Received: 8 February 2018 | Accepted: 18 July 2018

DOI: 10.1111/jfb.13758

JOURNAL OF FISH BIOLOGY fsbi

REGULAR PAPER

Updated checklist and DNA barcode-based species delimitations reveal taxonomic uncertainties among freshwater fishes from the mid-north-eastern Caatinga ecoregion, north-eastern Brazil

Waldir M. Berbel-Filho[1] | Telton P. A. Ramos[1] | Uedson P. Jacobina[2,3] | Danielle J. G. Maia[2] | Rodrigo A. Torres[2] | Sergio M. Q. Lima[1]

[1]Laboratório de Ictiologia Sistemática e Evolutiva, Departamento de Botânica e Zoologia, Universidade Federal do Rio Grande do Norte, Natal, Brazil

[2]Laboratório de Genômica Evolutiva e Ambiental, Departamento de Zoologia, Universidade Federal de Pernambuco, Recife, Brazil

[3]Laboratório de Ictiologia e Conservação, Universidade Federal de Alagoas, Penedo, Brazil

Correspondence
Waldir M. Berbel-Filho, Laboratório de Ictiologia Sistemática e Evolutiva, Departamento de Botânica e Zoologia, Universidade Federal do Rio Grande do Norte, 59078-900, Natal, Rio Grande do Norte, Brazil.
Email: waldirmbf@gmail.com

Funding information
Cnpq, Grant/Award Numbers: 233161/2014-7, 306290/2015-4; CNPq/FAPERN , Grant/Award Number: 350674/2012-4; Conselho Nacional de Desenvolvimento Científico e Tecnológico/Ministério de Ciências e Tecnologia, Grant/Award Number: 552086/2011-8; FACEPE, Grant/Award Number: BCT-0125-2.04/15; PNPD-CAPES , Grant/Award Number: 1273846; Universidade Federal do Rio Grande do Norte

The mid-north-eastern Caatinga is a semiarid freshwater ecoregion in North-eastern Brazil that is dominated by temporary rivers and is currently classified as one of the least ichthyologically-known ecoregions in the world. The present study aimed to provide an updated checklist of mid-north-eastern Caatinga ecoregion (MNCE) freshwater fish species and evaluate their taxonomic identity using morphology, DNA barcoding and multiple species delimitation approaches. After reviewing published studies and ichthyological collections, 119 species were identified. Among these were 94 putatively valid native and 14 non-native species, five undescribed native species, four new records for the MNCE, 11 potential cases of misidentification and 14 species listed as *inquirenda*. Additionally, 252 individuals from 49 species were barcoded, revealing three potential taxonomic synonyms. The combined molecular approaches estimated a total of 91 native species, although a finalized species list for the MNCE awaits additional taxonomic revisions and field surveys. This study provides the most up-to-date species checklist for the MNCE and a molecular reference database for identifying MNCE fishes with DNA barcodes. Results highlight the need to integrate traditional taxonomy with molecular approaches to correctly identify species, especially in taxonomically problematic ecoregions such as the MNCE.

KEYWORDS

Atlantic forest, Caatinga, genetic clusters delineation, ichthyofaunal survey, molecular systematic

1 | INTRODUCTION

In tropical regions, freshwater biodiversity often correlates with the biodiversity of adjacent terrestrial ecosystems, thus ecoregional classifications are valuable for conservation (Abell *et al*., 2010). In a global analysis, Abell *et al*. (2008) proposed ecoregions based on combinations of topography, watersheds, differences and similarities among freshwater fish fauna. Analysing the patterns of freshwater fish diversity within ecoregions and drainages at a global scale, Lévêque *et al*. (2008) and Tedesco *et al*. (2017) highlighted knowledge gaps in north-eastern Brazil and the need for further studies in this region. North-eastern Brazil comprises parts or all of four freshwater ecoregions: Maranhão-Piauí, mid-north-eastern Caatinga (between the São Francisco and Parnaíba Basins), São Francisco and north-eastern Atlantic Forest (southwards to the São Francisco River basin) (Lima *et al*., 2017; Rosa *et al*., 2003).

The mid-north-eastern Caatinga ecoregion (MNCE) is contained primarily within the semiarid Caatinga biome (dry forest), which is

 © 2018 The Fisheries Society of the British Isles

characterized by a dry climate and impermeable soil that result in intermittent and seasonal hydrologic regimes in rivers and streams (Rocha *et al*., 2012; Rosa *et al*., 2003). Because of this hydrologic regime, the MNCE has traditionally been thought to have few freshwater fish species, making it a historically neglected ecoregion for ichthyological studies (Rosa & Groth, 2004). As a result, the freshwater fish diversity of the MNCE remains poorly understood (Langeani *et al*., 2009) despite recent studies of Caatinga fishes (Lima *et al*., 2017). The systematic understanding of freshwater fishes from the MNCE is mainly based on brief descriptions, with an abundance of mistaken identifications and confusing taxonomy (*e.g.* some taxa are considered synonyms of described species) (Barros *et al*., 2011; Lima *et al*. 2017; Ramos *et al*. 2016; Rosa *et al*., 2003). Among these are 24 species described by Fowler (1915, 1941), many of which lack diagnostic features, with some probably having inaccurate locality data (Lima *et al*., 2017). The need for an updated list of the MNCE's freshwater fish fauna is especially urgent now because the four main drainages of the ecoregion (Jaguaribe, Piranhas-Açu, Apodi Mossoró and Paraíba do Norte Basins) will be affected by the ongoing São Francisco interbasin water transfer project (SFR-IWT), which is now under construction. This project may result in the introduction of exotic species via artificial canals and other major environmental changes (Berbel-Filho *et al*., 2016, Silva *et al*., 2017).

DNA barcodes have been especially useful in historically neglected geographical regions (Hubert *et al*., 2008; Ward *et al*., 2005). Indeed, this method is frequently used to estimate the inter and intraspecific diversity of freshwater fishes around the world (Carvalho *et al*., 2011; Díaz *et al*., 2016; Hubert *et al*., 2008; Pereira *et al*., 2011; Ramirez *et al*., 2017; Ward *et al*., 2005) highlighting groups that require taxonomic revision (Armstrong & Ball, 2005; Díaz *et al*., 2016; Hajibabaei *et al*., 2007).

DNA barcodes and other methods of single-locus species delimitation, have been largely used as a tool for species identification (Hajibabaei *et al*., 2007; Ramirez *et al*., 2017). Despite working as a useful starting point to identify molecular operational taxonomic units (mOTU), single-locus species delimitations have their constraints (*e.g.* incomplete lineage sorting and introgression) which can confuse or incorrectly delineate evolutionary lineages (Kekkoken & Hebert, 2014). Preferably, approaches that integrate different assumptions (morphological taxonomy, molecular analyses, behavioural and ecological traits) should be used to increase the accuracy of species delimitations (Carstens *et al*., 2013). However, given globally increasing rates of biodiversity loss, faster access to biodiversity information and more precise tools are needed for identification and conservation purposes (Smith *et al*., 2005).

New analyses have been used to increase the robustness of mOTUs identified by single-locus analyses, such as DNA barcoding based on portions of the mitochondrial cytochrome oxidase I gene (*coI*). Coalescent-based methods such as the generalized mixed yule-coalescent (GMYC) method use maximum-likelihood and an ultrametric gene tree to model transition points between inter and intraspecific branches (Fujisawa & Barraclough, 2013). The poisson tree process (PTP) method (Zhang *et al*., 2013) also looks for branching transitions, but PTP includes an expected number of mutations. Furthermore, distance-based methods, such automatic barcode gap discovery (ABGD), apply clustering algorithms to analyse partitions based on the genetic distance among groups of individuals. Both coalescent-based methods and distance-based methods have been extensively used in molecular systematics (Blair & Bryson, 2017; Kennoken & Hebert, 2014).

Given the extensive taxonomic uncertainties and frequently cryptic diversity among Neotropical freshwater fishes (Carvalho *et al*., 2011, 2012; Díaz *et al*., 2016; Pereira *et al*., 2011; Torres & Ribeiro, 2009), the present study aimed to provide an updated checklist of the poorly known MNCE ichthyofauna based on recent extensive surveys and comparative material in fish collections. Additionaly, we aimed to evaluate the taxonomic identity of MNCE fishes using DNA barcoding and multiple species delimitation approaches.

2 | MATERIALS AND METHODS

2.1 | Species records and taxonomic validation

Qualitative species records were originally obtained from the two largest collections of fishes from the MNCE: those located at the Universidade Federal da Paraíba and the Universidade Federal do Rio Grande do Norte. This material included many samples collected in recent years from drainages along the MNCE, including type localities whenever possible (Table 1). Specimens were morphologically identified to the lowest taxonomic level possible based on meristic and morphometric data provided by identifcation keys, systematic reviews, original descriptions and assistance of experts following the same criteria and nomenclature adopted in Silva *et al*. (2017) and Lima *et al*. (2017). Distribution data and species compositions were confirmed by crossreferencing several online databases including Specieslink (www.splink.org.br), NEODAT II (www.mnrj.ufrj.br/search1p.htm) and Global Biodiversity Information Facility (GBIF; gbif.org). Specialized taxonomic literature was also consulted (Britzke *et al*., 2016; Buckup *et al*., 2007; Costa, 2008; Costa & Vono, 2009; Fowler, 1941; Gurgel-Lourenço *et al*., 2013; Jerep & Malabarba, 2014; Novaes *et al*., 2013; Ramos *et al*., 2005, 2013; Reis *et al*., 2003; Rosa & Groth, 2004; Zawadzki *et al*., 2017) along with systematic compilations of MNCE freshwater fishes (Costa et al., 2017a, 2017b; Lima *et al*., 2017; Nascimento *et al*., 2014; Paiva *et al*., 2014; Rodrigues-Filho *et al*., 2016; Rosa *et al*., 2003; Silva et al., 2014, 2017).

