Current Zoology, 2020, 66(5), 527–538
doi: 10.1093/cz/zoaa035
Advance Access Publication Date: 11 September 2020
Article

Article

Does divergence from normal patterns of integration increase as chromosomal fusions increase in number? A test on a house mouse hybrid zone

Carmelo Fruciano[a,b,*], Paolo Colangelo[c], Riccardo Castiglia[d], and Paolo Franchini[e,*]

[a]Institut de Biologie de l'Ecole Normale Supérieure (IBENS), Ecole Normale Supérieure, CNRS, PSL Université Paris, Paris, 75005, France, [b]School of Biological Sciences, University of Portsmouth, Portsmouth, PO1 2DY, UK, [c]National Research Council, Research Institute on Terrestrial Ecosystems, Montelibretti (RM), 00010, Italy, [d]Department of Biology and Biotechnology "Charles Darwin", "La Sapienza" University of Rome, Rome, 00161, Italy, and [e]Department of Biology, Lehrstuhl für Zoologie und Evolutionsbiologie, University of Konstanz, Konstanz, 78457, Germany

*Address correspondence to Carmelo Fruciano and Paolo Franchini. E-mail: c.fruciano@unict.it and paolo.franchini@uni-konstanz.de

Handling editor: Martha Muñoz

Received on 6 April 2020; accepted on 3 July 2020

Abstract

Chromosomal evolution is widely considered an important driver of speciation because it can promote the establishment of reproductive barriers. Karyotypic reorganization is also expected to affect the mean phenotype, as well as its development and patterns of phenotypic integration, through processes such as variation in genetic linkage between quantitative trait loci or between regulatory regions and their targets. Here we explore the relationship between chromosomal evolution and phenotypic integration by analyzing a well-known house mouse parapatric contact zone between a highly derived Robertsonian (Rb) race ($2n = 22$) and populations with standard karyotype ($2n = 40$). Populations with hybrid karyotypes are scattered throughout the hybrid zone connecting the two parental races. Using mandible shape data and geometric morphometrics, we test the hypothesis that patterns of integration progressively diverge from the "normal" integration pattern observed in the standard race as they accumulate Rb fusions. We find that the main pattern of integration observed between the posterior and anterior part of the mandible can be largely attributed to allometry. We find no support for a gradual increase in divergence from normal patterns of integration as fusions accumulate. Surprisingly, however, we find that the derived Rb race ($2n = 22$) has a distinct allometric trajectory compared with the standard race. Our results suggest that either individual fusions disproportionately affect patterns of integration or that there are mechanisms which "purge" extreme variants in hybrids (e.g. reduced fitness of hybrid shape).

Key words: chromosomal races, geometric morphometrics, hybrid zone, integration, modularity, Robertsonian fusions

Understanding how the evolutionary process produces phenotypic disparity and species diversity—as well as their non-uniform distribution across space, time and lineages—is a major goal of evolutionary research. Chromosomal rearrangements, large-scale genomic modifications that involve portions or entire chromosomes, are considered key players in promoting the rapid establishment of reproductive barriers between diverging taxa (Coyne and Orr 2004; Charron et al. 2014; Merot et al. 2020). It is often the case that individuals bearing chromosomal rearrangements cannot interbreed with individuals possessing the ancestral karyotype or that, if they can, the hybrids thus obtained suffer from reduced fertility (Noor et al. 2001; Lowry and Willis 2010). Chromosomal rearrangements can also affect the phenotype by acting directly on the presence, number or functionality of quantitative trait loci (QTL; the regions of the genome containing the genes responsible for the variation in a trait) or by affecting the genetic independence of multiple QTL (Imprialou et al. 2017). For instance, the deletion of a QTL region can affect directly the phenotype by removing the genes controlling trait variation . Conversely, a fusion of two chromosomes each bearing a QTL for a distinct trait creates a physical linkage between the two QTL, which can in turn affects how independent the two traits are. In this last example, a chromosomal rearrangement produces a new constraint to phenotypic variation (if the two traits are not genetically independent anymore) or alters an existing constraint (if the two traits were already genetically non-independent for other reasons). By simultaneously facilitating the establishment of reproductive barriers and affecting phenotypic variation, chromosomal evolution can accelerate the accumulation of both new species and phenotypic diversity.

Here, we will focus on the effect of a common type of chromosomal rearrangement—Robertsonian (Rb) fusions—on the phenotype and on how different parts of an organism are associated to each other. We approach this topic using one of the most prominent biological systems for the study of chromosomal evolution—the house mouse. The Western European house mouse, *Mus musculus domesticus*, is characterized by 40 acrocentric chromosomes ($2n = 40$), the standard karyotype of the entire genus *Mus*. However, Rb fusions, a class of large-scale chromosomal rearrangements that consist in the joining of 2 non-homologous acrocentric chromosomes at their centromere, are highly common in this subspecies across its whole distribution area (Gropp and Winking 1981). The fixation of newly formed metacentric chromosomes has led to a mosaic of Rb populations (harboring a combination of Rb and acrocentric chromosomes) interspersed among mice with the ancestral "Standard" karyotype. In some cases, Rb fusions become fixed in homozygous state, entitling the population with such karyotype to be called a "chromosomal race". Numerous chromosomal races distinguished by different combinations of metacentric chromosomes in homozygous condition have been described in the house mouse (Pialek et al. 2005). This high karyotypic diversity has made natural populations of *M. m. domesticus* an ideal experimental setting to investigate the effect of large-scale translocations in establishing reproductive isolation and in producing phenotypic variation (Sage et al. 1993). The study of these evolutionary processes is further facilitated by the occurrence of several hybrid zones between house mouse populations differing in their karyotypes. In these hybrid zones, different stages of the "speciation continuum" can be observed, ranging from uninterrupted gene flow with ongoing hybridization to almost complete reproductive isolation. The hybrid zones identified so far involve secondary contacts between chromosomal races characterized by different combinations of metacentric chromosomes (Franchini et al. 2008; Hauffe et al. 2011), Rb chromosomal races and populations with standard karyotype (Capanna et al. 1975; Franchini et al. 2010) or primary intergradations areas with populations carrying heterozygous Rb chromosomes in which the fixation process has not been completed yet (Sans-Fuentes et al. 2009; Castiglia et al. 2011; Muñoz-Muñoz et al. 2011). In these hybrid zones, the extent of gene flow between the different populations is inversely proportional to their karyotypic incompatibility, that in turn leads to reduced fertility of hybrids (Hauffe et al. 2011).

Rb-fused chromosomes have been shown to locally reduce recombination rates, and thus gene flow between diverging populations (Rieseberg 2001; Panithanarak et al. 2004; Franchini et al. 2010). This acts as a mechanism which complements the hypofertility of hybrids due to the occurrence of aneuploid gametes in producing reproductive barriers and promoting phenotypic variation (Faria and Navarro 2010). The Rb-induced reduced recombination, mainly affecting pericentromeric regions of Rb chromosomes, can also promote adaptive evolution by building up linkage disequilibrium between advantageous allelic combinations or simply by altering the regulatory landscape of genes, and ultimately their expression (Lindholm et al. 2016). As a result, not only Rb fusions can induce changes in traits that are inherited in a population, but they can also alter their level of modularity and integration (Klingenberg 2010). In other words, changes in the inter-dependence of chromosomal regions may affect how and how strongly different traits co-vary (integration) and their grouping into suites of traits which are relatively independent from each other (modularity). For these reasons, the house mouse mandible has been intensively used as a focal trait in morphometric studies aimed at investigating the effect of Rb fusions in producing intra- and inter-population phenotypic differences (Corti and Rohlf 2001; Li 2011; Muñoz-Muñoz et al. 2011; Martínez-Vargas et al. 2014; Franchini et al. 2016; Martínez-Vargas et al. 2018).

Furthermore, the mouse mandible has long served as a primary model for developmental, genetic and morphological studies of complex traits (Klingenberg et al. 2003). Within this complex bony structure, several anatomically and functionally differentiated units, with distinct embryological origin, have been identified (Atchley and Hall 1991). In the last two decades, the genetic basis of mandible shape variation has been widely investigated revealing its polygenic architecture, and several QTL have been found to control different units (Ehrich et al. 2003; Klingenberg et al. 2004). In particular, two main regions of the mandible showed a certain degree of genetic independence, further corroborating the mouse mandible as a model in studies of morphological integration and modularity. These two relatively independent regions include an anterior, tooth-bearing zone called "alveolar region," and a posterior portion called "ascending ramus," where most of the muscles are inserted (Leamy 1993; Klingenberg et al. 2004) (Figure 1).

