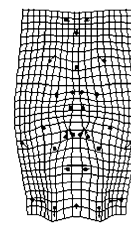
22.50



12.36

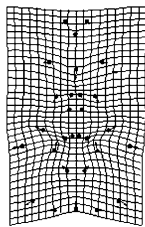


10.24

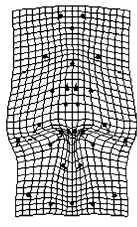


8.84

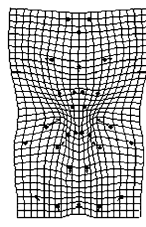
40 days



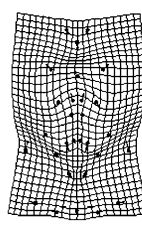
23.62



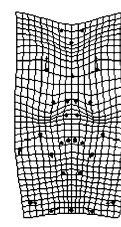
16.50



11.79



9.46



7.52

50 days

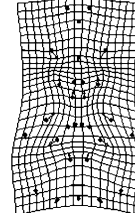
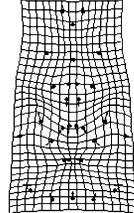
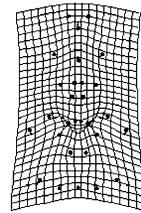
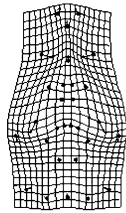
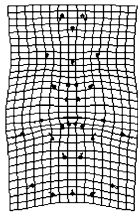


Fig. 5. (B). Continued

because the initial values are so low that they cannot decrease much further without resulting in correlations lower than expected from randomly related samples. In the case of house mice, a decrease is possible but not found.

Comparing the structure of variation

Comparisons limited to the smallest subspace that can be meaningfully compared suggest that the dominant features of variation are fairly stable (Table 3). In general, when comparisons are made in a three-dimensional space, age-classes do not differ significantly, a finding consistent with the patterns depicted in Figs. 3 and 4. Variance may be reallocated from one component to another, but the space defined by the three components does not change. When comparisons span larger subspaces, successive age-classes do tend to differ significantly. Most comparisons of successive age-classes of cotton rats require comparing subspaces of four or more dimensions, so they typically differ. In contrast, comparisons between successive age-classes of house mice often consider only three- or four-dimensional spaces, and with the exception of the comparisons involving the 20-day-olds, these do not differ. Including the component of shape variation correlated with size has very little impact on these conclusions; when that com-

ponent is included, the angle between the two youngest age-classes of mice (for the complete data set) is 101.58° (compared with 109.64° with the allometric component excluded). Interestingly, even a comparison of five-dimensional subspaces (i.e., between the 30- and 40-day-olds) indicates no statistically significant difference between these ages in house mice.

Comparisons encompassing 80% of the variance of cotton rats reveal significant change in patterns of variation (Table 4). In comparisons based on the reduced data set, the compared age-classes are sometimes no more similar to each other than would be expected by chance. House mice also undergo a restructuring of variance, although some age-classes do not differ statistically significantly. In the case of comparison between 20- and 25-day-olds, the difference is significant if the zygomatic landmark (ZA) is included, but (marginally) non-significant when that landmark is excluded. The oldest samples do not differ in pairwise comparisons, but they do when 30-day-olds are compared with 50-day-olds; for that comparison, the between-age angle is 132.82° and within-sample 95% ranges are 121.21° and 129.86° , respectively. In all of these comparisons age-classes exhibit more similarity than expected by chance.

The angle between the youngest age-class and all older ones is typically high, and so is the angle between the young-

Table 1. Correlations between each covariance matrix and 1000 bootstrap versions of itself ($R_{M(\text{self})}$), 95% confidence intervals (CI) on ($R_{M(\text{self})}$). The complete data set comprises all landmarks sampled on each species; the reduced data set comprises only landmarks common to both

