



# Life barcoded by DNA barcodes

Mali Guo<sup>1,2</sup> · Chaohai Yuan<sup>1,2</sup> · Leyan Tao<sup>1</sup> · Yafei Cai<sup>1,2</sup> · Wei Zhang<sup>1,2</sup> 

Received: 15 July 2021 / Accepted: 5 August 2022  
© The Author(s) 2022

## Abstract

The modern concept of DNA-based barcoding for cataloguing biodiversity was proposed in 2003 by first adopting an approximately 600 bp fragment of the mitochondrial *COI* gene to compare via nucleotide alignments with known sequences from specimens previously identified by taxonomists. Other standardized regions meeting barcoding criteria then are also evolving as DNA barcodes for fast, reliable and inexpensive assessment of species composition across all forms of life, including animals, plants, fungi, bacteria and other microorganisms. Consequently, global DNA barcoding campaigns have resulted in the formation of many online workbenches and databases, such as BOLD system, as barcode references, and facilitated the development of mini-barcodes and metabarcoding strategies as important extensions of barcode techniques. Here we intend to give an overview of the characteristics and features of these barcode markers and major reference libraries existing for barcoding the planet's life, as well as to address the limitations and opportunities of DNA barcodes to an increasingly broader community of science and society.

**Keywords** DNA barcode · *COI* · Reference libraries · Mini-barcode · DNA metabarcoding

## Introduction

In comparison to the universal bar code consisting of a series of vertical bars that are printed on commercial products, a DNA barcode, in a broad sense, refers to any DNA sequence adopted to identify species at any taxonomic level. To be more specific, a DNA barcode is one or more short gene sequences (generally 200–900 base pairs) taken from a standardized portion of the genome to aid species identification and discovery by employing sequence divergence based on nucleotide alignment (Emerson et al. 2011; Hebert et al. 2003a, 2004). Thus, the fundamental function of this genetic tool seeks to compare barcode sequences to reference databases to efficiently and effectively assign any biological sample to its species regardless of the visual classification of the sample (Fig. 1).

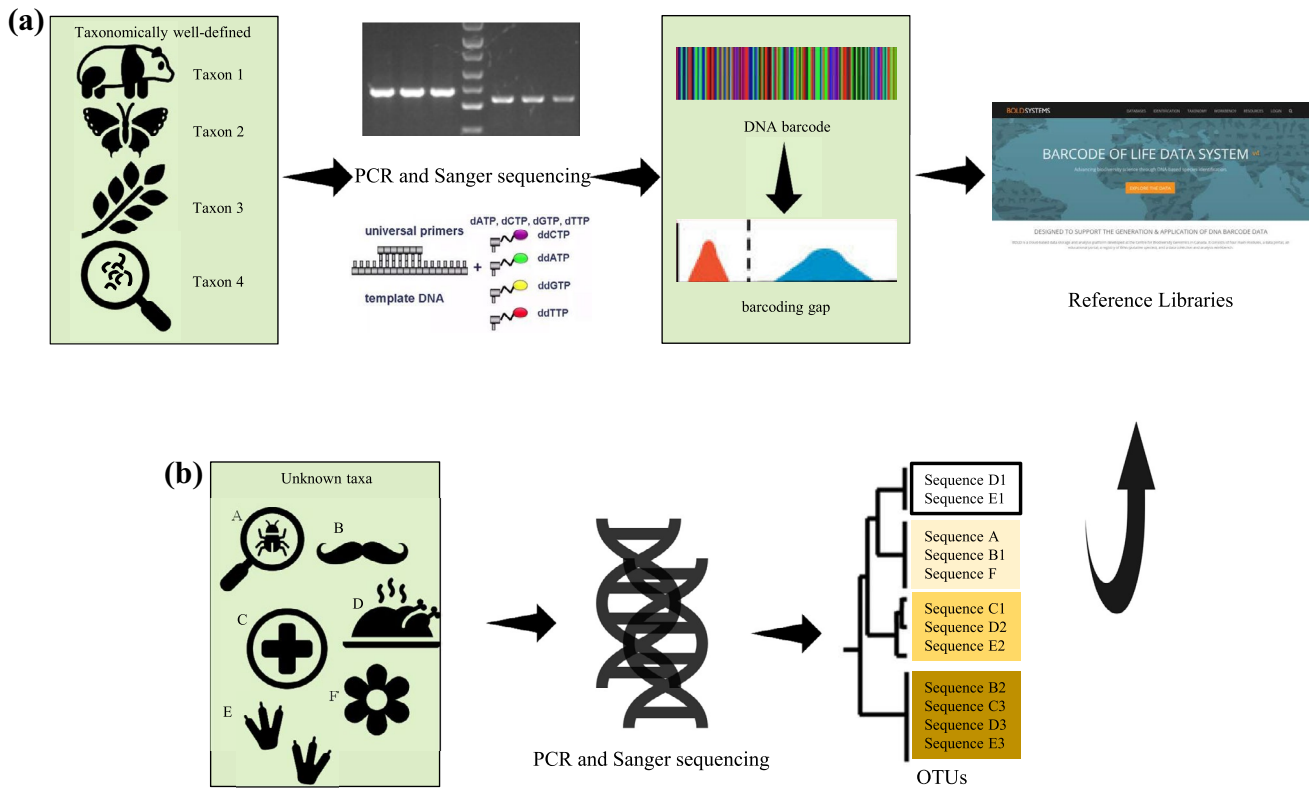
In fact, the history of genetic sequences applied in taxonomy research can be traced back to 1969 when Bicknell and Douglas found that the arrangement of species in yeast dependent on ribosomal RNA homologies in most cases agreed with the established taxonomic groupings via traditional measures (Bicknell and Douglas 1970). Before that, 250 years have been spent to catalogue about 1.2 million species by traditional taxonomic approaches through manual characterizations incorporating morphological features, which apparently is challenging and unrealistic when it is forecasted that around 7.52 million terrestrial species and 2.01 million species in the ocean still await description (Leray and Knowlton 2016; Scheffers et al. 2012). The advent of techniques for gene isolation, cloning and Sanger sequencing in the second half of last century allows the term “DNA barcode” to be first used in 1993 when length information of tandemly repeated DNA sequences from hypervariable alleles was barcoded to discriminate isolates of *Plasmodium falciparum* (Arnot et al. 1993). Nevertheless, the novel concept of DNA barcode relying upon nucleotide divergence was not formally proposed and established for species diagnosis until 2003 by Hebert PD et al. (Hebert et al. 2003b). Since then, international initiatives have been launched across hundreds of countries to evaluate the world's bio-diversities using this new taxonomic tool, and

---

✉ Wei Zhang  
weizhang@njau.edu.cn

<sup>1</sup> College of Animal Science and Technology, Nanjing Agricultural University, Nanjing 210095, China

<sup>2</sup> National Experimental Teaching Demonstration Center of Animal Science, Nanjing Agricultural University, Nanjing 210095, China



**Fig. 1** Basic workflow for getting barcode markers using Sanger sequencing. **a** Workflow for generating reference databases. **b** Workflow for taxonomic assignment of unknown samples by comparing barcode sequences with reference databases

more than 321 K species, covering animals, plants, fungi and others, have been barcoded so far (Jeanson et al. 2011; Ratnasingham and Hebert 2007, 2013).

In light of the rapid progress and vast application of DNA barcodes, the purpose of this paper is to review these genetic markers in a variety of living organisms and provide a snapshot glimpse of mini-barcodes and DNA metabarcoding, which are essential extensions of the regular barcodes. All these barcodes, however, are heavily relying on the presence of high-quality barcode sequence reference databases that are based on good taxonomy and barcode coverage (Ratnasingham and Hebert 2007). At the end, we will also summarize some of the most exciting prospects for using this modern taxonomic tool.

## How to get a barcode

The initial motivation to have DNA barcode is to group an individual with its conspecifics using simple molecular tools instead of morphology-based procedures, which are tedious tasks requiring experienced taxonomists (Giangrande 2003). Although it has been repeatedly called into question, the core idea behind current barcodes rests on the fact that

certain pieces of DNA, when aligned, can be found to vary only to a limited degree within species while this variation is much less than between species, which is referred to as the barcoding gap (Fig. 1) (Hill 2016; Liu et al. 2011; Meyer and Paulay 2005). Therefore, whether samples of target species can be differentiated largely depends on the choice of the short DNA segment. Gene regions that evolve slowly often do not differ among closely related organisms, whereas DNA sequences that evolve rapidly, on the other hand, may overwrite the traces of ancient affinities, but introduce more sequence diversity and increase the chance to distinguish between species (Cho et al. 2004; Steinke et al. 2016). In addition, the DNA section chosen must be standardized and accessible in various taxonomic groups with conservative primer binding sites so that the barcode marker technically can be robustly amplified and sequenced from a small amount of specimen through polymerase chain reaction (PCR).

Although enormous efforts have been made to find a single segment of DNA meeting all criteria outlined above, such a region has not been identified, and researchers start to realize that a single universal DNA barcode for all forms of life is unlikely to exist. This is largely because barcoding regions are not evolving neutrally since the time of

speciation, and more often are influenced by weak positive or negative selection, making them suitable in some species but not others. Under such circumstances, multi-locus barcodes aiming for different living taxa have been developed and examined with respect to both their ease of amplification and their capacity to resolve species as a part of the barcode validation process (Table 1).

## Barcodes in animals

Notwithstanding differences in evolutionary history between nuclear and mitochondrial DNA mean that a mitochondrial barcode is unlikely to be representative of nuclear divergence, in animals, regions from mitochondrial DNA (mtDNA) are preferred over nuclear genome for barcoding (Hill 2016, 2020). This is because mtDNA in most eukaryotes is known to be inherited uniparentally from the maternal parent, possessing circular DNA packaged into nucleoids without the protection of histone proteins. Despite rare recombination, mitochondrial genome, compared with nuclear DNAs, lacks sufficient DNA repair mechanisms, leading to a tenfold higher rate of nucleotide substitution in the presence of reactive oxygen species generated during the respiratory chain (Adamowicz et al. 2017). The rapid pace of sequence change in mtDNA consequently allows accumulation of differences between closely related species that have only been separated for brief periods of time.

