



Recovering missing data: estimating position and size of caudal vertebrae in *Staurikosaurus pricei* Colbert, 1970

ORLANDO N. GRILLO and SERGIO A.K. AZEVEDO

¹Departamento de Geologia e Paleontologia, Museu Nacional/UFRJ, Quinta da Boa Vista, s/n
São Cristóvão, 20940-040 Rio de Janeiro, RJ, Brasil

Manuscript received on August 24, 2010; accepted for publication on December 9, 2010

ABSTRACT

Missing data is a common problem in paleontology. It makes it difficult to reconstruct extinct taxa accurately and restrains the inclusion of some taxa on comparative and biomechanical studies. Particularly, estimating the position of vertebrae on incomplete series is often non-empirical and does not allow precise estimation of missing parts. In this work we present a method for calculating the position of preserved middle sequences of caudal vertebrae in the saurischian dinosaur *Staurikosaurus pricei*, based on the length and height of preserved anterior and posterior caudal vertebral centra. Regression equations were used to estimate these dimensions for middle vertebrae and, consequently, to assess the position of the preserved middle sequences. It also allowed estimating these dimensions for non-preserved vertebrae. Results indicate that the preserved caudal vertebrae of *Staurikosaurus* may correspond to positions 1-3, 5, 7, 14-19/15-20, 24-25/25-26, and 29-47, and that at least 25 vertebrae had transverse processes. Total length of the tail was estimated in 134 cm and total body length was 220-225 cm.

Key words: missing data, caudal vertebra, regression, *Staurikosaurus pricei*.

INTRODUCTION

Missing data is a major problem in the reconstruction of extinct taxa (Paul 1987, 1988, Czerkas 1997) and has several implications in phylogenetic analyses (Nixon and Davis 1991, Novacek 1992, Maddison 1993, Anderson 2001, Kearney and Clark 2003, Norell and Wheeler 2003) and also on biomechanical studies (Holliday et al. 2010). Comparisons to phylogenetic related taxa are often used to make assumptions about shape and dimensions of missing parts in order to allow the reconstruction of fossil taxa (Paul 1987, 1988, Czerkas 1997). However, when we compare closely related organisms we notice that, despite the close relationship, they can be distinguished not just by their autapomorphies, but also by some morphometric and biomechanical proper-

ties (Miller and Gross 1998). When we deal with incomplete materials, some of the divergences may have been lost as missing data. The use of data from closely related taxa to “fill these gaps” is a common strategy applied on most reconstructions of fossil taxa (Paul 1987, 1988, Czerkas 1997), but this approach, evidently, hides the possible differences among them and discards the taxon as a potential material for comparative studies.

Some authors have developed formulas to reconstruct general dimensions of extinct organisms, as the total body length, based on the dimension of hind limb elements (Christiansen 1999) or the skull length (Therrien and Henderson 2007). Unfortunately, total length is not sufficient to determine the size of specific segments of the body.

Lack of data also excludes several materials as potential samples for biomechanical studies, in which accurate dimensions of body segments and associated

Proceedings of the Third Gondwanan Dinosaur Symposium
Correspondence to: Orlando N. Grillo
E-mail: ongrillo@gmail.com

muscles are necessary to determine body mass, position of the center of mass and also to estimate muscle lengths and volumes (Gunga et al. 2007, Hutchinson et al. 2007). For example, the muscle *caudofemoralis longus*, which inserts in the caudal vertebrae, is the main femoral extensor in archosaurs with the exception of birds (Gatesy 1990). A larger *caudofemoralis longus* would move the center of mass to a more posterior position. Accordingly, a correct estimation of the size of this muscle is necessary to make assumptions about the biomechanics of archosaurs. A longer muscle would also lead to a larger muscle moment arms since it would be farther from the center of rotation of the femur (Hutchinson et al. 2005). Gatesy (1990) proposed that the length of the muscle *caudofemoralis longus* is related to the number of caudal vertebrae that possess transverse processes. Therefore, to estimate the size of this muscle in fossil taxa that have not preserved a complete series of caudal vertebrae, it is necessary to correctly estimate the position of the preserved caudal vertebrae and, also, to estimate the dimension of the missing ones, especially of those that have transverse processes.

In this work we present a method for estimating the position of the preserved caudal vertebrae and the dimensions (length and height of vertebral centra) of non-preserved vertebrae. As an example, we reconstructed the tail of the basal saurischian dinosaur *Staurikosaurus pricei* Colbert, 1970 (Santa Maria Formation, Late Triassic of Rio Grande do Sul, Brazil). This taxon is known by an almost complete skeleton that includes 35 preserved caudal vertebrae, of which six are isolated anterior vertebrae, a posterior sequence of 19 vertebrae (which lack transverse processes) and three middle portion sequences (with seven, two and two vertebrae, from which the first and second sequences have transverse processes) (Fig. 1). Accordingly, in order to accurately determine the number of vertebrae that have transverse processes, it is necessary to correctly determine the position of the middle sequences.

Some authors (Colbert 1970, Galton 1977) indicated that *Staurikosaurus* had about 20 vertebrae with transverse processes, but they did not provide an accurate estimation of the dimension of the vertebrae included in this sequence. An empirical approach is necessary to correctly quantify these values.

Also, the methodology presented in this work allowed a more precise estimation of the total length of *Staurikosaurus*. The vertebral column of *Staurikosaurus* holotype MCZ 1669 is almost complete: it lacks the atlas, axis and the third cervical vertebra, and about twelve vertebrae in the mid portion of the tail (Bittencourt and Kellner 2009). Consequently, as mandibles are also present, a correct estimation of the total length of this taxon is centered on a correct estimation of the size of the missing caudal vertebrae and on the total length of the tail.

