Acta Scientiarum

http://periodicos.uem.br/ojs
ISSN on-line: 1807-863X
Doi: 10.4025/actascibiolsci.v46i1.68981

ZOOLOGY

Comparison between different methods of estimating the developmental stages in a synanthropic scorpion *Tityus stigmurus* (Thorell, 1876)

Kleber Agari Campos[1*], Eunice Aparecida Bianchi Galati[2] and André Felipe Araujo Lira[3]

[1]Laboratório de Identificação e Pesquisa de Fauna Sinantrópica, Divisão de Vigilância de Zoonoses, Prefeitura de São Paulo, 02031-020, São Paulo, São Paulo, Brazil. [2]Departamento de Epidemiologia, Faculdade de Saúde Pública, Universidade de São Paulo, São Paulo, São Paulo, Brazil. [3]Colécción Nacional de Arácnidos, Universidad Autónoma Nacional de Mexico, Coyoacán, Ciudad de Mexico, Mexico. *Author for correspondence. E-mail: klebercampos@prefeitura.sp.gov.br

ABSTRACT. Determining developmental stages is an important tool for monitoring populations of medically important arthropods, including some scorpions species. Therefore, prediction tools for scorpion developmental stages may provide useful information for public health. Thus, this study compared different methods to estimate the developmental stages of the alien and medically important scorpion species *Tityus stigmurus* collected in São Paulo. We tested the applicability of the following morphometric methods: grouping (K-means and Ward's), theoretical, and distribution of frequencies for instar estimation in field-caught *T. stigmurus* individuals. The methods showed similar results up to the 5th instar. Our results corroborate the literature on the non-objectivity of the hierarchical clustering method in determining the exact number of size classes in the sample. We also found that the hierarchical and non-hierarchical clustering methods diverged regarding the two largest size classes. The peaks in the frequency curve of the carapace length and the theoretical calculation using a growth factor of 1.26 proved useful in determining the stages of development. Seven size classes were demonstrated, which provide evidence that *T. stigmurus* individuals can reach maturity at the 6th or 7th instar.

Keywords: Buthidae; arachnid; ontogeny; morphometry; public health

Received on July 20, 2023.
Accepted on December 06, 2023.

Introduction

Studies on the life cycle of animals are essential to understand the population dynamics of species, including those of medical importance. The immature stages in arthropods grow following geometric progression, increasing in size at a constant rate (e.g., Parra & Haddad, 1989; Francke & Sissom, 1984; Ambrosano, Igue, & Lourenção, 1997), with linear dimensions that grow at the rate of $\sqrt[3]{2} = 1.26$ until adulthood (Przibram & Megušar, 1912; Ambrosano et al., 1997). Thus, the developmental phase can be determined by applying morphometric techniques. In addition to recognition of species, a variety of biological studies raise morphometric questions, such as those related to life stages, allometric relations and population characterization (Marcus, 1990; Dujardin, 2008). Morphometric studies in scorpions can use bivariate scatter plots in order to observe the level of aggregation and gaps between points and infer instars and moults, respectively (Polis & Sissom, 1990). However, the arbitrariness in interpreting the results can compromise the final result because a very high number of individuals in the sample can make the clusters not visible in the graph (Francke & Sissom, 1984). However, few individuals may not make possible the perception of some instar (Polis & Sissom, 1990).

In some groups of arthropods, such as scorpions, the maintenance of captivity for the determination of post-embryonic development is hampered by a long time to reach adulthood (greater than one year) and high mortality in the laboratory (e.g., Francke & Sissom, 1984; Polis & Sissom, 1990; De Souza, Santana-Neto, Lira, & Albuquerque, 2016). Some scorpion species are responsible for accidents that can affect the daily routine of victims, even causing death (Albuquerque, Santana Neto, Amorim, & Pires, 2013; Isbister & Bawaskar, 2014; Coorg, Levitan, Gerkin, Muenzer, & Ruha, 2017; Ward, Ellsworth, & Nystrom, 2018). Most species are capable of inhabiting urban areas increasing their human encounter chances (Ward et al., 2018; Amado et al., 2021; Guerra-Duarte et al., 2023). Thus, scorpion envenomation is a public health concern in several countries, including Brazil (Brasil, 2009; Ward et al., 2018).

In Brazil, the primary species causing accidents belong to the genus *Tityus* Koch, 1836, especially *T. serrulatus* Lutz & Mello, 1922, *T. bahiensis* (Perty, 1833), *T. obscurus* (Gervais, 1843), and *T. stigmurus* (Thorell, 1876) (Brasil, 2009). The latter is a facultative parthenogenetic species (Ross, 2010; Foerster, Dionisio-da-Silva, Santos, Albuquerque, & Lira, 2021) with class III envenoming capacity (Lira-da-Silva, Amorim, & Brazil, 2000; Ward et al., 2018). *Tityus stigmurus* species was originally distributed in the states of Pernambuco, Bahia, Ceará, Paraíba, Alagoas, Rio Grande do Norte and Sergipe (Brasil, 2009; De Souza, Candido, Lucas, & Brescovit, 2009). However, in 2012, individuals of *T. stigmurus* were recorded in São Paulo, with three currently identified populations established in the city, two of them probably introduced by fruit transport from the northeast of Brazil (Bertani et al., 2018).

Furthermore, ontogenetic differences have been observed in the potency and composition of venom according to scorpion developmental stages (McElroy et al., 2017). These differences may directly impact the symptomatology of accidents caused by scorpions; thus, information on specimen developmental stages is relevant to public health. Additionally, determining the developmental stage of these arachnids collected in the field is often impossible. Therefore, this study aimed used various morphometric approaches to estimate the specimen development parameters of *T. stigmurus* collected in São Paulo City. The results are expected to provide subsidies for the development of a tool that will help determine the ontogeny from field caught *T. stigmurus* specimens.