A standard fragment of ~648 base pairs (bp) at the 5' end of the mitochondrial gene coding the cytochrome c oxidase subunit 1 (*COI*), a component of an enzyme complex essential for oxidative phosphorylation, is the first and so far the most broadly adopted molecular marker for barcoding animals (Hebert et al. 2003b; Kress et al. 2005; Steinke et al. 2016). By making use of universal primers for PCR amplification, *COI* barcode has been claimed to achieve high rates of success in identification of species in test assemblages of different animal groups, mainly insects, birds and fishes (Pratheepa et al. 2014; Prum et al. 2015; Ward 2012). These PCR primers, initially described for diverse metazoan invertebrates, are fundamental to the barcode field and prevalent even today, generating informative sequences for phylogenetic analyses at the species and higher taxonomic levels (Folmer et al. 1994). Yet some studies challenged the degree of universality for *COI* and its primers for a number of reasons. For instance, the high variability of nucleotide sequences at the *COI* priming sites hinders its application to a broader spectrum of animal species (Hawlitschek et al. 2016; Shearer et al. 2002; Zangl et al. 2020). To address this issue, cocktails of degenerate primer sets were proposed for barcoding species like reptiles and amphibians (Che et al. 2012; Lyra et al. 2017; Vences et al. 2005). Moreover, *COI* barcoding region were also found to provide insufficient species resolution when it comes to organisms such as sea

snails and corals because of limited nucleotide diversity (McFadden et al. 2011; Young et al. 2017). Such a shallow *COI* variation was also uncovered within many species of parasitoid wasps as the smallest interspecific divergence of only 1 bp was recorded between wasps that are known to parasitize different families of caterpillars (Smith et al. 2008). This exemplifies the integration of DNA barcoding with morphological, behavioral and ecological descriptions to improve the accuracy of species identification. In contrast, large intraspecific distance ranging from 0% to as much as 17.3% was noticed for *COI* genes in pseudoscorpions, which distorts the pattern of intra- and interspecific variation and spoils the existence of a barcode gap. This observation may result from undocumented species diversity, but also from anomalies in the *COI* evolution of these arachnids, indicating variable molecular change between species within different taxa (Muster et al. 2021).

Alternative barcode candidates for animals include segments from mitochondrial *cytochrome b* (*cytb*), *12S ribosomal RNA* (*rRNA*) and *16S ribosomal RNA*, and so on (Table 1) (Fernandes et al. 2020; Milan et al. 2020; Sun et al. 2019; Wong et al. 2004; Xia et al. 2012). Choices of these genetic markers are substantially due to practical reasons that a huge number of DNA sequences spanning these regions already exist in public databases before the barcoding methods became popular. Nevertheless, as a more comprehensive *COI* reference database is becoming feasible, it is argued that these alternatives may no longer perform equally well, even cases that *16S* is superior to *COI* in barcoding Arthropoda and amphibians, for example, are still reported at current stage (Sikes et al. 2017; Vences et al. 2012; Xia et al. 2012; Zangl et al. 2020).

## Barcodes in plants

As revealed in the previous section, animal mtDNAs are characterized by their rapid evolution in primary sequence, but there is a wide consensus that they are essentially invariant in gene order, especially among all vertebrates (Khan et al. 2007). However, this is not case for plants as their mtDNAs are postulated to have undergone extensive internal rearrangements, resulting in a high rate of length mutations rather than nucleotide substitution (Chevigny et al. 2020). Therefore, it is suggested that the point mutation rate in plant mtDNA is around 100 times slower than in animal mtDNA (Chevigny et al. 2020). Searching for suitable mitochondrial barcodes to delineate plant species thus has proved to be tricky and botanists thereby have focused on DNA sequences outside the mitochondrial genome.

So far, the nuclear-encoded ribosomal internal transcribed spacer (ITS) region and the chloroplast intergenic spacer *trnH-psbA* have emerged as candidates for barcoding plants, followed by others including coding sequences from plastid

**Table 1** Molecular markers routinely used for DNA barcoding studies

Organism	Region	Marker	Gene description	References	
Animals	Mitochondrion	<i>12S</i>	12S ribosomal RNA	(Kocher et al. 1989; Olmstead 1996)	
		<i>16S</i>	16S ribosomal RNA	(Palumbi 1991)	
		<i>atp6</i>	ATP synthase F0 subunit 6	(Haag et al. 2009; Trigo et al. 2008)	
		<i>COI</i>	Cytochrome c oxidase subunit I	(Hebert et al. 2003a, 2003b)	
		<i>cytb</i>	Cytochrome b	(Hardman 2004; Kocher et al. 1989; Maxfield et al. 2012; Tchaicka et al. 2007)	
		<i>D-loop</i>	Mitochondrial displacement loop region	(Hoelzel et al. 1991)	
		<i>ND1</i>	NADH dehydrogenase subunit 1	(Thacker 2003)	
		<i>ND2</i>	NADH dehydrogenase subunit 2	(Thacker 2003)	
		Nucleus	<i>28S</i>	28S ribosomal RNA	(Saux et al. 2004)
			<i>ITS</i>	Internal transcribed spacer	(Smith et al. 2008)
			<i>Rag1</i>	Recombination activating 1	(López et al. 2004)
			<i>Rag2</i>	Recombination activating 2	(Hardman 2004)
			<i>WG</i>	Wingless	(Fagan-Jeffries et al. 2018)
Plants	Nucleus	<i>ITS</i>	Internal transcribed spacer	(Chen et al. 2005; Michelangeli et al. 2004)	
		<i>ITS2</i>	The 2nd internal transcribed spacer	(Moorhouse-Gann et al. 2018)	
	Plastid	<i>atpF-atpH</i>	Non-coding atpF-atpH intergenic spacer region	(Marcelo et al. 2010; Reginato and Michelangeli 2016)	
		<i>matK</i>	Maturase K	(Fazekas et al. 2008; Parveen et al. 2012)	
		<i>psbK-psbI</i>	Non-coding psbK-psbI intergenic spacer region	(Marcelo et al. 2010)	
		<i>rbcL</i>	Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit	(Fazekas et al. 2008; Parveen et al. 2012)	
		<i>rpoB</i>	RNA polymerase beta subunit	(Fazekas et al. 2008; Parveen et al. 2012)	
		<i>rpoC1</i>	RNA polymerase beta' subunit	(Parveen et al. 2012)	
		<i>rps16</i>	Ribosomal protein S16	(Oxelman et al. 1997)	
		<i>trnC-rpoB</i>	Non-coding trnC-rpoB intergenic spacer region	(Ohsako and Ohnishi 2000)	
		<i>trnH-psbA</i>	Non-coding trnH-psbA intergenic spacer region	(Tate and Simpson 2003)	
		<i>trnL (UAA)</i>	tRNA trnL intron	(Chen et al. 2005; Taberlet et al. 2007)	
		<i>trnL (UAA)-trnF (GAA)</i>	tRNA trnL-trnF intergenic spacer region	(Sang et al. 1997)	
		<i>trnK (UUU)</i>	tRNA trnK intron	(Ohsako and Ohnishi 2000)	
		<i>ycf1</i>	Translocon at the inner envelope membrane of chloroplasts 214	(Dong et al. 2014)	
		<i>ycf5</i>	Cytochrome c biogenesis protein CcsA	(Kress and Erickson 2007)	
		Fungi	Mitochondrion	<i>atp6</i>	ATP synthase F0 subunit 6
<i>COI</i>	Cytochrome c oxidase subunit I			(Pino-Bodas et al. 2013; Vialle et al. 2009)	
<i>CO3</i>	Cytochrome c oxidase III			(Vialle et al. 2009)	
<i>nad6</i>	NADH dehydrogenase subunit 6			(Vialle et al. 2009)	
Nucleus	<i>LSU</i>			large ribosomal subunit gene D1/D2 domains	(Eberhardt 2012; Robert et al. 2011)
	<i>ACT</i>			Actin	(Carbone and Kohn 1999)
	<i>TUB</i>			β-tubulin	(Glass and Donaldson 1995)
	<i>CAL</i>		Calmodulin	(Carbone and Kohn 1999)	
	<i>EF1-α</i>		Translation elongation factor 1-alpha	(Pino-Bodas et al. 2013)	
	<i>H3</i>		Histone H3	(Crous et al. 2006)	
	<i>ITS</i>		internal transcribed spacer	(Vialle et al. 2009)	
<i>rpb2</i>	DNA-directed RNA polymerase II subunit		(Pino-Bodas et al. 2013)		
Bacteria	Nucleoid		<i>16S</i>	16S ribosomal RNA	(Lane 1991; Sundquist et al. 2007)
			<i>chaperonin-60</i>	60 kDa chaperonin	(Brousseau et al. 2001)

**Table 1** (continued)

Organism	Region	Marker	Gene description	References
		<i>rpoB</i>	RNA polymerase beta subunit	(Adékambi et al. 2009)
		<i>ITS</i>	Internal transcribed spacer	(Benga et al. 2020; Soltan Dallal et al. 2019)
Archaea	Nucleoid	<i>16S</i>	16S ribosomal RNA	(Bates et al. 2011)

genes *rbcL* and *matK*, two loci now the most commonly used for plants (Kress and Erickson 2007; Loera-Sánchez et al. 2020; Yao et al. 2010). Unfortunately, no single marker from them has been found to fully satisfy all of the desired characteristics required for DNA barcodes. For instance, *rbcL* fragment is easy to amplify, sequence and align, but only yields modest discriminatory power whereas the *matK* barcode, perhaps the closest plant analogue to *COI* in animal, is difficult to amplify due to the lack of competent primers of universality (Braukmann et al. 2017; de Vere et al. 2015; Fang et al. 2019; Li et al. 2015). Even so, it is worth mentioning that the ITS2 primers designed for the second internal transcribed spacer of nuclear ribosomal DNA, which have been used successfully for a number of applications with short amplicons of 187–387 bp and addressed many of the issues, though not all, levied against the other primers (Table 1) (Moorhouse-Gann et al. 2018).

Since a standard plant barcode has been complicated by the trade-off that arises between the high variability of sequences and high conservation of primers, it is then recommended to simultaneously utilize more than one marker as a compromise that best matches the barcoding criteria (Lahaye et al. 2008). As a consequence, combinations of multiple barcode markers were shown to improve the ability to classify plants maximally by 60% when compared to a single barcode, which has persuaded researchers out of botany to take the same measures when barcoding other organisms (Group 2009; Li et al. 2021; Nitta et al. 2020; Zhang et al. 2013).

### Barcodes in microorganisms

The microorganisms discussed in this section will mainly refer to fungi, bacteria and viruses, which are all around us, having an enormous biological and economic impact, but often invisible to our naked eyes. Species discrimination is often frustrated as microorganisms only occasionally exhibit the morphological characters needed for identification in natural ecosystems. Until now, our knowledge of microbial biodiversity has been severely restricted by relying on microorganisms that can be cultured while vast majority (>99%) cannot (Mendes et al. 2013). Luckily, PCR-based DNA barcoding techniques offer such a great opportunity to characterize microbial communities without prior cultivation.

In line with animal *COI*, the fungal counterpart is also officially recognized as an eligible barcode marker, yet it is usually excluded from consideration by mycologists due to the presence of mobile introns in the priming and sequencing regions (Seifert et al. 2007; Xu 2016; Yahr et al. 2016). Other reasons for exclusion include low nucleotide variation and a complete lack of mitochondria in some fungal lineages (Wickes and Wiederhold 2018). Instead, the ITS region and the D1/D2 region of the large subunit (*LSU*) rDNA, both belonging to the nuclear ribosomal RNA genes, are the most widely used in all fungi from diverse environments (Dulla et al. 2016; Schoch et al. 2012; Scorzetti et al. 2002). These two loci can be easily amplified using relatively universal primers, and have the largest amount of reference sequence data for fungi (Blackwell 2011; Suhr and Hallen-Adams 2015). In general, ITS is better at distinguishing closely related species than *LSU*, but ITS is more difficult to align because of length differences (Takashima et al. 2019). Other fragments, like nuclear  $\beta$ -tubulin, translation elongation factor 1- $\alpha$  and calmodulin, sometimes are also applied together for selected fungal genera (Table 1), consistent with the barcoding mixtures in plants (Panelli et al. 2013; Pino-Bodas et al. 2013; Quaedvlieg et al. 2012; Robba et al. 2006).