Taxonomic validation of species was done using Eschmeyer & Fong (2017). All species listed herein correspond to at least one voucher in the following institutions: The Academy of Natural Sciences (ANSP), U.S.A., Museu Nacional da Universidade Federal do Rio de Janeiro (MNRJ), Brazil, Museu de Zoologia da Universidade de São Paulo (MZUSP), Brazil, Universidade Federal da Paraíba (UFPB), Brazil and Universidade Federal do Rio Grande do Norte (UFRN), Brazil. If a species was only known by its type material and its identity was uncertain due to a poor diagnosis or unavailability of topotypes, it was classified as *inquirenda* to indicate the need for further taxonomic review to confirm its taxonomic validity, following the definition of Sigovini *et al*. (2016). Surveys of regional fish collections, recently sampled fish specimens, including topotypes (Table 1), were compared with digitized images of the holotypes and additional material

TABLE 1 Updated list of freshwater fish species of the mid-north-eastern Caatinga ecoregion (MNCE)

	Status	Voucher	Reference	Barcode
Order Osteoglossiformes				
Family Arapaimidae				
Arapaima gigas	NNA	–	LI	
Order Characiformes				
Family Parodontidae				
Apareiodon davisi	E	UFRN 0452	RO, LI	
Family Curimatidae				
Curimatella lepidura		UFRN 1833	RO, LI	
Psectrogaster rhomboides		UFRN 2252	RO, LI	X
Psectrogaster saguiru	E	MNRJ 9147	RO, LI	
Steindachnerina notonota		UFRN 0357	RO, LI	X
Family Prochilodontidae				
Prochilodus brevis		UFRN 0594	RO LI	X
Family Anostomidae				
Leporinus melanopleura	MIS		RO	
Leporinus piau		UFRN 0755	RO, LI	X
Leporinus taeniatus		UFRN 1836	LI	X
Megaleporinus obtusidens	NNA		RO	
Schizodon fasciatus		UFRN 3218	RO, LI	
Family Bryconidae				
Salminus hilarii	NNA	ANSP 69608	RO	
Family Erythrinidae				
Erythrinus erythrinus		UFRN 0082	LI	X
Hoplerythrinus unitaeniatus		UFPB 0351	RO, LI	
Hoplias aff. *malabaricus*		UFRN 1223	RO, LI	X
Family Serrasalmidae				
Colossoma macropomum	NNA	UFRN 1710	RO, LI	
Metynnis lippincottianus		UFRN 1036	LI	X
Myleus micans	NNA	UFPB 10319	LI	
Pygocentrus nattereri		UFPB 4457	RO, LI	
Pygocentrus piraya	MIS		RO	
Pristobrycon striolatus	MIS		RO	
Serrasalmus brandtii		UFPB 4456	RO, LI	X
Serrasalmus rhombeus		UFRN 1498	RO, LI	X
Serrasalmus spilopleura	NR	UFRN 1858	TS	
Family Hemiodontidae				
Hemiodus parnaguae	MIS		RO	
Family Characidae				
Astyanax aff. *bimaculatus*		UFRN 0123	RO, LI	X
Astyanax aff. *fasciatus*		UFRN 1282	RO, LI	
Cheirodon jaguaribensis	E	UFRN 1632	LI	X
Cheirodon macropterus	INQ	ANSP 69531	F1	
Compsura heterura		UFRN 1211	RO, LI	
Ctenobrycon spilurus		UFRN 1298	RO	X
Hemigrammus brevis		UFPB 4130	RO	
Hemigrammus guyanensis		UFRN 2533	LI	X
Hemigrammus marginatus		UFRN 1699	RO, LI	X
Hemigrammus rodwayi		UFRN 2815	LI	X
Hemigrammus unilineatus		UFRN 1467	LA	
Hyphessobrycon bentosi		UFRN 2827	LI	X
Hyphessobrycon iheringi	INQ	ANSP 69579	F1, RO	

TABLE 1 (Continued)

	Status	Voucher	Reference	Barcode
Hyphessobrycon piabinhas	INQ	ANSP 69580	F1, LI	
Hyphessobrycon parvellus		UFRN 2635	LI	
Moenkhausia costae		UFRN 1623	RO, LI	X
Moenkhausia intermedia		UFRN 2557	LI	
Moenkhausia lepidura	MIS		RO	X
Phenacogaster calverti		UFPB 7053	RO, LI	
Roeboides microlepis	MIS		RO	
Serrapinnus heterodon		UFRN 1304	RO, LI	X
Serrapinnus piaba		UFRN 2563	RO, LI	
Serrapinnus potiguar	E	UFRN 3419	LI	
Nanocheirodon insignis	MIS		RO	
Tetragonopterus argenteus		UFRN 1831	RO, LI	
Family Lebiasinidae				
Nannostomus beckfordi		UFRN 1913	LA, LI	
Family Triportheidae				
Triportheus signatus		UFRN 2280	RO, LI	X
Family Crenuchidae				
Characidium bimaculatum	E	UFRN 2197	RO, LI	X
Order Siluriformes				
Family Auchenipteridae				
Trachelyopterus cratensis	INQ	MNRJ 0947	MR, LI	
Trachelyopterus galeatus		UFRN 3430	RO, LI	X
Trachelyopterus striatulus	MIS		RO	
Family Heptapteridae				
Pimelodella dorseyi	E	UFRN 1808*	RO, LI	X
Pimelodella enochi	E	UFRN 1369*	RO, LI	X
Pimelodella gracilis	MIS		RO	
Pimelodella papariae	INQ	ANSP 69387	F1, RO, LI	
Pimelodella witmeri	INQ	ANSP 69383	F1, RO, LI	X
Pimelodella wolfi	INQ	ANSP 69388	F1, RO, LI	
Rhamdia quelen		UFRN 0633	RO, LI	X
Family Callichthyidae				
Aspidoras carvalhoi	E	MNRJ 5230	RO, LI	
Aspidoras depinnai	E	MZUSP 56214	RO, LI	
Aspidoras rochai	E	UFRN 1879	RO, LI	
Aspidoras menezesi	E	UFRN 3745*	RO, LI	X
Aspidoras spilotus	E	UFRN 1580*	RO, LI	
Callichthys callichthys		UFRN 2607	RO, LI	
Corydoras sp.	UND	UFRN 1604	LI	
Megalechis personata	MIS		RO	
Megalechis thoracata		UFRN 2363	RO, LI	
Family Loricariidae				
Hypostomus carvalhoi	INQ	UFPB 1810*	MR, RO, LI	
Hypostomus jaguribensis	INQ	UFRN 1802	F0, RO, LI	X
Hypostomus nudiventris	INQ	UFPB 7697*	F1, RO, LI	
Hypostomus papariae	INQ	UFRN 2421	F1, RO, LI	X
Hypostomus pusarum		UFRN 0293	RO, LI	X
Hypostomus salgadae	INQ	ANSP 69440*	F1, LI	
Hypostomus sertanejo	E	UFRN 1840	LI	X
Loricariichthys derbyi		UFRN 1837	RO, LI	X
Loricariichthys sp.	UND	UFRN 0586	LI	X

TABLE 1 (Continued)

	Status	Voucher	Reference	Barcode
Parotocinclus cearensis		UFRN 1132*	RO, LI	X
Parotocinclus cesarpintoi	E	UFRN 1149*	RO, LI	
Parotocinclus haroldoi	NR	UFRN 1294	TS	
Parotocinclus jumbo		UFRN 1587*	LI	
Parotocinclus seridoensis	E	UFRN 1588	LI	X
Parotocinclus sp. 1	UND	UFRN 2259	LI	X
Parotocinclus sp. 2	UND	UFRN 0428	LI	
Parotocinclus spilosoma	E	UFRN 1584*	RO, LI	
Parotocinclus spilurus	E	UFRN 1252*	RO, LI	X
Pseudancistrus genisetiger	E	UFRN 1477*	RO, LI	X
Pseudancistrus papariae	INQ	ANSP 69442	F1, RO, LI	
Aphanotorulus gomesi	INQ	ANSP 69409	F2, LI	
Order Gymnotiformes				
Family Sternopygidae				
Eigenmannia virescens		UFPB 0344	RO, LI	
Family Gymnotidae				
Gymnotus carapo		UFRN 1084	RO, LI	X
Order Cyprinodontiformes				
Family Cynolebiidae				
Anablepsoides cearensis	E	UFRN 2657*	LI	X
Cynolebias microphthalmus	E	MZUSP 42312	RO, LI	
Hypsolebias antenori	E	UFRN 3533	RO, LI	
Hypsolebias longignatus	E	UFRJ 6614	LI	
Hypsolebias martinsi	E	ZUEC 10791	LI	
Kryptolebias hermaphroditus	NR	UFRN 2541	TS	X
Family Poeciliidae				
Poecilia reticulata	NNA	UFRN 2195	RO, LI	
Poecilia sarrafae		UFRN 2575	LI	X
Poecilia vivipara		UFRN 0289	RO, LI	X
Xiphophorus cf. *helleri*	NNA	UFRN 1259	LI	
Order Synbranchiformes				
Family Synbranchidae				
Synbranchus sp.	UND	UFRN 1684	LI	X
Order Cichliformes				
Family Cichlidae				
Astronotus ocellatus	NNA	UFRN 1807	RO, LI	
Cichla kelberi	NNA	UFRN 0221	LI	
Cichla monoculus	NNA	UFPB 4417	RO, LI	
Cichla ocellaris	NNA	UFPB 2917	RO, LI	
Cichlasoma orientale		UFRN 1012*	RO, LI	X
Cichlasoma sanctifranciscense		UFRN 1715	LI	X
Coptodon rendalli	MIS		LI	
Crenicichla menezesi		UFRN 1522	RO, LI	X
Geophagus brasiliensis		UFRN 0719	RO, LI	X
Laetacara curviceps	NNA, NR	UFRN 0566	TS	
Oreochromis niloticus	NNA	UFRN 0674	RO, LI	
Parachromis managuensis	NNA	UFRN 1971	LI	

