

The brachymorph mouse and the developmental-genetic basis for canalization and morphological integration

Benedikt Hallgrímsson,^{a,*} Jevon J. Y. Brown,^a Alice F. Ford-Hutchinson,^b H. David Sheets,^c Miriam L. Zelditch,^d and Frank R. Jirik^b

^aDepartment of Cell Biology & Anatomy and the Joint Injury and Arthritis Research Group,

^bDepartment of Biochemistry & Molecular Biology and the Joint Injury and Arthritis Research Group, University of Calgary, 3330 Hospital Dr., Calgary, AB, Canada T2N 4N1

^cDepartment of Physics, Canisius College, 2001 Main Street, Buffalo, NY 14208, USA

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

*Address for correspondence (email: bhallgri@ucalgary.ca)

SUMMARY Although it is well known that many mutations influence phenotypic variability as well as the mean, the underlying mechanisms for variability effects are very poorly understood. The brachymorph (*bm*) phenotype results from an autosomal recessive mutation in the phospho-adenosine-phosphosulfate synthetase 2 gene (*Papps2*). A major cranial manifestation is a dramatic reduction in the growth of the chondrocranium which results from undersulfation of glycosaminoglycans (GAGs) in the cartilage matrix. We found that this reduction in the growth of the chondrocranium is associated with an altered pattern of craniofacial shape variation, a significant increase in phenotypic variance and a dramatic increase in morphological integration for craniofacial shape. Both effects

are largest in the basicranium. The altered variation pattern indicates that the mutation produces developmental influences on shape that are not present in the wildtype. As the mutation dramatically reduces sulfation of GAGs, we infer that this influence is variation among individuals in the degree of sulfation, or variable expressivity of the mutation. This variation may be because of genetic variation at other loci that influence sulfation, environmental effects, or intrinsic effects. We infer that chondrocranial development exhibits greater sensitivity to variation in the sulfation of chondroitin sulfate when the degree of sulfation is low. At normal levels, sulfation probably contributes minimally to phenotypic variation. This case illustrates canalization in a particular developmental-genetic context.

INTRODUCTION

It is commonly thought that deleterious mutations often produce increased phenotypic variance among mutants compared with that typical among individuals for the wildtype (Wagner 2003; Hermisson and Wagner 2004). However, the generality of this phenomenon is difficult to assess because variances are not routinely reported in the literature, even though there is substantial evidence for the claim that mutant phenotypes are often more variable than the wildtype (Scharloo 1991). That observation is the basis for the concept of canalization or robustness, which refers to the tendency for developmental systems to minimize the effects of genetic or environmental variation (Waddington 1942; Wagner et al. 1997; Wilkins 1997; de Visser et al. 2003). According to Waddington (1957), mutant phenotypes are more variable because they express genetic or environmental variation that is normally suppressed in the wildtype. The mechanisms by

which buffering suppresses variability, in developmental-genetic terms, remains poorly understood.

There are two contrasting views on the developmental basis for canalization. One is that distinct genes (and developmental mechanisms) promote canalization and are directly selected or evolve as a byproduct of stabilizing selection. Hsp90 and other molecular chaperones are reportedly examples of these (Rutherford and Lindquist 1998; Rutherford 2000). The other is that canalization arises as an emergent property of developmental-genetic architecture, such as the complexity of epistatic interactions, or redundancies among related genes, or thresholds and nonlinearities in relationships between developmental determinants and phenotypic outcomes (Siegal and Bergman 2002; Bergman and Siegal 2003). If this second hypothesis is true, then most genes could influence a developmental system's tendency to express phenotypic variation through the indirect effects of these emergent properties. Robustness would then evolve either by

stabilizing selection, as a byproduct of the complexity of epistatic interactions, of thresholds and nonlinearities, or through the transient redundancies that result from gene duplications—all of which may evolve independently of stabilizing selection (Siegal and Bergman 2002).

This study examines the effect of the brachymorph (*bm*) mutation on phenotypic variability in the mouse skull. The brachymorph phenotype results from an autosomal recessive mutation in the phosphoadenosine-phosphosulfate synthetase 2 gene (*Papss2*) (Kurima et al. 1998). The brachymorph mouse is an interesting model for studies of the developmental genetics of phenotypic variability for several reasons. First, although this mutation obviously affects morphology, the *Papss2* gene is not directly involved in the developmental-genetic control of development because it affects the composition of the extracellular matrix (as discussed below). In contrast to *Hsp90*, it is not likely to have a specific buffering function. Secondly, the mutation differentially affects the chondrocranium in the skull and would thus be expected to have differential effects on the variability of chondrocranium and dermatocranium. Finally, it is likely that the developmental context of the gene involves two aspects of developmental genetic architecture thought to influence variability. The first is the likely functional redundancy with the related gene *Papss1* based on sequence and functional similarity of these genes, and the fact that some sulfation occurs in the cartilage of *bm/bm* homozygotes (Orkin et al. 1976; Xu et al. 2000). Also, although it is not known whether there is genetic variation in PAPSS2 activity during formation and growth of cartilage in development, PAPSS2 activity is known to vary 18-fold in human liver without obvious phenotypic effects (Xu et al. 2001). If there is genetic variation for PAPSS2 activity, this suggests either a threshold or extreme nonlinearity in the relationship between the activity of this enzyme and phenotypic outcome.

If the *bm* mutation causes the phenotypic expression of variation that is unexpressed in the wildtype, this variation should be qualitatively different from the variation that characterizes the wildtype. This is a prediction that can be tested using current morphometric techniques. Should these predictions be met, we could conclude that the *Bm* mutation influences phenotypic variability via emergent properties of the developmental-genetic system rather than because of any direct canalizing effects of the PAPSS2 enzyme. In this study, we test the predictions that the *bm* mutation increases phenotypic variance and produces variation that is distinct in covariance structure from that which characterizes the wildtype.

THE BRACHYMORPH MUTATION

The *Papss2* gene codes for an enzyme involved in the sulfation pathway, catalyzing the synthesis of adenosine-phosphosul-

fate (APS) and the subsequent phosphorylation of APS to produce PAPS. Within the Golgi apparatus, PAPS is a sulfate donor for a variety of proteins including glycosaminoglycans (GAGs). The mutation is a base pair substitution which causes a Gly to Arg substitution within the APS-kinase domain of the protein and thus interferes with the second of the two functions of the enzyme (Kurima et al. 1998). The result is a dramatic decrease in the availability of PAPS and thus in the sulfation of GAGs (Orkin et al. 1976). The cartilage extracellular matrix of *bm/bm* mice is therefore affected in that GAGs are undersulfated and thus less negatively charged, making the proteoglycan aggregates smaller in size and less abundant (Orkin et al. 1976). As a result of this altered extracellular matrix, the growth of cartilage is dramatically reduced thus all skeletal elements that depend on cartilage growth are abnormally small. Long bones are reduced by 50% in length and articular surfaces are flared (Lane and Dickie 1968; Ford-Hutchinson et al. 2005). In the skull, the direct effects of the mutation should be confined to the chondrocranium because the growth of dermatocranial elements does not depend directly on cartilage growth (Kaufman and Bard 1999).

MATERIALS AND METHODS

The sample

Mice homozygous for the brachymorph mutation on a C57BL/6J background and the C57BL/6J wildtype controls were obtained from Jackson Labs (Bar Harbor, ME, USA). All individuals were housed in a virus antibody free barrier facility at the University of Calgary. Samples of 19 (8 females, 11 males) controls and 21 (10 females, 11 males) *bm/bm* mice were euthanized by CO₂ inhalation in accordance with the University of Calgary Animal Care Committee and Canadian Council on Animal Care guidelines. After euthanasia, heads were removed and stored at -20°C prior to analysis. Individuals in the sample were sacrificed at ages ranging from 180 to 366 days; mice were harvested initially as part of a study on age-dependent joint degeneration in the brachymorph mutant (Ford-Hutchinson et al. 2005). The age heterogeneity of the sample is not a concern for this study because there is minimal cranial growth in mice past 90 days of age and the *bm/bm* sample is age matched to the controls. The absence of age-related variation for craniofacial shape was verified using multiple regression for all Procrustes coordinates with age as the independent variable and genotype as a factor (df = 37, Hotelling's trace = 6552, NS at $\alpha = 0.05$).

