Herpetologists' League

AN ANALYSIS OF FLUCTUATING ASYMMETRY IN A HYBRID ZONE BETWEEN TWO CHROMOSOME RACES OF THE *SCELOPORUS GRAMMICUS* COMPLEX (SQUAMATA: PHRYNOSOMATIDAE) IN CENTRAL MEXICO

Dennis J. Dosselman[1,3], G. Bruce Schaalje[2], and Jack W. Sites, Jr.[1]

[1]*Department of Zoology and M. L. Bean Life Science Museum, and* [2]*Department of Statistics, Brigham Young University, Provo, UT 84602, USA*

ABSTRACT: We performed an analysis of fluctuating asymmetry (FA) on samples of lizards from a hybrid zone between fission 5 (F5) and multiple fission (FM2) chromosome races of the *Sceloporus grammicus* complex in central Mexico. Defining the zone are three diagnostic chromosome markers that demonstrate sharp concordant changes from one parental race to the other over a geographic distance of about 1 km. These chromosome markers allowed five classes of genotypes to be designated: parental F5, parental FM2, F_1 hybrid, $F_1 \times$ F5 backcross, and $F_1 \times$ FM2 backcross. We counted seven bilateral meristic characters on each individual lizard ($n = 427$). These characters were canthal scales, parietal border scales, post-rostral scales, nasal border scales, circumorbital scales, pes toe lamellae of fourth digit, and manus toe lamellae of fourth digit. We tested the prediction that "hybrid" individuals (F_1, $F_1 \times$ F5 backcross, $F_1 \times$ FM2 backcross) should show lower levels of developmental stability and thus higher levels of FA than parental individuals. Hybrids demonstrated significantly lower FA levels than parental classes in most pairwise tests for canthal, parietal border, and post-rostral scale characters. Parentals demonstrated a significantly lower level of FA than hybrids for just one pairwise test (canthal scale). Thus, in general, these results do not support the hypothesis that hybrids should demonstrate lower levels of developmental stability and corresponding higher levels of FA.

Key words: Coadapted gene complexes; Developmental stability; Fluctuating asymmetry; Hybrid zone; Squamata; *Sceloporus grammicus* complex

FLUCTUATING asymmetry (FA) is one of three types of bilateral asymmetry (Van Valen, 1962), and it is defined as that asymmetry which results from an organism's inability to develop in a genetically encoded ideal manner. Leary et al. (1992) noted that there are four potential sources of the total variability expressed by a character in a population: somatic mutation, developmental accidents, genotypic variation among individuals, and environmental variation. The phenotypic variability represented by FA is due only to developmental noise, which can be defined as a non-genetic phenomenon representing the cumulative effects of small random developmental perturbations or accidents (Palmer and Strobeck, 1992). Developmental stability seems to have a genetic basis (Clarke, 1993) and refers to an organism's ability to resist or correct for de-

velopmental perturbations or accidents. Mather (1953) demonstrated that FA could be used as a measure of developmental stability.

Normally, FA can be measured by subtracting the left side value of a quantitative bilateral character from the right side value. The values used are typically either meristic counts or metric measurements. The distribution of right-minus-left differences, hereafter referred to as (R-L), can be used to differentiate FA from the two other types of asymmetry: directional asymmetry and antisymmetry. Directional asymmetry (also known as handedness) is present when there is a consistent bias towards greater development of a bilateral character on one side of the body over the other (Van Valen, 1962). Directional asymmetry typically results in a normal distribution of (R-L) differences about a mean that is significantly greater or less than zero (Palmer and Strobeck, 1986). Antisymmetry is defined by the presence of a consistent enlargement of one member of

a bilateral character pair where the side that is enlarged occurs in roughly equal frequency among individuals of the sample. The (R-L) differences for antisymmetry in bilateral characters typically result in a platykurtic or bimodal distribution about a mean of zero. Fluctuating asymmetry can be distinguished from these other types of asymmetry by its characteristic (R-L) normal distribution about a mean of zero. As noted by Van Valen (1962), any two or all three types of asymmetry can occur together in the same character. Both directional asymmetry and antisymmetry generally result from normal development (asymmetry being the norm) whereas FA occurs when symmetry is the norm, but normal development is perturbed (Leary and Allendorf, 1989).

As noted by Palmer and Strobeck (1986), FA is unique in that it is applied to situations in which the ideal condition of the character state is known: perfect bilateral symmetry. This quality gives value to FA as an indicator of developmental stability. It is important to keep in mind, however, that departures from bilateral symmetry must not be due to the action of genes that direct the asymmetry to develop if FA is to be useful (Palmer and Strobeck, 1992). The level of developmental instability is reflected by the level of FA only if the asymmetry is of a kind that is deviant from the developmental "plan" of the genome, and cannot be corrected by the regulatory mechanisms of the organism.

Even though the FA observed in an organism is not coded for by its genome, it is important to note that there does seem to be a heritable basis for both the likelihood that an individual departs from bilateral symmetry and the degree to which it departs from bilateral symmetry in either direction (Palmer and Strobeck, 1986). Both genetic and environmental factors thought to affect developmental stability have been studied. Palmer and Strobeck (1992) listed several patterns pertaining to developmental stability that have emerged from the various FA studies. These patterns include reduced developmental stability due to an increase in homozygosity, hybridization between nominal species, extreme physical conditions, and pollution or declining habitat quality.

Of particular interest to our study is the relationship of decreasing developmental stability among hybrids as compared to that of parental classes. In an early study, Soulé (1967) hypothesized that the input of foreign chromosomes into an organism's genome will disrupt coadapted gene complexes. This should cause a deterioration in developmental homeostasis and, thus, increase the level of FA. This hypothesis has given rise to many studies of FA within hybrid zones testing the predictions of Soulé. Studies of vertebrate hybrid zones carried out by Graham and Felley (1985) on sunfish (*Enneacanthus*) and by Leary et al. (1985) on trout (*Salvelinus* and *Salmo*) demonstrated higher FA levels in hybrids when compared to parental groups. However, studies by Felley (1980) on bluegill sunfish (*Lepomis*), Jackson (1973) on lizards (*Sceloporus*), and Lamb et al. (1990) on treefrogs (*Hyla*) did not find elevated FA levels in hybrids as compared to parental groups. A few studies of vertebrates suggest an advantage in hybrids with respect to developmental stability. Alibert et al. (1994) demonstrated increased developmental stability in hybrids between two subspecies of *Mus musculus* from a hybrid zone in Denmark. Studies attempting to discover the mediator of developmental stability have suggested both level of genetic heterozygosity (Soulé, 1979) and level of genomic coadaptation (Graham and Felley, 1985) as key factors. It seems possible that developmental stability may depend upon a balance between heterozygosity and coadaptation; hybrid populations exhibit both increased heterozygosity and disruption of coadapted gene complexes, and thus hybrids may exhibit higher or lower levels of developmental stability (Graham, 1992).

