Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding

A. Valentini, P. Taberlet, C. Miaud, Raphaël Civade, J. Herder, P.F. Thomsen, E. Bellemain, A. Besnard, E. Coissac, F. Boyer, et al.

▶ **To cite this version:**

A. Valentini, P. Taberlet, C. Miaud, Raphaël Civade, J. Herder, et al.. Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. Molecular Ecology, 2015, 25 (4), pp.929-942. 10.1111/mec.13428 . hal-01419572

HAL Id: hal-01419572
https://hal.science/hal-01419572

Submitted on 21 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

OATAO
Open Archive Toulouse Archive Ouverte

Open Archive Toulouse Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible

This is an author's version published in: http://oatao.univ-toulouse.fr/23485

Official URL: https://doi.org/10.1111/mec.13428

To cite this version:

Valentini, Alice and Taberlet, Pierre and Miaud, Claude and Civade, Raphael and Herder, Jelger and Thomsen, Philip Francis and Bellemain, Eva and Besnard, Aurelien and Coissac, Eric and Boyer, Frederic and Gaboriaud, Coline and Jean, Pauline and Poulet, Nicolas and Roset, Nicolas and Copp, Gordon H. and Geniez, Philippe and Pont, Didier and Argillier, Christine and Baudoin, Jean-Marc and Peroux, Tiphaine and Crivelli, Alain J. and Olivier, Anthony and Acqueberge, Manon and Le Brun, Matthieu and Moller, Peter R. and Willerslev, Eske and Dejean, Tony *Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding.* (2015) Molecular Ecology, 25 (4). 929-942. ISSN 0962-1083

Any correspondence concerning this service should be sent to the repository administrator: tech-oatao@listes-diff.inp-toulouse.fr

Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding

ALICE VALENTINI,* PIERRE TABERLET,†‡ CLAUDE MIAUD,§ RAPHAËL CIVADE,¶ JELGER HERDER,** PHILIP FRANCIS THOMSEN,†† EVA BELLEMAIN,* AURÉLIEN BESNARD,§ ERIC COISSAC,†‡ FRÉDÉRIC BOYER,†‡ COLINE GABORIAUD,* PAULINE JEAN,* NICOLAS POULET,‡‡ NICOLAS ROSET,§§ GORDON H. COPP,¶¶*** PHILIPPE GENIEZ,§ DIDIER PONT,¶ CHRISTINE ARGILLIER,††† JEAN-MARC BAUDOIN,††† TIPHAINE PEROUX,††† ALAIN J. CRIVELLI,‡‡‡ ANTHONY OLIVIER,‡‡‡ MANON ACQUEBERGE,§§§ MATTHIEU LE BRUN,¶¶¶ PETER R. MØLLER,**** ESKE WILLERSLEV†† and TONY DEJEAN*

*SPYGEN, Savoie Technolac Bât. Koala, 17, Rue du Lac Saint-André BP 274, Le Bourget-du-Lac Cedex 73375, France, †Laboratoire d'Ecologie Alpine (LECA), CNRS, Grenoble 38000, France, ‡Laboratoire d'Ecologie Alpine (LECA), Univ. Grenoble Alpes, Grenoble 38000, France, §Laboratoire Biogéographie et Ecologie des Vertébrés, CEFE UMR 5175, Montpellier 34293, France, ¶Hydrosystems and Bioprocesses Research Unit, IRSTEA, Antony Cedex 92761, France, **RAVON, Postbus 1413, Nijmegen 6501 BK, The Netherlands, ††Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Øster Voldgade, Copenhagen, Denmark, ‡‡Direction de l'Action Scientifique et Technique, ONEMA, Vincennes 94300, France, §§Rhône-Alpes Regional Direction, ONEMA, Bron 69500, France, ¶¶Centre for Environment, Fisheries and Aquaculture Science, Pakefield Road, Lowestoft, Suffolk NR33 0HT, UK, ***Environmental and Life Sciences Graduate Program, Trent University, Peterborough, ON K9J 7B8, Canada, †††Pole ONEMA/IRSTEA Hydroécologie des plans d'eau, Centre d'Aix-en-Provence, IRSTEA UR HYAX, Aix-en-Provence 13182, France, ‡‡‡Le Sambuc, Tour du Valat, Arles 13200, France, §§§Agence Centre-Ouest, Ecosphère, Orléans 45000, France, ¶¶¶LNHE Department, EDF R&D, Chatou Cedex 78401, France, ****Natural History Museum of Denmark, University of Copenhagen, Universitetsparken 15, Copenhagen 2100, Denmark*

Abstract

Global biodiversity in freshwater and the oceans is declining at high rates. Reliable tools for assessing and monitoring aquatic biodiversity, especially for rare and secretive species, are important for efficient and timely management. Recent advances in DNA sequencing have provided a new tool for species detection from DNA present in the environment. In this study, we tested whether an environmental DNA (eDNA) metabarcoding approach, using water samples, can be used for addressing significant questions in ecology and conservation. Two key aquatic vertebrate groups were targeted: amphibians and bony fish. The reliability of this method was cautiously validated in silico, in vitro and in situ. When compared with traditional surveys or historical data, eDNA metabarcoding showed a much better detection probability overall. For amphibians, the detection probability with eDNA metabarcoding was 0.97 (CI = 0.90 0.99) vs. 0.58 (CI = 0.50 0.63) for traditional surveys. For fish, in 89% of the studied sites, the number of taxa detected using the eDNA metabarcoding approach was higher or identical to the number detected using traditional methods. We argue that the proposed DNA-based approach has the potential to become the next-generation tool for ecological studies and standardized biodiversity monitoring in a wide range of aquatic ecosystems.

