European Journal of Wildlife Research (2020) 66:80
https://doi.org/10.1007/s10344-020-01419-5

ORIGINAL ARTICLE

Caught red handed: iDNA points to wild source for CITES-protected contraband leeches

Kalani M. Williams[1,2] · Megan Barkdull[3] · Mai Fahmy[1,2] · Evon Hekkala[2,4] · Mark E. Siddall[1] · Sebastian Kvist[5,6]

Received: 12 November 2019 / Revised: 10 August 2020 / Accepted: 31 August 2020

Abstract

In October of 2018, Canada Border Services Agency (CBSA) at Pearson International Airport in Toronto notified the Wildlife Enforcement Branch of Environment and Climate Change Canada that a passenger had arrived aboard a flight from Russia with a large quantity of live leeches. The leeches had been discovered in the passenger's carry-on luggage. An enforcement officer with Environment and Climate Change Canada detained the leeches to identify the species in order to determine whether the import was lawful. We identified the leeches as *Hirudo verbana* and extracted DNA from the bloodmeals of a subsample of 240 leeches and used metabarcoding of 6 mitochondrial loci to determine the vertebrate host species on which the leeches had previously fed. Sixteen undomesticated vertebrate host species were identified from the bloodmeals of the imported leeches, indicating that these leeches were collected from wild habitats. Furthermore, the overlap of host species' distributions point to a possible collecting source in the Volga delta, the Danube delta, or the coastal area region on the east side of the Sea of Azov. Our findings support the utility of invertebrate-derived DNA (iDNA) as a valuable tool in forensic evaluation of trafficked wildlife and provide new evidence regarding illegal exploitation of the European medicinal leech.

Keywords iDNA · Leeches · Metabarcoding · Wildlife trafficking · CITES · *Hirudo verbana*

Introduction

On October 17, 2018, Canada Border Services Agency (CBSA) stationed at Toronto's Pearson International Airport discovered a passenger who arrived aboard a flight from Russia carrying 4788 live European medicinal leeches in his carry-on luggage. As per protocol, CBSA contacted the Wildlife Enforcement Branch of Environment and Climate Change Canada, who detained the leeches to determine if the import was lawful. The Convention on International Trade in Endangered Species (CITES), which oversees the movement of protected and endangered species across international borders, includes both *Hirudo medicinalis* and *Hirudo verbana* under their protection (IUCN 2019; CITES Appendices I, II and III 2017). Both species were erroneously considered *H. medicinalis* until reexamination of their morphology and genetics proved otherwise (Siddall et al. 2007). CITES regulates both *Hirudo* species (the only two species of annelids protected under CITES) under Appendix II, indicating that "trade must be controlled in order to avoid utilization incompatible with their survival." This entails obtaining permits for international exportation and importation, as well as professional preparation for shipment of live animals (CITES Appendices I, II and III 2017) and declarations regarding the source of the material. Upon detainment, it was unknown if the leeches had come from a breeder/supplier, or if they had been collected from the wild.

Since the advent of DNA fingerprinting (Jeffreys et al. 1985), molecular techniques have grown more sophisticated and standardized in forensics (Roewer 2013). Recently,

✉ Kalani M. Williams
kwilliams1@amnh.org

[1] Division of Invertebrate Zoology, American Museum of Natural History, New York, NY, USA

[2] Biological Sciences, Fordham University, Bronx, NY, USA

[3] Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY, USA

[4] Division of Vertebrate Zoology, American Museum of Natural History, New York, NY, USA

[5] Department of Natural History, Royal Ontario Museum, Toronto, ON, Canada

[6] Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, ON, Canada

Published online: 05 September 2020

genetic barcoding on mitochondrial loci has overtaken microsatellite data for species-level identifications pertaining to the multibillion-dollar illegal wildlife trade (Eaton et al. 2009). Mitochondrial loci are easier to amplify from samples that were not freshly collected from a living animal because of their circular structure (which puts them at lower risk for degradation than nuclear loci) and their high copy number (which provides additional starting copies of the loci for PCR) (Foran 2006). These traits make mitochondrial loci ideal for determining species identifications (and therefore IUCN and CITES status) of exotic game meat, pelts, furs, feathers, and other exports (Linacre and Tobe 2011; Speller et al. 2011; D'Amato et al. 2013; Pilli et al. 2014).

Use of invertebrate-derived DNA (iDNA) is a recent advancement in biodiversity monitoring that harnesses metabarcoding of mitochondrial genes to identify vertebrate species that invertebrates (such as biting flies and leeches) fed on by targeting DNA from the bloodmeal inside the invertebrate (Schnell et al. 2015; Tessler et al. 2018; Drinkwater et al. 2019; Fahmy et al. 2019; Fahmy et al. 2020). Thus far, iDNA studies using leeches have only leveraged terrestrial haemadipsid leeches. Haemadipsid leeches are generalists in host selection (Sawyer 1986; Rocha et al. 2012), therefore providing a broad view of vertebrate diversity (Calvignac-Spencer et al. 2013). Furthermore, because haemadipsid leeches are philopatric, unlike flying hematophagous insects, investigators can assume the species identified in the bloodmeal were recently present in the region of collection (Tessler et al. 2018; Fahmy et al. 2020). While these features have not been studied in the iDNA context for aquatic leeches such as *Hirudo*, the applicability of these leeches for iDNA studies can be estimated based on the wealth of research that underlies our knowledge of the natural history of *Hirudo* species. *Hirudo* leeches have experimentally been found to preserve high-quality host DNA in their gut for at least 4 months (Schnell et al. 2012). Like their terrestrial relatives, aquatic *Hirudo* leeches are known to feed off a wide variety of vertebrate hosts (Sawyer 1986). These include amphibians, birds, mammals, and fish (Wilkin 1989; Keim 1993). *Hirudo* also appear to be equivalent in their ability to preserve high-quality host DNA. Schnell et al. (2012) experimentally determined that both medicinal *Hirudo* leeches and terrestrial haemadipsid leeches could preserve host DNA for at least 4 months. While aquatic leeches are likely more mobile than terrestrial leeches, in most cases they are restricted to the body of water they inhabit and are hypothesized to rely on their hosts for dispersal to new bodies of water (Utevsky et al. 2010; Trontelj and Utevsky 2012). Thus, aquatic *Hirudo* leeches should be able to retain host DNA and represent the aquatic vertebrate fauna of the local area.