Age	Complete		Reduced	
	$R_{M(\text{self})}$	95% CI	$R_{M(\text{self})}$	95% CI
<i>Sigmodon fulviventer</i>				
1	0.768	0.597–0.895	0.778	0.555–0.912
10	0.766	0.613–0.893	0.744	0.579–0.888
20	0.748	0.583–0.880	0.735	0.537–0.878
30	0.752	0.562–0.891	0.738	0.519–0.891
40	0.735	0.547–0.891	0.732	0.543–0.883
50	0.704	0.473–0.917	0.764	0.509–0.941
<i>Mus musculus domesticus</i>				
10	0.860	0.674–0.942	0.862	0.693–0.945
15	0.747	0.595–0.875	0.743	0.569–0.88
20	0.792	0.545–0.928	0.796	0.554–0.933
25	0.794	0.566–0.923	0.790	0.571–0.923
30	0.783	0.614–0.917	0.772	0.615–0.908
40	0.803	0.657–0.904	0.798	0.652–0.901
50	0.789	0.656–0.878	0.785	0.664–0.878

est age-class in which variance has reached its minimum and all older samples (Fig. 7). In neither species do we see a progressive increase in the angle. In the case of cotton rats, we cannot reject the null hypothesis that the angle between the two youngest samples is equal to that between the two oldest; the confidence interval on the difference in angle, which rang-

Table 2. Matrix correlations between covariance matrices of the complete and reduced skull landmarks from samples of successive ages; correlations based on the data ($R_{M(\text{obs})}$) and adjusted to take into account the impact of sampling error ($R_{M(\text{adj})}$)

Age	Complete		Reduced	
	$R_{M(\text{obs})}$	$R_{M(\text{adj})}$	$R_{M(\text{obs})}$	$R_{M(\text{adj})}$
<i>Sigmodon fulviventer</i>				
1–10	0.142	0.186	0.189	0.248
10–20	0.101	0.134	0.136	0.185
20–30	0.096	0.128	0.153	0.208
30–40	0.047	0.063	0.092	0.125
40–50	0.162	0.225	0.091	0.122
<i>Mus musculus domesticus</i>				
10–15	0.228	0.284	0.279	0.349
15–20	0.300	0.390	0.318	0.414
20–25	0.019	0.023	0.021	0.026
25–30	0.050	0.063	0.008	0.011
30–40	0.377	0.475	0.307	0.391
40–50	0.371	0.466	0.296	0.374

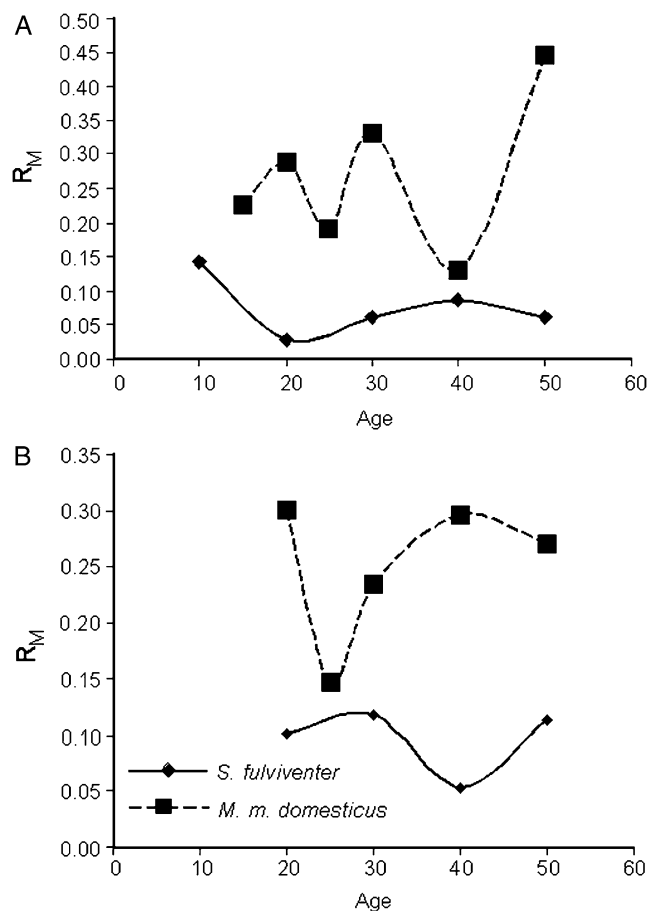


Fig. 6. Similarity between covariance matrices measured by matrix correlations, R_M . (A) Comparisons with the youngest sample (1-day-old cotton rats, 10-day-old house mice); (B) comparisons with the youngest sample in which variance has reached its minimum value (10-day-old cotton rats, 15-day-old house mice). A trend away from the covariance structure of the youngest age would be evident in a continuously decreasing R_M .

