J. Anat. (2004) **205**, pp501–517

Craniofacial variability and morphological integration in mice susceptible to cleft lip and palate

Benedikt Hallgrímsson,[1] Curtis J. Dorval,[2] Miriam Leah Zelditch[3] and Rebecca Z. German[4]

[1]*Department of Cell Biology & Anatomy and the Joint Injury and Arthritis Research Group, University of Calgary, Canada*
[2]*Department of Biological Sciences, University of Calgary, Canada*
[3]*Museum of Paleontology, University of Michigan, USA*
[4]*Department of Biological Sciences, University of Cincinnati, USA*

Abstract

A/WySnJ mice are an inbred strain that develops cleft lip with or without cleft palate (CL/P) with a frequency of 25–30% and a predominantly unilateral expression pattern. As in humans, the pattern of incomplete penetrance, and variable and frequent unilateral expression suggests a role for altered regulation of variability (developmental stability, canalization and developmental integration) during growth. We compared both mean and variability parameters for craniofacial shape and size among A/WySnJ mice, a strain that does not develop CL/P (C57BL/6J) and their F1 cross. We show that adult A/WySnJ mice that do not express cleft lip exhibit decreased morphological integration of the cranium and that the co-ordination of overall shape and size variation is disrupted compared with both C57BL/6J mice and the F1 cross. The decrease in integration is most pronounced in the palate and face. The absence of this pattern in the F1 cross suggests that it is determined by recessive genetic factors. By contrast, the shape differences between the strains, which are thought to predispose A/WySnJ mice to CL/P, show a range of dominance which suggests a polygenic basis. We suggest that decreased integration of craniofacial growth may be an aetiological factor for CL/P in A/WySnJ mice.

Key words A/WySnJ mice; canalization; cleft lip; craniofacial malformation; developmental stability; morphological integration.

Introduction

Clefting of the primary palate (CL) occurs when the maxillary prominence and the medial and lateral nasal prominences fail to fuse during the formation of the face (Kaufman & Bard, 1999). Aetiological heterogeneity and complex gene–environment interactions are clearly characteristic of the CL malformation (Wyszynski et al. 1996; Prescott et al. 2001). The mechanistic causes of non-syndromic CL fall into two groups. Some produce clefting by affecting growth rates of craniofacial components and thus reducing the time or area for fusion (Fraser & Pashayan, 1970; Hermann et al. 1999) whereas others interfere with the process of facial process fusion (Jara et al. 1995; Pezzetti et al. 1998; Sözen et al. 2001). Diverse environmental factors are associated with CL (Wyszynski & Beaty, 1996), and different genetic causes are likely to have different environmental interactions (Prescott et al. 2001). The expression of developmental variability may also play a role in this malformation. Concordance for cleft lip in monozygotic twins is less than 60% (Lin et al. 1999; Wyszynski & Beaty, 2002), suggesting a role for intrinsic sources of variation. In addition, many syndromes that include CL also exhibit increased phenotypic variation (Khan et al. 1986; Lacombe et al. 1995; Kondo et al. 2002). In humans, there is some evidence for an association between CL/P and developmental instability (Kobyliansky et al. 1999; Neiswanger et al. 2002) and parents of children with non-syndromic CL show increased phenotypic variances and asymmetry for craniofacial measurements (AlEmran et al. 1999). In mice, increased variance in the growth of the facial processes in CL/Fr embryos, which develop CL/P with elevated frequency, has been described but not quantified (Millicovsky et al. 1982).

Correspondence
Dr Benedikt Hallgrímsson, Department of Cell Biology & Anatomy, University of Calgary, 3330 Hospital Dr, Calgary, AB T2N 4 N1, Canada. T: +1 403 220 3060; F: +1 403 210 9747; E: bhallgri@ucalgary.ca

Accepted for publication *8 October 2004*

The regulation of variability in development may play an underappreciated role in determining phenotypic heterogeneity in the expression of dysmorphology. Mutations and environmental effects can alter not only the phenotypic mean but also the variance about the mean (Gibson & Wagner, 2000). Shifts in mean and changes in variance can interact to produce a malformation with incomplete penetrance. For CL, if a population mean is shifted towards the threshold, increased variation in the timing of the specific events in the fusion process can push individuals outside the critical window of time in which the event must occur. For example, if the variation in growth rates is very high, the facial prominences in some individuals will fail to achieve sufficient contact to fuse, even though the mean phenotype may be only shifted towards but not over the threshold for CL formation. A similar argument can be made for the co-ordination of development. During the formation of the primary palate, the nasal and maxillary processes are growing towards one another within a head that is expanding due to facial and neural tube (forebrain) growth. If the integration of growth in the nasal or maxillary prominences with facial width is reduced, some individuals will fail to attain the critical degree of apposition for successful primary palate formation.

Three well-documented evolutionary processes are relevant to the interplay between variability and the generation of dysmorphology. Canalization refers to the minimization of the effects of genetic or environmental variation on among-individual variation (Waddington, 1942) whereas developmental stability (DS) refers to the minimization of variation within the same genetic and environmental conditions (Van Valen, 1962). DS is measured by the subtle deviations from symmetry, or fluctuating asymmetry (FA) (Van Valen, 1962; Leamy et al. 2002), whereas canalization is measured by changes in the magnitude of among-individual variation that are attributable to genetic or environmental impacts. This variation, both within an individual and among individuals, has a genetic basis (Scharloo, 1991; Wagner et al. 1997; Rutherford & Lindquist, 1998; Rutherford, 2000; Hallgrímsson et al. 2002; Ozbudak et al. 2002). The third phenomenon, morphological integration (MI) refers to the degree of co-ordination among growth in size and shape among component parts of an organism (Olson & Miller, 1958; Cheverud, 1982, 1996; Hallgrímsson et al. 2002). MI is conventionally measured by the covariation of phenotypic measurements of size. Low integration reflects uncoordinated deviations from the average growth trajectories of developmental components responding to different genetic, developmental and/or environmental factors. Structures of the face, for example, are more integrated with each other than they are to structures of the neurocranium or basicranium (Cheverud, 1982, 1995; Ackermann & Cheverud, 2000; Hallgrimsson et al. 2004).

Four closely related 'A' strains of mice, including A/WySnJ, have high background rates of CL ranging, by strain, from about 4 to 30% of live births (Juriloff et al. 2001; Diewert & Lozanoff, 2002). Morphometric studies of embryonic development have shown that in A/WySnJ mice, the facial prominences are late to appose and exhibit a reduced degree of contact as compared with other strains (Wang & Diewert, 1992; Wang et al. 1995). Wang and Diewert (Diewert & Wang, 1992; Wang & Diewert, 1992) have shown that the reduced area and time of contact between the facial processes result from reduced growth of the maxillary process. Juriloff et al. (2001, 2004) have identified two recessive factors (*clf1* and *clf2*) that interact epistatically to produce CL in A strain mice, although it is not currently known how these alleles affect craniofacial development. A/WySnJ mice are homozygous for both known alleles. Although MSX1 is not a candidate for either *clf1* or *clf2*, it has been shown that this gene is misexpressed at the tips of the facial processes in all A/WySnJ mice and may thus be downstream of one of the *clf* factors (Gong, 2001). This gene is also implicated in human CL/P (Lidral et al. 1998).

It is not known why some A/WySnJ individuals develop CL/P whereas others do not or why the expression of the trait is often unilateral. Low penetrance and unilateral expression are also characteristic of CL/P in humans. The A/WySnJ strain has been maintained by brother–sister mating for 249 generations (www.jax.org). Genetic variation within the strain is therefore minimal. Phenotypic heterogeneity in the expression of CL/P including incomplete penetrance therefore must be due to the propensity of this strain to exhibit environmentally induced or internally generated phenotypic variation. In other words, phenotypic heterogeneity in A/WySnJ mice for the CL trait must be due to developmental instability, sensitivity to environmental variation or lack of integration among relevant craniofacial components.

There are two ways in which phenotypic variability can influence the expression of CL/P in A/WySnJ mice.

One possibility is that the phenotypic means for underlying determinants of the CL malformation are shifted so close to the threshold for formation of the dysmorphology that typical levels of variability are sufficient to push some individuals over threshold on one or both sides. The other possibility is that variability is actually elevated in this strain in addition to a shift of the phenotypic mean towards the threshold. This study addresses the hypothesis that phenotypic variability is relevant to the formation of cleft lip in mice. If increased phenotypic craniofacial variability is one of the factors that predisposes A/WySnJ mice to expressing CL, then this should be characteristic of the strain and not just those individuals that exhibit CL. We test this prediction in a sample of adolescent (30-day-old) A/WySnJ mice.

Craniofacial shape has been implicated as an aetiological factor for cleft lip in embryonic mice (Trasler, 1968; Juriloff & Trasler, 1976) and in humans (Fraser & Pashayan, 1970; Hermann et al. 1999). Wide faces in relation to cranial height and length are associated with increased susceptibility to CL both within (Yoon et al. 2003, 2004) and among (Chung & Kau, 1985) human populations. A secondary aim of this study was to determine whether the craniofacial shape components (i.e. wide faces and short crania) that characterize A/WySnJ mice and are thought to be associated with CL exhibit a recessive pattern of inheritance. This finding would be consistent with the likely possibility that those shape components are due to the effects of the *clf* loci.

