

Article

Detecting trends in body size: empirical and statistical requirements for intraspecific analyses

C. M. GIENGER^{a,*}, Ned A. DOCHTERMANN^b, and C. Richard TRACY^c

^aDepartment of Biology, Center of Excellence for Field Biology, Austin Peay State University, Clarksville, TN, USA,

^bDepartment of Biological Sciences, North Dakota State University, Fargo, ND, USA, and ^cDepartment of Biology, Program in Ecology, Evolution, and Conservation Biology, University of Nevada, Reno, NV, USA

*Address correspondence to C. M. Gienger. E-mail: giengerc@apsu.edu.

Handling editor: Yiming Li

Received on 2 August 2018; accepted on 9 October 2018

Abstract

Attributing biological explanations to observed ecogeographical and ecological patterns require eliminating potential statistical and sampling artifacts as alternative explanations of the observed patterns. Here, we assess the role of sample size, statistical power, and geographic inclusivity on the general validity and statistical significance of relationships between body size and latitude for 3 well-studied species of turtles. We extend those analyses to emphasize the importance of using statistically robust data in determining macroecological patterns. We examined intraspecific trends in body size with latitude in *Chelydra serpentina*, *Chrysemys picta*, and *Trachemys scripta* using Pearson's correlations, diagnostic tests for influential points, and resampling. Existing data were insufficient to ascertain a latitudinal trend in body size for *C. serpentina* or *T. scripta*. There was a significant relationship for *C. picta*, however, resampling analyses show that, on average, 16 of the 23 available independent populations were needed to demonstrate a significant relationship and that at least 20 of 23 populations were required to obtain a statistically powerful correlation between body size and latitude. Furthermore, restricting the latitudes of populations resampled shows that body size trends of *C. picta* were largely due to leveraging effects of populations at the edge of the species range. Our results suggest that broad inferences regarding ecological trends in body size should be made with caution until underlying (intraspecific) patterns in body size can be statistically and conclusively demonstrated.

Key words: Bergmann's rule, geographic variation, macroecology, resampling, sensitivity analysis, statistical power, turtle

The study of geographical patterns in body size (Brown 1999; Trussell 2000; Gilchrist et al. 2004), size dimorphism (Blanckenhorn et al. 2006), life history traits (Arnett and Gotelli 1999; Blanckenhorn and Demont 2004), species diversity (Brown and Maurer 1989) and mechanisms associated with these patterns serve as foundations to the physiological, ecological, and evolutionary paradigms explaining patterns in variation in body size in nature. However, our understanding of macroecological patterns, which are inherently interspecific,

depends on the validity of component intraspecific trends, which are part of macro-scale analyses.

Intraspecific trends in body size are difficult to discern because each population represents only a single data point. Thus, adequate sample sizes to reveal trends in size across populations are difficult to obtain, and statistics calculated from numerically small samples tend to have large sample variances (Zar 1999), are often biologically uninformative or have unstable parameter estimates (Simpson

et al. 1992), and low statistical power to detect trends (Thomas and Juanes 1996). In addition, with small samples, influential outliers can leverage regression trend lines and correlation analyses (Cook 1979; Gerrodette 1987; Belsley et al. 2004), and obscure actual biological relationships. The leveraging effect of geographic outliers (points that are geographically separated by large distances from other data points in analyses) on body size trends has been shown to be important empirically (Litzgus et al. 2004), and as a result, the geographic extent of population sampling has the potential to influence observed macroecological patterns.

One important macro-ecological pattern is Bergmann's rule; the observation that larger body sizes are often found in populations in cooler climates (Bergmann 1847). This rule was logically derived from mechanisms for retaining body heat in endothermic animals (James 1991; Watt et al. 2010). Similar patterns of body trends have been purported to occur widely among vertebrates (Ashton and Feldman 2003) and invertebrates (Van Voorhies 1996; Atkinson and Sibly 1997) and is potentially deeply rooted phylogenetically (de Queiroz and Ashton 2004).