MATERIALS AND METHODS

In order to determine the position of the preserved caudal vertebrae of *Staurikosaurus pricei* it was initially assumed that this taxon had a total number of vertebrae similar to *Herrerasaurus ischigualastensis*, which is considered its closest related taxon (Novas 1993, 1996, 1997, Sereno and Novas 1993, Sereno 1999, Rauhut 2003, Langer 2004, Langer and Benton 2006, Bittencourt and Kellner 2009). Novas (1993) indicated that the holotype of *Herrerasaurus* (PVL 2566) has 43 caudals preserved, with a hiatus separating the eight most posterior from the rest of the tail, and considered that a total of 47 caudals should have been present.

It was assumed that the first four preserved anterior caudals of *Staurikosaurus* correspond to positions 1, 2, 3 and 5 (Table I). The gap between caudals 3 and 5 was assumed because, despite both have centra with similar height, the prezygapophyses of the fourth preserved vertebra do not articulate correctly with the postzygapophyses of the third vertebra (Fig. 1). We believe that just one vertebra is missing between them, due to the very similar dimensions of the centrum and the general shape of these two vertebrae. The distal articulated sequence corresponds to positions 29 to 47 (Table I). All other preserved vertebrae were initially treated as of undetermined position (Table I).

The first preserved vertebra is considered the third sacral by some authors (Langer 2004, Bittencourt and Kellner 2009; but see Colbert 1970, Galton 1977, 2000), therefore two analyses were conducted: including and excluding this vertebra.

The sixth preserved caudal includes an incomplete centrum. Therefore it was ignored in the present ana-

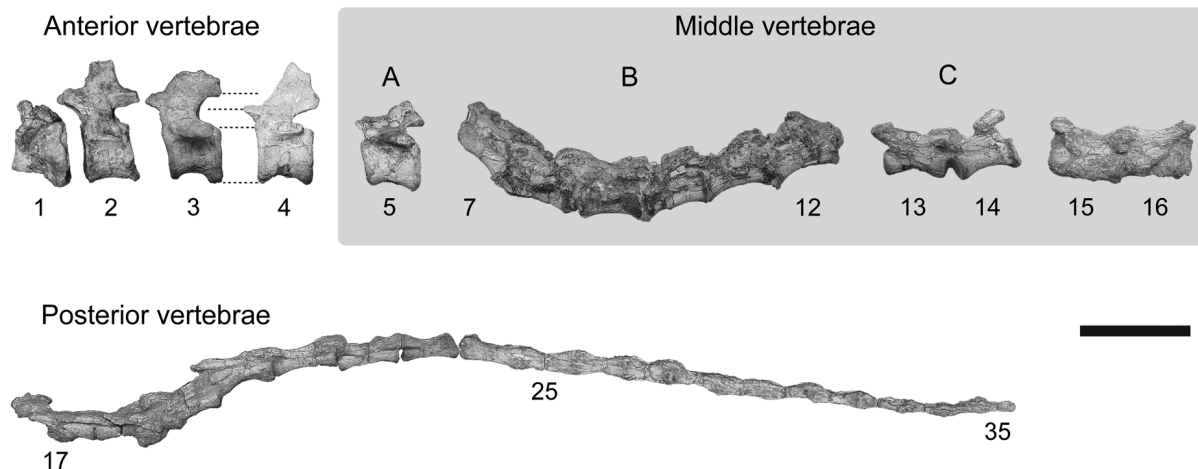


Fig. 1 – Preserved caudal vertebrae of *Staurikosaurus* and nomenclatures used in this paper. The positions are relative to the total number of preserved caudal vertebrae (fragment of the sixth preserved caudal not shown). Dashed lines between vertebrae 3 and 4 indicate the similar height of the centra, but misalignment between pre- and postzygapophyses indicate that they are not consecutive vertebrae. Sequence 15-16 was excluded from the analyses (see text). Scale bar equals 5 cm.

lysis. One of the middle sequences composed of two vertebrae (15-16 in Fig. 1) was also ignored because the bone surface is quite fragmented, and the general appearance of the bones and the disposition of the surface fragments suggest that the crystallization of calcite and hematite caused the break and dislocation of the natural bone structure, increasing the volume of these vertebrae (Holz and Schultz 1998). As proposed by Holz and Schultz (1998), fossils from the Mid-Triassic of south Brazil may be classified according to three levels of deformation/disruption: 1) non-deformed, 2) internally altered (with volumetric change of bones not always apparent, but with small fractures visible on the external layer of the compact bone), and 3) disrupted. External observations of the holotype of *Staurikosaurus* indicate that most bones belong to level 2, including all caudal vertebrae, except this middle sequence composed of two caudals, which was excluded from our analyses. As fractures on most caudal vertebrae are scarce, we consider volumetric changes absent or insignificant, not affecting our measurements.

Measurements of the centrum length and height at the anterior articular surface of all preserved vertebrae in the holotype MCZ 1669 were computed. Measurements of vertebrae 1-3, 5 (referred in this paper as ‘anterior vertebrae’) and 29-47 (referred as ‘posterior vertebrae’) were used to produce regressions that allowed

estimating these values for the non-preserved and non-positioned vertebrae (referred as ‘middle vertebrae’). The estimated values were compared to the observed values in the middle vertebrae using the deviation equation

$$D = (o - e)/e$$

in that o corresponds to the observed value and e corresponds to the estimated value. The position of the middle vertebrae was determined in order to obtain the best fit to all regressions (deviation smaller than $\pm 10\%$). The non-positioned vertebrae are referred in this paper as sequences A (5th preserved vertebra), B (7th to 12th preserved vertebrae) and C (13th to 14th preserved vertebrae).