es from -11.12° to 21.71° , includes 0° . Similarly, comparisons with the youngest sample in which variance has reached its minimum suggests no progressive compensation for the variation present at that stage; the confidence interval on the difference between the angle between the two youngest and two oldest samples ranges from -10.93° to 27° , which includes 0° . However, as when comparisons were made between successive age-classes, many of the angles lie within the range expected for five-dimensional subspaces of randomly generated data. In the case of house mice, the angles between the 10- and 15-day-old samples are numerically almost identical to that between the 10- and 50-day-old samples, hence it is not surprising that the difference between them is not significant; the confidence interval for the difference between the angles ranges from -21.08° to 18.96° , which includes 0° . Similarly, the comparison with the youngest age at which variance has reached its minimum indicates no progressive change in

Table 3. Differences in structure of variation, measured by the angles (in degrees) between the smallest subspaces that can be meaningfully compared

Age	Complete			Reduced		
	No. of PCs	<i>B</i>	<i>W</i> (1)/ <i>W</i> (2)	No. of PCs	<i>B</i>	<i>W</i> (1)/ <i>W</i> (2)
<i>Sigmodon fulviventer</i>						
1–10	3	114.72	113.08/116.74	4	114.62	115.86/123.61
10–20	4	121.14	125.07/125.64	5	131.68	122.70/126.91
20–30	4	147.17	129.43/129.11	5	138.11	126.02/131.84
30–40	4	138.32	129.63/134.05	5	138.41	132.09/136.69
40–50	5	140.87	136.12/135.31	4	126.61	120.46/122.36
<i>Mus musculus domesticus</i>						
10–15	4	109.63	112.43/120.31	4	107.40	111.18/121.56
15–20	5	131.73	128.62/120.29	4	128.20	118.47/120.49
20–25	4	122.09	118.26/120.86	3	106.72	107.16/109.08
25–30	3	103.80	107.81/111.89	4	118.24	113.98/120.60
30–40	5	126.86	128.92/131.88	4	109.68	115.09/122.10
40–50	3	99.01	106.03/110.19	3	103.54	105.03/106.19

Given are the number of principal components (PCs) compared (No. of PCs), between-age angle (*B*), within-age angles (*W*(1)/*W*(2)). When *B* exceeds both *W*(1) and *W*(2), the difference between the two subspaces is statistically significant (indicated by a bold font).

structure of variation; once again, the values for the comparisons between youngest and oldest are nearly numerically identical and the confidence interval on the difference in angle, which ranges from -25.90° to 16.0° , includes 0° .

DISCUSSION

The strikingly low matrix correlations between ages, plus the significant differences in structure of variation, indicate that

patterns of integration are temporally and spatially dynamic. Whereas the level of variation is constant from age to age, the structure continually changes. Those dynamics argue against the hypothesis that constancy of variance is due to processes preventing new variance from being generated in the first place. If that were the case, the patterns of variation would not differ from age to age by more than expected by chance. That stability in the structure and even relative amount of variance appears to be characteristic of the most highly variable features, which are presumably the least effectively

Table 4. Differences in structure of variation, measured by the angles (in degrees) between the subspaces encompassing 80% of the variation