Material and methods

Scorpions collection and morphometry

Tityus stigmurus specimens were collected in urban settlements during monitoring activities in the region corresponding to the Regional Prefecture of Penha (23°32'19"S 46°30'29"W) from September to December 2018, February to September 2019, November 2019, and May 2021. A total of 199 specimens were used (84 females and 115 juveniles), of which 179 (82 females and 97 juveniles) were used for ontogeny estimation using the methods described below. The other 20 specimens corresponded to two adult females that were kept in the laboratory until reproduction, generating 12 and 6 offspring, respectively. Offspring specimens were observed until the first molt. These 20 specimens were analyzed according to the theoretical method described below, and the results were used as comparison parameters for the 179 specimens collected in the field. The specimens were euthanized in the freezer at -18 °C. The lengths of the left pedipalp movable finger (MF), fifth segment of the metasoma (MetV), and the carapace (Car) of each individual were measured. Measurements were performed *post-mortem* using an analog caliper of 150 mm, with a precision of 0.02 mm, in a stereoscopic magnifying glass (Stahnke, 1970; Sissom, Polis, & Watt, 1990) (Figure 1). The specimen used in this study were identified in accordance with the proposal of Lourenço (2002) and deposited in the *Coleção de Fauna Sinantrópica* of the Municipality of São Paulo (CFS-MSP) of the *Laboratório de Identificação e Pesquisa em Fauna Sinantrópica da Divisão de Vigilância de Zoonoses*.

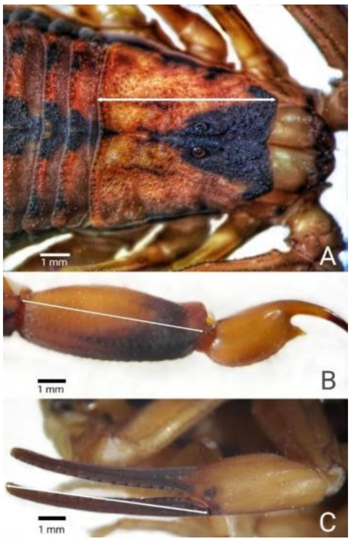

Figure 1. Standardized axes to obtain the dimensions of the structures used in the morphometric analysis of *Tityus stigmurus*. A= carapace, B= metasoma segment V and C= movable finger.

Data analysis

Grouping methods

The size classes of *T. stigmurus* were estimated using different morphometric analyses. Two methods were used for clustering: K-means and Ward's. The first analysis corresponds to a non-hierarchical method that performs grouping by partition, recommended when the number of groups to be formed is known (Moori, Marcondes, & Ávila, 2002). The procedure requires the establishment of arbitrary centroids from which objects are moved between groups according to their distance from each of these centroids, which are again determined based on the average values of the elements that make up the cluster of a given conglomerate momentarily. The distances are recalculated, elements are reclassified and moved between the groups, and so on, until the objects stop migrating between the conglomerates, at the moment when the process reaches stability (Manly, 2008). Six clusters (G2 to G7) were determined based on the maximum number of instars after the first moult, as described for the species in the literature (De Souza et al., 2016). However, Ward's method is a hierarchical and forms groups from the minimization of the sum of squares of the deviations of each of them (Tomaz, Matos, & Souza, 2017). The result is presented in the form of a dendrogram (Manly, 2008). To verify the robustness of the clusters, the results of the two methods were subjected to the kappa concordance coefficient test. Landis and Koch (1977) classify the levels of agreement as "non-existent" ($\kappa < 0$); "minimal" ($\kappa = 0–0.20$); "reasonable" ($\kappa = 0.21–0.40$); "moderate" ($\kappa = 0.41–0.60$); "substantial" ($\kappa = 0.61–0.80$) and "perfect" ($\kappa = 0.81–1.0$). Validation of the data obtained at the cluster analysis was performed by calculating the progression factor by applying the adjusted Rand index (ARI). This index determines the similarity between two parcels – P1 and P2, examining the pairs of parcels belonging to the two groups (Albuquerque & Barros, 2020). According to these authors, if the index is one (1), the two partitions are identical. Zero (0) indicates that the agreement occurred by chance. When it assumes negative values, it indicates that the agreement is lower than that given by chance. To compare the mean values of the three structures between the two methods, the Mann–Whitney's test (U test) was used because the normal distribution of the data was not observed.

Theoretical method

A theoretical progression factor of 1.26 was used for the calculation, which was admitted for scorpions (Francke & Sissom, 1984). The presumed number of instars was calculated according to Francke and Sissom (1984) using the equation $n = (logA - logY) \div logP$, where n corresponds to the number of juveniles, A is the dimension of the structure of the adult specimen, Y is the dimension of the same structure in a young specimen of a known instar, and P corresponds to the progression factor. The theoretical method was validated by multiple linear regression, observing the coefficient of determination (R^2) to estimate the percentage of total variability explained by the model. Partial regression coefficients were used to verify the structure that explained most of the variation (Ayres, Ayres Júnior, Ayres, & Santos, 2000). The mean values of the parameters obtained by the theoretical method were compared with those described in the literature for the species using the Student's t-test.

Determination of instars by the frequency distribution curve of the carapace length measurements

The distribution frequencies of the carapace measurements were obtained using multimodal curves, where the peaks were interpreted as possible instars (Ecole, Anjos, Michereff-Filho, & Picanço, 1999). Thus, four frequency classes were defined per peak from visual inspection of the curve, testing intervals between 0.1 and 0.2 mm, until satisfactory resolution of the histogram was obtained (Logan, Bentz, Vandygriff, & Turner, 1988; Chen & Seybold, 2013). The limits between the classes corresponding to the instars were established by the lowest points separating the peaks (Chen & Seybold, 2013) and by fitting the normal distribution of frequencies as much as possible using Microsoft Excel 2019 mass probability function.

The criterion for rejecting the hypothesis was the overlapping with the limits of the confidence interval ($p < 0.05$) of the means between neighbouring groups. The probability of error in the interpretation and determination of instars was calculated using the reduced normal distribution (standardized Z variable), where the normal curves of the size classes were reduced to standard curves with mean = 0 and deviation = 1 (Maia, Morais, & Oliveira, 2001). Tabulated values for Z were obtained. The assumption when using this approach is that the data have a normal distribution (Maia et al., 2001; Camacho, Silveira, Camargo, & Natale, 2012). Shapiro–Wilk's or Lilliefours' tests were used to verify the normality of the distributions, depending on the number of specimens in each size class (Ayres et al., 2000).

To compare the results obtained with data from the literature on specimens observed in the laboratory, the growth factor between the average values of the size classes was calculated by dividing the average value of the structure in a given size class by the corresponding average value of the same structure in the previous size class, according to Dyar (1890). The Z-test was used to compare the averages of the Carapace length between the theoretical method and the distribution of frequency classes in a multimodal curve. The Z test was chosen because the variances differed between the methods for one of the size classes, in addition to the two of the six classes that contained more than 30 samples (Ayres et al., 2000). All statistical analyses were performed using Bioestat v. 5.3 and Past v. 2.16 software.