Keywords: amphibian, detection probability, environmental DNA, fish, monitoring, wildlife management

Correspondence: Dr. Tony Dejean, Fax: +33(0)4 79 26 15 83;
E mail: tony.dejean@spygen.com

.

Introduction

Global biodiversity loss represents one of the most serious environmental crises of the 20th and 21st centuries, with considerable impact on both ecosystem services and the health of our planet (Pimm *et al.* 2014). An overall biodiversity decline of 52% was recorded between 1970 and 2010, and this loss was even higher for freshwater populations than for marine or terrestrial ecosystems (WWF 2014). While scientists are struggling to find explanations and to understand the gravity of this loss (Monastersky 2014), there is an urgent need to improve effectiveness of the strategies employed to halt global biodiversity loss and render them more integrative (Beumer & Martens 2013). It is commonly acknowledged that biodiversity is threatened by a variety of anthropogenic factors (Barnosky *et al.* 2011); however, the most critical issue is the sheer lack of efficient and reliable tools to document the remaining species and to assess biodiversity trends. Indeed, to monitor certain taxonomic groups, the existing methods appear to be inefficient, selective, destructive or strictly dependent on a declining taxonomic expertise (Wheeler *et al.* 2004). Recent technological advances have provided an alternative tool for species detection using DNA present in aquatic or terrestrial environments (environmental DNA or eDNA; Taberlet *et al.* 2012). Taxonomical inventories and assessment of geographical distribution of species based on eDNA analysis may help to improve environmental monitoring and influence management and policy decisions (Kelly *et al.* 2014a; Thomsen & Willerslev 2015). In contrast to conventional survey methods, eDNA offers the advantages of being noninvasive (no macroorganisms are caught, disturbed or killed during monitoring) and reduces the risk of unintentional secondary dispersal of alien species and diseases. Overall, eDNA methods used for species detection demonstrated higher detection capability and cost-effectiveness compared to traditional methods (e.g. Darling & Mahon 2011; Dejean *et al.* 2012). Two main approaches using eDNA have been proposed: eDNA barcoding (or species-specific approach), which aims at detecting a single species in the environment, and eDNA metabarcoding (or multispecific approach), which simultaneously identifies several taxa from an environmental sample without 'a priori' knowledge of the species likely to be present in the sampled ecosystem (Taberlet *et al.* 2012). The latter method is based on the use of primers specific for a given taxonomic group (thereafter referred as group-specific primers) coupled with next-generation sequencing (NGS). It has been largely applied in terrestrial ecosystems and for diet analysis (e.g. Pompanon *et al.* 2012; Yoccoz *et al.* 2012). However, applications of eDNA metabarcoding in aquatic environments are still in their infancy, with only six studies published to date (Thomsen *et al.* 2012a,b; Kelly *et al.* 2014b; Deiner *et al.* 2015; Evans *et al.* 2015; Miya *et al.* 2015), of which only the ones from Thomsen, Deiner and Miya were performed in natural environments. These studies have emphasized the challenges posed by the method, for example sampling strategy optimization and markers reliability, but also the great perspectives yet to be investigated.

The objective of this study was to test whether an eDNA approach using water samples can be used for addressing significant questions in ecology and conservation. More specifically, our goal is to assess the potential of an integrated eDNA metabarcoding approach for ecological studies and for monitoring aquatic biodiversity, and to demonstrate its reliability considering two key aquatic vertebrate groups: amphibians and bony fish. Characterized by many discrete, rare and recently extinct species, amphibians (Batrachia) represent one of the most vulnerable animal groups (Stuart *et al.* 2004). Their distributions are often poorly known because the detection probabilities using traditional survey methods can be very low and may vary because of local environmental conditions (Kéry & Schmidt 2008; Tanadini & Schmidt 2011). Bony fish (Teleostei) are generally less cryptic and are often used as indicators of the ecological status of rivers (Roset *et al.* 2007) and lakes (Argillier *et al.* 2013) or of the function of water bodies (Copp *et al.* 1991). Nevertheless, traditional fish inventories show their limits in large water bodies (e.g. large rivers or lakes) where stratified sampling methods and invasive fishing gear (e.g. gill nets) must be used. Furthermore, those methods are usually species and/or size-selective (Hudy 1985; Hubert *et al.* 2012) and suffer from a relatively high proportion of fish species misidentification, particularly for young stages (Daan 2001).

The eDNA metabarcoding approach proposed here was cautiously validated in silico (i.e. bioinformatically), in vitro (i.e. using DNA extracted from tissue samples) and in situ, on a wide range of aquatic ecosystems (including both stagnant and running waters) by comparing the eDNA metabarcoding results with traditional survey data, historical data and results obtained using a previous eDNA metabarcoding approach (Thomsen *et al.* 2012b).

Materials and methods

Design and in silico validation of group-specific primers

Group-specific primers were designed on mtDNA for Teleostei and Batrachia using the ECOPRIMERS software

(Riaz *et al.* 2011) which identifies barcode markers and their associated primers, considering several constraints, such as (i) the target taxonomic group, (ii) the minimum and maximum length of the amplicons, (iii) the maximum number of mismatches between the primers and their target sequences and (iv) the number of nucleotides with a perfect match on the 3′ end of the primers. Because DNA is susceptible to degradation in the environment, metabarcodes of <100 bp were targeted (Taberlet *et al.* 2012). ECOPRIMERS optimizes both the versatility of the primers for the target group and the taxonomic resolution of the amplified region. The designed primers were then tested by in silico PCR (ECOPCR program; Bellemain *et al.* 2010; Ficetola *et al.* 2010) against the entire set of DNA sequences available from the EMBL-European Nucleotide Archive (release 117, standard sequences). This step allowed the evaluation of the match between the designed primers and all potentially amplifiable sequences. The primers were then manually optimized, when necessary, by degenerating some nucleotide positions within the primers to better match the desired target group and by adding a few nucleotides on the 5′ end to homogenize the annealing temperatures of both the forward and reverse primers.