Materials and methods

C57BL/6J mice are an appropriate strain for comparison of variability as this strain has been the standard in previous studies of palate development (Ciriani & Diewert, 1986; Wang et al. 1995). Strains more closely related to A/WySnJ mice such as the other A strains probably share some of the genetic factors that may influence both the mean developmental configuration which predisposes these mice to clefts as well as the possible increase in variability which may underlie the phenotypic heterogeneity exhibited by these mice in the expression of dysmorphology.

A/WySnJ and C57BL/6J breeder pairs were obtained from Jackson Laboratories. The C57BL/6J mice are inbred generation F220, whereas the A/WySnJ are at F249. Both strains can be assumed to have minimal genetic variances. All mice were fed standard diets *ad libitum*, and housed in the same room at the University of Calgary Animal Resources Centre. Male/female pairs were mated and checked every 24 h for the appearance of a vaginal plug. To produce the F1 crosses, A/WySnJ males were mated with C57BL/6J females so as to exclude maternal effects on the CL trait as reported by Juriloff et al. (2001). Litters were weaned at 20 days and killed at 30 days by CO_2 asphyxiation. Heads were frozen and stored for micro-CT scanning. The sample analysed consists of groups of 50 individuals from the A/WySnJ and C57BL/6J parental strains and 29 individuals from the F1 cross. The number of litters in the A/WySnJ, C57BL/6J and F1 cross samples are 13, eight and three, respectively.

Heads were thawed at room temperature and then scanned using a Skyscan 1072 100-kV microtomograph using a protocol optimized for adult 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 two-dimensional (2D) reconstruction stage, global threshold values and post-alignment corrections were verified manually. The reconstructed 2D slices were then subjected to a two-pixel kernel median filter using a custom plug-in written for ImageJ.

The filtered image stacks were used to generate 3D reconstructions in Analyze 3D 4.0. Three-dimensional landmark co-ordinates were then digitized directly from the 3D reconstruction using a dual monitor setup in which the rendered surface could be viewed simultaneously with orthogonal sections. This method eliminates the possibility of perspective errors that can occur when 3D landmarks are obtained from 2D projections of 3D surface renderings. Figure 1 shows the 3D landmarks that were digitized for this study.

Each individual was digitized twice to allow assessment of measurement error. The landmark configurations were inspected for gross outliers due to measurement errors by visual inspection of superimposed Procrustes co-ordinates. Individuals with gross errors due to reversal of side or mislabelling of landmarks were re-digitized.

The statistical analysis focused on comparisons of size, mean shape, FA and among-individual variance for shape parameters and interlandmark distances, and morphological integration. The analysis of data employed both geometric–morphometric- and Euclidean Distance Matrix Analysis (EDMA)-based methods. Choice

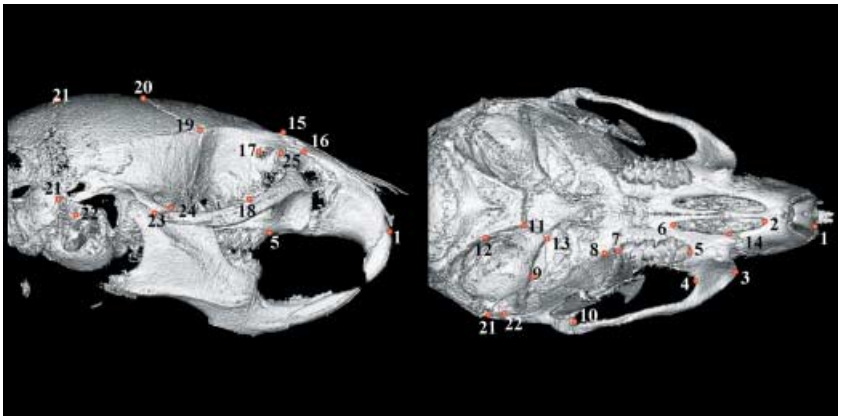

Fig. 1 Landmarks digitized shown on a 3D reconstruction from a computed microtomography scan.

of methods in this area is controversial. We take the position that both families of methods have advantages and limitations. Through Procrustes superimposed data, geometric morphometrics offer convenient ways to quantify and visualize overall shape variation across multiple groups. Procrustes superimposition can distribute large differences in the placement of particular landmarks across others in the configuration (the 'Pinocchio effect'), although this effect is probably rarely significant. EDMA provides an alternative means of localization of shape differences that is not affected by the distribution of variance across landmarks by the superimposition process. However, EDMA-based approaches also produce a very large number of variables that substantially inflates degrees of freedom. The debate over the relative merits of these methods will not be resolved in this paper. The statistical analysis employed in this study is quite complex. For this reason, Table 1 provides a list of methods and their objectives for the study.

Prior to data analysis, Procrustes superimposed landmark configurations were visually inspected for gross outliers that would be due to incorrectly identified landmarks or other kinds of gross error such as misplacement due to pseudoforamina. Individuals (comprising less than 5% of the initial sample) that showed gross errors were deleted and re-digitized.

Principal components analysis of general-least-squares Procrustes superimposed landmark configurations was used to visualize overall shape variation among and within groups. Morpheus (Slice, 1994–1999) was used to obtain the Procrustes values and Morphologika (O'Higgins & Jones, 1998) was used to visualize deformations of wireframes along principal components. Principal components analysis was performed on the Procrustes values averaged across trials for each individual.

For EDMA analysis of shape, it was necessary to obtain the average landmark configuration for the two trials and both sides within each individual. To this end, the left side was mirrored onto the right and the ESTIMATE module in EDMA was used to obtain the mean landmark configuration across sides and trials. The FORM (Lele, 1993) and SHAPE modules were then used to identify regional differences. This is done through pairwise comparisons of individual distances as well as the overall form and shape matrices for the three strains. Confidence intervals for individual distances and the statistical significance of the form matrix differences are obtained by means of a non-parametric bootstrap. We used the geometric mean of the inter-landmark distance set as the measure of size to scale the form matrices in the SHAPE analysis. For descriptions of these methods see papers by Lele, Richtsmeier and Cole (Lele, 1993; Lele & Cole, 1996; Lele & Richtsmeier, 2001; Lele & McCulloch, 2002).

To test for dominance, dominance deviations were obtained for each shape matrix element. Dominance deviations are significant departures from the expectation that the F1 cross phenotype is intermediate between the two parental strains. The deviations were tested for deviation from 0 by a *t*-test. To visualize the dominance distribution, dominance deviations were assigned direction and scaled to the halved difference between the two parental strains.

To compare overall FA across groups, we used the object asymmetry method (Klingenberg et al. 2002). We used the conventional model that assumes equal and isotropic variation across landmarks. In this method, the measure of asymmetry is the Procrustes distance between each individual landmark configuration and its reflection on itself and the degrees of freedom are adjusted appropriately for comparisons across groups

Table 1 Statistical procedures used in the analysis. These are listed in the order described and reported in the Materials and methods and Results sections

Statistical test or procedure	Objective
1. Comparisons of size and shape	
a. ANOVA for body mass	Comparison of mean body mass among strains.
b. Principal components analysis	This analysis is used to produce new variables that summarize shape variation in the Procrustes of Procrustes coordinate data coordinate data. The variables that explain the most shape variation are then plotted for the three strains. Shape variation along principal components can also be visualized by obtaining the predicted Procrustes values along each component.
c. EDMA FORM and SHAPE analysis	These tests compare the complete set of Euclidean distances (scaled and unscaled) among groups and provide boostrap based *P*-values for individual interlandmark distances.
2. Analysis of dominance	To test the hypothesis that the shape differences between the two parental strains exhibit significant dominance deviations.
3. Fluctuating asymmetry comparisons	
a. Comparison of overall FA using Klingberg's object asymmetry method.	This method tests for the significance of the overall FA difference between between groups.
b. EDMA Asymmetry Comparison	Tests for the significance of the FA difference for individual interlandmark distances.
4. Among individual variance	
a. Levene's test on Procrustes mean-shape deviations.	Tests for differences in among-individual variation in overall shape among strains.
b. Levene's test on interlandmark distances	Tests for differences in among individual variation for interlandmark distances among strains.
5. Integration of size and shape	
a. Regressions of principal components against centroid size	To compare the amount of variation explained by size for particular shape components.
b. Regression of shape mean-deviations against centroid size-deviations.	This method estimates the total percentage of shape variation that is explained by size.
c. Matrix correlations with Mantel's test for significance for both the Procrustes and the interlandmark distance matrices.	This test is used to test the null hypothesis that the pairs of covariation matrices among the three groups (and by sex) are unrelated.
d. ANOVA for z-transformed correlations	To compare overall magnitudes of size integration among groups.
e. Monte carlo test to compare the variances of eigenvalues.	To compare overall magnitudes of shape integration among groups.
f. EDMA-based test for integration.	To compare magnitudes of size integration for particular interlandmark distances.
g. Comparison of variances of eigenvalues for Procrustes centroid-size residuals.	To compare shape integration among groups after removing the allometric component of shape variation.

and for assessing the significance of FA above measurement error. We used Klingenberg et al.'s multivariate adaptation of Palmer and Strobeck's mixed-model ANOVA method (Palmer & Strobeck, 1986, 2003; Palmer, 1994) to test the significance of FA above measurement error as well as for directional asymmetry. For localization of FA variation, we performed pairwise comparisons of individual interlandmark distances using the EDMA ASYMMETRY module (Richtsmeier et al. 2002). All asymmetry distributions were tested for departure from normality using a one-sample Kolmogorov–Smirnov test.

