Applications of DNA barcoding & metabarcoding – Genetic species identification and possibilities for biodiversity monitoring 2.0 in the digital age

Marita Sacher, Vedran Bozicevic, Rene Tänzler & Jérôme Morinière

AIM Advanced Identification Methods GmbH

Zusammenfassung: Mittels DNA Barcoding können unbekannte Individuen, aber auch ganze Mischproben analysiert und identifiziert werden. Komplexe Biodiversitätsstudien können somit auch in Zeiten der Taxonomiekrise durchgeführt werden und die gesamte Biodiversität kann erfasst werden. Mittels integrativer Taxonomie können neue Arten bestimmt und deren genetischer Code kann in weltweit zugänglichen Datenbanken hinterlegt und katalogisiert werden. DNA Barcoding unterstützt somit die klassischen Methoden über deren Grenzen hinaus. Metabarcoding ist die genetische Analyse von Mischproben bestehend aus einer Vielzahl von Individuen. Die Organismen können in einem einzigen Analyseschritt erfasst und müssen nicht wie bei klassischen Methoden einzeln sortiert und identifiziert werden. Ganze Proben von tausenden Individuen können homogenisiert werden und mittels Hochdurchsatzmethode analysiert werden. Hiermit erhält man komplette Artenlisten und die Zusammensetzung der Arten innerhalb einer Probe in nur einem Arbeitsschritt. In den letzten Jahren konnten mittels DNA Barcoding diverse Biodiversitätsstudien unterstützt und somit grundlegende Hypothesen schnell und effizient untersucht werden.

Keywords: DNA barcoding, metabarcoding, biodiversity monitoring, genetic species identification

AIM Advanced Identification Methods GmbH, Spinnereistrasse 11, 04179 Leipzig, Germany; E-Mail: marita.sacher@aimethods-lab.com

Integrative taxonomy with the help of DNA barcoding

We are currently in a taxonomic crisis, but at the same time, the need for species identification is increasing. For years, samples from biomonitoring studies have been collected and are now stored at various institutes, museums, and collections. However, the funding to have all these mixed samples sorted and determined by taxonomists in the classical way has been, and still is, severely lacking.

DNA barcoding supports classical methods beyond their capabilities and can identify individuals of a species regardless of the life stage, the gender, or whether only small body parts are available for analysis. Although a full species determination is only possible if reference sequences of previously identified species are available (e.g. in BOLD: www.boldsystems.org), unknown specimens can be identified to order, family or genus level based on the DNA barcode framework present in the databases. DNA barcoding is ready to be used for species determination beyond the taxonomic and regional limitations and can be used to capture biodiversity in its entirety – not only focus groups. In the last 10 years, national DNA barcoding initiatives have successfully DNA barcoded more than 50% of the German fauna – approximately 23,000 species. However, gaps within the species catalogue of so called "dark taxa" in hyperdiverse groups such as parasitoid Hymenoptera or Diptera still need to be filled. Around the world, DNA barcoding initiatives are collecting and entering data into worldwide, freely accessible databases (such as BOLD). Therefore, molecular taxonomic expertise is now digitally available and can be coupled with genetic data.

Methodology

The name "barcoding" comes from product barcoding, which is known, for example, from the retail industry. Instead of a barcode to uniquely identify a product, a particular gene (cytochrome oxidase 1, or CO1) is

"scanned" here. For most species, this results in a very specific species barcode, consisting of particular sequence of base pairs A, C, T, and G. The species of land plants are identified via two different gene regions of the chloroplast (matK and rbcL). Mushrooms are identified via the ITS region. With the help of these short DNA barcodes, one can thus clearly identify a referenced known species, or in some cases determine a new, previously unknown species.

Throughout the world, scientists are currently working to verify and complete digital databases with DNA barcodes. Following an in-depth review by scientists within the barcoding community, these data are made available in a publicly available form. This makes it possible to analyze and identify unknown individuals, as well as entire mixed samples. Similarities of unknown sequences to existing codes are assessed and evaluated in the genetic analysis and the subsequent bioinformatics pipeline. Codes with 97-100% match to an already determined code can be uniquely assigned to one species. For lower-percentage matches, the specific codes have likely not yet been assigned to any species, but can be assigned to higher taxonomic levels – a genus, or a family, for example.

