JOURNAL OF
EMERGING INVESTIGATORS

Article

Friend or foe: Using DNA barcoding to identify arthropods found at home

Maggie E. Wang[1], Xiaoqun C. Zhang[2]
[1] Northport High School, Northport, New York
[2] Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

SUMMARY

Arthropoda make up the largest phylum in the animal kingdom. A common assumption is that many of these arthropods that people encounter in their homes are harmful, such as insects and spiders, and people react with the use of pesticides as a result. This means that many people are exposed to pesticides at home. Are arthropods our enemies who deserve to be exterminated at any cost? Our hypothesis was that all arthropods found at home are harmful to human health. We collected arthropods, whole or partial, found inside one residential house over 12 months. We used both morphological characters (field guide) and DNA barcoding to identify them. We checked identified species to see if it was on the pest lists provided by US government, but unexpectedly, none were. Therefore, we concluded that it is a misconception that arthropods found at home are harmful to humans. Furthermore, we suggest that DNA barcoding technology, if made readily available, would be an accurate method for citizens to identify arthropods at the species level, which may help people to avoid overusing pesticides and to reduce pollution. If our experimental strategy were employed at a larger scale, the data generated could help scientists to better understand the evolution of the largest group of animals on Earth and aid in mapping the ecosystem we live in.

INTRODUCTION

Arthropods are invertebrate animals with jointed legs and a skeleton outside of the body. Insects and arachnids are very common arthropods living among humans (1). Many people assume that arthropods found at home can hurt us and subsequently use pesticides to exterminate them. According to a recent report by the United States Environmental Protection Agency, 75% of American households used at least one pesticide indoors in 2020 (2). Most often, the pesticides used are for the purpose of killing arthropods. Numerous people have been exposed to up to 12 different pesticides in the air inside their homes, which could potentially be harmful to their health (2). In 1990, the American Association of Poison Control Centers (AAPC) reported that about 79,000 children were involved in common household pesticide poisonings or exposures (3). In 2008, pesticide poison was the 9th most common substance reported to poison control; 45% of poisoning cases involved children (4). In 2019, there were still 77,707 exposures recorded by AAPC; pesticide exposure was the 3rd most common reported poison associated with pregnancy (5).

We wanted to assess if arthropods found at home are harmful and should therefore warrant pesticide use by conducting a scientific investigation. We hypothesized that arthropods found at home are harmful to human health. We collected arthropods, whole or partial, found inside one residential house during a 12-month period. We used both traditional taxonomies, which requires the whole animal to exam the morphologic characters to identify the samples at the species level, and deoxyribonucleic acid (DNA) barcoding, a molecular method which requires only a small piece of tissue to identify the samples at the species level, to identify collected samples (6–11). We conducted literature research on each species identified to decide whether that species was a foe, which means it is harmful to human health, or a friend, which means it is not harmful to humans (12–14). An overview of the workflow is shown in **Figure 1**. To our surprise, our data does not support the hypothesis. Therefore, pesticide use is not necessary in most instances of arthropod sightings in homes.

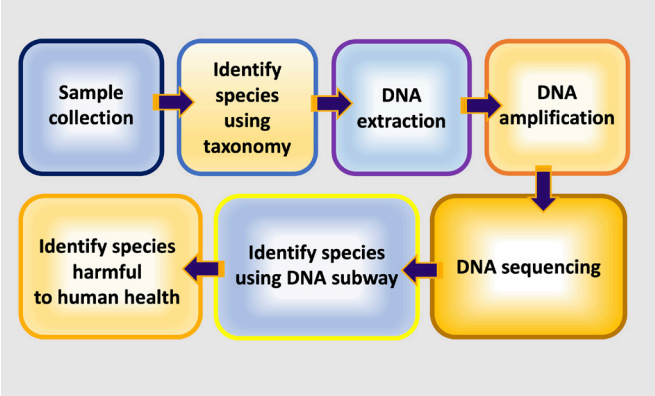

Figure 1: The experiment workflow. Schematic representation to illustrate the experiment workflow for sample collection and identification.