In order to identify the most likely origin of the confiscated aquatic medicinal leeches, we used molecular genomic metabarcoding of the leech bloodmeals. While iDNA surveys have previously been leveraged to characterize the vertebrate fauna in a particular region, the corollary, pinpointing a previously unknown collection region based on the diversity of fauna identified from iDNA, has not been attempted. Provided with mitochondrial signatures of a sufficiently large number of leeches, one should be able to estimate the geographic origin of the leeches. Similarly, the relative species richness of vertebrate hosts and relative abundance of domesticated animal DNA present in the iDNA should allow investigators to deduce whether the leeches were purchased from a licensed captive breeding supplier or were collected from wild populations. Collection of medicinal leeches from the wild has had detrimental effects on populations for centuries and recently renewed exploitation of leeches for both traditional and modern medical use has put these populations at risk of local extinction (Elliott and Kutschera 2011).

Materials and methods

One thousand live leeches were received at the American Museum of Natural History from Environment and Climate Change Canada on December 14, 2018, 59 days after having been impounded at Pearson International Airport, during which time they had not had the opportunity to feed. The leeches were identified as *Hirudo verbana* morphologically and molecularly (using the COI locus). Two hundred and forty of these leeches were relaxed and fixed in 95% ethanol.

The posterior one-third of each leech, anterior to the caudal sucker (the post-cecal crop region), was targeted for DNA isolation, regardless of whether or not a bloodmeal was observable. A 5-mm transverse slice was cut from each leech, and then trisected sagittally in order to exclude the medial intestine, retaining the bilateral gastric postcaeca. Pairs of gastric postcaeca from 10 leeches were pooled for DNA isolations, forming 24 pools in total, representing 240 leeches. These samples were agitated with 0.5 mm ceramic beads in a Fisherbrand Bead Mill 24 at 1.95 m/s for 1 min. To isolate DNA, 200 μL of liquid was removed from the agitated samples for overnight incubation with 720 μL ATL buffer and 80 μL proteinase K. The remainder of the extraction followed Qiagen's DNeasy Blood and Tissue kit protocol.

Six loci targeting a range of vertebrate taxa (Table 1) were PCR amplified for each of the 24 pools of DNA isolates. With the exception of the 16S rDNA primers for fishes, each primer set was constructed in two versions: with a forward Illumina adaptor (ACACTCTTTCCCTACACGACGCTCTTCCGATCT) and with a reverse Illumina adaptor (GACTGGAGTTCAGACGTGTGCTCTTCCGATCT) in order to overcome any amplification inefficiencies introduced by the addition of the adaptors and to avoid Illumina sequencing error bias between read 1 and read 2 (first optimized by Fahmy et al. 2020). Thus, there were two PCR reactions for

Table 1 Primers used for amplification of mitochondrial iDNA. Base pairs = bp. Both forward and reverse primers are included in the primer set column

Target taxa	Target gene	Primer set (5′ – 3′)	Product length (bp)	Reference
Tetrapoda	12S rRNA	CTGGGATTAGATACCCCACTAT GTCGATTATAGGACAGGTTCCTCTA	120	Poinar et al. (2016)
Mammalia	16S rRNA	CGAGGGCTTTACTGTCTCTT CCTATTGTCGATATGGACTCT	294	Caragiulo et al. (2014)
Aves	NAD2	GGNGGNTGAATRGGNYTNAAYCARAC ACTCTTRTTTAAGGCTTTGAAGGC	1059	Payne and Sorenson (2007)
Amphibia and Osteichthyes	16S rRNA	AGACGAGAAGACCCYDTGGAGCTT GATCCAACATCGAGGTCGTAA	250	Vences et al. (2016)
Squamata	COI	TNTTMTCAACNAACCACAAAGA ACTTCTGGRTGKCCAAARAATCA	664	Nagy et al. (2012)
Osteichthyes and Chondrichthyes	16S rRNA	GACCCTATGGAGCTTTAGAC CGCTGTTATCCCTADRGTAACT	200	Berry et al. (2017) Deagle et al. (2007)

each targeted locus: one using the forward primer with a forward Illumina adaptor in combination with the reverse primer with a reverse Illumina adaptor, and another with the forward primer with a reverse Illumina adaptor in combination with the reverse primer with a forward Illumina adaptor. The 16S rDNA fishes PCR was performed twice to prevent it from being underrepresented, as all of the other primer sets had two versions.

All PCR reactions were performed using Illustra PuReTaq Ready-To-Go PCR Beads (GE Healthcare Life Sciences) with 0.5 μL of each (10 μM) primer, 22 μL of water, and 2 μL of DNA template. For all 12S and 16S primer reactions, the following thermocycling conditions were used: an initial denaturation step of 94 °C for 1 min, followed by 40 cycles of 94 °C for 15 s, 54 °C for 30 s, and 72 °C for 45 s, with a final extension of 72 °C for 2 min. The NAD2 and CO1 loci followed the same conditions, but required an annealing temperature of 50 °C.

Following amplification, all PCR products generated from any single template were combined together, totaling 24 amplicon aggregates, each containing the product of 12 unique PCR amplifications. These aggregates were then purified using a 2:1 carboxylated bead (Faircloth and Glenn 2012) to amplicon volumetric ratio and submitted for paired-end 250 bp sequencing by Genewiz on an Illumina Mi-Seq platform. Raw sequencing reads are deposited on NCBI's Sequence Read Archive (BioProject **PRJNA608698**).

After sequencing, primer sequences were informatically removed from the raw Illumina sequences and the retained sequences were trimmed for quality, using a phred score cut off of 33, and requiring a minimum length of 100 nucleotides (nt) using Trimmomatic v. 0.38 (Bolger et al. 2014). Trimmed sequences representing nested substrings then were dereplicated and clustered by 98% sequence similarity with USEARCH v.5 (Edgar 2010). Reads remained unmerged following the assumption that the clustering tool unites overlapping reads. Clustered sequences then were compared with BLASTn to a local database of whole vertebrate mitochondrial genomes in order to exclude contaminants (e.g., leeches and bacteria). Matching clusters were compared with the NCBI non-redundant nt nucleotide database (updated September 6, 2019) using the BLASTn algorithm, requiring an *e*-value of e^{-15} or better (Altschul et al. 1990; Camacho et al. 2009; Sayers et al. 2019).

In order to be considered a candidate species-level taxonomic match to a retrieved hit description, queried clusters were required to have a percent identity match greater than or equal to 98% over at least 80% of the locus length (i.e., 200 nucleotides for all loci save the tetrapod primers with a 120 bp product). Species-level matches also were required to exhibit a percent identity match to the best hit description that was 3 percentage points better than to any other species recovered. Queried clusters that failed to meet these criteria were either evaluated at the genus level or discarded.