Figure 2 illustrates the data analysis, applying the proposed methods according to the specimen categories used in this study (observed in the laboratory × field).

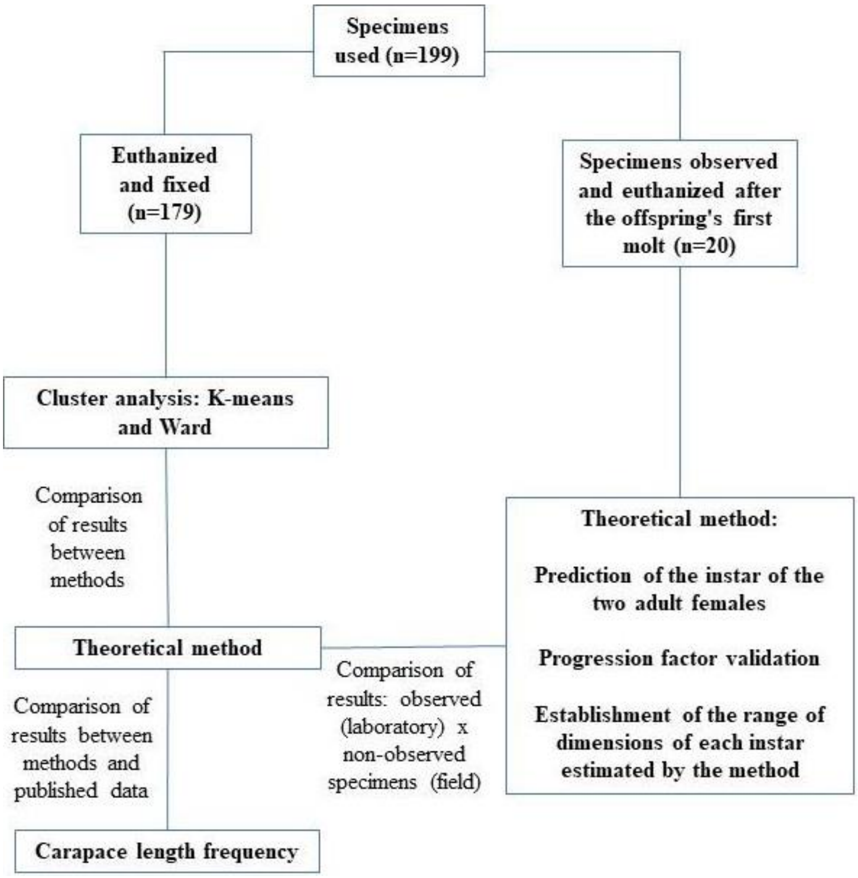

Figure 2. Flowchart of the data analysis carried out by applying the methods to specimens of *Tityus stigmurus* specimens used in the present work.

Results

Grouping methods

Convergence by K-means resulted in a grouping of 41 specimens, with the mean values of the analyzed structures approaching those of the individuals corresponding to instar 2. The other group contained 18 individuals compatible with instar 3. The following groups comprised 23, 18, 46, and 33 individuals compatible with instars 4, 5, 6, and 7, respectively. The Ward method gathered the same 41 individuals in one of the groups (Group 2): 19 individuals in Group 3, 28 in Group 4, 12 in Group 5, 61 in Group 6, and 18 in Group 7 (Figure 3).

Comparing the clusters formed by the two methods, it was possible to verify that there was correspondence in the clusters for 157 of the 179 specimens collected in the field (87.7%), and in group 2, there was 100% agreement (Figure 3 and Table 1).

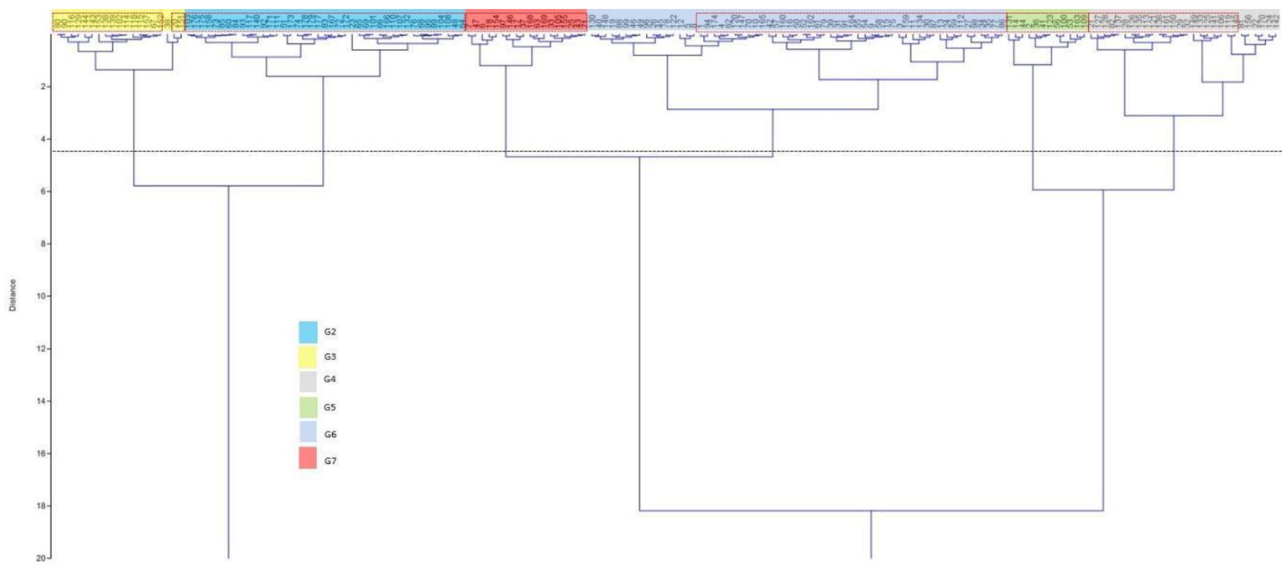

Figure 3. Dendrogram for the 179 specimens of *Tityus stigmurus* grouped by the Ward algorithm. The dashed polygons in red indicate the specimens consistent with the groupings by K-means. The distance selected for cutting the branches is indicated by the black dotted line.

Table 1. Level of agreement between the clusters of specimens (G2 to G7) originated from the K-means and Ward methods.