The designed primers were then extensively tested in silico using the ECOPCR program on: (i) a collection of all mitochondrial DNA sequences, and (ii) all the publicly available DNA sequences. Both data sets were based on release 117 (standard sequences) of the EMBL database. First, for each primer, a sequence logo was generated (Crooks *et al.* 2004) based on the in silico PCR results for the target group without enforcing a perfect match at the 3′-end nucleotides and with a maximum of three mismatches in the entire primer sequence. This logo illustrates the match between the primer and its target sequence within the target taxonomic group. Second, a mismatch analysis was performed, both for the target taxonomic group and for the nontarget group (maximum of three mismatches allowed, excluding the last two nucleotides on the 3′ end) to assess the specificity of the primer pairs. Finally, the length distribution of the amplified sequences (excluding primers) was analysed using the OBITOOLS package (http://metabarcoding.org/obitools; Boyer *et al.* 2015).

Additionally, the Teleostei primer performance was evaluated by comparison with the alternative primer pairs recently proposed by Kelly *et al.* (2014b) and by Thomsen *et al.* (2012b).

Reference database

Fifty-three amphibian and 86 fish species (64 European freshwater fish species and 22 marine fish species present in Danish coastal waters) were collected by experienced herpetologists and ichthyologists and used to construct the reference database (Tables S1 and S2, Supporting information, respectively). DNA was extracted from 10 mg of tissue sampled from 1 to 13 individuals per species, taken from hind leg muscles of amphibians and from dorsal or caudal fins of fish, using DNeasy Blood & Tissue kit (Qiagen GmbH, Hilden, Germany), following the manufacturer's instructions, in a room dedicated to DNA extraction from tissues. DNA amplifications were performed in a final volume of 25 μL, using 3 μL of DNA extract as template (ranging from 1 to 30 ng/μL). The amplification mixture contained 1 U of AmpliTaq Gold DNA Polymerase (Applied Biosystems, Foster City, CA, USA), 10 mM of Tris-HCl, 50 mM of KCl, 2.5 mM of $MgCl_2$, 0.2 mM of each dNTP, 0.2 μM of each group-specific primer designed in this study (Table 1), 4 μM of human blocking primers (i.e. a DNA oligo that preferentially binds to human DNA and that is modified to impede its amplification; Table 1) and 0.2 μg/μL of bovine serum albumin (Roche Diagnostic, Basel, Switzerland). For both taxonomic groups, the PCR mixture was denatured at 95 °C for 10 min, followed by 35 cycles of 30 s at 95 °C, 30 s at 55 °C and 1 min at 72 °C, followed by a final elongation at 72 °C for 7 min, in a room dedicated to amplified DNA, with negative air pressure and physically separated from the DNA extraction rooms.

Table 1 Group specific mitochondrial 12S primers and human blocking primers for Batrachia and Teleostei. The code corresponds to the position of the 3′ nucleotide on the *Rana nigromaculata* mitochondrial genome (Accession no.: NC 002805) for the batra F and batra R primers and on the *Cyprinus carpio* mitochondrial genome (Accession no.: NC 001606) for teleo F and teleo R primers

Name	Code	Sequence (5′ 3′)	Suggested annealing temperature
batra F	L3541	ACACCGCCCGTCACCCT	55 °C
batra R	H3596	GTAYACTTACCATGTTACGACTT	
batra blk		TCACCCTCCTCAAGTATACTTCAAAGGCA SPC3I	
teleo F	L1848	ACACCGCCCGTCACTCT	55 °C
teleo R	H1913	CTTCCGGTACACTTACCATG	
teleo blk		ACCCTCCTCAAGTATACTTCAAAGGAC SPC3I	

The PCR products were purified and sequenced (forward and reverse) using Sanger technology at the Eurofins MWG Operon sequencing facilities (Ebersberg, Germany). Sequences were aligned and primers were trimmed using GENEIOUS v6.0 (Biomatters, Auckland, New Zealand, http://www.geneious.com/).

In vitro validation of the designed primers

To validate the universality of the designed markers for the studied taxonomic groups, we assessed whether all target species used in the reference database could be amplified with these primers. For this purpose, two pools of DNA extracted from tissue samples were prepared, one for each taxonomic group (Batrachia and freshwater Teleostei) by mixing 3 μL of DNA from each species. The group-specific primers were 5′ labelled with a unique seven-nucleotide tag (with at least three differences between tags) allowing the assignment of sequences to the respective samples during the sequence analysis. Tags for forward and reverse primers were identical for each sample. A negative control (ultrapure water) was added during the PCR step. After amplification, the two samples were titrated using capillary electrophoresis (QIAxcel; Qiagen GmbH) and purified using a MinElute PCR purification kit (Qiagen GmbH). Before sequencing, purified DNA was titrated again using capillary electrophoresis. Several purified PCR products were pooled in equal volumes, to achieve an expected sequencing depth of 100 000 reads per amphibian DNA sample and 300 000 reads per fish DNA sample. Library preparation and sequencing were performed at Fasteris facilities (Geneva, Switzerland). Libraries were prepared using TruSeq Nano DNA genomic kit (Illumina, San Diego, CA, USA) and a pair-end sequencing (2×100 bp) was carried out using an Illumina MiSeq sequencer (Illumina) using the Pair-end MiSeq Reagent Kit V2 (Illumina) following the manufacturer's instructions. In total, eight MiSeq runs were performed.