On the other hand, *16S rRNA* gene was first advised by microbiologists as a phylogenetic tool to describe the evolutionary relationships among bacteria, archaea and eukaryotes in 1977, since when over 41 million *16S* sequences, much more than the 3 million *COI* sequences, have been deposited in GenBank (Woese and Fox 1977). However, the idea to use *16S* as the primary barcode nowadays only catches on in bacteria for a number of causes. First of all, this gene is frequently accessible in almost all bacterial species, either harmless or pathogenic. Secondly length of the gene is approximately 1500 bp, which is informative enough for analyses (Clarridge 2004). Finally, function of this gene has not changed, containing conserved sequences for universal PCR primers. Conversely, utility of *16S* is constrained in a broader taxonomic investigation by the prevalence of insertions and deletions that deeply complicate sequence alignments (Church et al. 2020; Yarza et al. 2014). Other options for barcoding bacteria include *chaperonin-60* and *RNA polymerase  $\beta$  subunit* (*rpoB*) gene, which can act as important supplementary markers to *16S* in appropriate cases (Pavan et al. 2012; Vancuren et al. 2020).

To date, detection and interpretation of virus, for example the SARS-CoV-2 responsible for the ongoing COVID-19 pandemic, is heavily dependent upon quantitative RT-PCR designed according to genomic sequences, which normally are assembled by overlapping 400 bp fragments from a serial amplicon generated in multiplexed PCR based on the latest ARTIC SARS-CoV-2 sequencing protocol (Asselah et al. 2021; Giri et al. 2021; Tyson et al. 2020; Weissleder et al. 2020). Nonetheless, viral genomes are still the most reliable source to estimate the rate of viral evolution and monitor circulating lineages. The difficulty with barcoding SARS-CoV-2 comes with the designation of target sites that are diagnostic of particular variants and, ideally, able to detect novel variants. Any attempt to capture the molecular identity of virus with standard barcoding unfortunately has turned out to be fruitless owing to the continuous introduction of new virus variants with random mutation and recombination (Bano et al. 2022). Thus, development of DNA barcode to understand viral diversity is still an open question in the field, unless multiple markers deployed to cover the whole viral genome are considered as an extension of the combination barcoding concept in plants.

## Sequence reference libraries

No matter it is conventional taxonomic approach or DNA barcoding method, the accuracy of species assignment and consequent taxonomic coverage are certainly relying on the availability of a well-curated and comprehensive reference system for judgement. A DNA barcode database is particularly vital for the latter because it fulfils the dual role of a library for data depository and a tool for monitoring the results and conclusions (Hawlitschek et al. 2016; Ratnasingham and Hebert 2007). Hence demand for high quality reference libraries has increased dramatically since the launching and extended utilization of barcoding technologies.

The Barcode of Life Data System (BOLD) is a bioinformatics platform serving for the acquisition, storage, analysis and publication of DNA barcode records (Liu et al. 2013). Core features of BOLD include open access to the entire biodiversity community, as well as the persistent linkage between a qualified barcode sequence and its source specimen with authoritative taxonomic identification. As such, BOLD workbench also implements a special analytical tool called Barcode Index Number (BIN) system, a molecular registry for codifying operational taxonomic units (OTUs) (Hausmann et al. 2013; Ratnasingham and Hebert 2013). The BIN system in principle uses well compiled algorithms and clusters similar sequences encountered in different studies into groups corresponding to presumptive species, but not necessarily actual species. Each BIN has an individual webpage displaying a unique alphanumeric identifier,

nearest neighbor, all member sequences, haplotype network, specimen images, sampling map and attribution details. At this moment, the Public Data Portal of BOLD is hosting more than 715 K BINs and 9 million barcodes, which must derive from 12 million verified specimen records within the data library, highlighting a key role of morphology-based diagnostics in barcoding (Ratnasingham and Hebert 2007). All information is free for download as reference so that large amounts of data would be screened concurrently, allowing an integrated comparison of specimens identified by both molecular and morphological characters.

As soon as results are ready for public release, a copy of all sequences and crucial specimen data from BOLD would migrate to major genomics repositories worldwide, such as Genbank database at the National Center for Biotechnology Information (NCBI). Besides the *16S* and *COI* sequences mentioned earlier, a fair portion of records in GenBank actually are generated by non-barcoding studies, lack connection to a voucher specimen, and thereby may not abide to the formal barcode data standards (Sayers et al. 2020). Compared to BOLD, however, much more nucleotide sequences, including erroneous sequences uploaded by people with poor taxonomic knowledge, are currently present in GenBank, constituting a useful resource that should be closely monitored but never overlooked. Furthermore, the Basic Local Alignment Search Tool (BLAST) is attached to NCBI so that any query sequence practically can be aligned against all Genbank libraries in one go through a user-friendly web interface (Altschul et al. 1997). In contrast, selected data have to be downloaded from BOLD before blast search using local softwares or online programs. Since similarity-based alignment is a central step for classifying DNA sequences, this is why Genbank is still the best-known one-stop solution for a quick species diagnosis.

## Development of DNA barcoding

As a matter of fact, DNA barcoding, similar to any other analytical method in science, brings some controversies and concerns too, especially in the field of taxonomy, as it does not always work as effectively as first claimed (Goldstein and DeSalle 2011; Knapp et al. 2004; Miller 2007). However, in recent years, remarkable progress towards optimizing this technology has been made to improve the efficiency and lower the cost.

## Challenges to DNA barcoding

The most serious challenges in practice probably come from the initial DNA preparation and extraction, a step which is very difficult to standardize because of the complexity and diversity of the biological samples encountered, each

representing different problems. It is lucky that many suspect samples, such as microorganisms that do not require prior cultivation (Sect. 2.3), can be directly boiled in reaction buffer as DNA template for PCR, while more often specimens might have been subjected to varied treatments, for instance, pH modification, high pressure, grinding or drying, which would damage DNA integrity and cause DNA degradation (Fode-Vaughan et al. 2001). Although multiple extraction methods, either in-house developed protocols or commercially available kits, are open for DNA purification on a case-by-case basis, it is still impossible to find a universal method that could be applied in all contexts and meanwhile guarantee the quality of DNA so that impurities potentially interfering downstream steps are eliminated. It should also be noticed that most DNA preparation courses conducted now are aiming to isolate genomic DNA, which could be problematic if the barcodes are targeting regions out of nucleus.

Additionally, the heteroplasmic conditions in mtDNA and the presence of nuclear pseudogenes of mitochondrial origin (numts) raise concerns as well, particularly when barcoding mitochondrial markers (D'Errico et al. 2004; Stefano et al. 2017). It is known that each eukaryotic cell contains approximately 500 to 6000 copies of mtDNA, which are tissue-specific (Friedman and Nunnari 2014). This leads to a phenomenon called heteroplasmy, where both wild-type and mutant mtDNA molecules co-exist within the same cell. The occurrence of heteroplasmic variants absolutely brings in ambiguous sequence reads, eventually influencing the accuracy of taxonomic description (Sobenin et al. 2014). To make matters worse, numts can be easily co-amplified with these mtDNA variants by using conserved PCR primers if an extraction method preferring nuclear DNA is carried out before (Guo et al. 2021). Due to the differences in genetic code between mitochondrial and nuclear genomes, numts are detected as non-functional copies of mtDNA with various sizes integrated into the nuclear chromosome naturally through unknown mechanisms. Once inserted into nucleus, numts decelerate their evolutionary rate and become molecular fossils of mtDNA, which to some extent could be indispensable for recovering ancient relationships (Mishmar et al. 2004). Nevertheless, as more eukaryotic genomes are sequenced and scanned, more numts are being discovered, which may cause misidentifications of species as numts, compared to mitochondrial barcodes, are undergoing a completely different inheritance pattern.

### Advancements of DNA barcoding

To overcome the degradation of samples with poor DNA preservation, shorter barcode regions, so-called mini-barcodes, have been developed in place of full-length barcodes over the past ten years. As an extension of DNA

barcoding, mini-barcodes can be amplified with higher efficiency than regular barcodes owing to their reduced size (typically  $\leq 200\text{--}300$  bp), but until now they are merely considered as short versions of the full barcode markers with no real standard or reference database for mini-barcodes adopted. In addition to the deficiencies associated with normal barcoding, the rate of taxonomic discrimination is remarkably curtailed as critical information may be missed in mini-barcodes due to the length constraint (Hajibabaei and McKenna 2012; Shokralla et al. 2015). As a result, mini-barcodes cannot achieve universal application for most species unless identification at the genus or family level is warranted (Gao et al. 2019).

However, when complex samples containing DNA of different origins have to be assessed, Sanger sequencing-based barcoding protocols, either mini-barcodes or normal barcodes, will be costly and laborious, and surely produce chimeric reads with little relevance to the taxa within the sample. Then, the advent of high-throughput sequencing (HTS) technologies facilitates the emergence of DNA metabarcoding and revolutionizes our ability to barcode life. DNA metabarcoding mainly refers to the use of barcode-based (or amplicon-based) HTS for genotyping multiple species in mixtures that may take the form of propagules, or an individual organism engaging parasites, mutualists, diet items, and symbionts (Kress et al. 2015). By taking advantage of the multiplex nature of next-generation sequencing (NGS) and the third-generation sequencing platform, metabarcoding not only enables assignment of multiple species using DNA barcodes in a mixed sample and makes the data output magnitudes more reliable, but also allows simultaneous processing of DNA barcodes for thousands of diverse specimens in a single sequencing run (Coissac et al. 2012; Piper et al. 2019). Starting with minimal amounts of materials, theoretically current NGS technology with a maximum read length of 300 bp is highly suitable for mini-barcodes, while complete barcode can be recovered through assembly of short overlapping reads, or alternatively by third generation sequencing, which provides read lengths superior to any previous sequencing technology (Behjati and Tarpey 2013). A recent work with the real-time MinION sequencer, a portable third generation sequencer, has just achieved great barcode sequencing throughput at a cost of less than 10 cents, showing a promising future in this direction (Srivathsan et al. 2021). Moreover, sequences for independent gene loci can be garnered in parallel on HTS platforms in order to improve the phylogenetic resolution of generated data, a strategy first recommended in plants as reviewed in Sect. 2.2, though individual barcode from these multi-locus combinations in this context cannot be linked together without the assistance of an extensive reference library (Taberlet et al. 2012).