Distribution data for each species-level identification was procured from the IUCN Global and European Red Lists (IUCN 2019), including both a list of countries where the species is extant and shapefiles of the global distribution of the species. *Hyla orientalis* and *Perccottus glenii* are not listed on the IUCN; therefore, country presence/absence data was supplied for these species from GBIF and FishBase (GBIF Secretariat 2019; Froese and Pauly 2019). The country presence data was used to construct a map showing what proportion of the species identified are known to be present in each country using Tableau, which harnesses OpenStreetMap (© Tableau Software 2019; © OpenStreetMap 2019). The distribution shapefiles for each species were overlaid to show regional hotspots using ArcGIS (ESRI 2011). The regions where all identified species overlap were highlighted as potential regions in which the leeches may have been collected.

Results

Clusters that were identified to the species-level are summarized in Table 2. Next-generation sequencing of the 24 pools

Table 2 Species identifications as determined by NCBI GenBank's BLASTn using the previously described criteria. For the species identified in each pool, the highest percent identity, longest alignment length, and lowest *e*-value for the best cluster (based on those criteria) are reported. Species denoted with * did not have full geographic distribution data available on the IUCN and are not included in Fig. 2

Scientific name	Common name	Gene	PID	Alignment length	*e*-value	Pool
Anas platyrhynchos	Mallard	12S rRNA	99.14	117	1.00E−49	3
Andrea alba	Great white egret	COI	99.60	251	1.00E−121	7
		NAD2	98.39	249	4.00E−117	7
Chlidonias hybrida	Whiskered tern	COI	99.60	250	2.00E−121	1, 13
		NAD2	98.80	250	9.00E−119	1, 13
Cygnus olor	Mute swan	COI	100.00	250	1.00E−121	1
		NAD2	100.00	250	2.00E−121	10
*Esox lucius**	Northern pike	COI	100.00	227	2.00E−110	1, 7, 9, 13
Fulica atra	Eurasian coot	COI	100.00	204	3.00E−98	2, 20
		NAD2	98.03	203	3.00E−92	20
*Hyla orientalis**	Shelkovnikov's tree frog	COI	99.17	245	6.00E−115	2, 22
Limosa limosa	Black-tailed godwit	COI	100.00	251	1.00E−121	2, 20
		NAD2	98.40	250	1.00E−117	2, 20
Microcarbo pygmeus	Pygmy cormorant	COI	100.00	250	1.00E−121	2, 24
		NAD2	100.00	250	9.00E−119	2, 24
*Nycticorax nycticorax**	Black-crowned night heron	12S rRNA	98.44	128	2.00E−54	1, 2, 3, 4
		NAD2	100.00	252	2.00E−123	1, 2
Pelobates fuscus	Common spadefoot toad	16S rRNA	100.00	252	2.00E−123	1, 2, 15, 20, 22
*Perccottus glenii**	Chinese sleeper	16S rRNA	100.00	251	2.00E−123	1, 13, 21, 24
		NAD2	100.00	250	4.00E−121	2, 21
Podiceps cristatus	Great crested grebe	16S rRNA	100.00	251	6.00E−121	16, 17, 18, 21, 22
		COI	100.00	252	2.00E−123	1, 2, 16, 21
		NAD2	100.00	250	4.00E−123	21
*Pseudorasbora parva**	Topmouth gudgeon	16S rRNA	100.00	251	2.00E−123	2
Rutilus rutilus	Common roach	16S rRNA	100.00	244	2.00E−119	9
Spatula querquedula	Garganey	NAD2	99.11	250	2.00E−117	3

generated a total of roughly 4 million reads, which were consolidated into 533,158 unique clusters of at least 98% intra-similarity. Of the clusters generated, 2923 were identifiable to the species-level using the stringent criteria described above collectively representing 16 distinct species (excluding *Homo sapiens*). These 16 species are comprised of 10 birds, 4 fish, and 2 amphibians. Of the 24 pools, 8 lacked any clusters yielding species-level identification. None of the species confidently identified were domesticated animals.

According to the IUCN, Bulgaria, Poland, Romania, Russia, and Serbia are the only countries in which all 16 species are extant (Fig. 1a). Full geographic distribution data that are more precise than country-level presence/absence are available for 11 of the 16 species identified (the species without these data are denoted in Table 2). A heat map representing the overlap of these 11 species shows them co-inhabiting in river deltas along the Black Sea, the Sea of Azov, and the Caspian Sea (Fig. 2a and b). More specifically, these deltas are in Moldova, Romania, and Ukraine along the Black Sea, Russia along the Sea of Azov, and Kazakhstan and Russia along the Caspian Sea. However, it is worth noting that Kazakhstan, Moldova, and Ukraine are not known to harbor all 16 species identified (Fig. 1a), suggesting all 16 species may only overlap in the Romanian delta on the Black Sea (the Danube delta), the Russian deltas on the Sea of Azov (between the Don delta and the Kuban delta), and the Russian delta on the Caspian Sea (the Volga delta).

Several pools of the sampled leech bloodmeals (5, 6, 8, 11, 12, 14, 19, 23, and) had no sequence reads that could be identified to the species-level. These pools tended to exhibit higher *e*-values and shorter alignment lengths for their vertebrate BLASTn hits. It is possible that pools containing little or no bloodmeal were overwhelmed by leech DNA during the PCR amplification process, or that the reads were simply too short for species identification (Hanya et al. 2019).

Some clusters that were not identified to the species-level nonetheless matched multiple species in the same genus