A well-known house mouse hybrid zone can be found in Central Italy between the Cittaducale (CD) chromosomal race ($2n = 22$) and the surrounding populations with all-acrocentric standard karyotype (Capanna et al. 1975). Rb karyotypes are represented in the literature with a notation indicating which acrocentric chromosomes found in the standard karyotype are fused in the Rb karyotype. For instance, "Rb(1.7)" indicates a Rb fusion between chromosomes 1 and 7, which are distinct in the standard karyotype. The CD race is characterized by 9 Rb chromosomes [Rb(1.7), Rb(2.18), Rb(3.8), Rb(4.15), Rb(5.17), Rb(6.13), Rb(9.16), Rb(10.11), Rb(12.14)]. Populations with hybrid karyotypes are scattered throughout an

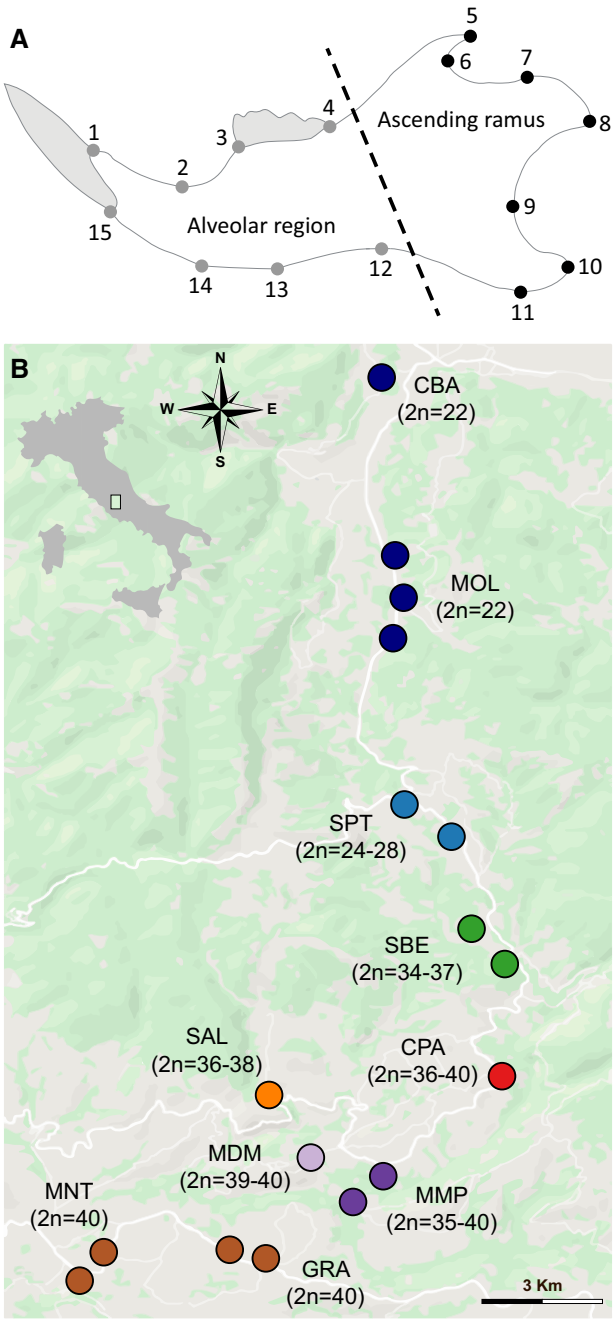

Figure 1. (**A**) Lingual side of the house mouse left mandible. The location of the 15 selected landmarks are shown. The dashed line separates the two main modules of the mandible, the alveolar region and the ascending ramus. **(B)** Map of the CD-Standard hybrid zone in Central Italy. The 17 sampling localities are indicated by full circles (different colors reflect different diploid number). Localities with same acronyms follow population nomenclature as in Franchini *et al.* (2010).

~10 km-wide transect connecting the two parental races following a clinal distribution displaying an increase of the number of metacentric chromosomes along a Standard-CD (South-North) direction (Figure 1). Previous research has shown a decreased fertility of hybrids, highlighting how the number of heterozygous metacentrics is directly related to the number of aneuploid gametes (Castiglia and Capanna 2000). Furthermore, genetic analysis carried out with microsatellite loci located in centromeric and telomeric regions of 5 Rb metacentrics and on chromosome 19 (the only acrocentric autosomal shared by the 2 parental races) showed that the hybrid zone acts as a semi-permeable barrier to gene flow, with stronger resistance to gene flow in centromeric regions of Rb chromosomes, characterized by reduced recombination (Franchini et al. 2010). Whereas previous genetic surveys clearly showed the occurrence of gene exchange between the 2 parental races, no conclusion could be drawn on the direction of gene flow. Taken together, these studies allowed us to demonstrate how the interplay between hybrid fertility and localized suppression of recombination could jointly drive genetic divergence between the focal races and, ultimately, speciation.

In this study, we go one step further and investigate the impact of chromosomal rearrangements in altering the morphology of mice mandibles from the CD-Standard hybridization area, with a special focus on the patterns of covariation between its main modules, the alveolar region and the ascending ramus. The chromosomal structure of this hybridization zone, characterized by the presence of two "pure" races ($2n = 22$ and $2n = 40$) and a wide range of hybrids carrying intermediate karyotypes (Figure 1), allowed us to test several hypotheses. In detail, we wanted to test for the effect of Rb translocations on mandible morphology, especially whether the increase in the number of metacentric chromosomes can be related to a progressive disruption of normal patterns of integration. Additionally, in consideration of the potential developmental and/or genetic disruption induced by Rb fusions in the ancestral all-acrocentric chromosome populations, we tested for an alteration of the amount of mandible morphological variation (i.e. disparity) in the highly derived CD race (as an effect of possible reduced fitness of individuals carrying maladapted mandible shapes).

Materials and Methods

Data collection and preparation

We examined a total of 116 mice that were live-trapped in 17 localities between 1996 and 1998, and between 2002 and 2005 in Central Italy (Figure 1A). At the time of sampling, mice were weighed and the karyotype of each individual was characterized using bone marrow cytogenetic analysis and G-banding techniques (Castiglia and Capanna 1999; Franchini et al. 2010). The sampling area includes a well-known house mouse parapatric contact zone between a highly derived Rb race, Cittaducale (CD, $2n = 22$) and populations with standard all-acrocentric chromosomes karyotype ($2n = 40$). Populations with hybrid karyotypes ($2n = 24$–39) are scattered throughout the hybrid zone, a narrow mountain valley connecting the 2 parental races ~10-km wide (Figure 1A and Supplementary Table S1). As the main aim of this study is to test the hypothesis that an increase of Rb fusions is associated to higher departure from normal patterns of integration, for several analyses we used as a variable the diploid number of chromosomes, which is the predictor that best reflects this hypothesis. However, the same number of chromosomes can be obtained with different fusions. To quantify the extent to which diploid number reflects karyotypic variation, we performed a preliminary analysis and found out that diploid number accounts for most of the variation in karyotype (Supplementary Note S1).

For each specimen, the left and right side of the mandible were separated at the mandibular symphysis and cleaned by Dermestid beetles. To assess size and shape variation, we used the left mandible, as it was the side with the highest number of intact samples. Photos of the lingual side of each mandible were collected and 15 landmarks were digitized (Figure 1B) as described in Franchini

et al. (2016). The configurations of points—all fixed landmarks—were subjected to a generalized Procrustes analysis (GPA) (Dryden and Mardia 1998) using the R package *Morpho* (Schlager 2017) and centroid size was computed.

General linear models

To test for the effect of a few predictors on shape variation, we fitted and tested the significance of a series of general linear models. In all these models, shape data were the dependent variable. The predictors of our most important model were mandible centroid size, diploid number (the number of chromosomes in the diploid karyotype), and their interaction. Testing this model equates to testing whether allometry (the dependency of a trait on size) has a significant impact on shape variation, as well as whether allometric trajectories are different between karyotypes. We also fitted the same model using only the specimens with $2n = 22$ and $2n = 40$. Furthermore, to exclude that sexual dimorphism was a confounding factor, we also separately tested a model with sex as a predictor, as well as a model with sex, centroid size and their interactions as predictors.