DNA metabarcoding to support biodiversity monitoring 2.0

The databases are ready for the application of metabarcoding, which allows the analysis of bulk samples derived from, among others, Malaise traps, soil samples, fecal samples, stomach contents, water samples, and processed foods. Hundreds of samples can be sequenced and analyzed in one single step. Entire biodiversities of the samples can be recorded using this method (Morinière & al. 2016). DNA barcoding allows biomonitoring 2.0 to be cost-effective and fast. With a fraction of the budget, and a significantly lower expenditure of time and personnel, a high percentage of the organisms in a sample can be identified.

Use cases

Integrative Taxonomy as a tool for the determination of species of the smaller arachnid orders

Integrative taxonomy combines morphological species determination by means of classical taxonomy with modern methodological developments like COI DNA barcoding. This helps to reevaluate species boundaries or search for cryptic species previously undetected by morphological analysis. In a recent study by Lehmann & Friedrich (2018), DNA barcoding could support in the determination of species of the smaller arachnid orders in Peru using integrative taxonomy. Two orders were of great interest to the scientists: the order Schizomida, commonly known as shorttailed whipscorpions, and the order of Amblypygi, commonly known as whip spiders, or tailless whip scorpions.

At first sight in the field, the specialists were not sure whether two morphologically different groups (morphs) belonged to two different species, or whether the smaller morphs were just juvenile forms of the larger ones. Individuals from two orders were collected and analyzed morphologically. Additionally, a leg, or a part of a leg of each specimen was sent to AIM for DNA barcoding service.

The molecular results from DNA barcoding affirmed the conclusion of the morphological analyses, and adjustments concerning some morphological features could be made for future identifications of the order Schizomida (shorttailed whipscorpions).

A combination of classical taxonomy and modern genetic methods should be used in any species description. Only in this way can the worldwide-accessible and digital catalog of species be completed.

DNA metabarcoding supports Europe-wide biodiversity study in beech forests

Much of Europe is covered by beech forests (dominated by beech trees - *Fagus sylvatica*). These forests provide a livelihood for many species, including thousands of arthropods. Mushrooms are an important component of functioning forests and are the most important organisms for wood decomposition. The fruit bodies and mycelia of fungi are an important source of food for many insects and other arthropods. Additionally, the fruit bodies provide a habitat for numerous organisms. The tinder fungus (*Formes formentarius*) is one of the main decomposers of wood in beech forests in Europe. Both its fruit body, as well as the resulting deadwood, provides shelter and food for many arthropods. Through deforestation and management measures, however, the fungus is threatened in many areas or even extinct locally. Therefore, the tinder fungus and the associated diversity of arthropods have been investigated in a recent study by Friess & al. (2019).

In this study, the total biodiversity within the fungal bodies was investigated. Beetles were determined using classical methods by taxonomists. The remaining animals were sent to AIM as mixed samples and analyzed there with the DNA metabarcoding service in just one analysis step at a time, and entire species lists were created for each sample.

In total, 216 different types of arthropods in the tinder fungi were identified in this study. Of these, 71 species of beetles were determined in the classical way by the taxonomists and 145 species of arthropods by the DNA metabarcoding service of AIM. The study also analyzed the level of the food chain in which the animals are to be arranged, since in a healthy system mostly representatives from all levels should be present. The holistic and integrative analysis of samples using classical methods and DNA metabarcoding identified 131 consumers, 68 predators, and 17 parasitic species. A few of the Operational Taxonomic Units (OTUs) generated by DNA metabarcoding could not be linked to any of the species already described. Most of these OTUs belong to the gall midges family (Cecidomyiidae). However, as there are too few specialists who are familiar with these species, these species have not yet been recorded digitally in the "Catalog of Life". If these species are cataloged in the future, a clear assignment to a species would be possible. The data from DNA metabarcoding projects will become more and more valuable over the years and growing data sets and digital libraries.