Age	Complete			Random	Reduced			Random
	No. of PCs	<i>B</i>	<i>W</i> (1)/ <i>W</i> (2)		No. of PCs	<i>B</i>	<i>W</i> (1)/ <i>W</i> (2)	
<i>Sigmodon fulviventer</i>								
1–10	6	148.26	131.26/134.13	161.59	6	147.30	122.82/127.13	149.49
10–20	6	153.83	137.78/134.10	161.44	6	150.27	127.10/126.30	148.05
20–30	6	156.16	137.72/140.12	161.84	6	149.69	128.53/131.85	149.99
30–40	5	153.36	135.87/140.05	150.41	5	138.41	130.22/135.42	140.21
40–50	5	140.87	135.57/132.42	148.46	5	148.99	129.53/125.72	138.67
<i>Mus musculus domesticus</i>								
10–15	6	126.82	124.16/124.73	152.63	6	136.53	124.27/124.48	148.98
15–20	5	131.73	120.68/124.81	142.65	5	131.63	122.05/126.82	139.72
20–25	5	131.42	123.08/123.94	142.49	5	125.08	123.05/125.57	140.21
25–30	5	137.59	121.61/125.88	141.37	5	140.11	118.57/126.45	140.87
30–40	6	133.20	130.95/137.14	142.48	6	135.51	128.61/135.71	148.54
40–50	7	110.83	127.6/125.73	160.51	7	115.63	128.15/130.99	156.69

Given are the number of principal components (PCs) compared (No. of PCs), between-age angle (*B*), within-age angles (*W*(1)/*W*(2)). When *B* exceeds both *W*(1) and *W*(2), the difference between the two subspaces is statistically significant (indicated by a bold font).

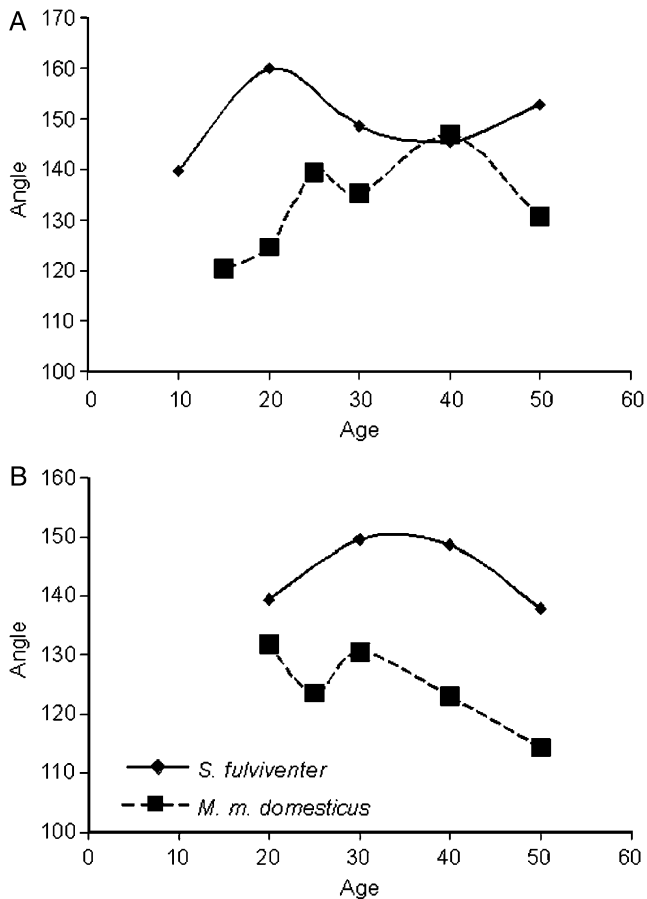


Fig. 7. Similarity between covariance matrices measured by angles between subspaces, measured in degrees ($^{\circ}$). (A) Comparisons with the youngest sample (1-day-old cotton rats, 10-day-old house mice); (B) comparisons with the youngest sample in which variance has reached its minimum value (10-day-old cotton rats, 15-day-old house mice). A trend away from the covariance structure of the youngest age would be evident in a continuously increasing angle.

canalized. Not only are the three most dominant features generally similar in structure, they also account for similar proportions of variance (50–60%) at nearly all ages, including that stage prior to the initial reduction in variance. The less variable features are the ones that appear to be restructured, but there is no evidence of any progressive trend away from the structure of variation present early in ontogeny, as might be expected if errors are gradually corrected. Rather, variation appears to be continually altered but not preferentially in any particular direction, suggesting that errors are corrected as rapidly as they are generated. Thus, our data more strongly support the hypothesis that constancy of variation results from a balance between processes that generate and remove variation.