Here, we report sensitivity analyses of intraspecific latitudinal patterns of body size. We evaluate the statistical requirements for demonstrating body size trends for 3 turtle species, discuss the shortcomings of some conventional statistical approaches, and provide an alternative diagnostic tool based upon statistical resampling. While we focus on the relationship between latitude and body size (which is often considered an indirect test of Bergmann's Rule), resampling can be easily extended to other environmental and geographic gradients (e.g., temperature, elevation, salinity, and pH) that function as useful predictors of, or explanations for, variation in phenotypic traits.

Materials and Methods

Intra-specific body size trends

We assembled data on body size for 3 species of semi-aquatic North American turtles, the Snapping Turtle *Chelydra serpentina* Linnaeus, the Pond Slider *Trachemys scripta* Schoepff, and the Painted Turtle *Chrysemys picta* Schneider. These 3 species are among the most abundant, and the most studied aquatic turtles of North America, and each has a broad geographic range extending over at least 20–30 degrees in latitude (Ernst and Barbour 1989; Ernst et al. 1994). Data were gathered from published reviews and reports (Iverson et al. 1997; Lindeman 1997; Tucker et al. 1998; Ashton and Feldman 2003; Cooley et al. 2003; Aresco 2004; Emer 2004) and verified against their original data sources to avoid duplication. Body size is reported as carapace length for *C. serpentina* and as plastron length for *C. picta* and *T. scripta*.

We restricted the analyses to the North American populations of these species above 27.5° north latitude. Both *C. serpentina* and *T. scripta* have additional and limited ranges in southern México and Central America (Ernst and Barbour 1989), but we chose not to include 4 populations of these species because those populations were geographically separated from the nearest adjacent population by more than 9° latitude and phylogenetic analyses suggest that those populations likely represent distinct taxa (Seidel et al. 1999; Starkey et al. 2003).

We restricted analyses to sites for which there were data on at least 10 individuals per population (following Ashton and Feldman 2003). To avoid obfuscating differences among populations due to sexual size dimorphism, we included only females in analyses (Gibbons and Lovich 1990; Lovich and Gibbons 1992). For the

well-studied Painted Turtle, *C. picta*, we restricted studies to those with at least 30 female individuals per population. It is at this sample size (where $df = 29$) that Student's t distribution closely approximates the standard normal z distribution (Student 1908) and the sample can be considered statistically large.

For each of the 3 turtle species, we calculated Pearson's correlation coefficient, r , for the relationship between mean body size of the individuals at each site and latitude of those sites, along with the corresponding probability value (P) at $\alpha = 0.05$. We diagnosed potentially influential outlier populations in each correlation by computing 3 measures commonly used in regression analyses to assess sample leverage and influence; DFFITS, the influence of i -th observation on the fitted value of \hat{Y}_i ; DFBETAS, the standardized difference for each individual coefficient estimate resulting from the exclusion of the i -th observation; Cook's D, an overall measure of the combined impact of the i -th observation on all of the estimated regression coefficients (Neter et al. 1990).

Sample size and power

To investigate the effect of sample size (number of sample populations) on body size trends, we used MATLAB (Mathworks, Natick, MA, USA) to iteratively resample (Crowley 1992; Roff 2006) the latitude-body size data. For each species, we started with 10,000 random draws (with replacement) of body sizes from 3 sample populations from the universe of populations, and calculated the correlation coefficients and associated probability values of each draw. From this resampling, we took the mean correlation coefficient for body size and latitude, the statistical significance of the mean value of r , by comparing it to the critical value (r_{crit}), and the number of statistically significant draws among the 10,000 total draws. We determined the statistical power ($1-\beta$) to detect a significant trend by counting the number statistically significant draws. Sufficient power was achieved when 80% of the 10,000 random draws were significant (Cohen 1988). We then increased the number of sample populations to be drawn by 1, and then repeated the process until all populations were analyzed the same way.

Geographic extent and outlier effects

To investigate the effect of the geographic extent of populations used in analyses, we randomly subsampled the data, but manipulated the ranges of latitudes included in the resampling. We restricted analyses to *C. picta* because it was the only species (of the 3) to show an overall significant size–latitude relationship (see “Results” below).