		K-means						
		G2	G3	G4	G5	G6	G7	Total
Ward	G2	41						41
	G3		18	1				19
	G4			22	6			28
	G5				12			12
	G6					46	15	61
	G7						18	18
	Total	41	18	23	18	46	33	179

The smallest divergence observed between the methods occurred between Groups 3 and 4 (only one specimen). The largest divergence occurred between Groups 6 and 7 (15 individuals). Perfect agreement (κ = 0.95) was observed between Groups 3 and 4. The agreement was substantial (κ = 0.69) in Groups 4 and 5 and moderate (κ = 0.58) among Groups 6 and 7. The average dimensions of each structure for each clustering algorithm are listed in Table 2.

Table 2. Mean and standard deviation of Carapace length (Car), fifth metasoma segment (MetV), and movable finger (MF) of *Tityus stigmurus* by size grouping/class and method (in mm).

Method	Parameters	II		III		IV		V		VI		VII	
		$\overline{X}$	σ	$\overline{X}$	σ	$\overline{X}$	σ	$\overline{X}$	σ	$\overline{X}$	σ	$\overline{X}$	σ
K-means	Car	2.35	0.16	3.08	0.14	3.96	0.19	4.86	0.26	5.84	0.17	6.21	0.23
	MetV	2.23	0.15	3.04	0.15	4.02	0.22	5.03	0.31	6.19	0.15	6.71	0.24
	MF	3.02	0.13	3.98	0.16	5.03	0.21	6.17	0.23	7.33	0.25	7.86	0.19
Ward	Car	2.35	0.16	3.10	0.17	4.11	0.29	5.01	0.17	5.89	0.16	6.35	0.21
	MetV	2.23	0.15	3.07	0.19	4.19	0.33	5.19	0.24	6.27	0.20	6.86	0.18
	MF	3.02	0.13	4.01	0.20	5.24	0.40	6.28	0.19	7.42	0.27	7.99	0.09
Theoretical	Car	2.34	0.15	3.03	0.11	3.86	0.21	4.70	0.27	5.93	0.24	6.61	0.15
	MetV	2.22	0.13	2.98	0.14	4.13	0.32	4.84	0.27	6.33	0.30	7.12	0.09
	MF	3.01	0.11	3.91	0.15	5.18	0.39	5.98	0.30	7.48	0.36	8.10	0.07

The U test presented a significant difference between the K-means and Ward's algorithms for the average dimensions of MetV in the G6 cluster ($p = 0.029$) and the three parameters in the G7 cluster: Car ($p = 0.009$), MetV ($p = 0.007$), and MF ($p = 0.007$). To some extent, overlap with the extreme values of the Carapace (Car) and mobile finger (MF) was observed between the groups, which varied depending on the method used. The K-means showed the highest percentage of specimens with overlap (17%) in one or both structures between G6 and G7. In the Ward method, the overlap, although smaller (5%), was distributed from G4 to G7.

The application of the theoretical method to predict the instars of the specimens collected in the field generated size classes similar to those of clusters up to G5 (Table 2). The correspondence percentage between the results of the theoretical method and the clustering algorithms was 79.3% (142 samples) and 83.8% (150 samples) for K-means and Ward, respectively (Tables 3 and 4).

Table 3. Correspondence between the results obtained by the theoretical method and the cluster analysis by the K-means method.

Grouping	Size class (Theoretical method)						Total
K-means	II	III	IV	V	VI	VII	
G2	40	1					41
G3		16	2				18
G4			20	3			23
G5				15	3		18
G6					46		46
G7					28	5	33
Total	40	17	22	18	77	5	179

Table 4. Correspondence between the results obtained by the theoretical method and the cluster analysis by the Ward algorithm.

Grouping	Size class (Theoretical method)						Total
Ward	II	III	IV	V	VI	VII	
G2	40	1					41
G3		16	3				19
G4			19	9			28
G5				9	3		12
G6					61		61
G7					13	5	18
Total	40	17	22	18	77	5	179

The discrepancies observed between the results obtained by the K-means and the theoretical methods were 2.4 (G2), 11.1 (G3), 13 (G4), and 16.7% (G5). Considering Ward's algorithm, the percentages were 2.4 (G2), 15.8 (G3), 32.1 (G4), and 25% (G5).

The sharpest discrepancy was concentrated in the G6 and G7 clusters for both algorithms compared with the theoretical method, with 28 examples (84.8%) being reclassified from the G7 cluster to size class 6 in the case of K-means and 13 (72.2%) in the same cluster as the Ward algorithm. The adjusted Rand index (ARI) values for validating the clusters in comparison to the size classes calculated by the theoretical method were 0.392 (Ward) and 0.299 (K-means).

The number of moults estimated by the theoretical method, using the log value average of the field specimen structures in the largest size class concerning the average of those located in the lower limit of the sample, was 5. Therefore, after five moults, the smallest specimens of the sample would reach the size of the largest. As the data was collected in specimens that had already passed the 1st instar, it was assumed that 2 (sample instar) + 5 (number of moults) = 7 instars. Considering the 20 specimens observed in the laboratory and the structures individually, we found n = 4.1 and 4.2 for two of them (Car and MF). Regarding MetV, the values obtained for n were 4.6 and 4.7 (Table 5). The values were rounded to the nearest whole number: Car, MF = 4, and MetV = 5.

Table 5. Prediction of the moults number (n) of the specimens observed in the laboratory using the equation n = (logX – logY) ÷ logP. Car: Carapace, MetV: metasoma, MF: movable finger. X: length of the structure in adult females A and B, $\bar{Y}$: average length of the same structure in the offspring, P: growth factor (1.26).

	A			B			$\bar{Y}$	
	mm	*log*	*n*	*mm*	*log*	*n*	*mm*	*log*
Car	6.08	0.784	4.2	6.02	0.779	4.1	2.34	0.368
MetV	6.6	0.819	4.7	6.4	0.806	4.6	2.22	0.345
MF	7.76	0.890	4.1	7.74	0.889	4.1	3.01	0.478
	$\bar{n}$		4.33	$\bar{n}$		4.27		

Using the mean of the structures as a reference, n = 4.33 and 4.27 were obtained for females A and B, respectively, in relation to the offspring. Therefore, it is estimated that four seedlings are necessary for the offspring to reach the size of their mothers. By deduction, the two females in relation to the offspring were in the 6th instar: 2 (instar known of the offsprings) + 4 (number of seedlings) = 6. Considering the dimensions of the structures of instar 2 specimens observed in the laboratory and assuming a progression factor of 1.26, it was verified that females A and B would belong to instar 7 in the case of MetV. Analogous to the prediction of the number of seedlings, the calculation showed that the dimensions of Car and MF were within the estimated amplitude for instar 6 in the case of both structures (Table 6).