Consistent with the rise of DNA metabarcoding, the past few years also represent a surge in bioinformatics advancement for taxonomic analysis. Since there are already many excellent reports on computational pipelines to process large quantities of sequence reads, here we will only briefly overview the fundamentals of key steps relevant to metabarcoding. Delimiting species basically is simple to implement if reference databases pre-exist and contain information from conspecifics. In this situation, query sequences from metabarcoding studies are directly matched to identified sequences of known species in references, commonly via similarity-based and tree-based algorithms that are frequently criticized though (DeSalle and Goldstein, 2019). Conversely, if reliable reference datasets are absent, query sequences would not be linked to a taxonomic name but would be binned together to form OTUs either according to their similarity (traditionally 97%), or based on their “true” biological sequences inferred using statistical models, which are also termed exact sequence variants (ESVs) or amplicon sequence variants (ASVs) in this context (García-López et al. 2021). These biological entities next can be compared with OTUs or ASVs in different studies, such as the BIN framework introduced by BOLD, to estimate the biodiversity of target samples. Yet biological interpretation of metabarcoding data can be seriously affected by the differences between the two methods: OTUs minimize the effects of slight variations in sequences that may or may not be of interest, but a small change, as in the case of parasitoid wasps, could be capturing actual differences between species; on the contrary, ASVs are defined as all “unique reads” within a metabarcoded dataset, often leading to a wrong differentiation between the SNPs of the same species, and in the same way making sequencing or PCR errors more prominent when compared to OTUs (Molik et al. 2020). By using simulations, it has been advised that approaches utilizing ASVs outperform OTUs only when the sequencing depth is sufficient to cover a biological complexity with low polymorphisms. Otherwise conclusions drawn from OTU analyses are more consistent (Joos et al. 2020). Therefore, which method would be chosen for the bioinformatic processing of metabarcoding should be dependent on the analysis desired, although OTUs currently seem to be less preferred with the continuous improvement of sequencing technologies.

When coming back to HTS technologies, it has been argued that the current barcoding practice could soon become obsolete and irrelevant as genomic data are created by untargeted shotgun sequencing with increasing ease (Taylor and Harris 2012). In one regard, the high-throughput nature of these techniques not only allows a genomic surveillance to avoid numts, but also enables a full coverage of mitochondrial heteroplasmy to distinguish functional alleles based upon length, translation and other criteria. From another perspective, it is impossible in essence to produce

a precise representation of organismal divergence using a genetic estimate taken from just parts of the genome. Taken together, it is most likely to happen that enthusiasm for DNA barcode to the end will transition to a larger endeavor of archiving accessible genomic data. Before that, however, more sophisticated bioinformatic modellings with user-friendly interfaces, as well as huge genome storages as references with information dedicated to taxonomic relatedness must be developed. At present, the question lies in whether the barcoding enterprise tends to take measures and evolve its methodologies to embrace novel techniques that are inevitably on the way? In fact, many botanists have conducted genome skimming for entire plastid genomes and nuclear ribosomal regions to cover all of the different standard plant barcoding regions as an extended barcode, while others assemble the whole organelle genomes as a resource for validating and designing short, informative barcode markers with diagnostic nucleotides (Coissac et al. 2016; Kreuzer et al. 2019). These attempts represent a stepping stone on the continuum between routine barcode movements and complete genome sequences. Along this path, we believe DNA barcoding will be capable of further exploring its potential and opportunities, and perhaps one day will encompass other “omics” approaches such as proteomics and metabolomics.

## Practical utilities of DNA barcoding

Undoubtedly, DNA barcoding is a chief component of the modern diagnostic toolbox with increasing applications in taxonomy, systems biology and ecological studies. Prior to barcoding, conventional approaches for classification of species mainly rely upon the characterization of distinguishable morphology while many organisms exhibit morphologically distinct stages controlled by gender or life cycles (Hall and Martín-Vega 2019). Furthermore, suspect specimens may be damaged or incomplete with only part of tissue feasible for identification. All these pitfalls would render morphological determination unclear or unlikely, but can be easily avoided with molecular barcoding. Besides traditional way by sampling separate individuals, barcode technology, especially metabarcoding, can be adopted for assessment to dietary items using gut contents and scats of animals, or utilized for analyzing environmental samples, namely samples from soil, water and even air that possibly contain DNA materials from life, for biomonitoring and disease screening (Chaves et al. 2012; Haag et al. 2009; Staats et al. 2016). Together with mini-barcodes, it could further mitigate problems with fragmented DNA present in the environment, gut contents or other sources of exogenous DNA (Prerna Vohra 2013). Another potentially valuable utility of combining metabarcoding with mini-barcodes is to analyze invertebrate-derived DNA (iDNA), where vertebrate genetic material is



extracted from diverse invertebrates, including terrestrial leeches, mosquitoes, midges, blow flies and ticks (Schnell et al. 2015). iDNA has recently been proposed as a powerful non-invasive tool to detect vertebrate species and to survey their population as long as the information regarding the biology, habitats, behaviors and diets of relevant invertebrates is secured.

The usefulness of DNA barcoding is not restricted to scientific research of biodiversity, but also concerns conservation, public health and biosafety (Fig. 2). It is not surprising that barcoding is highly desirable for customs and national authorities in the conservation area of rare wildlife. International conventions such as the Convention on International Trade in Endangered Species of Fauna and Flora (CITES) have categorized more than 35,000 species as threatened by extinction (Wyatt 2020). And DNA barcodes have been demonstrated to be helpful to monitor illegal collection and trade of protected species and their products when morphological characters were equivocal (Chapagain et al. 2021). Additionally, the risks of

pandemic spillover are higher than ever with increasingly intimate associations between humans and wildlife (or their meat), some of which might serve as hosts or vectors for medically important pathogens. For example, there are about 3500 species of mosquitoes, but only a handful of species spread malaria, dengue fever and other diseases in tropical areas (James 2007; Lee et al. 2018). DNA barcoding actually has been reported to successfully determine mosquitoes involved in disease transmission and public health. In the same way, genetic authentications using barcodes are also becoming more and more common in food adulteration and manufacture of drugs of natural origin (e.g. herbal products or mixtures in traditional Chinese medicine), misidentification of which sometimes could be poisonous and life-threatening (Kreuzer et al. 2019; Wu et al. 2019). Aside from herbal medicine, metabarcoding of pollen and fungal spores can also be incorporated into forensic palynology and security intelligence to link persons or objects with particular places and times, given pollen and fungal spores' ubiquity in the environment, their

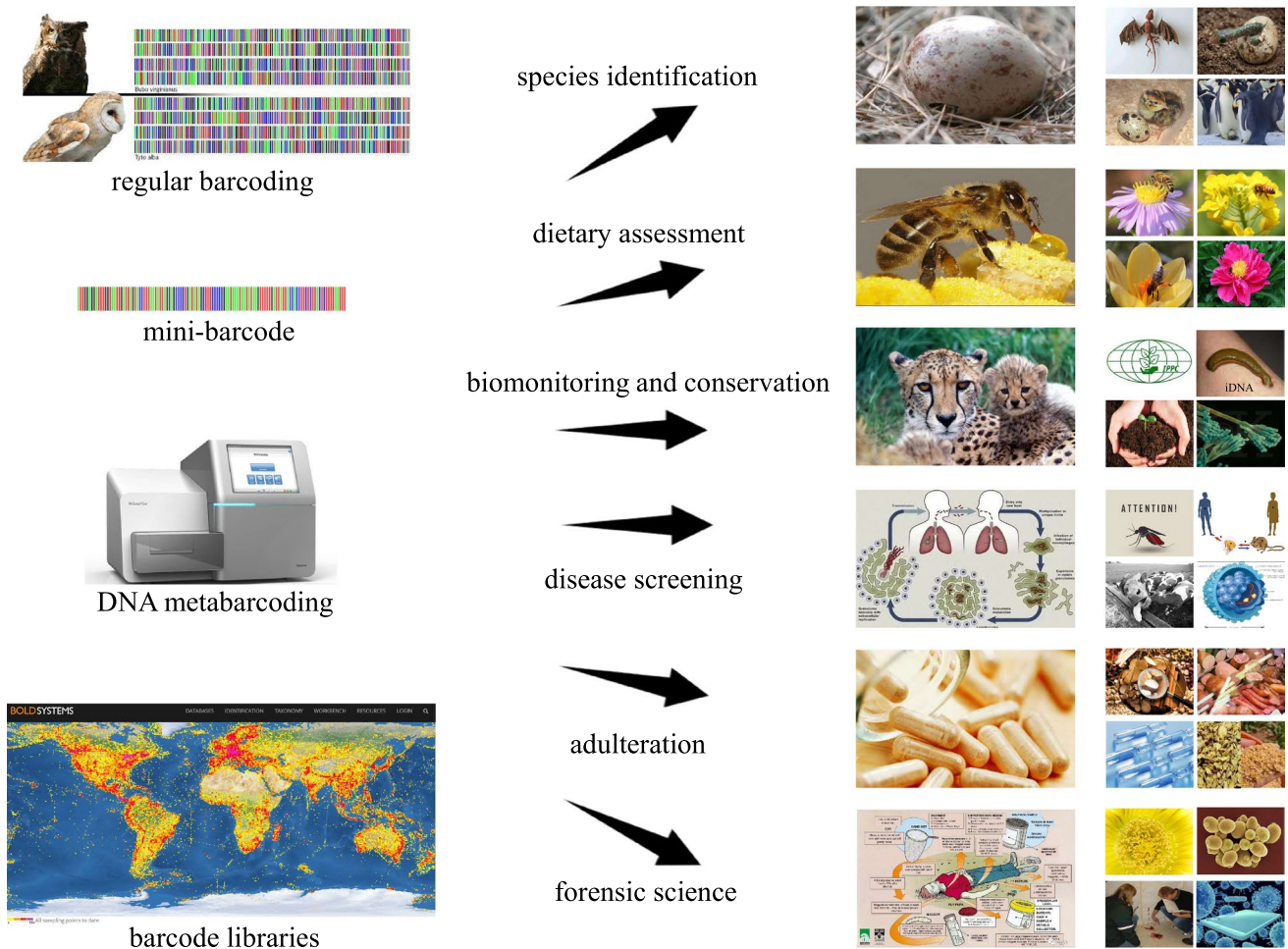


Fig. 2 Practical utilities of DNA barcoding technology

potential for geographic and temporal inference, and their long-term durability (Bell et al. 2016).

## Conclusions

By employing sequence divergence in several short and standardized gene fragments, DNA barcode and its library have become an invaluable addition to our suite of tools to understand life and nature. Although posing many controversies, DNA barcoding no doubt holds great promise for potentially widespread scientific and practical benefits. With the exploration of mini-barcodes and metabarcoding in DNA-based species delineation, it is believed that barcode techniques will be further integrated into a wider context of scientific, political, economic and social areas.