For *C. picta*, we first resampled the full data set of 23 populations, and then restricted resampling to geographic subsets: the middle 19 populations in latitude (the core of the species latitudinal range), middle 21, lower 21, upper 21, lower 22, and upper 22 (Table 1). For the full dataset and each of the subsets, we conducted the random draw as described above, starting with 10,000 draws of 3 populations and increasing number of drawn populations by 1 until all populations had been included.

Results

Intraspecific body size trends

There was no significant latitudinal trend in body size for *C. serpentina* ($n = 11$; $r = 0.310$; $P = 0.35$) or *T. scripta* ($n = 22$; $r = 0.360$; $P = 0.10$), but mean body sizes of *C. picta* populations correlated with latitude ($n = 23$; $r = 0.521$; $P = 0.01$; Figure 1). However, the

Table 1. Correlations (r) and resampling analyses to achieve statistically significant and sufficiently powerful latitude-body size trends in *C. picta*

Resampling Set	Populations Excluded From Analyses	N	r	P	N to Yield $1-\beta \geq 0.80$
All	None	23	0.521	0.01	20
Upper 22	A	22	0.443	0.04	22
Lower 22	W	22	0.366	0.09	NA
Upper 21	A, B	21	0.496	0.02	20
Lower 21	V, W	21	0.252	0.27	NA
Middle 21	A, W	21	0.233	0.31	NA
Middle 19	A, B, V, W	19	0.125	0.61	NA

Names of excluded populations refer to points in Figure 1. NA = resampling set never achieves sufficient statistical power ($1-\beta \geq 0.80$).

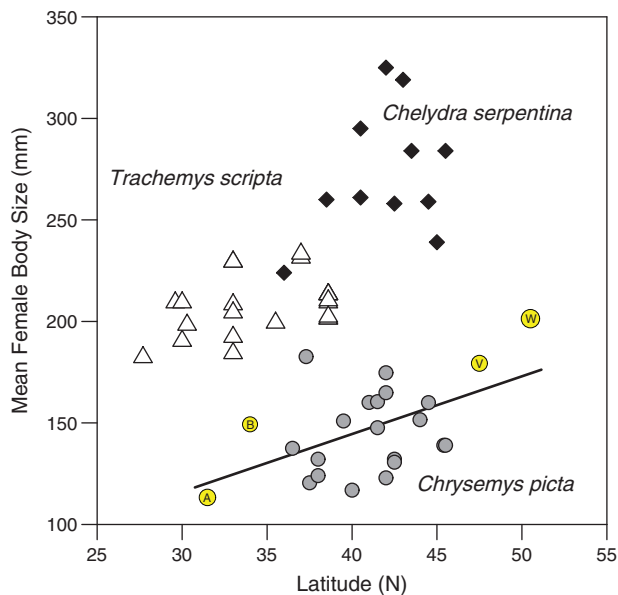


Figure 1. Correlations of mean body sizes of females (carapace or plastron length) in relation to latitude for 3 species of North American turtles. No significant correlation exists for the Snapping Turtle, *C. serpentina* ($N=11$ populations), or the Pond Slider, *T. scripta* ($N=22$ populations). The Painted Turtle, *C. picta* ($N=23$ populations) shows a significant increase of body size with latitude ($r=0.521$; $P=0.01$).

mean r (from 10,000 draws) for *C. picta* was only statistically significant ($r > r_{crit}$) after including 16 of the 23 total sample populations in resampling analyses (Figure 2), and the mean r for *T. scripta* or *C. serpentina* was never significant, regardless of how many populations were included (Figure 2).

Sample size and power

Sufficient power to detect a trend ($1-\beta \geq 0.80$) for *C. picta* was only achieved in 3 cases (Figure 3): when sampling at least 20 of the total 23 populations, when sampling at least 20 of the upper 21 populations, or when sampling all 22 of the upper 22 populations. Because it was not possible (on average) to observe a significant body size trend with respect to latitude for either *C. serpentina* or *T. scripta*, results from power analyses are not presented for those species.