E, endemic species; INQ, species *inquirenda* (doubtful); MIS, probable misidentification; NNA, non-native species; NR, new records for the MNCE; UND, probable undescribed species. References: MR, Miranda–Ribeiro (1937); F0, Fowler (1915); F1, Fowler (1941); F2, Fowler (1942); RO, Rosa *et al*. (2003); LA, Langeani *et al*. (2009); LI Lima *et al*. (2017); TS, this study. * Voucher numbers indicate specimens collected from the type locality (topotypes)

recorded by Starks (1913) and Fowler (1915, 1941, 1942), checking for updated identifications in the online database from Stanford University (SU) at the California Academy of Science (www.researcharchive.calacademy.org/research/Ichthyology/collection/index.asp) and from ANSP (www.clade.ansp.org/ichthyology/FTIP/search.php?mode=search&scope=Collection&contains=&contains_loc=&tbl=Specimens&Submit=Search+ANSP+Fish+Collection&gallery=ImageGallery). This study is part of a regional collaborative initiative to study the diversity and conservation of freshwater fishes from north-eastern Brazil, including the MNCE ichthyofauna (Costa *et al*., 2017a; Costa *et al*., 2017b; Lima *et al*. 2017; Lira *et al*., 2015; Paiva *et al*., 2014; Ramos *et al*., 2013, 2016; Silva *et al*., 2014, 2017; Teixeira *et al*., 2017; Zawadzki *et al*., 2017).

Additionally, some valid species listed by other authors as being present in the MNCE, but only found on their list without corresponding voucher material in regional collections, were indicated as potential misidentifications within our list until their existence is verified. New records were either reported to new, recently described species or additional records of valid species compared with previous lists (Lima *et al*., 2017; Rosa *et al*., 2003). These lists only describe fishes of the Caatinga and do not include records from species of the eastern-most strip of Atlantic Forest from Rio Grande do Norte to Alagoas States. In addition, these lists do not provide voucher material, making it difficult to check and update the dataset. To verify the MNCE endemicity of species, the main literature and databases checked were Reis *et al*. (2003), Rosa *et al*. (2003), Buckup *et al*. (2007), Eschmeyer & Fong (2017) and Lima *et al*. (2017).

2.2 | DNA extraction, amplification and sequencing

Tissue samples were obtained from fin-clips or muscle taken from the right side of specimens. Each tissue sample was recorded using a unique numerical code following the prefix TIUFRN and each corresponds to a formalin-fixed voucher specimen deposited either at the UFRN or UFPB fish collection. To minimize taxonomic confusion, samples were collected whenever possible from specimens with the most precise description of the type locality for each species (topotypes), usually matching the river basin and municipality. These samples were obtained from recent collections (2011–2016) and were mostly from species-rich genera (*e.g. Aspidoras* Ihering 1907, *Leporinus* Agassiz 1829, *Hemigrammus* Gill 1858, *Hypostomus* Lacépède 1803, *Parotocinclus* Eigenmann & Eigenmann 1889 and *Pimelodella* Eigenmann & Eigenmann 1888) and species described by Fowler (1915, 1941, 1942) from within this ecoregion. Whenever possible, samples were selected according to a species' geographic range, spreading sampling across different river basins to maximize intraspecific genetic diversity within the MNCE. Fish collection and euthanasia using eugenol (30 ml of a 10% eugenol alcohol solution in 970 ml of water) was done under permits 30532-1/2011 and 32656-1/2012, issued by ICMBio/SISBIO (Instituto Chico Mendes de Conservação da Biodiversidade /Sistema de Autorização e Informação em Biodiversidade).

DNeasy tissue kits (QIAGEN; www.qiagen.com) were used to extract genomic DNA, following the manufacturer's protocol. Polymerase chain reaction (PCR) was conducted using the primer combinations: *FishF1*–5′TCAACCAACCACAAAGACATTGGCAC3′, *FishF2*–5′TCGACTAATCATAAAGATATCGGCAC3′, *FishR1*–5′TAGACTTCTGGG TGGCCAAAGAATCA3′ and *FishR2*–5′ACTTCAGGGTG ACCGAAGAATCAGAA3′ (Ward *et al*., 2005). Twenty-five μl of PCR product was obtained using two PCR parameters, the first of which included 12.5 μl of 2X Taq master mix Vivantis (www.vivantechnologies.com), 10–30 ng μl^{-1} of DNA template, 0.5 μl (10 mM) of each primer and 9.5 μL of ultrapure water. These reactions were done using the following thermal regime: 95 °C for 2 min, 94 °C for 30 s, 57 °C for 2 min, 72 °C for 2 min (35×), 72 °C for 7 min as a final extension step. The second set of PCR parameters included 12.5 μl of 2X Taq master mix Vivantis, 10–30 ng μl^{-1} of DNA template, 0.3 μl (5 mM) of each primer and 9.9 μl of ultrapure water. The thermal regime for these reactions was: 95 °C for 5 min, 94 °C for 30 s, 50 °C for 30 s, 72 °C for 70 s (35×), 72 ° C for 7 min, 20 ° C for 2 min. PCR products were checked via 1.8% agarose gels using GelRed (Uniscience; www.uniscience.com) and then purified using ExoSap-IT (Affimetrix; www.thermofisher.com). Sequencing was done using an ABI 3130 sequencer (Applied Biosystems; www.appliedbiosystems.com). All sequences obtained were deposited in Genbank (Supporting Information Table S1).

2.3 | Phylogenetic analysis and species delimitation methods

A total of 49 taxonomically identified species (based on morphological identification and geographical distribution) were successfully amplified. Eletropherograms were checked and edited using Geneious 7.1 (Kearse *et al*., 2012). Alignment was done in MEGA 6.0 (Tamura *et al*., 2013), generating a final dataset with *coI* fragments of 465 bp. The edited sequences were then blasted (BLASTn) against the NCBI database to confirm their identity and detect putative pseudogenes. A Bayesian phylogenetic reconstruction was run in BEAST 1.7 (Drummond *et al*., 2012) using the haplotype data with substitution models partitioned by codon position (where each codon evolves under a generalized time reversible model with rate heterogeneity across sites being modelled by a gamma distribution; GTR + Γ). Substitution rates, rate heterogeneity and base frequencies across codons were unlinked (Yang, 1996) in BEAST. An uncorrelated relaxed lognormal model with estimated rate was used, with ucld.mean parameter set and uniform distribution (0 and 10 as lower and upper boundaries). Remaining parameters were set as default. Length of the Monte-Carlo Markov chain (MCMC) was set to 10 million generations with sampling every thousandth generation. Effective sample size (ESS) values > 200 were determined using Tracer 1.5 (Rambaut, 2009). The initial 2000 trees were discarded as burn-in and a consensus tree was constructed using TreeAnnotator 1.50.

Four single-locus species-delimitation analyses were performed: single and multiple-threshold generalized mixed yule-coalescent analyses (sGMYC and mGMYC; Fujisawa & Barraclough, 2013), Bayesian implementation of Poisson tree processes (bPTP; Zhang *et al*., 2013) and automatic barcode gap discovery (ABGD; Puillandre *et al*., 2012). An ultrametric tree generated from haplotypes used in the Bayesian phylogenetic analysis was used as an input file for both sGMYC and mGMYC. Previous studies have shown that these methods are consistent under different tree assumptions (*e.g.* priors and molecular rates;

Tavalera *et al*., 2013). These two analyses were conducted in R (www.r-project.org), using the package splits (Ezard *et al*., 2009). The bPTP analyses were performed using the online server (www.species.h-its.org) by transforming the Bayesian phylogenetic reconstruction derived from BEAST into a phylogram using the R package phangorn 2.2 (Schliep, 2010). The analysis was programed to run for 500,000 generations with sampling every 500th generation and the first 10% of results discarded as burn-in. Convergence was visualized using the log-likelihood plots of MCMC interactions. The ABGD distance-based analysis was done using a gap width value of 1.0 for all distances available (p-distance, Kimura-2-parameter and Jukes-Cantor distances). Concordance between OTU delimitation methods was evaluated by comparing cluster composition across all five methods.