Regressions were produced for length, height and height/length ratios using PAST v2.0 (Hammer et al. 2001). The type of regression curve (linear, logarithmic, binomial or polynomial) was chosen according to the R^2 correlation value. Values of R^2 were obtained using the RMA methodology because it is a more adequate one in the case the measurements may contain errors (Bohonak and Linde 2004).

After positioning the middle vertebrae, their values of height and length were included in the data matrix to produce new regressions that allowed estimating the dimensions of the non-preserved vertebrae.

TABLE I
Heights and lengths (mm) of preserved caudal centra in *Staurikosaurus pricei*. The positions are relative to the total number of preserved caudal vertebrae.

	Preserved position	Assumed position	Height	Length	
	1*	1	25.0	28.5	
Anterior vertebrae	2	2	26.0	24.0	
	3	3	23.5	22.5	
	4	5	23.0	23.5	
Middle vertebrae	A	5	?	22.0	
	Excluded	6**	-	-	
		7	?	17.0	25.5
		8	?	17.3	31.5
	B	9	?	17.0	28.0
		10	?	16.8	28.0
		11	?	16.5	28.5
		12	?	17.0	29.0
	C	13	?	13.0	29.0
		14	?	13.5	29.0
	Excluded	15**	-	-	-
	Excluded	16**	-	-	-
	Posterior vertebrae	17	29	11.25	30.5
		18	30	11.25	30.5
19		31	11.25	30.0	
20		32	10.5	26.5	
21		33	8.8	29.0	
22		34	8.8	28.5	
23		35	9.0	28.0	
24		36	8.5	27.0	
25		37	8.5	27.0	
26		38	8.0	22.5	
27		39	8.0	22.0	
28		40	7.3	23.5	
29		41	7.0	23.0	
30		42	6.7	20.0	
31		43	6.0	20.0	
32		44	5.0	18.5	
33		45	5.0	21.5	
34		46	5.0	15.5	
35		47	5.0	16.0	

*The first preserved caudal vertebra is considered the third sacral by some authors (Langer 2004, Bittencourt and Kellner 2009). **The sixth, fifteenth and sixteenth preserved vertebrae were excluded from the analysis due to difficulties to measure dimensions accurately (see text for explanation).

RESULTS

Height and length of preserved caudal vertebrae centra are indicated in Table I. Regression equations are presented in Table II. Regression graphs obtained from the data of the anterior and posterior vertebrae are shown in Figure 2.

Height of centrum – The analysis of the height values indicates that a linear or a binomial equation adequately predicts the observed values for the anterior and posterior vertebrae ($R^2 = 0.990$ and $R^2 = 0.995$, respectively).

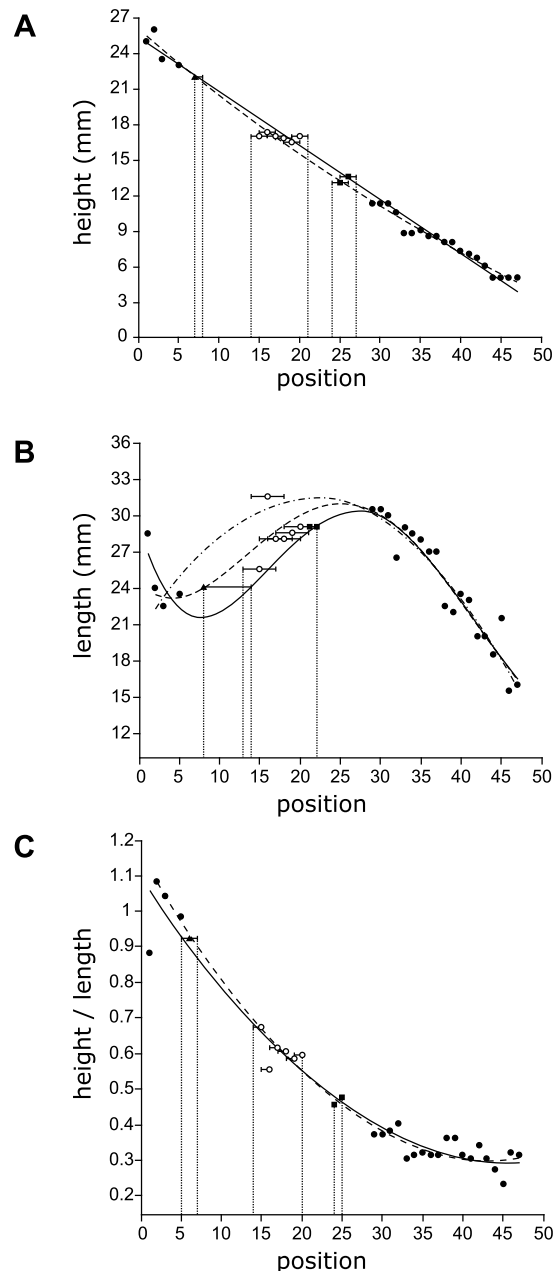


Fig. 2 – Positioning of the middle sequences A (triangle), B (open circles), and C (squares) based on regressions obtained from data of anterior and posterior vertebrae (filled circles). A – Linear (full line) and binomial (dashed line) regressions for height values. B – Fourth order polynomial regressions including the first caudal (full line) and excluding the first caudal (dashed line), and binomial regression excluding first caudal (dash and dot line). C – Binomial regressions including the first caudal (full line) and excluding the first caudal (dashed line). Alternative positions for the middle sequences are indicated by the line brackets. Dotted lines indicate the maximum anterior and posterior position that each sequence may occupy in each case.

TABLE II
Regression equations obtained from the analyses. The x value corresponds to the position of the vertebra.