The brachymorph sample and the C57BL/6J controls can be assumed to be comparable in all respects except the effects of the *bm* mutation. The C57BL/6J strain, on which the *bm* mutation is expressed, has been inbred for 220 generations and therefore has minimal genetic variance (www. Jax.org). A difference in variance between the two groups is, therefore, not the result of environmental or genetic influences that differ between groups but rather of the degree to which the two developmental configurations

differing only in the *bm* mutation respond to the same genetic and environmental influences.

Computed microtomography and landmark data

Whole heads were defrosted and scanned twice using a Skyscan 1072 100 kV computed microtomograph (Skyscan, Aartselaar, Belgium) at an isotropic resolution of 19.43 μm using a protocol optimized for mouse heads (no filter, 100 kV, 0.9° rotation step, 1.9-ms exposure time, three frame averaging). Flat field corrections were performed prior to each batch of scans. At the 2D reconstruction stage, global threshold values and postalignment corrections were verified manually. The reconstructed 2D slices were then subjected to a 2 pixel kernel 3D median filter using a custom plug-in written for ImageJ.

The morphometric analysis was based on 3D landmarks obtained directly from volumetric renderings of the CT data (Fig. 1). Table 1 provides a key to these landmarks and brief anatomical descriptions. Volumetric renderings were produced using Analyze 3D (AnalyzeDirect, Lenexa, KS, USA). As each individual was scanned twice, the entire process from scanning through filtering to digitization is repeated, allowing assessment of measurement error. The landmark configurations were inspected for gross outliers resulting from measurement error by visual inspection of superimposed Procrustes coordinates. Individuals with gross errors (because of reversal of side or mislabeling of landmarks) were re-digitized. The landmarks (Fig. 1A) are the same as those used in earlier studies (Hallgrímsson et al. 2004a, b). For analysis of regional variation and covariation structure, landmarks were divided

into groups based on the developmental organization of the skull (Fig. 1B). The basicranial region is composed of elements that develop from the chondrocranium; the neurocranium is dermatocranial in origin and is influenced by the growth of the brain; the zygomatic region is sampled by a set of landmarks in the regions most directly affected by the masseter muscle whereas the facial region encompasses the remaining landmarks facial landmarks. All regions are known to have statistically significant internal correlation based on a previous analysis of modularity in the mouse skull (Hallgrímsson et al. 2004b).

Statistical methods

The statistical analysis focused on (1) comparing among-individual variation for size and shape, (2) fluctuating asymmetry (FA), (3) the covariation structure in the two groups, and (4) the level of morphological integration. Detailed analysis of the effect of the mutation on the mean phenotype is the subject of a separate study. We used both geometric morphometric and Euclidean distance-based methods. For the geometric morphometric methods, the data were subjected to Procrustes superimposition in either Morphueus (Slice 1994–1999) or the integrated morphometrics package (IMP) written by sheets.

The use of Procrustes superimposition has been criticized on the grounds that the variance co-variance matrix is non-identifiable (Lele and Richtsmeier 2001; Lele and McCulloch 2002). This area of debate, reviewed cogently by Slice (2005), is far beyond this article to resolve. However, we deal with this uncertainty in two ways. Firstly, we avoid the assumption throughout that varian-

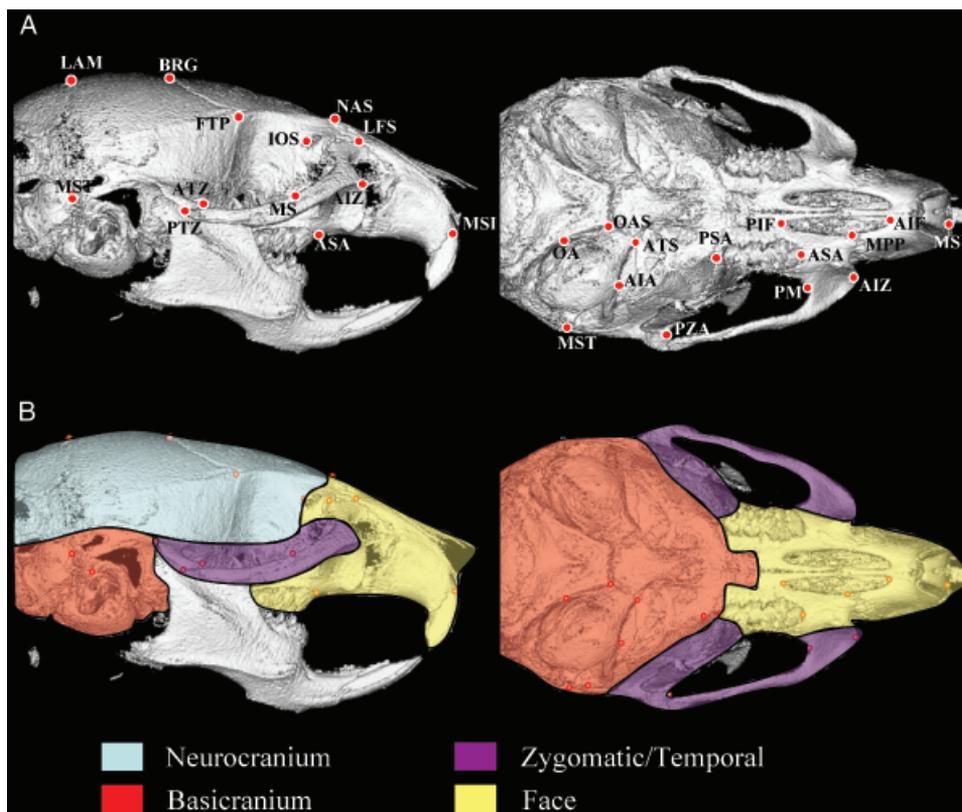


Fig. 1. Landmarks used in this study (A) and developmentally based anatomical regions used for regional comparisons (B).

Table 1. List of landmarks shown in Fig. 1 with brief descriptions

Code	Description
AIA	Anterior process of auditory bulla (ectotympanic)
AIF	Anterior margin of anterior palatine foramen
AIZ	Anteriormost point along inferior zygomatic rim of the maxilla
ASA	Anterior edge of the alveolus of the upper first molar
ATS	Antero-medial projection of ectotympanic in basicranial view
ATZ	Anterior temporal–zygomatic junction on zygomatic arch
BRG	Bregma—the midline intersection of the coronal and sagittal sutures
FTP	Fronto-temporal-parietal junction (anterior projection of parietal in superior view)
IOS	Intersection of fronto-maxillary suture with orbital rim
LAM	Lambda—the midline intersection of the sagittal and lambdoidal sutures
LFS	Lateral extremity of frontonasal suture
MMP	Junction of premaxillary suture and the anterior palatine foramen
MS	Superior margin of suture of temporal and zygomatic processes of zygomatic arch
MSI	Midline of superior incisor along anterior alveolar margin
MST	Posterior edge of post-tympanic hook of temporal bone
NAS	Nasion—midline junction of frontal and nasal bones
OA	Posterior edge of ectotympanic along its margin with the basioccipital
OAS	Junction of basioccipital, ectotympanic and basisphenoid
PIF	Posterior margin of anterior palatine foramen
PM	Point of greatest curvature on the posterior margin of the zygomatic process of the maxilla
PSA	Posterior edge of the alveolus of the upper third molar
PTZ	Posterior temporal–zygomatic junction on zygomatic arch
PZA	Point of greatest curvature along posterior edge of zygomatic process of temporal bone

ces in Procrustes data can be localized to individual landmarks. Secondly, we conduct a parallel analysis based on linear interlandmark distances. These distances are obtained from the raw landmark data prior to Procrustes superimposition.

Among-individual variation

In comparisons of variance, one or a few outliers resulting from misplacement of landmarks, pseudoforamina in the 3D reconstruction, or a pathology unrelated to the one in question can significantly impact statistical comparisons. For this reason, we used Grubb's test for outliers to remove statistically significant ($\alpha = 0.05$) outliers for the among-individual, asymmetry and the measurement error distributions as recommended by Palmer and Strobeck (2003). This resulted in the removal of 53 individual \times landmark \times trial data points out of the total of 13,680 in the dataset (or 0.4% of the total data). The remaining distributions are free of significant outliers

We used two different methods to test the prediction that among-individual variation is increased in the brachymorph

mouse. Following Zelditch et al. (2004), we used the distribution of Procrustes distances from the sample mean as a measure of shape variation. For this analysis, the left side was reflected on to the right and the bilaterally homologous landmarks were averaged. For each individual, we thus calculated an average of two trials and two sides. The variance of the Procrustes deviations from the sex-specific mean landmark configuration of each sample is a measure of shape variance that can be compared across groups using parametric methods such as ANOVA or *t*-tests (i.e., Levene's test for differences in variance). As the Procrustes distances tend not to be normally distributed we also use a permutation test in which the set of calculated residuals is assigned randomly to the two groups. The null hypothesis is that the two groups are drawn randomly from the combined set of Procrustes mean deviations. At each iteration, a difference (ΔV) between the permuted variances is obtained. For comparisons of regional variation, the Procrustes superimposition step was repeated for the landmark set defining each region.