We compared among-individual variation across groups in two ways. First, we followed Zelditch et al. (2004) in using the distribution of Procrustes distances from the sample mean as a measure of shape variation. The average Procrustes distance from the mean configuration of each sample is a measure of shape variance

that can be compared across groups using ANOVA or *t*-tests (i.e. Levene's test for differences in variance). We also used Levene's test on the raw interlandmark distances, comparing across groups (strain, sex), individual and interlandmark distance. In both cases, the mean deviations were calculated separately by sex so as not to confound among-sex with among-individual variation in size and shape.

Size is the single most important determinant of cranial shape in most mammalian species and a strong correlation between measures of size and shape is thus a central feature of integrated phenotypic variation in shape (Frost et al. 2003). In natural populations, it is not currently known to what extent age variation contributes to the correlation between size and shape. To visualize the relationship between size and shape within each group, we performed regressions of Procrustes-based shape principal component scores on cranial centroid size. To estimate the proportion of shape variation explained by size, we performed regressions of Procrustes distance from the mean configuration (at each size) against mean deviations for centroid size within each group. This is similar to the method developed by Monteiro (1999) and used by Zelditch et al. (2004) who regressed shape (Procustes distance mean deviations) on centroid size, and estimated the variation not explained by size from the Procrustes distance between each individual and the expected shape for its size.

Completely satisfactory methods for comparing levels of morphological integration across groups have not yet been developed. For this reason, we employed a battery of methods and interpreted the results in terms of the limited views provided by each one. Sample heterogeneity is an issue for any study of morphological integration. In our sample, the only within-group factor of concern is sex because all individuals are the same age (30 days) and genetically homogeneous. Following Cheverud (1982), we adjusted for sex differences by z-transforming all data within sex and group prior to further analysis of morphological integration.

Covariation patterns were compared across groups using Mantel's test for both the Procrustes matrices and the interlandmark distance matrices. Overall magnitudes of morphological integration for the entire landmark set and for specific regions were compared in two ways. First, we compared mean correlations after Fisher z-transformation for interlandmark distances for entire matrices and specific regions. Means of z-transformed correlations were compared among or between groups by analysis of variance or *t*-tests. Second, principal components analysis was performed on both the Procrustes data and the interlandmark distance matrices. We followed Wagner (1989) in using the variance of the eigenvalues as the measures of overall integration. We compared this metric among groups using two different randomization tests. To test the hypothesis that the three groups are drawn from a population with the same covariance structure, we used a Monte-Carlo randomization. In this test, the assignment of individuals from the original dataset to the three groups was randomly shuffled at each iteration for 1000 iterations. A PCA analysis was run for each iteration and the variance of the eigenvalues was saved. We then examined this distribution to determine how frequently the observed variances could have been obtained from this pooled sample by chance. The second method tests the hypothesis that the differences in integration among the three groups could have been obtained by chance. In this case, we used within-sample bootstraps in which each group was resampled with replacement 1000 times. We then compared the distributions of resampled variances of eigenvalues for each group and calculated a *P*-value as the number of times overlap in values was obtained by chance.

We also compared interlandmark correlation matrices and correlations for individual interlandmark distances using the EDMA-based bootstrap method developed by Cole & Lele (2002). For this, we used the MIBoot software developed by Cole (2002; Palmer & Strobeck, 2003). This method tests the null hypothesis that the difference matrix [$\mathbf{D} = \mathbf{R}(A) - \mathbf{R}(B)$] is equal to 0. The bootstrap generates confidence intervals for each element of the difference matrix **D**. The bootstrap analysis was performed only for the distances that fall within defined anatomical regions because the complete interlandmark correlation matrices are very large. Distances were accepted as significantly differently correlated if the 95% confidence intervals for the correlations are non-overlapping, which is a very conservative criterion.

Finally, to compare integration patterns after allometric effects of size were removed, we used linear regressions for the landmark coordinate values against centroid size within each group to generate size-independent residuals. This is equivalent to regressing any complete set of shape variables, such as partial warps plus the scores on the uniform component, on size; the results do not depend on the choice among shape variables (Zelditch et al. 2004).

Results

The two strains and the F1 cross show small but significant differences in body mass (ANOVA, d.f. = 620, F = 136, $P < 0.001$) and cranial size (ANOVA, d.f. = 128, F = 21, $P < 0.001$). Sexes differed significantly in body mass (ANOVA, d.f. = 620, F = 307, $P < 0.001$) but not in cranial centroid size (ANOVA, d.f. = 128, F = 0.03, P = NS). For both measures of size, the A/WySnJ mice are very slightly smaller on average than the other two groups.

The three groups differ significantly in shape, with the F1 cross demonstrating an intermediate morphology. Figure 2 plots the first two principal component scores for shape based on principal components analysis of the Procrustes superimposed landmarks. The first principal component explains 27% of the variation in shape and the second explains 14%. However, visual inspection of variation along the other components reveals that all of the variation that relates to differences among the strains resides on the first principal component when a principal components analysis is done for the whole sample. Figure 3 shows the shape transformations that correspond to the components plotted in Fig. 2. The overall impression here is of a longer cranium and lower cranial vault in C57BL/6J mice and of anterior and lateral displacement of some of the facial landmarks in A/WySnJ mice. There is no evidence of sexual dimorphism in any of the three strains for the first and second principal components for shape.

A complementary but somewhat clearer representation of the shape differences is provided by EDMA form and shape analysis. Both the form and the shape analysis demonstrated highly significant shape differences among all three groups (non-parametric bootstrap, 1000 iterations, $P < 0.01$). The total set of 168 interlandmark distances was compared across the three groups. Of these, 132 or 76% of the form matrix differences fell outside the 90% confidence intervals for the shape comparison between A/WySnJ and C57BL/6J. The shape comparisons between the F1 cross and A/WySnJ and C57BL/6J revealed 103 (61%) and 81 (48%) distances fell outside the 90% confidence intervals. Figure 4 shows the subset of distances that differed in relative length (scaled to geometric mean) by more than 3%. This reveals that A/WySnJ mice differ from C57BL/6J mice in that they have relatively shorter neurocrania as well as wider faces. The comparisons between the F1 cross and the two parental strains reveals an intermediate phenotype for both principal shape differences between the parental strains.

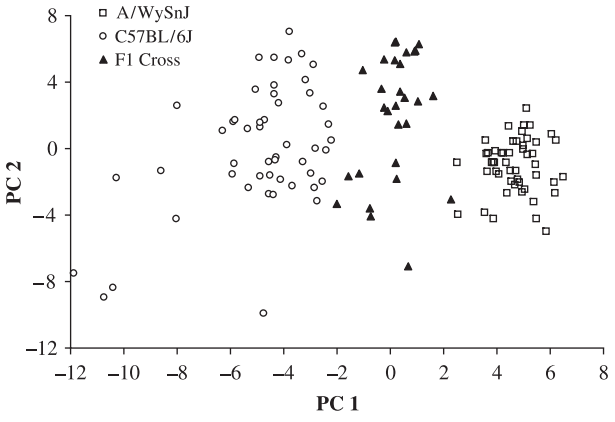

Fig. 2 Plot of principal components 1 and 2 based on Procrustes analysis of 3D landmark coordinates for all three groups.

We tested for dominance deviations in the 132 distances that differed significantly between the two parental strains. A one-sample t-test for deviation from 0 showed the presence of significant dominance deviations across traits in the F1 cross phenotype (T = 22, d.f. = 149, $P < 0.001$). Figure 5 shows these distances scaled to half the difference between the two parental strains and assigned the correct polarity. This distribution reveals that whereas there are significant dominance effects in the genetic differences between the strains, they are a mixture of direction and magnitude. The same is true when the analysis of dominance is limited to only the distances shown in Fig. 4.

The analysis of object asymmetry using Klingenberg et al. (2002) indicated that FA variation for shape is significant in all three samples. Table 2 shows the full Procrustes ANOVA table based on the isotropic variation model for object asymmetry analysis. These results also show that directional asymmetry (DA) is present in all three samples. Moreover, DA is significantly more pronounced in the A/WySnJ sample than in the other two samples (F-test, F = 1.6, $P < 0.05$ for A/WySnJ vs. C57BL/6J comparison and F = 2.4, $P < 0.01$ for A/WySnJ vs. F1 cross comparison). The three groups do not differ significantly in overall FA (F = 1.2, P = NS). Only one out of the 168 asymmetry distributions showed significant departure from normality. Antisymmetry is thus unlikely to be a significant factor in these distributions. There is, however, a significant size dependence for asymmetry as indicated by analysis of covariance with interlandmark distance as the factor (d.f. = 11610, F = 68, $P < 0.001$).