Although our data argue against the hypothesis that the constancy of variance is due to processes that forestall any new variation from being produced, we cannot conclusively

reject it because we measured different individuals at different ages and our samples are small. It is possible that our samples do not adequately represent the range of variation present at each age and that even statistically significant differences are artifacts of sampling. Convincing evidence for the hypothesis that variation is continually generated but fails to accumulate because it is equally rapidly removed, requires larger samples, or ideally, measurements made on the same individuals repeatedly throughout growth. Such evidence is provided by one longitudinal analysis of variation of cranial shape of laboratory rats (*Rattus norvegicus*), sampled in sagittal view (data published in Bookstein 1991, Appendix A.4.5., available electronically at <http://life.bio.sunysb.edu/morph>). In that population, individuals who are most deviant at one age, later approach the mean, whereas those who are near the mean, later depart from it, and the same individuals deviate in different features at different ages (Zelditch 2005). Because only a few landmarks were sampled, and the variation is too nearly random to allow for comparisons of its structure, this is not convincing evidence that the processes responsible for maintaining the level of variance in dynamic equilibrium also restructure it. However, one crucial element of that hypothesis—the balance between generation and removal of variance, is well supported. Moreover, compensatory differential growth has been detected in longitudinal studies of fluctuating asymmetry, suggesting that it is also important in regulating developmental noise, although in that case, compensatory differential growth is not balanced by processes generating new variance (Kellner and Alford 2003).

Variance, overall, restructures, but the most poorly canalized (hence most variable) features tend to be temporally stable, and qualitatively similar between cotton rats and house mice. They even resemble features found to be highly variable in other mammals, suggesting that their variability is due to developmental processes shared by mammals in general. For example, one of the most highly variable features in both cotton rats and house mice is the location of the juncture of three bones: the alisphenoid, squamosal, and frontal (AS). Studies of primates also find high variability of lengths of orbital traits (which would include lengths measured from AS), and these are negatively correlated, suggesting that orbital landmarks vary randomly in their locations around an invariant-length orbital rim (Cheverud 1995). In effect, the space occupied by the orbit is canalized but the proportions of the bones filling that space are not. In such cases, adjacent bones (or regions within bones) may evince a spatial form of compensatory growth in that any deficit of one is balanced by an excess of another. The high variability of the landmark at the FO may have a similar explanation—the location of the foramen is unconstrained so long as it lies within the alisphenoid. Similarly, the width of the cranial base, another consistently variable feature, may be determined partly by the space to be filled between bullae. As that space diminishes, the

absolute variability of these features diminishes, but their relative variability does not.

Notwithstanding the stability of the least canalized features, the overall structure of variation differs notably from age to age. The matrix correlations between age-classes that are separated by only 5–10 days are remarkably low, especially in context of other studies that also use geometric methods of shape analysis. They are similar to values reported for comparisons between covariance matrices of skull landmarks of two distantly related species, rhesus macaques and house mice ($R_M = 0.127, 0.169$ for females and males, respectively; Hallgrímsson et al. 2004). They are sometimes exceeded by matrix correlations between patterns of variation and fluctuating asymmetry ($R_M = 0.12$, cichlid pharyngeal jaws, Klingenberg et al. 2002; $R_M = 0.37$, mandibles of house mice, Klingenberg et al. 2003; and $R_M 0.030–0.220$, for skulls of *M. m. domesticus*, *M. m. musculus*, and hybrids, Debat et al. 2000) even though such low correlations are taken to mean that the underlying processes of canalization and developmental stability differ.

The low correlations between ages might result from continual changes in the processes generating variance, perhaps because of differential gene expression (Atchley et al. 1981), or changes in the spatial organization of bone deposition and resorption because of changes in the forces generated by growth and muscle activity. Both the magnitude and orientation of masticatory forces change over time because of increasing muscle strength, differential bone growth, or both (Sun et al. 2004). Those changes in forces can even evoke changes in gene expression because mechanical forces induce the expression of some growth-related genes (Fitzgerald and Hughes-Fulford 1999; Hatton et al. 2003). The spatiotemporal dynamics of these forces may explain the changes in the spatial distribution of variation over the skull. However, the temporal behavior of variance cannot be explained solely by the processes generating variance because as much variation is removed as is generated. It is possible that the processes removing variance are spatially uniform and therefore have no impact on the structure of integration. That, however, seems unlikely.