3 | RESULTS

3.1 | Species records and taxonomic validation

After reviewing literature and examining the main regional ichthyological collections and online databases containing information about freshwater fishes from the MNCE, 119 species belonging to 65 genera, 23 families and 7 orders were listed. Among these, 14 were non-native and 11 represented cases of misidentification (*i.e.* species listed in published studies with no corresponding voucher specimens), resulting in a total 94 native species. From these nominal native species, 36 had their type locality within the MNCE, of which 16 were compared with sampled topotypes for morphology and 9 were barcoded (Table 1). Among the 14 cases of *species inquirenda*, molecular data indicated at least three potential synonyms (*Pimelodella witmeri* Fowler 1941 as a junior synonym of *Pimelodella dorseyi* Fowler 1941 and *Hypostomus jaguribensis* (Fowler 1915) and *Hypostomus papariae* as Fowler 1941 junior synonyms of *Hypostomus pusarum* (Starks 1913); see below), resulting in a conservative estimate of 91 native species. These 91 species belonged to 48 genera, 19 families and 6 orders, of which 23 species (28%) are endemic, four represent new records for the MNCE and five of those are putatively undescribed: *Corydoras* sp., *Loricariichthys* sp., *Parotocinclus* sp.1, *Parotocinclus* sp. 2 and *Synbranchus* sp. (Table 1).

3.2 | DNA barcode and species delimitation

A total of 252 sequences from 49 species (based on morphological identification and geographical distribution of putatively valid species), belonging to 33 genera, 17 families and six orders, were barcoded at 52 localities across 18 river basins (Figure 1, Table 2 and Supporting Information Table S1). This represents 52.1% of the potential 94 native species listed for MNCE (Figure 1 and Table 1).

The average K2P distances among specimens within species was 0.18%, ranging from zero (*Ctenobrycon spilurus* Valenciennes 1850), *Metynnis lippincottianus* (Cope 1870), *Leporinus taeniatus* Lütken 1875, *Steindachnerina notonota* (Miranda Ribeiro 1937), *Loricariichthys derbyi* Fowler 1915, *Parotocinclus spilurus* (Fowler 1941), *Parotocinclus cearensis* Garavello 1977, *Aspidoras menezesi* Nijssen & Isbrücker 1976, *P. dorseyi*, *Rhamdia quelen* (Quoy & Gaimard 1824), *Trachelyopterus galeatus* (L. 1766), *Kryptolebias hermaphroditus* Costa 2011, *Poecilia sarrafae* Bragança & Costa 2011, *Characidium bimaculatum* Fowler 1941, *Cheirodon jaguaribensis* Fowler 1941 and *Hemigrammus guyanensis* Géry 1959) to 1.2% (*Serrapinnus heterodon* (Eigenmann 1915)). The average interspecific K2P distance within genera was 6.70%, or about 37 times greater than the within species average (0.18%). Average K2P distances continued to increase across higher taxonomic levels (Table 2).

A majority of the species were discriminated based on their barcode sequences. However, *P. dorseyi* (Rio Salgado at Icó) and *P. witmeri* (Rio Jaguaribe at Orós), which were both sampled at their type localities in the Jaguaribe River basin, shared the same haplotype. Also, *H. pusarum* from its type locality (the Ceará-Mirim River basin) shared a haplotype with *H. papariae* within a clade that also included *H. jaguribensis*. Mean genetic distance within this clade was 0.3%.

From the 49 species barcoded, three out of four species delimitation methods (sGMYC, bPTP, ABGD) indicated 44 as the best number of partitions, while mGMYC showed 47 as the most likely number of OTUs (splitting *Cichlasoma orientale* Kullander 1983 and *Astyanax* aff. *bimaculatus* (L. 1758) into three and two mOTUs, respectively; Figure 2 and Table 3). Overall, all four methods similarly discriminated the taxonomically identified species, with few exceptions. The *Hypostomus pusarum* cluster (comprising *H. pusarum*, *H. jaguribensis* and *H. papariae*) represented only one genetic cluster. Additionally, the *Pimelodella dorseyi* cluster (comprising *P. dorseyi* and *P. witmeri*) represented only one OTU. Finally, two cases of morphologically distinguishable species merged into the same genetic cluster in the genus *Parotocinclus*. All methods indicated that *Parotocinclus cearensis* Garavello 1977 and *Parotocinclus* sp. 1 belong to the same mOTU. The same result was found for *P. spilurus* and *Parotocinclus seridoensis* Ramos *et al*., 2013 (Figure 2).

4 | DISCUSSION

The present study provided an updated checklist of freshwater fish species from the mid-north-eastern Caatinga ecoregion, as well as an evaluation of taxonomic consistency across its species and drainages using DNA barcodes and species delimitation methods. Our list included 119 nominal species. Of these, 14 species were classified as non-native and 11 represented potential cases of misidentification, resulting in 94 native species. However, this number still includes 14 cases of species *inquirenda* (Table 1), which potentially represent non-valid species, as corroborated by the barcode and species delimitation results. These results indicate that at least three of these species are potential synonyms, which would decrease native species richness to a more conservative estimate of 91 species. Of these, four species represent new records for the MNCE (including the non-native species *Laetacara curviceps* Ahl 1923) and five species appear to be undescribed (Table 1). Although our native species richness estimate is similar to the 88 species proposed by Albert & Reis (2011), the composition of our list is both different and more taxonomically robust, as the present list is based on extensive surveys that take into consideration both morphological and molecular evidence. In terms of endemism, the percentage of endemic species in our list (28.0%) is

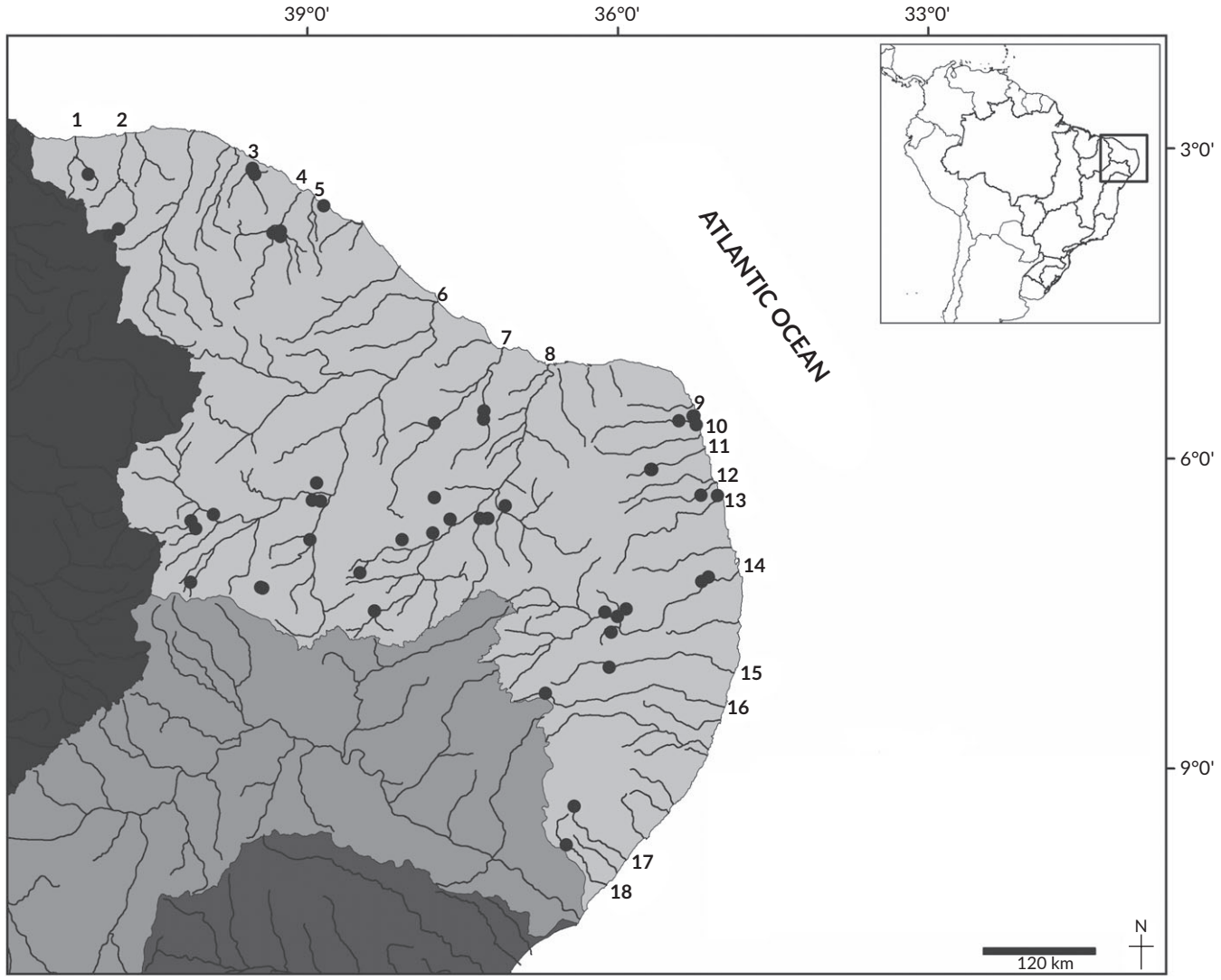

FIGURE 1 Map of the sampling localities for 252 fish specimens from the mid-north-eastern Caatinga ecoregion that were barcoded in this study. 1, Acaraú River basin; 2, Coreaú River basin; 3, Mundaú River basin; 4, Curu River basin; 5, Cauípe River basin; 6, Jaguaribe River basin; 7, Apodi-Mossoró River basin; 8, Piranhas-Açu River basin; 9, Pratagí River basin; 10, Ceará-Mirim River basin; 11, Trairí River basin; 12, Catu River basin; 13, Curimataú River basin; 14, Paraíba do Norte River basin; 15, Capibaribe River basin; 16, Ipojuca River basin; 17, Coruripe River basin; 18, Paraíba do Meio River basin. (●) Sampling points, (■) Maranhão-piaui ecoregion, (■) Mid-north-eastern Caatinga ecoregion, (■) São Francisco ecoregion, and (■) North-eastern Mata Atlântica ecoregion