Table 6. Calculated amplitude of the dimensions of the three structures of *Tityus stigmurus* from the specimens of instar 2 observed in the laboratory; assuming the progression factor 1.26. Car: Carapace. MetV: metasoma. MF: movable finger (in mm).

Instar	Car		MetV		MF	
II	2.08	2.44	2.02	2.32	2.8	3.08
III	2.62	3.07	2.55	2.92	3.53	3.88
IV	3.30	3.87	3.21	3.68	4.45	4.89
V	4.16	4.88	4.04	4.64	5.60	6.16
VI	5.24	6.15	5.09	5.85	7.06	7.76
VII	6.61	7.75	6.42	7.37	8.89	9.78
Female A	6.08		6.6		7.76	
Female B	6.02		6.4		7.74	

The mean values of the structures of the 179 field specimens obtained using the theoretical method for each size class were compared with the assumed amplitude by calculating the progression factor applied to the 20 specimens observed in the laboratory.

In some cases, the length of MetV (instars 3, 4, 5, and 6) and MF (instars 3, 4, and 7) was intermediate between the size classes. However, the Carapace length was within the limits of the estimated range for all of them (Tables 2 and 6, respectively). The multiple linear regression model between the analyzed structures (independent variables x) and size classes (dependent variable y) obtained by the theoretical method applied to field samples displayed a highly positive relationship ($R^2 = 0.975$, $p = 0$). The predictor variable Carapace (Car) had the highest statistical significance (t = 3.899, $p = 0.0001$) among the partial regression coefficients. The t-test indicated that the average dimensions of the structures obtained by the theoretical method were significantly higher than those described in the literature for the species by De Souza et al. (2016), except for Car at instar 2 ($p = 0.040$) and instar 7 ($p = 0.022$), and MetV at instar 3 ($p = 0.054$), instar 4 ($p = 0.197$), and instar 5 ($p = 0.4425$).

The frequency distribution curve using measurements of the Carapace length of the specimens showed six peaks, in agreement with the presumed number of instars for the species (Figure 4).

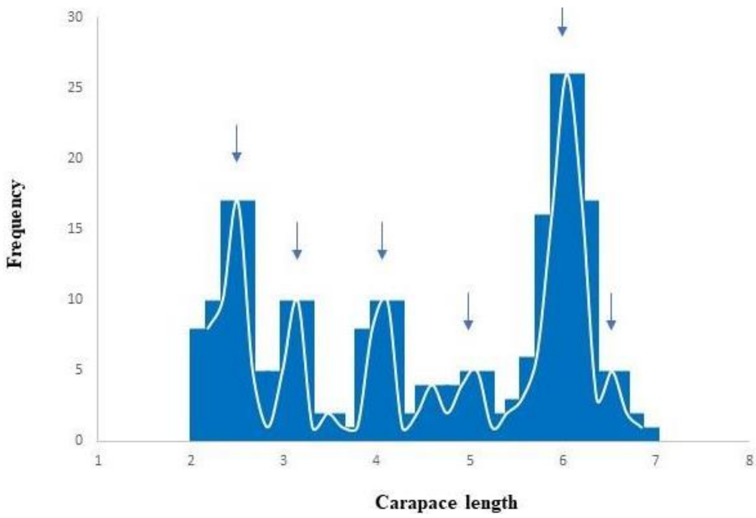

Figure 4. Multimodal distribution of the Carapace length (in mm) of *Tityus stigmurus* individuals showing six peaks, suggesting the stages of development (instars).

Adjusting the limits of the frequency classes by the normal distribution produced results close to those obtained by the theoretical method, with no overlap of the confidence intervals between the peaks being observed (Table 7). All frequency classes were normally distributed ($p > 0.05$).

Table 7. Number of instars, amplitude, and confidence interval of the frequency classes for the Carapace length of *Tityus stigmurus* (in mm).

Instar	Range	Mean	CI (p< 0.05)
II	2.02 - 2.81	2.35	2.31 - 2.40
III	2.82 - 3.61	3.06	3.01 - 3.11
IV	3.62 - 4.41	3.94	3.89 - 3.99
V	4.42 - 5.21	4.84	4.71 - 4.97
VI	5.22 - 6.41	5.95	5.90 - 5.99
VII	>6.42	6.56	6.45 - 6.67

The values obtained for the carapace length between the two methods did not differ significantly (Table 8).

Table 8. Z test for comparison of the average carapace length of *Tityus stigmurus* (in mm) by size classes, obtained by the theoretical method and frequency distribution in a multimodal curve. Decision level: one-sided α = 0.01. GF: Growth factor ($i_{i+1} \div i$).

Methods	Size class					
	II	III	IV	V	VI	VII
Theoretical	2.34±0.15	3.03±0.11	3.86±0.21	4.70±0.27	5.93±0.24	6.61±0.15
	GF	1.29	1.28	1.22	1.26	1.12
Frequency	2.35±0.16	3.06±0.10	3.94±0.11	4.84±0.31	5.95±0.18	6.56±0.16
	GF	1.30	1.29	1.23	1.23	1.10
Z (unilateral)	p = 0.3864	p = 0.2850	p = 0.0580	p = 0.0611	p = 0.2750	p = 0.2711

The calculated probability of error in classifying a given instar into an above- or below size class using the carapace length frequency distribution was less than 1% in most cases, except in the case of classes 5 to 6 (1.49%), with the highest percentage related to class 7 in relation to 6 (15.7%) (Table 9).

Table 9. Probability (%) of error in the determination of *Tityus stigmurus* instars based on the method of distributing the frequencies of Carapace length.

Instar *i*	Instar *i* to *i -1*	Instar *i* to *i + 1*
II	-	0.032
III	0.15	0.0
IV	0.0	0.004
V	0.1	1.49
VI	0.41	0.52
VII	15.7	-

Discussion

In this study, we evaluated the effectiveness of different morphometric methods for determining the size classes using the *Tityus stigmurus* scorpion as a model. Our results indicated that the K-means and Ward clustering methods had a correspondence percentage of 87.7%. Similarly, Seidel, Moreira, Ansuj, and Noal (2008) used both methods and observed correspondence of 82.54% considering the highly concordant results, validating the robustness of the clusters formed by both. Thus, we can consider that the high correspondence found in this work represents a robust cluster.