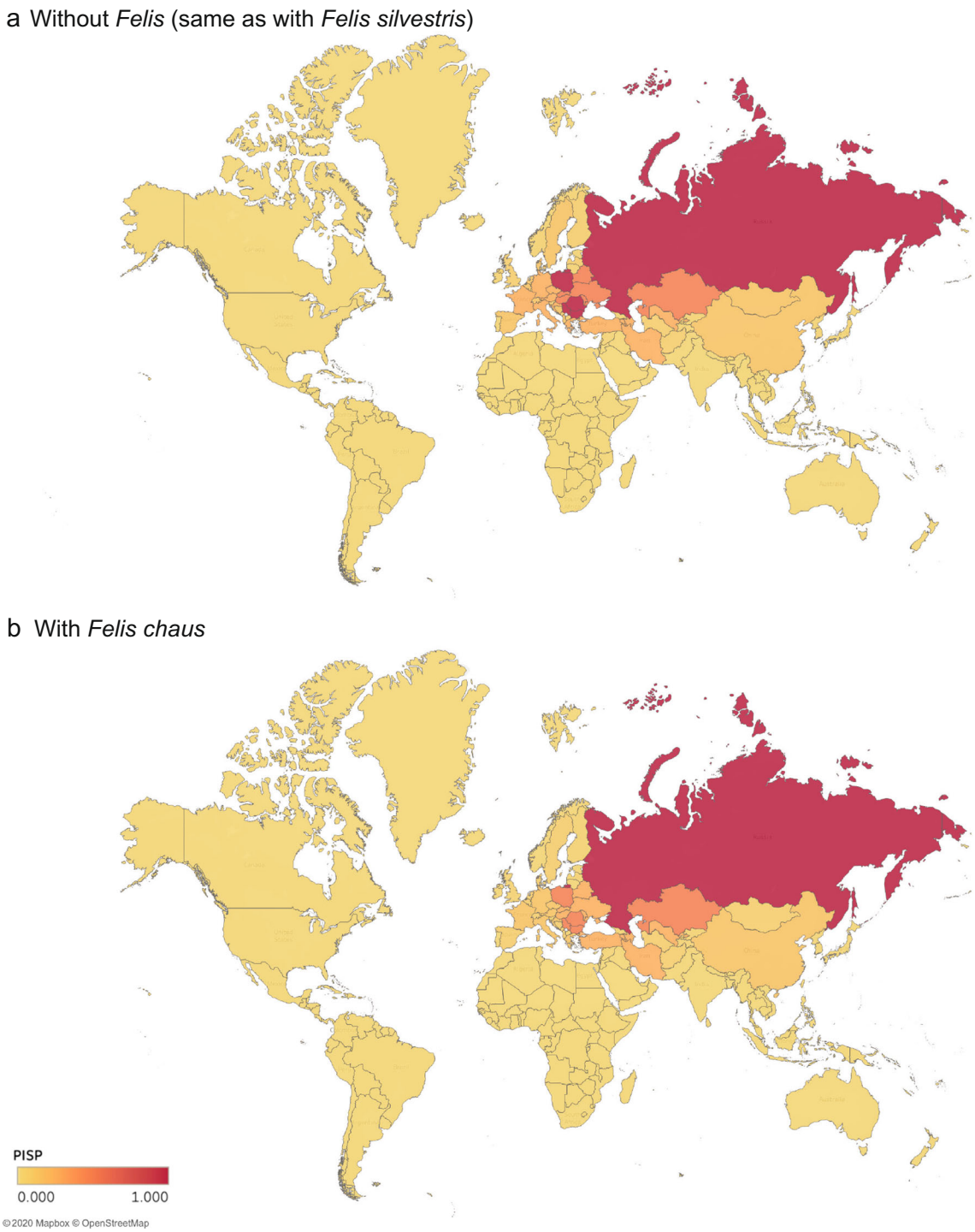

Fig. 1 Percent of identified species present by country, according to IUCN, GBIF, and FishBase distribution data (Dybowski 1877; IUCN 2019; GBIF Secretariat 2019; Froese and Pauly 2019). PISP = percent identified species present. Colorization is exponential to clarify the differences between countries with more species present. **a** The percent of identified species present by country excluding any *Felis* species (see text for further details). This map does not change if *Felis silvestris* is added. **b** The percent of identified species present by country including *Felis chaus*

equally well at the specified taxonomic identification criteria. *Anas acuta* could not be distinguished from *Anas crecca*, *Anas clyptea*, *Anas cyanoptera*, *Anas formosa*, or *Anas querquedula* on the basis of 16S rRNA, COI, or NAD2 in pools 3, 5, and 6. In pools 1, 2, 3, 5, and 7, *Anas platyrhynchos* could not be distinguished from *Anas poecilorhyncha* on the basis of 16S rRNA or NAD2. In pools 1, 2, and 8, several species of *Pelophylax* and *Rana* could not be distinguished on