TABLE 2 Summary of Kimura-2-parameter (K2P) genetic distances of freshwater fish taxa from the mid-north-eastern Caatinga ecoregion at various taxonomic levels

		K2P distance (%)			
Category	**Taxa**	**Minimum**	**Mean**	**Maximum**	**S.E.**
Within species	41	0	0.18	1.20	0.02
Within genus	10	2.37	6.70	15.30	0.09
Within families	8	6.54	12.04	17.76	1.28
Within orders	3	15.99	18.17	20.16	3.95

lower than those found in previous studies, (40.7% in Rosa *et al*., 2003; 43.1% in Albert & Reis, 2011). This discrepancy could have resulted from an improved understanding of the geographic distribution of fishes of the MNCE and adjacent ecoregions, mainly due to more extensive recent field surveys (Ramos *et al*., 2014).

From our four new records for the MNCE, three (*Kryptolebias hermaphroditus* Costa 2011, *L. curviceps* Include (Ahl 1923), *Parotocinclus haroldoi* Garavello 1988 and *Serrasalmus spilopleura* Kner 1858) are restricted to a single or few records in restricted areas of the MNCE, suggesting a marginal occurrence (*P. haroldoi* in the westernmost Timonha River basin; Rodrigues-Filho *et al*., 2016) or even anthropogenic introdutions (*S. spilopleura* and *L. curviceps* in a single small coastal basins of Rio Grande do Norte State, each, both in the Atlantic Forest area). The mangrove killifish *K. hermaphroditus* is broadly distributed within mangrove microhabitas along the Brazilian coast (Lira *et al*., 2015; Tatarenkov *et al*. 2017).

Human introduction might also have influenced the current distribution of Amazonian species, such as *Hemigrammus guyanensis* Géry 1959, *Hemigrammus rodwayi* Durbin 1909, *Hyphessobrycon bentosi* Durbin 1908, *Nannostomus beckfordi* Günther 1872, mainly due to these species' importance to the ornamental aquarium fish trade

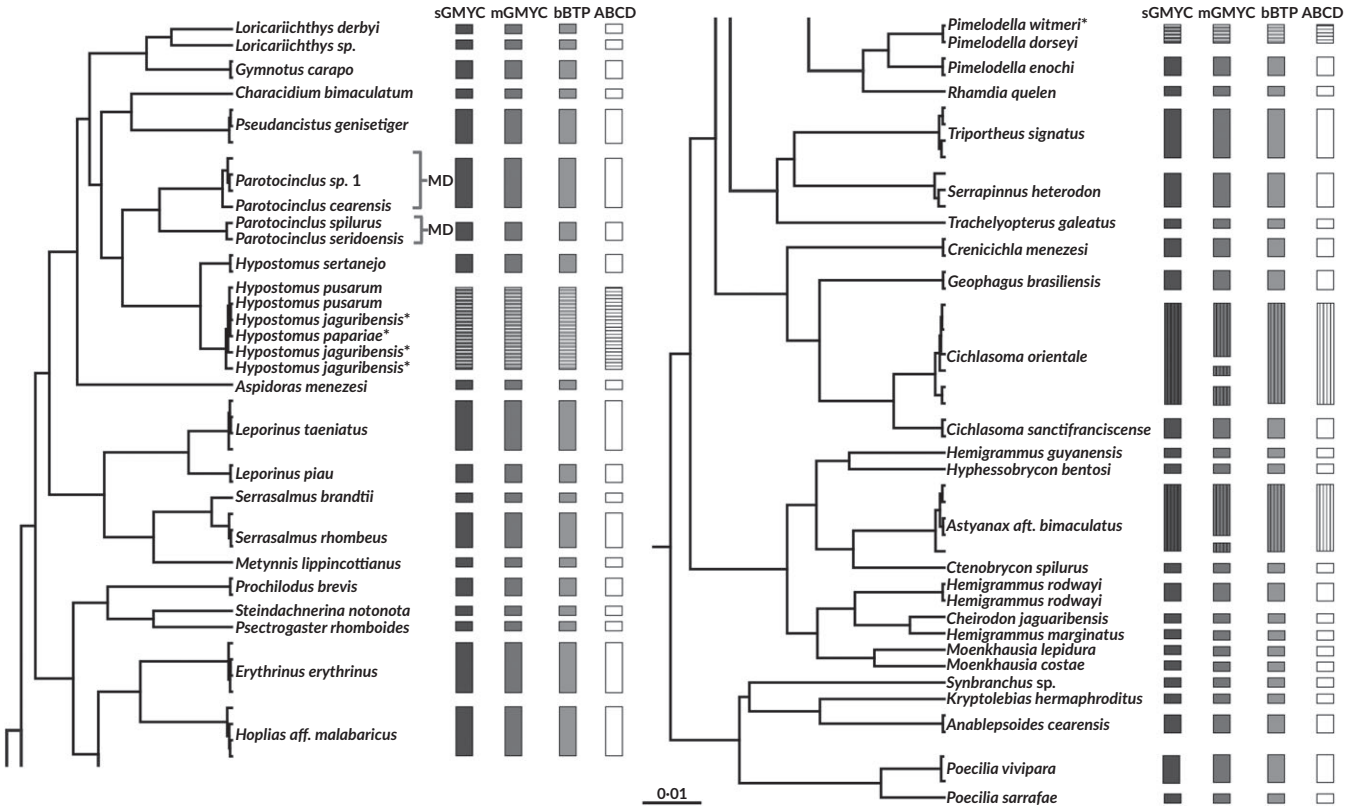

FIGURE 2 Ultrametric Bayesian tree illustrating relationships among 49 freshwater fish species from the mid-north-eastern Caatinga ecoregion that were barcoded in this study. Column boxes indicate the genetic clusters assigned by each species delimitation method; horizontally-striped bars represent disagreement among clustering methods; vertically-striped bars indicate potential synonyms. MD, morphologically distinguishable species that were assigned to the same genetic cluster; *, potential senior synonyms for topotype specimens

TABLE 3 Variation in richness estimates, based on four species delimitation analyses (C.I. in parentheses) for a sequenced subset of fishes from the mid-north-eastern Caatinga ecoregion

	sGMYC	mGMYC	bPTP	ABGD
Clusters	44 (44–48)	47 (37–47)	44 (43–49)	44 (39–45)
Matched	0.89	0.89	0.89	0.89
Merged	0.18	0.18	0.18	0.18
Splits	0	0.04	0	0

Matched, the proportion of delimited species matching valid species; Merged, the proportion of taxonomic species classified within a delimited species; Splits, the proportion of taxonomic species split by each delimitation method.

(Benzaquem *et al*., 2015; Marinho *et al*. 2016); however, their natural occurrence cannot be ruled out (Lima *et al*., 2017). All these species, with the exception of *H. rodwayi*, have been sampled from coastal Atlantic Forest basins of the MNCE, supporting the hypothesis of historical connections between Atlantic Forest and Amazonian biomes (Menezes *et al*., 2007; Sobral-Souza *et al*., 2015; Wang *et al*., 2004).

While the present study provides a newly updated list of MNCE freshwater fishes, there is still an urgent need for taxonomic revisions to determine the exact number of species, as well as percentage endemism. Such reviews should focus on genera with several nominal species (*e.g. Hypostomus*, *Leporinus*, *Pimelodella*), as well as genera with high species richness (*e.g. Aspidoras*, *Parotocinclus*). Additionaly, 14 of the 94 native species listed here were classified as *inquirenda*, with three of those potentially representing synonyms according to the barcode results (Figure 2). These 14 species belong to seven genera and were described by either Fowler (12) or Miranda-Ribeiro (2), whose descriptions were usually made based on few specimens and sometimes based only on juveniles (Lima *et al*., 2017; Ramos *et al*., 2016).

Rosa *et al*. (2003) registered 11 species from the MNCE that were not detected in this study: *Leporinus melanopleura* Günther 1864, *Pygocentrus piraya* (Cuvier 1819), *Pristobrycon striolatus* (Steindachner 1908), *Hemiodus parnaguae* Eigenmann & Henn 1916, *Roeboides microlepis* (Reinhardt 1851), *Serrapinnus* sp., *Trachelyopterus striatulus* (Steindachner 1908), *Salminus hilarii* Valenciennes 1850, *Moenkhausia lepidura* (Kner 1858), *Pimelodella gracilis* (Valenciennes 1835) and *Megalechis thoracata* (Valenciennes 1840). It may be that these records represent misidentifications. Additionally, seven of the 14 species *inquirenda* are only known from their type material (*Hypostomus salgadae* (Fowler 1941), *Cheirodon macropterus* Fowler 1941, *Hyphessobrycon iheringi* Fowler 1941, *Hyphessobrycon piabinhas* Fowler 1941, *Trachelyopterus cratensis* (Miranda Ribeiro 1937), *Pimelodella papariae* (Fowler 1941) and *Aphanotorulus gomesi* (Fowler 1942)), despite the extensive fieldwork that has been done in this region. The lack of available material, a history of poor diagnosis and imprecise or inaccurate type-locality descriptions (Lima *et al*., 2017; Rosa *et al*., 2003) combine to hamper the investigation of the real identity of these species. Although

unlikely, we cannot rule out the possibility that these species may be extinct.