For comparisons of variance for groups of landmarks, the Procrustes superimposition was repeated from raw landmark data. The reason for this is concern over how the Procrustes superimposition distributes variation across landmarks (Lele and McCulloch 2002). The same procedure explained above was then used to compare shape variances for specific regions. We do not assume that variances can be compared directly for individual landmarks or even localized to specific landmarks.

We also compared variances for interlandmark distances, log-transformed to eliminate the size dependence of the variance (Van Valen 1978; Palmer and Strobeck 2003). We then obtained the absolute mean deviation for each distance and averaged this value across distances within individuals for the whole skull or the regions within the skull. The average absolute mean deviation is then compared using Levene's test. For both the interlandmark distance and the Procrustes distance methods, mean deviations were calculated separately by sex so as not to confound among-sex with among-individual variation in size and shape.

FA

For comparisons of FA, we used Klingenberg's (2002) object asymmetry method, assuming equal and isotropic variation across landmarks. This method measures asymmetry by the Procrustes distance between a side and its reflection for each individual. The advantage of the method is that it does not assume that the midline landmarks actually fall on the geometric midline; it thus allows the assessment of asymmetry for landmarks that are not bilateral. To test the statistical significance of differences in FA, the object FA method adapts the Palmer and Strobeck mixed model ANOVA method (Palmer and Strobeck 1986, 2003; Palmer, 1994) with the appropriate adjustment for the degrees of freedom (Klingenberg et al. 2002). Significance is tested using a permutation method as described by Klingenberg (2002). This analysis was run using the SAGE 3D program developed by Marquez (2004b).

Integration and covariation structure

We compared covariation structure for both Procrustes coordinates and interlandmark distances using matrix correlations. This analysis determines whether the increase in variance in the brachymorph mouse is because of variation that is qualitatively different from that seen in the wildtype, which would be manifest by a

significant change in the phenotypic covariance structure. To test the hypothesis that covariance matrices are significantly different, we obtained the distribution of matrix correlations between each matrix and its resampled self, with 1000 resampling iterations with replacement. The P -value is calculated as the number of iterations in which the resampled R_M for the matrix and itself is equal or lower than the observed R_M for the matrices to be compared. The P -value is calculated independently for each of the two matrices and the average is reported. Mantel's test is used to compare the significance of matrix correlations, that is, to test the hypothesis that the covariance structures are unrelated (Mantel 1967). For Procrustes residuals, we used Klingenberg's variant of the Mantel's test which shuffles the x , y , z coordinates for each landmark as single blocks and thus takes into account the non-independence of the three coordinates for each landmark. These procedures were performed using MACE 3D (Marquez 2004a) or equivalent procedures in R for the distance data.

To compare the pattern of covariation among anatomical regions, we used the trace correlation, r_t (discussed in more detail in Mardia et al. 1979), which estimates the strength of the relationship between two blocks of multivariate data measured on the same individuals. One block of variables is assumed to depend on the other, a relationship modeled by the regression equation

$$Y = XB + U \quad (1)$$

where X and Y are the two blocks of variables, B is a vector of regression coefficients, and U is a matrix describing the residual variance. The maximum likelihood estimates of B and U are calculated as

$$B = (X'X)^{-1} \times X'Y \quad (2)$$

$$U = Y - XB \quad (3)$$

The trace correlation r_t is then

$$r_t^2 = (1/p) \text{trace}(\mathbf{I} - D) \quad (4)$$

where \mathbf{I} is the identity matrix and

$$D = (Y'Y)^{-1}U'U \quad (5)$$

and p is the number of response variables in Y , which is $2k - 4$ where k is the number of landmarks. The matrix D is analogous to the unexplained variation so $\mathbf{I} - D$ is analogous to the proportion of the variance in Y that is explained by its dependence on X and it ranges from 0 to 1.

The analysis can be done using the Procrustes residuals (as in Klingenberg et al. 2003) or partial warp scores (including scores on the uniform component). Both are equally valid shape variables. We use partial warp scores because the covariance matrix is of full rank, which is not the case for the matrix of Procrustes residuals. The statistical significance of r_t can be determined by a permutation test (Klingenberg et al. 2003). The null hypothesis is that the correlation between the two blocks is no higher than expected by chance.

To compare overall magnitudes of morphological integration, we followed Wagner (1989) in using the variance of the eigenvalues for the variance-covariance matrix as the measure of integration. We compared integration for Procrustes coordinates before and after removing the allometric component of size (done by regressing each coordinate on centroid size) as well as for the

set of inter-landmark distances defined by the landmark set. These three analyses capture different but overlapping aspects of integration and experience shows that they can produce opposing results. To determine the significance of differences in eigenvalues, we obtained confidence intervals (CIs) for this metric by resampling each dataset with replacement for 1000 iterations; the statistical significance of the difference between two values is calculated from the number of iterations in which the difference between the two resampled values exceeds 0 divided by the total number of iterations.

RESULTS

Figure 2 shows 3D reconstructions from micro-CT scans of representative age-matched individuals in the control and brachymorph groups. Detailed analysis of the changes in mean phenotype is presented elsewhere but a comparison of these individuals clearly shows the dramatic alteration in mean shape produced by the brachymorph mutation.

The comparisons of among-individual variance is also clear: The brachymorph mutation is associated with a dramatic increase in morphological variance (Fig. 3); the Procrustes mean deviations, after adjustment for sex differences within groups, are significantly higher in the brachymorph group compared with the wildtype whether the statistical analysis is based on Levene's test ($df = 39$, $F = 29$, $P < 0.001$) or the permutation test (observed $\Delta V = 0.00097$, 1600 permutations, $P < 0.001$). Similar results were obtained when the data were not sex standardized. Figure 4A shows the landmark scatters for the symmetric component of the variance. The larger landmark scatters in the brachymorph sample reveal the higher variance in this sample compared with the wildtype. Analysis of the Procrustes mean deviation by region (after repeating the Procrustes superimposition step) show that the largest differences in variances occur in the basicranium (brachymorph variance: 0.00014 vs. control variance: 0.000005; 1600 permutations, $P < 0.001$) with significant differences also in the face (brachymorph variance: 0.000003 vs. control variance: 0.000001, 1600 permutations, $P < 0.001$). Figure 5 shows the scatters for the basicranial landmarks for both groups, which demonstrates both the difference in average shape and the greater variances in the brachymorph sample.

Comparisons of variances for interlandmark distances reveal similar results. As shown in Fig. 6, the mean variances of the log-transformed distances are greater in the brachymorph mice both for the whole skull and anatomical regions, with the strongest effect being localized to the basicranial region. Based on Levene's test ANOVA on the mean deviations, averaged across distances within individuals, shows that the brachymorphs have significantly higher variances for the whole skull ($F = 14.8$, $n = 40$, $P < 0.001$), the basicranium