Then, as the barcoded reference species expands across the tree of life, ultimately one must ask whether it is possible to barcode all life on Earth. In theory, the barcoding process is able to yield 100% accuracy of species delimitation as long as robust thresholds defining species boundaries are established, which is truly tough to settle for all of life (Matute and Sepúlveda 2019). Also, it would be naïve to portray a species or infer a phylogeny without any corroborating evidence other than certain pieces of DNA sequences. In the absence of other evidence, DNA barcoding creates hypotheses regarding new species rather than outright discovering them (Taylor and Harris 2012). More importantly, it should be noted that barcoding must supplement morphological data for species description, which usually fails to break into the mainstream of barcoding studies despite the fact that morphological identification laid the foundation of all barcode databases. To sum up, what we know today is that no single classification approach can be applied universally for all species. DNA barcodes in conjunction with traditional taxonomic tools for sure are more rapid and more reliable than any method alone for disclosing cryptic and overlooked biodiversity.

**Acknowledgements** This work is supported by the National Key R&D Program of China (2018YFC1200201) and funding from Hai Nan, Qinghai Province (2021-KT07-B and 2022-KT09-D).

## Declarations

**Conflict of interest** The authors have declared that no conflicts of interest exist.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are

included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Adamowicz SJ, Hollingsworth PM, Ratnasingham S, van der Bank M (2017) International Barcode of Life: Focus on big biodiversity in South Africa. *Genome* 60:875–879. <https://doi.org/10.1139/gen-2017-0210>
- Adékambi T, Drancourt M, Raoult D (2009) The rpoB gene as a tool for clinical microbiologists. *Trends Microbiol* 17:37–45. <https://doi.org/10.1016/j.tim.2008.09.008>
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402. <https://doi.org/10.1093/nar/25.17.3389>
- Arnot DE, Roper C, Bayoumi RA (1993) Digital codes from hyper-variable tandemly repeated DNA sequences in the *Plasmodium falciparum* circumsporozoite gene can genetically barcode isolates. *Mol Biochem Parasitol* 61:15–24. [https://doi.org/10.1016/0166-6851\(93\)90154-p](https://doi.org/10.1016/0166-6851(93)90154-p)
- Asselah T, Durantel D, Pasmant E, Lau G, Schinazi RF (2021) COVID-19: Discovery, diagnostics and drug development. *J Hepatol* 74:168–184. <https://doi.org/10.1016/j.jhep.2020.09.031>
- Bano I, Sharif M, Alam S (2022) Genetic drift in the genome of SARS COV-2 and its global health concern. *J Med Virol* 94:88–98. <https://doi.org/10.1002/jmv.27337>
- Bates ST, Berg-Lyons D, Caporaso JG, Walters WA, Knight R, Fierer N (2011) Examining the global distribution of dominant archaeal populations in soil. *Isme j* 5:908–917. <https://doi.org/10.1038/ismej.2010.171>
- Behjati S, Tarpey PS (2013) What is next generation sequencing? *Arch Dis Child Educ Pract Ed* 98:236–238. <https://doi.org/10.1136/archdischild-2013-304340>
- Bell KL, Burgess KS, Okamoto KC, Aranda R, Brosi BJ (2016) Review and future prospects for DNA barcoding methods in forensic palynology. *Forensic Sci Int Genet* 21:110–116. <https://doi.org/10.1016/j.fsigen.2015.12.010>
- Benga L, Benten PM, Engelhardt E, Köhrer K, Hueber B, Nicklas W, Christensen H, Sager M (2020) Differentiation among rodent-bacter species based on 16S–23S rRNA internal transcribed spacer analysis. *Comp Med* 70:487–491. <https://doi.org/10.30802/aalas-cm-99-990085>
- Bicknell JN, Douglas HC (1970) Nucleic acid homologies among species of *Saccharomyces*. *J Bacteriol* 101:505–512. <https://doi.org/10.1128/jb.101.2.505-512.1970>
- Blackwell M (2011) The fungi: 1, 2, 3 ... 5.1 million species? *Am J Bot* 98:426–438. <https://doi.org/10.3732/ajb.1000298>
- Braukmann TW, Kuzmina ML, Sills J, Zakharov EV, Hebert PD (2017) Testing the Efficacy of DNA Barcodes for Identifying the Vascular Plants of Canada. *PLoS ONE* 12:e0169515. <https://doi.org/10.1371/journal.pone.0169515>
- Brousseau R, Hill JE, Préfontaine G, Goh SH, Harel J, Hemmingsen SM (2001) *Streptococcus suis* serotypes characterized by analysis of chaperonin 60 gene sequences. *Appl Environ Microbiol* 67:4828–4833. <https://doi.org/10.1128/aem.67.10.4828-4833.2001>

- Carbone I, Kohn LMJM (1999) A method for designing primer sets for speciation studies in filamentous ascomycetes. *Mycologia* 91:553–556. <https://doi.org/10.2307/3761358>
- Chapagain DJ, Meilby H, Baniya CB, Budha-Magar S, Ghimire SK (2021) Illegal harvesting and livestock grazing threaten the endangered orchid *Dactylorhiza hatagirea* (D. Don) Soó in Nepalese Himalaya. *Ecol Evol* 11:6672–6687. <https://doi.org/10.1002/ece3.7520>
- Chaves PB, Graeff VG, Lion MB, Oliveira LR, Eizirik E (2012) DNA barcoding meets molecular scatology: short mtDNA sequences for standardized species assignment of carnivore noninvasive samples. *Mol Ecol Resour* 12:18–35. <https://doi.org/10.1111/j.1755-0998.2011.03056.x>
- Che J, Chen HM, Yang JX, Jin JQ, Jiang K, Yuan ZY, Murphy RW, Zhang YP (2012) Universal COI primers for DNA barcoding amphibians. *Mol Ecol Resour* 12:247–258. <https://doi.org/10.1111/j.1755-0998.2011.03090.x>
- Chen S, Xia T, Wang Y, Liu J, Chen S (2005) Molecular systematics and biogeography of *Crawfordia*, *metagentiana* and *tripterospermum* (Gentianaceae) based on nuclear ribosomal and plastid DNA sequences. *Ann Bot* 96:413–424. <https://doi.org/10.1093/aob/mci188>
- Chevigny N, Schatz-Daas D, Lotfi F, Gualberto JM (2020) DNA repair and the stability of the plant mitochondrial genome. *Int J Mol Sci* 21:328. <https://doi.org/10.3390/ijms21010328>
- Cho Y, Mower JP, Qiu YL, Palmer JD (2004) Mitochondrial substitution rates are extraordinarily elevated and variable in a genus of flowering plants. *Proc Natl Acad Sci U S A* 101:17741–17746. <https://doi.org/10.1073/pnas.0408302101>
- Church DL, Cerutti L, Gürtler A, Griener T, Zelazny A, Emler S (2020) Performance and application of 16S rRNA gene cycle sequencing for routine identification of bacteria in the clinical microbiology laboratory. *Clin Microbiol Rev* 33:e00053–e119. <https://doi.org/10.1128/cmr.00053-19>
- Clarridge JE (2004) Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. *Clin Microbiol Rev* 17:840–862. <https://doi.org/10.1128/cmr.17.4.840-862.2004>
- Coissac E, Riaz T, Puillandre N (2012) Bioinformatic challenges for DNA metabarcoding of plants and animals. *Mol Ecol* 21:1834–1847. <https://doi.org/10.1111/j.1365-294X.2012.05550.x>
- Coissac E, Hollingsworth PM, Lavergne S, Taberlet P (2016) From barcodes to genomes: extending the concept of DNA barcoding. *Mol Ecol* 25:1423–1428. <https://doi.org/10.1111/mec.13549>
- Crous PW, Groenewald JZ, Rissède JM, Simoneau P, Hyde KD (2006) *Calonectria* species and their *Cylindrocladium* anamorphs: species with clavate vesicles. *Stud Mycol* 55:213–226. <https://doi.org/10.3114/sim.55.1.213>
- de Vere N, Rich TC, Trinder SA, Long C (2015) DNA barcoding for plants. *Methods Mol Biol* 1245:101–118. [https://doi.org/10.1007/978-1-4939-1966-6\\_8](https://doi.org/10.1007/978-1-4939-1966-6_8)
- D'Errico I, Gadaleta G, Saccone C (2004) Pseudogenes in metazoa: origin and features. *Brief Funct Genomic Proteomic* 3:157–167. <https://doi.org/10.1093/bfpgp/3.2.157>
- DeSalle R, Goldstein P (2019) Review and interpretation of trends in DNA barcoding. *Front Ecol Evol*. <https://doi.org/10.3389/fevo.2019.00302>
- Dong W, Liu H, Xu C, Zuo Y, Chen Z, Zhou S (2014) A chloroplast genomic strategy for designing taxon specific DNA mini-bar-codes: a case study on ginsengs. *BMC Genet* 15:138. <https://doi.org/10.1186/s12863-014-0138-z>
- Dulla EL, Kathera C, Gurijala HK, Mallakuntla TR, Srinivasan P, Prasad V, Mopati RD, Jasti PK (2016) Highlights of DNA Bar-coding in identification of salient microorganisms like fungi. *J Mycol Med* 26:291–297. <https://doi.org/10.1016/j.mycmed.2016.05.006>
- Eberhardt U (2012) Methods for DNA barcoding of fungi. *Methods Mol Biol* 858:183–205. [https://doi.org/10.1007/978-1-61779-591-6\\_9](https://doi.org/10.1007/978-1-61779-591-6_9)
- Emerson BC, Cicconardi F, Fanciulli PP, Shaw PJ (2011) Phylogeny, phylogeography, phylobetadiversity and the molecular analysis of biological communities. *Philos Trans R Soc Lond B Biol Sci* 366:2391–2402. <https://doi.org/10.1098/rstb.2011.0057>
- Fagan-Jeffries EP, Cooper SJB, Bertozzi T, Bradford TM, Austin AD (2018) DNA barcoding of microgastrine parasitoid wasps (Hymenoptera: Braconidae) using high-throughput methods more than doubles the number of species known for Australia. *Mol Ecol Resour*. <https://doi.org/10.1111/1755-0998.12904>
- Fang T, Liao S, Chen X, Zhao Y, Zhu Q, Cao Y, Wang Q, Zhang S, Gao Z, Yang Y, Wang Y, Zhang J (2019) Forensic drowning site inference employing mixed pyrosequencing profile of DNA barcode gene (rbcL). *Int J Legal Med* 133:1351–1360. <https://doi.org/10.1007/s00414-019-02075-4>
- Fazekas AJ, Burgess KS, Kesanakurti PR, Graham SW, Newmaster SG, Husband BC, Percy DM, Hajibabaei M, Barrett SC (2008) Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS ONE* 3:e2802. <https://doi.org/10.1371/journal.pone.0002802>
- Fernandes TJR, Amaral JS, Mafra I (2020) DNA barcode markers applied to seafood authentication: an updated review. *Crit Rev Food Sci Nutr*. <https://doi.org/10.1080/10408398.2020.1811200>
- Fode-Vaughan KA, Wimpee CF, Remsen CC, Collins ML (2001) Detection of bacteria in environmental samples by direct PCR without DNA extraction. *Biotechniques* 31:598. <https://doi.org/10.2144/01313rr04>
- Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol* 3:294–299
- Friedman JR, Nunnari J (2014) Mitochondrial form and function. *Nature* 505:335–343. <https://doi.org/10.1038/nature12985>
- Gao Z, Liu Y, Wang X, Wei X, Han J (2019) DNA mini-barcoding: a derived barcoding method for herbal molecular identification. *Front Plant Sci* 10:987. <https://doi.org/10.3389/fpls.2019.00987>
- García-López R, Cornejo-Granados F, Lopez-Zavala AA, Cota-Huizar A, Sotelo-Mundo RR, Gómez-Gil B, Ochoa-Leyva A (2021) OTUs and ASVs produce comparable taxonomic and diversity from shrimp microbiota 16S profiles using tailored abundance filters. *Genes (basel)*. <https://doi.org/10.3390/genes12040564>
- Giangrande A (2003) Biodiversity, conservation, and the “Taxonomic impediment.” *Aquatic Conservation-Marine and Freshwater Ecosystems* 13:451–459. <https://doi.org/10.1002/aqc.584>
- Giri B, Pandey S, Shrestha R, Pokharel K, Ligler FS, Neupane BB (2021) Review of analytical performance of COVID-19 detection methods. *Anal Bioanal Chem* 413:35–48. <https://doi.org/10.1007/s00216-020-02889-x>
- Glass NL, Donaldson GC (1995) Development of primer sets designed for use with the PCR to amplify conserved genes from filamentous ascomycetes. *Appl Environ Microbiol* 61:1323–1330. <https://doi.org/10.1128/aem.61.4.1323-1330.1995>
- Goldstein PZ, DeSalle R (2011) Integrating DNA barcode data and taxonomic practice: determination, discovery, and description. *BioEssays* 33:135–147. <https://doi.org/10.1002/bies.201000036>
- Group C.P.W (2009) A DNA barcode for land plants. *Proc Natl Acad Sci U S A* 106:12794–12797. <https://doi.org/10.1073/pnas.0905845106>
- Guo M, Peng Z, Zhang X, Yuan C, Lu H, Zhang K, Cai Y, Zhang W (2021) Isolation and characterization of DNA barcodes from distinctive and rare terrestrial animals in China using universal