Anterior and posterior sequences only			
	Equation	R ²	Type
Height	$H = -0.4579x + 25.398$	0.990	Linear; including 1 st caudal
	$H = -0.4574x + 25.377$	0.988	Linear; excluding 1 st caudal
	$H = 0.0026x^2 - 0.5742x + 25.9385$	0.995	Binomial; including 1 st caudal
	$H = 0.0028x^2 - 0.5851x + 26.1109$	0.993	Binomial; excluding 1 st caudal
Length	$L = 5.356E-5x^4 - 0.0061x^3 + 0.2105x^2 - 2.2355x + 28.9536$	0.899	4 th order polynomial; including 1 st caudal
	$L = 0.0243x^2 + 1.043x + 20.2811$	0.898	Binomial; excluding 1 st caudal
	$L = 3.154E-5x^4 - 0.0035x^3 + 0.105x^2 - 0.6770x + 24.4428$	0.912	4 th order polynomial; excluding 1 st caudal
Height/length	$y = 0.0004x^2 - 0.0349x + 1.0912$	0.965	Binomial; including 1 st caudal
	$y = 0.0005x^2 - 0.0398x + 1.1624$	0.988	Binomial; excluding 1 st caudal
Positioned middle sequences included*			
	Equation	R ²	Type
Height	$H = -0.4467x + 24.853$	0.987	Linear; including 1 st caudal
	$H = -0.4437x + 24.747$	0.985	Linear; excluding 1 st caudal
	$H = 0.0028x^2 - 0.5787x + 25.8325$	0.993	Binomial; including 1 st caudal
	$H = 0.0029x^2 - 0.5862x + 25.9312$	0.992	Binomial; excluding 1 st caudal
Length	$L = 3.386E-5x^4 - 0.0038x^3 + 0.1242x^2 - 1.0822x + 26.7966$	0.871	4 th order polynomial; including 1 st caudal
	$L = 1.908E-5x^4 - 0.0023x^3 + 0.0667x^2 - 0.2885x + 23.6239$	0.892	4 th order polynomial; excluding 1 st caudal
Height/length	$y = 0.0004x^2 - 0.0348x + 1.0896$	0.961	Binomial; including 1 st caudal
	$y = 0.0005x^2 - 0.0398x + 1.1580$	0.982	Binomial; excluding 1 st caudal

* Middle sequences A, B, and C were considered to be in positions 7, 14-19, and 24-25, respectively.

The exclusion of the first caudal produced very similar R² values: 0.988 for the linear regression and 0.993 for the binomial regression.

These results indicate that the middle sequence A could be at positions seven or eight (Fig. 2A). The sequence B would correspond to positions 14-19 or 15-20 according to the binomial regression, and to positions 15-20 or 16-21 according to the linear regression. Sequence C could be at positions 24-25 to 26-27, being the position 25-26 a better approximation to both regressions (Fig. 2A).

Length of centrum – The second vertebra in sequence B deviated largely from other vertebrae in this sequence and could not be aligned with any regression (Fig. 2B). Therefore, this vertebra was ignored when trying to determine the position of the sequence B.

The analysis of the length values including the first caudal indicates that a fourth order polynomial (R² = 0.899) was the only equation capable of predicting the observed values for anterior and posterior vertebrae with a R² value close to 0.9. The resultant from this regression suggests that middle sequences A and B correspond, respectively, to positions 14 and 16-21 or 17-22. Sequence C would also fit position 21-22, since lengths are equal to those of the most posterior vertebrae of sequence B (Fig. 2B).

Excluding the first caudal, a binomial equation predicted the values with a similar accuracy (R² = 0.899), but the estimated positions for sequences A, B and C (4, 6-11, and 12-13) are incompatible with the size and morphology of these vertebrae in relation to the anterior ones. A fourth order polynomial (also excluding the first caudal) produced a R² value of 0.912 and could better

TABLE III
Estimated height and length (mm) of non-preserved caudal centra in *Staurikosaurus pricei*.

Estimated position	4	6*	8	9	10	11	12	13	20	21	22	23	26	27	28
Height	23.6	22.5	21.4	20.8	20.3	19.8	19.3	18.8	15.4	14.9	14.4	14.0	12.7	12.2	11.8
Length	23.4	23.8	24.5	24.9	25.3	25.8	26.3	26.7	29.5	29.8	30.0	30.1	30.2	30.1	30.0

*Only a fragment of the centrum of the sixth vertebra was preserved, so that its dimensions were also estimated.

predict the position of the middle sequences. This suggests that sequences A, B, and C correspond to positions 8, 13-18, and 16-17/17-18, and also resulted in overlapping the position of sequences B and C (Fig. 2B).

Height/Length ratio – Binomial regressions predicted the values for anterior and posterior vertebrae with a R^2 of 0.965 (including the first caudal) and 0.988 (excluding the first caudal).

The first caudal vertebra deviated largely from the predicted value (16-22%), but its inclusion or exclusion does not influence the results for the middle sequences, except for sequence A (Fig. 2C). Sequence A could be at position six or seven (based on the equation that excludes the first caudal), or at position five (based on the equation that includes the first caudal), overlapping with the fifth already positioned vertebra. The sequences B and C fit to positions 14-19/15-20 and 24-25, respectively (Fig. 2C).

Size of non-preserved vertebrae – Considering that the most reasonable positions for sequences A, B, and C are, respectively, 7, 14-19/15-20 and 24-25/25-26 (see discussion), new regressions including the data from these vertebrae (Table II) allowed estimating the length and height values for the centrum of non-preserved caudal vertebrae (Table III). The new regressions were generated in order to obtain a single and more precise value for each non-preserved vertebra because the positioning of middle vertebrae relied on two regressions for length and four for height. Also, using these regressions to estimate the size of non-preserved middle vertebrae would produce less precise values because data from preserved middle sequences are not considered. The estimated values were obtained using the regressions with larger R^2 values, that is, the binomial equation that includes data from the first caudal (for the height value) and the poly-

nomial equation that excludes data from the first caudal (for the length value).