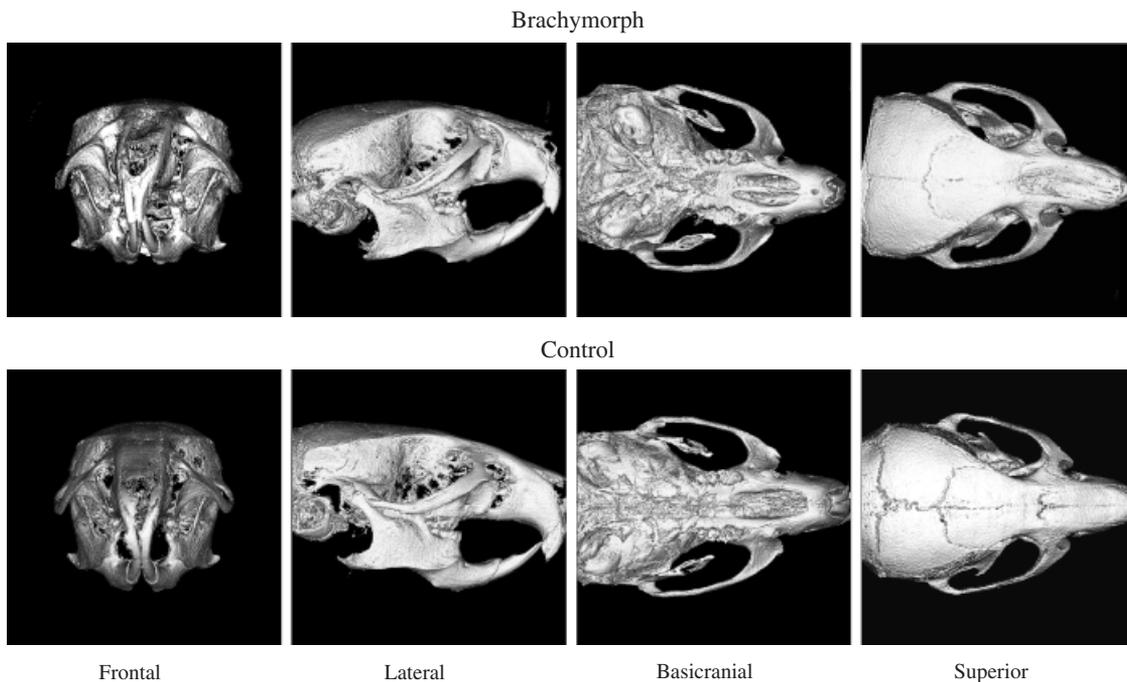


Fig. 2. 3D reconstructions from micro-CT scans of representative individuals from the *bm/bm* and C57BL/6J control groups. Both individuals are shown at the same scale. *bm*, brachymorph.

($F = 18, n = 40, P < 0.001$), the face ($F = 6.1, n = 40, P < 0.05$), and the neurocranium ($F = 7.9, n = 40, P < 0.01$) but not for the zygomatic region ($F = 3.8, n = 40, P = 0.06$).

Analysis of FA using Klingenberg’s object FA method revealed significant FA in both groups as well as small but statistically significant directional asymmetry. The difference in overall object FA between the two groups is not significant (brachymorph: 0.0051 vs. control = 0.0039, $F = 1.3$). Performing independent Procrustes analysis and tests for FA within regions revealed significantly higher FA for facial

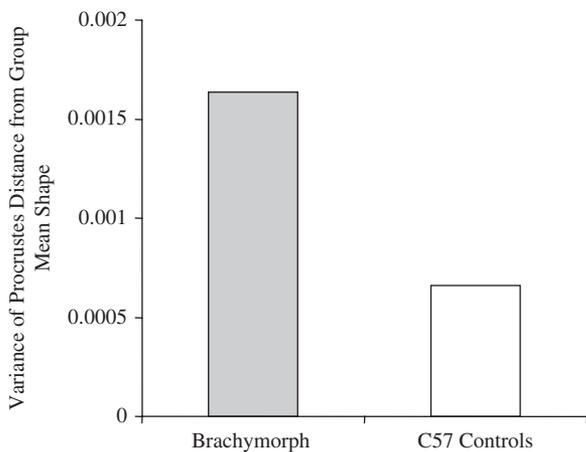


Fig. 3. Variances of Procrustes distance from the mean shape (adjusted for sex) for *bm/bm* mice and C57BL/6J controls. *bm*, brachymorph.

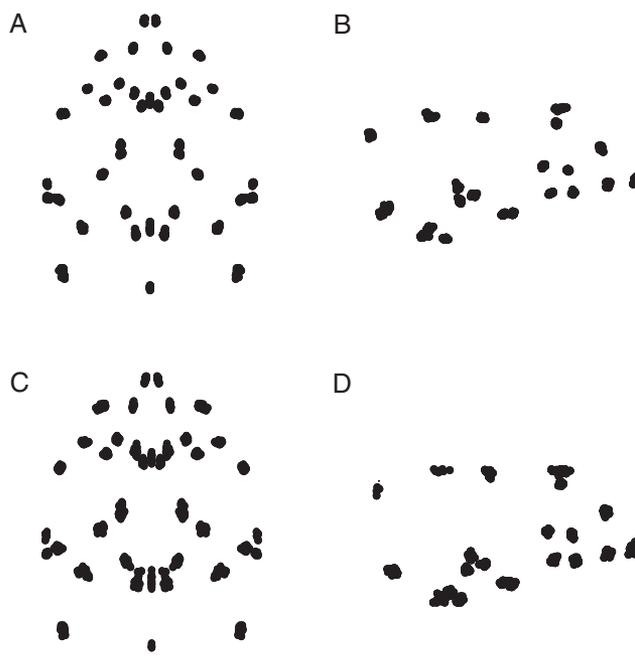


Fig. 4. Landmark scatters after Procrustes superimposition for the averages of the two measurement trials. (A) and (B) show superoinferior and lateral views of the wildtype control whereas (C) and (D) show the same views for the brachymorph group. The data for both sides are shown to aid visualization.

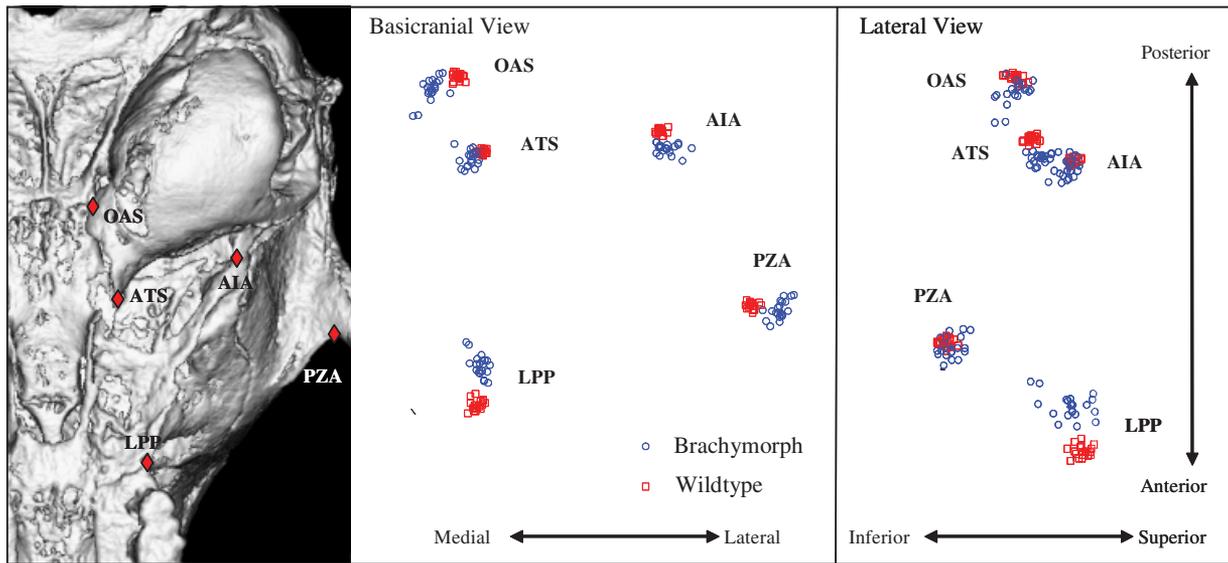


Fig. 5. Scatters of Procrustes coordinates in the basicranial region for the brachymorph and wildtype control samples showing both the differences in mean configuration and the variance about the mean. Each point within each landmark scatter represents a single individual and is obtained by reflecting the left side on to the right and averaging the two sides and the two replicates for each individual. The raw data for this landmark subset for the two samples were combined and subjected to a single general least-squares Procrustes superimposition.

landmarks ($F = 76$, $P < 0.001$) and neurocranial landmarks ($F = 2.14$, $P < 0.05$) in the brachymorph group. Basicranial FA could not be compared as FA was not detectable above measurement error in the brachymorph mice for this region. Thus, there is some evidence for increased FA in the brachymorph group, but the effect is weak and the pattern is not clear and must be interpreted cautiously because these sample sizes are quite small for analyses of FA.