The calculation for predicting the maximum number of instars (theoretical method) presented a similar result to the cluster analysis, observing a high percentage of correspondence for the size classes. However, the adjusted Rand index showed that the groupings by the K-means and Ward methods, when compared with the expected results calculated by the theoretical method, presented a weak validation level because of the high percentage of the discrepancy between the largest size classes (G6 and G7). These results suggest that both methods had greater difficulty forming these two clusters from the dataset. The allometric and lower growth ratio between the last two developmental stages of *T. stigmurus*, as well as the reduced number of specimens existing in the last size class (only five individuals) as evidenced by the theoretical method and carapace length, likely yielded similar objects for the algorithms, which tended to group them; this resulted in discordant clustering when comparing the two algorithms and the other methods. Multivariate clustering algorithms appear to work better when groups are more distinct, commonly presenting solutions that do not reflect the natural structure of the data when there is an overlap between groups (Manly, 2008). De Souza et al. (2016) found an overlap in the dimensions of the three structures between extremes of the developmental stages of *T. stigmurus*. They stated that despite the variation, it was possible to distinguish the instars based on the length of the structures with 95% accuracy. Size variation in scorpions has been reported in the literature, mainly in representatives of the genus *Tityus* (Lourenço, 1979; Lourenço & Cloudsley-Thompson, 2010; Albuquerque & Lira, 2016). Some studies conducted in the laboratory have shown that MetV and Car show proportionality in dimensions in all phases of *T. stigmurus*, *T. fasciolatus* Pessoa, 1935, and *T. neblina* Lourenço, 2008, with MetV on average equivalent or slightly greater than Car (Lourenço, 1979; Lourenço & Cloudsley-Thompson, 2010; De Souza et al., 2016). In contrast, the same structures mentioned above showed an inverse proportion in the early stages of development in the *T. stigmurus* individuals analyzed in this study. A similar situation was observed in *Androctonus crassicauda* (Olivier, 1807) adults, depending on the population analyzed (Ozkan, Adiguzel, & Kar, 2006; Ebrahimi, Azizi, Moemenbellah-Fard, Fakoorziba, & Soltani, 2015).

In this study, seven instars were observed for *T. stigmurus*, both by the theoretical method and the frequency distribution of carapace length. Matthiesen (1971) observed five ecdyses in the laboratory, following which the specimen died. In a study that proved the occurrence of parthenogenesis in the species, females reached maturity and gave birth to a litter in the 6th instar (Ross, 2010). De Souza et al. (2016) observed the occurrence of an additional instar, suggesting that maturity was reached after the 6th ecdysis. Using our data from immature specimens, we applied the theoretical method to adults of *T. stigmurus* analyzed by De Souza et al. (2009), obtaining the prediction of seven instars for males. Records of different instar numbers reaching maturity are numerous in the Buthidae family (Polis & Sissom, 1990). Lourenço (1979) observed that *T. fasciolatus* reached maturity in a variable number of instars, establishing the possibility of the coexistence of large and small males. Similar results have been described for *T. pusillus* Pocock, 1893, which displayed an extra moult to reach sexual maturity in some specimens of both sexes, being more common in males (Albuquerque & Lira, 2016).

It should be noted that, in the case of clustering algorithms (K-means and Ward's), pre-existing knowledge about the maximum number of instars for the species influenced the interpretation of the dendrogram generated by the Ward method and was decisive for the definition of the desired number, for clusters using K-means clustering. According to Logan et al. (1998), it is desirable to have evidence that corroborates the number of instars of a given species for better refinement in the interpretation of the algorithm results. However, our results support the claim of Francke and Sissom (1984) that hierarchical clustering methods do not solve the critical problem of determining the number of size classes in a sample. Finally, we found that the analysis results of the peaks in a multimodal curve using the carapace length of the specimens were similar to those obtained by the calculation of the theoretical method for the same structure, with a remote probability of error in determining the instar in most cases (< 1%), which was in agreement with the percentages described in the literature for other groups of arthropods (Chen & Seybold, 2013; Castañeda-Vildózola et al., 2016).

Conclusion

In summary, the robustness of the cluster analysis methods used in this study (K-means and Ward) was insufficient to guarantee the desired precision in separating groups, considering the intended purpose. Furthermore, the Ward algorithm corroborated Francke and Sissom's (1984) statements about their non-objectivity in determining the number of instars. Without this prior knowledge, the interpretation of the results becomes impaired and prone to errors. As reproductive capacity has already been demonstrated in the penultimate size class and considering that there are no reports in the literature about post-reproductive seedlings in scorpions (Polis & Sissom, 1990), it seems reasonable to assume that sexual maturity is reached in the case of *T. stigmurus* in the 6th instar (or 5th ecdysis). Some specimens may have undergone an additional moult before reaching maturity. Based on the results obtained, it is believed that the theoretical method can be used in a complementary way or combined with the observation of the peaks of the length distribution of the carapaces in a multimodal curve, providing greater precision for defining the limits between the frequency classes. As the sexual dimorphism of *T. stigmurus* is expressed in structures and/or axes different from those analyzed in this study (De Souza et al., 2009), with the carapace being a structure less affected by allometry, there appear to be no limitations related to the sexual dimorphism of the species in using these two methods for determining the developmental stage of individuals. The elaboration of an instrument as a test for use on a bench can provide greater agility and practicality in the visual determination of the developmental stages of *T. stigmurus* from the carapace length of the specimens. Analysis of other populations of *T. stigmurus* from urban areas of other localities, as well as natural populations, compared to the results of this study, can confirm the observed variation in relation to the analyzed parameters.

Acknowledgements

To Julia Vono Alvarez Figueiredo for the authorization to carry out the work within the *Laboratório de Identificação e Pesquisa de Fauna Sinantrópica*. Sylvio Cesar Rocco and Marcia Bicudo de Paula for the valuable suggestions. Alessandra Bergamo Araújo for the indication and availability of statistical software. Marco Otávio de Matos Júnior and Leandra Sayuri Isa Hernandes for providing the collection deposit of the specimens used in the work. Sérgio Cominato Ferraz, Jailson Apóstolo dos Santos and Giovanni Balaton Pupin for their encouragement and assistance in starting the project. To the *Núcleo de Vigilância e Controle de Animais Sinantrópicos - DVZ* and *Unidade de Vigilância em Saúde Penha*, whose monitoring of the area where the specimens came from made the work possible. Marisa Mayumi Toda for information about the collection area. Wendy Gallo and Marcelo Munhoz Di Petta for the guidance on configuring the images. This work was supported by *Dirección General de Asuntos del Personal Académico (DGAPA*) postdoctoral fellowship from the *Universidad Nacional Autónoma de Mexico* to AFAL. We would like to thank Editage for English language editing.

