



## Article

# DNA Barcoding for Scorpion Species from New Valley Governorate in Egypt Reveals Different Degrees of Cryptic Speciation and Species Misnaming

Khaled Mohammed-Geba <sup>1,\*</sup>, Ahmad Hamed Obuid-Allah <sup>2</sup>, Naser Abdellatif El-Shimy <sup>2</sup>,  
Mohamed Abd El-Moez Mahbob <sup>3</sup>, Rouwaida Saadawy Ali <sup>3</sup> and Shima Mohamed Said <sup>2</sup>

<sup>1</sup> Zoology Department, Faculty of Science, Menoufia University, Shebin El-Kom 32952, Egypt

<sup>2</sup> Zoology Department, Faculty of Science, Assiut University, Assiut 71515, Egypt; obuidallah1155@gmail.com (A.H.O.-A.); nshimy2@yahoo.com (N.A.E.-S.); shima.mohamed82@aun.edu.eg (S.M.S.)

<sup>3</sup> Zoology and Entomology Department, Faculty of Science, New Valley University, El-Kharga 72511, Egypt; mahbobent@sci.nvu.edu.eg (M.A.E.-M.M.); Rowida.sadawy@sci.nvu.edu.eg (R.S.A.)

\* Correspondence: Khaled.mohammed@icman.csic.es or khaledspain@yahoo.com; Tel.: +20-1022144485

**Abstract:** (1) Background: Scorpions (Arthropoda: Arachnida) represent a diverse group of invertebrates, accounting for a significant proportion of earth's predators and ecosystems' modulators. Surviving mostly in hardly reachable nests, and representing key hazards to human health, they attracted major interest for characterizing their eco-, morpho-, and genotypes. (2) Methods: Four scorpion species were collected from the New Valley governorate in Upper Egypt, where a high level of scorpionism and related neurological symptoms are found, that were *Leiurus quinquestriatus*, *Androctonus amoreuxi*, *Orthochirus innesi*, *Buthacus leptochelys*. They were DNA barcoded, genetically and phylogenetically analyzed through PCR amplification and sequencing of the mitochondrial cytochrome oxidase subunit 1 (COI) gene hypervariable 5' region. (3) Results: New, morphologically authenticated scorpion barcodes could be added to the barcoding databases. However, several discrepancies and barcode database inadequacies could be revealed. Moreover, taxon-specific patterns for nitrogenous bases' distribution could be identified, resulting in a significantly high percentage of COI barcode guanine in scorpionids, in comparison to araneids and opilions. (4) Conclusions: For a group of animals where both cryptic speciation and a high risk of human envenomation are evident, the findings of the current study strongly recommend continuous and comprehensive research efforts dealing with morphogenetic authentication for different species of scorpions.

**Keywords:** *Androctonus amoreuxi*; *Buthacus leptochelys*; DNA barcoding; *Leiurus quinquestriatus*; misnaming; morphogenetic; *Orthochirus innesi*



**Citation:** Mohammed-Geba, K.; Obuid-Allah, A.H.; El-Shimy, N.A.; Mahbob, M.A.E.-M.; Ali, R.S.; Said, S.M. DNA Barcoding for Scorpion Species from New Valley Governorate in Egypt Reveals Different Degrees of Cryptic Speciation and Species Misnaming. *Conservation* **2021**, *1*, 228–240. <https://doi.org/10.3390/conservation1030018>

Academic Editor: Svein Øivind Solberg

Received: 31 July 2021

Accepted: 26 August 2021

Published: 31 August 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Scorpions are arachnid arthropods of key environmental and medical importance. They constitute one of the oldest animal groups in the world, with a fossil history dated to more than 300 million years ago. They inhabit mainly hot and dry environments, where they represent the most important taxa of terrestrial predators in terms of density and biomass [1]. Some early estimates for their biomass identified that they exceed most key terrestrial arthropod taxa, except ants and termites [2]. They control major ecosystem processes such as community structure and function via predation and feeding competition [1,3]. Since scorpions are predators of small arthropods and feed infrequently across multi-year lifespans, their high biomass could be attributed to the depressed metabolic rate [4]. They exhibit long life spans, (2–25 years, with a mean of 4–8 years), late maturation period (6 months to 7 years), long gestation period (1.5 months to 2 years), and a characteristic pattern of maternal care of young [5]. Moreover, scorpions are tolerant to

extreme environmental conditions such as ionizing radiation, e.g., some species of the genus *Androctonus* tolerate levels of 400–800 grays, i.e., more than 100 times the dose that causes human blood and digestive problems; extreme temperatures, e.g., 40–50 °C, as well as drought, food scarcity, and infections [5]. However, these capabilities are species-specific. Their distribution is associated with areas of climatic, topographic, and geological complexity [6]. At local scales, scorpions' distribution is governed mainly by temperature, precipitation, substrate (soil hardness and texture; amount of stone or litter cover) and vegetation physiognomy [7,8]. All scorpions have a venomous sting and several thousands of people die every year from scorpion stings [9].

The scorpion fauna of Egypt is represented by four families: Buthidae, Euscorpionidae, Hemiscorpiidae and Scorpionidae [10]. Valuable, recent knowledge about scorpion fauna of Egypt and the region has been compiled, reviewed and edited by several authors (e.g., [10–12]). Moreover, numerous works were carried out to elucidate composition, ecology and biogeography of Middle Eastern scorpions, including Egypt (for examples, see [13–17]). These works pointed to the presence of a total of 35 scorpion species in Egypt, all of which were listed in [15,16].

The New Valley governorate in Egypt is the largest Egyptian governorate, and the one with the most extended desert areas. The state of genetic diversity of scorpions in this governorate, however, is still understudied. Further, extensive sampling of this biodiverse region is needed to recover the complete genetic phylogeographic pattern of scorpions in this area.

DNA barcoding is a rapidly expanding protocol for molecular identification of different animal species from different environments, even the ones with highly similar morphologies, or with variable degrees of integrity. The 5', inter-specifically hypervariable region in the barcode of life gene, i.e., the mitochondrial cytochrome oxidase subunit 1 (COI) gene, is typically included in barcoding studies as a global bio-identification system [18,19]. The efficiency of DNA barcoding is almost complete upon coupling it to a thorough morphological description of the species [20]. The present authors previously provided full morphological descriptions of four major scorpion species in the New Valley, Egypt: *Leiurus quinquestriatus*, *Androctonus amoreuxi*, *Orthochirus innesi*, *Buthacus leptochelys* [21,22]. The aim of the current study was to complete the morphogenetic authentication of these species using DNA barcoding, a task that seems crucial since some of these species are the most notorious in terms of scorpionism, especially in the New Valley governorate, and in Egypt in general [12,23]. However, there are still many deficiencies regarding the DNA barcoding and phylogenetic relationships of these species, which create many serious issues regarding the appropriate use of DNA barcoding databases for scorpion species identification.