Geographic extent and outlier effects

Our resampling analysis demonstrated a strong influential role of one population (“Population W,” Figure 1). In no case was a

significant trend detected when W was excluded from analyses. This population was not identified as influential according to conventional regression diagnostics. Population W had a Cook’s D of 0.23, DFFITS of 1.00 and DFBETAS of 0.90. Cutoff values to judge the influence of a data point are values >2 for DFFITS and DFBETAS (Belsley et al. 2004) and 0.5 for Cook’s D (Neter et al. 1990). Even under more conservative DFFITS and DFBETAS criteria of 1 (Neter et al. 1990), population W would not be considered likely to exert a substantial effect on regression results. Thus, conventional regression diagnostics failed to detect that a single data point determined the significant clinal relationship. In contrast, the influence of this population was always detected by our resampling approach (Table 1, Figure 3).

Discussion

With available data on body sizes of turtles, it may not be possible to confirm a significant intraspecific body size trend in relation to latitude in well-studied species (*T. scripta* and *C. serpentina*). Even in the very well-studied *C. picta*, 16 sample populations were required to determine a statistically significant mean trend (Figure 2). To have sufficient statistical power ($1-\beta > 0.8$) to detect a significant trend in *C. picta* required sampling across $\sim 20^\circ$ latitude.

In our analyses, the most northerly population of *C. picta* ultimately determined whether a statistically significant trend could be found. A single population effectively dominated every analysis, and when it was deleted from the total dataset, it was not possible to obtain a significant latitude-body size correlation, or to accumulate sufficient statistical power to detect a trend. This population would not be identified as individually influential when using traditional diagnostics, but resampling demonstrated its large leveraging effect.

It is unclear whether the influence of the leveraging population (W) represents a real biogeographical relationship or whether a unique evolutionary history led to that single population’s large mean size. However, assuming that statistical power accumulates in other species as it does for *C. picta*, these results provide initial estimates for the approximate sample sizes and geographic representation needed to detect and describe intraspecific body size trends.

Difficulties in detecting body size trends, and the accumulation of statistical power, are often overlooked issues in ecological and biogeographic investigations (Gerrodette 1987; Thomas and Juanes 1996; Gotelli and Ellison 2004). Our analyses also demonstrate that the geographic extent of sampling importantly influences the ability to detect body size trends. The effect of influential populations has also been observed empirically. Litzgus et al. (2004) found that a body size trend in relation to latitude in the Spotted Turtle *Clemmys*