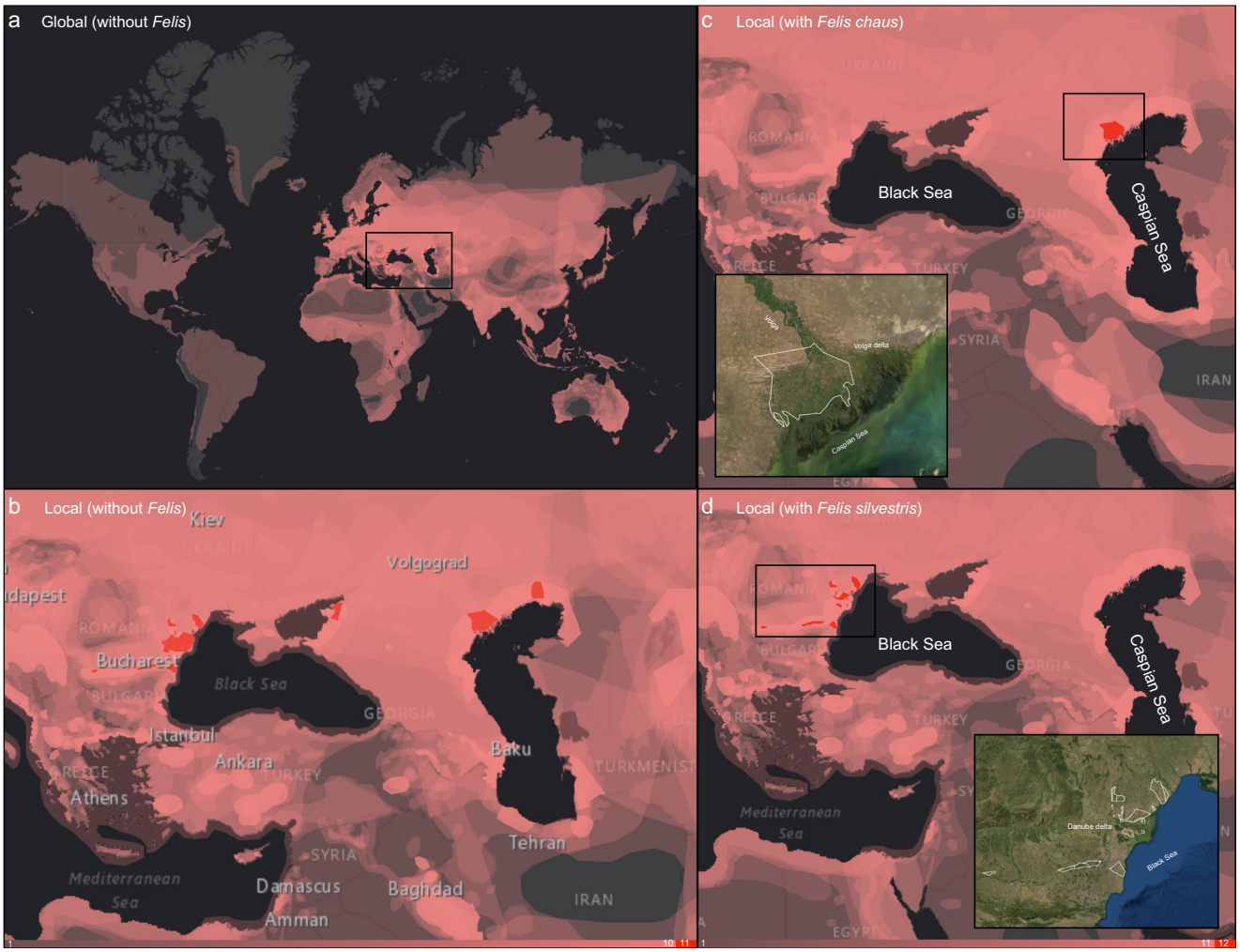

Fig. 2 Distribution overlays (data from IUCN 2019) of the vertebrate species recovered from iDNA gut contents of 240 individual *Hirudo verbana*. Coincidence of increasing numbers of vertebrates proportional to shade at a global scale (**a**) and regional scale (**b**). **a** and **b** are the distribution overlays not including any *Felis* species. Distribution overlays are shown at the local level again including *Felis chaus* (**c**) and *Felis silvestris* (**d**). The geography of the local region employing Earthstar Geographics TerraColor imagery is shown as an inset in **c** and **d**. The red polygons (**a** and **b**) and white polygons (**c**) demarcate the overlapping distributions of the species based on the IUCN data

the basis of 16S rRNA or COI. In all pools except pools 3 and 8, the carp genera *Carassius* and *Cyprinus* were present and indistinguishable from each other on the basis of COI within 2% similarity.

Among the clusters that could not be identified to the species-level on the basis of the 16S rRNA locus are three members of the *Felis* genus, present in pool 4. Alignments to our 250 bp clusters placed *Felis catus* (the domestic cat), *Felis chaus* (the jungle cat), and *Felis silvestris* (the European wildcat) within 2% similarity of each other, varying between 98.08% identity and 96.13% identity. Given the geographic decisiveness of two of these species (*F. chaus* and *F. silvestris*), the implications of each species being the true match to our sequences were further investigated. The domestic cat, much like their human caretakers, has spread across every continent but Antarctica and thus does not impact the distributions discussed earlier. *F. chaus* overlaps with the other 16 species identified in Russia only (Fig. 1b), specifically at the Volga delta (Fig. 2c). Meanwhile, *F. silvestris* is present in many countries, therefore not altering the country presence/absence map shown in Fig. 1a, but within these countries is only found to overlap with all of the other species at the Danube delta in Romania (Fig. 2d).

Discussion

The wide variety of species captured by metabarcoding of bloodmeals of the confiscated leeches indicates that the specimens were wild caught, not captive bred. Leeches bred in supply companies are typically fed exclusively on blood of cows, pigs, or other common livestock. The International Medical Leech Center in Udelnaya, Russia, for example, exclusively feeds the leeches they breed on bovine blood from

certified meat-packing factories (see http://www.leech.ru/en/). Not one of the species confidently identified from the bloodmeals of these leeches were domesticated animals, suggesting that the leeches were wild caught. The only data indicative of a potentially domestic animal comes from a felid, which is an unlikely candidate for the feeding of captive leeches. This result would be more parsimoniously attributed to leech attachment during the feeding of one of the wild species matched, *F. chaus* or *F. silvestris*, both specifically known to hunt waterfowl (Heptner 1992).

The 11 species identified with detailed geographic distribution data overlapped in river deltas along the Black Sea, the Sea of Azov, and the Caspian Sea (Fig. 2a and b). However, because Kazakhstan, Moldova, and Ukraine are not known to harbor all 16 species identified (Fig. 1a), all 16 species likely overlap only in the Romanian delta on the Black Sea (the Danube delta), the Russian deltas on the Sea of Azov (between the Don delta and the Kuban delta), and the Russian delta on the Caspian Sea (the Volga delta). While the host species' ranges would not need to overlap if the leeches were collected from multiple sites, the complete overlap of distributions seen in Fig. 2 suggests that it is possible that they were all collected in one location. Moreover, these data are consistent with published reports regarding the habitat and diet of *H. verbana*. All vertebrate host species identified are extant in the regions where *H. verbana* is found. All of the countries and regions identified in Figs. 1 and 2 that are inhabited by all identified species are within *H. verbana*'s distribution (Utevsky et al. 2010; Saglam et al. 2016). Furthermore, this species of medicinal leech is known to feed on the blood of a wide variety of fish and amphibians, mammals that spend time in aquatic environments, and nesting waterfowl, with a penchant for congregating around these nests (Kutschera and Elliott 2014).

Many of the 16 host species are birds with fairly global or circumboreal distributions. Of the 11 species with full geographic distribution data (9 of which are birds), 3 particular species helped limit the potential region to the Eurasian land surrounding the Black and Caspian Seas. Unsurprisingly, 2 of these 3 were non-avian taxa: the toad *Pelobates fuscus* and the fish *Rutilus rutilus*. The remaining avian species, *Microcarbo pygmaeus*, is capable of flight, but is primarily sedentary and does not migrate long distances (Waldenström et al. 2017). Of these 3 species with ranges primarily limited to Eurasia, the particularly limited and patchy distribution of *Microcarbo* dramatically reduced the potential range. However, some of the species with global or circumboreal distributions, such as *Cygnus olor*, also had patchy distributions within Eurasia, which further defined the boundary. The accuracy of sequence identification for range-restricted species in particular (Table 2), as well as the accuracy of the IUCN geographic distributions for each of these species (Fig. 2), is central to the validity of the inferred leech origin. However, the fact that the regions identified as the potential sources of the wild leech acquisition are in Southeastern Russia is consistent with their having been brought to Canada on a flight from Russia.

Some host taxa identified were unable to be identified to the species-level. Limited resolution for the taxa is due to the clusters matching equally well to multiple species. In these cases, the amplified region of the loci is insufficiently variable to distinguish among these species. The genus *Anas* is known to have diversified recently and rapidly (Sun et al. 2017) to the extent that closely related species cannot be resolved via the metabarcoding loci used herein. Similarly, the frog genera *Pelophylax* and *Rana* are thought to have rapidly diversified at the beginning of the Paleogene (Feng et al. 2017), and the carp genera *Carassius* and *Cyprinus* are known to be difficult to distinguish with a single locus (Cherfas et al. 1994; Yang et al. 2010). While this could explain all of the instances where clusters were unable to be identified to the species-level, it is also possible that the sequence for the gene of the species detected is not available in the reference database. When this occurs, a sequence of the next closest species (usually of the same genus or family) will be matched, though at a lower percent identity. Thus, analysis of iDNA can only be as precise as the reference database allows (Siddall et al. 2019).