It is well known that geometric morphometric data cannot be directly used for testing general linear models because the covariance matrix of shape variables is singular (which, in turn, is due to the loss of dimensions when removing the effect of rotation, translation and scale in GPA). To circumvent this problem, we used both a traditional, and conservative, approach and a recently developed approach. In particular, the traditional approach consists in performing a principal component analysis (PCA) and retaining only axes with nonzero eigenvalues, using the scores on these axes as dependent variables in the linear model. The second approach (Clavel et al. 2019), implemented in the R package *mvMORPH* (Clavel et al. 2015), uses a penalized likelihood framework to circumvent the problem of singularity of covariance matrices. This approach is more powerful than both multivariate analysis of variance (MANOVA) on PCA scores and distance-based approaches (Clavel and Morlon in press). Tests of significance of models fitted in *mvMORPH* were performed using 1,000 permutations. For both the traditional and the penalized likelihood approach, we used Pillai's trace (Pillai 1955) as the test statistic.

To visualize predictions from general linear models, we used plots of the scores along the first two PC axes which for the penalized likelihood approach were obtained by subjecting the predictions from the linear model to a PCA. We also plotted shape change between predictions at extreme observed sizes for the $2n = 22$ and $2n = 40$ races.

Analyses of disparity

We also performed a series of analyses of disparity (also called "morphospace occupation" in the literature) on the specimens with $2n = 22$ and $2n = 40$, to test the null hypothesis that these two races have the same levels of mandibular morphological variation. We also included in these analyses another group with all the intermediate karyotypes. Analyses in a group with a mixture of karyotypes are generally hard to interpret and should be considered at best exploratory. However, if such group had particularly high disparity compared with the pure races, then that may indicate that stabilizing selection has not yet acted to reduce total diversity. Conversely, if such group had particularly low disparity, then that may point out to developmental perturbations producing a lower number of viable phenotypes in this group. All the analyses of disparity were performed using the R package *GeometricMorphometricsMix* (Fruciano 2020). In particular, we used permutation tests to compare the disparity between groups employing two test statistics: multivariate variance and mean pairwise Euclidean distances. These are statistics widely used to characterize the amount of variation in both traditional and geometric morphometric samples (Ciampaglio et al. 2001; Kaplan 2004; Fruciano et al. 2014; Fruciano et al. 2016). In addition to these, we also computed in *Geometric Morpho-metricsMix* rarefied estimates of the following statistics: multivariate variance, mean pairwise Euclidean distances, proper variance (Claramunt 2010), and multidimensional convex hull volume. All of these, except proper variance, are widely used measures of disparity and computing race-specific rarefied estimates and confidence intervals allow further comparison between groups of the amount of dispersion around mean shape. In the procedure implemented in *GeometricMorphometricsMix*, resampling is performed with replacement (like in bootstrap, but resampling at a smaller sample size than the observed). As to the best of our knowledge, proper variance (Claramunt 2010) has never been used with geometric morphometric data, it is worth noticing that this statistic accounts not only for the level of dispersion (like multivariate variance does), but also for how this variance is spread across dimensions in terms of eigenvalues of the sample covariance matrix (thereby accounting for integration). A linear shrinkage estimator of the covariance matrix (Ledoit and Wolf 2004) was used to compute proper variance, using the implementation in *GeometricMorphometricsMix*. In all cases except convex hull volume, we used the Procrustes-aligned coordinates and rarefied at a sample size of 15. For convex hull volume, computed using the Quickhull algorithm (Barber et al. 1996), we used the scores along the first 4 principal components and rarefied to a sample size of 10. The choice of using a subset of principal components is common in morphometric studies computing convex hull volume (Drake and Klingenberg 2010; Fruciano et al. 2012, 2014), and it is necessary because computing a convex hull requires fewer dimensions than cases and because the resampling with replacement allows for cases where the same specimen is sampled multiple times (thus further reducing the number of dimensions that can be used).

We also performed the analyses of disparity outlined above—except for the computation of multivariate convex hull rarefied estimates—on data corrected for allometry, obtained as the residuals of regressions of mandible shape on mandible centroid size performed separately between the two races. We excluded from the analysis of disparity on allometry-corrected data the intermediate karyotypes because the results of the general linear models and the lower number of specimens for several diploid number suggest that it would not be possible to meaningfully correct for allometry these specimens.

Analyses of modularity and integration within and between the two main races

For all the analyses of modularity and integration, we divided the configuration of landmarks in two sets, corresponding to the ascending ramus (posterior part of the mandible) and alveolar region (anterior portion) (Figure 1). We first compared the levels of association of ascending ramus and alveolar region between the two races $2n = 22$ and $2n = 40$. This was obtained with the permutation test developed by Fruciano et al. (2013), which is based on a permutation of observations between groups, as implemented in *GeometricMorpho-metricsMix*. This test uses as statistic the Escoufier RV coefficient (Escoufier 1973), which is a measure of the association between two blocks of variables (a multivariate analog of the correlation coefficient). A smaller value of Escoufier RV can, then, be interpreted as a lower association between the two parts of

the mandible and, therefore, corresponds to higher levels of modularity (independence between parts). Fruciano et al. (2013) have unequivocally shown that the value of Escoufier RV depends on sample size and, for this reason, its value should not be compared directly between groups with unequal number of observations (the two chromosomal races in our case). However, both its use as the test statistic in permutation tests and the computation of rarefied estimates (rarefying both groups to the same sample size) are appropriate, as both approaches account for the dependency of the test statistic on sample size (Fruciano et al. 2013). Therefore, in addition to performing the permutation test, we also computed in *GeometricMorphometricsMix* rarefied estimates of Escoufier RV for both races, rarefying them to a sample size of 15. We performed these analyses both using data containing allometric variation and using allometry-corrected data, obtained by fitting separate regressions of shape on centroid size for each race and using the residuals in subsequent analyses. For the data containing allometric variation, we also performed these analyses both using a single GPA for the whole mandible and with separate GPAs for each part. This is because it is well known that GPA introduces strong non-independence between parts of the landmark configuration. This, in turn, has the potential of being particularly problematic for analyses of modularity and integration by introducing spurious association between modules (Cardini 2019).

To further test the patterns of integration between ascending ramus and alveolar region, we used partial least squares (PLS) analysis (Rohlf and Corti 2000) which allows one to describe how one part co-varies with the other. First, we fitted PLS models separately for the two races $2n = 40$ and $2n = 22$ and tested the significance of the association (i.e. how strongly the parts co-vary) using a permutational procedure with Escoufier RV as test statistic (1,000 permutations). As typical for this kind of analysis testing for the association between two blocks of variables in one sample, the observations of one of the blocks are permuted to reflect the null hypothesis of independence between the blocks. We performed the same analysis both using a single alignment for both blocks (the full mandible) and using separate GPAs. In addition to testing for global association between the two parts of the mandible using Escoufier RV as test statistic, we also used the same permutation scheme to test for the significance of individual PLS axis pairs based on both correlations of PLS scores (i.e. for each pair of axes, correlation between left PLS scores and right PLS scores) and singular values (Rohlf and Corti 2000). Next, we tested whether the patterns of co-variation detected by PLS (i.e. how shape of the two parts co-vary) were different between the two races. To this aim, as PLS identifies pairs of axes (one per module; called "singular axes"), we performed in *GeometricMorphometricsMix* an angular comparison of the first PLS axis (singular axis) obtained for the $2n = 40$ race with the corresponding axis obtained for the $2n = 22$ race. We did this analysis for both the data with a single GPA and the data with separate GPAs and used for the testing the method in Klingenberg and Marugán-Lobón (2013), which uses the formulas in Li (2011). With this approach, a significant test implies that the angle is small enough that the two multivariate vectors being tested are "significantly similar," whereas a non-significant test suggests that the vectors have different directions (i.e. reflect distinct shape changes).