THE SCELOPORUS GRAMMICUS COMPLEX

One important model system used in vertebrate hybrid zone studies has been the *Sceloporus grammicus* complex. This complex of "iguanid" lizards (family Phrynosomatidae sensu Frost and Etheridge,

1989) ranges from southern Texas through most of mainland Mexico and consists of eight distinct chromosome races (cytotypes; 2N range = 32–46) that form seven parapatric areas of contact in central Mexico (Arévalo et al., 1991; Porter and Sites, 1986). Of particular interest in this complex is the Tulancingo hybrid zone of eastern Hidalgo, Mexico (Sites et al., 1993). This is a hybrid zone between the F5 (2N = 34) and FM2 (2N = 44–46) cytotypes which, on the basis of phylogenetic evidence, is believed to be a secondary contact (i.e., the races are recovered as nonsister taxa: Arévalo et al., 1994). Sites et al. (1995) demonstrated that this hybrid zone is approximately 1 km in width and is defined by three diagnostic chromosome markers that demonstrate sharp concordant changes from one parental race to the other.

Pure F5 individuals are characterized by a fixation (100% frequency) of biarmed elements for chromosome pairs 1, 2, and 6, and pure FM2 individuals are fixed for acrocentric elements for these same pairs. These rearrangements permit unambiguous identification of F_1-like (heterozygous for fissions at all three pairs) and several classes of backcross genotypes (Sites et al., 1995). In addition to chromosomal data, data for morphological, allozyme, mitochondrial, and ribosomal-DNA restriction sites have been collected for the F5 and FM2 races along the Tulancingo transect (Sites et al., 1993), and much progress has been made in understanding some factors contributing to the maintenance of this particular zone (Reed and Sites, 1995; Reed et al., 1995a,b; Sites et al., 1995, 1996).

The Tulancingo transect provides an opportunity to test Soulé's (1967) hypothesis in a well defined hybrid zone through the application of FA analysis. Over 400 preserved specimens with known genotypes have been mapped to precise locations on a microgeographic scale, and they are available for FA analysis (Sites et al., 1995). The karyotypic data permit unambiguous placement of all individuals into one of five classes of genotypes: parental F5, parental FM2, F_1-like hybrid, $F_1 \times$ F5

backcross, and $F_1 \times$ FM2 backcross. If hybridization decreases developmental stability within the Tulancingo zone, then "hybrid" individuals (F_1-like hybrid, $F_1 \times$ F5 backcross, and, or, $F_1 \times$ FM2 backcross) should experience decreased levels of developmental stability due to input of foreign chromosomes and, therefore, will show higher levels of FA when compared to those in the parental classes. Hybrid individuals are predicted to show higher levels of FA than parentals, because it is expected that the mixing of foreign genes from the two very distinct parental chromosome races should disrupt the coordination of gene complexes that were coadapted within the parentals. This in turn should decrease the developmental stability within the hybrids and result in higher FA levels.

MATERIALS AND METHODS

Previous karyotypic screening of lizards collected in the Tulancingo transect area (Sites et al., 1993) served as a guide for a more intense sampling effort and organizing the placement of sampling quadrats in the hybrid zone area. A total of 28 quadrats, each roughly 0.5 × 0.8 km in area, were placed along the Tulancingo transect (Highway 130: Fig. 1) in all areas suitable for sampling. Quadrats were set up on either side of the highway and were not set up in areas that had been cleared for pasture. A total of 24 quadrats were sampled within the hybrid zone as defined by Sites et al. (1995), and a total of four quadrats were sampled from outside the zone. Lizards were captured by noose, stunning with rubber bands, or grabbing by hand, and their precise locations were mapped as described by Sites et al. (1995). We assumed that lizards collected within a quadrat were a random sample of all the lizards in that quadrat. All specimens were preserved as vouchers, and chromosomes from each were analyzed to assign genotypes. For a more detailed description of sampling methods, genotype assignment, and voucher deposition see Sites et al. (1995).

Genotypically defined parental classes were each partitioned into two separate
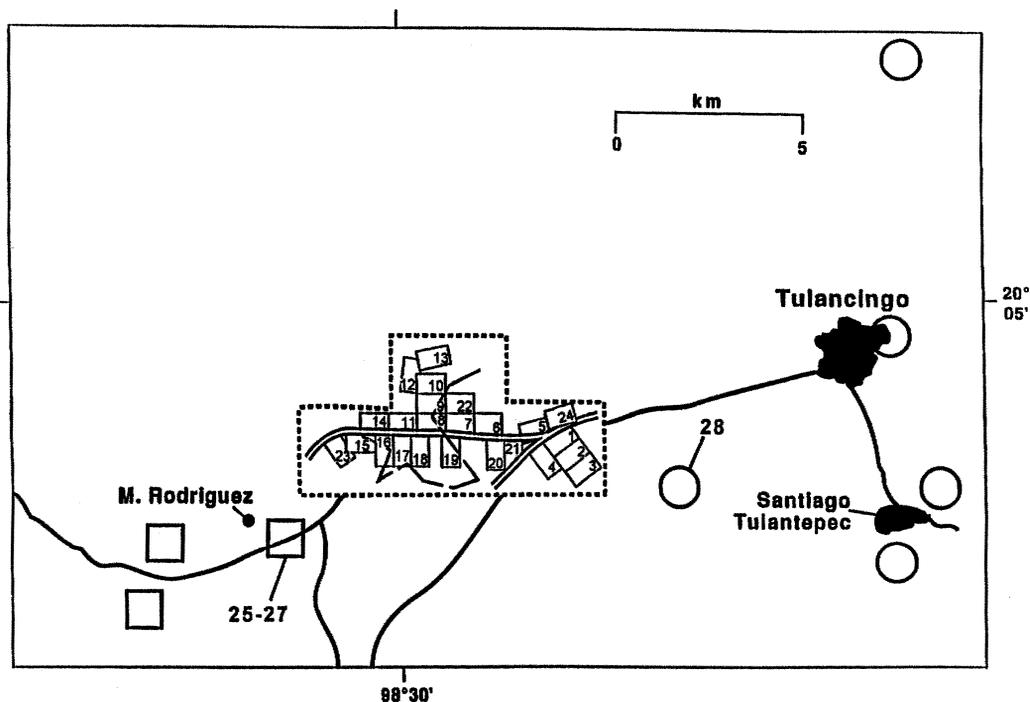
FIG. 1.—Map of Tulancingo transect in the Mexican state of Hidalgo. The transect is located along Highway 130, which is shown extending from M. Rodriguez to Tulancingo. Circles represent chromosomally pure populations of the F5 cytotype, while squares denote the same for FM2 (see Arévalo et al., 1991, for more detail on distributions). Quadrats 1–24 are in or adjacent to the hybrid zone—indicated here by the irregular line that summarizes the approximate null point (the region where alternative rearrangements have a frequency of about 0.50) of the clines for all three chromosome markers, while quads 25–27 and 28 show locations of pure parental genotypes outside of the zone.

classes: parentals from within the hybrid zone and parentals from outside the zone. We did this to test for differences between pure and potentially introgressed parental groups, because genotypes scored as "parental" on the basis of chromosome markers might nevertheless be introgressed with other genetic markers, especially in or close to the hybrid zone. Bilateral meristic characters were therefore counted on individuals of S. grammicus assigned a priori to one of the following classes: parental F5 inside ($F5_{in}$) and outside the hybrid zone ($F5_{out}$), parental FM2 inside ($FM2_{in}$) and outside the hybrid zone ($FM2_{out}$), F_1-like (F_1), $F_1 \times F5$ backcross ($F5_{bk}$), and $F_1 \times$ FM2 backcross ($FM2_{bk}$). Table 1 summarizes the numbers of lizards, by genotype and location, used in this study ($n = 427$).