Sample Condition	Common Name	Taxonomy Identification	DNA Barcoding Identification
Partial, dead	Lady bug	*Coleoptera*	*Hippodamia convergens*
Partial, dry	Woodlouse	*Crustaceans*	*Porcellio scaber*
Partial, dry	Woodlouse	*Crustaceans*	*Porcellio scaber*
Whole, dead	Bug	*Artheneidae*	*Chilacis tyhae*
Partial, dead	Spider	*Araneae*	*Pholcus manueli*
Partial, dead	Beetle	*Coleoptera*	*Anthrenus verbasci*
Whole, dead	Stink bug	*Pentatomidae*	*Halyomorpha halys*
Partial, dry	Click beetle	*Elateridae*	*Melanotus communis (Gyllenhal, 1817)*
Partial, dead	Cricket	*Orthopteran*	*Velarifictorus micado*
Partial, dry	Stink bug	*Pentatomidae*	*Halyomorpha halys*
Whole, dead	Spider	*Araneida*	*Steatoda triangulosa*
Whole, dead	Spider	*Araneida*	*Steatoda triangulosa*
Whole, dead	Moth	*Lepidoptera*	*Amphipyra pyramidoides*
Partial, shell	Millipede	*Diplopoda*	*Oxidus gracilis*
Whole, dead	Mayfly	*Ephemeroptera*	*Cloeon dipterum*
Whole, dead	Beetle	*Coleoptera*	*Xylopinus saperdioides*
Whole, dead	Carpenter Ant	*Camponotus*	*Camponotus pennsylvanicus*
Partial, dead	Spider	*Araneida*	*Pholcus manueli*
Whole, dead	Click beetle	*Elateridae*	*Hemicrepidius memnonius*
Whole, dead	Spider	*Araneida*	*Trochosa ruricola*

Table 1: Identification of 20 arthropod corpses using both taxonomy and DNA barcoding techniques. Table showing samples identified with their condition and common names. Books and field guides from Iowa State University were used to identify samples with the morphologic characteristics for taxa identification. For DNA barcoding identification: high quality COI DNA sequences generated from samples were matched with GenBank database provided by the National Institute of Science to identify species via DNA Subway system developed by DNALC.

RESULTS

A total of 34 samples were collected from one home located on Long Island, New York, over 12 months. All of the samples collected were corpses of arthropods and demonstrated various levels of decomposition. If the sample body was complete, we labeled it as "whole". Many of the samples were not complete animals, and we labeled them as "partial". Depending on the condition, some of the samples were identified to the order level, and some of the samples were identified to the family level by checking the morphological characters only **(Table 1)**. We then moved on to using a DNA barcoding strategy to identify all the samples. We chose the mitochondrial cytochrome c oxidase subunit 1 (COI) region for barcoding. We successfully amplified the COI region in 27 of the 34 samples by polymerase chain reaction (PCR), which was shown by a single clear band sized between 600bp-700bp **(Figure 2)**. Of these, we were able to identify 20 samples to the species level by using high-quality sequences of the PCR products and the Basic Local Alignment Search Tool (BLAST) in DNA Subway **(Table 1)**.

Of the DNA sequences we generated, 18 sequences between 468bp and 652bp were published in GenBank, which is part of the International Nucleotide Sequence Database Collaboration, and also available through the DNA DataBank of Japan (DDBJ) and the European Nucleotide Archive (ENA). Among the 18 published DNA COI sequences, 3 of them were new entries to the GenBank sequence data base at the time they were submitted. The *Xylopinus saperdioides* sequence generated from our sample (sample ID: 5D) had a total of 8bp of mismatch across the COI sequence when compared with the same species sequences already existing in the GenBank nucleotide database. Similarly, the *Melanotus communis* sequence generated from our sample (sample ID: 502) had a total of 3bp of mismatch, and the *Hemicrepidius memnonius* sequence from our sample (sample ID: 5J) had a total of 6bp of mismatch when submitted **(Table 2)**.