Deviations from patterns of integration in the standard race

The central question of this study is whether accumulating chromosomal rearrangements results in a progressive departure from "normal" patterns of integration. To test this question, we used as reference the pattern of integration between the alveolar region and the ascending ramus observed in the standard race ($2n = 40$). The expectation is, then, that as chromosomal arrangements accumulate, the individuals bearing them will have a mandible shape progressively more dissimilar from the one predicted by the model of integration in the standard race. In this respect, it is worth noticing that each individual of the standard, reference, race will show some level of deviation from the general pattern predicted for the race as a whole (Figure 2). When data from new specimens with a different karyotype are projected onto the multivariate space spanned by the first pair of PLS axes, these new observations will themselves show some deviation from the general pattern (Figure 2). It is possible that the deviations of the new observations will be larger than the deviations from the original observations simply because the model has been computed using the original observations. However, even in this case, under the null hypothesis that the deviation from the "normal" pattern of integration is not associated to the accumulation of chromosomal rearrangements, one would not observe a progressive increase in deviations from the model. In our case, as we are using the standard karyotype ($2n = 40$) as reference and chromosomal rearrangements accumulate up to the extreme case of $2n = 22$, we expect under the hypothesis we want to test (accumulating chromosomal rearrangements result in a progressive departure from "normal" patterns) an inverse relationship between diploid number and deviation from the "normal" model. Here, based on the PLS analysis on the standard ($2n = 40$) race, we define the model as the major axis of the PLS scores for each block. This can be

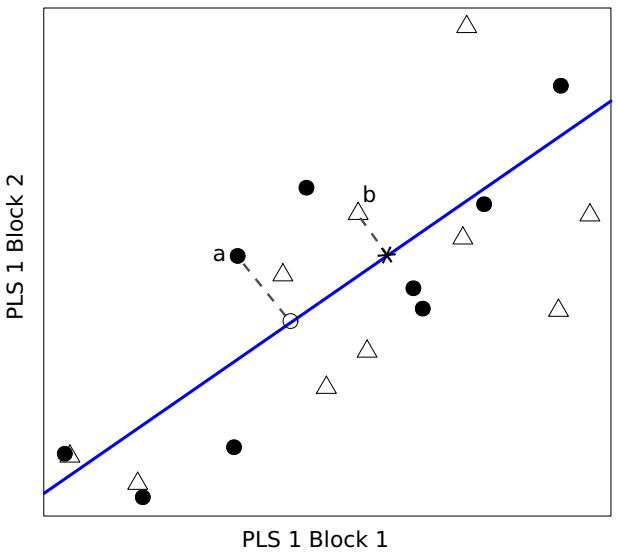

Figure 2. Conceptual representation of the computation of individual-level prediction from the PLS model, as used in our analysis of divergence from expectations under the model observed in standard house mice. Horizontal and vertical axes represent the first pair of PLS axes (one axis per module), as computed using only a subset of the specimens. Filled circles indicate the specimens based on which the PLS model has been computed. Empty triangles indicate new specimens projected on the PLS 1 space computed based on the specimens represented by the circles. The solid blue line is prediction line of the PLS model computed either as the first principal component of PLS 1 scores for filled circles or using major axis regression. The empty circle is the individual-level prediction for individual a, obtained by projecting (dashed line) the scores for individual on the prediction line. The asterisk represents the individual-level prediction for individual b, obtained by projecting (dashed line) the scores for individual b on the prediction line.

equivalently computed as the first principal component of such PLS scores or using major axis regression. Once this predicted model is inferred, the individual scores along this axis can be computed by projecting each observation on the axis itself (Figure 2). Each individual projection on this axis can then be expressed as a shape by first back-transforming this projection into scores on the first pair of PLS axis and then back-transforming these scores into shape variables. This can be interpreted as the individual-level predicted shape under the model of integration. For each individual, one can then compute the deviation between the original shape of that individual and the predicted shape for that individual under the fitted model. Here, we express this deviation as the Procrustes distance between the original shape and the prediction. We perform this analysis both using the data from a single GPA and using the data from separate GPAs. To facilitate the use of this approach in future studies, this method has been implemented in *GeometricMorphometricsMix*.

Furthermore, to more quantitatively asses this potential relationship, we compute and test for significance the correlation between the deviation from the PLS model computed on the standard race (expressed in Procrustes distance) and diploid number. To allow for both linear and non-linear relationships, we employ both the usual Pearson product moment correlation coefficient, which is best suited for linear relationships, and a recently proposed correlation coefficient *Xi* (Chatterjee 2020) which is asymptotically comprised between 0 and 1 and is useful for non-linear and non-monotonic relationships. As *Xi* is not a symmetric coefficient (Chatterjee 2020), we computed it in the R package *XICOR* (Chatterjee 2020) considering Procrustes distance from standard integration as dependent on diploid number (as per our biological hypothesis).

To test whether the association detected by PLS analysis could be simply explained by allometry, we performed a further angular test. For this test, we considered the case with separate GPAs and for each module we tested whether the vector of coefficients of a multivariate regression of shape on size for all specimens (including also individuals with intermediate karyotype) was different from the corresponding vector of PLS coefficients for the standard race. This means asking whether the shape change due to allometry is approximately the same as the shape change due to "normal" co-variation between modules.

Results

Chromosomal structure of the hybridization zone

Two parental chromosomal races are located at the extremes of our sampling area, the CD race with 9 Rb fusions ($2n = 22$) at the north end and the "Standard" ($2n = 40$) all acrocentric chromosomes at the south end. The chromosome number generally increases along the study transect, that includes 17 sampling localities, from the CD to the standard area (Figure 1). All the Rb chromosomes that characterize the CD race have frequencies that show a gradual decline from north (Rb area) to south (standard area). The hybrid individuals have karyotypic number ranging from 24 to 28 in the northern localities, and from 34 to 39 in the southern localities (Figure 1 and Supplementary Table S1).

General linear models

Both the traditional approach of fitting and testing general linear models on scores along non-zero PC axes and the more recent approach based on penalized likelihood produced consistent results. For this reason, the results based on the penalized likelihood framework will be provided in the main text, whereas the ones based on the traditional approach will be provided in the Supplementary Material.

Table 1. MANOVA based on a penalized-likelihood approach and type III sum of squares

	All specimens		Two main races only	
Term	Pillai's trace	*P*-value	Pillai's trace	*P*-value
(Intercept)	1	**0.001**	1	**0.001**
Centroid size	1.291	**0.044**	1.5316	0.114
Diploid number	1.25	**0.018**	1.4027	0.059
Centroid size x Diploid number	1.224	**0.011**	0.4804	**0.045**

Significant *P*-values are in bold.

Sex has no significant effect on shape, neither when tested alone ($P = 0.37$ with both approaches) nor when tested in conjunction with size (Supplementary Table S2).

Conversely, both diploid number and centroid size—as well as their interaction—had a significant effect on mandible shape (Table 1 and Supplementary Table S3). This result was broadly confirmed when testing exclusively the $2n = 22$ and $2n = 40$ specimens. This suggests that allometry has a significant impact on mandible shape, and that allometric trajectories vary depending on chromosomal number. This is evident when examining plots of PCA scores for predictions (Figure 3), as well as when comparing predicted allometric shape change within the 2 main chromosomal races (Figure 4). It is worth noticing that, as the diploid number changes, we do not observe a gradual change in the predicted slope of the relationship between shape and size (Figure 3). Rather, we observe a fairly abrupt difference between low diploid number and high diploid number. It is, however, possible that this pattern is driven by small sample sizes for intermediate diploid numbers, as well as the absence of karyotypes with diploid number comprised between 28 and 34.

Analyses of disparity

In all cases, the two races ($2n = 22$ and $2n = 40$) did not show substantial differences in disparity. Specifically, all permutation tests were nonsignificant and, therefore, we failed to reject the null hypothesis of equal disparity between the two races (Table 2). In a similar fashion, rarefied estimates of disparity statistics do not suggest substantial differences between the races, as 95% confidence intervals are largely overlapping and the mean of one group is always within 2 standard deviations from the mean of the other (Supplementary Table S4). When the focal races were compared with individuals characterized by intermediate karyotypes (Supplementary Table S4), the latter showed always the lowest point estimates of disparity measures. However, tests of differences were always not significant and confidence intervals for rarefied estimates always overlapped (Supplementary Table S4).

Analyses of modularity and integration within and between the two main races

All the analyses comparing levels of modularity between the chromosomal races $2n = 22$ and $2n = 40$ fail to reveal any substantial difference between them. Indeed, all permutation tests based on the difference in Escoufier RV between the two races were nonsignificant (single GPA, observed absolute difference in RV = 0.086,

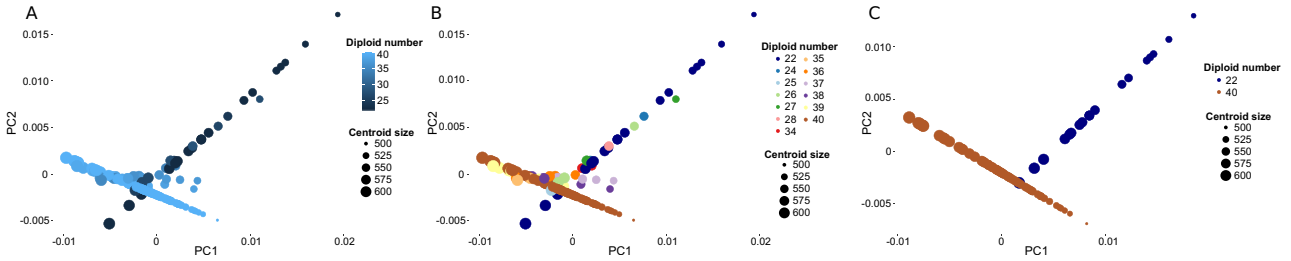

Figure 3. Scores along the first two principal components obtained subjecting to a PCA the predictions of the linear models fit using shape variables as dependent variables and both diploid number and size (centroid size) as predictors. (**A**) Using a continuous color scale for diploid number. (**B**) Using a discrete color scale for diploid number. (**C**) Restricting the analysis to the standard ($2n = 40$) and the CD ($2n = 22$) races.