The bilateral meristic characters counted on each individual and analyzed for FA were canthal scales, parietal border scales,

post-rostral scales, nasal border scales, circumorbital scales, pes toe lamellae of the fourth digit, and manus toe lamellae of the fourth digit. The first author made the counts using a Wild-Heerbrugg dissecting microscope, and he selected characters for their unambiguous and consistent expression. All lizards were originally scored for 15 characters. However, eight characters had to be eliminated a posteriori due to an unanticipated high level of character ambiguity. This character ambiguity was due to the difficulty in determining if certain scales belonged to the scale series of interest. The seven remaining characters (several measured in a method slightly modified from the original screening to lessen ambiguity) were again scored for all individuals. We assumed that error was reduced by the replicate scoring of most characters, the resulting increased familiarity with idiosyncrasies of the scale char-

TABLE 1.—Summary of sample sizes for *S. grammicus*, by quadrat and genotype, from which meristic data were collected for FA analysis. See text for descriptions of genotypic classes and Fig. 1 for location of sampling quadrats.

Quadrat	Genotypes in contact zone (quads 1–24)					Genotypes outside of zone	
	$FM2_{in}$	$FM2_{bk}$	F_1	$F5_{bk}$	$F5_{in}$	$FM2_{out}$	$F5_{out}$
1	0	0	0	2	5	0	0
2	0	0	0	3	12	0	0
3	0	0	0	0	6	0	0
4	0	0	0	0	9	0	0
5	0	0	1	10	11	0	0
6	0	0	0	6	15	0	0
7	0	0	0	0	1	0	0
8	4	6	4	10	1	0	0
9	9	6	0	1	0	0	0
10	14	6	1	0	0	0	0
11	16	14	1	0	0	0	0
12	8	2	0	1	0	0	0
13	11	4	0	0	0	0	0
14	10	5	0	1	0	0	0
15	0	2	1	7	10	0	0
16	3	2	1	3	2	0	0
17	11	6	1	0	1	0	0
18	6	7	2	0	0	0	0
19	4	0	1	0	0	0	0
20	0	1	1	0	0	0	0
21	2	0	0	4	26	0	0
22	0	3	4	6	9	0	0
23	0	0	0	12	7	0	0
24	0	0	0	4	16	0	0
25	0	0	0	0	0	13	0
26	0	0	0	0	0	12	0
27	0	0	0	0	0	11	0
28	0	0	0	0	0	0	10
Totals	98	64	18	70	131	36	10

acters of *S. grammicus*, and the inherent lower likelihood of measurement error when taking meristic counts rather than metric measurements. Genotypic class assignments were not made known until after all individual lizards were scored. The seven characters included in the study are defined below, and the five head scale characters are illustrated in Fig. 2.

Post-rostral scales (character 1 of Fig. 2A) are those immediately posterior to the rostral and that are located entirely medial to the border formed by the lateral-most edge of the rostral. If a post-rostral scale occupies equal space on both the right and left side of an individual, as determined by visualizing a medial line down the length of the individual, then the scale is recorded as present on both sides. Nasal border scales (character 2) are those encircling and making contact with the nasal scale, excluding the canthal and post-rostral scales. Canthal scales (character 3) are those along the canthus rostralis, located between the most anterior superciliary and the most posterior scales bordering the nasal. A scale is considered a canthal if there is both a dorsal portion and a lateral portion present and if it is evident that the canthal is not fused to the subnasal. Circumorbital scales (character 4) are those forming a semi-circle series above the eye, beginning with the scale contacting the most anterior and lateral superciliary and ending with the most posterior circumorbital to make contact with the parietal. If the most anterior circumorbital does not contact the most anterior superciliary, then the scale that joins these two is counted as a circumorbital also. Small granular scales in the series and scales that are located medial to and bridging together two other circumorbitals where they abut are also counted as circumorbital scales. Parietal border scales (character 5) are those forming a border posterior to the parietal scale, beginning with the scale that makes contact with both the parietal and interparietal, and ending with the scale that makes contact with both the parietal and the circumorbital scale series. Pes toe lamellae scales (not illustrated) are those on the fourth digit of the foot extending from the angle formed between the union of the third and fourth digit to the distal end of the toe. All scales entirely distal to this angle (toward the tip of the toe) are considered pes toe lamellae. Manus toe lamellae scales (not illustrated) are those on the fourth digit of the hand extending from the angle formed between the union of the third and fourth digit to the distal end of the toe. All scales entirely distal to this angle (toward the tip of the toe) are considered manus toe lamellae. We recorded as missing observations those characters that could not be counted by following these specific guidelines.

We examined data for character independence, size effects, antisymmetry, sex influence, and directional asymmetry before proceeding to the analysis of FA. We
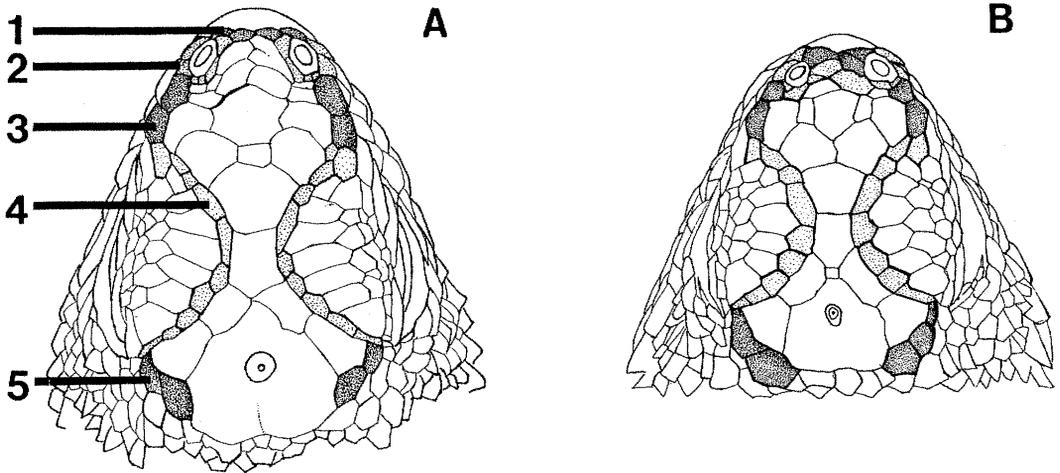
FIG. 2.—Bilateral meristic characters examined in this study, with number identifying the following series: 1 = post-rostrals, 2 = nasal border, 3 = canthals, 4 = circumorbitals, 5 = parietal border. Note that pes and manus toe lamellae are not illustrated. The individual in panel A (BYU 39810) is symmetrical for all five characters, while the one in panel B is asymmetrical for all characters except canthals (panel B drawing based on BYU 39942, with nasal border asymmetry added from another specimen for illustrative purposes). Methods of counting each of these characters are described in the text.

performed all statistical analyses using the SAS statistical software package (SAS, 1996).