Euclidean distance matrix-based analysis of asymmetry revealed that A/WySnJ mice showed significantly

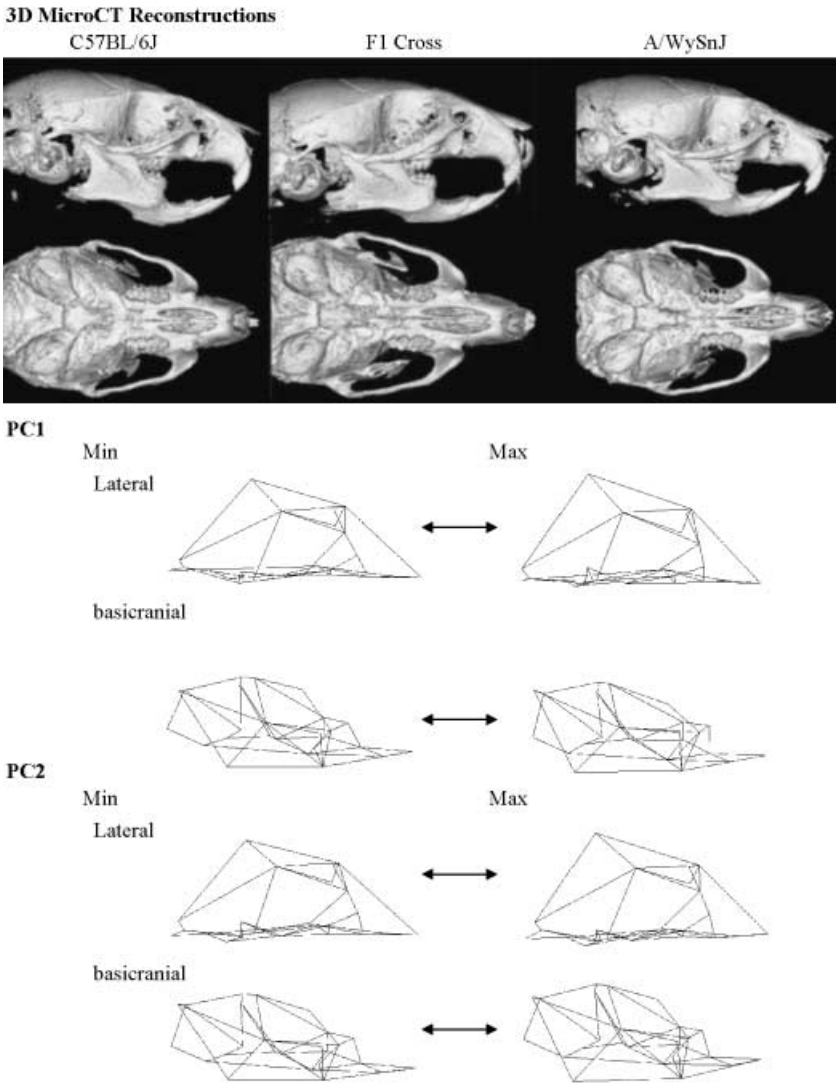

Fig. 3 Shape variation among the three groups. The micro-CT reconstructions show typical examples from each group. The wireframes show shape changes along principal components 1 and 2 based on Procrustes analysis in lateral and basicranial views.

higher FA than C57BL/6J for 19 out of the 168 traits but significantly lower FA for 40 out of the 168 traits. Comparisons with the F1 cross resulted in similar conflicting and weak results. This analysis thus reveals no evidence for higher FA in A/WySnJ compared with C57BL/6J mice.

Comparing within-group variances of shape using the Procrustes metric did show a significant difference in variance among groups (Levene's test ANOVA, d.f. = 125, $F = 11$, $P < 0.001$). This difference, however, was in the opposite direction predicted, with C57BL/6J mice having the highest shape variances and A/WySnJ the lowest. Sexes did not differ significantly in shape variation.

Comparison of interlandmark distance variances by Levene's test ANOVA exhibited the same result. There is a very small but significant strain effect in the comparison of the three strains (ANOVA, d.f. = 125, $F = 2.9$, $P < 0.05$). However, the A/WySnJ variances are the lowest among the three groups. There was no significant sex effect for the comparison of interlandmark distance variances.

The integration of size and shape or the allometric component of shape variation differs significantly among the three strains. For principal components analysis of Procrustes data performed separately within each group, plots of the first principal component scores for shape variation against centroid size reveal very different patterns in the three groups (Fig. 6). In both the C57BL/6J and the F1 cross, the pattern is that typically seen in vertebrate cranial morphology with the main axis of shape variation showing a strong allometric component. In the A/WySnJ mice, however, the first axis of shape variation is completely unrelated to size. Comparison of overall shape and size variation, by

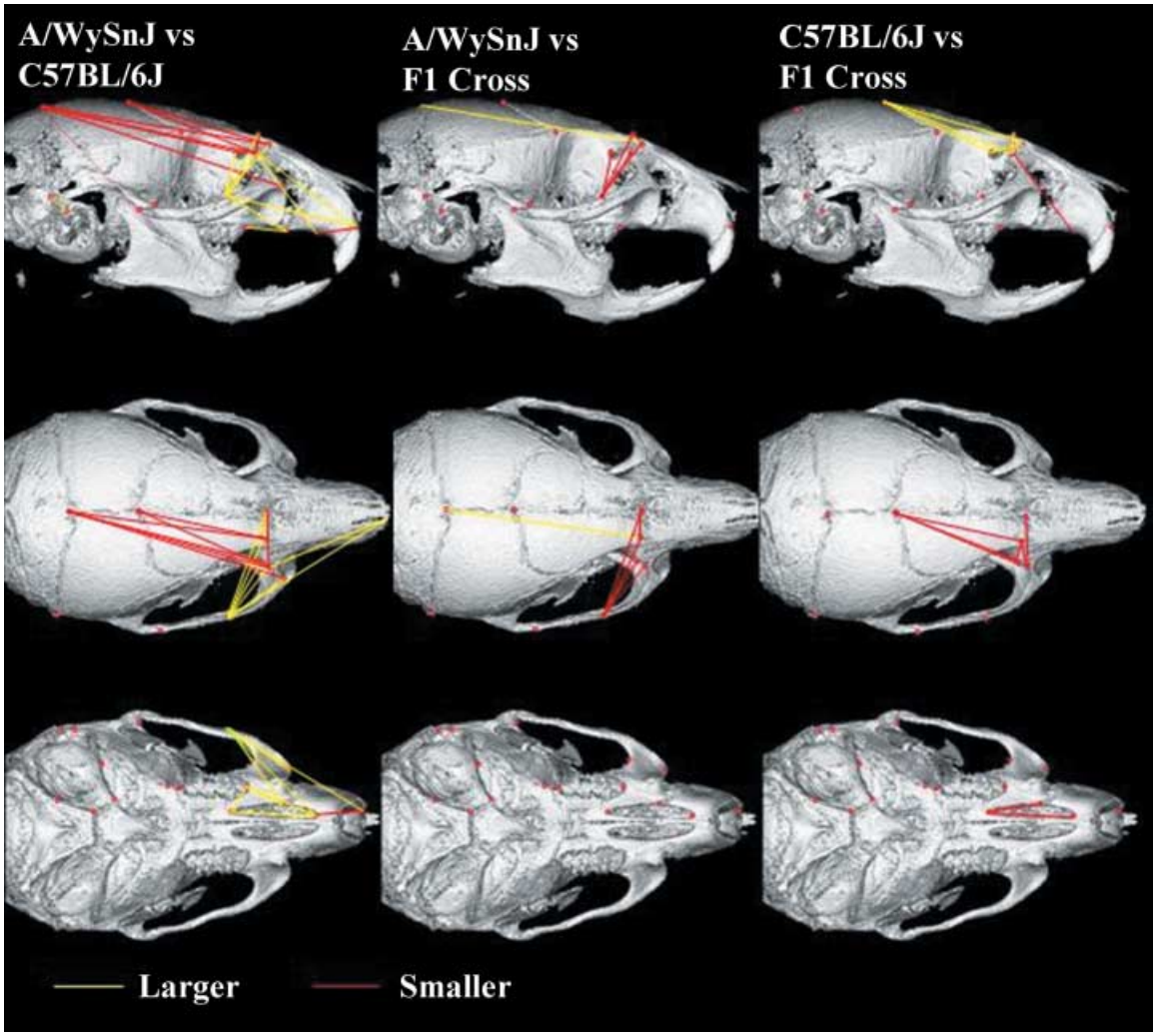

Fig. 4 EDMA Shape matrix distances that differ by 3% or more among the three strains. All distances shown are significantly different at $P < 0.001$ using a non-parametric bootstrap.