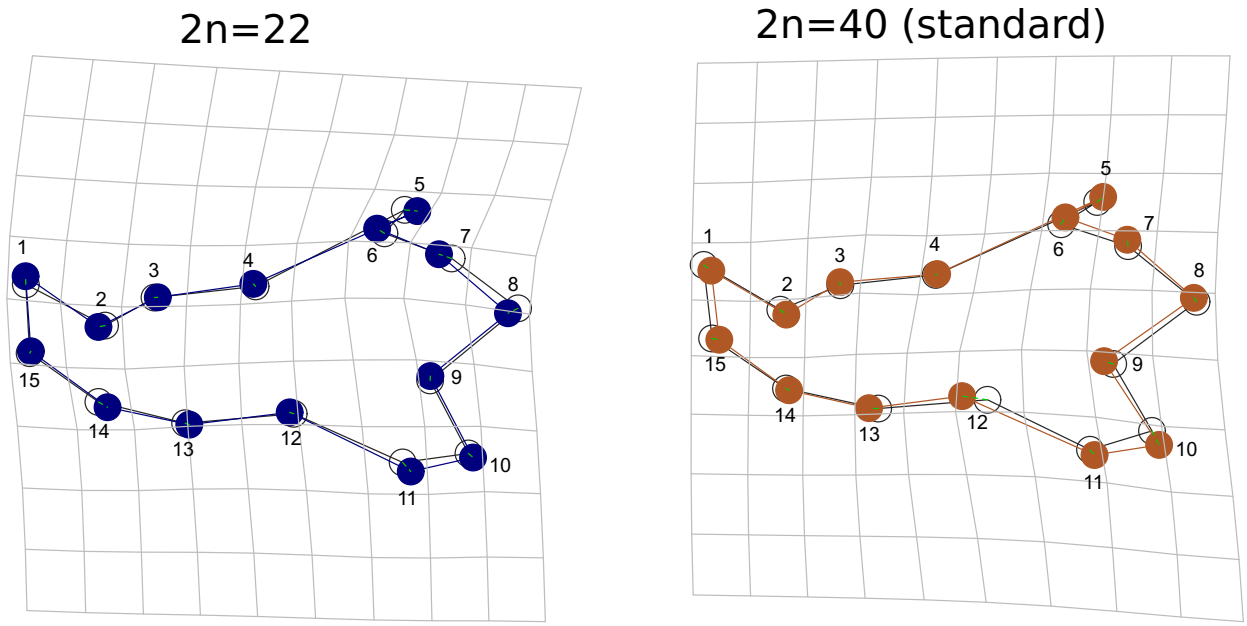

Figure 4. Allometric shape variation predicted for the two main races studied here. The predictions are obtained by first fitting a linear model on specimens belonging to the two races. The models use as dependent variable shape variables and as predictors size, diploid number (race) and their interaction. Within each of the two races, the predicted shapes at the extreme observed value of centroid size are then obtained and plotted against each other. Empty gray circles:

Table 2. Tests for difference in disparity between the two main races

Using the original data				
	Observed $2n = 22$	Observed $2n = 40$	Difference	*P*-value
Multivariate variance	0.002	0.002	<0.001	0.145
Mean pairwise Euclidean distance	0.064	0.058	0.006	0.150
Using data after allometric correction				
	Observed $2n = 22$	Observed $2n = 40$	Difference	*P*-value
Multivariate variance	0.002	0.002	<0.001	0.069
Mean pairwise Euclidean distance	0.061	0.055	0.006	0.084

$P = 0.843$; separate GPAs, observed absolute difference in RV = 0.097, $P = 0.802$; single GPA and allometry correction, observed absolute difference in RV = 0.204, $P = 0.295$). The same pattern is confirmed by computing rarefied estimates of Escoufier RV, which overall show similar levels of association between ascending ramus and alveolar region, as well as largely overlapping 95% confidence intervals (Supplementary Table S5).

The analyses of the association between ascending ramus and alveolar region provided evidence of significant association between these modules in the standard race ($2n = 40$; single GPA, RV = 0.314, $P < 0.001$; separate GPAs, RV = 0.164, $P = 0.008$). Conversely, the results for the $2n = 22$ race were mixed (single GPA, RV = 0.4, $P = 0.043$; separate GPAs, RV = 0.261, $P = 0.252$). Permutation tests of the significance of individual axes broadly

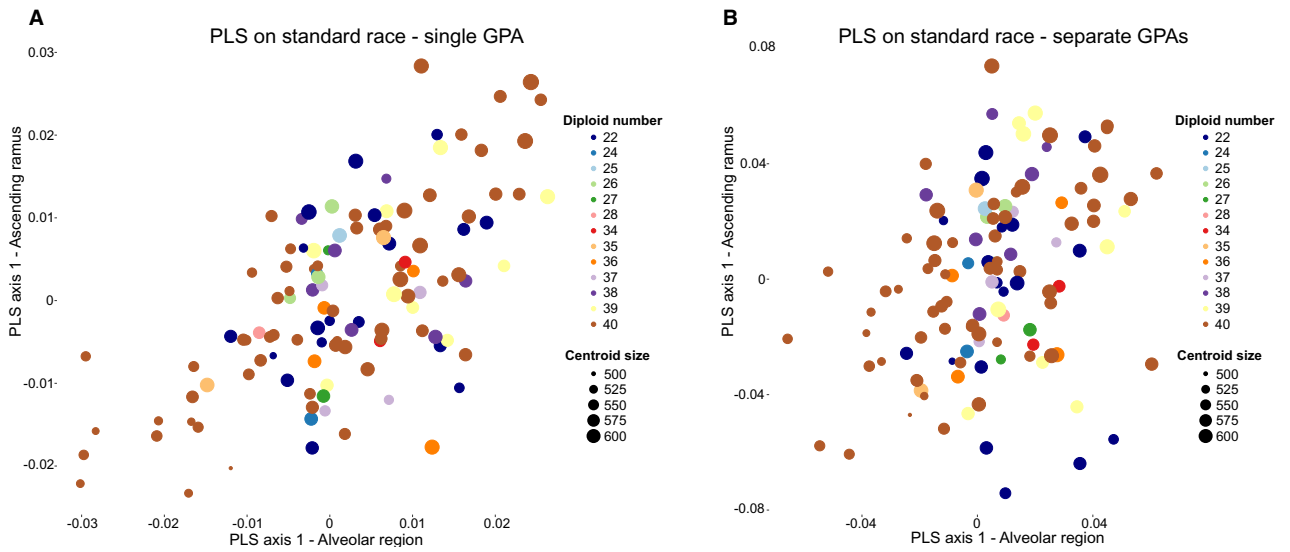

Figure 5. Scores of all studied specimens along the first pair of PLS axes computed using only the standard race ($2n = 40$). (**A**) Single GPA for all the mandible and (**B**) performing separate GPAs for each module. Notice that in this case, all specimens have been centered to the mean used for the computation of the PLS axes prior to projection onto the axes for the computation of scores.

confirmed the results of the analyses based on Escoufier RV, with fewer axes significant when performing separate GPAs (Supplementary Table S6). This result is expected under the reasonable assumption that a common GPA is more likely to induce spuriously significant association between modules (Cardini 2019).

The angular comparisons performed on the result of separate PLS fits for the two races were always nonsignificant whether performing a single GPA (alveolar region, angle = 69.33°, $P = 0.082$; ascending ramus, angle = 73.22°, $P = 0.148$) or separate GPAs (alveolar region, angle = 77.45°, $P = 0.201$; ascending ramus, angle = 76.23°, $P = 0.196$). The critical angle—the angle below which the test would be significant, which depends on the dimensionality of the multivariate vectors being tested—was 65.65° for the alveolar region, and 63.84° for the ascending ramus. This means that the patterns of co-variation between parts of the mandible detected by PLS analysis are different between the two main races studied here.