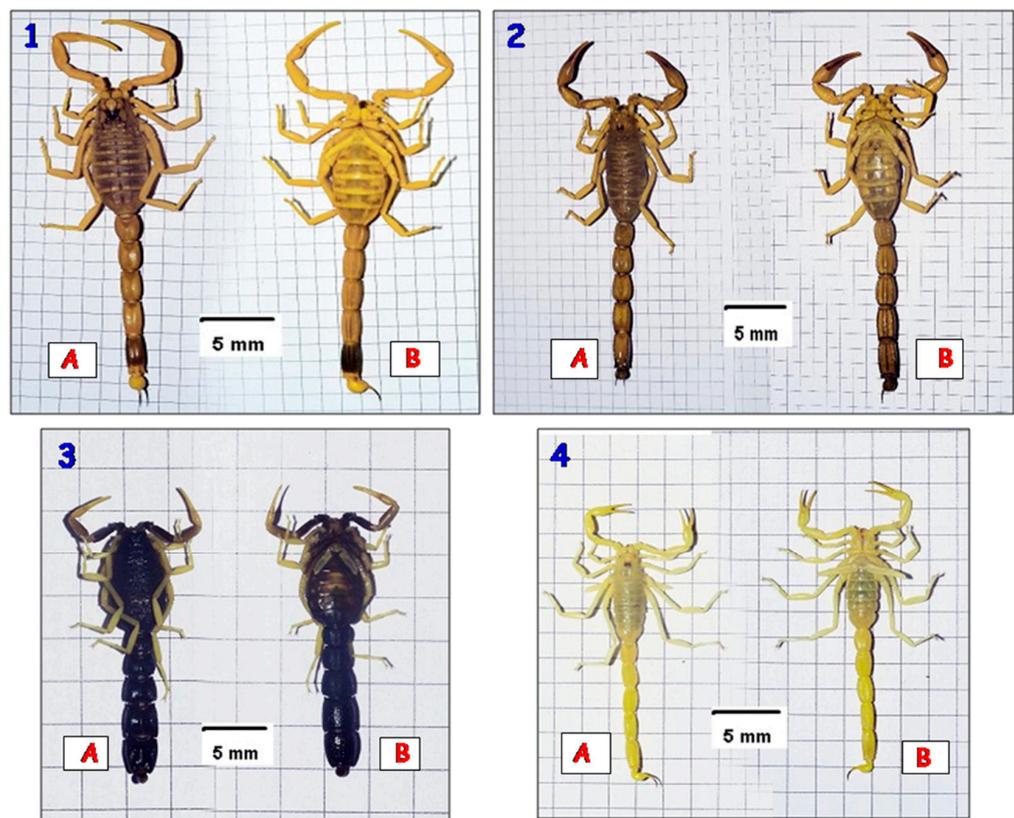
## 2. Materials and Methods

### 2.1. Sample Collection and Preservation

Three to five samples from four different scorpion species were collected by professional hunters from El Kharga oasis (25.4390° N, 30.5586° E), El Dakhla oasis (25.5167° N, 29.1667° E) Balat area (25.560742° N, 29.265589° E), and Farafra oasis (27.0567° N, 27.9703° E) (Figure 1). These samples belonged to four species: *Leiurus quinquestriatus*, *Androctonus amoreuxi*, *Orthochirus Innesi*, *Buthacus leptochelys* (Figure 2). The samples were preserved in absolute ethanol, then transferred to the Molecular Biology and Biotechnology Laboratory in the Faculty of Science of Menoufia University (Shebeen El-Kom City, Egypt). About 100 mg of tail musculature were removed from each specimen and stored in absolute ethyl alcohol in a –20 °C freezer until being used in DNA extraction.



**Figure 1.** Location map of New Valley Governorate, Egypt showing the study areas of El Kharga oasis (25.4390° N, 30.5586° E), El Dakhla oasis (25.5167° N, 29.1667° E) Balat area (25.560742° N, 29.265589° E), and Farafra oasis (27.0567° N, 27.9703° E). Photo credits: GoogleMaps™ (below the map).



**Figure 2.** Dorsal (A) and ventral (B) views for the four scorpion species collected in the current study: *Leiurus quinquestriatus* (1), *Androctonus amoreuxi* (2), *Orthochirus Innesi* (3), *Buthacus leptochelys* (4).

## 2.2. DNA Purification and COI-Based Polymerase Chain (PCR)

Total genomic DNA was purified from 30 mg of tail musculature of scorpions using the method described in Mohammed-Geba et al., (2016) [24]. Briefly, the samples were lysed individually using 200  $\mu\text{L}$  of TNES-urea buffer [25] and 2.4 U  $\text{mL}^{-1}$  Proteinase K solution (Thermo Fischer Scientific, Waltham, MA, USA), with incubation at 55 °C for 30 min. Later on, 54  $\mu\text{L}$  of 5 M NaCl was added, the tubes were thoroughly mixed by inversion, then centrifuged at 4000 $\times g$  for 10 min. The supernatant from each sample was transferred to another 1.5 mL Eppendorf tube, and the DNA was then precipitated by adding 200  $\mu\text{L}$  of cold isopropanol (at  $-20$  °C) with shaking by inversion. The tubes were centrifuged at 11,000 $\times g$  for 10 min, and the supernatant was completely removed. The DNA pellet was washed by 400  $\mu\text{L}$  of 70% ethanol, centrifuged for 5 min at 11,000 $\times g$  and poured completely from ethanol, then 30  $\mu\text{L}$  of tris EDTA buffer (pH 8) was added for DNA pellet resuspension. DNA quality was checked by running 5  $\mu\text{L}$  of the genomic DNA with 1  $\mu\text{L}$  of 6  $\times$  DNA loading buffer (0.25% *w/v* bromophenol blue, 40% *w/v* sucrose), in 1% agarose gels stained by 0.5  $\mu\text{g mL}^{-1}$  ethidium bromide (Thermo Fisher Scientific). DNA samples were used directly for amplification of partial barcode region of the COI gene by PCR. A 658 bp target from the 5' end of COI was amplified using the primer pair LCO1490(5'-GGTCAACAAATCATAAAGATATTGG-3') and HCO2198(5'-TAAACTTCAGGGTGACCAAAAAATCA-3') [26]. The amplification reactions were performed in a total volume of 25  $\mu\text{L}$ . The reaction mixture consisted of 2  $\mu\text{L}$  of template DNA (~50 ng), 0.5  $\mu\text{M}$  of each primer, 25  $\mu\text{L}$  of 2  $\times$  of My Taq red master mix (Bioline, London, UK), and completed to 50  $\mu\text{L}$  with PCR-grade water. PCR amplifications were carried out in the thermal cycler TC512 (Techne, Stone, UK). The PCR program contained an initial denaturation step at 94.6 °C for 10 min, followed by 40 cycles of 94 °C for 1 m, 46 °C for 1 min, 72 °C for 1 min, and a final extension of 72 °C for 10 min. PCR products were electrophoresed in a 1% agarose gel. Adequately sized PCR products were sent to Macrogen Inc. (Seoul, South Korea) for sequencing.