4.1 | DNA barcode and species delimitation

Our DNA barcode dataset assessed 49 (52.1%) of the 94 native species in our updated list from the MNCE. Overall, both DNA barcode and species delimitation analyses discriminated the majority of the taxonomically identified species, with only 3 of 49 (6.1%) having low enough interspecific genetic distances to merge into the same genetic cluster, suggesting potential cases of taxonomic synonyms (Figure 2). The average intraspecific K2P distance was 0.18%, while average divergence among congeners was 6.70% (Table 2). These values were similar to, but slightly lower than values found by other studies (Carvalho *et al*., 2011; Pereira *et al*., 2013; Ward, 2009). These low intraspecific values could be related to a limited sampling of the geographic and genetic variation for many species. Although we tried to maximize the sampling of genetic variation within species (average of 5.1 specimens per species), 15 of the 49 species showed no genetic variation. Thus, a larger geographical sampling might increase average conspecific values. The low congeneric variation could be explained by taxonomic issues found for some species already treated as *inquirenda*. Indeed, two out of the 10 comparisons among genera (*Pimelodella* and *Hypostomus*) contained species *inquirenda* indicating possible synonyms, which would decrease average congeneric K2P distances.

When compared with other Neotropical freshwater fish studies, K2P distances were lower than those found by Pereira *et al*. (2011) and Carvalho *et al*. (2011). Importantly in the present study we have made comparisons among more genera (10) than those previous studies (four and six genera, respectively). The low congeneric variation was similar to that observed by Pereira *et al*. (2013), who made comparisons among 19 genera. These authors asserted that a larger number of congeneric comparisons is likely to decrease average congeneric divergence, which was reinforced by our data. Additionally, some studies suggest some lineages of Neotropical ichthyofauna have recently radiated (Montoya-Burgos, 2003; Hubert *et al*. 2007), which may further explain low values of congeneric variation when compared with non-Neotropical fish fauna, such as those in Australia (9.9%; Ward *et al*., 2005) and Canada (8.4%; Hubert *et al*., 2008).

The usual threshold value for species delimitation in barcode studies is 2% (Carvalho *et al*., 2011; Pereira *et al*., 2011, 2013; Ward, 2009). However, this value could just be indicative of mean divergence among species and species delimitation should take into account other factors, such as evolutionary history, morphology, ecology and behaviour (Hajibabaei *et al*., 2007; Pereira *et al*., 2013). Our results showed some cases of low interspecific genetic variation that may be related to taxonomic uncertainty. *Pimelodella dorseyi* and *P. witmeri*, for example, shared the same haplotype and genetic cluster across all species delimitation methods. This was despite both species being sampled from their respective type localities (Fowler, 1941; Supporting Information Table S1), which are separated by only approximately 30 km within the Jaguaribe River basin. The genetic distance within this clade of two haplotypes was 0.1%, or 67 times less than the conspecific average, suggesting that *P. witmeri* represents a junior synonym of *P. dorseyi* (according to the rules of nomenclatural priority). Descriptions of these species were brief and based on few specimens (two for *P. dorseyi*, three for *P. witmeri*) from the same drainage. Additionally, it is not possible to distinguish these species based on morphological data provided in the original description (Fowler, 1941), even when the topotypes collected herein were compared.

Most drainages in the MNCE featured only a single *Hypostomus* morphotype. Analysis of some putative dark-spotted species, including topotypes of *H. pusarum*, *H. papariae* and *H. jaguribensis*, along with specimens from the same coastal drainages, formed a single mOTU clade with little divergence among species (0.3% between *H. pusarum* and *H. papariae* and 0.4% between *H. pusarum* and *H. jaguribensis*). Divergence among these species and the sister clade (*Hypostomus sertanejo* Zawadzki, Ramos & Sabaj 2017) was about 6.7%, suggesting that *H. papariae* and *H. jaguribensis*, both described by Fowler (1915, 1941), might be junior synonyms of *H. pusarum*. We also recommend further investigation of *Hypostomus carvalhoi* (Miranda Ribeiro 1937), *Hypostomus nudiventris* (Fowler 1941) and *Hypostomus salgadae* (Fowler 1941), since all these putative dark-spotted species have their type localities within the MNCE and most were proposed by Fowler (1941). Recently, *Hypostomus eptingi* Fowler 1941, another dark-spotted species from the MNCE described by Fowler (1941), was formally synonymised with *H. jonhii* (Ramos *et al*., 2017). A formal taxonomic review, integrating both morphological and molecular data, is also necessary to better define the number of *Hypostomus* species and their distribution in the MNCE.

In addition to these examples of possible taxonomic synonyms in which no morphological or molecular differences were detected, some low levels of genetic variation could to be related to recent species divergence or slow mutation rates within particular taxa (Ward, 2009), which seems to be the case for the genus *Parotocinclus* in the MNCE. *Parotocinclus seridoensis* was only 0.9% divergent from *P. spilurus*; despite the former species having a naked abdomen region while the later presents an abdomen mostly covered by rounded dermal plates, in addition to other morphological differences (Ramos *et al*., 2013). *Parotocinclus cearensis* and *Parotocinclus* sp.1 also showed low genetic divergence (1.6%), but significantly differed in morphological traits traditionally used to diagnose *Parotocinclus* species. Such cases highlight the need for complementary data when assessing the accuracy of DNA barcode analyses (Pereira *et al*., 2013).

In conclusion, this study provided the first broad taxonomic study using both morphological and molecular data to determine freshwater fish species composition of the MNCE, including endemic, undescribed and non-native species. The study also highlights taxa that should be further reviewed and suggests a standardized nomenclature, mainly for some dubious species described by Fowler (Fowler, 1915, 1941, 1942) and Miranda-Ribeiro (1937). DNA barcode sequences that were generated for approximately 52% of native species also constitute an important contribution to studies of the systematics, biogeography and evolution of these mostly semi-arid lineages that evolved in temporary rivers historically connected to perennial drainages of adjacent forested Amazonian and Atlantic ecoregions.

ACKNOWLEDGEMENTS

Authors are grateful to everyone involded in the sampling, especially those from Universidade Federal do Rio Grande do Norte and Universidade Federal da Paraíba, for all the field trips made during several long years to make this study possible.

Authors' contributions

W.M.B-F. contributed with data collection, taxonomic and molecular analyses. T.P.A.R. contributed with data collection, taxonomic analyses, list composition and writing. U.P.J. contributed with data collection and molecular analyses. D.J.G.M. contributed with data collection. R.A.T. contributed with data collection, molecular analyses and funding. S.M.Q.L. contributed with data collection, taxonomic data, molecular analyses and funding. All authors contributed to writing the manuscript.

ORCID

Waldir M. Berbel-Filho http://orcid.org/0000-0001-6991-4685
Telton P. A. Ramos http://orcid.org/0000-0002-3808-8701

REFERENCES

Abell, R., Thieme, M. L., Revenga, C., Bryer, M., Kottelat, M., Bogutskaya, N., … Petry, P. (2008). Freshwater ecoregions of the world: A new map of biogeographic units for freshwater biodiversity conservation. *Bioscience*, *58*, 403–414.

Abell, R., Thieme, M. L., Ricketts, T. H., Owlero, N., Ng, R., Petry, P., … Hoeskstra, J. (2010). Concordance of freshwater and terrestrial biodiversity. *Conservation Letters*, *2*, 127–136.

Albert, J. S., & Reis, R. E. (2011). Introduction to Neotropical freshwaters. In J. S. Albert & R. E. Reis (Eds.), *Historical biogeography of Neotropical freshwater fishes* (pp. 1–19). Los Angeles, CA: University of California Press.

Armstrong, K. F., & Ball, S. L. (2005). DNA barcodes for biosecurity: Invasive species identification. *Philosophical Transactions of the Royal Society B*, *360*, 1813–1823.

Barros, M. C., Fraga, E. C., & Birindelli, J. L. O. (2011). Fishes from the Itapecuru River basin, state of Maranhão, northeast Brazil. *Brazilian Jounal of Biology*, *71*, 375–380.

Benzaquem, D. C., Oliveira, C., Batista, J. S., Zuanon, J., & Porto, J. I. R. (2015). DNA barcoding in pencilfishes (Lebiasinidae: *Nannostomus*) reveals cryptic diversity across the Brazilian Amazon. *PLoS One*, *10*, e0123363.

Berbel-Filho, W. M., Martinez, P. A., Ramos, T. P. A., Torres, R. A., & Lima, S. M. Q. (2016). Inter- and intra-basin phenotypic variation in two riverine cichlids from north-eastern Brazil: Potential eco-evolutionary damages of São Francisco interbasin water transfer. *Hydrobiologia*, *766*, 43–56.