- COI and 16S primers. *Stemmedicine* 2:e95-106. <https://doi.org/10.37175/stemmedicine.v2i7.95>
- Haag T, Santos AS, De Angelo C, Srbek-Araujo AC, Sana DA, Morato RG, Salzano FM, Eizirik E (2009) Development and testing of an optimized method for DNA-based identification of jaguar (*Panthera onca*) and puma (*Puma concolor*) faecal samples for use in ecological and genetic studies. *Genetica* 136:505–512. <https://doi.org/10.1007/s10709-008-9347-6>
- Hajibabaei M, McKenna C (2012) DNA mini-barcodes. *Methods Mol Biol* 858:339–353. [https://doi.org/10.1007/978-1-61779-591-6\\_15](https://doi.org/10.1007/978-1-61779-591-6_15)
- Hall MJR, Martín-Vega D (2019) Visualization of insect metamorphosis. *Philos Trans R Soc Lond B Biol Sci* 374:20190071. <https://doi.org/10.1098/rstb.2019.0071>
- Hardman M (2004) The phylogenetic relationships among Noturus catfishes (Siluriformes: Ictaluridae) as inferred from mitochondrial gene cytochrome b and nuclear recombination activating gene 2. *Mol Phylogenet Evol* 30:395–408. [https://doi.org/10.1016/s1055-7903\(03\)00203-3](https://doi.org/10.1016/s1055-7903(03)00203-3)
- Hausmann A, Godfray HC, Huemer P, Mutanen M, Rougerie R, van Nieukerken EJ, Ratnasingham S, Hebert PD (2013) Genetic patterns in European geometrid moths revealed by the Barcode Index Number (BIN) system. *PLoS ONE* 8:e84518. <https://doi.org/10.1371/journal.pone.0084518>
- Hawllitschek O, Morinière J, Dunz A, Franzen M, Rödder D, Glaw F, Haszprunar G (2016) Comprehensive DNA barcoding of the herpetofauna of Germany. *Mol Ecol Resour* 16:242–253. <https://doi.org/10.1111/1755-0998.12416>
- Hebert PD, Cywinska A, Ball SL, deWaard JR (2003a) Biological identifications through DNA barcodes. *Proc Biol Sci* 270:313–321. <https://doi.org/10.1098/rspb.2002.2218>
- Hebert PD, Ratnasingham S, deWaard JR (2003b) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc Biol Sci* 270(Suppl 1):S96–S99. <https://doi.org/10.1098/rsbl.2003.0025>
- Hebert PD, Stoeckle MY, Zemlak TS, Francis CM (2004) Identification of birds through DNA barcodes. *PLoS Biol* 2:e312. <https://doi.org/10.1371/journal.pbio.0020312>
- Hill GE (2016) Mitonuclear coevolution as the genesis of speciation and the mitochondrial DNA barcode gap. *Ecol Evol* 6:5831–5842. <https://doi.org/10.1002/ece3.2338>
- Hill GE (2020) Genetic hitchhiking, mitonuclear coadaptation, and the origins of mt DNA barcode gaps. *Ecol Evol* 10:9048–9059. <https://doi.org/10.1002/ece3.6640>
- Hoelzel AR, Hancock JM, Dover GA (1991) Evolution of the cetacean mitochondrial D-loop region. *Mol Biol Evol* 8:475–493. <https://doi.org/10.1093/oxfordjournals.molbev.a040662>
- James AA (2007) Preventing the spread of malaria and dengue fever using genetically modified mosquitoes. *J vis Exp*. <https://doi.org/10.3791/231>
- Jeanson ML, Labat JN, Little DP (2011) DNA barcoding: a new tool for palm taxonomists? *Ann Bot* 108:1445–1451. <https://doi.org/10.1093/aob/mcr158>
- Joos L, Beirinckx S, Haegeman A, Debode J, Vandecasteele B, Baeyen S, Goormachtig S, Clement L, De Tender C (2020) Daring to be differential: metabarcoding analysis of soil and plant-related microbial communities using amplicon sequence variants and operational taxonomical units. *BMC Genomics* 21:733. <https://doi.org/10.1186/s12864-020-07126-4>
- Khan SM, Smigrodzki RM, Swerdlow RH (2007) Cell and animal models of mtDNA biology: progress and prospects. *Am J Physiol Cell Physiol* 292:C658–C669. <https://doi.org/10.1152/ajpcell.00224.2006>
- Knapp S, Lamas G, Lughadha EN, Novarino G (2004) Stability or stasis in the names of organisms: the evolving codes of nomenclature. *Philos Trans R Soc Lond B Biol Sci* 359:611–622. <https://doi.org/10.1098/rstb.2003.1445>
- Kocher TD, Thomas WK, Meyer A, Edwards SV, Pääbo S, Villablanca FX, Wilson AC (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proc Natl Acad Sci U S A* 86:6196–6200. <https://doi.org/10.1073/pnas.86.16.6196>
- Kress WJ, Erickson DL (2007) A two-locus global DNA barcode for land plants: the coding rbcL gene complements the non-coding trnH-psbA spacer region. *PLoS ONE* 2:e508. <https://doi.org/10.1371/journal.pone.0000508>
- Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH (2005) Use of DNA barcodes to identify flowering plants. *Proc Natl Acad Sci U S A* 102:8369–8374. <https://doi.org/10.1073/pnas.0503123102>
- Kress WJ, García-Robledo C, Uriarte M, Erickson DL (2015) DNA barcodes for ecology, evolution, and conservation. *Trends Ecol Evol* 30:25–35. <https://doi.org/10.1016/j.tree.2014.10.008>
- Kreuzer M, Howard C, Adhikari B, Pendry CA, Hawkins JA (2019) Phylogenomic approaches to DNA barcoding of herbal medicines: developing clade-specific diagnostic characters for berberis. *Front Plant Sci* 10:586. <https://doi.org/10.3389/fpls.2019.00586>
- Lahaye R, van der Bank M, Bogarin D, Warner J, Pupulin F, Gigot G, Maurin O, Duthoit S, Barraclough TG, Savolainen V (2008) DNA barcoding the floras of biodiversity hotspots. *Proc Natl Acad Sci U S A* 105:2923–2928. <https://doi.org/10.1073/pnas.0709936105>
- Lane DJ (1991) 16S/23S rRNA Sequencing. *Nucleic Acid Techniques in Bacterial Systematics*. *Nucleic Acid Techniques in Bacterial Systematics* pp: 115–175.
- Lee H, Halverson S, Ezinwa N (2018) Mosquito-borne diseases. *Prim Care* 45:393–407. <https://doi.org/10.1016/j.pop.2018.05.001>
- Leray M, Knowlton N (2016) Censusing marine eukaryotic diversity in the twenty-first century. *Philos Trans R Soc Lond B Biol Sci*. <https://doi.org/10.1098/rstb.2015.0331>
- Li X, Yang Y, Henry RJ, Rossetto M, Wang Y, Chen S (2015) Plant DNA barcoding: from gene to genome. *Biol Rev Camb Philos Soc* 90:157–166. <https://doi.org/10.1111/brv.12104>
- Li H, Xiao W, Tong T, Li Y, Zhang M, Lin X, Zou X, Wu Q, Guo X (2021) The specific DNA barcodes based on chloroplast genes for species identification of Orchidaceae plants. *Sci Rep* 11:1424. <https://doi.org/10.1038/s41598-021-81087-w>
- Liu J, Li Q, Kong L, Yu H, Zheng X (2011) Identifying the true oysters (Bivalvia: Ostreidae) with mitochondrial phylogeny and distance-based DNA barcoding. *Mol Ecol Resour* 11:820–830. <https://doi.org/10.1111/j.1755-0998.2011.03025.x>
- Liu D, Liu L, Guo G, Wang W, Sun Q, Parani M, Ma J (2013) BOLD-Mirror: a global mirror system of DNA barcode data. *Mol Ecol Resour* 13:991–995. <https://doi.org/10.1111/1755-0998.12048>
- Loera-Sánchez M, Studer B, Kölliker R (2020) DNA barcode trnH-psbA is a promising candidate for efficient identification of forage legumes and grasses. *BMC Res Notes* 13:35. <https://doi.org/10.1186/s13104-020-4897-5>
- López JA, Chen W-J, Ortí G (2004) Esociform Phylogeny. *Copeia* 2004:449–464. <https://doi.org/10.1643/CG-03-087R1%J%J>
- Lyra ML, Haddad CFB, de Azeredo-Espin AML (2017) Meeting the challenge of DNA barcoding Neotropical amphibians: polymerase chain reaction optimization and new COI primers. *Mol Ecol Resour* 17:966–980. <https://doi.org/10.1111/1755-0998.12648>
- Marcelo R, Michelangeli FA, Renato GJ (2010) Phylogeny of Pleiochiton (Melastomataceae, Miconieae): total evidence. *Botanical J Linnean Soc* 162:423–434. <https://doi.org/10.1111/j.1095-8339.2009.01022.x>