The continual restructuring of (co)variation found in cotton rats directly contradicts the conclusions of a prior study of this same population (Zelditch and Carmichael 1989). In that study, a significant change in integration was detected prior to weaning; thereafter the pattern was constant. However, our present study analyzed different parts of the skull in that we exclude the mandible, but include more details of palatal morphology and analyze the face together with the braincase. Also, we use geometric rather than traditional morphometric methods, and perhaps more importantly, we use a different method to compare patterns of integration. In the previous study, models of integration were evaluated for their goodness-of-fit to the data, and samples were compared by testing

the hypothesis that the samples were drawn from a single homogeneous population with the factor structure predicted by the model. Prior to weaning, the best-fitting model postulates only a size factor with the remaining variation being randomly distributed around it; shortly before weaning, a second factor emerges, indicating integration between mandible and upper jaw (and teeth). That one masticatory factor might indeed be stable, but our present analysis argues against such a simple model of skull integration.

The hypothesis of compensatory differential growth implies that skull shape is actively monitored, with deviations from the norm actively corrected, either by up-regulating bone growth where too little bone was previously deposited, or by down-regulating it, or both. An alternative explanation is that shape is not directly regulated at all but rather something else is, such as bone strain. Bone is modeled in response to physical forces so that bone does not bend when loaded by muscles, primarily by modulating rates of periosteal apposition (Chamay and Tchantz 1972; Lanyon 1984; Lieberman and Pearson 2001; for a review of the biomechanics of bone growth, see Skerry 2000). The processes canalizing skull shape may be unresponsive to deviations in shape, reacting instead to present levels of strain. The restructuring of subtle features, such as the tapering of the ZSs and the relative growth rate of the premaxilla relative to maxilla (within the palatine process) may be less a matter of error correction than overwriting old deviations with new ones.

The ontogenetic dynamics of (co)variance structure suggests that morphological integration does not require coordinated growth of parts but can instead arise via compensatory interactions that correct for locally disproportionate growth. This is a developmental analogue to a pattern commonly seen in phylogenetic studies—characters may be correlated within species, but they need not have undergone correlated evolutionary transformations. Our analysis does not dispute the idea that pleiotropy is an important source of integration but rather it argues that that coordinated growth caused by pleiotropy does not explain either the constancy of variance or details of skull integration. We suggest that compensatory differential growth, balanced by newly generated variance, explains both. Whether shape is actively regulated, or instead it is strain that is regulated so that deviations from the normal shape are incidentally corrected when bone adapts to local strains, is an important question that we cannot answer. However, our results suggest that (co)variation patterns are continually restructured by processes that equilibrate variance, and thus canalization plays a critical role in molding patterns of integration.

Acknowledgment

T. G. was supported by NSG grant IBN-0212567. We thank D. Polly and an anonymous reviewer for helpful criticisms of this manuscript.