References

Albuquerque, C. M. R., & Lira, A. F. A. (2016). Insights into reproductive strategies of *Tityus* (*Archaeotityus*) *pusillus* Pocock, 1893 (Scorpiones, Buthidae). *Comptes Rendus Biologies*, *339*(5), 179-184. DOI: https://doi.org/10.1016/j.crvi.2016.03.003

Albuquerque, C. M. R., Santana Neto, P. L., Amorim, M. L. P., & Pires, S. C. V. (2013). Pediatric epidemiological aspects of scorpionism and report on fatal cases from *Tityus stigmurus* stings (Scorpiones: Buthidae) in State of Pernambuco, Brazil. *Revista da Sociedade Brasileira de Medicina Tropical*, *46*(4), 484-489. DOI: https://doi.org/10.1590/0037-8682-0089-2013

Albuquerque, M. A., & Barros, K. N. N. O. (2020). Determinação do número de grupos em análise de agrupamento via de raio de influência. *Brazilian Journal of Development*, *6*(6): 38342-38355. DOI: https://doi.org/10.34117/bjdv6n6-396

Amado, T. F., Moura, T. A., Riul, P., Lira, A. F. A., Badillo-Montaño, R., & Martinez, P. A. (2021) Vulnerable areas to accidents with scorpions in Brazil. *Tropical Medicine and International Health,* 26(5), 591–601. DOI: https://doi.org/10.1111/tmi.13561

Ambrosano, M. B. G., Igue, T., & Lourenção, A. L. (1997). Using the segmented regression in determining instar number of insects. *Anais da Sociedade Entomológica do Brasil*, *26*(2), 371-374. DOI: https://doi.org/10.1590/S0301-80591997000200020

Ayres, M., Ayres Júnior, M., Ayres, D. L., & Santos, A. S. (2000). *Bioestat 2.0: Aplicações estatísticas nas áreas das Ciências Biológicas e Médicas*. Belém, PA: Sociedade Civil Mamirauá.

Bertani, R., Bonini, R. K., Toda, M. M., Isa, L. S., Figueiredo, J. V. A., Santos, M. R., & Ferraz, S. C. (2018). Alien scorpions in the Municipality of São Paulo, Brazil – evidence of successful establishment of *Tityus stigmurus* (Thorell, 1876) and first records of *Broteochactas parvulus* Pocock, 1897, and *Jaguajir rochae* (Borelli, 1910). *Bioinvasions*, *7*(1), 89–94. DOI: https://doi.org/10.3391/bir.2018.7.1.14

Brasil (2009). *Manual de controle de escorpiões*. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Vigilância Epidemiológica. Brasília: Ministério da Saúde.

Camacho, M. A., Silveira, M. V., Camargo, R. A., & Natale, W. (2012). Faixas normais de nutrientes pelos métodos ChM, DRIS e CND e nível crítico pelo método de distribuição normal reduzida para Laranjeira-Pera. *Revista Brasileira de Ciência do Solo*, *36*(1), 193-200. DOI: https://doi.org/10.1590/S0100-06832012000100020

Castañeda-Vildózola, A., González-Hernández, H., Equihua-Martínez, A., Valdez-Carrasco, J., Peña, J. E., Cazado, L. E., & Franco-Mora, O. (2016). Head capsule width is useful for determining larval instar in *Heilipus lauri* (Coleoptera: Curculionidae). *Florida Entomologist*, *99*(4), 822-825. DOI: https://doi.org/10.1653/024.099.0448

Chen, Y., & Seybold, S. J. (2013). Application of a frequency distribution method for determining instars of the beet armyworm (Lepidoptera: Noctuidae) from widths of cast head capsules. *Journal of Economic Entomology*, *106*(2), 800-806. DOI: https://doi.org/10.1603/EC12367

Coorg, V., Levitan, R. D., Gerkin, R. D., Muenzer, J., & Ruha, A. M. (2017). Clinical presentation and outcomes associated with different treatment modalities for pediatric bark scorpion envenomation. *Journal of Medical Toxicology*, *13*(1), 66–70. DOI: https://doi.org/10.1007/s13181-016-0575-3

De Souza, A. M., Santana Neto, P. L., Lira, A. F. A., & Albuquerque, C. M.R. (2016). Growth and developmental time in the parthenogenetic scorpion *Tityus stigmurus* (Thorell, 1876) (Scorpiones: Buthidae). *Acta Scientiarum. Biological Sciences*, *38*(1), 85-90. DOI: https://doi.org/10.4025/actascibiolsci.v38i1.28235

De Souza, C. A. R., Candido, D. M., Lucas, S. M., Brescovit, A. D. (2009). On the *Tityus stigmurus* complex (Scorpiones, Buthidae). *Zootaxa*, 1987, 1-38. DOI: https://doi.org/10.5281/zenodo.185451

Dujardin, J. P. (2008). Morphometrics applied to medical entomology. *Infection, Genetics and Evolution*, *8*(6), 875-890. DOI: https://doi.org/10.1016/j.meegid.2008.07.011

Dyar, H. G. (1890). The number of molts of lepidopterous larvae. *Psyche*, *5*, 420-422. DOI: https://doi.org/10.1155/1890/23871

Ebrahimi, M., Azizi, K., Moemenbellah-Fard, M. D., Fakoorziba, M. R., & Soltani, A. (2015). Morphometry indices of the black fat-tailed scorpion *Androctonus crassicauda* (Scorpiones: Buthidae), from Fras Province, Southern Iran. *Journal of Entomology, 12*(1), 39-47. DOI: https://doi.org/10.3923/je.2015.39.47

Ecole, C. C., Anjos, N., Michereff-Filho, M., & Picanço, M. (1999). Determinação do número de ínstares larvais em *Plutella xylostella* (L.) (Lepidoptera: Yponomeutidae). *Acta Scientiarum. Biological Sciences, 21*(2), 331-335. DOI: https://doi.org/10.4025/actascibiolsci.v21i0.4447

Foerster, S. Í. A., Dionisio-da-Silva, W., Santos, A. B., Albuquerque, C. M. R., & Lira, A. F. A. (2021). Notes on the courtship behavior of the parthenogenetic scorpion *Tityus stigmurus*. *Acta Arachnologica*, *70*(2), 83-86. DOI: https://doi.org/10.2476/asjaa.70.83

Francke, O. F., & Sissom, W. D. (1984). Comparative review of the methods used to determine the number of molts to maturity in scorpions (Arachnida), with analysis of the post-birth development of *Vaejovis coahuilae* Williams (Vaejovidae). *The Journal of Arachnology, 12*(1), 1-20. Retrieved from http://www.jstor.org/stable/3705098

Guerra-Duarte, C., Saavedra-Langer, R., Matavel, A., Oliveira-Mendes, B. B. R., Chavez-Olortegui, C., & Paiva, A. L. B. (2023). Scorpion envenomation in Brazil: Current scenario and perspectives for containing an increasing health problem. *PLoS Neglected Tropical Diseases, 17*(2): e0011069. DOI: https://doi.org/10.1371/journal.pntd.0011069

Isbister, K. G., & Bawaskar, H. S. (2014). Scorpion envenomation. *The New England Journal of Medicine*, *371*(5), 457–463. DOI: https://doi.org/10.1056/nejmra1401108

Landis, J. R., & Koch, G. G. (1977). An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*, *33*(2), 363-374.