The only mammal DNA captured in the study was unable to be resolved to the species-level based on our 16S rRNA sequence data. The three candidate matches provided by GenBank, *Felis catus*, *Felis chaus*, and *Felis silvestris* were all within 2% identity matches of each other, suggesting that this portion of the 16S rRNA locus is unlikely to resolve these species of *Felis* to the species-level. However, these three species are not equally likely candidates for *Hirudo verbana* feeding due to their ecology and behavior. *F. chaus* inhabits wetlands where it has ample opportunity to come into contact with aquatic leeches (Gray et al. 2014). While *F. chaus* opportunistically preys on rodents, reptiles, amphibians, and even fish (Ogurlu et al. 2010), in southern Russia, the jungle cat is known to prey primarily on waterfowl, hunting them at their nesting sites along unfrozen rivers and marshes (Gray et al. 2016), where leeches also are known to feed on waterfowl. Meanwhile, *F. silvestris* primarily hunts small mammals in woodland areas; however, the population surrounding the Black Sea and the Azov Sea are specifically known for hunting waterfowl in these riparian and coastal areas (Heptner 1992). Thus, both *F. chaus* and *F. silvestris* likely spend more time in and near water than the average feral domestic cat. While this hardly rules out the domestic cat (a 2% closer match to our sequences) as the prey of the leeches, it is also important to consider the database from which these three potential species matches originated. NCBI's BLAST tool is known to exhibit biases towards species that are over-represented in the database (Cameron et al. 2007). GenBank's nt database, the database used in this study, typically has many more representative sequences from domestic species (and model organisms) than from their close non-domesticated

relatives, leading to a more complete view of the population diversity of domestic and model species than wild species. *F. catus*, the domestic cat, has tenfold more representation of 16S sequences in GenBank's nt database than *F. chaus*, potentially hundredfold or thousandfold more than *F. chaus* and *F. silvestris* if unannotated whole mitochondrial genomes and shotgun sequenced genomes are included. Thus, inherent bias towards the more represented domestic species is likely. If more individuals of *F. chaus* and *F. silvestris* were uploaded to the database, it is possible that this could more readily evince the species-level identity of our *Felis* sequences. An alternative possibility, although untestable with the current data, is that the *Felis* sequence represents domestic cat mitochondrial introgression into a "hybrid" wild cat with the behavioral characteristic of *F. chaus* or *F. silvestris*. Hybridization has been documented between domestic cats and *F. silvestris* (Pierpaoli et al. 2003). While it is unclear whether or not *F. chaus* interbreeds with domestic cats in the wild, they have been intentionally bred as pets and are even recognized as an exotic breed by The International Cat Association (Murphy 2015). Additional testing of nuclear markers would be required to confirm if our sequence comes from a hybrid felid.

Conservation implications

The species identified from the bloodmeals of the leeches in this study have pinpointed potential collection sites where all species overlap to the Volga delta and the Danube delta, as well as the coastal area region on the east side of the Sea of Azov. The Volga delta is the largest river delta in Europe and has grown by 24,000 km^2 since 1880 due to the rising water levels of the Caspian Sea (Coleman et al. 2008). In order to protect the many migratory species of waterfowl, raptors, and passerines that use the delta for staging, the region has been under protection through the Astrakhan Nature Reserve since 1919, making it among Russia's oldest nature preserves. Despite this protection, the region is under threat from algal blooms forming in the delta, fed by agricultural runoff in the Volga river, a problem that is amplified by the development of damming, reservoirs, and water diversion (Bhattacharyya 2013; Morley 2007). *Hirduo verbana*, as well as *Hirudo medicinalis*, have previously become locally extinct in areas with excessive algal blooming, such as the Srebarna Nature Reserve, a UNESCO World Heritage Site in Bulgaria (Michev et al. 1998). The other major delta identified in this study, the Danube delta, a UNESCO World Heritage Site in Romania, has also sustained eutrophication due to increased agriculture along the Danube river banks across Europe and erosion from riparian reed harvesting surrounding the delta itself (Galatchi and Tudor 2006). Ecosystem damage is a major threat to *H. medicinalis* and *H. verbana*, but overharvesting of medicinal leeches for blood-letting and traditional medicinal purposes has already led to their decline since the mid-1800s (Elliott and Kutschera 2011). While *H. verbana* is not yet catalogued by the IUCN, it is protected by CITES, which oversaw the exportation of 132,093 live leeches for medicinal purposes in 2018 alone (CITES Trade Database 2018). Our success using iDNA metabarcoding to determine the potential origin of the leeches serves as a model for the use of this strategy elsewhere.

In this study, the species identified from the bloodmeals of 240 leeches formed an assemblage of species so unique that there are only three deltas worldwide where all of them coexist. While this study did not directly aim to comment on the use of leeches for iDNA biodiversity surveys, it further lends credence to the potential of this system. With 16 species-level identifications from 240 leeches, approximately 1 species was identified per 15 leeches. The number of species identified per leech surveyed has the potential to be used as a proxy for biodiversity evenness. Furthermore, this study is the first to show that iDNA can be captured from aquatic leeches, such as *Hirudo*, and that a diverse array of organisms (fish, amphibians, birds, and mammals) can be detected in these bloodmeals. Of the vertebrate species that have historically inhabited the Astrakhan Nature Reserve in the Volga delta, 88 birds, 13 fish, and 7 mammals (including *Felis chaus*) are rare or endangered at that locality. Furthermore, while *Felis chaus* is assumed to inhabit the Volga delta region according to the IUCN 2016 Assessment, there has been no official siting or record of the species in the Astrakhan State Reserve (which lies within the Volga delta) since the 1980s, giving the species a threatened status in this region (Gray et al. 2016). Thus, an iDNA biodiversity survey conducted in the Volga delta region with a primer set better suited to species-level identification of felids may be in order to confirm the continued presence of this threatened species and other endangered fauna of the region.

Acknowledgments This work would not have been possible without the dedication of Wildlife Officer Mark McIntyre and the Wildlife Enforcement Branch of Environment and Climate Change Canada, for the seizure, care, and delivery of the leeches. Lily Berniker assisted in the processing and maintenance of leeches at the American Museum of Natural History.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Berry TE, Osterrieder SK, Murray DC, Coghlan ML, Richardson AJ, Grealy AK, Stat M, Bejder L, Bunce M (2017) DNA metabarcoding for diet analysis and biodiversity: a case study using the endangered Australian sea lion (*Neophoca cinerea*). Ecol Evol 7:5435–5453

Bhattacharyya R (2013) The Volga River. In: Howarth RW (ed) Biomes and ecosystems: an encyclopedia, Salem Press: A Division of EBSCO. Publishing Ipswich, Massachusetts (ISBN: 978-1-4298-3813-9)