## 2.3. Genetic and Phylogenetic Analyses for New Valley Scorpions

After careful revision of the obtained sequences to assure correct base recall, these sequences were translated into the primary amino acids sequence using MEGA7 software, as a way to assure that there were no premature stop codons in the obtained COI sequences, that usually mark the amplification of false nuclear copies of mitochondrial genes (NuMTs) [27]. Then, each sequence was individually compared to the GenBank database using BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi> (accessed on 25 June 2021)), and to the Barcode of Life Database (BOLD) using the BOLD Identification System IDS ([http://www.boldsystems.org/index.php/IDS\\_OpenIdEngine](http://www.boldsystems.org/index.php/IDS_OpenIdEngine) (accessed on 25 June 2021)). In order to calculate the pairwise distances, and to construct the phylogenetic tree, COI sequences belonging to the same and closely related scorpion species (Table 1) were retrieved from the GenBank database. All sequences were aligned using CLUSTAL W Multiple alignment implemented in MEGA7 software. A prior identification of the best nucleotide substitution model was carried out using J Model Test software V. 2.1.10 [28]. This alignment, with the determined substitution model, was applied to calculate the genetic pairwise distances. Later, the alignment was uploaded to Mr Bayes 3.2.1 software [29] for constructing a Bayesian inference (BI)-based phylogenetic tree. Next, four Markov chain Monte Carlo (MCMC) chains were analyzed for 10 million generations, saving a tree each 1000 generations. The subsequent analyses were carried out after assuring an average standard deviation of split frequencies below 0.001. The number of burn-ins was identified using Tracer 1.7 [30]. Tracer 1.7 exhibited that 25% of the saved trees are to be discarded as burn-ins. This information was transferred to Mr Bayes 3.2.1. for constructing the summarized tree, which was then viewed using the Interactive Tree of Life online platform (iTOL: [31]). Finally, nucleotide compositions for the sequenced COI fragments were analyzed to test whether there are taxonomic trends among different arachnids.

**Table 1.** Accession numbers, species and locations of sequences retrieved from BOLD/GenBank database for construction of phylogenetic tree.

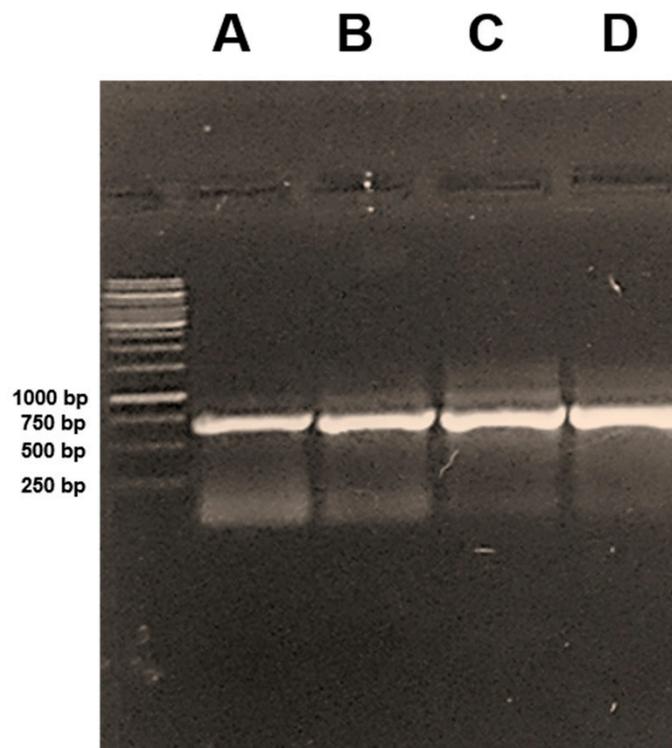
BOLD Accession	Species	Country	GenBank Accession	
GACO2126-19	<i>Leiurus quinquestriatus</i>	Sudan	KX648423	
GACO2127-19			KX648424	
GBCH7621-13			JQ514258	
GBMIN116614-17			KX648421	
GACO2125-19			KX648422	
GBMNC71230-20			MT636860	
GACO2124-19			KX648420	
GBA14470-14			Algeria	KJ538427
GBA14469-14				KJ538423
GACO2547-19				KJ538294
GBA14726-14	KJ538282			
GBA14727-14	KJ538286			
GBA14731-14	KJ538290			
GBA14732-14	KJ538297			
GBA14733-14	KJ538305			
GBA14738-14	KJ538301			
GBA14739-14	KJ538313			
GBA14740-14	Morocco	KJ538317		
GBA14741-14		KJ538309		
GBA14750-14		KJ538275		
GBA14751-14		<i>Androctonus amoreuxi</i>	KJ538279	
GBA14752-14			KJ538436	
GBA14756-14			KJ538429	
GBA14757-14			KJ538447	
GBA14758-14			KJ538433	
GBA14765-14			KJ538450	
GBA14997-14			Tunisia	KJ538439
GBA14998-14	KJ538443			
GBA16571-14	Egypt		KJ538478	
GBA16576-14			KJ538492	
GBA16572-14		KJ538480		
GBA16573-14		KJ538483		
GBA16574-14		KJ538486		
GBA16575-14		KJ538489		
GBCH7634-13		<i>Orthochirus innesi</i>	Origin_Unknown	JQ514244
GBMNC71229-20	<i>Buthacus arenicola</i>	Origin_Unknown	MT636861	
GBMNB48836-20	<i>Buthacus macrocentrus</i>	Iraq	MT229838	
GBA16582-14	<i>Buthacus</i> sp.	Egypt	KF548116	

Furthermore, in order to reveal possible differences in the nitrogenous bases' percentages among different arachnid taxa, COI sequences from aranean and opilion (Orders:

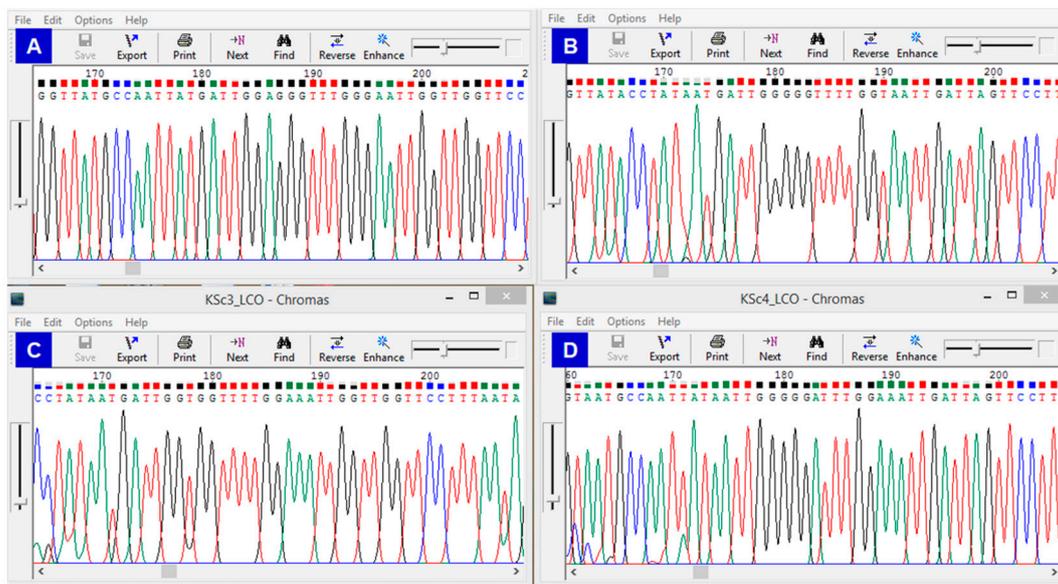
Aranea and Opiliones) species that showed high barcode similarity with the identified scorpion species in the current study were retrieved from GenBank, i.e., *Afrarchaea royalensis* (acc. No. KP201255), *Coelotes terrestris* (acc. No. KX039130.1), *Pireneitega tianchiensis* (acc. No. KY778995.1), *Saltonia incerta* (acc. No. KY017701), *Galeodes bacillatus* (acc. No. KX420494.1), *Metanonychus nigricans* (acc. No. MN125378.1), *Hesperonemastoma modestum* (acc. No. EF108588.1), *Ischyropsalis adamii* (acc. No. KP224378.1), *Neopachylus bellicosus* (acc. No. MG769244.1), and *Sabacon briggsi* (acc. No. JX573647.1). These sequences were appended to the same alignment previously produced using MEGA7 for phylogenetic analysis. Then, MEGA7 software was also applied for calculation of percentages of composite nucleotides in the analyzed scorpions, aranean, and opilion species. These percentages were statistically analyzed among the three orders and were compared using one-way analysis of variance (ANOVA), applying LSD as post-hoc. Differences were considered significant at  $p < 0.05$ . All statistical analyses were carried out using Statgraphics Centurion XVI software.