The sequence reads were analysed using the programs implemented in the OBITOOLS package (http://metabarcoding.org/obitools; Boyer *et al.* 2015) as described in De Barba *et al.* (2014). Program ILLUMINAPAIREDEND was used to assemble forward and reverse reads corresponding to a single molecule. No special threshold was applied after the alignment step, the bad alignments being removed implicitly during the following filtration steps. Subsequently, NGSFILTER identified primers and tags and assigned the sequences to each sample. This program was used with its default parameters tolerating two mismatches for each of the two primers and no mismatch for the tags. A separate data set was created for each sample by splitting the original data set in several files using OBISPLIT. After this step, each sample was analysed individually before merging the taxon list for the final ecological analysis. Strictly identical sequences were clustered together using OBIUNIQ. Sequences shorter than 20 bp, or with occurrence lower than 10 were excluded using the OBIGREP program. The OBICLEAN program was then run to assign the status of 'head', 'internal' or 'singleton' to each sequence, within a PCR product. All sequences labelled 'internal' that correspond most likely to PCR substitutions and indel errors were discarded. The taxonomic assignment of MOTUs was performed using the program ECOTAG, with both the local reference database of Batrachia and Teleostei built for this study and the sequences extracted from the release 118 (standard sequences) of the EMBL database using the ECOPCR program (Bellemain *et al.* 2010; Ficetola *et al.* 2010). MOTUs showing <98% similarity with either the local or the EMBL reference databases were removed. Taxa were preferentially assigned based on the local reference database, except if the similarity was higher for the EMBL reference database. Finally, to take into account bad assignation of a few numbers of sequences to the wrong sample (Schnell *et al.* 2015), all sequences with a frequency of occurrence below 0.001 per taxon and per run for amphibians and below 0.003 per taxon and per run for fish were discarded. These threshold were empirically determined to clear the controls included in our global data production procedure (De Barba *et al.* 2014).

In situ validation of the eDNA metabarcoding approach

To assess the efficiency and reliability of this proposed eDNA metabarcoding approach, a comparative study was performed on 62 sites corresponding to a wide range of aquatic ecosystems (Tables S3 and S4, Supporting information). The eDNA sampling for all sites was conducted on the same day as the traditional method with some exceptions (see Tables S3 and S4, Supporting information). At each site, the list of species obtained using eDNA was compared to that from traditional surveys (n 57) or historical data (n 4). Six sites were chosen to test the presence of false positives: two small channels in a saltwater marsh along the Mediterranean Sea in 2012 because they were unsuitable for amphibian life (sites 18 and 19, Table S3, Supporting information), and four mountain lakes, two without fish (Sites 1 and 2, Table S4, Supporting information) and two where the brown trout *Salmo trutta* was the only species present (sites 3 and 4, Table S4, Supporting information). To evaluate the performance of the metabarcoding approach proposed here (primers, sequencing device

and analysis of NGS data) without the effect of sampling or extraction protocol, the same DNA extracts used by Thomsen *et al*. (2012b; sea samples) were also analysed.

Traditional field survey. For the comparative study, 39 sites were surveyed for amphibian species in 2012 (16 sites) and 2013 (23 sites), consisting of cattle ponds, ditches and natural field depressions (i.e. temporary ponds). During the day, a visual encounter survey was performed on the borders of each pond. Amphibians (larvae and adults) were also caught using a 4-mm mesh hand-net. Sampling duration varied from 10 to 30 min depending on pond area. From approximately 10:00 pm (sunset) to 12:00 pm, the ponds were visited again and a calling survey was performed. The ponds were approached quietly to a distance of approximately 20 m and, after waiting for 5 min, a 10-min auditory survey was conducted. Depending on the site typology, sometimes dip netting and visual survey with a 340-lumens torch was also performed to detect newts. All these methods were grouped as 'traditional survey methods' in the following analysis.

Fish surveys were performed from 2012 to 2014 by experienced fish biologists in different aquatic ecosystems (ponds, ditches, streams, lakes and rivers) in France and in the Netherlands. According to the ecosystem and the Water Framework Directive (European Council 2000), recommendations for fish monitoring, electrofishing and/or netting protocols (fyke, seine, gill) were implemented (Table S4, Supporting information).

eDNA metabarcoding survey. For small still-water ecosystems (ponds, ditches and mountain lakes), sampling was performed following the protocol described in Biggs *et al*. (2015). For running water (streams and rivers) or large ecosystems (lakes) where DNA is diluted in large water volumes, a new sampling strategy was tested. This strategy is based on the filtration of up to 100 L of water on site, using a filtration capsule (Enviroche HV 1 μm; Pall Corporation, Ann Arbor, MI, USA), sterile tubing and a peristaltic pump (1.67 L/min) for each sample. The number of filtrations per site depended on the size of the water ecosystem (Table S4, Supporting information).

All sampling kits were prepared in a dedicated 'DNA-free' room. This laboratory was equipped with positive air pressure, UV treatment and frequent air renewal. Laboratory personnel wore full protective clothing (disposable coveralls, hood, mask, laboratory-specific shoes, overshoes and two pair of gloves) that was put on in an airlock foyer before entering the processing room.