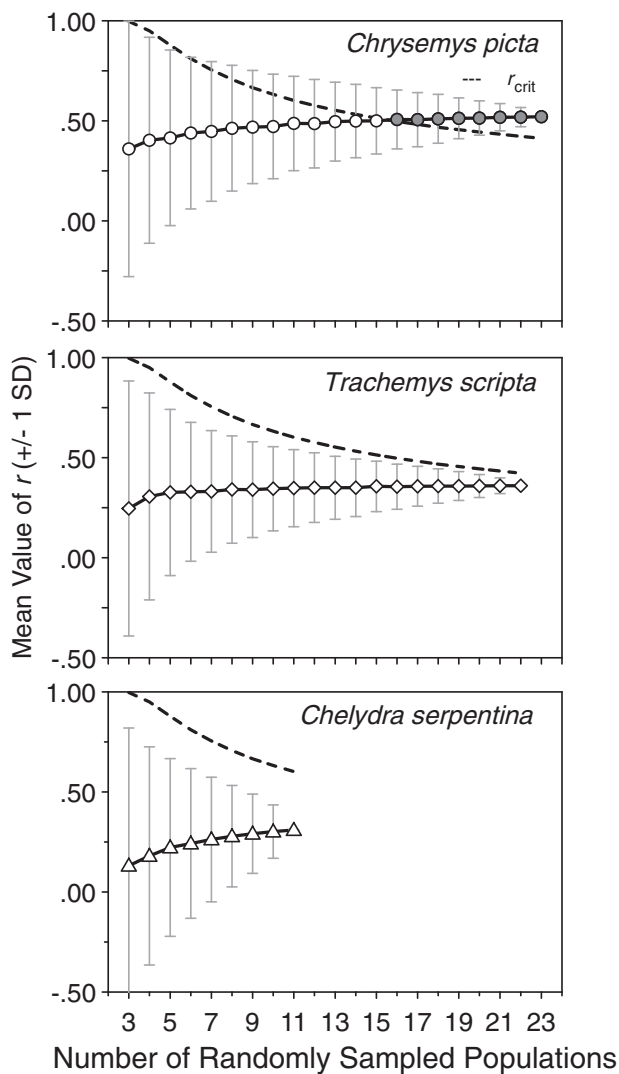


Figure 2. Resampling analyses (10,000 random draws per point) of the mean value of the correlation coefficient, r , between body size and latitude in relation to sample size in 3 species of North American turtles (error bars are ± 1 SD). For *T. scripta* and *C. serpentina*, it is not possible to obtain a significant mean correlation where the calculated r is larger than the critical value of r (dashed line); for *C. picta*, 16 or more populations (filled circles) must be part of the sample to yield a significant mean correlation.

guttata was only supported by including a single extreme northern population. Geographic extent of data used in analyses has been important in detecting other commonly referenced biogeographic trends, and [Harcourt \(2000\)](#) reported that the influence of geographic outliers largely determined the ability to demonstrate significant Rapoport trends (species having larger geographic ranges at higher latitudes) in primates.

Our results indicate that individual intraspecific trends should be considered when interpreting interspecific patterns and/or macroecological patterns across broad or phylogenetically disparate groups. [Ashton \(2004\)](#) reports that interspecific latitude-body size trends calculated from a wide range of intraspecific data across all tetrapods were robust to variation in both sampling (sample size) and geographic inclusivity (range of sample latitude). Our analyses do not contradict those interspecific findings, but draw attention to the importance of biological variation and the need for statistical

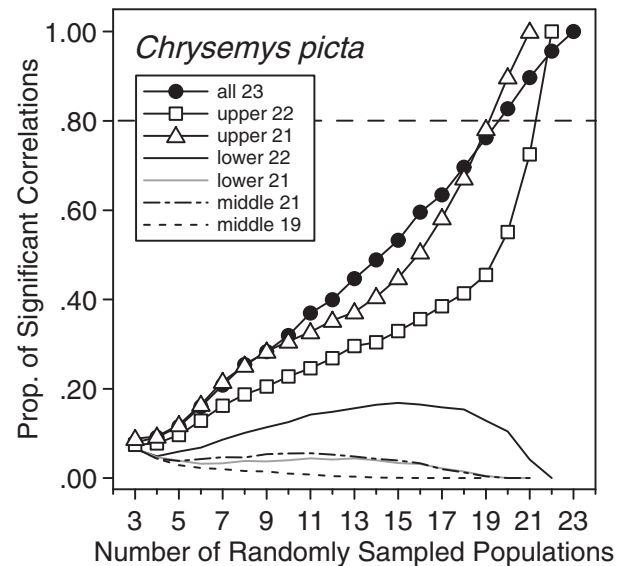


Figure 3. Resampling analyses of the power to detect a body size trend with respect to latitude in Painted Turtles (*C. picta*). Each data point represents the proportion of significant correlations from 10,000 random draws of given sample size, and latitudinal subset (subset details in Table 1 and Figure 1). A sufficiently powerful correlation ($1-\beta \geq 0.80$) is only obtained when 20 or more populations are included when the sampling is done from all 23 populations (circles), or upper 21 populations (triangles), or when sampling all 22 populations of the upper 22 (squares).

rigor in individual intraspecific studies, which are often overlooked in broad analyses (necessarily so as the goals of macro and meta-analyses are to draw broad inference).

Numerous diagnostic techniques are available to detect outliers and leveraging ([Martin and Roberts 2006](#)), however, our results demonstrate that common diagnostic procedures may not always be able to detect problems with real world data. In particular, commonly used regression diagnostics (DFFITS, DFBETAS, and Cook's D) can fail to recognize influential data points that, in and of themselves, can determine whether a significant body size trend with respect to latitude can be demonstrated. We, therefore, propose using resampling as an additional diagnostic tool to evaluate sample size and statistical power, as well as to identify influential observations and potential outliers in intraspecific analyses.