### 3. Results

PCR amplification for the barcode region of the mitochondrial COI gene resulted in a 650 bp specific band in each of the assessed species, without any apparent length variation (Figure 3). Sequencing of the produced PCR amplicons from all the four species' samples resulted in good quality sequence chromatograms (Figure 4). Manual trimming of these sequences resulted in about 600 good quality nucleotide peaks in all the assessed species. In silico translation of the sequences using the invertebrates' mitochondrial DNA codons table resulted in a complete absence of premature stop codons, i.e., no NUMTs were amplified instead of the expected COI sequences. As all sequences from each species belonged to a single haplotype, the haplotype sequence of each species was deposited to GenBank database under the accession numbers MZ669859 for *L. quinquestriatus* MZ669860 for *A. amoreuxi*; MZ669861 for *O. innesi*; and MZ669862 for *B. leptochelys*. This latter one was the first barcode for this species ever in the GenBank database.



**Figure 3.** 650 bp COI-based PCR amplicons for example sampled from *Leirus quinquestriatus* (A), *Androctonus amoreuxi* (B), *Orthochirus Innesi* (C), *Buthacus leptochelys* (D).



**Figure 4.** COI sequences' chromatograms for *Leirus quinquestriatus* (A), *Androctonus amoreuxi* (B), *Orthochirus Innesi* (C), *Buthacus leptochelys* (D). Red peaks: Thymine, Green peaks: Adenine, Blue peaks: Cytosine, Black peaks: Guanine.

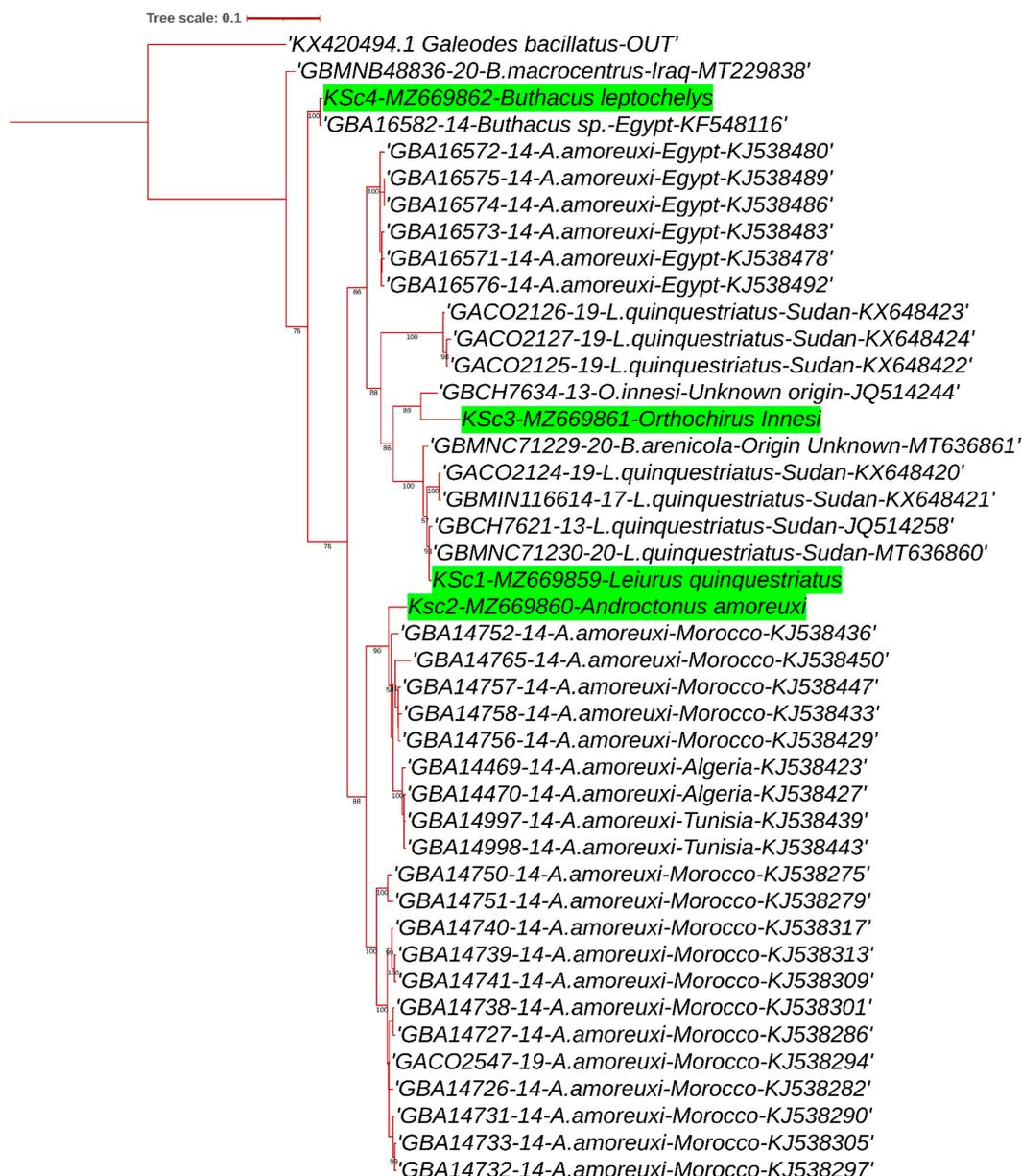
BLAST and BOLD comparisons exhibited the absence of accurate species-level barcodes for two of the four assessed species. Moreover, some discrepancies were found for the species nomenclature available in the GenBank. *Leirus quinquestriatus* showed homogenous species identify levels (i.e., 98–100%) with samples for the same species that were previously collected from the South of Egypt and the North of Sudan and deposited into the GenBank under the accession numbers (acc. No.) KX648420, KX648421.1, and JQ514258.1. However, unexpected high identity level of COI sequence (99.40%) was found between *L. quinquestriatus* barcoded in the current study and a sample of *Buthacus arenicola* from Egypt (acc. No. MT636861.1).