Total caudal length and total body length – The total length of the tail can be estimated by the sum of the length of centra from all vertebrae plus the distance among vertebrae that is occupied by cartilage discs. This distance corresponds to about 10% of the length of the vertebrae (Paul 1988). The estimation, including real values for preserved vertebrae and estimated values for non-preserved ones, indicates that the tail was 1343 mm long. The length of the portion of the tail that comprises vertebrae with transverse processes (vertebrae 1 to ~25) is 738 mm.

The total length of the body of *Staurikosaurus* can be estimated by the sum of the estimated caudal length and the length of the anterior portion of the body (head to pelvis). As the holotype of *Staurikosaurus* has an almost complete pre-sacral vertebrae series, lacking the three anteriormost cervical vertebrae (Bittencourt and Kellner 2009), the length can be estimated from the length of the preserved vertebrae (accounting for the natural curvature of the trunk and neck), considering also the space of the cartilage discs, plus the length of the mandibles and of the missing cervical vertebrae (it was used the length of the fourth cervical vertebra as a reference). This calculation provided a value of 85-90 cm for the head-pelvis, depending on the inclination of the trunk (10° to 25° ; see Fig. 3), so that the total length of *Staurikosaurus* can be estimated as 220 to 225 cm.

DISCUSSION

The positioning of vertebrae in an incomplete series is often accomplished by the analysis of the general morphology and size of each fossilized vertebra. However,

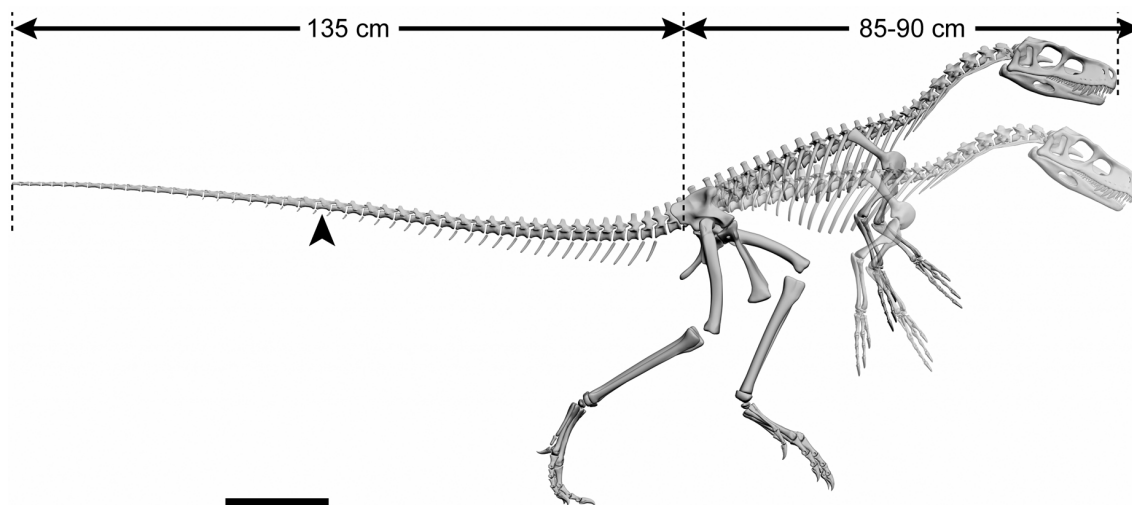


Fig. 3 – Reconstruction of the skeleton of *Staurikosaurus* indicating how alternative postures may change total length estimations. Arrowhead indicates the position of the most posterior preserved vertebra that has transverse processes (position 25). Scale bar equals 20 cm.

this procedure does not provide accurate and testable ways of predicting vertebral position and, also, does not provide a simple way of estimating dimensions of non-preserved vertebrae in the sequence. The results obtained from the regressions in this work allowed an accurate positioning of the caudal vertebrae in *Staurikosaurus pricei*.

Comparing results predicted by the fourth order polynomial regressions for length values that include and exclude the first caudal, sequence A could be positioned anywhere between positions 8 and 14. These equations also predict that the first vertebra of sequence B would be in position 13 to 17. The centrum height and width of the vertebra of sequence A are much larger than those of the first vertebra in sequence B. Also, the vertebra of sequence A is quite similar in size to the fifth caudal vertebra and has a general morphology similar to the more anterior vertebrae. This indicates that sequence A should correspond to a more anterior position, suggesting that the results predicted by the equations that exclude the first caudal are more accurate. This result is also in agreement with the regression of the height values that indicates that this vertebra corresponds to position seven or eight. Also, analyzing the regression for the height/length ratio, it can be observed that the first caudal deviates by a large extent from predicted values, even if it is included in the analysis. Ac-

cordingly, regressions that do not include the first caudal should provide more accurate results. This does not necessarily suggest that this vertebra should be considered a sacral; this conclusion needs to be based on anatomical studies since sacral vertebrae have modified transverse processes and ribs that allow them to connect to the ilium (Williston 1925, Romer 1956). Several authors, based on anatomical analyses of vertebrae and also pelvic bones, discussed the classification of this vertebra as the third sacral (Langer 2004, Langer and Benton 2006, Bittencourt and Kellner 2009), or the first caudal (Colbert 1970, Galton 1977, 2000) and new herrerasaurid material (Alcober and Martinez 2010) may contribute to clarify this question. We consider that the divergence in dimensions compared to predicted values is associated with the difference in dimension of the centrum of the second sacral vertebra relative to the second caudal (as defined in this work), being much larger in the former. The dimensions of the centrum of the first caudal (or third sacral) correspond to a transition point between these two vertebrae.