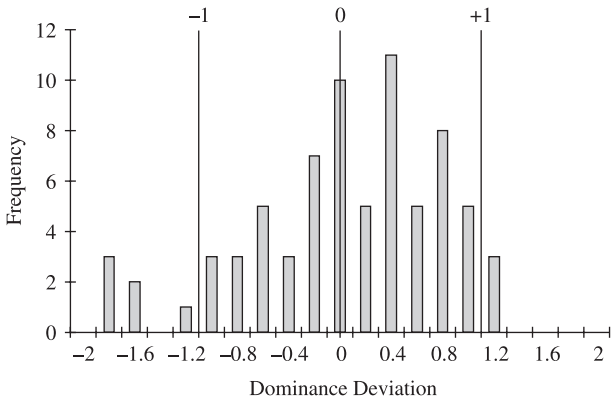

Fig. 5 Histogram of dominance deviations for 168 interlandmark distances for the F1 cross between A/WySnJ and C57BL/6J strains. The dominance deviations are standardized so that −1 is dominance for the A/WySnJ phenotype and +1 is dominance for the C57BL/6J phenotype. Values that exceed ±1 indicate overdominance.

regressing the Procrustes within-group mean deviation against the centroid size mean deviation, reveals that size and overall shape variation are significantly related in the sample as a whole (ANCOVA, $r^2 = 0.495$, $P < 0.01$, d.f. = 128). This indicates that size explains a substantial portion of the variation in shape within groups. However, the correlations between size and shape are significantly different in the three groups, with 36% of the variation in shape explained by size in C57BL/6J mice and only 8% of shape variation explained by size in A/WySnJ mice. For this comparison, the F1 cross is much closer to the A/WySnJ strain than to the other parental strain.

Comparisons of overall covariation patterns revealed broadly similar patterns in the three groups. Table 3 shows the matrix correlations for comparisons of correlation matrices for both interlandmark distances and Procrustes

Table 2 Procrustes ANOVA table for the object asymmetry analysis. This table follows the analysis laid out by Klingenberg (Hallgrimsson et al. 2002)

Group and Source	d.f.	SS	MS	*F*	*P*	FA10
A/WySnJ						
Individual	3294	0.08672	0.00002632			
Reflection	69	0.00774	0.00011214	10.03	< 0.001	
Individual × reflection	3202	0.03580	0.00001118	188.72	< 0.001	0.0015
Measurement error	6636	0.00039	0.00000006			
F1 cross						
Individual	1704	0.06761	0.00003967			
Reflection	69	0.00319	0.00004629	4.45	$P = 0.035$	
Individual × reflection	1656	0.01723	0.00001040	57.21	< 0.001	0.0015
Measurement error	3501	0.00064	0.00000018			
C57BL/6J						
Individual	3059	0.14794	0.00004836			
Reflection	69	0.00472	0.00006837	4.01	$P < 0.010$	
Individual × reflection	2973	0.05065	0.00001704	54.80	< 0.001	0.0019
Measurement error	6172	0.00192	0.00000031			

Table 3 Matrices of matrix correlations for both interlandmark distance matrices and Procrustes landmark coordinate matrices for the three strains by sex. The upper right hand portions of the two matrices contain *P*-values for Mantel's tests

		A/WySnJ F	A/WySnJ M	F1 Cross F	F1 Cross M	C57BL/6J F	C57BL/6J M
Interlandmark distance data							
A/WySnJ	F	–	$P < 0.01$	NS	$P < 0.01$	$P < 0.01$	$P < 0.01$
A/WySnJ	M	0.601	–	$P < 0.01$	$P < 0.01$	$P < 0.01$	$P < 0.01$
F1 Cross	F	0.436	0.246	–	$P < 0.01$	$P < 0.01$	$P < 0.01$
F1 Cross	M	0.682	0.459	0.694	–	$P < 0.01$	$P < 0.01$
C57BL/6J	F	0.687	0.371	0.459	0.638	–	$P < 0.01$
C57BL/6J	M	0.662	0.408	0.426	0.653	0.772	–
Procrustes aata							
A/WySnJ	F	–	$P < 0.01$	NS	$P < 0.01$	$P < 0.01$	$P < 0.01$
A/WySnJ	M	0.364	–	$P < 0.01$	$P < 0.01$	$P < 0.01$	$P < 0.01$
F1 Cross	F	0.087	0.183	–	$P < 0.01$	$P < 0.01$	$P < 0.01$
F1 Cross	M	0.237	0.157	0.496	–	$P < 0.01$	$P < 0.01$
C57BL/6J	F	0.323	0.173	0.289	0.358	–	$P < 0.01$
C57BL/6J	M	0.27	0.166	0.283	0.331	0.609	–

data for all groups divided by sex. The patterns seen in these two comparative matrices are very similar, with the matrix correlation between the interlandmark distance comparative matrix and the Procrustes comparative matrix being 0.88 (Mantel's test, $P < 0.01$). In both cases, comparisons of A/WySnJ strain to either C57BL/6J or the F1 cross show the lowest matrix correlations.

Overall magnitudes of integration show very different patterns in the three groups. Figure 7(A) shows the variances of the eigenvalues by strain. The hypothesis that the three strains are drawn from a population with the same covariance structure was tested using a randomization in which the assignment of individuals to the three groups was reshuffled for 1000 iterations. This analysis showed that the observed variances of eigenvalues for all three groups were lower than any generated by the randomization. The higher integration in the reshuffled data reflects sample heterogeneity. These results reject the null hypothesis that the three groups share a common covariance structure. To test the hypothesis that the observed differences between the three groups are significant, we used a within-sample bootstrap with replacement to generate confidence intervals for the variance of eigenvalue estimates for each group. Figure 7(B) shows the distributions of the bootstrapped values for C57BL/6J and A/

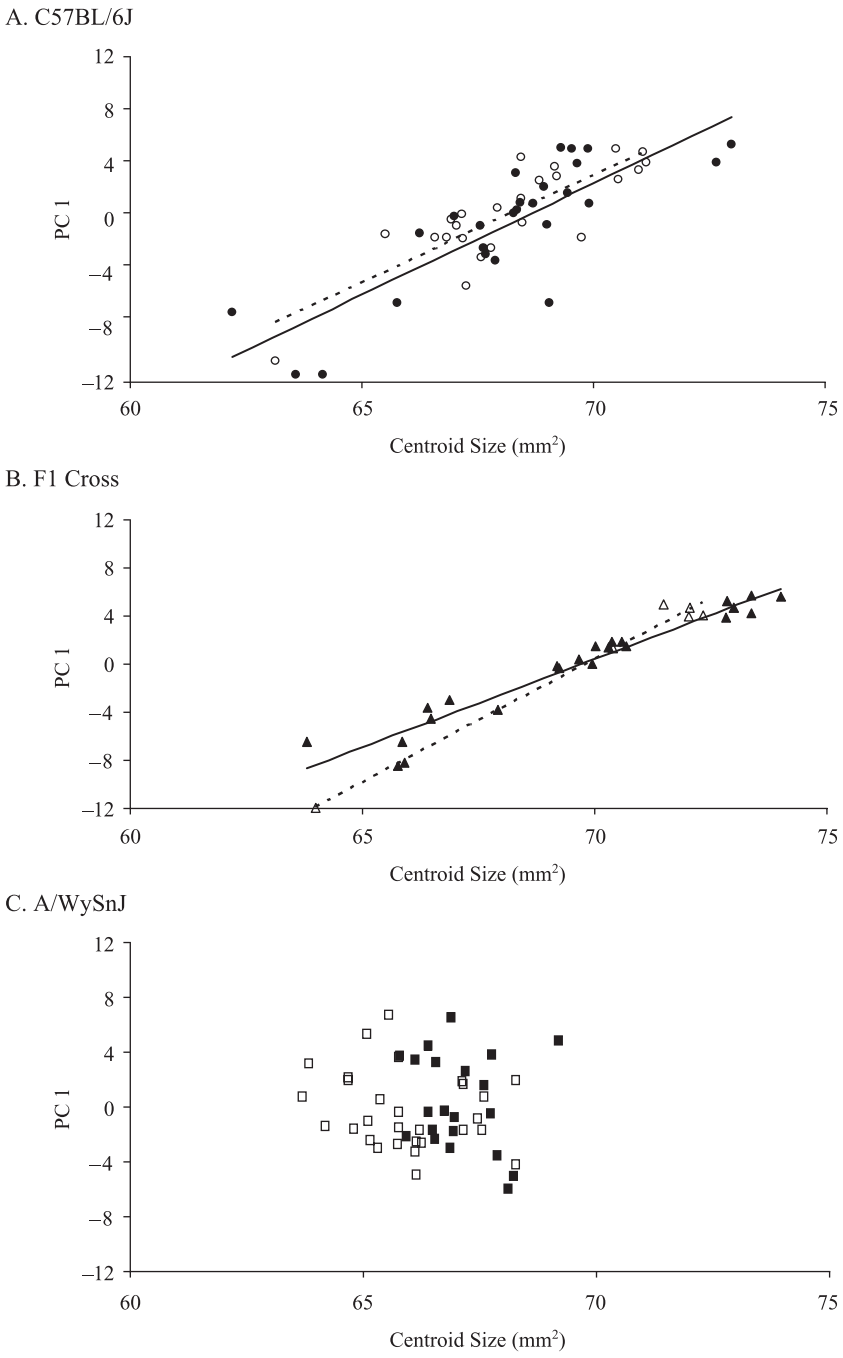

Fig. 6 Regression of Principal component 1 against centroid size for all three groups. Solid symbols are males and unfilled symbols females.

WySnJ mice. There is no overlap between these distributions, indicating that the difference between the two parental strains is significant at $P < 0.01$.