For ponds, ditches and mountain lake samples, DNA extraction was performed following the method described in Tréguier *et al*. (2014) in a room dedicated to the analysis of 'rare DNA' samples, with the same requirements as the 'DNA-free' room. For the other sites, the filtration capsules were transported at 4 °C to the DNA extraction room and then stored at −20 °C. Later, they were rinsed with 125 mL of resuspension buffer (50 mM Tris, 10 mM EDTA) and shaken by hand for 5 min. The buffer was emptied into three 50-mL tubes, and the procedure was repeated once. In total, 250 mL was retrieved in five tubes and was centrifuged at 15 000 ***g*** for 15 min, and the supernatant was removed with a sterile pipette, leaving 15 mL of liquid at the bottom of the tube. Subsequently, 33 mL of ethanol and 1.5 mL of 3 M sodium acetate were added to each tube. After manual shaking, the tube was immediately placed at −20 °C overnight. The remaining DNA extraction steps were performed following the protocol described in Tréguier *et al*. (2014). Negative extraction controls were performed at each DNA extraction session (2 for amphibian and 6 for fish samples). Those controls were amplified and sequenced in the same way and in parallel to the samples to monitor possible contaminations.

DNA amplification, sequencing and analysis of NGS results were performed following the protocol described in the 'In vitro validation of the designed primers' section, using 12 PCR replicates per DNA sample and 50 PCR cycles instead of 35. PCR-negative controls (with 12 replicates as well) were performed and sequenced to detect potential contamination. DNA from each filter capsule was extracted, amplified and sequenced separately, but sequencing results were pooled for each site. The three DNA extracts used by Thomsen *et al*. (2012b) were also pooled before the amplification, and then, they were amplified and analysed as described above.

The results of the traditional surveys were compared with those from the eDNA metabarcoding approach. For this purpose, a unit was considered as an identified taxon (family, genus or species) for eDNA metabarcoding and as a species for traditional surveys.

Site occupancy analysis

Detection probabilities were calculated for each amphibian species and both methods using site occupancy models (MacKenzie *et al*. 2002). All statistical analyses were performed using R.3.0.2 (R Core Team 2014) and the package 'UNMARKED' (Fiske & Chandler 2011). Site occupancy models were performed for the 39 amphibian sites sampled in 2012 2013 (Table S3, Supporting information) with the two approaches (traditional, eDNA metabarcoding) considered as temporal replicates and

the 39 water bodies as spatial replicates. Analyses were performed for all species merged into a single data set and for each species separately to assess whether or not the relative performance of each method differed among species and for all species combined. Models that systematically included differences in occupancy probability between species ('species' effect on occupancy probability in the model) were fitted. On the detection probability, four different models were fitted (constant, differences between species, differences between species and methods in addition, differences between species and methods in interaction). The relative performance of these five models was compared using Akaike information criteria (Burnham & Anderson 2004).

The number of sites per ecosystem type included in the fish survey was found to be insufficient; therefore, these data were excluded from this site occupancy analysis.

Historical data. To confirm the amphibian detectability observed in this study using conventional survey methods (one visit only), historical survey data for this group were examined (55 sites). From 1983 to 2013, in the Languedoc-Roussillon region, 983 water bodies were submitted to repeated amphibian inventories (at least one visit in two different years or two visits during the same year) using traditional methods (e.g. diurnal or nocturnal visual encounters, dip netting and calling surveys, with any combination of these methods). All data are stored in a regional database 'MALPOLON' hosted at CEFE UMR 5175, Biogeography and Vertebrate Ecology team, in Montpellier (France) and used for the regional atlas (e.g. Geniez & Cheylan 2012). Among these 983 inventoried water bodies, 55 sites were selected near the 30 water bodies surveyed in 2012 2013 for the comparative study (Site 1 30, Table S3, Supporting information). Those 55 sites were visited at least 10 times since 1995, and at least two amphibian species were detected per site. A classical site occupancy data set was constructed in which each visit was considered to be a temporal replicate and each water body/year as a spatial replicate. The data set was analysed for all species separately and constructed only using the constant model on occupancy and detection probability. Once the detection probability was obtained, the number of visits required to detect each species with a 95% probability was calculated.

Results

Design, in silico and in vitro validation of group-specific primers

Given the defined constraints, a single primer pair was identified by ECOPRIMERS for Batrachia and Teleostei, hereafter referred as 'batra' and 'teleo' (Table 1). They both amplify a short fragment of the 12S rRNA region. For each primer pair, Fig. 1 shows (i) sequence logos illustrating the match between the different primers and their target sequences for the considered target group, (ii) statistics about the mismatches for the target and nontarget groups and (iii) the length of the amplified fragment (excluding primers). The in silico analysis demonstrates that the designed primers show a high taxonomical coverage (i.e. the proportion of species amplified in the target group; 0.9928 for Batrachia and 0.9855 for Teleostei) and high taxonomical discrimination (Table S5, Supporting information). Furthermore, the 'teleo' primers perform better concerning taxonomical coverage and discrimination than the alternative primer pairs recently proposed in other metabarcoding studies (Thomsen *et al*. 2012b; Kelly *et al*. 2014b). For comparing the different primer pairs, both the *Bc* index and the taxonomic discrimination have to be considered. For example, despite having a better taxonomic discrimination (4295 different sequences for 3811 species) when compared to the 'teleo' primers, the primers 'Thomsen2' have a low coverage (*Bc*) leading to the nonamplification of more than half of the target fish sequences. The comparison between the 'Kelly' and the 'teleo' primers is also interesting: the 'teleo' primers not only have a slightly higher coverage, but also show a better taxonomic discrimination, despite amplifying a shorter fragment (Table S5, Supporting information).