Lira-da-Silva, R. M., Amorim, A. M. D., & Brazil, T. K. (2000) Envenenamento por *Tityus stigmurus* (Scorpiones; Buthidae) no Estado da Bahia, Brasil. *Revista Brasileira de Medicina Tropical*, *33*(3), 239-245. DOI: https://doi.org/10.1590/S0037-86822000000300001

Logan, J. A., Bentz, B. J., Vandygriff, J. C., & Turner, D. L. (1998). General program for determining instar distributions from head capsule widths: Example analysis of mountain pine beetle (Coleoptera: Scolytide) data. *Environmental Entomology*, *27*(3), 555-563. DOI: https://doi.org/10.1093/ee/27.3.555

Lourenço, W. R. (1979). La biologie sexuelle et development postembryonnaire du scorpion Buthidae: *Tityus trivittatus fasciolatus* Pessôa, 1935. *Revista Nordestina de Biologia*, *2*(1), 49-96.

Lourenço, W. R. (2002). *Scorpions of Brazil*. Paris, FR: Les Éditions l'If.

Lourenço, W. R., & Cloudsley-Thompson, J. L. (2010). The cycle of *Tityus* (*Atreus*) *neblina* Lourenço 2008 (Scorpiones, Buthidae) in Cerro de La Neblina, Brazil/Venezuela. *Boletín de la Sociedade Entomologíca Aragonesa*, *47*, 293-298.

Maia, C. E., Morais, E. R. C., & Oliveira, M. (2001). Nível crítico pelo critério da distribuição normal reduzida: uma nova proposta para interpretação de análise foliar. *Revista Brasileira de Engenharia Agrícola e Ambiental*, *5*(2), 235-238. DOI: https://doi.org/10.1590/S1415-43662001000200010

Manly, B. J. F. (2008). Análise de agrupamentos. In B. J. F. Manly (Ed.), *Métodos estatísticos multivariados: Uma introdução* (p. 139-155). Porto Alegre, RS: Bookman.

Marcus, L. F. (1990). Traditional Morphometrics. In F. J. Rohlf, & F. L. Bookstein (Ed.), *Proceedings of the Michigan Morphometrics Workshop*. Ann Arbour, Michigan: The University of Michigan Museum of Zoology.

Matthiesen, F. A. (1971). Observation on four species of brazilian scorpions in captivity. *Revista Brasileira de Pesquisa Médicas e Biológicas*, *4*(4), 301-302.

McElroy, T., McReynolds, C. N., Gulledge, A., Knight, K. R., Smith, W. E., & Albrecht, E. A. (2017) Differential toxicity and venom gland gene expression in *Centruroides vittatus*. *PLoS ONE, 12*(10). DOI: https://doi.org/10.1371%2Fjournal.pone.0184695

Moori, P. G., Marcondes, R. C., & Ávila, R. T. (2002). A análise de agrupamentos como instrumento de apoio à melhoria da qualidade. *Revista de Administração Contemporânea*, *6*(1), 63-84. DOI: https://doi.org/10.1590/S1415-65552002000100005

Ozkan, O., Adiguzel, S., & Kar, S. (2006). Parametric values of *Androctonus crassicauda* (Oliver, 1807) (Scorpiones: Buthidae) from Turkey. *Journal of Venomous Animals and Toxins including Tropical Diseases*, *12*(4), 549-559. DOI: https://doi.org/10.1590/S1678-91992006000400003

Parra, J. R. P., & Haddad, M. L. (1989). *Determinação do número de ínstares de insetos*. Piracicaba, SP: Fundação de Estudos Agrários Luiz de Queiroz.

Polis, G. A., & Sissom, W. D. (1990). Life history. In G. A. Polis (Ed.), *The biology of scorpions* (p. 161-223). California, CA: Stanford University Press.

Przibram, H., & Megušar, F. (1912). Wachstumsmessungen an *Sphodromantis bioculata* Burm. *Archiv für Entwicklungsmechanik der Organismen*, *34*, 680–741. DOI: https://doi.org/10.1007/BF02292211

Ross, L. K. (2010). Confirmation of the parthenogenesis in the medically significant, synanthropic scorpion *Tityus stigmurus* (Thorell, 1876) (Scorpiones: Buthidae). *Revista Ibérica de Aracnología*, *18*(1), 115-121. Retrieved from http://sea-entomologia.org/Publicaciones/RevistaIbericaAracnologia/RIA18/RIA18.htm

Seidel, E. J., Moreira Junior, F. J., Ansuj, A. P., & Noal, M. R. C. (2008). Comparação entre o método Ward e o método K-means no agrupamento de produtores de leite. *Ciência e Natura*, *30*(1), 7-15. DOI: https://doi.org/10.5902/2179460X9737

Sissom, W. D., Polis, G. A., & Watt, D. D. (1990). Field and Laboratory Methods. In G. A. Polis (Ed.), *The Biology of Scorpions* (p. 445-461). California, CA: Stanford University Press.

Stahnke, H. L. (1970). Scorpion nomenclature and mensuration. *Entomological News*, 81, 291-316.

Tomaz, F. S. C., Matos, K., & Souza, E. (2017). Profile of customers of a provider of cellular services company: Application of Ward's clustering method. *Multi-Science Journal*, *1*, 92-95.

Ward, M. J., Ellsworth, S. A., & Nystrom, G. S. (2018). A global accounting of medically significant scorpions: Epidemiology, major toxins, and comparative resources in harmless counterparts. *Toxicon*, *151*, 137-155. DOI: https://doi.org/10.1016/j.toxicon.2018.07.007