Comparison of the covariance structure showed the covariation structure in the two groups, whereas significantly correlated ($R_M = 0.53$, $P < 0.01$), is also significantly different. The correlation between the two groups falls outside the 95% CI of the values for each matrix repeatedly compared with itself (brachymorph: 0.85 (95% CI = 0.68–0.94), control = 0.84 (95% CI = 0.64–0.94)). We can thus reject the null hypothesis that the two groups are drawn from a pop-

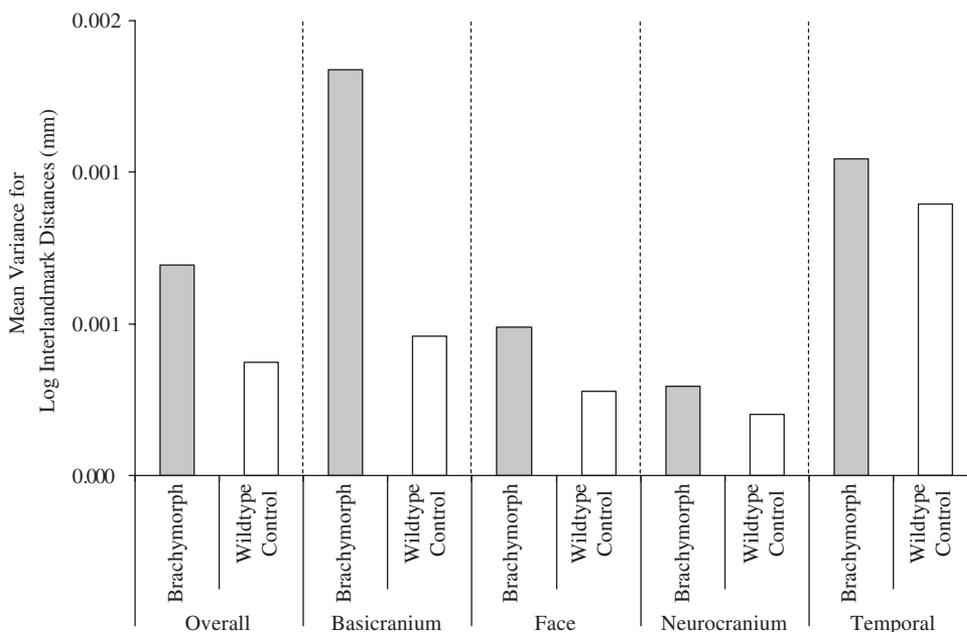


Fig. 6. Mean variances of log inter-landmark distances for both the entire set of distances for the skull and the sets of distances for the four anatomical regions.

ulation with a single covariation structure; the *bm* mutation significantly affects the structure of covariation among morphological components within the cranium.

To visualize the change in covariation structure, Fig. 7 shows wireframe deformations corresponding to variation along the first three principal components obtained from the Procrustes data. For this comparison, the ordering of the first three components is not statistically significant in either sample. Consequently, the relevant comparison between the groups is among all three components taken together. In the brachymorphs, variation in the height of the neurocranium and length and flexion of the basicranium appears to

dominate. By contrast, in the wildtype, there is subtle variation in neurocranial height but more dramatic variation in the length of the face and the anterior neurocranium. These results support the view that the additional variation seen in the brachymorphs is qualitatively different from that which characterizes the wildtype.

Comparison of the pattern of trace correlations between anatomical regions reveals a subtle change in correlations among anatomical regions (Table 2). The groups differ in which correlations are statistically significant; whereas the strongest significant trace correlation in the wildtype is between the basicranium and the face, with a weaker but still

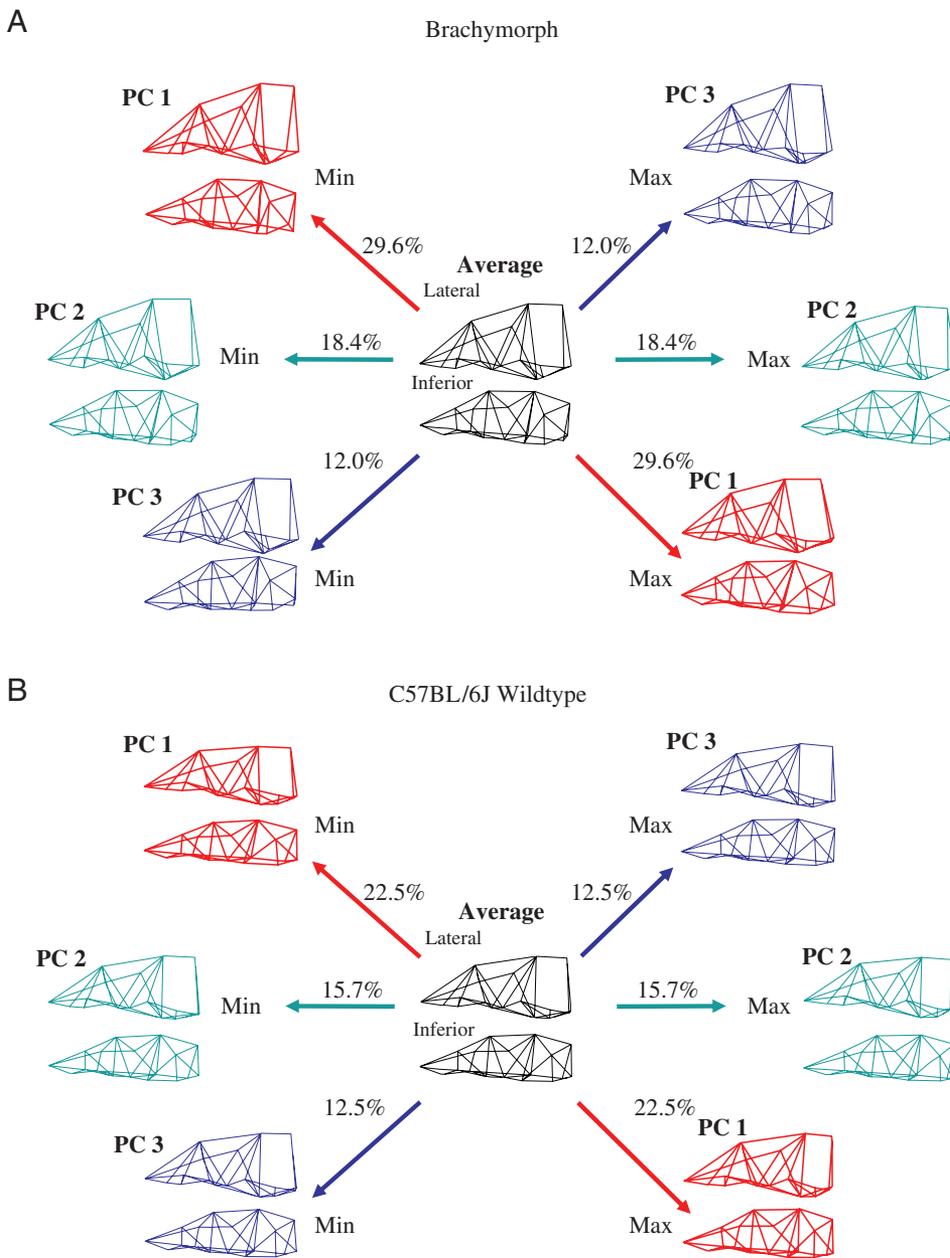


Fig. 7. Visualization of covariance structures for the two groups based on principal components analysis. The wireframe deformations represent variation along principal components. The average landmark configuration is in the middle and the PCs are read diagonally across.

Table 2. Trace correlations between anatomical regions for the Brachymorph and wildtype control groups

Block 1	Block 2	<i>r</i>	<i>P</i>	Block 1	Block 2	<i>r</i>	<i>P</i>
<i>Control</i>							
Basi	Face	0.636	0.150	Face	Basi	0.962	0.160
Basi	Zygomatic	0.557	0.860	Zygomatic	Basi	0.421	0.850
Basi	Neural	0.677	0.020	Neural	Basi	0.810	0.010
Neural	Face	0.760	0.038	Face	Neural	0.961	0.044
Neural	Zygomatic	0.692	0.828	Zygomatic	Neural	0.437	0.838
Zygomatic	Face	0.485	0.114	Face	Zygomatic	0.969	0.124
<i>Brachymorph</i>							
Basi	Face	0.603	0.040	Face	Basi	0.912	0.040
Basi	Zygomatic	0.652	0.104	Zygomatic	Basi	0.493	0.108
Basi	Neural	0.639	0.018	Neural	Basi	0.764	0.010
Neural	Face	0.693	0.435	Face	Neural	0.876	0.411
Neural	Zygomatic	0.482	0.082	Zygomatic	Neural	0.482	0.720
Zygomatic	Face	0.479	0.001	Face	Zygomatic	0.958	0.001

Significant P-values ($\alpha < 0.05$) indicated in bold.

significant correlation between the basicranium and neurocranium, the significant correlations in the brachymorph are between basicranium and neurocranium, zygomatic region and face, and basicranium and face. Nevertheless, despite these differences in the correlations that are judged statistically significant, the estimated trace correlations are very similar. Evidently, the *Papps2* mutation produces morphological variation different from that normally seen in the wildtype and either alters the covariation patterns or else increases the range of correlations in the brachymorph and thereby complicates statistical testing.