The in vitro validation demonstrated that the two primer pairs successfully amplified the DNA of all amphibian and fish species in the pooled samples, confirming their universality for the studied taxonomic group. Species-level identification was possible for all analysed amphibians except for species of the genus *Pelophylax*, which are well known to hybridize (e.g. Plötner *et al*. 2008). For freshwater fish species, 81% were unambiguously identified to species level, 14% to genus level (i.e. *Carassius*, *Cottus*, *Leuciscus*, *Salvelinus*) and 5% to family level (i.e. *Chondrostoma nasus*, *Chondrostoma toxostoma* and *Telestes souffia)*. A taxon (species, genus or family) was 'unambiguously identified' if all the sequences associated with this taxon were not found in any other taxon. For marine fish species analysed, 68% were identified to species level and 32% to family level (i.e. *Hippoglossoides platessoides*, *Limanda limanda*, *Platichthys flesus*, *Pleuronectes platessa*, *Ammodytes marinus*, *Ammodytes tobianus*, *Hyperoplus lanceolatus*).

In situ validation of the eDNA metabarcoding approach

Sequencing data analysis. In total, 20 211 018 reads were obtained (7 655 154 for amphibian samples and

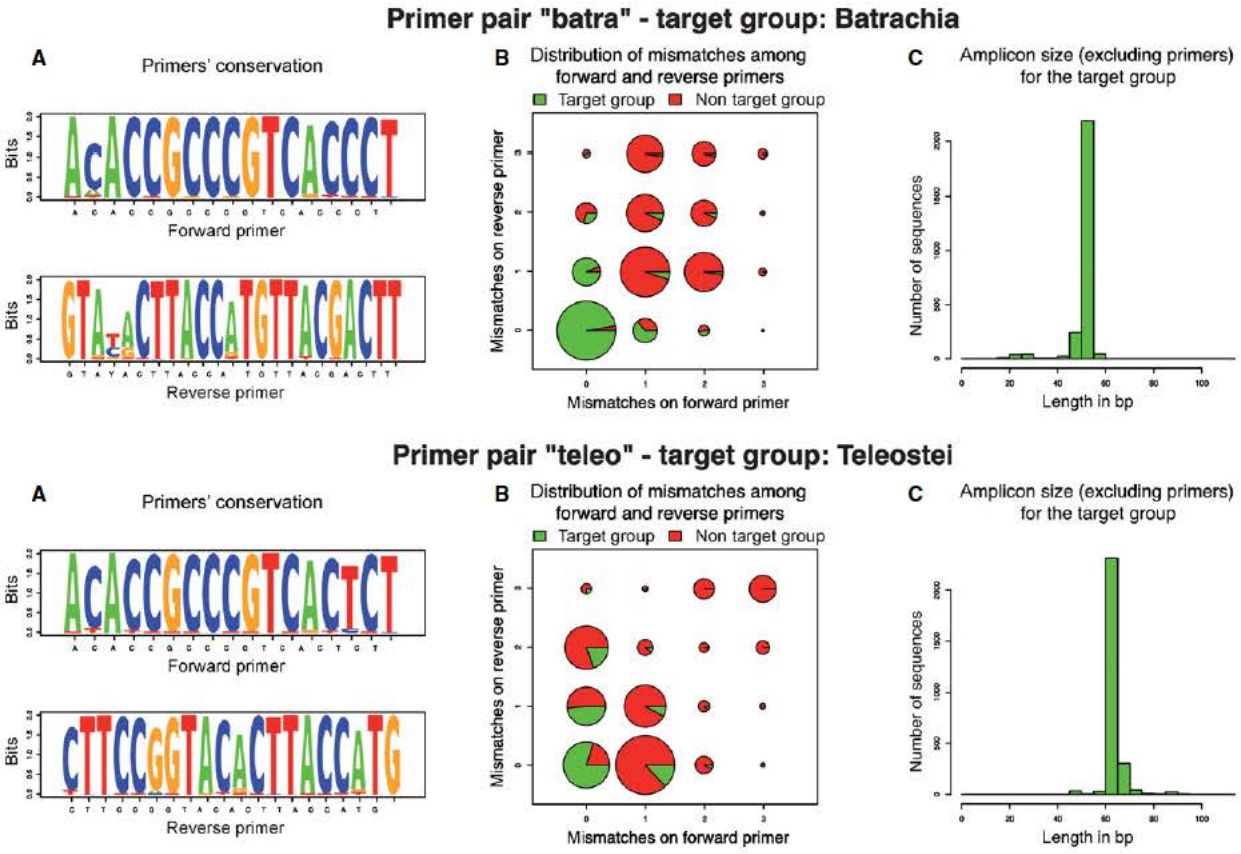

Fig. 1 Results of the in silico validation of the 'batra' and 'teleo' primer pairs. (A) sequence logo of the primer illustrating the quality of the match between the primer and its target sequence; (B) combined mismatch analysis of the forward and reverse primers (target group in green, nontarget group in red); and (C) length of the amplified fragment (excluding primers).

12 369 765 for fish samples), corresponding to an average of 253 626 per sample. After filtering, 11 994 052 sequences were retrieved (3 300 203 for amphibian samples and 8 527 193 for fish samples). After the filtering, all extraction and PCR controls analysed were negative. The raw number of reads obtained per sample, the number of reads retrieved after bioinformatics filtering per sample as well as the associated run number are shown in Table S6 (Supporting information).