RESULTS

Character Correlation and Size Effects

In general, correlation analysis revealed independence for the meristic scale characters. All pairwise correlation tests yielded correlation coefficients of $r \leq 0.21$ except for the correlation coefficient for manus toe lamellae with pes toe lamellae ($r = 0.63$). The correlation coefficient for manus and pes toe lamellae suggests that these two characters are not independent of each other, so conclusions drawn about the loci responsible for these characters must be made with caution. Size plots of (R-L) counts versus snout–vent lengths were generated for each of the seven characters for each of the seven genotypically defined class of individuals. These plots indicated that there were no significant relationships between body size and magnitude of (R-L) differences.

Character Normality and Kurtosis

Normality tests were used to assess antisymmetry and to consider the appropriateness of subsequent analyses. Antisym-

metry was difficult to assess. Results of Kolmogorov-Smirnov tests of normality and normal probability plots indicated that the distributions of (R-L) values were nonnormal. However, this rejection of normality resulted from the discrete nature of the data. We then analyzed the distributions of (R-L) differences for skewness and kurtosis to detect true antisymmetry more accurately. We calculated standard errors (SE) for skewness and kurtosis values and assessed the significance by determining how far away (in units of SE) each skewness or kurtosis value fell from the expected value for a normal distribution. In almost all cases, the differences did not exhibit significant skewness (Table 2). The most extreme skewness value fell 6.55 SE away from the expected skewness value for a normal distribution. We judged this and the other high skewness values, upon closer inspection, to be highly inflated. Distribution plots (data not shown) revealed a single observation causing the high skewness value. The value was inflated due to this single observation and the discrete nature of the data. For subsequent analyses, this skewness was not a problem. The central limit theorem generally holds for sample sizes >30 (Ott, 1993), and the sample

TABLE 2.—Presence or absence of antisymmetry and directional asymmetry. Antisymmetry results are based on normality tests (ranges of kurtosis values are given in parentheses) and directional asymmetry results are based on analysis of variance tests. F-test significance levels are given for directional asymmetry components. Specific chromosomal classes containing the asymmetry in each character are described in the text.

Character	Antisymmetry	Directional asymmetry
Canthal scales	Absent (2.98, 20.41)	Absent
Parietal border scales	Absent (−0.73, 2.36)	Present ($P = 0.028$)
Post-rostral scales	Absent (1.79, 12.74)	Absent
Nasal border scales	Absent (−0.31, 0.60)	Present ($P = 0.012$)
Circumorbital scales	Present (−0.91, 1.36)	Present ($P = 0.026$)
Pes toe lamellae	Absent (−0.27, 1.96)	Absent
Manus toe lamellae	Absent (−0.16, 4.87)	Present ($P = 0.005$)

size of this seemingly problematic distribution was $n = 69$. We used analysis of variance and F-tests in subsequent analyses, because these are not very sensitive to non-normality for large samples (Moore, 1995). Steel and Torrie (1980) noted that biological data rarely fit the normal distribution perfectly, but statistical tests performed on these distributions are usually valid. Therefore, we included all observations, including those responsible for high skewness values, in the study. We judged that kurtosis was acceptable in all but one case (circumorbital character for $FM2_{bk}$), in which a large negative kurtosis value and shape of the distribution plot indicated bimodality, with the kurtosis value falling 1.46 SE from that expected in a normal distribution (Table 2). The bimodal distribution necessitates caution in drawing any inferences from this particular character class, because the presence of antisymmetry is indicated.

Tests for Asymmetry

For each of the seven classes of lizards, a two-way mixed linear model was fitted, with sex, sides (right and left), and the sides-by-sex interaction as fixed effects, and individuals as random effects. The effects of sex, sides, and the sides-by-sex interaction were estimated, as were variance components due to individuals and the combination of random error and possible non-directional asymmetry (FA and antisymmetry). The sides factor represents variation due to directional asymmetry. The individuals factor represents the random variation due to differences among individuals. A significant sides-by-sex in-

teraction would indicate that the sex factor was responsible for some of the variation otherwise assigned to the sides factor. We found this sides-by-sex interaction to be significant for only one of the 49 tests (canthal scales for class $FM2_m$, $P = 0.002$), and thus we assumed it to represent a type I error rather than a real sex factor influence. Therefore, we removed the sex and sides-by-sex interaction factors from the model and estimated and analyzed the remaining factors.

We used Levene's test, as recommended by Palmer and Strobeck (1992), to evaluate differences in FA levels among classes for each particular character. Levene's test is a one-way analysis of variance on the absolute value of (R-L) differences among the different classes for a particular character. If no directional asymmetry or antisymmetry is present, the absolute value of the (R-L) difference for a character is equivalent to the absolute value of the (R-L) difference due only to developmental noise and any error present. Thus, Levene's test should compare FA levels if no error is present.