This Europe-wide study has identified tinder fungus as an important host for a large number of beech forest animals. For the European beech forests, therefore, the retention of trees with tinder fungi and the promotion of reintroduction of the fungus where it has declined is a promising conservation strategy. These recommendations can be used to actively promote and protect native species and farm animals. Thus, a targeted contribution to the preservation of domestic biodiversity can be made.

Ecological vs. conventional farming

In a recent study by Hausmann & al. (2020), supported by HiPP GmbH, a German baby foods producer, DNA metabarcoding was used to assess the insect communities on conventionally versus ecologically used fields. Malaise traps were used to capture most of the flying insects over the period of one season. The biomass, as well as the species richness was significantly higher on the ecological plots compared to the conventional plots. Over the season, we could also show the beneficent effect of ecological farming on different groups of insects and species. The surveys of the fields are continued in the coming seasons. Within a few weeks certain hypotheses can now be tested and answered using DNA metabarcoding.

DNA barcoding makes routine species identification automatable

- Assessment of complete fauna and flora possible (e.g., animals, plants, bacterial fauna, soil, dead wood, fungal communities)
- Efficient, rapid and cheap analysis of bulk samples possible (e.g. lists of 3,000 – 5,000 species in a few days/weeks)
- Early warning system for invasive insects and pests (borders, airports, train stations, greenhouses etc.)
- Long-term-monitoring, ecological analyses (e.g. climate change, insect decline, disturbance gradients) with huge sample sizes
- Detecting cryptic species

AIM – Advanced Identification Methods GmbH

AIM Advanced Identification Methods GmbH is a scientific spinoff originating in the scientific environment of the DNA barcoding projects in Bavaria and Germany. With 10 years of expertise in DNA barcoding and metabarcoding and analysis of several thousand bulk samples in many different projects, we were able to help resolve many scientific questions from various scopes of application (e.g., Biodiversity Exploratories, The Jena Experiment). With sophisticated bioinformatics and extensive reporting we can now assess animal, plant, fungal and microbial faunas quickly and efficiently in various environmental contexts.

References

Friess, N., Müller, J.C., Aramendi, P., Bässler, C., Brändle, M., Bouget, C., Brin, A., Bussler, H., Georgiev, K.B., Gil, R., Gossner, M.M., Heilmann-Clausen, J., Isacsson, G., Krištín, A., Lachat, T., Larrieu, L., Magnanou, E., Maringer, A., Mergner, U., Mikoláš, M., Opgenoorth, L., Schmidl, J., Svoboda, M., Thorn, S., Vandekerkhove, K., Vrezec, A., Wagner, T., Winter, M.-B., Zapponi, L., Brandl, R. & Seibold, S. (2019): Arthropod communities in fungal fruitbodies are weakly structured by climate and biogeography across European beech forests. – Diversity and Distributions **2019**(25): 783-796. https://doi.org/10.1111/ddi.12882.

Hausmann, A., Segerer, A.H., Greifenstein, T., Knubben, J., Morinière, J., Bozicevic, V., Doczkal, D., Günter, A., Ulrich, W. & Habel, J. C. (2020): Toward a standardized quantitative and qualitative insect monitoring scheme. – Ecology and Evolution **10**(9): 4009-4020. https://doi.org/10.1002/ece3.6166.

Lehmann, T. & Friedrich, S. (2018): DNA barcoding the smaller arachnid orders from ACP Panguana, Amazonian Peru. – Spixiana **41**(2): 169-172.

Morinière, J., Cancian de Araujo, B., Lam, A.W., Hausmann, A., Balke, M., Schmidt, S., Hendrich, L., Doczkal, D., Fartmann, B., Arvidsson, S. & Haszprunar, G. (2016): Species Identification in Malaise Trap Samples by DNA Barcoding Based on NGS Technologies and a Scoring Matrix. – PLoS ONE **11**(5): e0155497. https://doi.org/10.1371/journal.pone.0155497