Amphibians. Amphibian species were detected in all studied water bodies using the eDNA metabarcoding approach and in 89.2% (33/37) using traditional survey methods (Table S3, Supporting information). The eDNA approach missed the detection of two species (*Triturus marmoratus* and *Pelophylax* sp.) identified using the traditional survey, in one and two ponds, respectively. By contrast, in 64 cases, the traditional method missed the detection of species identified using eDNA (all species and water bodies together). The detection probability for each species using site occupancy models was always 1.00 using the eDNA metabarcoding approach (Fig. 2 and Table S7, Supporting information), except for *Pelophylax* sp. (*P* 0.89) and *T. marmoratus* (*P* 0.91). The detection probability for the traditional method was highly variable among species (*P* 0.20 1.00). Both methods show the same detection probability (*P* 1.00) for *Alytes obstetricans* and *Discoglossus pictus*, two species inhabiting a single water body. For the remaining nine species, the detection probability using the eDNA metabarcoding method was always higher (*P* 0.00013, Student's *t*-test). When all species were considered together, the best fitted model included the sampling method covariate in addition to the species effect. This additive effect demonstrates that, considering all sites, eDNA metabarcoding improves the detection of all species. Overall, the detection probability with eDNA metabarcoding was 0.97 (CI 0.90 0.99) vs. 0.58 (CI 0.50 0.63) for traditional surveys. Thus, four successive visits at a water body are required using traditional methods to achieve the same detection probability obtained with a single visit using the eDNA metabarcoding approach (1 (1 0.58)^4). The analysis of historical data using occupancy models demonstrated that the detection probabilities

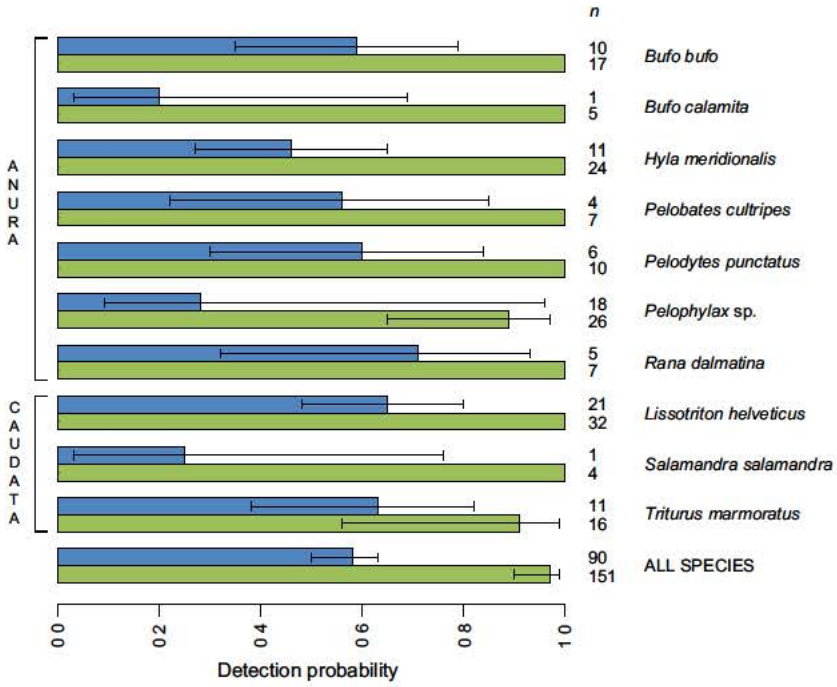

Fig. 2 Detection probabilities using the traditional survey (blue) and eDNA metabarcoding approach (green) for each amphibian species and for all species combined. Error bars represent the confi dence intervals of the analyses, and *n* represents the number of detections for each method and for each species or all species combined.

for each species during a single visit varied between 0.25 and 0.65 (Table S7, Supporting information). When the species detection probability is 0.25, 11 visits are required to reach a 95% chance of detection, whereas four visits are required when the species detection probability reaches 0.65. No amphibian species were detected using the eDNA metabarcoding and traditional approaches in the ponds that were not suitable for amphibians.

Fish. In 89% of the studied sites, the number of fish taxa detected using the eDNA metabarcoding approach was higher (~47% of the sites) or identical (~42% of the sites) to the number species detected using traditional methods. When a taxon was not detected using the genetic approach, only a few individuals of this taxon were generally retrieved using the traditional survey. However, on many occasions, eDNA detected the presence of the taxon even when it was present at low effectives (~66% when $n < 6$). Sometimes, the combination of the two methods led to a higher number of taxa detected (Fig. 3 and Tables S8 S15, Supporting information). When comparing our results with the ones from (Thomsen *et al*. 2012b), using the same DNA samples, 24 taxa were identified vs. 15, respectively (Fig. 4 and Table S15, Supporting information).

In the fishless mountain lakes, no fish species were detected using eDNA, and only *Salmo trutta* was detected in the two other control sites. Therefore, no false-positive records were generated using the eDNA metabarcoding approach.

Discussion

Biodiversity assessments are at the basis of numerous ecology and conservation issues. In the present study, a novel eDNA metabarcoding approach was tested for bony fish and amphibian monitoring in a variety of water systems. The reliability and efficiency of this new method was assessed using in silico, in vitro and in situ validations, which emphasized its strengths and limitations.