*Androctonus amoreuxi* collected in the current study showed 92.5–100% sequence identity with samples from the same species that were collected from different areas in Morocco. An increasing tendency for sequence similarity could be identified when the sample was collected more towards the North of Morocco, i.e., 92.51% similarity, with the sample whose acc. No. was KJ538282.1 and Latitude-Longitude (lat\_lon) = "28.25° N 9.33° W"; 92.71% similarity with KJ538294 coming from lat\_lon = "28.61° N 9.43° W"; 92.51% similarity with KJ538309.1 coming from lat\_lon = "29.73° N 7.97° W"; 92.51% similarity with KJ538313 coming from lat\_lon = "29.63° N 8.01° W"; 92.71% similarity with KJ538290.1 coming from lat\_lon = "28.77° N 9.46° W"; 92.71% similarity with KJ538297 coming from lat\_lon = "29.05° N 8.78° W"; 92.71% similarity with KJ538317 coming from lat\_lon = "29.68° N 7.98° W"; 96.20% similarity with KJ538450.1 coming from lat\_lon = "33.89° N 2.02° W"; 97.00% similarity with KJ538423.1 coming from lat\_lon = "32.44° N 3.74° E"; 97.00% similarity with KJ538427.1 coming from lat\_lon = "32.44° N 3.74° E"; 97.60% similarity with KJ538433.1 coming from lat\_lon = "31.14° N 4.02° W"; 97.80% similarity with J538429.1 coming from lat\_lon = "32.48° N 1.72° W"; and 97.80% similarity with KJ538447.1 coming from lat\_lon = "32.48° N 1.72° W".

For *O. innesi*, the closest similarity was 91% with a sample from the same species with acc. No. JQ514244.1, that was collected from a non-native pet shop. *Buthacus leptochelys* showed 99.6% sequence identity with *Buthacus* sp. specimen that was previously isolated from Egypt, DNA barcoded and registered in GenBank under the accession number KF548116. The closest species (90%) was with *Buthacus macrocentrus*, with acc. No. MT229838.1.

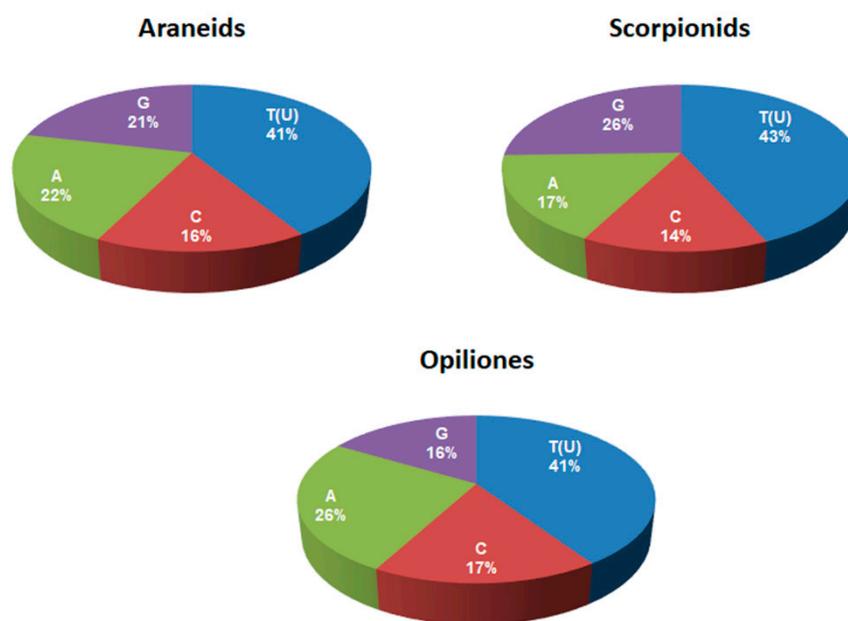
The Phylogenetic analysis was concordant with the barcoding databases comparisons' results. In most cases, the scorpion samples analyzed in the current study existed

in the same clades with their BOLD/GenBank references. However, there were several unexpected contradictions in the phylogenetic tree. First, *A. amoreuxi* appeared to be split into two completely separate clades, one of which was closer to the clade encompassing *L. quinquestriatus* and *O. innesi* samples and references, and the other one that was completely independent and encompassing the samples and the references for this species from Morocco, Tunisia and Algeria (Figure 5). Second, the clade that encompassed *O. innesi* was more closely related to one of the two *L. quinquestriatus* clades, which also encompassed the sampled *L. quinquestriatus* in the current study, while another *L. quinquestriatus* clade appeared to be diverged from the first one (Figure 5). Genetic pairwise distances agreed with these intraspecific phylogenetic divergences, being the distances ranged from 0 to 0.21 for *L. quinquestriatus*; 0 to 0.12 for *A. amoreuxi*; and 0 to 0.17 for *Buthacus* sp. (Supplementary Table S1).



**Figure 5.** BI phylogenetic analysis for the scorpions barcoded in the current study (Green-highlighted) in relation to BOLD and GenBank references. Four Markov chain Monte Carlo (MCMC) chains were analyzed for ten million (ngen = 10,000,000) generations. Bootstrap values are shown above the branches. References' labels are ordered as BOLD BIN-Genus, species-Origin-GenBank accession number. The sequence of COI barcode of the spider *Galeodes bacillatus* was used as an outgroup.

Nitrogenous bases' composition for the barcode DNA fragment analyzed from the COI gene varied slightly among the four species, yet they were not the same among any of them. The order of abundance of these bases was fixed in all species, thymine (T) being the most abundant, followed by guanine (G), then fewer amounts of adenine (A) and cytosine (C). The percentages for T, G, A, C were 42.5%, 25.9%, 17.2%, and 14.4%, for *L. quinquestriatus*; 43.1%, 25.6%, 16.9%, and 14.4%, for *A. amoreuxi*; 42.8%, 24.1%, 19.1%, and 14.1% for *O. innesi* and 43.8%, 25%, 16.6%, and 14.7% for *B. leptochelys*, respectively. The average percentages in the scorpionids covered in the current study and these whose sequences are available in the GenBank database were 43% T, 25% G, 17% of A, and 14% for C. These percentages were similar among different scorpion species that were covered by the phylogenetic analysis in the current study. However, they differed in comparison to the other free living arachnid orders (Figure 6, Table 2). The average nucleotide composition for the same COI fragment in different araneids were 41% T, 21% for both A and G, and 16% for C. The opilion's nucleotide composition percentages were 41% T, 26% A, 17% of C, and 16% for G (Table 2). Pairwise comparisons, using Student's *t*-test, exhibited that the differences in percentage of purines (A and G) were highly significant ( $p = 0.0$ ) between scorpionids and araneids. The percentage of A was higher in scorpionids than in araneids, while A was lower in scorpionids than in araneids. Cytosine was significantly lower in scorpionids than in araneids ( $p = 0.04$ ). Thymine percentages did not vary significantly between the two groups (Table 2). ANOVA exhibited that pyrimidines (T, C) did not vary significantly among scorpionids, araneids, and opiliones. However, purines (A, G) were significantly different among the three arachnid groups ( $p = 0.0$ ). Guanine in scorpionids was the highest among the three orders, while adenine was the least abundant among the three of them.



**Figure 6.** Percentages of nucleotides in the COI fragment in the three free living arachnid orders. T: Thymine, A: Adenine, C: Cytosine, and G: Guanine.

**Table 2.** Average nucleotide composition in different free-living arachnid orders. Asterisk refers to significant inter-group difference ( $p = 0.0$ ).