Deviations from patterns of integration in the standard race

We found no evidence of a gradual departure of shape from the "normal" pattern of integration as chromosomal rearrangements accumulate. Indeed, by fitting a PLS model describing the covariation between the alveolar region and the ascending ramus in the standard race ($2n = 40$) and then projecting all observations in the space described by the first pair of PLS axes does not show any clear pattern (Figure 5). In particular, this plot does not show evidence of larger deviations in the other karyotypes (Figure 5). The only pattern that is somewhat evident is some association between PLS scores and mandible centroid size (Figure 5), which is suggestive of the fact that the main pattern of integration between parts of the mandible is largely driven by allometry. An exploratory plot of individual-level deviations from overall patterns of integration described in the standard race (Figure 6) further confirms the absence of a clear pattern of larger deviations at lower diploid numbers (lower diploid number in our case means more chromosomal rearrangements departing from the standard $2n = 40$ race). This result holds whether we use data from a single GPA for the whole mandible or separate GPAs for each module as starting data. When more explicitly testing for the correlation between diploid number and the Procrustes distance of each individual shape from its predicted shape under the PLS model computed for the standard race, this was always small and not significant, except for the ascending ramus in the analysis with separate GPAs. Indeed, using a single GPA, Pearson correlation was −0.09 ($P = 0.335$) and Xi was 0.04 ($P = 0.245$). Using separate GPAs, for the alveolar region, Pearson correlation was −0.12 ($P = 0.19$) and Xi was 0.06 ($P = 0.14$) although for the ascending ramus Pearson correlation was −0.22 ($P = 0.019$) and Xi was 0.1 ($P = 0.04$). However, when computing the correlation using separate GPAs on the ascending ramus but excluding the two main races, this became positive and not significant with Pearson correlation equal to 0.1 ($P = 0.538$) and Xi equal to 0.06 ($P = 0.28$).

The comparison of the allometric patterns found for the full sample with the patterns of co-variation between modules obtained in the standard ($2n = 40$) race supported the idea that the "normal" pattern of co-variation can be largely explained by allometry (alveolar region, angle = 30.04°, $P = 3.57\text{E-}6$; ascending ramus, angle = 45.87°, $P = 0.002$).

Discussion

The role of chromosomal evolution in promoting genetic and phenotypic divergence is a hotly debated topic in evolutionary biology (Coyne and Orr 2004). Among the different types of chromosomal rearrangements (e.g. translocations, inversions, duplications, and fusions), Rb fusions are a prominent class of structural variants often associated with speciation (Faria and Navarro 2010; Kirkpatrick 2017). Rb fusions are likely the most common chromosomal rearrangements in animals (King 1993) and they occur in ~0.1% of human meiotic divisions (Evans et al. 1982; Song et al. 2016). These fusions are also frequent in sheep, cattle (Pagacova et al. 2011), and in the Western European house mouse, a model system in chromosomal speciation (Pialek et al. 2005). Here, we analyzed morphological variation in the house mouse mandible in a hybrid zone between a highly derived chromosomal race (CD race:

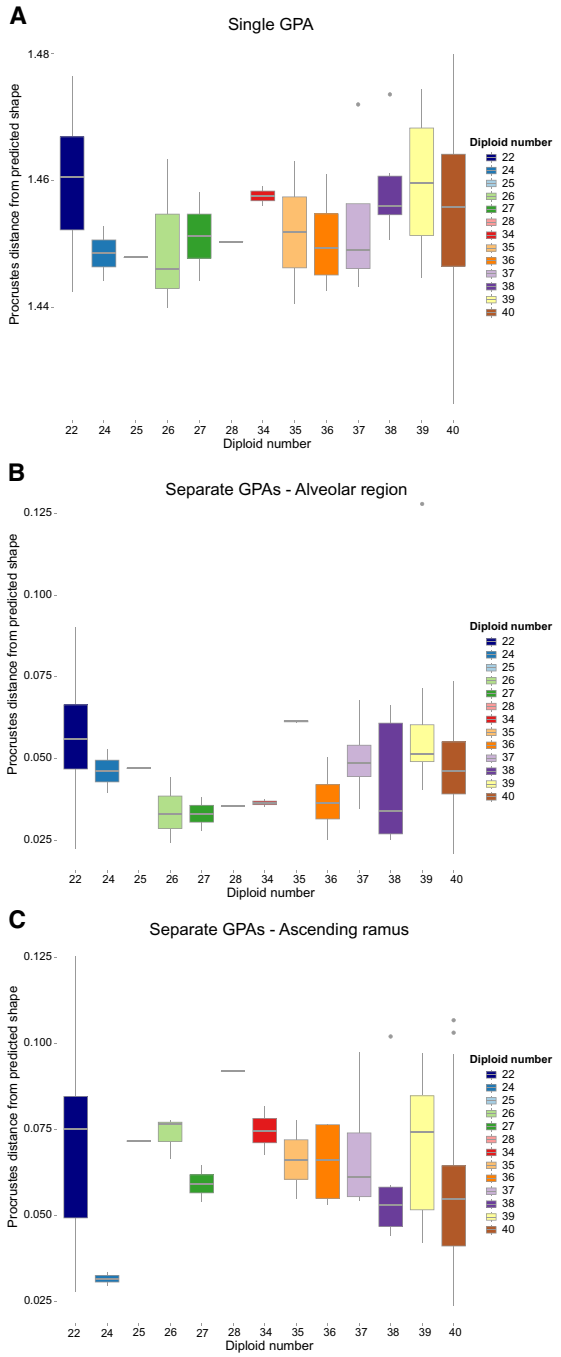

Figure 6. Individual-level deviations from the pattern of integration predicted by the first pair of PLS axes computed on the standard race for (**A**) a single GPA and (**B** and **C**) separate GPAs for each of the two modules. The box plot visualizes intra-karyotype variation. The box delimits the first and third quartile, whereas the dark line within the box is the median. The lower and upper whiskers refer to the smallest observation greater than or equal to lower hinge −1.5 * interquartile range (IQR) and the largest observation less than or equal to upper hinge + 1.5 * IQR. All observations with values larger or smaller than the one at the whiskers are marked as outliers (dot). Notice that this is just a visual representation of the individual deviations from patterns predicted by the first pair of PLS axis and no statistical testing has been performed using diploid number as a categorical predictor. Notice also that to compute the predicted shape under the model, the mean subtracted prior to projection (see Figure 5 legend) is added back so that the difference between the actual shape and the shape estimated under the model is not overestimated.

$2n = 22$, 9 Rb fused chromosomes) and surrounding all-acrocentric standard populations (Capanna et al. 1975). The most striking pattern we observe in our results is that individual-level deviations from "normal" patterns of integration do not simply increase as the chromosomal fusions increase. Conversely, we do observe an important effect of allometry in driving the main patterns of integration and a difference in allometric trajectories between the standard ($2n = 40$) populations and the most derived race studied here ($2n = 22$). In particular, as in this study we analyzed mandibles from adult specimens only, these patterns can be attributed to static allometry (Klingenberg 1998).

More in detail, the two main races do not appear to have neither different levels of overall disparity (amount of mandible morphological variation) nor different levels (strength) of integration between modules. As the two main races studied here have similar disparity, we can reject the reasonable hypothesis that a high number of Rb fusions produces—through developmental or genetic disruptions—a reduced number of viable phenotypes. It is interesting to notice that the lowest point estimates of disparity were obtained when analyzing intermediate karyotypes. This is particularly interesting as almost all individuals with intermediate karyotypes are heterozygous for at least one fusion. However, as these differences are not significant, this result should be confirmed in future studies with a more homogeneous sampling of intermediate karyotypes. This result is interesting as previous research suggested how Rb fusions in a heterozygous state might affect developmental stability of several house mouse tooth traits (Chatti et al. 1999; Auffray et al. 2001). While contrasting results on the developmental disruption in population differing by a single Rb fusion have been observed, an additive and/or a threshold relationship would be expected between the number of Rb fusions and disruption of genomic coadaptation that can affect hybrid viability (Graham 1992). In our study, the 2 focal parental races are highly different in their chromosome number, and thus a developmental disruption reflected in different level of overall disparity was a reasonable hypothesis to test. At least two explanations, not mutually exclusive, might suggest why we did not observe the expected pattern of difference in disparity between the two races. First, even though the occurrence of developmental instability might have been prominent during the fixation of the Rb fusions, the relatively old age of the CD race could have allowed multiple potential evolutionary forces to shape a number of coadapted genotypes. Second, the type of Rb fusions characterizing the CD race may have had little and not easily detectable effect on normal development.

