



Annual Research & Review in Biology

Volume 39, Issue 7, Page 40-45, 2024; Article no.ARRB.117857
ISSN: 2347-565X, NLM ID: 101632869

(Past name: Annual Review & Research in Biology, Past ISSN: 2231-4776)

Spider Species Identification: Bridging Traditions with Molecular and Deep Learning Approaches

Aakriti Shrivastava ^{a*} and V. K. Sharma ^a

^a Department of Zoology, Govt. Holkar Science College, Indore, India.

Authors' contributions

This work was carried out in collaboration between both authors. Both authors read and approved the final manuscript.

Article Information

DOI: <https://doi.org/10.9734/arrb/2024/v39i72097>

Open Peer Review History:

This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: <https://www.sdiarticle5.com/review-history/117857>

Review Article

Received: 15/04/2024

Accepted: 18/06/2024

Published: 04/07/2024

ABSTRACT

This article provides an in-depth review of various methods employed in the identification and sequencing of spiders, highlighting the advancements and challenges in the field. With the increasing importance of spiders in ecological studies, medical research, and biodiversity conservation, accurate identification and genetic analysis have become crucial. This review discusses traditional and modern techniques, shedding light on their applications, limitations, and future prospects.

The exploration begins with an analysis of taxonomists' etymological choices, examining patterns in naming conventions across continents and centuries. Traditional morphological identification, anchored in backbone taxonomy, dichotomous keys, and statistical analyses, highlights the advantages and challenges of relying on observable features. The study transitions to molecular techniques, elucidating the applications and challenges of DNA barcoding, Next-Generation Sequencing (NGS), and metabarcoding in spider identification. The integration of deep learning models, exemplified by the YOLOv7-based Spider Identification APP, represents a landmark in

*Corresponding author: E-mail: aakritishrivastava@yahoo.com;

Cite as: Shrivastava, Aakriti, and V. K. Sharma. 2024. "Spider Species Identification: Bridging Traditions With Molecular and Deep Learning Approaches". *Annual Research & Review in Biology* 39 (7):40-45. <https://doi.org/10.9734/arrb/2024/v39i72097>.

computer vision for efficient and user-friendly spider species recognition. The study's multifaceted approach provides a nuanced understanding of spider taxonomy, bridging historical practices with state-of-the-art technologies, and lays the groundwork for future advancements in the field.

Keywords: Spider identification; morphological taxonomy; DNA barcoding; next-generation sequencing; metabarcoding; deep learning models; biodiversity analysis.

1. INTRODUCTION

Spiders play a vital role in ecosystems, and their identification and sequencing have gained significance in recent years. This section introduces the importance of spider identification and the role of genetic sequencing in understanding their biology, behavior, and evolutionary relationships. Understanding the complexities of the natural world has been a longstanding human endeavor, characterized by the systematic classification of elements [1]. This tradition, evident in Mendeleev's periodic table and stellar categorization, extends to biological taxonomy. Linnaeus's binomial nomenclature, proposing species identification through genus and species epithets, remains a lasting system. This study explores taxonomists' etymological choices in spider taxonomy, scrutinizing patterns across time and continents, with a focus on potential cultural influences. The investigation traverses traditional methods, emphasizing the significance of morphological identification and molecular techniques, offering a holistic perspective on spider taxonomy.

2. DIFFERENT EXPERIMENTAL APPROACH FOR SPECIES IDENTIFICATION

2.1 Traditional Methods

Throughout history, humans have organized natural elements like stars and species into systematic frameworks. Linnaeus introduced binomial nomenclature, defining species through genus and species names. Taxonomists creatively name species, drawing inspiration from traits, habitats, individuals, or wordplay (Scerri, 2019; Langer & Kudritzki, 2014; Linnaeus, 1751, 1758; Lehman, 1967; Hey, 2001; Slater, 2016) [2]. Our study examines these naming patterns in spider taxonomy to reveal potential cultural influences.

2.2 Morphological Identification

Traditional taxonomic methods rely on morphological characteristics, including body structure, coloration, and genital morphology. In

this study, spider samples were collected from Ratapani Wildlife Sanctuary Bhopal (M.P.). The range of sanctuary covers maximum area of Raisen and Sehore districts, with approximate geographical coordinates of Latitude: 23.4346° N and Longitude: 77.4980° E.

2.3 Sample Organism

The specimen I collected belongs to the Nephila family. The Nephila family of spiders is commonly found in the area surrounding Ratapani Wildlife Sanctuary. The Nephila spider, also known as the Giant Wood Spider or Golden Orb Weaver, belongs to a family known for its large, colorful spiders. These spiders are famous for their impressive size, large size of webs, and striking appearance. These spiders are the major part of food chain to minimize the population of insects.