Direct comparisons of mean correlations after Fisher-z transformation reveals lower overall integration in the A/WySnJ mice, especially in the face and palate. Analysis of variance for the z-transformed correlations reveals significant variation among strains ($F = 64$, $P < 0.001$) as well as anatomical regions ($F = 86$, $P < 0.001$) and that some regions differ more among strains than others (strain × region interaction, $F = 9.1$, $P < 0.001$). Figure 8 shows these results by anatomical region. This method, which captures positive covariation in size, shows the A/WySnJ strain to be the least and the F1 cross to be the most integrated.

Cole & Lele's (2002) bootstrap method for comparing individual correlations across correlation matrices also reveals a pattern of lower integration in the A/WySnJ strain, particularly in the palate. For the palate, 26 out of 44 correlations were significantly lower when A/

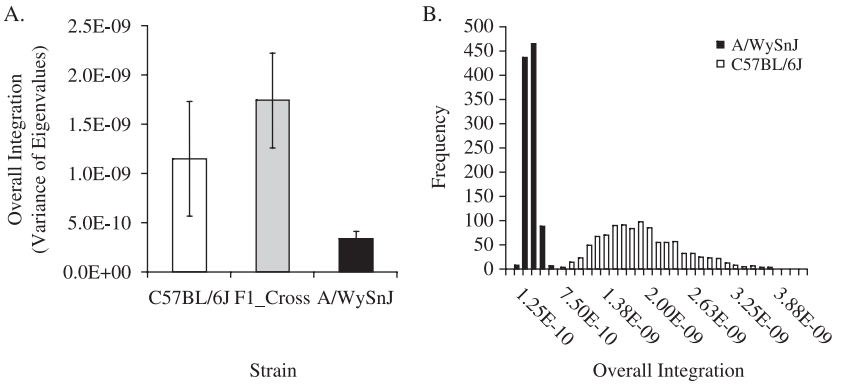

Fig. 7 Variances of eigenvalues by strain with standard deviations of the bootstrapped values (A). (B) The frequency distributions of the bootstrapped values for the C57BL/6J and A/WySnJ strains.

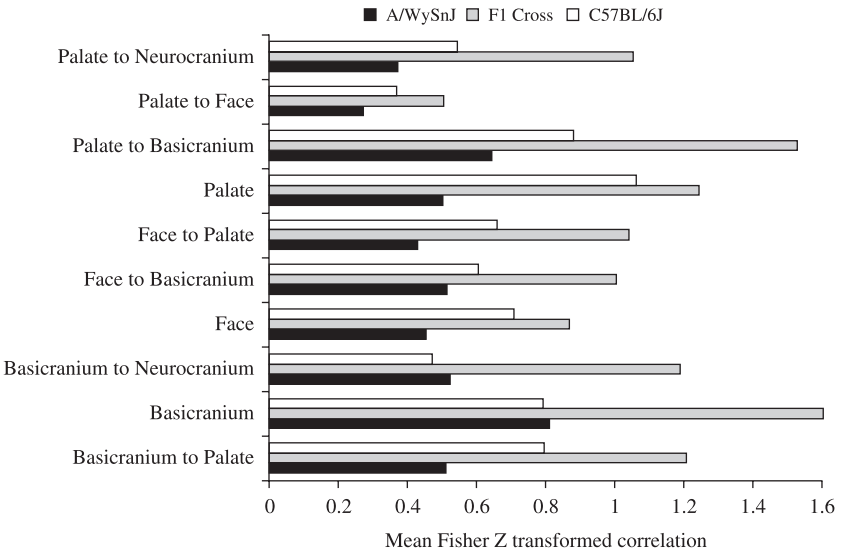

Fig. 8 Mean Fisher z-transformed correlations by region for the three strains.

WySnJ are compared with C57BL/6J mice and 35 out of 44 are lower for the comparison with the F1 cross. For the remainder of the cranium, the same pattern holds but to a lesser degree for the face; 20 out of 210 correlations were lower in the comparison with C57BL/6J mice and 98 out of 210 in the F1 cross comparison. No distances were more highly correlated in the A/WySnJ mice in either the face or palate. For the neurocranium and basicranium, 36 out of 165 distances were less integrated in A/WySnJ than in C57BL/6J mice and 125 out of 165 distances were less integrated in A/WySnJ mice compared with the F1 cross. Twenty-two out of the 165 distances were more highly correlated in A/WySnJ mice than C57BL/6J but none was higher in the comparison with the F1 cross.

The result that A/WySnJ mice are less integrated than the other two strains and also show less integration with size raises the question of whether the difference in integration is due to a difference in the allometric component of shape variation. To test this possibility, we generated residuals for the regression of the Procrustes coordinates on centroid size separately within each group to generate coordinate data free from allometric variation. As shown in Fig. 9, when the allometric component of the Procrustes data variation is removed, the F1 cross changes from the most to the least integrated. However, the difference between the two parental strains remains, with the bootstrap revealing that the C57BL/6J mice have a higher, albeit only marginally significant, level of integration (bootstrap, 1000 iterations, $P < 0.05$).

Discussion

The analysis of craniofacial shape differences between C57BL/6J and A/WySnJ mice shows that A/WySnJ mice differ in exactly those components of shape that are thought to contribute to predisposition to CL. A/WySnJ mice have wider faces and shorter neurocrania. However, in overall shape the F1 cross mice exhibit an intermediate

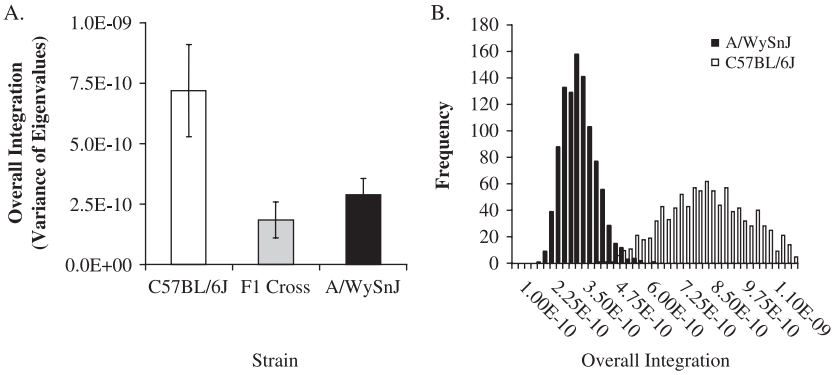

Fig. 9 Variances of eigenvalues by strain for the residuals of regression of Procrustes coordinates against centroid size (A). (B) The frequency distributions of the bootstrapped values for the C57BL/6J and A/WySnJ strains.

morphology between the two parental strains. As shown by the EDMA analysis, this includes the aspect of shape associated with the predisposition to CL. Moreover, the measures associated with these shape features do not show any specific pattern of dominance. Thus, shape features of A/WySnJ that may predispose these mice to CL are probably not due to a small number of recessive genetic factors such as the *clf1* and *clif2* loci. These results do not mean that the genetic factors that cause cleft lip in A/WySnJ mice do not also influence adult craniofacial shape. However, they do indicate that these factors are unlikely to be the major determinants of the increased facial width in these mice. Our findings are consistent with the view that the phenotypic differences between the two parental strains have a complex genetic basis that extends significantly beyond the few factors that directly influence the formation of cleft lip and palate in the A/WySnJ strain.

It is not clear from our results how the known reduction in the growth of the maxillary process during facial formation in A/WySnJ compared with C57BL/6J mice (Wang & Diewert, 1992) relates to observed shape differences in the postnatal (30-day-old) sample used for this study. Sorting out the developmental basis for the reduction in maxillary process growth and its relation to the *clf* factors is an important issue for understanding the aetiology of CL in these mice, which will be addressed in future studies. Determining how these developmental factors influence facial shape throughout ontogeny will also have important evolutionary implications. In particular, relating the shorter adult faces of A/WySnJ mice to the developmental basis for the reduction in the growth of the maxillary process may have important implications for primate evolution. Evolutionary changes in facial prognathism feature prominently in some primate lineages.

Although our results do not show increased within- or among-individual variation in A/WySnJ mice, morphological integration is significantly decreased in this strain compared with both C57BL/6J mice and the F1 cross. Most of this difference is due to the allometric component of integration, or to the tendency of shape to correlate with size. When allometry is removed, in fact, the F1 cross is less integrated than the A/WySnJ mice. These results do indicate, however, that A/WySnJ craniofacial development is characterized by a reduced co-ordination of growth of craniofacial components leading to a lower tendency for morphological covariation in the adult phenotype. The finding that the F1 cross mice show a pattern of shape and size integration that is very similar to C57BL/6J mice and is probably quite typical of mice in general indicates that this disruption in the allometric component of integration may be due to recessive genetic factors. Our results suggest that the developmental configuration seen in A/WySnJ mice that produces an elevated incidence of cleft lip is also characterized by a reduced co-ordination of growth among facial and palatal structures. At least some component of this reduction in co-ordination is due to the effects of recessive genetic factors.