- Matute DR, Sepúlveda VE (2019) Fungal species boundaries in the genomics era. *Fungal Genet Biol* 131:103249. <https://doi.org/10.1016/j.fgb.2019.103249>
- Maxfield JM, Van Tassell JL, St Mary CM, Joyeux JC, Crow KD (2012) Extreme gender flexibility: using a phylogenetic framework to infer the evolution of variation in sex allocation, phylogeography, and speciation in a genus of bidirectional sex changing fishes (Lythrypnus, Gobiidae). *Mol Phylogenet Evol* 64:416–427. <https://doi.org/10.1016/j.ympev.2012.04.016>
- McFadden CS, Benayahu Y, Pante E, Thoma JN, Nevarez PA, France SC (2011) Limitations of mitochondrial gene barcoding in Octocorallia. *Mol Ecol Resour* 11:19–31. <https://doi.org/10.1111/j.1755-0998.2010.02875.x>
- Mendes R, Garbeva P, Raaijmakers JM (2013) The rhizosphere microbiome: significance of plant beneficial, plant pathogenic, and human pathogenic microorganisms. *FEMS Microbiol Rev* 37:634–663. <https://doi.org/10.1111/1574-6976.12028>
- Meyer CP, Paulay G (2005) DNA barcoding: error rates based on comprehensive sampling. *PLoS Biol* 3:e422. <https://doi.org/10.1371/journal.pbio.0030422>
- Michelangeli FA, Penneys DS, Giza J, Soltis D, Skean MHHD, Taxon JJ (2004) A preliminary phylogeny of the tribe Miconieae (Melastomataceae) based on nrITS sequence data and its implications on inflorescence position. *Taxon* 53:279. <https://doi.org/10.2307/4135608>
- Milan DT, Mendes IS, Damasceno JS, Teixeira DF, Sales NG, Carvalho DC (2020) New 12S metabarcoding primers for enhanced Neotropical freshwater fish biodiversity assessment. *Sci Rep* 10:17966. <https://doi.org/10.1038/s41598-020-74902-3>
- Miller SE (2007) DNA barcoding and the renaissance of taxonomy. *Proc Natl Acad Sci U S A* 104:4775–4776. <https://doi.org/10.1073/pnas.0700466104>
- Mishmar D, Ruiz-Pesini E, Brandon M, Wallace DC (2004) Mitochondrial DNA-like sequences in the nucleus (NUMTs): insights into our African origins and the mechanism of foreign DNA integration. *Hum Mutat* 23:125–133. <https://doi.org/10.1002/humu.10304>
- Molik DC, Pfrender ME, Emrich SJ (2020) Uncovering effects from the structure of metabarcoding sequences for metagenetic and microbiome analysis. *Methods Protoc*. <https://doi.org/10.3390/mps3010022>
- Moorhouse-Gann RJ, Dunn JC, de Vere N, Goder M, Cole N, Hiperson H, Symondson WOC (2018) New universal ITS2 primers for high-resolution herbivory analyses using DNA metabarcoding in both tropical and temperate zones. *Sci Rep* 8:8542. <https://doi.org/10.1038/s41598-018-26648-2>
- Muster C, Spelda J, Rulik B, Thormann J, von der Mark L, Astrin JJ (2021) The dark side of pseudoscorpion diversity: The German Barcode of Life campaign reveals high levels of undocumented diversity in European false scorpions. *Ecol Evol* 11:13815–13829. <https://doi.org/10.1002/ece3.8088>
- Nitta JH, Ebihara A, Smith AR (2020) A taxonomic and molecular survey of the pteridophytes of the Nectandra Cloud Forest Reserve, Costa Rica. *Plos One* 15:e0241231. <https://doi.org/10.1371/journal.pone.0241231>
- Ohsako T, Ohnishi O (2000) Intra- and interspecific phylogeny of wild Fagopyrum (Polygonaceae) species based on nucleotide sequences of noncoding regions in chloroplast DNA. *Am J Bot* 87:573–582. <https://doi.org/10.2307/2656601>
- Olmstead R.G. (1996) *Molecular Systematics*, second edition.—David M. Hillis, Craig Moritz, and Barbara K. Mable (eds.) 1996. Sinauer, Sunderland, Massachusetts. 655 pp. \$4995 (paper). *Systematic Biology* 45:607–609. DOI: <https://doi.org/10.1093/sysbio/45.4.607>
- Oxelman B, Liden M, Berglund D (1997) Chloroplast rps16 intron phylogeny of the tribe Sileneae (Caryophyllaceae). *Plant Syst Evol* 206:393–410. <https://doi.org/10.1007/bf00987959>
- Palumbi S.R. (1991) *The Simple Fool's Guide to PCR*. Version 2.
- Panelli S., Brambati E., Bonacina C., Feligini M. (2013) Diversity of fungal flora in raw milk from the Italian Alps in relation to pasture altitude. *Springerplus* 2:405. <https://doi.org/10.1186/2193-1801-2-405>
- Parveen I, Singh HK, Raghuvanshi S, Pradhan UC, Babbar SB (2012) DNA barcoding of endangered Indian Paphiopedilum species. *Mol Ecol Resour* 12:82–90. <https://doi.org/10.1111/j.1755-0998.2011.03071.x>
- Pavan ME, Robles C, Cairó FM, Marcellino R, Pettinari MJ (2012) Identification of Corynebacterium pseudotuberculosis from sheep by PCR-restriction analysis using the RNA polymerase  $\beta$ -subunit gene (rpoB). *Res Vet Sci* 92:202–206. <https://doi.org/10.1016/j.rvsc.2011.02.007>
- Pino-Bodas R, Martín MP, Burgaz AR, Lumbsch HT (2013) Species delimitation in Cladonia (Ascomycota): a challenge to the DNA barcoding philosophy. *Mol Ecol Resour* 13:1058–1068. <https://doi.org/10.1111/1755-0998.12086>
- Piper AM, Batovska J, Cogan NOI, Weiss J, Cunningham JP, Rodoni BC, Blacket MJ (2019) Prospects and challenges of implementing DNA metabarcoding for high-throughput insect surveillance. *Gigascience*. <https://doi.org/10.1093/gigascience/giz092>
- Pratheepa M, Jalali SK, Arokiajaraj RS, Venkatesan T, Nagesh M, Panda M, Pattar S (2014) Insect barcode information system. *Bioinformation* 10:98–100. <https://doi.org/10.6026/97320630010098>
- Prum RO, Berv JS, Dornburg A, Field DJ, Townsend JP, Lemmon EM, Lemmon AR (2015) A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature* 526:569–573. <https://doi.org/10.1038/nature15697>
- Quaedvlieg W, Groenewald JZ, de Jesús Yáñez-Morales M, Crous PW (2012) DNA barcoding of Mycosphaerella species of quarantine importance to Europe. *Persoonia* 29:101–115. <https://doi.org/10.3767/003158512x661282>
- Ratnasingham S., Hebert P.D. (2007) The Barcode of Life Data System (<http://www.barcodinglife.org>). *Mol Ecol Notes* 7:355–364. DOI: <https://doi.org/10.1111/j.1471-8286.2007.01678.x>.
- Ratnasingham S, Hebert PD (2013) A DNA-based registry for all animal species: the barcode index number (BIN) system. *PLoS ONE* 8:e66213. <https://doi.org/10.1371/journal.pone.0066213>
- Reginato M, Michelangeli FA (2016) Untangling the phylogeny of Leandra s.str. (Melastomataceae, Miconieae). *Mol Phylogenet Evol* 96:17–32. <https://doi.org/10.1016/j.ympev.2015.11.015>
- Robba L, Russell SJ, Barker GL, Brodie J (2006) Assessing the use of the mitochondrial cox1 marker for use in DNA barcoding of red algae (Rhodophyta). *Am J Bot* 93:1101–1108. <https://doi.org/10.3732/ajb.93.8.1101>
- Robert LV, Szöke S, Eberhardt U, Cardinali G, Meyer W, Seifert KA, Lévesque C, Lewis CT (2011) The quest for a general and reliable fungal DNA Barcode. *Open Applied Informatics Journal* 5:45–61. <https://doi.org/10.2174/1874136301105010045>
- Sang T, Crawford D, Stuessy T (1997) Chloroplast DNA phylogeny, reticulate evolution, and biogeography of Paeonia (Paeoniaceae). *Am J Bot* 84:1120. <https://doi.org/10.2307/2446155>
- Saux C, Fisher BL, Spicer GS (2004) Dracula ant phylogeny as inferred by nuclear 28S rDNA sequences and implications for ant systematics (Hymenoptera: Formicidae: Amblyoponinae). *Mol Phylogenet Evol* 33:457–468. <https://doi.org/10.1016/j.ympev.2004.06.017>
- Sayers EW, Cavanaugh M, Clark K, Ostell J, Pruitt KD, Karsch-Mizrachi I (2020) GenBank. *Nucleic Acids Res* 48:D84–d86. <https://doi.org/10.1093/nar/gkz956>
- Scheffers BR, Joppa LN, Pimm SL, Laurance WF (2012) What we know and don't know about Earth's missing biodiversity. *Trends*