N. pilipes exhibit female gigantism and male dwarfism, with significant size differences between males and females, particularly notable among terrestrial animals. This sexual dimorphism is likely a result of evolutionary selection favoring females with higher fecundity [3]. Female *N. pilipes* invest heavily in parental care, including egg production and web construction, demonstrating substantial commitment to their progeny.

Female: Female *N. pilipes* are usually 30–50 mm in size, with a cephalothorax measuring about 15 mm long and 10 mm wide, and an abdomen around 30 mm long and 15 mm wide. They are mostly dark yellow-brown with yellow stripes, black or brown tergum covered in dense hairs, and long black and yellow legs lacking apparent hairs. The color contrast aids in foraging success with visually-oriented prey.

Male: Male *N. pilipes* typically measure 5–6.5 mm, with a cephalothorax about 2.5 mm long and 2 mm wide, and an abdomen approximately 4 mm long and 1.5 mm wide. They have light brown legs with some hairs, larger front eyes than rear eyes, and a yellow carapace with very few hairs.



Image 1. Ratapani Wildlife Sanctuary Latitude: 23.4346° N Longitude: 77.4980° E.(Image taken by Google Map).



Image 2. Image of *Nephila pilipes* from Ratapani Wildlife Sanctuary, Bhopal (M.P.), highlighting the morphological features that illustrate sexual dimorphism between the female and male.

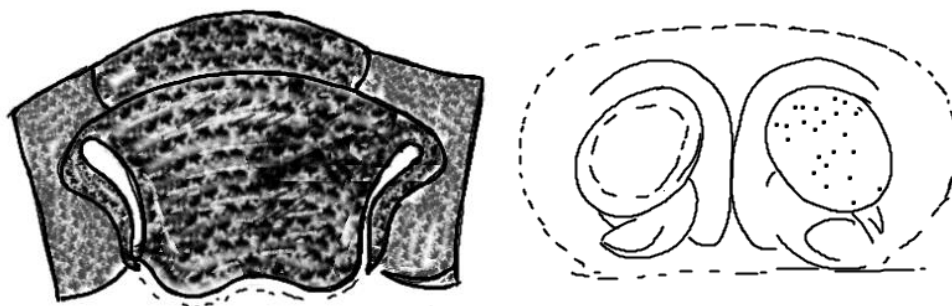


Image 3. Recreate a ray diagram of the female genital organ of *Nephila pilipes* as described by Kuntner et al. [4]

2.4 Backbone Taxonomy

The backbone taxonomy was derived from the World Spider Catalog (2020) with 48,464 taxonomic entities, spanning back to Carl Alexander Clerck's 1757 work. To align with zoological nomenclature, Clerck's publication

date was adjusted to January 1, 1758 (Clerck, 1757; Linnaeus, 1758).

2.5 Etymology Classification and Validation

Species etymologies were categorized into six groups, covering morphology, ecology,

geography, people, modern/past culture, and others. Etymologies were assigned based on original descriptions, supplemented by inference using language expertise, etymological references, and collaboration [5,6]. Cross-validation of 400 randomly selected inferred etymologies demonstrated internal consistency, with 57.75% matching original descriptions.

2.6 Statistical Analyses

All analyses were conducted in R v.4.0.3 (R Core Team, 2021) using 'ggplot2' v.3.3.4 (Wickham, 2016). Hypotheses were tested through regression models with a significance threshold set at 0.001. Temporal trends in etymology