The variances of the eigenvalues of the covariance matrices show that the *bm* mutation dramatically increases the magnitude of morphological integration in the skull; integration is significantly higher ($P < 0.01$) for all three matrices tested (Fig. 8). The nearly identical results obtained before and after removing the allometric component of shape (Fig. 8, A and B) shows that the increase in integration is not related to allometric integration. The similar results obtained from the interlandmark distance data indicates that the increase in integration is true for both the shape and size of morphological components.

DISCUSSION

This study tests the hypothesis that a mutation which alters craniofacial shape by decreasing cartilage growth via interference with the extracellular matrix molecule sulfation process also increases phenotypic variance and alters phenotypic covariances in the mouse skull. The two groups compared in this study differ with respect to a homozygous base-pair substitution in the *Papps2* gene, which leads to a dramatic decrease in cartilage matrix molecule sulfation and abnormal

cartilage growth. In mice homozygous for the *bm* mutation, phenotypic variance for both size and shape is dramatically increased, especially in the basicranium. Additionally, phenotypic covariation structure is significantly altered by the *bm* mutation; the brachymorph phenotype is associated with qualitatively new morphological variation, even though major aspects of covariation structure are basically conserved. Moreover, phenotypic integration is substantially increased meaning that the increased variation is not random but rather highly coordinated throughout the skull.

The primary caveat on these results is that the brachymorph and wildtype samples comprise mice with same genetic background, raised under the same conditions, but maternal or among-litter effects could potentially confound these results. However, given that the increase found in both variance and morphological integration is larger than that found in other comparisons of mouse strains, or even distantly related mammalian species (Hallgrímsson et al. 2004a, b), it is extremely unlikely that these factors contribute to the observed differences.

Increased phenotypic variances because of mutations have been reported in the literature, but only rarely tied to a specific developmental cause. The best examples come from *Drosophila* genetics. Repeating Waddington's (1942) classic bithorax experiment, Gibson and Hogness (1996) showed that a null mutation in the *Ubx* protein, which produces the ultrabithorax phenotype, generates the expression of genetic variance not found in the wildtype. Similarly, Dworkin (2005) shows that the *Dll* null mutation increases genetic variance within the mutant group for the prothoracic leg. In an earlier study, Dworkin et al. (2003) linked a gain of function mutation in *Egfr* to an increase in the eye morphology variation resulting from polymorphisms in the gene itself. Finally, Trotta et al. (2005) showed using random genomic insertions

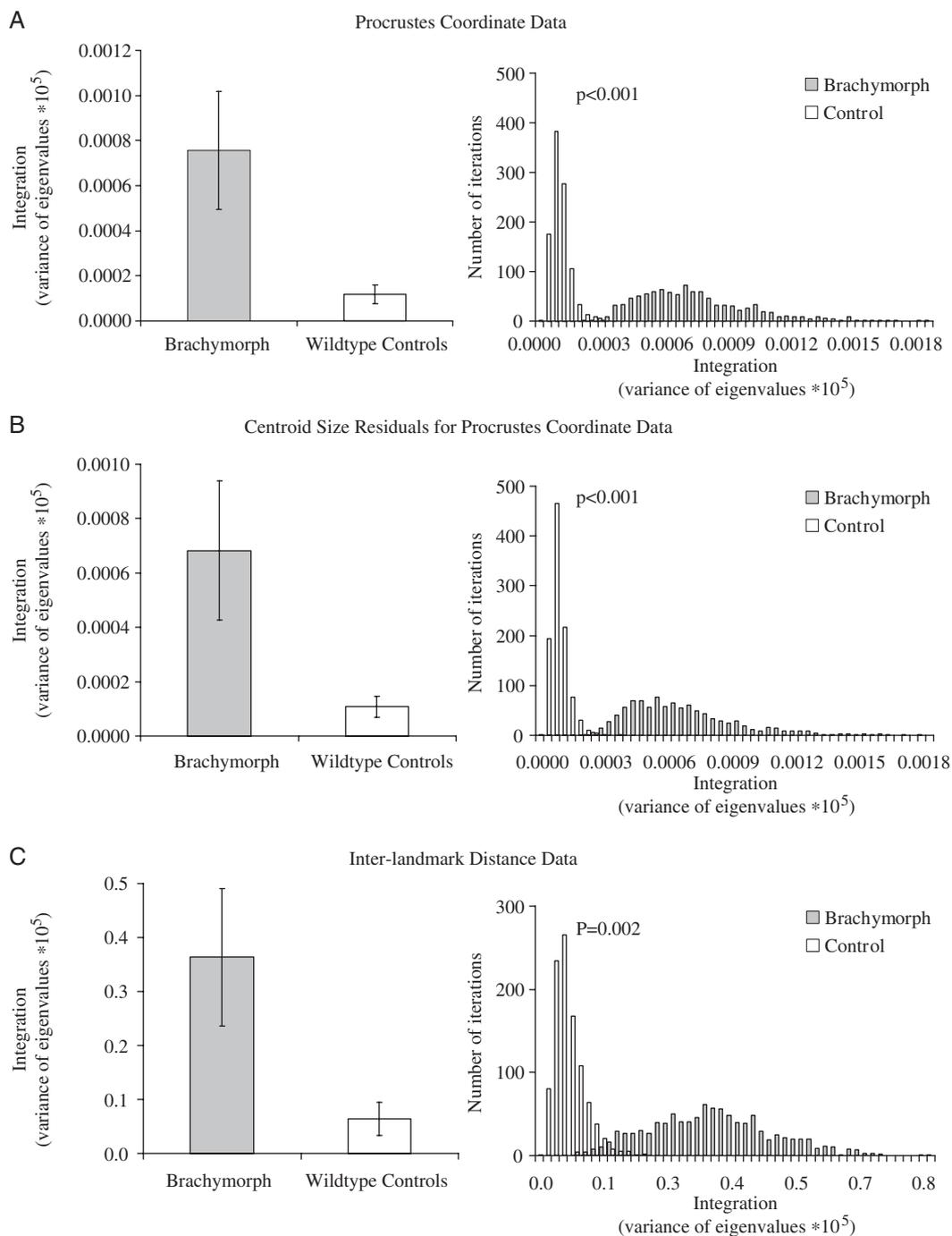


Fig. 8. Comparisons of three aspects of morphological integration. (A) Integration of shape as measured by Procrustes coordinate data. (B) Shape integration after removal of the allometric component of shape and (C) form (size+shape) integration based on interlandmark distance data. The bar chart shows the integration values for the two groups whereas the histogram shows the resampling distributions for both samples. All comparisons are significant at $P < 0.01$ based on 1000 resampling iterations.

that the magnitude of developmental instability produced depends on the magnitude of the morphological perturbation induced. In all these cases, differences between wildtype and mutants in degree of canalization or developmental stability arise from the developmental architecture of the trait itself.

That contrasts with the hypothesis of Rutherford and Lindquist (1998) which suggests an organism-wide canalizing mechanism based on molecular chaperone proteins. Hsp90 and related molecular chaperones are plausible candidates for that role because they encode for proteins that increase pro-

tein structural integrity and thus stabilize diverse developmental pathways (Rutherford 2000; Ruden et al. 2005). For example, Hsp90 stabilizes the function of integrin linked kinase that is involved in signaling pathways regulating cell proliferation and cell adhesion in various developmental contexts (Aoyagi et al. 2005). *Drosophila* genetics thus provides plausible examples of both local and global mechanisms that may modulate phenotypic variability.