Blair, C., & Bryson, R. W. (2017). Cryptic diversity and discordance in single-locus species delimitation methods within horned lizards (Phrynosomatidae: Phrynosoma). *Molecular Ecology Resources*, *17*, 1168–1182.

Britzke, R., Nielsen, D. T. B., & Oliveira, C. (2016). Description of two new species of annual fishes of the *Hypsolebias antenori* speces group (Cyprinodontiformes: Rivulidae), from Northeast Brazil. *Zootaxa*, *4114*, 123–138.

Buckup, P. A., Menezes, N. A., & Ghazzi, M. S. (2007). *Catálogo das Espécies de Peixes de Água Doce do Brasil*. Rio de Janeiro, Brazil: Museu Nacional.

Carstens, B. C., Pelletier, T. A., Reid, N. M., & Satler, J. D. (2013). How to fail at species delimitation. *Molecular Ecology*, *22*, 4369–4383.

Carvalho, D. C., Oliveira, D. A. A., Beheregaray, L., & Torres, R. A. (2012). Hidden genetic diversity and distinct evolutionarily significant units in an commercially important Neotropical apex predator, the catfish *Pseudoplatystoma corruscans*. *Conservation Genetics*, *13*, 1671–1675.

Carvalho, D. C., Oliveira, D. A. A., Pompeu, P. S., Leal, C. G., Oliveira, C., & Hanner, R. (2011). Deep barcode divergence in Brazilian freshwater fishes: The case of the São Francisco River basin. *Mitochondrial DNA*, *22*, 80–86.

Costa, N. K. R., Paiva, R. E. C., Silva, M. J., Ramos, T. P. A., & Lima, S. M. Q. (2017b). Ichthyofauna of Ceará-Mirim River basin, Rio Grande do Norte State, north-eastern Brazil. *ZooKeys*, *715*, 39–51. https://doi.org/10.3897/zookeys.715.13865

Costa, S. Y. L., Viana, L. G., Barbosa, J. E. L., & Ramos, T. P. A. (2017a). Composition of the ichthyofauna in Brazilian semiarid reservoirs. *Biota Neotropica*, *17*(3), e20170334. https://doi.org/10.1590/1676-0611-BN-2017-033

Costa, W. J. E. M. (2008). *Simpsonichthys longignatus*, a new seasonal killifish of the S. Flammeus group from the Pacoti river basin, north-eastern Brazil (Cyprinodontiformes:Rivulidae). *Ichthyological Exploration of Freshwaters*, *19*, 155–159.

Costa, W. J. E. M., & Vono, V. (2009). *Rivulus cearensis*, a new aplocheiloid killifish from north-eastern Brazil (Cyprinodontiformes: Rivulidae). *Ichthyological Exploration of Freshwaters*, *20*, 7–11.

Díaz, J., Villanova, G. V., Brancolini, F., del Pazo, F., Posner, V. M., Grimberg, A., & Arranz, S. E. (2016). First DNA barcode reference library for the identification of south American freshwater fish from the lower Paraná River. *PLoS One*, *11*, e0157419. https://doi.org/10.1371/journal.pone.0157419

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution, 29, 1969-1973.

Eschmeyer, W. N. & Fong, J. D. (2017). *Species by family/subfamily*. Retrieved from www.researcharchive.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp

Ezard, T., Fujisawa, T., & Barraclough, T. (2009). splits: SPecies' LImits by Threshold Statistics. R package version 1.0-11/r29.

Fowler, H. R. (1915). Cold-blooded vertebrates from Florida, the west INDIES, Costa Rica and eastern Brazil. *Proceedings of the Academy of Natural Sciences of Philadelphia*, *67*, 244–269.

Fowler, H. R. (1941). A collection of fresh-water fishes obtained in eastern Brazil by Rodolpho von Ihering. *Proceedings of the Academy of Natural Sciences of Philadelphia*, *93*, 123–199.

Fowler, H. W. (1942). An account of two new fishes from Brazil. *The Fish Culturist*, *22*, 1–2.

Fujisawa, T., & Barraclough, T. G. (2013). Delimiting species using single-locus data and the generalized mixed yule coalescent approach: A revised method and evaluation on simulated data sets. *Systematic Biology*, *62*, 707–724.

Gurgel-Lourenço, R. C., Sousa, W. A., Sánchez-Botero, J. I., & Garcez, D. S. (2013). Ichthyofauna of two reservoirs in the middle Acaraú river basin, Ceará, north-eastern Brazil. *Check List*, *9*, 1391–1395.

Hajibabaei, M., Singer, G. A. C., Hebert, P. D. N., & Hickey, D. A. (2007). DNA barcoding: How it complements taxonomy, molecular phylogenetics and population genetics. *Trends in Genetics*, *23*, 167–172.

Hubert, N., Duponchelle, F., Nunez, J., Garcia-Davila, C., Paugy, D., & Renno, J. F. (2007). Phylogeography of the piranha genera *Serrasalmus* and *Pygocentrus*: Implications for the diversification of the Neotropical ichthyofauna. *Molecular Ecology*, *16*, 2115–2136.

Hubert, N., Hanner, R., Holm, E., Mandrak, N. E., Taylor, R., Burridge, M., … Bernatchez. (2008). Identifying Canadian freshwater fishes trough DNA barcodes. *PLoS One*, *3*, e2490. https://doi.org/10.1371/journal.pone.0002490

Jerep, F. C., & Malabarba, L. R. (2014). A new species of *Serrapinnus* Malabarba 1998 (Teleostei: Characidae: Cheirodontinae) from Rio Grande do Norte State, north-eastern Brazil. *Neotropical Ichthyology*, *12*, 301–308.

Kekkonen, M., & Hebert, P. D. N. (2014). DNA barcode-based delineation of putative species: Efficient start for taxonomic workflows. *Molecular Ecology*, *14*, 706–715.

KKearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., … Thierer, T. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*, 1647–1649

Langeani, F. L., Buckup, P. A., Malabarba, L. R., Py-Daniel, L. H. R., Lucena, C. A., Rosa, R. S., … Gomes-Filho, G. (2009). Peixes de água Doce. In R. M. Rocha & W. A. P. Boeger (Eds.), *Estado da arte e perspectivas para a Zoologia no Brasil* (pp. 211–230). Curitiba, Brazil: Editora UFPR.

Lévêque, C., Oberdoff, T., Paugy, D., Stiassny, M. L. J., & Tedesco, P. A. (2008). Global diversity of fish (Pisces) in freshwater. *Hydrobiologia*, *595*, 545–567.

Lima, S. M. Q., Ramos, T. P. A., Silva, M. J., & Rosa, R. S. (2017). Diversity, distribution and conservation of the Caatinga fishes: Advances and challenges. In J. M. C. Silva, I. R. Leal, & M. Tabarelli (Eds.), *Caatinga* (pp. 97–131). Basel, Switzerland: Springer International Publishing.

Lira, M. G. S., Paiva, R. E. C., Ramos, T. P. A., & Lima, S. M. Q. (2015). First record of *Kryptolebias hermaphroditus* Costa, 2011 (Cyprinodontiformes: Rivulidae) in the extreme north Atlantic Forest mangroves, Rio Grande do Norte state, Brazil. *Checklist*, *11*, 1656.

Marinho, M. M. F., Dagosta, F. C. P., Camelier, P., & Oyakawa. (2016). A name for the 'blueberry tetra', an aquarium trade popular species of *Hyphessobrycon* Durbin (Characiformes, Characidae), with comments on fish species descriptionslacking accurate type locality. *Journal of Fish Biology*, *89*, 510–521.

Menezes, N. A., Weitzman, S. H., Oyakawa, O. T., Lima, F. C. T., Castro, R. M. C., & Weitzman, M. J. (2007). *Peixes de água doce da Mata Atlântica: lista preliminar das espécies e comentários sobre conservação de peixes de água doce neotropicais*. São Paulo, Brazil: Museu de Zoologia da Universidade de São Paulo.

Miranda-Ribeiro, A. (1937). Sobre uma collecção de vertebrados do nordeste brasileiro. Primeira parte: Peixes e batrachios. *O Campo*, *1*, 54–56.

Montoya-Burgos, J. I. (2003). Historical biogeography of the catfish genus *Hypostomus* (Siluriformes: Loricariidae), with implications on the diversification of Neotropical ichthyofauna. *Molecular Ecology*, *12*, 1855–1867.

Nascimento, W. S., Barros, N. H. C., Araújo, A. S., Gurgel, L. L., Canan, B., Molina, W. F., … Chellappa, S. (2014). Composição da ictiofauna das bacias hidrográficas do Rio Grande do Norte, Brasil. *Biota Amazônia*, *4*, 126–131.

Novaes, J. L. C., Moreira, S. I. L., Freire, C. E. C., Sousa, M. M. O., & Costa, R. S. (2013). Fish assemblage in a semi-arid Neotropical reservoir: Composition, structure and patterns of diversity and abundance. *Brazilian Journal of Biology*, *74*, 290–301.

Paiva, R. E. C., Lima, S. M. Q., Ramos, T. P. A., & Mendes, L. (2014). Fish fauna of Pratagi River coastal microbasin, extreme north Atlantic Forest, Rio Grande do Norte State, north-eastern Brazil. *Check List*, *10*, 968–975.