The most commonly identified class of arthropods were arachnids (spiders). Of the five spider samples collected, we identified three species: *Trochosa ruricola*, *Pholcus manueli*, and *Steatoda triangulosa*. We also identified four beetle species found inside the house: *Melanotus communis*, *Hemicrepidius memnonius*, *Xylopinus saperdioides*, and *Hippodamia convergens*. We also identified one ant species (*Camponotus pennsylvanicus*), one moth species (*Amphipyra pyramidoides*), one mayfly species (*Cloeon dipterum*), one cricket species (*Velarifictorus micado*), one species of stink bug (*Halyomorpha halys*) and one species of woodlouse (*Porcellio scaber*) **(Table 2)**. None of the species we identified were on lists of pests with significant public health and/or clinical importance (12-14). In summary, none of the species we observed were identified in any of the primary references of clinically important insect pests.

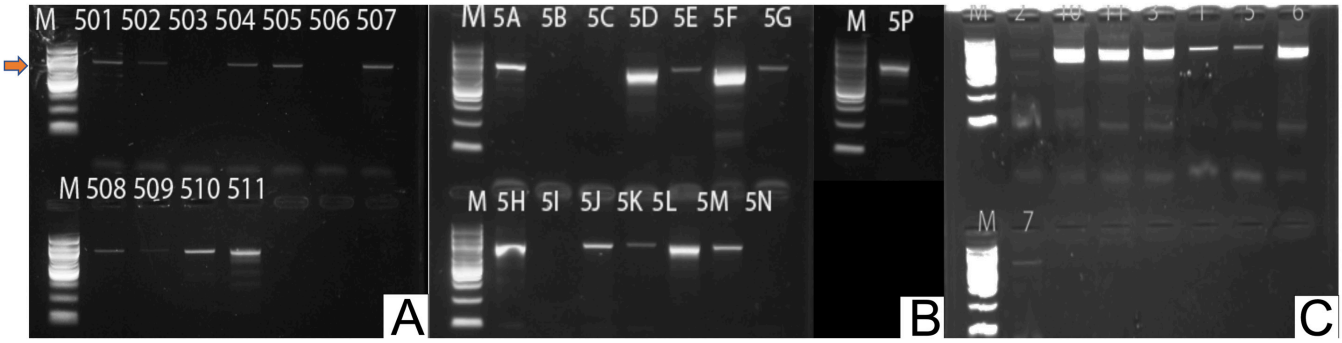

Figure 2: DNA amplification results. Panel A, B, C showing gel electrophoresis results from total 34 samples studied. Sample ID is on the top of each line. "M" is for 100bp DNA marker. The arrow points to the 600bp DNA fragment.

Sample ID	Species Name Identified by DNA Barcoding	Published GenBank Accession No.	Sequence Length (bp)
1	*Hippodamia convergens*	MK580639.1	473
3	*Porcellio scaber*	MK585618.1	626
5	*Porcellio scaber*	MN190299.1	577
7	*Pholcus manueli*	MN190301.1	514
11	*Halyomorpha halys*	MK585620.1	631
502	*Melanotus communis (Gyllenhal, 1817)*	MN610009.1	468
504	*Velarifictorus micado*	MN597064.1	501
505	*Halyomorpha halys*	MN597065.1	508
507	*Trochosa ruricola*	MN319605.1	652
509	*Steatoda triangulosa*	MN610011.1	371
511	*Amphipyra pyramidoides*	MN597066.1	649
5L	*Oxidus gracilis*	MN319604.1	648
5A	*Cloeon dipterum*	MN319599.1	637
5D	*Xylopinus saperdioides*	MN319600.1	646
5F	*Camponotus pennsylvanicus*	MN319601.1	583
5H	*Pholcus manueli*	MN319602.1	583
5J	*Hemicrepidius memnonius*	MN319603.1	580
5P	*Trochosa ruricola*	MN610010.1	516

Table 2: Research-quality DNA sequences published in GenBank. Table showing the COI sequences published in GenBank nucleotide database with their sample ID, GenBank accession number, and length.