	T(U) %	C %	A %	G %
Scorpiones	43	14	17 *	25 *
Aranea	41	16	21 *	21 *
Opiliones	41	17	26 *	16 *

#### 4. Discussion

The current work provided the molecular pillar for morphogenetic authentication of four scorpion species, i.e., *L. quinquestriatus*, *A. amoreuxi*, *O. innesi*, and *B. leptochelys*. This work came next to previous, detailed and descriptive analyses that were provided by our research group for elucidating the main morphological aspects for these species [22,23]. Herein, it was possible to provide new DNA barcodes for these four species from the New Valley governorate. This governorate is one of the key Egyptian locations where abundant diversity of scorpions is present. Additionally, it suffers from elevated levels of scorpionism [32,33].

Scorpionism in Upper Egypt is the cause of many cases of human systemic symptoms, among which the reported prevalence of neurological manifestation after envenomation was about 78–85% [33,34]. *Androctonus* and *Leiurus*, that were subjected to DNA barcoding in the current study, and morphological description in the previous ones, are among few representative scorpion genera that represent a real hazard for human health [33,35,36]. Internationally, annual cases of more than 1.2 million stings are reported, of which Middle Eastern and North African scorpion stings account for 42% of the total cases [9]. Such percentages of scorpionism raised the attention towards the fundamentality of providing accurate clues for scorpion species identification.

However, genetic and morphological discrimination of scorpion species are not only limited to toxicological response, but also a major component of interest is diverted to species-specific toxins' structures and variations, ecological niche delimitation, conservation, and evolution. DNA barcoding and molecular identification of species could resolve some problems related to the taxonomy of scorpions. Several species appear morphologically and ecologically similar, at least in some age classes such as juveniles, but genetically distinct [37–39]. Additionally, subspecies-dependent variations in toxin production and structures were found in several scorpion species before, including *L. quinquestriatus* [33,40].

The limited dispersal capability of scorpions leads to limited genetic diversity patterns and reduced gene flow among the species [39]. For instance, European scorpions were thought to be highly distinct from their counterparts from the North African populations [41]. More in-depth genetic analyses and diversified sampling for these populations produced better knowledge regarding the presence of cryptic and unknown species that exhibit clear genetic separation [38]. Likewise, in Egypt, the scorpion fauna distribution is highly discontinuous due to the patchy habitat distribution. This results from wide areas of highly arid deserts, interrupted by two terrestrial corridors of lower aridity, that are the Nile River plain and the Mediterranean belt [10]. DNA barcoding could provide an excellent tracking tool for these variations and hidden diversities within and among different taxa. For example, high genetic diversity and possible cryptic speciation within *L. quinquestriatus* populations in Egypt could be identified [12]. Application of DNA barcoding could detect a clear interspecific phylogenetic relationship among different species of the genus *Buthacus* (Scorpiones: Buthidae) in Egypt and Saudi Arabia [42]. Genetic and morphological analyses elucidated that the Australian endemic scorpion *Urodacus yaschenkoi* (Scorpiones: Urodacidae) is a species complex [43]. Morphogenetic identification and ecological niche modeling were proven crucial to delimit the boundary of the Chinese Przewalski's scorpion *Mesobuthus martensii* (Scorpiones: Buthidae) in arid regions of China and Mongolia [44].

In the current study, comparison of our morpho-genetically authenticated samples and the references in the GenBank database exhibited several discrepancies. For instance, a sequence of *Buthacus arenicola* showed a high identity level with our *L. quinquestriatus*. Additionally, our *O. innesi* exhibited 91% similarity with a sample deposited in the GenBank under the same species designation. Improper scorpion species identification has led to the presence of misidentified species with GenBank accessions (for example, see Reference [45]). Furthermore, the unexpected phylogenetic placements of *L. quinquestriatus*, *A. amoreuxi*, and *O. innesi* samples can either detect the presence of cryptic species, or some morphological identification errors during sampling owing to external morphological similarities. In

general, misidentification and errors in classification/identification are very common in scorpions, even in the specialized literature [46]. This can eventually lead to confusion and inadequacy in the treatment of problems caused by dangerous scorpion species [46]. Several studies reported misidentifications of the species analyzed in the current study, despite some of them being notorious, highly toxic species. These misidentifications seem to affect both species records and barcode databases' registration. Hendrixson et al. (2006) [47] reported a misidentification of *A. amoreuxi* from Medina in Saudi Arabia early in the 20th century. Lourenço (2020) [48] referred to misidentified samples of *L. quinquestriatus* that were collected from Mali and Algeria, i.e., out of the natural range of that species. *Orthochirus innesi*, whose geographical range is naturally in the North of Africa in close relation to oases, has been referred to in some works as a Sudanese species, despite not having been confirmed [49].

Moreover, we could identify a trend among species for taxon-specific patterns of purine-containing nucleotide, which can be suggested as a taxonomic criterion for characterizing different orders belonging to the class Arachnida. Scorpions exhibited the highest percentage of guanine in the COI barcode. Despite this being the first time, to the best of the authors' knowledge, that this issue in scorpions has been reported, some reports in other animal groups identified this type of deviation as a reflection of separate divergence of these groups from their ancestors and sister groups [50–52]. More future work is expected to ensue from this finding in order to assess the taxonomic significance of this guanine-skewed composition of COI barcode region in scorpions among all arachnid taxa.

## 5. Conclusions

In conclusion, we provided morphologically authenticated DNA barcodes for four scorpion species in Egypt, two of which are of direct medical importance and health hazards. The provided barcodes can be considered of high international importance for appropriate calibration of DNA barcoding databases, especially secondary to some discrepancies that could be identified there. Possibilities of cryptic speciation leading to morphological misidentifications can be strongly suggested in light of the present findings. Appearance of guanine as the most prevalent base within the COI gene barcode region and in a manner that was significantly different among the three free-living arachnid orders suggests the need for more future work in identifying its taxonomic significance. A direct recommendation for the current study is to provide more comprehensive studies for morphogenetic authentication of different scorpion species, especially in a region of the world where high diversity of this animal group is present.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/conservation1030018/s1>, Table S1: Genetic pairwise distances between different samples analyzed in the current study and their BOLD/GenBank references.

**Author Contributions:** Conceptualization, A.H.O.-A. and N.A.E.-S.; methodology, K.M.-G., R.S.A. and S.M.S.; software, K.M.-G.; validation, A.H.O.-A., N.A.E.-S., M.A.E.-M.M.; formal analysis, K.M.-G., R.S.A. and S.M.S.; investigation, K.M.-G.; resources, K.M.-G., R.S.A. and S.M.S.; data curation, K.M.-G., R.S.A. and S.M.S.; writing—original draft preparation, K.M.-G., R.S.A. and S.M.S.; writing—review and editing, A.H.O.-A., N.A.E.-S., M.A.E.-M.M.; visualization, K.M.-G., S.M.S.; supervision, A.H.O.-A., N.A.E.-S., M.A.E.-M.M.; project administration, A.H.O.-A., N.A.E.-S.; funding acquisition. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Ethics Committee) of Faculty of Science Menoufia University (protocol code MS/315 and 1 May 2021).