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Calvignac-Spencer S, Leendertz FH, Gilbert MTP, Schubert G (2013) An invertebrate stomach's view on vertebrate ecology: certain invertebrates could be used as "vertebrate samplers" and deliver DNA-based information on many aspects of vertebrate ecology. BioEssays 35(11):1004–1013. https://doi.org/10.1002/bies.201300060

Camacho C et al (2009) BLAST+: architecture and applications. BMC Bioinformatics 10:1–9

Cameron M, Bernstein Y, Williams HE (2007) Clustered sequence representation for fast homology search. J Comput Biol 14:594–614. https://doi.org/10.1089/cmb.2007.R005

Caragiulo A, Dias-Freedman I, Clark JA, Rabinowitz S, Amato G (2014) Mitochondrial DNA sequence variation and phylogeography of Neotropic pumas (*Puma concolor*). Mitochondrial DNA 25:304–312

Cherfas NB, Gomelsky BI, Emelyanova OV, Recoubratsky AV (1994) Induced diploid gynogenesis and polyploidy in crucian carp, *Carassius auratus gibelio* (Bloch), x common carp, *Cyprinus carpio*, hybrids. Aquac Res 25:943–954

CITES Appendices I, II and III (2017) Available at: https://www.cites.org/eng/app/appendices.php

CITES Trade Database (2018) Available at: https://trade.cites.org/

Coleman JM, Huh OK, Braud D (2008) Wetland loss in world deltas. J Coast Res 1:1–14. https://doi.org/10.2112/05-0607.1

D'Amato ME, Alechine E, Cloete KW, Davison S, Corach D (2013) Where is the game? Wild meat products authentication in South Africa: a case study. Investig Genet 4(1):1–13. https://doi.org/10.1186/2041-2223-4-6

Deagle BE, Gales NJ, Evans K, Jarman SN, Robinson S, Trebilco R, Hindell MA (2007) Studying seabird diet through genetic analysis of faeces: a case study on macaroni penguins (*Eudyptes chrysolophus*). PLoS One 2(9):e831. https://doi.org/10.1371/journal.pone.0000831

Drinkwater R, Schnell IB, Bohmann K, Bernard H, Veron G, Clare E, Gilbert MTP, Rossiter SJ (2019) Using metabarcoding to compare the suitability of two blood-feeding leech species for sampling mammalian diversity in North Borneo. Mol Ecol Resour 19:105–117

Dybowski (1877) in GBIF Secretariat (2019). *Perccottus glenii.* GBIF Backbone Taxonomy. Checklist dataset. https://doi.org/10.15468/39omei accessed via GBIF.org on 2019-10-28

Eaton MJ, Meyers GL, Kolokotronis SO, Leslie MS, Martin AP, Amato G (2009) Barcoding bushmeat: molecular identification of Central African and South American harvested vertebrates. Conserv Genet 11(4):1389–1404. https://doi.org/10.1007/s10592-009-9967-0

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 26:2460–2461. https://doi.org/10.1093/bioinformatics/btq461

Elliott JM, Kutschera U (2011) Medicinal leeches: historical use, ecology. Genet Conserv Freshw Rev 4:21–41. https://doi.org/10.1608/frj-4.1.417

ESRI (2011) ArcGIS Online. Environmental Systems Research Institute, Redlands

Fahmy M, Ravelomanantsoa NAF, Youssef S, Hekkala E, Siddall M (2019) Biological inventory of Ranomafana National Park tetrapods using leech-derived iDNA. Eur J Wildl Res 65. https://doi.org/10.1007/s10344-019-1305-3

Fahmy M, Williams K, Tessler T, Weiskopf S, Hekkala E, Siddall M (2020) Assessing iDNA metabarcoding on multilocus pooled sampling strategies from transcontinental terrestrial leech collections. J Parasitol In press

Faircloth BC, Glenn TC (2012) Not all sequence tags are created equal: designing and validating sequence identification tags robust to indels. PLoS One 7:e42543. https://doi.org/10.1371/journal.pone.0042543

Feng YJ, Blackburn DC, Liang D, Hillis DM, Wake DB, Cannatella DC, Zhang P (2017) Phylogenomics reveals rapid, simultaneous diversification of three major clades of Gondwanan frogs at the Cretaceous–Paleogene boundary. Proc Natl Acad Sci U S A 114:E5864–E5870

Foran DR (2006) Relative degradation of nuclear and mitochondrial DNA: an experimental approach. J Forensic Sci 51(4):766–770. https://doi.org/10.1111/j.1556-4029.2006.00176.x

Froese R, Pauly D eds (2019) FishBase. World Wide Web electronic publication. www.fishbase.org, version (08/2019)

Galatchi LD, Tudor M (2006) Europe as a source of pollution – the main factor for the eutrophication of the Danube Delta and the Black Sea. Chem Intention Accident Glob Environ Threats. https://doi.org/10.1007/978-1-4020-5098-5_5

Gray TNE, Channa P, Chanrattanak P, Sovanna P (2014) The status of jungle cat and sympatric small cats in Cambodia's Eastern Plains. Cat News 8:19–23

Gray TN, Timmins RJ, Jathana D, Duckworth JW, Baral H, Mukherjee S (2016) *Felis chaus*, jungle cat. The IUCN red list of threatened species 2016: E.T8540A50651463, 8235. https://doi.org/10.2305/IUCN.UK.2016-2.RLTS.T8540A50651463.en

Hanya G, Morishima K, Koide T, Otani Y, Hongo S, Honda T, Okamura H, Higo Y, Hattori M, Kondo Y, Kurihara Y, Jin S, Otake A, Shiroisihi I, Takakuwa T, Yamamoto H, Suzuki H, Kajimura H, Hayakawa T, Suzuki-Hashido N, Nakano T (2019) Host selection of hematophagous leeches (*Haemadipsa japonica*): implications for iDNA studies. Ecol Res 34(6):842–855. https://doi.org/10.1111/1440-1703.12059

Heptner VG (1992) Mammals of the Soviet Union volume II part 2. Vysshaya ShkolaPublishers

Hyla orientalis Bedriaga, 1890 in GBIF Secretariat (2019) GBIF backbone taxonomy. Checklist dataset https://doi.org/10.15468/39omei accessed via GBIF.org on 2019-10-28

IUCN (2019) The IUCN red list of threatened species. Version 2019-2. https://www.iucnredlist.org

Jeffreys AJ, Brookfield JFY, Semeonoff R (1985) Positive identification of an immigration test-case using human DNA fingerprints. Nature 317(6040):818–819. https://doi.org/10.1038/317818a0

Keim A (1993) Studies on the host specificity of the medicinal blood leech *Hirudo medicinalis*. Parasitol Res 79(3):251–255. https://doi.org/10.1007/BF00931901