There are several caveats to these findings and their implications for the aetiology of CL in both mice and humans. The first is that interpretation of the dramatically lower integration in A/WySnJ mice compared with C57BL/6J mice is somewhat ambiguous without a larger comparative sample of inbred mouse strains. However, a strong relationship between shape and size variation (allometry) is commonly found in natural populations of mammals. Our analyses of similar datasets in rhesus macaques shows that a very high proportion of shape variation is explained by size (Hallgrimsson et al. 2004) and the same is true in chimpanzees and

humans (Penin et al. 2002). Strong size and shape relationships were found in a detailed study of nine Cercopithecine species (Frost et al. 2003). All of these studies used only adult specimens, but some portion of the shape size covariation may still be due to age. In contrast, Zelditch et al. (2004) found a much lower correlation between size and shape variation in mice and cotton rats of the same age in days. More study of natural variation in both magnitudes of integration, and the integration of size and shape is needed to place these results in a comparative context.

Secondly, this study is based on an analysis of 30-day-old mice and not on embryos from the period during which the primary palate forms. It is possible therefore that the decrease in integration occurs due to factors acting subsequent to the formation of the primary palate. Studies of phenotypic variability in both morphology and measures that relate to relevant developmental mechanisms, such as the expression of genes that regulate mesenchymal growth in the maxillary process, are needed to provide a firm link between a decrease in developmental integration and the cleft lip malformation. These studies are currently underway.

The third caveat is that the reduction in integration in A/WySnJ mice could be coincidental with respect to the aetiology of CL. This study shows that an inbred mouse strain that develops a craniofacial malformation with incomplete penetrance, variable expressivity and predominant unilateral expression also shows a low level of morphological integration. Although suggestive, this establishes an association but not causation. Experimental manipulation of integration in controlled developmental contexts will be necessary to establish causation, but this requires a much more advanced understanding of the developmental genetics of phenotypic variability than we currently possess.

Fourthly, because 20% of A/WySnJ mice develop CL and die perinatally, the remaining sample may be biased. If the CL formation is correlated with shape and size variables, then this effect would likely curtail a portion of the distributions for the relevant shape and size variables. This may explain why A/WySnJ mice show reduced phenotypic variances for size and shape. The effect, however, would not explain the reduction in morphological integration except in the exceptional case in which measurement error variances are so large that a reduction in variance significantly alters the relative proportions of covariances to measurement error.

Finally, there is the obvious caveat that although the pattern of expression of CL in A/WySnJ mice is remarkably similar to that seen in humans, the aetiology of the malformation in the two species may differ significantly.

Keeping these caveats in mind, we suggest that the disruption of integrated craniofacial growth plays a role in the aetiology of CL in A/WySnJ mice and thus perhaps in humans. If true, this would help explain why a group of genetically identical individuals show such variable expression of the trait. Low integration reflects high independent variances for the component structures of the craniofacial complex. Thus even though overall variances are not high in A/WySnJ mice, the variances of structures relative to one another are. In a developmental configuration already predisposed to CL due to some other factors, low integration would push some individuals and, in many cases, particular sides within individuals over the threshold for CL formation. This might occur, for instance, if the rate of maxillary process growth is very low relative to the growth of the lateral nasal process or to overall cranial size. The pattern of incomplete penetrance and variable expressivity in a genetically and environmentally homogenous sample could thus be explained as the consequence of low integration of craniofacial development.

Very little is known presently about how genetic factors influence morphological integration in specific developmental contexts. In natural populations covariation among structures is determined by pleiotropy, linkage, as well as developmental and functional interactions during development (Cheverud, 1996; Wagner, 1996). These patterns tend to be relatively stable both among populations within species and among closely related species. In humans, for instance, patterns of integration in craniofacial structures are stable and not correlated with genetic distance among human populations (Gonzalez-Jose et al. 2004). Similar patterns of covariation in craniofacial structures are seen among closely related species of primates (Ackermann & Cheverud, 2000; Marroig & Cheverud, 2001) and even between mice and macaques (Hallgrimsson et al. 2004).

In the groups compared in this study, within-group genetic variances can be assumed to be minimal due to long histories of inbreeding in the two parental strains. The patterns of integration seen in these strains therefore should mostly reflect direct developmental and functional interactions during development. Differences in integration will mostly be due to differences in how environmental variation in one developmental

component is transmuted into variation in another. There are many possible ways in which genetic perturbations could affect covariation patterns in the absence of genetic variation. One example might be a mutation that affects the physical interaction between developing components. For example, a mutation might reduce the growth of the maxillary process such that the degree of spatial and temporal apposition of the major facial processes is reduced during the formation of the face. This is thought to occur in A/WySnJ mice (Wang et al. 1995). In this case, variation in the size of the maxillary process during the period of face formation would have less impact on the development of frontal process structures simply because the degree of physical contact is reduced. To the extent that face and palate covariation patterns are determined during the formation of the face, such a mutation would reduce covariation among facial structures.

If increased variability plays a role in some malformations, this role may often be secondary to a change in the phenotypic mean. We hypothesize that this is the case in A/WySnJ mice. The reduction in the growth of the maxillary process relative to other facial components (Wang et al. 1995) represents the shift in phenotypic mean which predisposes these mice to develop CL. The reduction in developmental integration, however, helps explain why, in this genetically and relatively environmentally homogenous group, some individuals develop the trait whereas others do not. Similar roles for phenotypic variability may exist in other complex malformations.

Conclusions

The regulation of phenotypic variability, although familiar to evolutionary biologists, is largely ignored as an aetiological factor for dysmorphology. This study shows that A/WySnJ mice that develop cleft lip with incomplete penetrance and variable expressivity also show reduced integration of craniofacial structures as compared with another inbred strain (C57BL/6J). Furthermore, this strain shows an unusual degree of disassociation of shape and size variation. More research is necessary to establish a firm causal link between variation in integration and the aetiology of cleft lip. However, we suggest that decreased integration of craniofacial development may be an aetiological factor for CL in these mice and that reduced co-ordination of development contributes to the observed pattern of incomplete penetrance and frequent unilateral expression. If true, the regulation of phenotypic variability would very likely be a relevant aetiological factor for cleft lip formation in humans.

Acknowledgements

We are grateful to Wei Liu for technical assistance and to Wendy Verwey for editing. We also thank Mae Chung, Michelle Dymond and Jon Chung for discussions and assistance with various technical issues. We thank Brian Hall and John Matyas for discussions. This work was supported by NSERC grant 238992-02, CFI grant #3923 and Alberta Innovation and Science grant #URSI-01-103-RI to B.H.

References

Ackermann RR, Cheverud JM (2000) Phenotypic covariance structure in tamarins (genus *Saguinus*): a comparison of variation patterns using matrix correlation and common principal component analysis. *Am. J. Phys. Anthropol.* **111**, 489–501.

AlEmran SE, Fatani E, Hassanain JE (1999) Craniofacial variability in parents of children with cleft lip and cleft palate. *J. Clin. Pediatr. Dent.* **23**, 337–341.

Cheverud JM (1982) Phenotypic, genetic, and environmental integration in the cranium. *Evolution* **36**, 499–516.

Cheverud JM (1995) Morphological integration in the saddleback tamarin (*Saguinus fuscicollis*) cranium. *Am. Nat.* **145**, 63–89.

Cheverud JM (1996) Developmental integration and the evolution of pleiotropy. *Am. Zool.* **36**, 44–50.

Chung CS, Kau MC (1985) Racial differences in cephalometric measurements and incidence of cleft lip with or without cleft palate. *J. Craniofac. Genet. Dev. Biol.* **5**, 341–349.

Ciriani D, Diewert VM (1986) A comparative study of development during primary palate formation in A/WySn, C57BL/6, and their F1 crosses. *J. Craniofac. Genet. Dev. Biol.* **6**, 369–377.

Cole TM, 3rd (2002) *Miboot: Software for Bootstrap Comparison of Morphological Integration Patterns*. Kansas City: University of Missouri – Kansas City School of Medicine.

Cole TM, 3rd, Lele S (2002) Bootstrap-based methods for comparing morphological integration patterns. *Am. J. Phys. Anthropol. Suppl.* **35**: 55.

Diewert VM, Lozanoff S (2002) Animal models of facial clefting: experimental, congenital, and transgenic. In *Understanding Craniofacial Anomalies: Etiopathogenesis of Craniosynostoses and Facial Clefting* (eds Mooney MM, Siegel MI), pp. 251–272. New York: Wiley-Liss.

Diewert VM, Wang KY (1992) Recent advances in primary palate and midface morphogenesis research. *Crit. Rev. Oral Biol. Med.* **4**, 111–130.

Fraser FC, Pashayan H (1970) Relation of face shape to susceptibility to congenital cleft lip. A preliminary report. *J. Med. Genet.* **7**, 112–117.

Frost SR, Marcus LF, Bookstein FL, Reddy DP, Delson E (2003) Cranial allometry, phylogeography, and systematics of large-bodied papionins (primates: Cercopithecinae) inferred from geometric morphometric analysis of landmark data. *Anat. Rec.* **275A**, 1048–1072.

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

Gong SG (2001) Phenotypic and molecular analyses of A/WySn mice. *Cleft Palate. Craniofac. J.* **38**, 486–491.

Gonzalez-Jose R, Van Der Molen S, Gonzalez-Perez E, Hernandez M (2004) Patterns of phenotypic covariation and correlation in modern humans as viewed from morphological integration. *Am. J. Phys. Anthropol.* **123**, 69–77.