The lack of any significant increase in FA because of the *bm* mutation is surprising. This may result from the low statistical power for comparisons of FA in relatively small samples. However, even if that is the case, it suggests that differences in FA, if present, are likely to be fairly subtle. Evidently, canalization is decreased but developmental stability is not. This contrast is important in light of the debate regarding the relationship between canalization and developmental stability (e.g., Debat et al. 2000; de Visser et al. 2003; Willmore et al. 2005).

In mice, a study of a mouse model for Down syndrome shows that increased variation of craniofacial structures results from a significant chromosomal defect (Richtsmeier et al. 2000). Only one specific mutation with a known developmental function has previously been shown to influence phenotypic variability in mice or other vertebrates. The tabby mutation reduces the numbers of vibrissae and generates increased variation in the mutant group (Dun and Fraser 1958). The tabby mouse carries a mutation in ectodysplasin-A (*Eda*), which is a TNF-related ligand involved in the epithelial–mesenchymal interactions in the formation of ectodermal outgrowths (Srivastava et al. 1997) including teeth (Kangas et al. 2004). Although it is not clear that variation *Eda* expression influences variability of dental morphology, the supplementary analysis provided with the Kangas et al. study suggests that they may.

We show that a lack-of-function mutation in the *Paps2* gene dramatically increases both phenotypic variance and morphological integration. More importantly, both those increases can be reasonably traced to the function of the gene. Sulfation of proteoglycans is dramatically reduced in the brachymorph mouse; production of PAPS through the APS kinase pathway is reduced to 10% of the wildtype, and PAPS synthetase activity is reduced to just 2% (Kurima et al. 1998). Normal cartilage growth depends on sulfation of proteoglycans and multiple mutations can reduce cartilage growth via this mechanism. For example, interfering with the function of the *SLC26A2* sulfate transporter gene, which is required for cellular uptake of inorganic sulfate, also reduces cartilage growth and results in a phenotype apparently similar to that of the brachymorph mouse (Forlino et al. 2005). We assume that degree of sulfation of proteoglycans in cartilage, like any other phenotypic trait, varies among individuals in both mutant and wildtype populations, which is demonstrated in the case of the *SLC26A2* mutation (Forlino et al. 2005). That

variation has not been reported in the case of brachymorph mice because only means are given in the literature (Sugahara and Schwartz 1979; Kurima et al. 1998). We hypothesize that variation within the wildtype in degree of sulfation does not translate into phenotypic variation because the level of sulfation is sufficient for normal growth; in contrast, sulfation in the mutants is reduced to a level incompatible with normal cartilage growth. Therefore, the amount of cartilage growth achieved depends on the degree of sulfation, which varies among individuals, and that variation is translated into variation in cartilage growth and thus to variation in cranial shape. The source of inter-individual variation in sulfation is not known. If variation in sulfation does not translate to variation in cartilage growth in the wildtype, this variation is normally cryptic. Variances are additive (Van Valen 1978), and thus the addition of sulfation related variation in the brachymorph phenotype increases overall phenotypic variation.

An alternative but complementary possibility is that the degree of sulfation may impact growth to different degrees in different individuals for mechanical reasons. In other words, there may be a stochastic element in the degree to which reduced sulfation reduces growth. This mechanism, however, should reduce integration and increase FA as the resulting variation would be randomly distributed throughout the skull which is clearly not the case in this study.

In these scenarios, integration increases for reasons directly related to the increase in variance, namely, that the variation introduced into cartilage growth by the brachymorph mutation has secondary (epigenetic) effects throughout the skull. Examination of the covariation pattern within the brachymorph sample shows that the degree of cranial vaulting is correlated with the degree to which the growth of the basi-cranium is retarded. These results suggest that the distinctive features of the brachymorph cranial shape are all consequences of the reduced growth of the chondrocranium, and the degree to which those features develop depends on the degree to which chondrocranial growth is retarded.

The scenario outlined above, if correct, provides two important insights into the developmental-genetics of phenotypic variability. The first is that the genetic and developmental mechanisms that increase the expressed phenotypic variance (and thus reduce canalization) do not have an impact on variability in general but rather are specific to a particular developmental context. In this case, the genetic basis for increased variance is the same as that responsible for the altered mean phenotype. Secondly, the developmental basis for the increased integration and decreased canalization is quite straightforward and even mundane; variation that is cryptic in the wildtype (i.e., the degree of sulfation) is “revealed” in the mutant, but that revelation stems in our explanation simply from a threshold effect or a nonlinear relationship between degree of sulfation and cartilage growth. Similarly, the in-

crease in integration flows obviously from the developmental relationships among craniofacial components.

There is another potentially important factor not captured by the scenario outlined above, which may contribute to the increased variability of the brachymorph mutant. That is the loss of redundancy in developmental regulation. Redundancy and complexity potentially reduce variation by increasing the number of independently varying components that regulate and determine the behavior of a system. Redundancy offers protection against a single mutation because of the additional routes to normal (wildtype) outcome. Although PAPS is the sole sulfate donor for cartilage proteoglycans (Xu et al. 2003), there is some redundancy between PAPPS2 and the closely related enzyme, PAPPS1 in the PAPS synthesis pathway (Xu et al. 2003). This redundancy may account for the fact that some sulfation does occur in the brachymorph mouse.

We have shown that a specific mutation with a known developmental effect that alters the phenotype increases both variance and integration and alters some aspects of covariation but does not affect fluctuating in the mouse skull. All these outcomes can be traced to the loss-of-function mutation. By linking gene function to variational properties we provide insights into two important questions in evolutionary-developmental biology: (1) how do developmental systems modulate the expression genetic variation and (2) how does developmental architecture determine variability? These are important questions because natural selection acts on phenotypes, but its results are interpreted and transmitted from generation to generation at the genetic level. In the example presented here, phenotypic variability arises from the same developmental mechanisms that produce the alteration in phenotypic mean. That does not imply that stabilizing selection acts on those mechanisms and thereby produces canalization. However, it does support the more general view that variability can arise from intrinsic properties of developmental genetic architecture such as nonlinearities, epistatic complexity, thresholds and redundancy, as well as the roles played by genes within networks (Siegal and Bergman 2002). Our findings, although not incompatible with the existence of global canalizing mechanisms of the kind found by Rutherford and Lindquist (1998) and Rutherford (2000), show that mechanisms of canalization need not be either global or special. If canalization results from developmental architecture rather than special canalizing processes, it can be affected by any mutation that significantly alters the relevant properties of developmental architectural. As the outcome of an epigenetic system, canalization can only be understood in terms of the roles that genes play within a developmental system.

Acknowledgments

We are grateful to Wei Liu for technical assistance and to Mitzi Murray for administrative support. We thank Eladio Marquez for writing and making available the Mace3D program and Andrew Ah-

Seng for writing a 3D median filtering plug-in for ImageJ and Peter Kublik for writing the StackPro utility. We thank Brian Hall, Daniel Lieberman, Nathan Young, Katherine Willmore, David Cooper, Virginia Diewert, Rebecca German, and John Matyas for discussions.

Supporting Grant Information: National Science and Engineering grant 238992-02, Canadian Foundation for Innovation grant #3923, Alberta Innovation and Science grant #URSI-01-103-RI to B. H. and a Canadian Institutes of Health Research Grant (to F. R. J.). F. R. J. held a Canada Research Chair as well as an Alberta Heritage Foundation for Medical Research Award.