**Data Availability Statement:** All data are available upon request.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Nime, M.F.; Casanoves, F.; Mattoni, C.I. Scorpion diversity in two different habitats in the Arid Chaco, Argentina. *J. Insect Conserv.* **2014**, *18*, 373–384. [[CrossRef](#)]
2. Polis, G.A.; Yamashita, T. The ecology and importance of predaceous arthropods in desert communities. In *The Ecology of Desert Communities*; Polis, G.A., Ed.; University of Arizona Press: Tucson, AZ, USA, 1991; Chapter 17; pp. 180–222.
3. Shehab, A.H.; Amr, Z.S.; Lindsell, J.A. Ecology and biology of scorpions in Palmyra, Syria. *Turk. J. Zool.* **2011**, *35*, 333–341.
4. Lighton, J.; Brownell, P.; Joos, B.; Turner, R. Low metabolic rate in scorpions: Implications for population biomass and cannibalism. *J. Exp. Biol.* **2001**, *204*, 607–613. [[CrossRef](#)]
5. Stockmann, R. Introduction to Scorpion Biology and Ecology. In *Scorpion Venoms*; Gopalakrishnakone, P., Ferroni Schwartz, E., Possani, L., Rodríguez de la Vega, R., Eds.; Springer: Dordrecht, The Netherlands, 2013.
6. Prendini, L. Scorpion diversity and distribution in southern Africa: Pattern and process. In *African Biodiversity: Molecules, Organisms, Ecosystems*; Huber, B.A., Sinclair, B.J., Lampe, K.-H., Eds.; Springer: Boston, MA, USA, 2005; pp. 25–68.
7. Druce, D.; Hamer, M.; Slotow, R. Patterns of millipede (Diplopoda), centipede (Chilopoda) and scorpion (Scorpionida) diversity in savanna habitats within the Greater Makalali Conservancy, South Africa. *Afr. Zool.* **2007**, *42*, 204–215. [[CrossRef](#)]
8. Foord, S.H.; Gelebe, V.; Prendini, L. Effects of aspect and altitude on scorpion diversity along an environmental gradient in the Soutpansberg, South Africa. *J. Arid Environ.* **2015**, *113*, 114–120. [[CrossRef](#)]
9. Chippaux, J.-P.; Goyffon, M. Epidemiology of scorpionism: A global appraisal. *Acta Trop.* **2008**, *107*, 71–79. [[CrossRef](#)]
10. Saleh, M.; Younes, M.; Badry, A.; Sarhan, M. Zoogeographical analysis of the Egyptian scorpion fauna. *Al-Azhar Bull. Sci.* **2017**, *28*, 1–14. [[CrossRef](#)]
11. Badry, A.; Younes, M.; Sarhan, M.M.; Saleh, M. On the scorpion fauna of Egypt, with an identification key (Arachnida: Scorpiones). *Zoöl. Middle East* **2018**, *64*, 75–87. [[CrossRef](#)]
12. Sarhan, M.; Badry, A.; Younes, M.; Saleh, M. Genetic diversity within *Leiurus quinquestriatus* (Scorpiones: Buthidae) populations in Egypt as inferred from 16S mtDNA sequence analysis. *Zoöl. Middle East* **2020**, *66*, 269–276. [[CrossRef](#)]
13. Kinzelbach, R.K.; Lutz, H. Stylopid larva from the Eocene—A spotlight on the phylogeny of the stylopid (Strepsiptera). *Ann. Entomol. Soc. Am.* **1985**, *78*, 600–602. [[CrossRef](#)]
14. Moustafa, A.A. Studies on Scorpions in ST. Catherine and Wadi Feiran Regions, South Sinai. Master's Thesis, Faculty of Science, Suez Canal University, Ismailia, Egypt, 1988.
15. El-Hennawy, H.K. A catalogue of the scorpions described from the Arab countries (1758–1990) (Arachnida: Scorpionida). *Serket* **1992**, *2*, 95–123.
16. El-Hennawy, H.K. Updated List of Scorpions of Egypt. *Serket* **2014**, 1–4. Available online: [http://hishamelhennawycv.blogspot.com/eg/p/blog-page\\_19.html](http://hishamelhennawycv.blogspot.com/eg/p/blog-page_19.html) (accessed on 1 July 2021).
17. Teruel, M.; McKinney, C.; Balsa, A.; Pascual-Salcedo, D.; Rodriguez-Rodriguez, L.; Ortiz, A.M.; Gómez-Vaquero, C.; González-Gay, M.A.; Smith, M.; Witte, T.; et al. Association of CD247 polymorphisms with rheumatoid arthritis: A replication study and a meta-analysis. *PLoS ONE* **2013**, *8*, e68295. [[CrossRef](#)] [[PubMed](#)]
18. Hebert, P.D.; Ratnasingham, S.; De Waard, J.R. Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. R. Soc. B Biol. Sci.* **2003**, *270*, S96–S99. [[CrossRef](#)] [[PubMed](#)]
19. Sheir, S.K.; Galal-Khallaf, A.; Mohamed, A.H.; Mohammed-Geba, K. Morphological and molecular clues for recording the first appearance of *Artemia franciscana* (Kellogg, 1906) in Egypt. *Heliyon* **2018**, *4*, e01110. [[CrossRef](#)] [[PubMed](#)]
20. Mohammed-Geba, K.; Sheir, S.K.; Hamed, E.A.E.A.; Galal-Khallaf, A. Molecular and morphological signatures for extreme environmental adaptability of the invasive mussel *Brachidontes pharaonis* (Fischer, 1870). *Mol. Cell. Probes* **2020**, *53*, 101594. [[CrossRef](#)]
21. Obuid-Allah, A.H.; El-Shimy, N.A.; Mahbob, M.A.; Ali, R.S. Survey and Morphological Studies on Scorpions Inhabiting New Valley Governorate, Egypt. *Egypt. Acad. J. Biol. Sci. B Zoöl.* **2020**, *12*, 227–238. [[CrossRef](#)]
22. Said, S.; Obuid-Allah, A.H.; El-Shimy, N.A.; Mahbob, M.A.; Ali, R.S. Ultrastructure Study on The Exo-Morphology of Four Species of Scorpion Inhabiting New Valley, Egypt. *Egypt. Acad. J. Biol. Sci. B Zoöl.* **2021**, *13*, 17–29. [[CrossRef](#)]
23. Amr, Z.S.; Baker, M.A.A.; Al-Sarairah, M.; Warrell, D.A. Scorpions and scorpion sting envenoming (scorpionism) in the Arab Countries of the Middle East. *Toxicon* **2020**, *191*, 83–103. [[CrossRef](#)]
24. Mohammed-Geba, K.; Hassab El-Nabi, S.E.; El-Desoky, M.S. Development of cytochrome-c-oxidase 1 specific primers for genetic discrimination of the European eel *Anguilla anguilla* (Linnaeus, 1758). *J. Biosci. Appl. Res.* **2016**, *2*, 258–262. [[CrossRef](#)]
25. Wasko, A.P.; Marlio, C.; Oliveira, C.; Forresli, F. Non-destructive genetic sampling in fish. An improved method for DNA extraction from fish fins and scales. *Hereditas* **2003**, *138*, 161–165. [[CrossRef](#)]
26. Folmer, O.; Black, M.; Hoeh, W.; Lutz, R.; Vrijenhoek, R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.* **1994**, *3*, 294–299.
27. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [[CrossRef](#)] [[PubMed](#)]
28. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. jModelTest 2: More models, new heuristics and parallel computing. *Nat. Methods* **2012**, *9*, 772. [[CrossRef](#)] [[PubMed](#)]