- Ecol Evol 27:501–510. <https://doi.org/10.1016/j.tree.2012.05.008>
- Schnell IB, Sollmann R, Calvignac-Spencer S, Siddall ME, Yu DW, Wilting A, Gilbert MT (2015) iDNA from terrestrial haematophagous leeches as a wildlife surveying and monitoring tool - prospects, pitfalls and avenues to be developed. *Front Zool* 12:24. <https://doi.org/10.1186/s12983-015-0115-z>
- Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA, Chen W (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proc Natl Acad Sci U S A* 109:6241–6246. <https://doi.org/10.1073/pnas.1117018109>
- Scorzetti G, Fell JW, Fonseca A, Statzell-Tallman A (2002) Systematics of basidiomycetous yeasts: a comparison of large subunit D1/D2 and internal transcribed spacer rDNA regions. *FEMS Yeast Res* 2:495–517. <https://doi.org/10.1111/j.1567-1364.2002.tb00117.x>
- Seifert KA, Samson RA, Dewaard JR, Houbraken J, Lévesque CA, Moncalvo JM, Louis-Seize G, Hebert PD (2007) Prospects for fungus identification using CO1 DNA barcodes, with *Penicillium* as a test case. *Proc Natl Acad Sci U S A* 104:3901–3906. <https://doi.org/10.1073/pnas.0611691104>
- Shearer TL, Van Oppen MJ, Romano SL, Wörheide G (2002) Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria). *Mol Ecol* 11:2475–2487. <https://doi.org/10.1046/j.1365-294x.2002.01652.x>
- Shokralla S, Hellberg RS, Handy SM, King I, Hajibabaei M (2015) A DNA mini-barcoding system for authentication of processed fish products. *Sci Rep* 5:15894. <https://doi.org/10.1038/srep15894>
- Sikes DS, Bowser M, Morton JM, Bickford C, Meierotto S, Hildebrandt K (2017) Building a DNA barcode library of Alaska's non-marine arthropods. *Genome* 60:248–259. <https://doi.org/10.1139/gen-2015-0203>
- Smith MA, Rodriguez JJ, Whitfield JB, Deans AR, Janzen DH, Hallwachs W, Hebert PD (2008) Extreme diversity of tropical parasitoid wasps exposed by iterative integration of natural history, DNA barcoding, morphology, and collections. *Proc Natl Acad Sci U S A* 105:12359–12364. <https://doi.org/10.1073/pnas.0805319105>
- Sobenin IA, Mitrofanov KY, Zhelankin AV, Sazonova MA, Postnov AY, Revin VV, Bobryshev YV, Orekhov AN (2014) Quantitative assessment of heteroplasmy of mitochondrial genome: perspectives in diagnostics and methodological pitfalls. *Biomed Res Int* 2014:292017. <https://doi.org/10.1155/2014/292017>
- Soltan Dallal MM, Validi M, Douraghi M, Bakhshi B (2019) Molecular typing of cytotoxin-producing *Klebsiella oxytoca* isolates by 16S–23S internal transcribed spacer PCR. *New Microbes New Infect* 30:100545. <https://doi.org/10.1016/j.nmni.2019.100545>
- Srivathsan A, Lee L, Katoh K, Hartop E, Kutty SN, Wong J, Yeo D, Meier R (2021) ONTbarcoder and MinION barcodes aid biodiversity discovery and identification by everyone, for everyone. *BMC Biol* 19:217. <https://doi.org/10.1186/s12915-021-01141-x>
- Staats M, Arulandhu AJ, Gravendeel B, Holst-Jensen A, Scholtens I, Peelen T, Prins TW, Kok E (2016) Advances in DNA metabarcoding for food and wildlife forensic species identification. *Anal Bioanal Chem* 408:4615–4630. <https://doi.org/10.1007/s00216-016-9595-8>
- Stefano GB, Bjenning C, Wang F, Wang N, Kream RM (2017) Mitochondrial Heteroplasmy. *Adv Exp Med Biol* 982:577–594. [https://doi.org/10.1007/978-3-319-55330-6\\_30](https://doi.org/10.1007/978-3-319-55330-6_30)
- Steinke D, Prosser SW, Hebert PD (2016) DNA barcoding of marine metazoans. *Methods Mol Biol* 1452:155–168. [https://doi.org/10.1007/978-1-4939-3774-5\\_10](https://doi.org/10.1007/978-1-4939-3774-5_10)
- Suhr MJ, Hallen-Adams HE (2015) The human gut mycobiome: pitfalls and potentials—a mycologist's perspective. *Mycologia* 107:1057–1073. <https://doi.org/10.3852/15-147>
- Sun CH, Liu DW, Huang YL, Zhou YW, Hou SL, Lu CH (2019) Genetic diversity analysis of Peking gecko (*Gekko swinhonis*) in mid-Eastern China based on mitochondrial COI and Cyt b gene sequences. *Mitochondrial DNA B Resour* 4:2156–2158. <https://doi.org/10.1080/23802359.2019.1623724>
- Sundquist A, Bigdeli S, Jalili R, Druzin ML, Waller S, Pullen KM, El-Sayed YY, Taslimi MM, Batzoglou S, Ronaghi M (2007) Bacterial flora-typing with targeted, chip-based Pyrosequencing. *BMC Microbiol* 7:108. <https://doi.org/10.1186/1471-2180-7-108>
- Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, Vermet T, Corthier G, Brochmann C, Willerslev E (2007) Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Res* 35:e14. <https://doi.org/10.1093/nar/gkl938>
- Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (2012) Towards next-generation biodiversity assessment using DNA metabarcoding. *Mol Ecol* 21:2045–2050. <https://doi.org/10.1111/j.1365-294X.2012.05470.x>
- Takashima M, Suh SO, Bai FY, Sugita T (2019) Takashi Nakase's last tweet: what is the current direction of microbial taxonomy research? *FEMS Yeast Res*. <https://doi.org/10.1093/femsyr/foz066>
- Tate JA, Simpson BB (2003) Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploid species. *Syst Bot* 28:723–737. <https://doi.org/10.1043/02-64.1>
- Taylor HR, Harris WE (2012) An emergent science on the brink of irrelevance: a review of the past 8 years of DNA barcoding. *Mol Ecol Resour* 12:377–388. <https://doi.org/10.1111/j.1755-0998.2012.03119.x>
- Tchaïcka L, Eizirik E, De Oliveira TG, Cândido JF Jr, Freitas TR (2007) Phylogeography and population history of the crab-eating fox (*Cerdoconyx thous*). *Mol Ecol* 16:819–838. <https://doi.org/10.1111/j.1365-294X.2006.03185.x>
- Thacker CE (2003) Molecular phylogeny of the gobioid fishes (Teleostei: Perciformes: Gobioidae). *Mol Phylogenet Evol* 26:354–368. [https://doi.org/10.1016/s1055-7903\(02\)00361-5](https://doi.org/10.1016/s1055-7903(02)00361-5)
- Trigo TC, Freitas TR, Kunzler G, Cardoso L, Silva JC, Johnson WE, O'Brien SJ, Bonatto SL, Eizirik E (2008) Inter-species hybridization among Neotropical cats of the genus *Leopardus*, and evidence for an introgressive hybrid zone between *L. geoffroyi* and *L. tigrinus* in southern Brazil. *Mol Ecol* 17:4317–4333. <https://doi.org/10.1111/j.1365-294X.2008.03919.x>
- Tyson JR, James P, Stoddart D, Sparks N, Wickenhagen A, Hall G, Choi JH, Lapointe H, Kamelian K, Smith AD, Prystajeky N, Goodfellow I, Wilson SJ, Harrigan R, Snutch TP, Loman NJ, Quick J (2020) Improvements to the ARTIC multiplex PCR method for SARS-CoV-2 genome sequencing using nanopore. *bioRxiv*. <https://doi.org/10.1101/2020.09.04.283077>
- Vancuren SJ, Dos Santos SJ, Hill JE (2020) Evaluation of variant calling for cpn60 barcode sequence-based microbiome profiling. *PLoS ONE* 15:e0235682. <https://doi.org/10.1371/journal.pone.0235682>
- Vences M, Thomas M, Bonett RM, Vieites DR (2005) Deciphering amphibian diversity through DNA barcoding: chances and challenges. *Philos Trans R Soc Lond B Biol Sci* 360:1859–1868. <https://doi.org/10.1098/rstb.2005.1717>
- Vences M, Nagy ZT, Sonet G, Verheyen E (2012) DNA barcoding amphibians and reptiles. *Methods Mol Biol* 858:79–107. [https://doi.org/10.1007/978-1-61779-591-6\\_5](https://doi.org/10.1007/978-1-61779-591-6_5)
- Vialle A, Feau N, Allaire M, Didukh M, Martin F, Moncalvo JM, Hamelin RC (2009) Evaluation of mitochondrial genes as DNA barcode for Basidiomycota. *Mol Ecol Resour* 9:99–113. <https://doi.org/10.1111/j.1755-0998.2009.02637.x>
- Vohra P (2013) DNA barcoding: current advances and future prospects—a review. *Asian J Biol Life Sci* 2:185–189

- Ward RD (2012) FISH-BOL, a case study for DNA barcodes. *Methods Mol Biol* 858:423–439. [https://doi.org/10.1007/978-1-61779-591-6\\_21](https://doi.org/10.1007/978-1-61779-591-6_21)
- Weissleder R, Lee H, Ko J, Pittet MJ (2020) COVID-19 diagnostics in context. *Sci Transl Med*. <https://doi.org/10.1126/scitranslmed.abc1931>
- Wickes BL, Wiederhold NP (2018) Molecular diagnostics in medical mycology. *Nat Commun* 9:5135. <https://doi.org/10.1038/s41467-018-07556-5>
- Woese CR, Fox GE (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci U S A* 74:5088–5090. <https://doi.org/10.1073/pnas.74.11.5088>
- Wong KL, Wang J, But PP, Shaw PC (2004) Application of cytochrome b DNA sequences for the authentication of endangered snake species. *Forensic Sci Int* 139:49–55. <https://doi.org/10.1016/j.forsciint.2003.09.015>
- Wu Y, Yang Y, Liu M, Wang B, Wang Y, Wang H (2019) Applying COI barcode to identify animal origin of food. *J Food Sci* 84:1256–1265. <https://doi.org/10.1111/1750-3841.14627>
- Wyatt T (2020) Canada and the convention on international trade in endangered species of wild fauna and flora (CITES): lessons learned on implementation and compliance. *Liverp Law Rev*. <https://doi.org/10.1007/s10991-020-09267-8>
- Xia Y, Gu HF, Peng R, Chen Q, Zheng YC, Murphy RW, Zeng XM (2012) COI is better than 16S rRNA for DNA barcoding Asiatic salamanders (Amphibia: Caudata: Hynobiidae). *Mol Ecol Resour* 12:48–56. <https://doi.org/10.1111/j.1755-0998.2011.03055.x>
- Xu J (2016) Fungal DNA barcoding. *Genome* 59:913–932. <https://doi.org/10.1139/gen-2016-0046>
- Yahr R, Schoch CL, Dentinger BT (2016) Scaling up discovery of hidden diversity in fungi: impacts of barcoding approaches. *Philos Trans R Soc Lond B Biol Sci*. <https://doi.org/10.1098/rstb.2015.0336>
- Yao H, Song J, Liu C, Luo K, Han J, Li Y, Pang X, Xu H, Zhu Y, Xiao P, Chen S (2010) Use of ITS2 region as the universal DNA barcode for plants and animals. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0013102>
- Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, Schleifer KH, Whitman WB, Euzéby J, Amann R, Rosselló-Móra R (2014) Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* 12:635–645. <https://doi.org/10.1038/nrmicro3330>
- Young RG, Abbott CL, Therriault TW, Adamowicz SJ (2017) Barcode-based species delimitation in the marine realm: a test using Hexanauplia (Multicrustacea: Thecostraca and Copepoda). *Genome* 60:169–182. <https://doi.org/10.1139/gen-2015-0209>
- Zangl L, Daill D, Schweiger S, Gassner G, Koblmüller S (2020) A reference DNA barcode library for Austrian amphibians and reptiles. *PLoS ONE* 15:e0229353. <https://doi.org/10.1371/journal.pone.0229353>
- Zhang W, Fan X, Zhu S, Zhao H, Fu L (2013) Species-specific identification from incomplete sampling: applying DNA barcodes to monitoring invasive solanum plants. *PLoS ONE* 8:e55927. <https://doi.org/10.1371/journal.pone.0055927>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.