REFERENCES

- Aoyagi, Y., Fujita, N., and Tsuruo, T. 2005. Stabilization of integrin-linked kinase by binding to Hsp90. *Biochem. Biophys. Res. Commun.* 331: 1061–1068.
- Bergman, A., and Siegal, M. L. 2003. Evolutionary capacitance as a general feature of complex gene networks. *Nature* 424: 549–552.
- de Visser, J. A., et al. 2003. Perspective: evolution and detection of genetic robustness. *Evolution* 57: 1959–1972.
- 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.
- Dun, R. B., and Fraser A. S. 1958. Selection for an invariant character—“vibrissae number”—in the house mouse. *Nature* 181: 1018–1019.
- Dworkin, I. 2005. Evidence for canalization of Distal-less function in the leg of *Drosophila melanogaster*. *Evol. Dev.* 7: 89–100.
- Dworkin, I., Palsson, A., Birdsall, K., and Gibson, G. 2003. Evidence that Egrf contributes to cryptic genetic variation for photoreceptor determination in natural populations of *Drosophila melanogaster*. *Curr. Biol.* 13: 1888–1893.
- Ford-Hutchinson, A., et al. 2005. Degenerative knee joint disease in mice lacking 3'-phosphoadenosine 5'-phosphosulfate synthetase 2 (Paps2) activity: a putative model of human PAPSS2 deficiency-associated arthritis. *Osteoarthritis Cartilage* 13: 418–425.
- Forlino, A., et al. 2005. A diastrophic dysplasia sulfate transporter (SLC26A2) mutant mouse: morphological and biochemical characterization of the resulting chondrodysplasia phenotype. *Hum. Mol. Genet.* 14: 859–871.
- Gibson, G., and Hogness, D. S. 1996. Effect of polymorphism in the *Drosophila* regulatory gene Ultrabithorax on homeotic stability. *Science* 271: 200–203.
- Hallgrímsson, B., Dorval, C. J., Zelditch, M. L., and German, R. Z. 2004a. Craniofacial variability and morphological integration in mice susceptible to cleft lip and palate. *J. Anat.* 205: 501–517.
- Hallgrímsson, B., Willmore, K., Dorval, C., and Cooper, D. M. 2004b. Craniofacial variability and modularity in macaques and mice. *J. Exp. Zool. Part B Mol. Dev. Evol.* 302: 207–225.
- Hermisson, J., and Wagner, G. P. 2004. The population genetic theory of hidden variation and genetic robustness. *Genetics* 168: 2271–2284.
- Kangas, A. T., Evans, A. R., Thesleff, I., and Jernvall, J. 2004. Nonindependence of mammalian dental characters. *Nature* 432: 211–214.
- Kaufman, M. H., and Bard, J. B. L. 1999. *The Anatomical Basis of Mouse Development*. Academic Press, San Diego.
- Klingenberg, C. P., et al. 2003. Developmental integration in a complex morphological structure: how distinct are the modules in mouse mandible? *Evol. Dev.* 5: 522–531.
- Klingenberg, C. P., Barluenga, M., and Meyer, A. 2002. Shape analysis of symmetric structures: quantifying variation among individuals and asymmetry. *Evol. Int. J. Org. Evol.* 56: 1909–1920.
- Kurima, K., et al. 1998. A member of a family of sulfate-activating enzymes causes murine brachymorphism. *Proc. Natl. Acad. Sci. USA* 95: 8681–8685.
- Lane, P. W., and Dickie, M. M. 1968. Three recessive mutations producing disproportionate dwarfing in mice: achondroplasia, brachymorphic, and stubby. *J. Hered.* 59: 300–308.

- Lele, S., and McCulloch, C. E. 2002. Invariance, identifiability and morphometrics. *J. Am. Stat. Assoc.* 97: 796–806.
- Lele, S., and Richtsmeier, J. T. 2001. *An Invariant Approach to the Statistical Analysis of Shapes*. Chapman & Hall, Boca Raton.
- Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27: 209–220.
- Mardia, K. V., et al. 1979. *Multivariate Analysis*. Academic Press, London.
- Marquez, E. 2004a. *MACE 3D*. University of Michigan, Ann Arbor, MI.
- Marquez, E. 2004b. *SAGE 3D*. University of Michigan, Ann Arbor, MI.
- Orkin, R. W., Pratt, R. M., and Martin, G. R. (1976). Undersulfated chondroitin sulfate in the cartilage matrix of brachymorphic mice. *Dev. Biol.* 50: 82–94.
- Palmer, A. R. 1994. Fluctuating asymmetry analyses: a primer. In T. A. Markow (ed.), *Developmental Instability: Its Origins and Evolutionary Implications*. Kluwer Academic Publishers, Dordrecht, pp. 355–364.
- Palmer, A. R., and Strobeck, C. 1986. Fluctuating asymmetry: measurement, Analysis, Patterns. *Ann. Rev. Ecol. Systematics* 17: 391–421.
- Palmer, A. R., and Strobeck, C. 2003. Fluctuating asymmetry analysis unplugged. In M. Polak (ed.), *Developmental Instability (DI): Causes and Consequences*. Oxford University Press, Oxford, pp. 279–319.
- Richtsmeier, J. T., Baxter, L. L., and Reeves, R. H. 2000. Parallels of craniofacial maldevelopment in Down syndrome and Ts65Dn mice. *Dev. Dyn.* 217: 137–145.
- Ruden, D. M., Xiao, L., Garfinkel, M. D., and Lu, X. 2005. Hsp90 and environmental impacts on epigenetic states: a model for the trans-generational effects of diethylstilbestrol on uterine development and cancer. *Hum. Mol. Genet.* 14 (Spec. No. 1): R149–R155.
- Rutherford, S. L. 2000. From genotype to phenotype: buffering mechanisms and the storage of genetic information. *Bioessays* 22: 1095–1105.
- Rutherford, S. L., and Lindquist, S. 1998. Hsp90 as a capacitor for morphological evolution. *Nature* 396: 336–342.
- Scharloo, W. 1991. Canalization: genetic and developmental aspects. *Ann. Rev. Ecol. Systematics* 22: 65–93.
- Siegal, M. L., and Bergman, A. (2002). Waddington's canalization revisited: developmental stability and evolution. *Proc. Natl. Acad. Sci. USA.* 99: 10528–10532.
- Slice, D. E. 1994–1999. *Morpheus*. Stony Brook, NY.
- Slice, D. E. 2005. Modern morphometrics. In D. E. Slice (ed.), *Modern Morphometrics in Physical Anthropology*. Kluwer, New York, pp. 1–45.
- Srivastava, A. K., et al. 1997. The Tabby phenotype is caused by mutation in a mouse homologue of the EDA gene that reveals novel mouse and human exons and encodes a protein (ectodysplasin-A) with collagenous domains. *Proc. Natl. Acad. Sci. USA* 94: 13069–13074.
- Sugahara, K., and Schwartz, N. B. 1979. Defect in 3'-phosphoadenosine 5'-phosphosulfate formation in brachymorphic mice. *Proc. Natl. Acad. Sci. USA.* 76: 6615–6618.
- Trotta, V., Garoia, F., Guerra, D., Pezzoli, M. C., Grifoni, D., and Cavicchi, S. 2005. Developmental instability of the *Drosophila* wing as an index of genomic perturbation and altered cell proliferation. *Evol. Dev.* 7: 234–243.
- Van Valen, L. M. 1978. The statistics of variation. *Evol. Theory* 4: 33–43.
- Waddington, C. H. 1942. The canalisation of development and the inheritance of acquired characters. *Nature* 150: 563.
- Waddington, C. H. 1957. *The Strategy of the Genes*. MacMillan Company, New York.
- Wagner, G. P. 1989. A comparative study of morphological integration in *Apis mellifera* (Insecta, Hymenoptera). *Z. Zool. Syst. Evol.forsch.* 28: 48–61.
- Wagner, G. P. 2003. Evolutionary genetics: the nature of hidden genetic variation unveiled. *Curr. Biol.* 13: R958–R960.
- Wagner, G. P., Booth, G., and Bagheri-Chaichian, H. 1997. A population genetic theory of canalization. *Evolution* 51: 329–347.
- Wilkins, A. S. 1997. Canalization: a molecular genetic perspective. *Bioessays* 19: 257–262.
- Willmore, K. E., Klingenberg, C. P., and Hallgrímsson, B. 2005. The relationship between fluctuating asymmetry and environmental variance in rhesus macaque skulls. *Evol. Int. J. Org. Evol.* 59: 898–909.
- Xu, Z., Wood, T. C., Adjei, A. A., and Weinshilboum, R. M. 2001. Human 3'-phosphoadenosine 5'-phosphosulfate synthetase: radiochemical enzymatic assay, biochemical properties, and hepatic variation. *Drug Metab. Dispos.* 29: 172–178.
- Xu, Z. H., et al. 2000. Human 3'-phosphoadenosine 5'-phosphosulfate synthetase 1 (PAPSS1) and PAPSS2: gene cloning, characterization and chromosomal localization. *Biochem. Biophys. Res. Commun.* 268: 437–444.
- Xu, Z. H., Thoma, B. A., Eckloff, B. W., Wieben, E. D., and Weinshilboum, R. M. 2003. Pharmacogenetics of human 3'-phosphoadenosine 5'-phosphosulfate synthetase 1 (PAPSS1): gene resequencing, sequence variation, and functional genomics. *Biochem. Pharmacol.* 65: 1787–1796.
- 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.