We found significant directional asymmetry in four character classes: parietal border scales for class $FM2_{bk}$, nasal border scales for class $F5_m$, circumorbital scales for class F_1, and manus toe lamellae for class $F5_m$ (Table 2, Figs. 3, 4). We then performed a multivariate analysis of variance test to detect any directional asymmetry that may have gone undetected by univariate methods. There was no indication of any directional asymmetry missed by the univariate methods. Therefore, all statistical comparisons of FA levels using
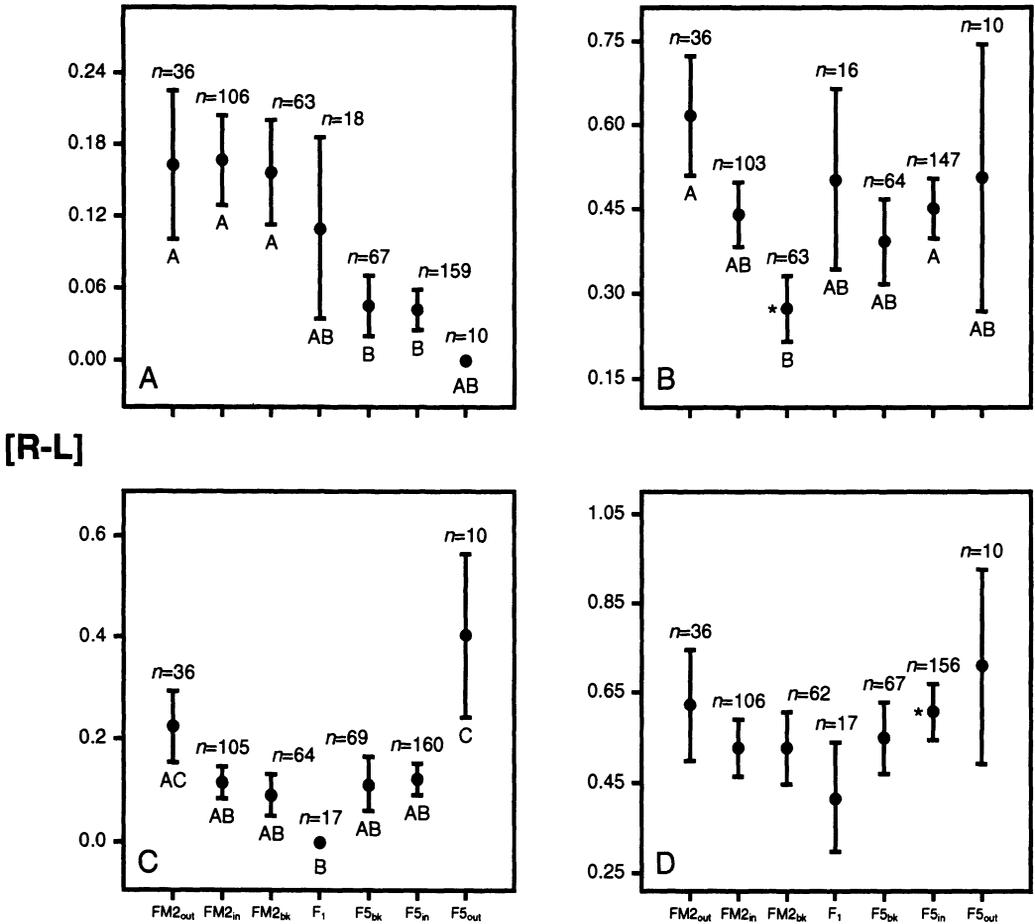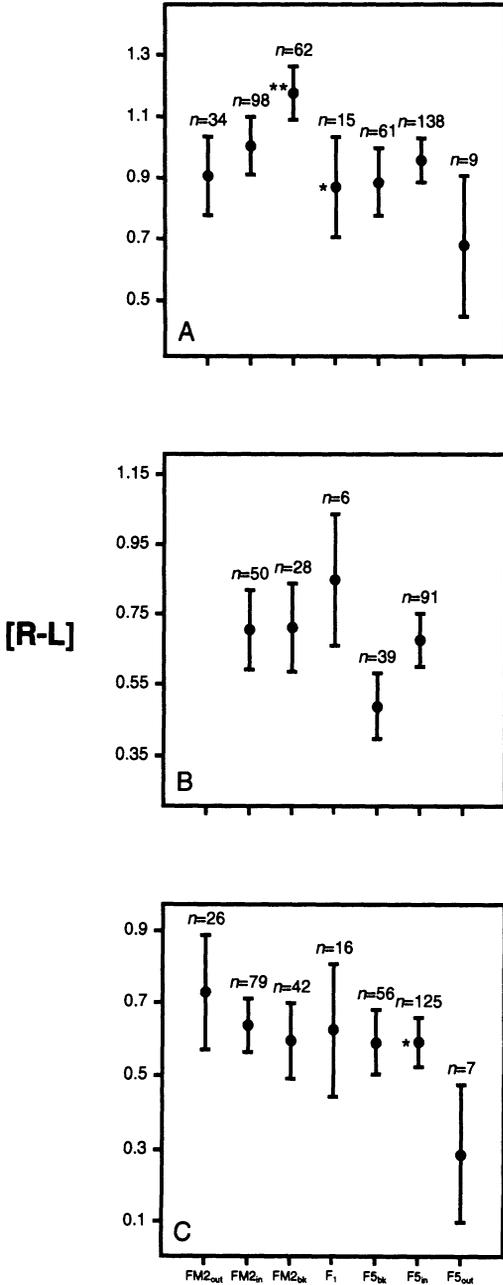
CHROMOSOMAL CLASS

FIG. 3.—Means (solid circles) (±1 SE, vertical lines) of absolute value for right minus left counts (|R-L|) for selected bilateral meristic characters, by genotypic class. Note: for A, B, and C, means that do not share letters are significantly different ($\alpha \leq 0.05$) for Levene's test. There are no means significantly different for D. The single asterisk denotes a significant directional asymmetry component (panels B, D). (A) Canthal scales, (B) parietal border scales, (C) post-rostral scales, (D) nasal border scales.

Levene's test were judged to be valid except those comparisons involving these particular character genotype classes containing directional asymmetry (marked by an asterisk in Figs. 3B,D and 4A,C).

Table 3 summarizes the means of |R-L| counts for each character in each genotype class, and these are graphically displayed in Figs. 3 and 4. Hybrids demonstrated markedly lower levels of FA than both parentals for post-rostral scales, and lower or equivalent levels of FA for the parietal

border scales. One parental class ($F5_{in}$) demonstrated lower levels of FA than hybrids for canthal scales (Fig. 3, Table 3). Of the tests demonstrating statistically significant differences in level of FA, only the two comparisons demonstrating lower FA levels in hybrids for the parietal border scale character involved a directional asymmetry component. Directional asymmetry indicates the presence of a genetic asymmetry influence, and thus (R-L) differences are not an indication of develop-

FIG. 4.—Means (solid circles) (±1 SE, vertical lines) of absolute value for right minus left counts (|R-L|) for selected bilateral meristic characters, by genotypic class. There are no significant differences for Levene's test ($\alpha \leq 0.05$) among means for A, B, or C. The single asterisk denotes a significant directional asymmetry component (panels A, C), and the double asterisk denotes antisymmetry component (panel A,

mental instability for this particular character class. However, because Levene's test involves comparisons in magnitude of (R-L) differences, if the genetic directional asymmetry component were to be removed, there should be an even lower absolute value of (R-L) in the hybrid class for this particular character. No significant differences were demonstrated for the circumorbital scale character, which indicated the presence of antisymmetry in kurtosis and distribution tests.

Sample Sizes and Statistical Power

Because some of the characters did not have statistically significant differences between |R-L| means, but patterns did seem to be evident, we performed power analyses to assess the possible influence that small sample size may have had on the outcome of comparisons between means (Table 4).