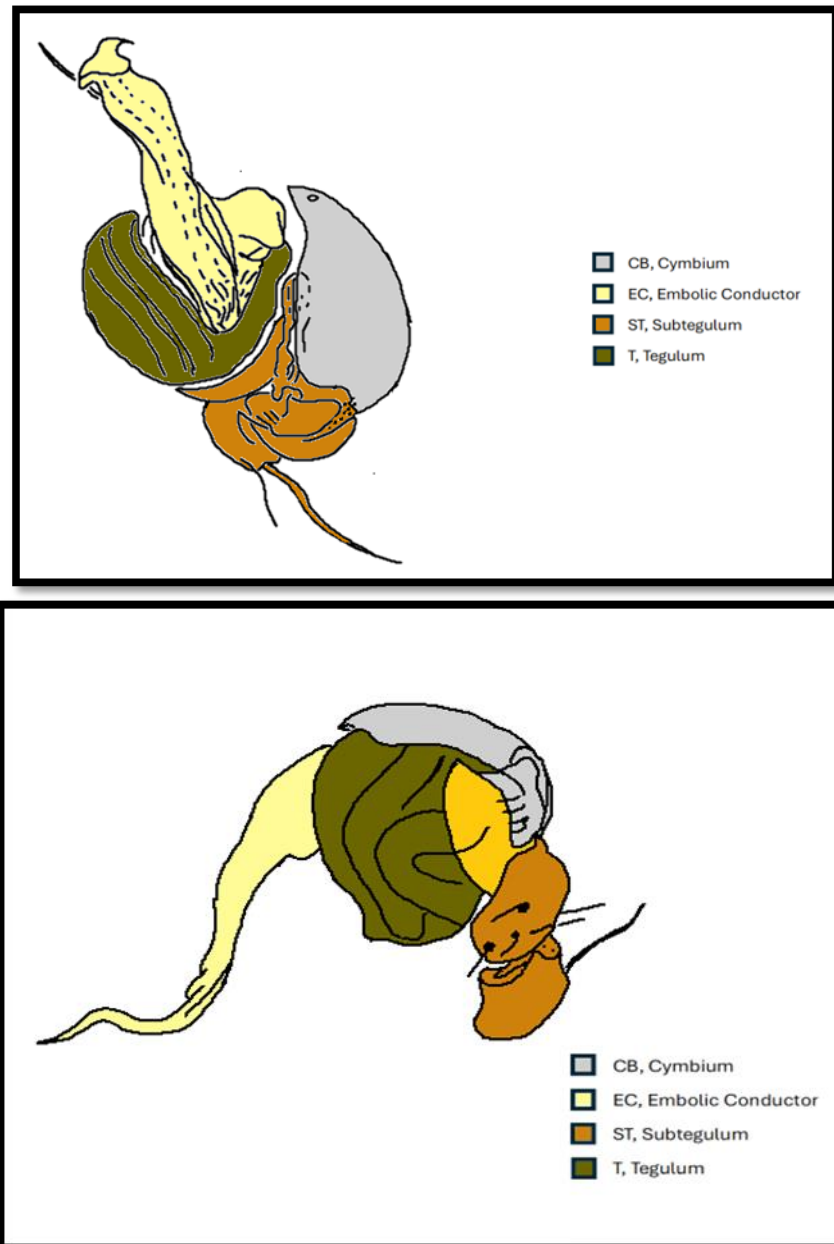


Image 4. Recreate the diagram of the male genital organ of *Nephila pilipes*, using different colors to label the following parts: CB (cymbium), E (embolus), EC (embolic conductor), Ecp (proximal part of the embolic conductor), Ecd (distal part of the embolic conductor), ST (subtegulum), and T (tegulum) as shown by Kuntner et al. [4]

choice were analyzed using a generalized additive model. A second model assessed temporal trends across continents, considering a binomial distribution. Generalized linear models examined the likelihood of taxonomists selecting a specific etymology category across continents (Wood, 2004; Zuur et al., 2009) [7].

2.7 Species Geographic Distribution

Species distributions were classified at the continental level using World Spider Catalog data and converted into ISO country codes [8,9].

2.8 Dichotomous Keys

Dichotomous keys, such as Levi's orb-weaving spider key [10], guide users through choices leading to species identification based on observable characteristics. Their value lies in structured decision-making, aiding accurate identifications [10].

3. MOLECULAR TECHNIQUES

3.1 DNA Barcoding

DNA barcoding uses specific DNA sequences for species identification. Spider specimens were collected, and their genomic DNA was extracted. The COI gene was amplified via PCR and sequenced. Sequences were compared to BOLD and GenBank databases using BLAST for species identification, following Hebert et al. [11] protocols.

3.2 Next-generation Sequencing (NGS)

NGS technologies have revolutionized genomics, enabling large-scale spider sequencing studies. Specimens were collected, and DNA extracted using a commercial kit (Manufacturer's Manual). Libraries were prepared with the Nextera XT Kit and sequenced on an Illumina HiSeq platform. Data were processed with Trimmomatic, assembled with SPAdes, and annotated using BLAST [12].

3.3 Metabarcoding

Metabarcoding enables simultaneous identification of multiple species from environmental DNA samples. In spider community analysis, DNA is non-invasively collected, extracted, amplified using universal primers, and sequenced. Bioinformatics pipelines process the data for community analysis using R or QIIME software (Edgar, 2013; Caporaso et al., 2010) [13,14].