Hallgrímsson B, Willmore K, Hall BK (2002) Canalization, developmental stability, and morphological integration in primate limbs. *Am. J. Phys Anthropol (Yearbook)* **119** (S35), 131–158.

Hallgrimsson B, Willmore K, Dorval C, Cooper DML (2004) Craniofacial variability and modularity in macaques and mice. *J. Exp. Zool.* **302B**, 225.

Hermann NV, Jensen BL, Dahl E, Bolund S, Kreiborg S (1999) A comparison of the craniofacial morphology in 2-month-old unoperated infants with unilateral complete cleft lip and palate, and unilateral incomplete cleft lip. *J. Craniofac. Genet. Dev. Biol.* **19**, 80–93.

Jara L, Blanco R, Chiffelle I, Palomino H, Carreno H (1995) Association between alleles of the transforming growth factor alpha locus and cleft lip and palate in the Chilean population. *Am. J. Med. Genet.* **57**, 548–551.

Juriloff DM, Trasler DG (1976) Test of the hypothesis that embryonic face shape is a causal factor in genetic predisposition to cleft lip in mice. *Teratology* **14**, 35–42.

Juriloff DM, Harris MJ, Brown CJ (2001) Unravelling the complex genetics of cleft lip in the mouse model. *Mamm. Genome* **12**, 426–435.

Juriloff DM, Harris MJ, Dewell SL (2004) A digenic cause of cleft lip in A-strain mice and definition of candidate genes for the two loci. *Birth Defects Res. Part A Clin. Mol Teratol.* **70**, 509–518.

Kaufman MH, Bard JBL (1999) *The Anatomical Basis of Mouse Development*. San Diego: Academic Press.

Khan SN, Hufnagle KG, Pool R (1986) Intrafamilial variability of popliteal pterygium syndrome: a family description. *Cleft Palate J.* **23**, 233–236.

Klingenberg CP, Barluenga M, Meyer A (2002) Shape analysis of symmetric structures: quantifying variation among individuals and asymmetry. *Evolution* **56**, 1909–1920.

Kobyliansky E, Bejerano M, Yakovenko K, Katznelson MB (1999) Relationship between genetic anomalies of different levels and deviations in dermatoglyphic traits. Part 6: Dermatoglyphic peculiarities of males and females with cleft lip (with or without cleft palate) and cleft palate – family study. *Coll. Antropol.* **23**, 1–51.

Kondo S, Schutte BC, Richardson RJ, et al. (2002) Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes. *Nat. Genet.* **32**, 285–289.

Lacombe D, Pedespan JM, Fontan D, Chateil JF, Verloes A (1995) Phenotypic variability in van der Woude syndrome. *Genet. Couns.* **6**, 221–226.

Leamy LJ, Routman EJ, Cheverud JM (2002) An epistatic genetic basis for fluctuating asymmetry of mandible size in mice. *Evolution* **56**, 642–653.

Lele S (1993) Euclidean distance matrix analysis of landmark data: estimation of mean form and mean form difference. *Mathemat. Geol.* **25**, 573–602.

Lele S, Cole TM III (1996) A new test for shape differences when variance-covariance matrices are unequal. *J. Hum. Evol.* **31**, 193–212.

Lele S, Richtsmeier JT (2001) *An Invariant Approach to the Statistical Analysis of Shapes*. Boca Raton: Chapman & Hall.

Lele S, McCulloch CE (2002) Invariance, identifiability and morphometrics. *J. Am. Statist. Assoc.* **00,** 796–806.

Lidral AC, Romitti PA, Basart AM, et al. (1998) Association of MSX1 and TGFB3 with nonsyndromic clefting in humans. *Am. J. Hum. Genet.* **63**, 557–568.

Lin YC, Lo LJ, Noordhoff MS, Chen YR (1999) Cleft of the lip and palate in twins. *Changgeng Yi Xue Za Zhi* **22**, 61–67.

Marroig G, Cheverud JM (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.

Millicovsky G, Ambrose LJ, Johnston MC (1982) Developmental alterations associated with spontaneous cleft lip and palate in CL/Fr mice. *Am. J. Anat.* **164**, 29–44.

Monteiro LR (1999) Multivariate regression models and geometric morphometrics: the search for causal factors in the analysis of shape. *Syst. Biol.* **48**, 192–199.

Neiswanger K, Cooper ME, Weinberg SM, et al. (2002) Cleft lip with or without cleft palate and dermatoglyphic asymmetry: evaluation of a Chinese population. *Orthod. Craniofac. Res.* **5**, 140–146.

O'Higgins P, Jones N (1998) *Morphologika*. London: University College London.

Olson EC, Miller RA (1958) *Morphological Integration*. Chicago: University of Chicago Press.

Ozbudak EM, Thattai M, Kurtser I, Grossman AD, Van Oudenaarden A (2002) Regulation of noise in the expression of a single gene. *Nat. Genet.* **31**, 69–73.

Palmer AR, Strobeck C (1986) Fluctuating asymmetry. measurement, analysis, patterns. *Annu. Rev. Ecol. Syst* **17**, 391–421.

Palmer AR (1994) Fluctuating asymmetry analyses: a primer. In *Developmental Instability: its Origins and Evolutionary Implications* (ed. Markow TA), pp. 355–364. Dordrecht: Kluwer Academic Publishers

Palmer R, Strobeck C (2003) Fluctuating asymmetry analysis unplugged. In *Developmental Instability (DI): Causes and Consequences* (ed. Polak M), pp. 279–319. Oxford: Oxford University Press.

Penin X, Berge C, Baylac M (2002) Ontogenetic study of the skull in modern humans and the common chimpanzees: neotenic hypothesis reconsidered with a tridimensional Procrustes analysis. *Am. J. Phys. Anthropol.* **118**, 50–62.

Pezzetti F, Scapoli L, Martinelli M, et al. (1998) A locus in 2p13-p14 (OFC2), in addition to that mapped in 6p23, is involved in nonsyndromic familial orofacial cleft malformation. *Genomics* **50**, 299–305.

Prescott NJ, Winter RM, Malcolm S (2001) Nonsyndromic cleft lip and palate: complex genetics and environmental effects. *Ann. Hum. Genet.* **65**, 505–515.

Richtsmeier JT, Cole TM, 3rd, Lindsay E, et al. (2002) Studying asymmetry with Euclidean distance matrix analysis. *Am. J. Phys. Anthropol. Suppl.* **34**, 131.

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

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

Scharloo W (1991) Canalization: genetic and developmental aspects. *Annu. Rev. Ecol. Syst.* **22**, 65–93.

Slice DE (1994–99) *Morpheus.* Stony Brook, State University of New York (SUNY).

Sözen MA, Suzuki K, Tolarova MM, Bustos T, Fernandez Iglesias JE, Spritz RA (2001) Mutation of PVRL1 is associated with sporadic, non-syndromic cleft lip/palate in northern Venezuela. *Nat. Genet.* **29**, 141–142.

Trasler DG (1968) Pathogenesis of cleft lip and its relation to embryonic face shape in A-J and C57BL mice. *Teratology* **1**, 33–49.

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

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

Wagner GP (1989) A comparative study of morphological integration in *Apis mellifera* (Insecta, Hymenoptera). *Z. Zool. Syst. Evol. – Forsch.* **28**, 48–61.

Wagner GP (1996) Homologues, natural kinds and the evolution of modularity. *Am. Zool.* **36**, 36–43.

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

Wang K-Y, Diewert VM (1992) A morphometric analysis of craniofacial growth in cleft lip and noncleft mice. *J. Craniofac. Genet. Dev. Biol.* **12**, 141–154.

Wang KY, Juriloff DM, Diewert VM (1995) Deficient and delayed primary palatal fusion and mesenchymal bridge formation in cleft lip-liable strains of mice. *J. Craniofac. Genet. Dev. Biol.* **15**, 99–116.

Wyszynski DF, Beaty TH (1996) Review of the role of potential teratogens in the origin of human nonsyndromic oral clefts. *Teratology* **53**, 309–317.

Wyszynski DF, Beaty TH, Maestri NE (1996) Genetics of nonsyndromic oral clefts revisited. *Cleft Palate. Craniofac. J.* **33**, 406–417.

Wyszynski DF, Beaty TH (2002) Phenotypic discordance in a family with monozygotic twins and nonsyndromic cleft lip and palate: follow-up. *Am. J. Med. Genet.* **110**, 182–183.

Yoon YJ, Perkiomaki MR, Tallents RH, et al. (2003) Association of nasomaxillary asymmetry in children with unilateral cleft lip and palate and their parents. *Cleft Palate. Craniofac. J.* **40**, 493–497.

Yoon YJ, Perkiomaki MR, Tallents RH, et al. (2004) Transverse craniofacial features and their genetic predisposition in families with nonsyndromic unilateral cleft lip and palate. *Cleft Palate. Craniofac. J.* **41**, 256–261.

Zelditch ML, Lundrigan BL, Garland T (2004) Developmental regulation of skull morphology. I. Ontogenetic dynamics of variance. *Evol. Dev.* **6**, 194–206.