Kutschera U, Elliott JM (2014) The European medicinal leech *Hirudo medicinalis*: morphology and occurrence of an endangered species. Zoosystem Evol 90:271–280. https://doi.org/10.3897/zse.90.8715

Linacre A, Tobe SS (2011) An overview to the investigative approach to species testing in wildlife forensic science. Investig Genet 2(1):1–9. https://doi.org/10.1186/2041-2223-2-2

Michev T (1998) Biodiversity of the Srebarna biosphere reserve : checklist and bibliography. Bulgaria, Ministry of Environment and Waters. Sofia: Context & Pensoft, Bulgaria

Morley NJ (2007) Anthropogenic effects of reservoir construction on the parasite fauna of aquatic wildlife. EcoHealth 4:374–383. https://doi.org/10.1007/s10393-007-0130-4

Murphy W (2015) Genetic analysis of feline interspecies hybrids. Tufts' Canine and Feline Breeding and Genetics Conference

Nagy ZT, Sonet G, Glaw F, Vences M (2012) First large-scale DNA barcoding assessment of reptiles in the biodiversity hotspot of Madagascar, based on newly designed COI primers. PLoS One 7(3):e34506. https://doi.org/10.1371/journal.pone.0034506

Ogurlu I, Gundogdu E, Yildirim IC (2010) Population status of jungle cat (*Felis chaus*) in Egirdir lake, Turkey. J Environ Biol 31(1–2):179–183

Payne R, Sorenson M (2007) Integrative systematics at the species level: plumage, songs and molecular phylogeny of quailfinches Ortygospizan. Bull Br Ornithol Club 127:4–26

Pierpaoli M, Birò ZS, Herrmann M, Hupe K, Fernandes M, Ragni B, Szemethy L, Randi E (2003) Genetic distinction of wildcat (*Felis silvestris*) populations in Europe, and hybridization with domestic cats in Hungary. Mol Ecol 12:2585–2598. https://doi.org/10.1046/j.1365294X.2003.01939.x

Pilli E, Casamassima R, Vai S, Virgili A, Barni F, D'Errico G, Berti A, Lago G, Caramelli D (2014) Pet fur or fake fur? A forensic approach. Investig Genet 5(1):1–18. https://doi.org/10.1186/2041-2223-5-7

Poinar HN et al (2016) Molecular coproscopy: dung and diet of the extinct ground sloth *Nothrotheriops shastensis*. Science 281:402–406

Rocha R, Borda E, Andreone F, Rosa GM (2012) First reports of leech parasitism in Malagasy anurans. Comp Parasitol 79:352–356. https://doi.org/10.1654/4546.1

Roewer L (2013) DNA fingerprinting in forensics: past, present, future. Investig Genet 4(1):1. https://doi.org/10.1186/2041-2223-4-22

Saglam N, Saunders R, Lang SA, Shain DH (2016) A new species of *Hirudo* (Annelida: Hirudinidae): historical biogeography of Eurasian medicinal leeches. BMC Zool 1. https://doi.org/10.1186/s40850-016-0002-x

Sawyer RT (1986) Leech biology and behaviour: feeding biology, ecology, and systematics. Clarendon Press, Oxford

Sayers EW, Cavanaugh M, Clark K, Ostell J, Pruitt KD, Karsch-Mizrachi I (2019). GenBank. Nucleic Acids Res doi:https://doi.org/10.1093/nar/gky989, 47, D94, D99

Schnell IB, Thomsen PF, Wilkinson N, Rasmussen M, Jensen LRD, Willerslev E, Bertelsen MF, Gilbert MTP (2012) Screening mammal biodiversity using DNA from leeches. Curr Biol 22:R262–R263

Schnell IB et al (2015) iDNA from terrestrial haematophagous leeches as a wildlife surveying and monitoring tool - prospects, pitfalls and avenues to be developed. Front Zool 12:1–14

Siddall ME, Trontelj P, Utevsky SY, Nkamany M, Macdonald KS (2007) Diverse molecular data demonstrate that commercially available medicinal leeches are not *Hirudo medicinalis*. Proc R Soc B Biol Sci 274:1481–1487. https://doi.org/10.1098/rspb.2007.0248

Siddall ME, Barkdull M, Tessler M, Brugler MR, Borda E, Hekkala E (2019) Ideating IDNA: lessons and limitations from leeches in legacy collections. PLoS One 14(2):1–15. https://doi.org/10.1371/journal.pone.0212226

Speller CF, Nicholas GP, Yang DY (2011) Feather barbs as a good source of mtDNA for bird species identification in forensic wildlife investigations. Investig Genet 2(1):1–7. https://doi.org/10.1186/2041-2223-2-16

Sun Z et al (2017) Rapid and recent diversification patterns in Anseriformes birds: inferred from molecular phylogeny and diversification analyses. PLoS One 12:1–21

Tessler M, Weiskopf SR, Berniker L, Hersch R, Mccarthy KP, Yu DW, Siddall ME (2018) Bloodlines: mammals, leeches, and conservation in southern Asia. Syst Biodivers 16:488–496

Trontelj P, Utevsky SY (2012) Phylogeny and phylogeography of medicinal leeches (genus *Hirudo*): fast dispersal and shallow genetic structure. Mol Phylogenet Evol 63(2):475–485. https://doi.org/10.1016/j.ympev.2012.01.022

Utevsky S, Zagmajster M, Atemasov A, Zinenko O, Utevska O, Utevsky A, Trontelj P (2010) Distribution and status of medicinal leeches (genus *Hirudo*) in the western palaearctic: anthropogenic, ecological, or historical effects? Aquat Conserv Mar Freshw Ecosyst 20:198–210. https://doi.org/10.1002/aqc.1071

Vences M, Lyra ML, Perl RGB, Bletz MC, Stanković D, Lopes CM, Jarek M, Bhuju S, Geffers R, Haddad CFB, Steinfartz S (2016) Freshwater vertebrate metabarcoding on Illumina platforms using double-indexed primers of the mitochondrial 16S rRNA gene. Conserv Genet Resour 8:323–327

Waldenström J, Kuiken T, Wille M (2017) Narrative overview on wild bird migration in the context of highly pathogenic avian influenza incursion into the European Union. EFSA Support Publ 14(10). https://doi.org/10.2903/sp.efsa.2017.en-1283

Wilkin PJ (1989) The medicinal leech, *Hirudo medicinalis* (Hirudinea: Gnathobdellae), at Dungeness, Kent. Bot J Linn Soc 101(1):45–57. https://doi.org/10.1111/j.1095-8339.1989.tb00135.x

Yang L, Mayden RL, Sado T, He S, Saitoh K, Miya M (2010) Molecular phylogeny of the fishes traditionally referred to *Cyprinini sensu stricto* (Teleostei: Cypriniformes). Zool Scr 39:527–550