Power analysis tests were conducted based on a significance level of $\alpha = 0.05$, and they are based on the assumption that there are equal sample sizes for the two means being compared. These requirements were not always met, so the harmonic mean of the sample sizes, $\tilde{n}$, was used as the sample size for power analysis. For each character, there were seven treatments (genotype classes), and the pooled SE was used as σ, the estimate of the standard deviation of the residual error. The largest observed difference between two |R-L| means for a given character (expressed as a percentage of the overall |R-L| mean, μ) was calculated, and denoted as "δ%obs" in Table 4. Possible values for δ%, the smallest difference of interest between two |R-L| means for a given character (expressed as a percentage of the overall |R-L| mean), were generated along with the corresponding levels of power, 1 − β. This allows the interpretation of the power level for *F*-tests between |R-L|

←

see text). (A) Circumorbital scales, (B) pes lamellae scales. Note: Means for FM2$_{out}$ and F5$_{out}$ are not shown due to lack of observations. (C) Manus lamellae scales.

TABLE 3.—Means of |R-L| counts for each character and genotype. Standard errors are shown in parentheses and ranges below, and an asterisk denotes missing observation.

Character	Genotypes						
	$FM2_{out}$	$FM2_{in}$	$FM2_{bk}$	F_1	$F5_{bk}$	$F5_{in}$	$F5_{out}$
Canthal	0.167 (±0.063) 0–1	0.170 (±0.037) 0–1	0.159 (±0.046) 0–1	0.111 (±0.076) 0–1	0.045 (±0.026) 0–1	0.044 (±0.016) 0–1	0.000 (±0.000) 0
Parietal border	0.611 (±0.107) 0–2	0.437 (±0.058) 0–3	0.270 (±0.056) 0–1	0.500 (±0.158) 0–2	0.391 (±0.073) 0–2	0.442 (±0.050) 0–3	0.500 (±0.167) 0–1
Post-rostral	0.222 (±0.070) 0–1	0.114 (±0.031) 0–1	0.094 (±0.037) 0–1	0.000 (±0.000) 0	0.116 (±0.044) 0–2	0.125 (±0.028) 0–2	0.400 (±0.163) 0–1
Nasal border	0.611 (±0.121) 0–2	0.519 (±0.059) 0–2	0.516 (±0.079) 0–2	0.412 (±0.123) 0–1	0.537 (±0.078) 0–2	0.596 (±0.054) 0–3	0.700 (±0.213) 0–2
Circumorbital	0.912 (±0.129) 0–3	1.010 (±0.093) 0–5	1.177 (±0.091) 0–3	0.867 (±0.165) 0–2	0.886 (±0.110) 0–4	0.957 (±0.076) 0–4	0.667 (±0.236) 0–2
Pes lamellae	*	0.709 (±0.112) 0–4	0.714 (±0.124) 0–2	0.833 (±0.167) 0–1	0.488 (±0.093) 0–2	0.674 (±0.074) 0–3	*
Manus lamellae	0.731 (±0.162) 0–4	0.633 (±0.070) 0–2	0.595 (±0.108) 0–2	0.625 (±0.180) 0–2	0.589 (±0.087) 0–2	0.592 (±0.064) 0–3	0.286 (±0.184) 0–1

means if one accepts that the largest observed difference between two means is biologically meaningful. Caution must be taken here, however, because the estimated power level is entirely dependent upon the magnitude of difference that is desirable to detect. With the canthal scale character, for example, if one desires to detect a difference between two means that is equal to 100% of the overall |R-L| mean, the power level is >0.80 for our sample size and confidence in the F-test is warranted. On the other hand, if as small a difference as 45% of the overall mean is considered biologically meaningful, the power to detect statistical significance with the available sample ($\tilde{n} = 30$) is only 0.20. The δ%obs value is thus subjective as used to interpret the level of power, but we use it conservatively (by insisting on large differences only) in these comparisons.

Assuming that a difference equal to 100% of the overall mean is desirable to detect, then F-tests for pes and manus lamellae have power levels of 0.60 and 0.70, respectively (Table 4). The other five characters have power levels above 0.80. These reasonably large power levels suggest that had there been large, statistically significant differences among means for nasal, circumorbital, pes, and manus scale characters, our sample sizes were adequate to detect them at least 80% of the time.

DISCUSSION

The prediction that hybrid classes (including backcrosses) should show higher levels of FA than parental classes was not supported in tests for three of the seven characters included in the study (canthal scales, parietal border scales, and post-rostral scales), and the remaining four characters showed no statistically significant differences between hybrid and parental FA levels. Hybrids demonstrated significantly lower FA levels than parental classes for a total of eight tests. Parentals demonstrated significantly lower levels of FA for just one test (between $F5_{in}$ and $FM2_{bk}$) for the canthal scale character. Thus, in general, these results do not confirm Soulé's hypothesis that hybrids should dem-

TABLE 4.—Power analysis results based on $\alpha = 0.05$. The harmonic mean of the sample size is denoted by $\tilde{n}$; the largest observed difference between two |R-L| means for a given character (expressed as a percentage of the overall |R-L| mean) is denoted by $\delta\%$ obs; values for the smallest difference of interest between two |R-L| means for a given character (expressed as a percentage of the overall |R-L| mean) are denoted by $\delta\%$; the level of power is denoted by $1 - \beta$. See text for discussion.

$\tilde{n}$ $\delta\%$ obs	Canthal 30 1.70	Parietal 29 0.80	Post-rostral 30 3.17	Nasal 30 0.52	Circumorbital 27 0.54	Pes 19 0.52	Manus 23 0.72
$1 - \beta$	$\delta\%$	$\delta\%$	$\delta\%$	$\delta\%$	$\delta\%$	$\delta\%$	$\delta\%$
0.80	0.97	0.99	0.97	0.97	1.03	1.23	1.12
0.70	0.88	0.90	0.88	0.88	0.93	1.12	1.01
0.60	0.80	0.81	0.80	0.80	0.84	1.01	0.92
0.50	0.72	0.74	0.72	0.72	0.76	0.92	0.83
0.40	0.64	0.66	0.64	0.64	0.68	0.82	0.74
0.30	0.56	0.57	0.56	0.56	0.59	0.71	0.64
0.20	0.45	0.46	0.45	0.45	0.48	0.57	0.52
0.10	0.28	0.29	0.28	0.28	0.30	0.36	0.33

onstrate lower levels of developmental stability and, as a result, higher levels of FA.

Of interest to us are three models to explain the maintenance of stable hybrid zones in nature. One model, known as the "dynamic equilibrium" (tension zone) model, proposes that hybrid zones are maintained by a balance between the influx of parental alleles and inherent inferiority of hybrids due to disruption of coadapted gene complexes. This model infers that the zone is maintained primarily by "endogenous" selection acting against hybrids; the inherent genetic nature of the hybrid is selected against, irrespective of the environment (Freeman et al., 1995). The alternative models, the "bounded hybrid superiority" (BHS) and "mosaic" models, propose that the hybrid zones are maintained by "exogenous" selection along environmental gradients. The BHS model proposes that hybrids are selectively favored over parentals in the ecotonal area that they occupy, whereas the "mosaic" model assumes that parental genotypes are selectively favored in their own environments, relative to those on the opposite side of the hybrid zone (these models are reviewed in detail by Arnold, 1997).