DISCUSSION

Identifying the arthropods living in our homes is important for us in order to find the proper way to handle them. At the species level, the identification information will help us to decide which arthropods are harmful and which arthropods are harmless. However, taxa identification could be difficult for untrained eyes. Additionally, sometimes we may only find decomposed parts of the animal, which could make it more difficult to do taxonomic identification at the species level (6,7). At home, we may see part of an arthropod, such as a head or leg lying in the dust and wonder: what kind of arthropods do these parts belong to? Should we call exterminators or buy heavy duty pesticide chemicals to clean up the house in order to protect our family? For ordinary people not trained as entomologists, it could be hard to answer these questions.

Recent research shows that DNA barcoding can be used to identify species using very small amounts of the sample tissue (8–11). As an independent citizen-science project, we identified arthropods at the species level using DNA barcoding so that we could distinguish if they were friends or foes. Surprisingly, all 17 of the species we identified are considered harmless to human health, and no foes were detected (12–14). Therefore, it is not necessary to use pesticide spray inside the house that we studied. Our findings suggest that, if we make DNA barcoding technology available to ordinary people, it could be an accurate way to identify arthropods at the species level and to avoid overusing pesticides, thereby reducing chemical pollution. A service using DNA barcoding method to help identify pests may be developed as a business to help communities without access to such technology.

Some species we identified may look intimidating, but actually are our friends and could be used to benefit humans. For example, *Porcellio scaber*, was first found in Britain, but now is widespread in Europe, Asia, North America, South Africa, and Australia. Because of its worldwide distribution and its high metal accumulation capacity, a previous study proposed that *Porcellio scaber* could be used as a suitable bioindicator for metal contaminated soils (15). Some of the arthropods we identified could even protect us. For example, *Steatoda triangulosa* can prey on ticks and several other spiders believed to be harmful to humans (16). Some of the arthropods we identified could be model organisms for science research. For example, research showed that *Cloeon dipterum* could be established as a new model system to investigate insect evolution (17). Some of the insects we identified could be our allies. For example, *Hippodamia convergens* is commonly known as the ladybug and are used for the biological control of some pests (18).

Although some of the species we identified are not harmful to human health, they are considered pests in other settings. For example, *Halyomorpha halys* could be harmful to certain plant crops and may need pest control on farms (14). If *Camponotus pennsylvanicus* nests are found in buildings, they could be very harmful to wooden structures. However, if there are no nests in the building, it may not be necessary to use chemical pesticide in the house (19).

Because our findings were limited to one household, our study is preliminary. Scaling up the study in multiple locations for in the future may allow us to draw more solid conclusions. It has been shown that our indoor communities of arthropods are more strongly influenced by the environment outside the house than by how tidily we live inside the house (20). If we can expand the collection of samples to sites all over the country, future research results might help us to understand the arthropods in North American residential areas. Such data could also help in mapping the ecosystem we live in. Popularizing citizen-science projects like ours will also help to educate people about not being afraid of arthropods, thereby avoiding overuse of pesticides. Thus, we could reduce chemical pollution, promote biodiversity, and protect our ecosystem and human well-being.

MATERIALS AND METHODS

Sample Collection

Insect samples were collected from one residential house (Long Island, NY) where we have the full accessibility over a one-year period. The samples were placed in individual 1.5-ml microcentrifuge tubes labeled with the sample ID and stored at a -20°C freezer until DNA extraction. Photos of the samples with the collection time, condition, location, and other information such as latitude, longitude, and altitude were saved.

Taxonomy Identification

Photos of each sample were taken with rulers next to it using a dissecting microscope. Books and internet guidance

were used for taxonomy identification (6,7).