REFERENCES

- Anderson, T. W. 1963. Asymptotic theory for principal components. *Ann. of Math. Stat.* 34: 122–148.
- Atchley, W. R., Rutledge, J. J., and Cowley, D. C. 1981. Genetic components of size and shape. II. Multivariate covariance patterns in the rat and mouse skull. *Evolution* 35: 1037–1055.
- Bookstein, F. L., 1991. *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press, Cambridge.
- Chamay, A., and Tchantz, P. 1972. Mechanical influences in bone remodeling. Experimental research on Wolff's Law. *J. Biomech.* 5: 173–180.
- Cheverud, J. M. 1982. Phenotypic, genetic and environmental integration in the cranium. *Evolution* 36: 499–512.
- Cheverud, J. M. 1984. Quantitative genetics and developmental constraints on evolution. *J. Theor. Biol.* 101: 155–171.
- Cheverud, J. M. 1995. Morphological integration in the saddle-back tamarin (*Saguinus fuscicollis*) cranium. *Am. Nat.* 145: 63–89.
- Cheverud, J. M. 1996. Developmental integration and the evolution of pleiotropy. *Am. Zool.* 36: 44–50.
- Cheverud, J. M., Wagner, G. P., and Dow, M. M. 1989. Methods for the comparative analysis of variation patterns. *Syst. Zool.* 38: 201–213.
- Debat, V., Alibert, P., David, P., Paradis, E., and Auffray, J.-C. 2000. Independence between developmental stability and canalization in the skull of the house mouse. *Proc. R. Soc. London B* 267: 423–430.
- Fitzgerald, J., and Hughes-Fulford, M. 1999. Mechanically induced *c-fos* expression is mediated by cAMP in MC3T3-E1 osteoblasts. *FASEB J.* 13: 553–557.
- Flury, B. 1987. Two generalizations of the common principal component model. *Biometrika* 74: 59–69.
- Flury, B. 1988. *Common Principal Components and Related Multivariate Models*. Wiley, New York.
- Hallgrímsson, B., Willmore, K., Dorval, C., and Cooper, D. M. L. 2004. Craniofacial variability and modularity in macaques and mice. *J. Exp. Zool.* 302B: 207–225.
- Hatton, J. P., Pooran, M., Li, C.-F., Luzzio, C., and Hughes-Fulford, M. 2003. A short pulse of mechanical force induces gene expression and growth in MC3T3-E1 osteoblasts via an ERK 1/2 pathway. *J. Bone Miner. Res.* 18: 58–66.
- Kellner, J. R., and Alford, R. A. 2003. The ontogeny of fluctuating asymmetry. *Am. Nat.* 161: 931–947.
- Klingenberg, C. P., Barluenga, M., and Meyer, A. 2002. Shape analysis of symmetric structures: quantifying variation among individuals and asymmetry. *Evolution* 56: 1909–1920.
- Klingenberg, C. P., Mebus, K., and Auffray, J.-C. 2003. Developmental integration in a complex morphological structure: how distinct are the modules in the mouse mandible? *Evol. Dev.* 5: 522–531.
- Krzyszowski, W. J. 1979. Between-groups comparisons of principal components. *J. Am. Stat. Assoc.* 74: 703–701, Correction (1981) 76: 1022.
- Krzyszowski, W. J. 1982. Between-groups comparisons of principal components—some sampling results. *J. Stat. Comput. Simulation* 15: 141–154.
- Lanyon, L. E. 1984. Functional strain as a determinant for bone remodeling. *Calcified Tissues Int.* 37: 119–124.
- Lieberman, D. E., and Pearson, O. M. 2001. Trade-off between modeling and remodeling responses to loading in the mammalian limb. *Bull. Museum Comp. Zool.* 156: 269–282.
- Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27: 209–220.
- Marquez, E. 2004. MACE: Matrix correlations for landmark data, ver. 1.1. Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA.
- Marroig, G., and 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.
- Mathworks. 2000. *MATLAB*. The Mathworks, Natick, MA.
- Olson, E. C., and Miller, R. L. 1958. *Morphological Integration*. University of Chicago Press, Chicago.
- Polly, P. D. 2005. Development and phenotypic correlations: the evolution of tooth shape in *Sorex araneus*. *Evol. Dev.* 7: 29–41.
- Rohlf, F. J., and Slice, D. 1990. Extensions of the procrustes method for the optimal superimposition of landmarks. *Syst. Zool.* 39: 40–59.
- Schlosser, G., and Wagner, G. P. 2004. *Modularity in Development and Evolution*. University of Chicago Press, Chicago.
- Siegal, M. L., and Bergman, A. 2002. Waddington's canalization revisited: developmental stability and evolution. *Proc. Natl. Acad. Sci. USA* 99: 10528–10532.
- Skerry, T. 2000. Biomechanical influences on skeletal growth and development. In P. O'Higgins and M. J. Cohn (eds.), *Development, Growth and Evolution: Implications for the Study of the Hominid Skeleton*. Academic Press, New York, pp. 29–39.
- Stephan, S. J. 1997. Phylogenetic analysis of phenotypic covariance structure. 2. Reconstructing matrix evolution. *Evolution* 51: 587–594.
- Sun, Z., Lee, E., and Herring, S. W. 2004. Cranial sutures and bones: growth and fusion in relation to masticatory strain. *Anatom. Rec. Part A* 276A: 150–161.
- Waddington, C. H. 1942. Canalization of development and the inheritance of acquired characters. *Nature* 150: 563–565.
- Waddington, C. H. 1952. Canalization of the development of a quantitative character. In C. H. Waddington (ed.), *Quantitative Inheritance*. Her Majesty's Stationary Office, London, pp. 43–46.
- Zelditch, M. L. 1988. Ontogenetic variation in patterns of phenotypic integration in the laboratory rat. *Evolution* 42: 28–41.
- Zelditch, M. L. 2005. Developmental regulation of variability. In B. Hallgrímsson and B. K. Hall (eds.), *Variation: A Central Concept in Biology*. Elsevier, London, pp. 249–276.
- Zelditch, M. L., and Carmichael, A. C. 1989. Ontogenetic variation in patterns of developmental and functional integration in skulls of *Sigmodon fulviventer*. *Evolution* 43: 1738–1747.
- Zelditch, M. L., Lundrigan, B. L., and Garland, T. 2004. Developmental regulation of skull morphology. I. Ontogenetic dynamics of variance. *Evol. Dev.* 6: 194–206.
- Zelditch, M. L., Lundrigan, B. L., Sheets, H. D., and Garland, T. Jr. 2003. Do precocial mammals develop at a faster rate? A comparison of rates of skull development in *Sigmodon fulviventer* and *Mus musculus domesticus*. *J. Evol. Biol.* 16: 708–720.