Freeman et al. (1995) conducted a study of developmental stability in a hybrid zone between two subspecies of sagebrush (*Artemisia tridentata*). They assessed levels of developmental stability using both morphological and biochemical data and found that hybrids were less developmentally stable than parentals in only two of 28 comparisons (one morphological and one biochemical trait), and were actually more developmentally stable for two of the comparisons. They concluded that the results contradict the predictions of the dynamic equilibrium model and, thus, the sagebrush hybrid zone does not appear to be maintained by endogenous selection. Paralleling this sagebrush hybrid zone study are the findings of Alibert et al. (1994), who examined FA of molar dentition in a hybrid zone of *Mus musculus* in Europe. They found that hybrid mice formed between two different subspecies had higher levels of developmental stability than parentals. Thus, their results also seem to indicate that exogenous selection, not endogenous selection, was operating to maintain the hybrid zone. The higher levels of developmental stability in the hybrid lizards of the present study appears consistent with the bounded hybrid superiority model; the hybrid genotypes (including backcrosses) may be selectively favored within an ecotonal area. This interpretation assumes that levels of FA for characters used in this study are inversely related to fitness of the genotypic classes (see Clarke, 1995, for a review of studies establishing a relationship between developmental stability and fitness).

The Tulancingo hybrid zone is positioned along a transition from oak-pine forest to a drier habitat characterized by *Agave, Opuntia,* and other xeric-adapted

plants, although agricultural activity has modified the original vegetation. Sites et al. (1995) demonstrated a statistically significant association between F5 genotypes and oak habitats (and conversely between FM2 genotypes and the xeric habitat), and they suggested that exogenous selection was likely operating on this zone. However, studies of fitness correlates of hybrid genotypes revealed that in males, F_1 genotypes manifested a high degree meiotic malassortment for chromosome pair 2, which potentially could reduce the fitness of these individuals (Reed et al., 1995b). A more direct estimate of fitness was the demonstration of significantly reduced litter sizes in females with F_1 genotypes (Reed and Sites, 1995), suggesting that endogenous selection was also responsible for the maintenance of the observed genetic structure of this hybrid zone. For these reasons, the patterns of FA described here are not adequately explained by the bounded superiority model of hybrid zone structure.

Arnold (1997) reviewed hybrid zone studies of both plants and animals and concluded that empirical studies of putative tension zones did not consistently demonstrate patterns expected for this model. Instead, he proposed a new model for the maintenance of hybrid zones. This model incorporates endogenous selection against certain hybrid genotypes, exogenous selection for or against different hybrid genotypes, rarity of F_1 genotype formation, and the invasion of parental or novel habitats by a subset of hybrid individuals more fit than the parentals. It seems possible that such a model may apply to the Tulancingo hybrid zone, which provides evidence for both environment-dependent and environment-independent selection.

One other observation is relevant here. Reed et al. (1995a) also demonstrated that normal meiotic segregation prevailed in males in all combinations of backcross genotypes studied, in both directions (back to F5 and to FM2 parental genotypes). Similarly, females with backcross genotypes had litter sizes similar to those of parental genotypes of the same body size

(Reed and Sites, 1995), and these observations collectively suggest that, if F_1 offspring mature and reproduce, there is opportunity for extensive recombination in backcross generations on both sides of the zone, and this is expected to break up the association among all but the most tightly linked characters.

In a hybrid zone study between two subspecies of grasshopper (*Chorthippus parallelus*), Virdee and Hewitt (1994) examined patterns of testis dysfunction and suggested that one possible outcome of the disruption of linkage among loci responsible for dysfunction would be the formation of independent and non-concordant clines for different characters. Similar mechanisms operating in the Tulancingo transect could explain the different patterns of developmental stability shown for the meristic characters used in this study, and these patterns would sharply contrast with the three concordant clines for the chromosomal markers used to define this zone. In other words, the recombination generated by backcrossing in the Tulancingo hybrid zone is expected to permit a dissociation among unlinked loci (or sets of loci), and developmental stability clines for different characters will move apart and eventually reach different equilibrium positions. The patterns evident in Figs. 3 and 4 suggest that the loci (or linkage groups) controlling the expression of these characters are (1) not tightly linked to each other and (2) have been shuffled independently within the zone. If the second interpretation is correct, the patterns of developmental stability expressed in the characters used in this study either reflect (1) independent responses to a balance between selection and developmental pressures on each character or (2) if these traits are approximately neutral with respect to fitness, then the patterns reflect independent responses to gene flow and drift.

As one example, Fig. 3A (canthal scales) shows a pattern of increasing FA levels as one moves from individuals of the $F5_{out}$ class to individuals of the $FM2_{out}$ class. One explanation for this pattern may be a superior F5 "allele" that, with respect to

the FM2 allele, confers a higher level of developmental stability for canthal scales in the F5 and F5-backcross classes. The observation that this effect has not introgressed toward the FM2 side of the zone suggests that it may either be selected against on this side (for reasons that outweigh its developmental advantages) or it is tightly linked to another locus that is selected against.

As a second example, Fig. 3C (post-rostral scales) shows a clear decrease in FA levels as the individuals become more F_1-like. It seems that this general type of pattern would be in agreement with that expected from an ecotonal hybrid zone. The higher level of developmental stability in hybrids may be due to an advantage of hybridity in the ecotone, whereas outside the zone hybrids may be strongly selected against. The increased FA levels in the parentals supports this hypothesis, as does the narrow nature of the Tulancingo hybrid zone (approximately 1 km in width), which requires strong selection against hybrids outside of the zone.

As Felley (1980) noted, one possible explanation for the finding that hybrids do not demonstrate higher levels of FA than parentals in some characters may be that the two parental races diverged so slightly in the past that, upon secondary contact, hybrids never suffered from significant levels of genomic disruption. Results for other characters seem to contradict this possibility, however, because genetic divergence between the FM2 and F5 parental races is suggested by the statistically significant difference in level of developmental stability (FA level) for the canthal scale character (Fig. 3A). For this character, the $F5_{in}$ mean is significantly different from both the $FM2_{out}$ and $FM2_{in}$ classes at $\alpha = 0.05$. Alternatively, the two parental races may have been in secondary contact long enough to have allowed selection for gene complexes that successfully work together to produce developmentally stable individuals. It would not be surprising to find, for some characters at least, that negative effects of coadapted gene complex disruption in hybrids have been corrected over time by natural selection.

In summary, the absence of pattern in FA levels for certain characters may reflect selective "adjustment" to maintain normal bilateral development across all chromosomal classes for those characters. The presence of significant differences in FA levels for other characters implies that independent selective forces are operating on unlinked "loci" responsible for the bilateral expression of these traits. Further, the selective forces acting on some characters showing significant mean differences (Fig. 3B,C) seem to be independent of those maintaining the chromosome marker clines defining the hybrid zone, while those operating on the canthal scale character (Fig. 3A) may indeed have an endogenous component (i.e., operating against F_1-like genotypes). What remains uncertain is the relationship between FA in the meristic characters and reproductive performance among different genotypes.