By drawing attention to, and addressing the importance of these issues at the intraspecific level, macroecological analyses (interspecific) combined with intraspecific data should be far more statistically defensible. Although we have shown how ignoring sample size, statistical power, and geographic extent can lead to potentially misleading conclusions for body size trends with respect to latitude in turtles, these considerations should be addressed in all studies attempting to identify broad ecological and biogeographic patterns. Moreover, the general diagnostic framework of resampling techniques provided here ([Supplementary Appendix S1](#) provides an R version of the code) can easily be extended to other predictors of and explanations for geographic variation in phenotypic traits.

Acknowledgments

This paper came from several critical discussions with students and faculty at the University of Nevada, Reno. Mike Sears provided useful encouragement and advice concerning resampling analyses, Kyle Ashton and Chris Feldman provided useful data tables and appendices. We thank members of the EECB

peer review group for comments on an early version of the manuscript. The UNR Program in Ecology, Evolution, and Conservation Biology, and the Biological Resources Research Center provided funding.

Supplementary Material

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

References

- Aresco MJ, 2004. Reproductive ecology of *Pseudemys floridana* and *Trachemys scripta* (testudines: emydidae) in northwestern Florida. *J Herpetol* 38: 249–256.
- Arnett AE, Gotelli NJ, 1999. Geographic variation in life-history traits of the ant lion *Myrmeleon immaculatus*: evolutionary implications of Bergmann's rule. *Evolution* 53:1180–1188.
- Ashton KG, 2004. Sensitivity of intraspecific latitudinal clines of body size for tetrapods to sampling, latitude and body size. *Integr Comp Biol* 44: 403–412.
- Ashton KG, Feldman CR, 2003. Bergmann's rule in nonavian reptiles: turtles follow it, lizards and snakes reverse it. *Evolution* 57:1151–1163.
- Atkinson D, Sibly RM, 1997. Why are organisms usually bigger in colder environments? Making sense of a life history puzzle. *Trends Ecol Evol* 12: 235–239.
- Belsley DA, Kuh E, Welsch RE, 2004. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. Hoboken (NJ): Wiley.
- Bergmann C, 1847. Über die verhältnisse der wärmeökonomie der thiere zu ihrer grösse. *Göttinger Studien* 3:595–708.
- Blanckenhorn WU, Demont M, 2004. Bergmann and converse Bergmann latitudinal clines in arthropods: two ends of a continuum? *Integr Comp Biol* 44:413–424.
- Blanckenhorn WU, Stillwell RC, Young KA, Fox CW, Ashton KG, 2006. When Rensch meets Bergmann: does sexual size dimorphism change systematically with latitude? *Evolution* 60:2004–2011.
- Brown JH, 1999. Macroecology: progress and prospect. *Oikos* 87:3–14.
- Brown JH, Maurer BA, 1989. Macroecology: the division of food and space among species on continents. *Science* 243:1145–1150.
- Cohen J, 1988. *Statistical Power Analysis for the Behavioral Sciences*, 2nd edn. Hillsdale (NJ): Lawrence Erlbaum.
- Cook RD, 1979. Influential observations in linear regression. *J Am Stat* 74: 169–174.
- Cooley CR, Floyd AO, Dolinger A, Tucker PB, 2003. Demography and diet of the painted turtle *Chrysemys picta* at high-elevation sites in southwestern Colorado. *Southwest Nat* 48:47–53.
- Crowley PH, 1992. Resampling methods for computation-intensive data analysis in ecology and evolution. *Ann Rev Ecol Systematics* 23:405–447.
- de Queiroz A, Ashton KG, 2004. The phylogeny of a species-level tendency: species heritability and possible deep origins of Bergmann's rule in tetrapods. *Evolution* 58:1674–1684.