Considering that regressions that exclude the first caudal are more accurate, sequence A would most probably belong to positions six to eight, as it is predicted by regressions of the three analyses (height, length and height/length). The position seven better fits the regressions since position six would deviate more for pre-

dicted height values, and position eight would not fit the height/width ratio estimation. Changing it to a more anterior or posterior position would reduce the deviation for one regression, but increases the deviation for the other.

The position of sequence B is also better predicted by the regressions that exclude the first caudal. The binomial regressions for height and height/length values both indicate that the position of sequence B is either 14-19 or 15-20. Linear regression for height values suggests a slightly posterior position (16-21), but the regression of length values indicates a more anterior position (13-18). Furthermore, both equations are predicting the position of sequence B with less accuracy than binomial regressions for height and height/length values, and the sequence B should, therefore, be considered as corresponding to vertebrae 14-19 or 15-20.

The sequence C is positioned as 24-25 or 25-26 according to regressions of height and height/length values, but the regression of length values indicates a much more anterior position (16-17 or 17-18). When the first caudal is included in the length value analysis, the regression indicates that sequence C should be at position 21-22. These results for length values are inconsistent with the position predicted for sequence B (14-19 or 15-20) because an obviously impossible overlap would occur. Thus, the most probable scenario would be that of the vertebrae in sequence C being shorter than the prediction, and they shall belong to position 24-25 or 25-26.

Considering the most probable positioning for sequences A, B and C as being 7, 14-19/15-20 and 24-25/25-26, respectively, it is possible to conclude that the binomial regression for height values produced the most precise results, even if the first caudal is included in the analysis. Also, the binomial regression of height/length values provided accurate results for the three sequences when the first caudal was excluded.

The position of the middle sequence C, which is the most posterior sequence in which transverse processes are present, indicates that at least ~25 caudal vertebrae have this structure. The vertebrae that follow this sequence (26-28/27-28) are not preserved (and/or were excluded from the analysis due to poor preservation), and vertebrae 29-47 do not have transverse pro-

cesses. According to Galton (1977), *Staurikosaurus* had a total of 45 or more caudal vertebrae from which about 20 had transverse processes, which differs from the present results.

Gatesy (1990) suggested that transverse processes correspond to the area of insertion of the muscle *caudofemoralis longus*, so that the number of vertebrae that have this process would be an indicative of the length of this muscle. In *Staurikosaurus* this muscle would insert in the 25 most proximal caudal vertebrae and, consequently, it was larger than it was supposed by previous studies. Considering that *Staurikosaurus* had just 20 vertebrae with transverse processes (according to Colbert 1970, Galton 1977), the area of insertion of the muscle *caudofemoralis longus* would be 543 mm long, which corresponds to 74% of the estimated length in the present study. This muscle is generally the bulkier in archosaurs and it helps to balance the anterior part of the body (Gatesy 1990). Consequently, it has a major importance in the question of positioning the center of mass because a longer muscle would move the center of mass posteriorly. Also, a longer insertion area may increase the moment arm for this muscle, and this also may have implications in posture and locomotion patterns adopted by *Staurikosaurus*. Biomechanical studies are required to evaluate these hypotheses, but these observations point to the importance of the correct estimation of each body part for biomechanical studies that require the estimation of total body mass and, more importantly, the correct positioning of the center of mass (Allen et al. 2009). We believe that our method may be applied to dorsal and cervical vertebrae as well. In this case, it would also contribute to correctly estimate dimensions of body parts such as the trunk and neck in other taxa, increasing the precision in the skeletal reconstructions used in biomechanical studies.

This methodology may also be useful to estimate some character states used in systematic studies. Gauthier (1986) and Rauhut (2003) included in their analysis data of the number of caudal vertebrae that have transverse processes. Our result on the number of transverse processes in the caudals of *Staurikosaurus*, increasing from 20 to 25, would not change the results of Rauhut (2003) because he considered the presence of 16 or more transverse processes the plesiomorphic

condition, and smaller numbers as derived conditions. Gauthier (1986) does not mention the number of vertebrae that have transverse processes in the plesiomorphic condition, just indicating that theropods have less than half vertebrae with this structure, and that transverse processes are restricted to the base of the tail in Maniraptora. Considering the definition made by Gauthier (1986), the condition of *Staurikosaurus* would be derived if it had 20 of 47 vertebrae with transverse processes, but our results point that it has the plesiomorphic condition, with more than half vertebrae having this structure.

These observations suggest that the correct positioning of preserved vertebrae may be important to define a character state on some taxa that do not have a complete series of vertebrae, because characters, other than presence of transverse processes, may vary along vertebral series and may be included in systematic analyses.

Considering the dimensions of other skeletal elements (mandible and cervical, dorsal, and sacral vertebrae), it was estimated that *Staurikosaurus* had a total body length of 220–225 cm, from which 60% would correspond to the tail. By comparison, in *Herrerasaurus* the tail corresponds to 54%, which was calculated from the skeletal reconstruction presented by Sereno (1993). The total body length estimated here is much larger than it was supposed by Colbert (1970), who proposed a length of 160 cm for *Staurikosaurus*. However, it is close to what was estimated by Paul (1988). He estimated that *Staurikosaurus* was 208 cm long, and the tail would have 125 cm (in contrast to ~135 cm in the present work), so that the 10 cm difference in tail length from the value presented in this study accounts for most of the divergence.

Other length estimation methods (Christiansen 1999, Therrien and Henderson 2007), which are based on the length of the femur, tibia or head (here estimated from the mandible length), suggest that *Staurikosaurus* had a total length varying from 1.99 to 2.57 m. All these methods are based on data from neotheropod dinosaurs, but the present results fit well in the middle of this range, and may confirm the validity of these methodologies for a rough estimation of the total length of theropods and some basal saurischians, such as herrerasaurids.