3.4 Deep Learning Models

Our YOLOv7-based Spider Identification APP (SpiderID_APP) efficiently identifies spider genera in Taiwan, leveraging deep learning. Trained on 24,000 iNaturalist images, it provides 120 genus classifications. Complementary methods like DNA barcoding are needed for species-level identification. The APP features a user-friendly GUI, integrating YOLO detection scripts [13,14].

4. RESULTS AND DISCUSSION

The comprehensive study delves into spider identification and sequencing methods, encompassing traditional approaches and cutting-edge technologies. The analysis of taxonomists' etymological choices revealed temporal trends and internal consistency, reflecting a balance between historical adherence and evolving conventions. Backbone taxonomy exploration and dichotomous keys highlighted the significance of morphological identification, recognizing its value while emphasizing the need for continuous updates. In this study, I observed *Nephila* spiders found in Ratapani Wildlife Sanctuary and conducted morphological identification based on previous scientific research. Molecular techniques, including DNA barcoding, Next-Generation Sequencing (NGS), and metabarcoding, showcased their efficacy in spider genomics and biodiversity studies. The YOLOv7-based Spider Identification APP represented a milestone in computer vision, seamlessly integrating with traditional methods. Challenges such as sample contamination and incomplete databases were identified, emphasizing the ongoing need for methodological refinement. In summary, the results provide a multifaceted understanding of spider taxonomy, bridging historical practices with state-of-the-art technologies and paving the way for future advancements in the field.

5. CONCLUSION

This review explores various methods used to identify and sequence spiders. As technology advances, combining traditional and molecular approaches improves accuracy and efficiency in spider taxonomy and genomics. Integrating these methods is crucial for understanding spider diversity, ecology, and evolution. Despite challenges like sample contamination and limited databases, the field progresses. Future advancements, including artificial intelligence and portable sequencing devices, promise

greater strides. Embracing these technologies will deepen our understanding of spider diversity and evolution. The interdisciplinary nature of spider identification ensures a strong foundation for uncovering the complexities of these fascinating creatures.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Bjerge K, Mann HM, Høye TT. Real-time insect tracking and monitoring with computer vision and deep learning. *Remote Sensing in Ecology and Conservation*. 2022;8(3):315-27.
2. Vendetti J, Garland R. Species name formation for zoologists: A pragmatic approach. *Journal of Natural History*. 2019;53(47-48), 2999-3018.
3. Kuntner M, Elgar MA. Evolution and maintenance of sexual size dimorphism: Aligning phylogenetic and experimental evidence. *Frontiers in Ecology and Evolution*. 2014;2:26.
4. Kuntner M, Coddington JA, Schneider J. Intersexual arms race? Genital coevolution in nephilid spiders (Araneae, Nephilidae). *Evolution*. 2009;63(6):1451–1463.
5. Bonnet P. *Bibliographia Araneorum*. Methodical analysis of all araneological literature until 1939. Volume III. Toulouse: Douladoure Printing Company; 1961.
6. Nilsson AN. All diving beetle specific and subspecific names explained. *Skörvnöpparn*, Umeå Supplement. 2010;1:1–42.
7. Zuur AF, Ieno EN. A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*. 2016;7(10):1121–1131.
8. Cardoso P, Pekár S. wscmap: A world spider catalog map. *Biodiversity Data Journal*. 2022;10:e79648.
9. Arel-Bundock V, Enevoldsen N, Yetman CJ. Country code: Convert country names and country codes. *Journal of Open Source Software*. 2018;3(28):1025.
10. Levi HW. The spider genera *Enoplognatha*, *Theridion*, and *Paidisca* in America north of Mexico. *American Museum Novitates*. 1957;1832:1-39.
11. Hebert PD, Ratnasingham S, deWaard JR. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*. 2003;270(1512): 313-321.
12. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *Journal of Molecular Biology*. 1990;215(3):403-410.
13. Elbrecht V, Leese F. PrimerMiner: An R package for development and in silico validation of DNA metabarcoding primers. *Methods in Ecology and Evolution*. 2017;8(5):622-626.
14. Deagle BE, Thomas AC, Shaffer SA, Trites AW, Jarman SN. DNA metabarcoding and the cytochrome c oxidase subunit I marker: Not a perfect match. *Biological Letters*. 2014;10(9): 20140562.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of the publisher and/or the editor(s). This publisher and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

© Copyright (2024): Author(s). The licensee is the journal publisher. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history:

The peer review history for this paper can be accessed here:

<https://www.sdiarticle5.com/review-history/117857>