- Emer S, 2004. Growth of an introduced population of *Trachemys scripta elegans* at Fox Pond, Eckerd College, Pinellas County. *Florida Herp Rev* 35:34–35.
- Ernst CH, Barbour RW, 1989. *Turtles of the World*. Washington (DC): Smithsonian Institution Press.
- Ernst CH, Barbour RW, Lovich JE, 1994. *Turtles of the United States and Canada*. Washington (DC): Smithsonian Institution Press.
- Gerrodette T, 1987. A power analysis for detecting trends. *Ecology* 68: 1364–1372.
- Gibbons JW, Lovich JE, 1990. Sexual dimorphism in turtles with emphasis on the slider turtle *Trachemys scripta*. *Herpetol Monog* 4:1–29.
- Gilchrist GW, Huey RB, Balanya J, Pascual M, Serra L, 2004. A time series of evolution in action: a latitudinal cline in wing size in South American *Drosophila subobscura*. *Evolution* 58:768–780.
- Gotelli NJ, Ellison AM, 2004. *A Primer of Ecological Statistics*. Sunderland (MA): Sinauer Associates, Inc.
- Harcourt AH, 2000. Latitude and latitudinal extent: a global analysis of the Rapoport effect in a tropical mammalian taxon: primates. *J Biogeogr* 27: 1169–1182.
- Iverson JB, Higgins H, Sirulnik A, Griffiths C, 1997. Local and geographic variation in the reproductive biology of the snapping turtle *Chelydra serpentina*. *Herpetologica* 53:96–117.
- James FC, 1991. Complementary descriptive and experimental studies of clinal variation in birds. *Am Zool* 31:694–706.
- Lindeman PV, 1997. Does life-history variation in the turtle *Chrysemys picta* have a subspecific component? *J Herpetol* 31:155–161.
- Litzgus JD, DuRant SE, Mousseau TA, 2004. Clinal variation in body and cell size in a widely distributed vertebrate ectotherm. *Oecologia* 140: 551–558.
- Lovich JE, Gibbons JW, 1992. A review of techniques for quantifying sexual size dimorphism. *Growth Develop Aging* 56:269–281.
- Martin MA, Roberts S, 2006. An evaluation of bootstrap methods for outlier detection in least squares regression. *J Appl Stat* 33:703–720.
- Neter J, Wasserman W, Kutner MH, 1990. *Applied Linear Regression Models*. Homewood (IL): Irvin.
- Roff DA, 2006. *Introduction to Computer-Intensive Data Analysis in Biology*. New York: Cambridge University Press.
- Seidel ME, Stuart JN, Degenhardt WG, 1999. Variation and species status of slider turtles (Emydidae: *trachemys*) in the southwestern United States and adjacent Mexico. *Herpetologica* 55:470–487.
- Simpson DG, Ruppert D, Carroll RJ, 1992. On one-step GM estimates and stability of inferences in linear-regression. *J Am Stat Assn* 87:439–450.
- Starkey DE, Shaffer HB, Burke RL, Forstner MRJ, Iverson JB et al. 2003. Molecular systematics, phylogeography, and the effects of pleistocene glaciation in the painted turtle *Chrysemys picta* complex. *Evolution* 57: 119–128.
- Student, 1908. The probable error of a mean. *Biometrika* 6:1–25.
- Thomas L, Juanes F, 1996. The importance of statistical power analysis: an example from animal behaviour. *Anim Behav* 52:856–859.
- Trussell GC, 2000. Phenotypic clines, plasticity, and morphological trade-offs in an intertidal snail. *Evolution* 54:151–166.
- Tucker JK, Paukstis GL, Janzen FJ, 1998. Annual and local variation in reproduction in the red-eared slider *Trachemys scripta elegans*. *J Herpetol* 32: 515–526.
- Van Voorhies WA, 1996. Bergmann size clines: a simple explanation for their occurrence in ectotherms. *Evolution* 50:1259–1264.
- Watt C, Mitchell S, Salewski V, 2010. Bergmann's rule; a concept cluster? *Oikos* 119:89–100.
- Zar JH, 1999. *Biostatistical Analysis*. Upper Saddle River (NJ): Prentice Hall.