DNA Extraction

Sterilized razor blades and tweezers were used to remove animal tissue. Soft tissues of approximately 10 mg, or ⅛- to ¼-inch diameter from the leg or neck were used for DNA extraction to avoid massive damage to the sample while yielding the best results for DNA extraction. Each sample was placed in a clean 1.5 mL microcentrifuge tube labeled with an identification number. Lysis buffer (6M Guanidine Hydrochloride [GuHCl], 300 µl) was added into the tube. After twisting a clean plastic pestle (VWR, Pennsylvania, USA, catalog number: 47750-354) against the inner surface of the 1.5 mL tube to forcefully grind the tissue for 2 minutes, the tubes were placed in a water bath at 65°C for 10 minutes. The tubes were then placed in a centrifuge to spin for one minute at 12,100 X g to pellet the cellular debris. The clear supernatant (150 µl) was transferred to a fresh tube labeled with the sample ID. Silica resin (Sigma-Aldrich, Missouri, USA, catalog number: S5631-100G) (3 µl) was added to the tube, then mixed well, and the tube was incubated for 5 minutes in a water bath at 57°C. The tube was centrifuged for 30 seconds, and the supernatant removed. Ice-cold wash buffer (0.05M NaCl, 0.02M Tris, 0.001M EDTA, 50% EtOH) (500 µl) was added to the pellet, and then the solution was mixed well by pipetting up and down, spun down for 30 seconds, and the supernatant removed. The silica pellet was washed with the ice-cold wash buffer again and the supernatant removed. Distilled water (100 µl) was added to the silica resin and mixed well by vortexing or by pipetting up and down. The mixture was incubated at 57°C for 5 minutes, then spun down for 30 seconds, after which 50 µl of the supernatant, containing DNA, was transferred into a new tube.

DNA Amplification

The COI primers mix (0.26µM primers: LCO1490:5'TGTAAAACGACGGCCAGTGGTCAACAAATCATAAAGATATTGG-3', HC02198:5'-CAGGAAACAGCTATGACTAAACTTCAGGGTGACCAAAAAATCA-3', with 40% Cresol Red loading dye) (23 µl) was added to each ready-to-go PCR tube (VWR, Pennsylvania, USA, catalog number:89497-136), and then the beads were allowed to dissolve for 1 minute at ambient temperature. DNA solution (2 µl) was then added into the PCR tube, and the tube was placed into a thermocycler to run the following program for PCR. Initial step was at 94°C for 1 minute, then followed by 35 cycles of the following profile: 95°C for 30 seconds, 50°C for 30 seconds and 72° C for 45 seconds. One final step was used to preserve the sample: 4°C for infinitum.

DNA Sequencing

PCR product (5 µl) was loaded and run on a 2% agarose gel with Gel Red. 100bp DNA ladder (New England Biolabs, Massachusetts, USA, catalog number:B7025) was used as the reference on each gel. If a visible clean DNA band between 600-700bp was detected under the UV light, the rest of the PCR product was sent to GENEWIZ (now under the new company name as Azenta) for Sanger DNA sequencing.

DNA Sequence Analysis

The sequence results were analyzed in the DNA Subway Blue Line. The DNA trace files were analyzed by Phred software. Nucleotides were only called if the Phred scores met or exceeded the quality cutoff (Phred score of 20, or greater than 99% accuracy). The sequences were trimmed to remove the "Ns" on the 5' and 3' prime ends. Consensus sequences were built using *Merger* at DNA Subway. After all the cleaning steps, the sequences were identified with a BLAST search on the GenBank website: If the highest percent identity score was >=95%, the sample was labeled as "identified"; high quality sequences of >=95% identity were submitted directly to GenBank for invertebrate distribution information (9,11).

Evaluation of Arthropod Impact on Human Health

Species names were used to check if there was a match on the pest lists provided by US government publications and reports. We used the "list of pests of significant public health importance" posted by the United State Environmental Protection Agency (12). The list of "disease vectors and pests" (Chapter 4, Healthy Housing Reference Manual) published by the U.S Department of Health and Human Services (13). The book "Harmful Non-indigenous Species in the United States" published by U.S. Congress, Office of Technology Assessment (14). If the species was not on any of the lists to harm human health, we considered it not a foe.

ACKNOWLEDGMENTS

This project was supported by the DNA Learning Center (DNALC) at Cold Spring Harbor Laboratory. We would like to extend our gratitude to Bruce Nash (DNALC) and Maria Smit (Cold Spring Harbor Press) for their comments on our manuscript. We would also like to thank Jeffry Petracca (Long Island Aquarium) who helped to morphologically identify samples and Lauren Correri (DNALC) for proofreading. Lastly, we would like to thank *Journal of Emerging Investigators* reviewers and editors for constructive suggestions.