The methodology employed here to estimate the missing data in the holotype of *Staurikosaurus* may be a useful way of estimating the total body length as well as parts of the body of other fossil vertebrates. The lack of complete vertebral series is extremely common, and an empirical methodology for the estimation of missing data is necessary to obtain better and more precise measurements.

Despite the benefits of the use of our methodology, it is important to point out that it has limitations and requirements to be properly applied:

- 1) The problem of determining the total number of vertebrae is the first limitation of the technique. This assumption is necessary and needs to be based on closely related taxa and, consequently, represents a consideration that may hide possible variation among taxa;
- 2) The preserved vertebrae must comprise the anterior and most posterior vertebrae, but the necessary number in each sequence is uncertain and must be determined by the analysis of complete specimens from the most exclusive supraspecific taxon that includes the specimen being analyzed. Our analyses with *Staurikosaurus* indicated regressions based on height data including just the posterior sequence (19 or even just the 17 posteriormost vertebrae), provided accurate estimations for the preserved anterior vertebrae. However, this was not observed with length data. In this case it was necessary to include data on at least three anterior vertebrae to obtain a regression compatible with the results suggested by height values;
- 3) Deviation of dimensions relative to predicted values contributes to reduce or increase the accuracy of the methodology. As deviation increases, it also increases the number of vertebrae that must be preserved in order to obtain accurate results. We observed that deviations in length values are much larger than that of heights in *Staurikosaurus*, especially in the posterior sequence, and the results were less precise with the former data.

We believe that each specific or supraspecific taxon will have different limitations, since variation among vertebrae is not constant and also the type of regression

equation may vary from one taxon to the other. Variation in length values should not be necessarily described by a fourth order polynomial equation. If a binomial equation describes accurately length values and deviation is small, data from just posterior vertebrae may be sufficient, even though the analysis with other taxa is necessary to evaluate this supposition.

ACKNOWLEDGMENTS

The Museum of Comparative Zoology is thanked for the loan of the *S. pricei* holotype to Alexander Kellner, Museu Nacional/UFRJ. This research is part of the Graduation Monograph of one of the authors (ONG) at the Museu Nacional, Universidade Federal do Rio de Janeiro, funded by a fellowship of the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and of the Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ). The Monograph was examined by doctors Leila Maria Pessoa, Luciano Artemio Leal, Oscar Rocha Barbosa, Sergio Potsch de Carvalho e Silva, who are thanked, together with an anonymous referee, for providing useful comments that greatly improved the final version of this paper. We also thank Jonathas Bittencourt for the caudal vertebrae photos used in Figure 1 and Tiago Simões for the revision of the English text.

RESUMO

Dados lacunares são um problema comum na paleontologia. Eles dificultam a reconstrução acurada de táxons extintos e limitam a inclusão de alguns táxons em estudos comparativos e biomecânicos. Particularmente, estimar a posição de vértebras em séries incompletas tem sido feito com base em métodos não empíricos que não permitem estimar corretamente as partes ausentes. Neste trabalho apresentamos uma metodologia que permite estimar a posição de sequências médias preservadas de vértebras caudais no dinossauro saurísquio *Staurikosaurus pricei*, com base no comprimento e altura dos centros das vértebras anteriores e posteriores preservadas. Equações de regressão foram usadas para estimar essas dimensões para as vértebras médias e, conseqüentemente, para posicionar as sequências médias preservadas e para estimar o tamanho das vértebras não preservadas. Os resultados indicam que as vértebras caudais preservadas de *Staurikosaurus* corresponderiam

às posições 1-3, 5, 7, 14-19/15-20, 24-25/25-26 e 29-47, e que pelo menos 25 vértebras possuíam processos transversos. O comprimento total da cauda foi estimado em 134 cm e o comprimento total do corpo em 220-225 cm.

Palavras-chave: dados lacunares, vértebra caudal, regressão, *Staurikosaurus pricei*.

REFERENCES

- ALCOBER OA AND MARTINEZ RN. 2010. A new herrerasaurid (Dinosauria, Saurischia) from the Upper Triassic Ischigualasto Formation of northwestern Argentina. *ZooKeys* 63: 55–81.
- ALLEN V, PAXTON H AND HUTCHINSON JR. 2009. Variation in center of mass estimates for extant sauropsids and its importance for reconstructing inertial properties of extinct archosaurs. *Anat Rec* 292: 1442–1461.
- ANDERSON JS. 2001. The phylogenetic trunk: maximal inclusion of taxa with missing data in an analysis of the Lepospondyli (Vertebrata, Tetrapoda). *Syst Biol* 50: 170–193.
- BITTENCOURT JS AND KELLNER AWA. 2009. The anatomy and phylogenetic position of the Triassic dinosaur *Staurikosaurus pricei* Colbert, 1970. *Zootaxa* 2079: 1–56.
- BOHONAK AJ AND LINDE K. 2004. RMA: Software for Reduced Major Axis regression, Java version. <http://www.kimvdlinde.com/professional/rma.html>.
- CHRISTIANSEN P. 1999. Long bone scaling and limb posture in non-avian theropods: Evidence for differential allometry. *J Vert Paleontol* 19: 666–680.
- COLBERT EH. 1970. A saurischian dinosaur from the Triassic of Brazil. *Am Mus Novit* 2405: 1–60.
- CZERKAS SJ. 1997. Reconstruction and restoration. In: CURRIE PJ AND PADIAN K (Eds), *Encyclopedia of Dinosaurs*, San Diego: Academic Press, California, USA, p. 626–629.
- GALTON PM. 1977. *Staurikosaurus pricei*, an early saurischian dinosaur from the Triassic of Brazil, with notes on the Herrerasauridae and Poposauridae. *Palaeontol Z* 51: 234–245.
- GALTON PM. 2000. Are *Spondylosoma* and *Staurikosaurus* (Santa Maria Formation, Middle-Upper Triassic, Brazil) the oldest saurischian dinosaurs? *Palaeontol Z* 74: 393–423.
- GATESY SM. 1990. Caudofemoral musculature and the evolution of theropod locomotion. *Paleobiology* 16: 170–186.