29. Ronquist, F.; Teslenko, M.; Van Der Mark, P.; Ayres, D.L.; Darling, A.; Höhna, S.; Larget, B.; Liu, L.; Suchard, M.A.; Huelsenbeck, J.P. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. *Syst. Biol.* **2012**, *61*, 539–542. [[CrossRef](#)]
30. Rambaut, A.; Drummond, A.J.; Xie, D.; Baele, G.; Suchard, M.A. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst. Biol.* **2018**, *67*, 901–904. [[CrossRef](#)] [[PubMed](#)]
31. Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **2019**, *47*, W256–W259. [[CrossRef](#)]
32. Ahmed, A.E.; Abdel-Baseer, K.A.; Saad, K.; Hassan, A.F.; El-Houfey, A.A. Endocrinological and biochemical changes of scorpionism in children in Upper Egypt. *Ther. Adv. Endocrinol. Metab.* **2015**, *6*, 210–216. [[CrossRef](#)]
33. Saad, K.; El-Hamed, M.A.A.; Abo-Elela, M.G.M.; Ahmed, A.E.; Abdel-Baseer, K.A.; Aboul-Khair, M.D.; Metwalley, K.A.; El-Houfey, A.A.; Hasan, G.M.; El-Shareef, A.M. Neurologic Complications in Children with Scorpionism: A Retrospective Study in Upper Egypt. *J. Child Neurol.* **2017**, *32*, 537–542. [[CrossRef](#)] [[PubMed](#)]
34. Bahloul, M.; Rekik, N.; Chabchoub, I.; Chaari, A.; Ksibi, H.; Kallel, H.; Damak, H.; Chaari, A.; Ben Hamida, C.; Chelly, H.; et al. Neurological complications secondary to severe scorpion envenomation. *Med Sci. Monit.* **2005**, *11*, CR196–CR202.
35. Amin, R.; Faiz, A. Scorpion sting and envenomation. In *Clinical Toxinology in Asia Pacific and Africa*; Gopalakrishnakone, P., Faiz, A., Fernando, R., Gnanathasan, C., Habib, A., Yang, C.C., Eds.; Toxinology; Springer: Dordrecht, The Netherlands, 2014; Volume 2, pp. 455–482.
36. Barros, R.M.; Pasquino, J.A.; Peixoto, L.R.; Targino, I.T.G.; Sousa, J.A.D.; Leite, R.D.S. Clinical and epidemiological aspects of scorpion stings in the northeast region of Brazil. *Ciência Saúde Coletiva* **2014**, *19*, 1275–1282. [[CrossRef](#)]
37. Al-Asmari, A.K.; Al-Saif, A.A.; Abdo, N.M. Morphological identification of scorpion species from Jazan and Al-Medina Al-Munawara regions, Saudi Arabia. *J. Venom. Anim. Toxins Incl. Trop. Dis.* **2007**, *13*, 821–843. [[CrossRef](#)]
38. Sousa, P.; Froufe, E.; Alves, P.C.; Harris, D.J. Genetic diversity within scorpions of the genus *Buthus* from the Iberian Peninsula: Mitochondrial DNA sequence data indicate additional distinct cryptic lineages. *J. Arachnol.* **2010**, *38*, 206–211. [[CrossRef](#)]
39. Štundlová, J.; Šmíd, J.; Nguyen, P.; Šťáhlavský, F. Cryptic diversity and dynamic chromosome evolution in Alpine scorpions (Euscorpiidae: Euscorpium). *Mol. Phylogenetic. Evolut.* **2019**, *134*, 152–163. [[CrossRef](#)] [[PubMed](#)]
40. Omran, M.A.; McVean, A. Intraspecific variation in scorpion *Leiurus quinquestriatus* venom collected from Egypt (Sinai and Aswan deserts). *J. Toxicol. Toxin Rev.* **2000**, *19*, 247–264. [[CrossRef](#)]
41. Gantenbein, B.; Largiader, C.R. The phylogeographic importance of the Strait of Gibraltar as a gene flow barrier in terrestrial arthropods: A case study with the scorpion *Buthus occitanus* as model organism. *Mol. Phylogenet. Evol.* **2003**, *28*, 119–130. [[CrossRef](#)]
42. Alqahtani, A.R.; Badry, A. Interspecific phylogenetic relationship among different species of the genus *Buthacus* (Scorpiones: Buthidae) inferred from 16S rRNA in Egypt and Saudi Arabia. *Zool. Middle East* **2020**, *66*, 178–185. [[CrossRef](#)]
43. Luna-Ramirez, K.; Miller, A.D.; Rašić, G. Genetic and morphological analyses indicate that the Australian endemic scorpion *Urodacus yaschenkoii* (Scorpiones: Urodacidae) is a species complex. *PeerJ* **2017**, *5*, e2759. [[CrossRef](#)] [[PubMed](#)]
44. Zhang, X.S.; Liu, G.M.; Zhang, D.X.; Shi, C.M. Genetic analysis and ecological niche modelling delimit species boundary of the Przewalski's scorpion (Scorpiones: Buthidae) in arid Asian inland. *BioRxiv* **2020**, 652024. [[CrossRef](#)]
45. Soltan-Alinejad, P.; Rafinejad, J.; Dabiri, F.; Onorati, P.; Terenius, O.; Chavshin, A.R. Molecular analysis of the mitochondrial markers COI, 12S rDNA and 16S rDNA for six species of Iranian scorpions. *BMC Res. Notes* **2021**, *14*, 40. [[CrossRef](#)]
46. Lourenço, W.R. Scorpion incidents, misidentification cases and possible implications for the final interpretation of results. *J. Venom. Anim. Toxins Incl. Trop. Dis.* **2016**, *22*, 1–25. [[CrossRef](#)] [[PubMed](#)]
47. Hendrixson, B.E. Buthid scorpions of Saudi Arabia, with notes on other families (Scorpiones: Buthidae, Liochelidae, Scorpionidae). *Fauna Arab.* **2006**, *21*, 120.
48. Lourenço, W.R. Why does the number of dangerous species of scorpions increase? The particular case of the genus *Leiurus* Ehrenberg (Buthidae) in Africa. *J. Venom. Anim. Toxins Incl. Trop. Dis.* **2020**, *26*, e20200041. [[CrossRef](#)] [[PubMed](#)]
49. Sadine, S.E. New locality of *Orthochirus innesi* Simon, 1910 in Algeria (Scorpiones: Buthidae). *Serket* **2020**, *17*, 171–175.
50. Husemann, M.; Schmitt, T.; Stathi, I.; Habel, J.C. Evolution and Radiation in the Scorpion *Buthus elmoutaouakili* Lourenço and Qi 2006 (Scorpiones: Buthidae) at the Foothills of the Atlas Mountains (North Africa). *J. Hered.* **2012**, *103*, 221–229. [[CrossRef](#)] [[PubMed](#)]
51. Kartavtsev, Y.P.; Rozhkovan, K.V.; Masalkova, N.A. Phylogeny based on two mtDNA genes (Co-1, Cyt-B) among Sculpins (Scorpaeniformes, Cottidae) and some other scorpionfish in the Russian Far East. *Mitochondrial DNA Part A* **2016**, *27*, 2225–2240. [[CrossRef](#)]
52. Devi, S.R.; Thomas, A.; Rebijith, K.B.; Ramamurthy, V.V. Biology, morphology and molecular characterization of *Sitophilus oryzae* and *S. zeamais* (Coleoptera: Curculionidae). *J. Stored Prod. Res.* **2017**, *73*, 135–141. [[CrossRef](#)]