APPENDIX A

To determine the angle between subspaces defined by the first L principal components (PCs) of each group, we construct the matrix \mathbf{V}_A such that its N columns are the PCs (eigenvectors) of the variance-covariance matrix for sample A and the equivalent matrix, \mathbf{V}_B , for sample B .

The projection matrices \mathbf{Q} and \mathbf{R} are operators that project an arbitrary vector \mathbf{X} in the original $M \times M$ variable space onto the subspace defined by the first L eigenvectors of A and B , respectively, and are calculated

$$\mathbf{Q} = \mathbf{V}_A \mathbf{P} \mathbf{V}_A^{-1} \quad (\text{A1})$$

$$\mathbf{R} = \mathbf{V}_B \mathbf{P} \mathbf{V}_B^{-1} \quad (\text{A2})$$

where \mathbf{P} is the matrix:

$$\mathbf{P} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (\text{A3})$$

such that \mathbf{I} is a $L \times L$ identity matrix; the zeros are necessary to make \mathbf{P} an $M \times M$ matrix. \mathbf{P} serves to extract only the first L vectors of \mathbf{A} and \mathbf{B} .

The operator \mathbf{J} is defined as the difference between \mathbf{Q} and \mathbf{R}

$$\mathbf{J} = \mathbf{Q} - \mathbf{R} \quad (\text{A4})$$

and an eigenvector decomposition determines the angular change implied by \mathbf{J} . The eigenvalues of \mathbf{J} are paired positive and negative values of the form $(J_1, -J_1, J_2, -J_2, J_3, -J_3, \dots)$, expressing the angles of rotation (in orthogonal two-dimensional subspaces) that produce the smallest rotation of one L dimensional subspace into another. To compute the total angular distance we compute the square root of the summed squared angles of rotation

$$A_{\text{Distance}} = \sqrt{\arcsin(J_1)^2 + \arcsin(J_2)^2 + \arcsin(J_3)^2 + \dots + \arcsin(J_K)^2} \quad (\text{A5})$$

where all angles are in radians (when reporting angles, we convert them into degrees).

Krzanowski (1979) measured the difference between subspaces by the summed squared cosines of the angles

$$K_{\text{Distance}} = \sum \cos^2(\arcsin(J_i)) \quad (\text{A6})$$

a metric that ranges from 0 (when all angles are 90°) to L (when all angles are 0°).