- GAUTHIER J. 1986. Saurischian monophyly and the origin of birds. In: PADIAN K (Ed), *The origin of birds and the evolution of flight*. San Francisco: Mem California Acad Sci 8: 1–55.
- GUNGA H-C, SUTHAU T, BELLMANN A, FRIEDRICH A, SCHWANEBECK T, STOINSKI S, TRIPPEL T, KIRSCH K AND HELLWICH O. 2007. Body mass estimations for *Plateosaurus engelhardti* using laser scanning and 3D reconstruction methods. *Naturwissenschaften* 94: 623–630.
- HAMMER Ø, HARPER DAT AND RYAN PD. 2001. PAST: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontologia Electronica* 4(1): 9. http://palaeo-electronica.org/2001_1/past/issue1_01.htm.
- HOLLIDAY CM, RIDGELY RC, SEDLMAYR JC AND WITMER LM. 2010. Cartilaginous epiphyses in extant archosaurs and their implications for reconstructing limb function in dinosaurs. *PLoS ONE* 5(9): e13120. doi:10.1371/journal.pone.0013120.
- HOLZ M AND SCHULTZ CL. 1998. Taphonomy of the south Brazilian Triassic herpetofauna: fossilization mode and implications for morphological studies. *Lethaia* 31: 335–345.
- HUTCHINSON JR, ANDERSON FC, BLEMKER SS AND DELP SL. 2005. Analysis of hindlimb muscle moment arms in *Tyrannosaurus rex* using a three-dimensional musculoskeletal computer model: implications for stance, gait, and speed. *Paleobiology* 31: 676–701.
- HUTCHINSON JR, NG-THOW-HING V AND ANDERSON FC. 2007. A 3D interactive method for estimating body segmental parameters in animals: Application to the turning and running performance of *Tyrannosaurus rex*. *J Theor Biol* 246: 660–680.
- KEARNEY M AND CLARK JM. 2003. Problems due to missing data in phylogenetic analyses including fossils: a critical review. *J Vert Paleontol* 23: 263–274.
- LANGER MC. 2004. Basal Saurischia. In: WEISHAMPEL DB, DODSON P AND OSMÓLSKA H (Eds). *The Dinosauria*, 2nd ed., Berkeley: University of California Press, p. 25–46.
- LANGER MC AND BENTON MJ. 2006. Early dinosaurs: a phylogenetic study. *J Syst Palaeontol* 4: 309–358.
- MADDISON WP. 1993. Missing data *versus* missing characters in phylogenetic analysis. *Syst Biol* 42: 576–581.
- MILLER JA AND GROSS MM. 1998. Locomotor advantages of Neandertal skeletal morphology at the knee and ankle. *J Biomech* 31: 355–361.
- NIXON KC AND DAVIS JI. 1991. Polymorphic taxa, missing values and cladistic analysis. *Cladistics* 7: 233–241.
- NORELL MA AND WHEELER WC. 2003. Missing entry replacement data analysis: a replacement approach to dealing with missing data in paleontological and total evidence data sets. *J Vert Paleontol* 23: 275–283.
- NOVACEK MJ. 1992. Fossils, Topologies, Missing Data, and the Higher Level Phylogeny of Eutherian Mammals. *Syst Biol* 41: 58–73.
- NOVAS FE. 1993. New information on the systematics and postcranial skeleton of *Herrerasaurus ischigualastensis* (Theropoda: Herrerasauridae) from the Ischigualasto Formation (Upper Triassic) of Argentina. *J Vert Paleontol* 13: 400–423.
- NOVAS FE. 1996. Dinosaur monophyly. *J Vert Paleontol* 16: 723–741.
- NOVAS FE. 1997. Herrerasauridae. In: CURRIE PJ AND PADIAN K (Eds), *Encyclopedia of Dinosaurs*, San Diego: Academic Press, California, USA, p. 303–311.
- PAUL GS. 1987. The science and art of restoring the life appearance of dinosaurs and their relatives: a rigorous how-to guide. In: CZERKAS SJ AND OLSEN EC (Eds), *Dinosaurs past and present*. Seattle: University of Washington Press, USA, p. 4–49.
- PAUL GS. 1988. *Predatory dinosaurs of the world: A complete illustrated guide*. New York: Simon & Schuster, 464 p.
- RAUHUT O. 2003. The interrelationships and evolution of basal theropod dinosaurs. *Spec Papers Palaeontol* 69: 1–213.
- ROMER AS. 1956. *Osteology of the Reptiles*. Chicago: Univ Chicago Press, 772 p.
- SERENO PC. 1993. The pectoral girdle and forelimb of the basal theropod *Herrerasaurus ischigualastensis*. *J Vert Paleontol* 13: 425–450.
- SERENO PC. 1999. The evolution of dinosaurs. *Science* 284: 2137–2147.
- SERENO PC AND NOVAS FE. 1993. The skull and neck of the basal theropod *Herrerasaurus ischigualastensis*. *J Vert Paleontol* 13: 451–476.
- THERRIEN F AND HENDERSON DM. 2007. My theropod is bigger than yours... or not: Estimating body size from skull length in theropods. *J Vert Paleontol* 27: 108–115.
- WILLISTON SW. 1925. *The osteology of the reptiles*. Cambridge: Harvard University Press, 304 p.