Similarly, the two races do not display significant differences in the strength of association between ascending ramus and alveolar region when tested directly (permutation tests and rarefied estimates based on Escoufier RV). When testing for the association between ascending ramus and alveolar region within the two main races ($2n = 40$, $2n = 22$) we do get mixed results, but these are most likely due to different statistical power caused by the substantial difference in sample sizes between these races. In fact—although performing separate alignments certainly protects from spuriously significant results due to the non-independence induced by the GPA (Cardini 2019)—it also makes it harder to detect the co-variation. Considering that all the other analyses directly comparing the two races for the levels of association between modules show no evidence of differences between races, a lower statistical power for the $2n = 22$ race appears the most likely explanation for the discrepancy of these within-race analyses. The modular behavior of the mandible has been investigated in previous studies targeting contact areas

involving different chromosomal races and Rb fusions. For instance, a previous morphometric survey of house mouse mandibles in the "Barcelona" Rb system has reported a significant association between variation in levels of modularity and variation in karyotype for data not subjected to allometric correction. This association vanished when accounting for allometry (Martínez-Vargas et al. 2014). However, similar to what we have found in this study, no significant differences in the strength of association between modules was observed in a contact area involving two highly metacentric races in Central Italy (the same CD race of this study and the ACR race with 8 Rb fusions) (Franchini et al. 2016).

While the two races do not have clear differences in the levels of integration between ascending ramus and alveolar region (i.e. how strong is the co-variation between these two parts), this does not mean that the patterns of co-variation (i.e. in which way these two parts co-vary) have also to be the same. Indeed, this is exactly what we find in our analyses. The two parts co-vary in a different way across the two races because we find that the PLS axes describing this co-variation are different between races. We do find a similar pattern when testing for allometry. That is, we observe different allometric trajectories for the two races (significant interaction term between diploid number and centroid size in general linear models). The impact of different environmental pressures on the morphology of individuals with different karyotypes cannot be completely ruled out as an explanation for this pattern. For instance, it is known that the bioclimatic conditions at the sampling sites where we collected the standard and CD races are slightly different (Castiglia and Capanna 1999). Italy covers two macrobioclimates (i.e. Mediterranean and Temperate). The Mediterranean macrobioclimate is characterized by at least two consecutive arid summer months, whereas the Temperate macrobioclimate does not have any summer aridity (Pesaresi et al. 2017). The hybrid zone is at the border between the two macrobioclimates, where populations with $2n = 22$ and $2n = 40$ karyotype are settled in the Temperate and in the Mediterranean zone, respectively. However, it is doubtful whether this would have a substantial effect on the house mouse, which is a commensal of humans and in this area is found associated to farms (and therefore subjected to a comparable diet regime). Additionally, a comprehensive study on the role of Rb rearrangements on the growth pattern of the house mouse mandible over early postnatal ontogeny showed significant differences between Rb and standard mice (Martínez-Vargas et al. 2018). Notably, these differences were found to be more evident in individuals around sexual maturity. The two groups of experimental mice were reared under similar laboratory conditions, thus allowing the authors to exclude morphological variability induced by different environmental pressures. As in our study we used adult mice, it is reasonable to assume that the different patterns of integration we detected could be underlined, at least in part, by a genetic component (i.e. in the form of presence/absence of Rb rearrangements). Furthermore, our results point to the fact that most likely individual Rb fusions have a disproportionate effect on adherence to "normal" patterns of integration, as opposed to a gradual change at increasing number of fusions. Finally, we also highlight that even a highly derived Rb race might have adjusted its development to produce a perfectly viable new race with a distinct allometric trajectory. In the light of all these results, future studies combining high-resolution genomic information and advanced morphometrics will have to tease apart which regions of the genome are most responsible for the observed change in trajectories and whether these are regions of reduced gene flow during incipient chromosomal divergence.

Acknowledgments

We thank the Guest Editor and the two anonymous reviewers who, with their suggestions, allowed us to greatly improve this manuscript.

Funding

This work was supported by the Ministry of Scientific Research and Education, Italy (MIUR).

Ethics Declarations

The authors declare no competing interests.

Supplementary Material

Supplementary material can be found at https://academic.oup.com/cz.

References

Atchley WR, Hall BK, 1991. A model for development and evolution of complex morphological structures. *Biol Rev Camb Philos Soc* **66**:101–157.

Auffray JC, Fontanillas P, Catalan J, Britton-Davidian J, 2001. Developmental stability in house mice heterozygous for single Robertsonian fusions. *J Hered* **92**:23–29.

Barber CB, Dobkin DP, Huhdanpaa H, 1996. The quickhull algorithm for convex hulls. *ACM Trans Mathe Softw* **22**:469–483.

Capanna E, Cristaldi M, Perticone P, Rizzoni M, 1975. Identification of chromosomes involved in the 9 Robertsonian fusions of the Apennine mouse with a 22–chromosome karyotype. *Experientia* **31**:294–296.

Cardini A, 2019. Integration and modularity in procrustes shape data: is there a risk of spurious results? *Evolut Biol* **46**:90–105.

Castiglia R, Capanna E, 2000. Contact zone between chromosomal races of *Mus musculus domesticus*. 2. Fertility and segregation in laboratory-reared and wild mice heterozygous for multiple Robertsonian rearrangements. *Heredity* **85**:147–156.

Castiglia R, Capanna E, 1999. Contact zones between chromosomal races of *Mus musculus domesticus*. 1. Temporal analysis of a hybrid zone between the CD chromosomal race (2n = 22) and populations with the standard karyotype. *Heredity* **83**:319–326.

Castiglia R, Gornung E, Cividin M, Cristaldi M, 2011. High diversity of centric fusions with monobrachial homology in an area of chromosomal polymorphism of *Mus musculus domesticus*. *Biol J Linnean Soc* **103**:722–731.

Charron G, Leducq JB, Landry CR, 2014. Chromosomal variation segregates within incipient species and correlates with reproductive isolation. *Mol Ecol* **23**:4362–4372.

Chatterjee S, 2020. A new coefficient of correlation. *J Am Stat Assoc* 1–21. doi: 10.1080/01621459.2020.1758115.

Chatti N, Said K, Catalan J, Britton-Davidian J, Auffray JC, 1999. Developmental instability in wild chromosomal hybrids of the house mouse. *Evolution* **53**:1268–1279.

Ciampaglio CN, Kemp M, McShea DW, 2001. Detecting changes in morphospace occupation patterns in the fossil record: characterization and analysis of measures of disparity. *Paleobiology* **27**:695–715.

Claramunt S, 2010. Discovering exceptional diversifications at continental scales: the case of the endemic families of neotropical suboscine passerines. *Evolution* **64**:2004–2019.

Clavel J, Aristide L, Morlon H, 2019. A penalized likelihood framework for high-dimensional phylogenetic comparative methods and an application to new-world monkeys brain evolution. *Syst Biol* **68**:93–116.

Clavel J, Escarguel G, Merceron G, 2015. *mvMORPH*: an R package for fitting multivariate evolutionary models to morphometric data. *Methods Ecol Evol* **6**:1311–1319.

Clavel J, Morlon H, in press. Reliable phylogenetic regressions for multivariate comparative data: illustration with the MANOVA and application to the

effect of diet on mandible morphology in Phyllostomid bats. *Syst Biol* doi: 10.1093/sysbio/syaa010.

Corti M, Rohlf FJ, 2001. Chromosomal speciation and phenotypic evolution in the house mouse. *Biol J Linn Soc* **73**:99–112.

Coyne JA, Orr HA, 2004. *Speciation*. Sunderland (MA): Sinauer Associates, Inc.

Drake AG, Klingenberg CP, 2010. Large-scale diversification of skull shape in domestic dogs: disparity and modularity. *Am Nat* **175**:289–301.

Dryden IL, Mardia KV, 1998. *Statistical Shape Analysis*. Chichester: Wiley Chichester.

Ehrich TH, Vaughn TT, Koreishi S, Linsey RB, Pletscher LS et al., 2003. Pleiotropic effects on mandibular morphology I. Developmental morphological integration and differential dominance. *J Exp Zool B Mol Dev Evol* **296b**:58–79.