LITERATURE CITED

ALIBERT, P. S. R., B. DOD, F. BONHOMME, AND J. AUFFRAY. 1994. Fluctuating asymmetry in the *Mus musculus* hybrid zone: a heterotic effect in disrupted co-adapted genomes. Proceedings of the Royal Society of London 258:53–59.

ARÉVALO, E., S. K. DAVIS, AND J. W. SITES, JR. 1994. Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosome races of the *Sceloporus grammicus* complex (Phrynosomatidae) in central Mexico. Systematic Biology 43:387–418.

ARÉVALO, E., C. A. PORTER, A. GONZÁLEZ, F. MENDOZA, J. L. CAMARILLO, AND J. W. SITES, JR. 1991. Population cytogenetics and evolution of the *Sceloporus grammicus* complex (Iguanidae) in central Mexico. Herpetological Monographs 5:79–115.

ARNOLD, M. L. 1997. Natural Hybridization and Evolution. Oxford University Press, New York, New York, U.S.A.

CLARKE, G. M. 1993. The genetic basis of develop-

mental stability. I. Relationships between stability, heterozygosity and genomic coadaptation. Genetica 89:15–23.

———. 1995. Relationships between developmental stability and fitness: application for conservation biology. Conservation Biology 9:18–29.

FELLEY, J. 1980. Analysis of morphology and asymmetry in bluegill sunfish (*Lepomis macrochirus*) in the southeastern United States. Copeia 1980:18–29.

FREEMAN, D. C., J. H. GRAHAM, D. W. BYRD, E. D. MCARTHUR, AND W. A. TURNER. 1995. Narrow hybrid zone between two subspecies of big sagebrush, *Artemisia tridentata* (Asteraceae). III. Developmental instability. American Journal of Botany 82: 1144–1152.

FROST, D. R., AND R. ETHERIDGE. 1989. A phylogenetic analysis and taxonomy of iguanian lizards (Reptilia: Squamata). University of Kansas Museum of Natural History Miscellaneous Publications 81:1–65.

GRAHAM, J. H. 1992. Genomic coadaptation and developmental stability in hybrid zones. Zoologica Fennica 191:121–131.

GRAHAM, J. H., AND J. D. FELLEY. 1985. Genomic coadaptation and developmental stability within introgressed populations of *Enneacanthus gloriosus* and *E. obesus* (Pisces, Centrarchidae). Evolution 39:104–114.

JACKSON, J. F. 1973. A search for the population asymmetry parameter. Systematic Zoology 22:166–170.

LAMB, T., J. M. NOVAK, AND D. L. MAHONEY. 1990. Morphological asymmetry and interspecific hybridization: a case study using hylid frogs. Journal of Evolutionary Biology 3:295–309.

LEARY, R. F., AND F. W. ALLENDORF. 1989. Fluctuating asymmetry as an indicator of stress: implications for conservation biology. Trends in Ecology and Evolution 4:214–217.

LEARY, R. F., F. W. ALLENDORF, AND K. L. KNUDSEN. 1985. Developmental instability and high meristic counts in interspecific hybrids of salmonid fishes. Evolution 39:1318–1326.

———. 1992. Genetic, environmental, and developmental causes of meristic variation in rainbow trout. Zoologica Fennica 191:79–95.

MATHER, K. 1953. Genetical control of stability in development. Heredity 7:297–336.

MOORE, D. S. 1995. The Basic Practice of Statistics. W. H. Freeman, New York, New York, U.S.A.

OTT, R. L. 1993. An Introduction to Statistical Methods and Data Analysis. Duxbury Press, Belmont, California, U.S.A.

PALMER, A. R., AND C. STROBECK. 1986. Fluctuating asymmetry: measurement, analysis, patterns. Annual Review of Ecology and Systematics 17:391–421.

PALMER, A. R., AND C. STROBECK. 1992. Fluctuating asymmetry as a measure of developmental stability: implications of non-normal distributions and power of statistical tests. Zoologica Fennica 191:57–72.

PORTER, C. A., AND J. W. SITES, JR. 1986. Evolution of *Sceloporus grammicus* complex (Sauria: Iguanidae) in central Mexico: population cytogenetics. Systematic Zoology 35:334–358.

REED, K. M., I. F. GREENBAUM, AND J. W. SITES, JR. 1995*a*. Cytogenetic analysis of chromosomal intermediates from a hybrid zone between two chromosome races of the *Sceloporus grammicus* complex (Sauria, Phrynosomatidae). Evolution 49:37–47.

———. 1995*b*. Dynamics of a novel chromosomal polymorphism within a hybrid zone between two chromosome races of the *Sceloporus grammicus* complex (Sauria, Phrynosomatidae). Evolution 49: 48–60.

REED, K. M., AND J. W. SITES, JR. 1995. Female fecundity in a hybrid zone between two chromosome races of the *Sceloporus grammicus* complex (Sauria, Phrynosomatidae). Evolution 49:61–69.

SAS INSTITUTE INC. 1996. SAS Proprietary Software Release 6.11. SAS Institute Inc., Cary, North Carolina, U.S.A.

SITES, J. W., JR., N. H. BARTON, AND K. M. REED. 1995. The genetic structure of a hybrid zone between two chromosome races of the *Sceloporus grammicus* complex (Sauria, Phrynosomatidae) in central Mexico. Evolution 49:9–36.

SITES, J. W., JR., C. BASTEN, AND M. A. ASMUSSEN. 1996. Cytonuclear genetic structure of a hybrid zone in lizards of the *Sceloporus grammicus* complex (Sauria, Phrynosomatidae). Molecular Ecology 5:379–392.

SITES, J. W., JR., S. K. DAVIS, D. W. HUTCHISON, B. A. MAURER, AND G. LARA. 1993. Parapatric hybridization between chromosome races of the *Sceloporus grammicus* complex (Phrynosomatidae): structure of the Tulancingo transect. Copeia 1993: 373–398.

SOULÉ, M. 1967. Phenetics of natural populations. II. Asymmetry and evolution in a lizard. American Naturalist 101:141–160.

SOULÉ, M. 1979. Heterozygosity and developmental stability: another look. Evolution 33:396–401.

STEEL, R. G. D., AND J. H. TORRIE. 1980. Principles and Procedures of Statistics: A Biometrical Approach, 2nd ed. McGraw-Hill, New York, New York, U.S.A.

VAN VALEN, L. 1962. A study of fluctuating asymmetry. Evolution 16:125–142.

VIRDEE, S. R., AND G. M. HEWITT. 1994. Clines for hybrid dysfunction in a grasshopper hybrid zone. Evolution 48:392–407.