Received: September 16, 2021
Accepted: March 4, 2022
Published: March 14, 2022

REFERENCES

1. "Insect Encounters". *Florida Department of Agriculture and Consumer Service*. doi:www.fdacs.gov/content/download/9755/file/Entomology%20brochure.pdf
2. "Pesticides' Impact on Indoor Air Quality". *United State Environmental Protection Agency.* www.epa.gov/indoor-air-quality-iaq/pesticides-impact-indoor-air-quality
3. "Pesticides and Child Safety". *Maryland Department of*

Agriculture. doi:www.mda.maryland.gov/plants-pests/Pages/pesticide_child_safety.aspx
4. Roberts, James R ,et al. "Pesticide Exposure in Children". *Pediatrics*. 131(5):1013-4,May 2013. doi: 10.1542/peds.2012-2758.
5. Gummin, David, et al. "(2020) 2019 Annual Report of the American Association of Poison Control Centers' National Poison Data System (NPDS): 37th Annual Report, Clinical Toxicology". 58:12, 1360-1541. doi: 10.1080/15563650.2020.1834219
6. Borror and Delong's *Introduction to the Study of Insects*, 7th edition. Books/cole cengage learning. 2004.
7. "Phylum Arthropoda-arthropods". *BugGuide. Iowa State University, Department of Entomology*. doi: www.bugguide.net/node/view/3 h
8. Hebert, Paul, et al. "Biological Identifications Through DNA Barcodes". *Proceedings of the Royal Society B Biological Science*. 2003,270(1512):313–3. doi:10.1098/rspb.2002.2218.
9. Marizzi, Christine, et al. "DNA Barcoding Brooklyn (New York): A First Assessment of Biodiversity in Marine Park by Citizen Scientists". *PLOS one*, 2018
10. Hebert, Paul, et al . "Barcoding Animal Life: Cytochrome C Oxidase Subunit 1 Divergences Among Closely Related Species". *Proceedings of the Royal Society B Biological Science*. 2003;270(supplement 1):S96–S99. doi: 10.1098/rsbl.2003.0025.
11. "Use DNA Barcoding to Identify and Classify Living Things". *DNA Learning Center, Cold Spring Harbor*. Doi:www.dnabarcoding101.org/lab/bioinformatics.html
12. "List of Pests of Significant Public Health Importance". *United State Environmental Protection Agency*. November 3, 2020. doi: www.epa.gov/insect-repellents/list-pests-significant-public-health-importance
13. "Healthy Housing Reference Manual". *U.S Department of Health and Human Services, U.S. Department of Housing and Urban Development*. 2012.
14. "Harmful Non-indigenous Species in the United States". *U. S. Staff Congress of the U.S., Washington, DC. Office of Technology Assessment*. 1993.
15. Paoletti, Maurizio, et al. "Woodlice (Isopoda oniscidea): Their Potential for Assessing Sustainability and Use as Bioindicators". *Agriculture, Ecosystems and Environment*. 74 (1999) 157-165.
16. "Triangulate Cobweb Spider". *Arthropod Museum, The University of Arkansas*. doi: www.arthropod.uark.edu/triangulate-cobweb
17. Almudi, Isabel, et al. "Establishment of the Mayfly Cloeon Dipterum as a New Model System to Investigate Insect Evolution". *Evo Devo* 2019 Apr 2;10. doi: 10.1186/s13227-019-0120-y. eCollection 2019.
18. "Featured Creatures, Entomology and Nematology". *Entomology and Nematology Department*, University of Florida, February 2018. www.entnemdept.ufl.edu/creatures/BENEFICIAL/convergent_lady_beetle.html
19. Jacobs, Steven. "Carpenter Ants". *Penn State Extension*, January 2014. doi:www.extension.psu.edu/carpenter-ants
20. Leong, Misha. et al. "The Habitats Humans Provide: Factors affecting the diversity and composition of arthropods in houses". *Scientific Reports*, 2017; 7 (1). doi: 10.1038/s41598-017-15584-2