Escoufier Y, 1973. Le traitement des variables vectorielles. *Biometrics* **29**: 751–760.

Evans JA, Devonflindt R, Greenberg C, Ramsay S, Hamerton JL, 1982. A cytogenetic survey of 14,069 newborn–infants. Further follow–up on the children with sex–chromosome anomalies. *Birth Defects Orig Artic Ser* **18**: 169–184.

Faria R, Navarro A, 2010. Chromosomal speciation revisited: rearranging theory with pieces of evidence. *Trends Ecol Evol* **25**:660–669.

Franchini P, Castiglia R, Capanna E, 2008. Reproductive isolation between chromosomal races of the house mouse *Mus musculus domesticus* in a parapatric contact area revealed by an analysis of multiple unlinked loci. *J Evol Biol* **21**:502–513.

Franchini P, Colangelo P, Meyer A, Fruciano C, 2016. Chromosomal rearrangements, phenotypic variation and modularity: a case study from a contact zone between house mouse Robertsonian races in Central Italy. *Ecol Evol* **6**:1353–1362.

Franchini P, Colangelo P, Solano E, Capanna E, Verheyen E et al., 2010. Reduced gene flow at pericentromeric loci in a hybrid zone involving chromosomal races of the house mouse *Mus musculus domesticus*. *Evolution* **64**:2020–2032.

Fruciano C, 2020. Geometric morphometrics mix: miscellaneous functions useful for geometric morphometrics. version 0.0.8. Available from: https://github.com/fruciano/GeometricMorphometricsMix (accessed June 20 2020).

Fruciano C, Franchini P, Meyer A, 2013. Resampling-based approaches to study variation in morphological modularity. *PLoS ONE* **8**:e69376.

Fruciano C, Franchini P, Raffini F, Fan S, Meyer A, 2016. Are sympatrically speciating Midas cichlid fish special? Patterns of morphological and genetic variation in the closely related species *Archocentrus centrarchus*. *Ecol Evol* **6**:4102–4114.

Fruciano C, Pappalardo AM, Tigano C, Ferrito V, 2014. Phylogeographical relationships of Sicilian brown trout and the effects of genetic introgression on morphospace occupation. *Biol J Linn Soc* **112**:387–398.

Fruciano C, Tigano C, Ferrito V, 2012. Body shape variation and colour change during growth in a protogynous fish. *Environ Biol Fish* **94**:615–622.

Graham JH, 1992. Genomic coadaptation and developmental stability in hybrid zones. *Acta Zool Fenni* **191**:121–131.

Gropp A, Winking H, 1981. Robertsonian translocations: cytology, meiosis, segregation pattern and biological consequences of heterozygosity. *Sympe Zool Soc Lond* **47**:141–181.

Hauffe HC, Gimenez MD, Vega R, White TA, Searle JB, 2011. Properties of a hybrid zone between highly distinct chromosomal races of the house mouse *Mus musculus domesticus* in Northern Italy, and comparisons with other hybrid zones. *Cytogenet Genom Res* **134**:191–199.

Imprialou M, Kahles A, Steffen JG, Osborne EJ, Gan X et al., 2017. Genomic rearrangements in arabidopsis considered as quantitative traits. *Genetics* **205**:1425–1441.

Kaplan P, 2004. Morphology and disparity through time. In: eLS. John Wiley & Sons Ltd, Chichester. http://www.els.net/. doi: 10.1038/npg.els.0001637.

King M, 1993. *Species Evolution: The Role of Chromosome Change*. Cambridge: Cambridge University Press.

Kirkpatrick M, 2017. The evolution of genome structure by natural and sexual selection. *J Hered* **108**:3–11.

Klingenberg CP, 2010. Evolution and development of shape: integrating quantitative approaches. *Nat Rev Genet* **11**:623–635.

Klingenberg CP, 1998. Heterochrony and allometry: the analysis of evolutionary change in ontogeny. *Biol Rev* **73**:79–123.

Klingenberg CP, Leamy LJ, Cheverud JM, 2004. Integration and modularity of quantitative trait locus effects on geometric shape in the mouse mandible. *Genetics* **166**:1909–1921.

Klingenberg CP, Marugán-Lobón J, 2013. Evolutionary covariation in geometric morphometric data: analyzing integration, modularity, and allometry in a phylogenetic context. *Syst Biol* **62**:591–610.

Klingenberg CP, Mebus K, Auffray JC, 2003. Developmental integration in a complex morphological structure: how distinct are the modules in the mouse mandible? *Evol Dev* **5**:522–531.

Leamy L, 1993. Morphological integration of fluctuating asymmetry in the mouse mandible. *Genetica* **89**:139–153.

Ledoit O, Wolf M, 2004. A well–conditioned estimator for large–dimensional covariance matrices. *J Multivariate Anal* **88**:365–411.

Li S, 2011. Concise formulas for the area and volume of a hyperspherical cap. *Asian J Mathe Stat* **4**:66–70.

Lindholm AK, Dyer KA, Firman RC, Fishman L, Forstmeier W et al. 2016. The ecology and evolutionary dynamics of meiotic drive. *Trends Ecol Evol* **31**:315–326.

Lowry DB, Willis JH, 2010. A widespread chromosomal inversion polymorphism contributes to a major life–history transition, local adaptation, and reproductive isolation. *PLoS Biol* **8**:e1000500.

Martínez-Vargas J, Muñoz-Muñoz F, Lopez-Fuster MJ, Cubo J, Ventura J, 2018. Multimethod approach to the early postnatal growth of the mandible in mice from a zone of robertsonian polymorphism. *Anat Rec* **301**: 1360–1381.

Martínez-Vargas J, Muñoz Muñoz J, Medarde N, López Fuster MJ, Ventura J, 2014. Effect of chromosomal reorganizations on morphological covariation of the house mouse mandible: insights from a Robertsonian system of *Mus musculus domesticus*. *Front Zool* **11**:51.

Merot C, Oomen RA, Tigano A, Wellenreuther M, 2020. A roadmap for understanding the evolutionary significance of structural genomic variation. *Trends Ecol Evol* **35**:561–572.

Muñoz-Muñoz F, Sans-Fuentes MA, Lopez-Fuster MJ, Ventura J, 2011. Evolutionary modularity of the mouse mandible: dissecting the effect of chromosomal reorganizations and isolation by distance in a Robertsonian system of *Mus musculus domesticus*. *J Evol Biol* **24**:1763–1776.

Noor MAF, Grams KL, Bertucci LA, Reiland J, 2001. Chromosomal inversions and the reproductive isolation of species. *Proc Natl Acad Sci USA* **98**: 12084–12088.

Pagacova E, Cernohorska H, Kubickova S, Vahala J, Rubes J, 2011. Centric fusion polymorphism in captive animals of family Bovidae. *Conserv Genet* **12**:71–77.

Panithanarak T, Hauffe HC, Dallas JF, Glover A, Ward RG et al., 2004. Linkage-dependent gene flow in a house mouse chromosomal hybrid zone. *Evolution* **58**:184–192.

Pesaresi S, Biondi E, Casavecchia S, 2017. Bioclimates of Italy. *J Maps* **13**: 955–960.

Pialek J, Hauffe HC, Searle JB, 2005. Chromosomal variation in the house mouse. *Biol J Linn Soc* **84**:535–563.

Pillai KCS, 1955. Some new test criteria in multivariate analysis. *Ann Mathe Stat* **26**:117–121.

Rieseberg LH, 2001. Chromosomal rearrangements and speciation. *Trends Ecol Evol* **16**:351–358.

Rohlf FJ, Corti M, 2000. Use of two-block partial least-squares to study covariation in shape. *Syst Biol* **49**:740–753.

Sage RD, Atchley WR, Capanna E, 1993. House mice as models in systematic biology. *Syst Biol* **42**:523–561.

Sans-Fuentes MA, Ventura J, Lopez-Fuster MJ, Corti M, 2009. Morphological variation in house mice from the Robertsonian polymorphism area of Barcelona. *Biol J Linn Soc* **97**:555–570.

Schlager S, 2017. Shape analysis in *R*: R -packages for geometric morphometrics, shape analysis and surface manipulations. In: Zheng G, Li S, Székely G, editors. *Statistical Shape and Deformation Analysis*. Cambridge (MA): Academic Press, 217–256.

Song JP, Li X, Sun L, Xu SQ, Liu N et al., 2016. A family with Robertsonian translocation: a potential mechanism of speciation in humans. *Mol Cytogenet* **9**:48.