Pereira, L. H. G., Hanner, R., Foresti, F., & Oliveira, C. (2013). Can DNA barcoding accurately disciminate megadiverse Neotropical freshwater fish fauna? *BMC Genomics*, *14*, 20. https://doi.org/10.1186/1471-2156-14-20

Pereira, L. H. G., Maia, G. M. G., Hanner, R., Foresti, F., & Oliveira, C. (2011). DNA barcodes discriminate freshwater fishes from the Paraíba do Sul river basin, São Paulo, Brazil. *Mitochondrial DNA*, *22*, 71–79.

Puillandre, N., Lambert, A., Brouillet, S., & Achaz, G. (2012). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Molecular ecology, *21*, 1864–1877.

Rambaut, A. (2009). TRACER v1. 5. http://tree. bio. ed. ac. uk/software/tracer/.

Ramirez, J. L., Birindelli, J. L., Carvalho, D. C., Affonso, P. R. A. M., Venere, P. C., Ortega, H., … Galetti, P. M., Jr. (2017). Revealing hidden diversity of the underestimated Neotropical ichthyofauna: DNA barcoding in the recently described genus *Megaleporinus* (Characiformes: Anostomidae). *Front Genetics*, *8*, 149. https://doi.org/10.3389/fgene.2017.00149

Ramos, R. T. C., Ramos, T. P. A., Rosa, R. S., Beltrão, G. B. M., & Groth, F. (2005). Diversidade de Peixes (Ictiofauna) da bacia do rio Curimataú, Paraíba. In F. S. Araújo, M. J. N. Rodal, & M. R. V. Barbosa (Eds.), *Análise das variações da biodiversidade do bioma caatinga: suporte das estratégias regionais de conservação* (pp. 291–318). Brasília, Brazil: Ministério do Meio Ambiente.

Ramos, T. P. A., Barros-Neto, L. F., Britski, H. A., & Lima, S. M. Q. (2013). *Parotocinclus seridoensis*, a new hypoptopomatine catfish (Siluriformes: Loricariidae) from the upper rio Piranhas-Açu basin, north-eastern Brazil. *Neotropical Ichthiology*, *11*, 787–796.

Ramos, T. P. A., Lehmann, P. A., Barros-Neto, L. F., & Lima, S. M. (2016). Redescription of the endangered hypoptopomatine catfish *Parotocinclus spilurus* (Fowler 1941) (Siluriformes: Loricariidae) from the upper rio Jaguaribe basin, north-eastern Brazil. *Neotropical Ichthyology*, *14*, 1.

Ramos, T. P. A., Ramos, R. T. C., & Ramos, S. A. Q. A. (2014). Ichthyofauna of the Parnaíba river Basin, north-eastern Brazil. *Biota Neotropica*, *14*, 1–8.

Ramos, T. P. A., Zawadzki, C. H., Ramos, R. T. C., & Britski, H. A. (2017). Redescription of *Hypostomus johnii*, a senior synonym of *Hypostomus eptingi* (Siluriformes: Loricariidae), North-eastern Brazil. *Neotropical Ichthyology*, *15*, e160064.

Reis, R. E., Kullander, S. O., & Ferraris, C. J. (2003). *Check list of the freshwater fishes of south and Central America*. Porto Alegre, Brazil: EDIPUCRS.

Rocha, L. G., Medeiros, E. S. F., & Andrade, H. T. A. (2012). Influence of flow variability on macroinvertebrate assemblages in an intermittent stream of semi-arid Brazil. *Journal of Arid Environments*, *85*, 33–40.

Rodrigues-Filho, C. A., Gurgel-Lourenço, R. C., Bezerra, L. A., Sousa, W. A., Garcez, D. S., Lima, S. M. Q., … Botero, J. I. S. (2016). Ichthyofauna of the humid forest enclaves in the tablelands of Ibiapaba and Araripe, North-eastern Brazil. *Biota Neotropica*, *16*, e20160273.

Rosa, R. S., & Groth, F. (2004). Ictiofauna dos ecossistemas de Brejos de Altitude de Pernambuco e Paraíba. In K. C. Porto, J. J. P. Cabral, & M. Tabarelli (Eds.), *Brejos de Altitude em Pernambuco e Paraíba, História Natural, Ecologia e Conservação* (pp. 201–210). Brasília, Brazil: Ministério do Meio Ambiente.

Rosa, R. S., Menezes, N. A., Britski, H. A., Costa, W. J. E. M., & Groth, F. (2003). Diversidade, padrões de distribuição e conservação dos peixes da Caatinga. In I. R. Leal, M. Tabarelli, & J. M. C. Silva (Eds.), *Ecologia e Conservação da Caatinga* (pp. 135–162). Recife, Brazil: Editora Universitária da UFPE.

Schliep, K. P. (2010). Phangorn: Phylogenetic analysis in R. *Bioinformatics*. p.btq706.

Schliep, K. P. (2011). phangorn: phylogenetic analysis in R. Bioinformatics, *27*(4), 592.

Sigovini, M., Keppel, E., & Tagliapietra, D. (2016). Open nomenclature in the biodiversity era. *Methods in Ecology and Evolution*, *30*, 1217–1255.

Silva, M. J., Neves, C. P., Ramos, T. P. A., Pinto, M. P., & Lima, S. M. Q. (2017). Efetividade das unidades de conservação para a ictiofauna das bacias hidrográficas envolvidas na transposição do rio São Francisco. In W. Mantovani, R. F. Monteiro, L. dos Anjos, & M. O. Cariello (Eds.), *Pesquisas em unidades de conservação no domínio da Caatinga: subsídios à gestão* (pp. 181–206). Fortaleza, Brazil: Edições UFC.

Silva, M. J., Ramos, T. P. A., Diniz, V. D., Ramos, R. T. C., & Medeiros, E. S. F. (2014). Ichthyofauna of Seridó/Borborema: A semi-arid region of Brazil. *Biota Neotropica*, *14*, 1–6.

Smith, M. A., Fisher, B. L., & Hebert, P. D. (2005). DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: The ants of Madagascar. *Philosophical Transactions of the Royal Society B*, *360*, 1825–1834.

Sobral-Souza, T., Lima-Ribeiro, M. S., & Solferini, V. N. (2015). Biogeography of Neotropical rainforests: Past connections between Amazon and Atlantic Forest detected by ecological niche modeling. *Evolutionary Ecology*, *29*, 643–655.

Starks, E. C. (1913). *The fishes of the Stanford expedition to Brazil* (pp. 1–77). Standford, CA: Leland Stanford Junior University Publications University Series.

Talavera, G., Dincă, V., & Vila, R. (2013). Factors affecting species delimitations with the GMYC model: insights from a butterfly survey. *Methods in Ecology and Evolution*, *4*, 1101–1110.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, *30*, 2725–2729.

Tatarenkov, A., Lima, S. M., Earley, R. L., Berbel-Filho, W. M., Vermeulen, F. B., Taylor, D. S., … Avise, J. C. (2017). Deep and concordant subdivisions in the self-fertilizing mangrove killifishes (*Kryptolebias*) revealed by nuclear and mtDNA markers. *Biological Journal of the Linnean Society*, *122*, 558–578.

Tedesco, P. A., Beauchard, O., Bigorne, R., Blanchet, S., Buisson, L., Conti, L., … Thierry, O. (2017). A global database on freshwater fish species occurrence in drainage basins. *Scientific Data*, *4*, 170141.

Teixeira, F. K., Ramos, T. P. A., Paiva, R. E. C., Távora, M. A., Lima, S. M. Q., & Rezende, C. F. (2017). Ichthyofauna of Mundaú river basin, Ceará State, North-eastern Brazil. *Biota Neotropica*, *17*, 1e20160174.

Torres, R. A., & Ribeiro, J. (2009). The remarkable species complex *Mimagoniates microlepis* (Characiformes: Glandulocaudinae) from the southern Atlantic rain forest (Brazil) as revealed by molecular systematic and population genetic analyses. *Hydrobiologia*, *617*, 157–170.

Wang, X., Auler, A. S., Edwards, R. L., Cheng, H., Cristalli, P. S., Smart, P. L., … Shen, C. (2004). Wet periods in north-eastern Brazil over the past 210kyr linked to distant climate anomalies. *Nature*, *432*, 740–743.

Ward, R. D. (2009). DNA barcode divergence among species and genera of birds and fishes. *Molecular Ecology Resources*, *9*, 1077–1085.

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., & Hebert, P. D. (2005). DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B*, *360*, 1847–1857.

Yang, Z. (1996). Maximum-likelihood models for combined analyses of multiple sequence data. *Journal of Molecular Evolution*, *42*, 587–596.

Zawadzki, C. H., Ramos, T. P. A., & Sabaj, R. (2017). *Hypostomus sertanejo* (Siluriformes: Loricariidae), new armoured catfish species from north-astern Brazil. *Journal of Fish Biology*, *91*, 317–330.

Zhang, J., Kapli, P., Pavlidis, P., & Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*, *29*, 2869–2876.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Berbel-Filho WM, Ramos TPA, Jacobina UP, Maia DJG, Torres RA, Lima SMQ. Updated checklist and DNA barcode-based species delimitations reveal taxonomic uncertainties among freshwater fishes from the mid-north-eastern Caatinga ecoregion, north-eastern Brazil. *J Fish Biol*. 2018;93:311–323. https://doi.org/10.1111/jfb.13758