A powerful tool for aquatic species detection

Several advantages of the eDNA metabarcoding approach proposed here make it a powerful tool for ecological studies and aquatic biodiversity monitoring. First, species detectability was shown to be superior to traditional surveys: the number of species detected per site using eDNA was identical or higher than conventional survey methods in all cases for amphibians and in 89% of the cases for fish. Second, the monitoring effort required to infer the amphibian and fish communities is lower using the eDNA approach, allowing an increased monitoring efficiency. For instance, the fish sampling effort required to reach an identical number of detected species in lakes was 3 days using 88 gill nets vs. 4 h using the eDNA metabarcoding approach

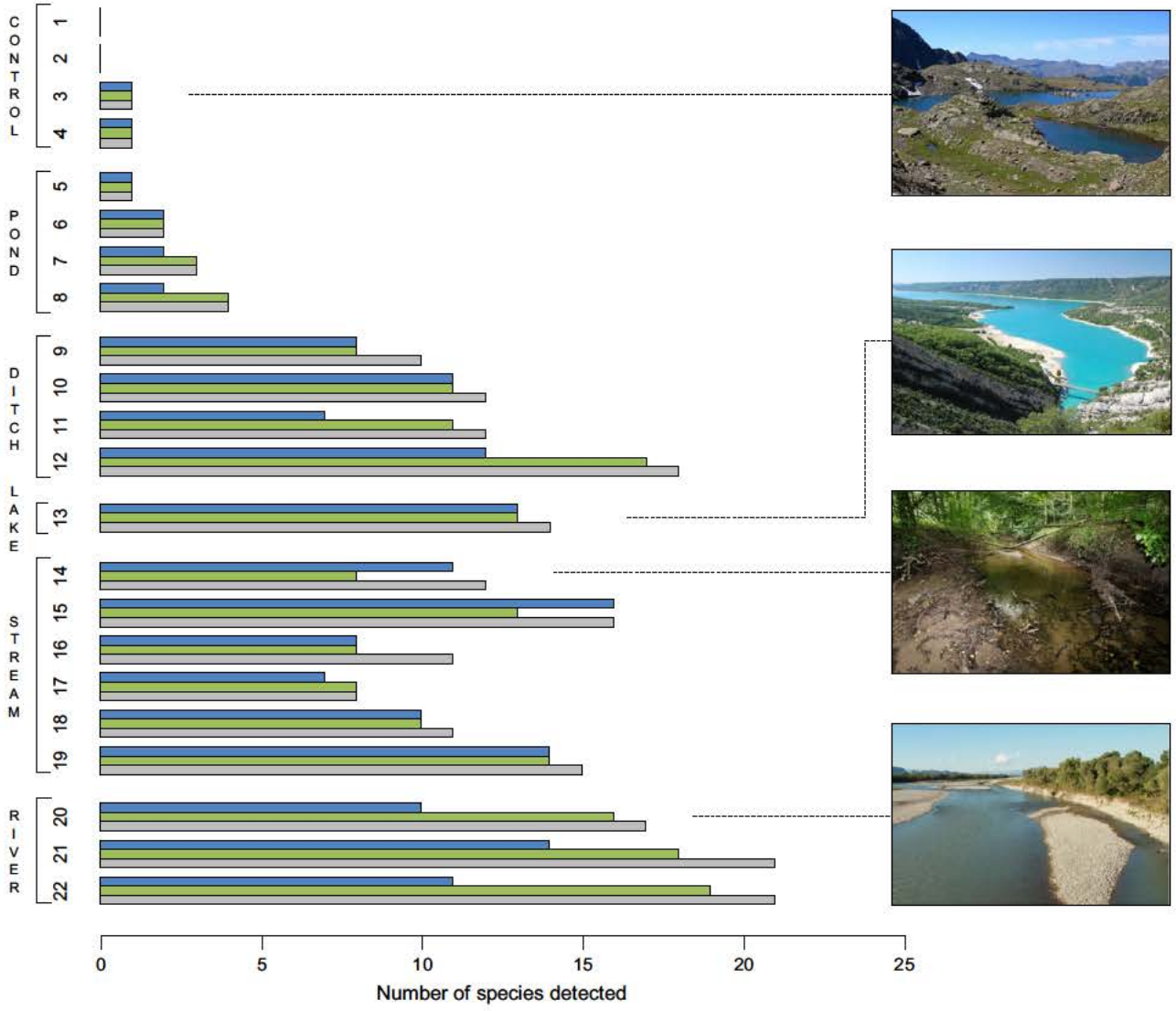

Fig. 3 Number of fish species detected in the seven freshwater ecosystems using historical data (sites 1 4) and traditional surveys data (sites 5 22) in blue and environmental DNA metabarcoding approach in green. Gray bars show the total number of species detected with the combination of both approaches.

(three water-filtration samples). For amphibian monitoring in Mediterranean ponds, four visits would be necessary to obtain similar detectability than with a single eDNA analysis. Third, this approach causes no disturbance to the ecosystem or to the target species, a top concern regarding conservation or restoration measures and for social and ethic valuation of scientific surveys of biodiversity. This noninvasive method, using sterile and disposable sampling material, also limits the risk of translocations of invasive alien species and pathogens during the field step. Fourth, it allows the detection of virtually all species of a target taxonomic group without a priori knowledge of their presence in the water body. This emphasizes the role of eDNA metabarcoding for environmental monitoring, including the early detection of alien species. Fifth, eDNA metabarcoding can also prove very advantageous in habitats where traditional methods cannot be implemented because of logistic constraints (e.g. low accessibility to the aquatic site) or have limited effectiveness (e.g. electric fishing cannot be performed in low conductivity waters; Allard *et al.* 2014). Sixth, as the primers designed in this study are universal for all amphibians and all bony fish species in the world, they can be used for global biodiversity assessment, allowing the large-scale temporal and spatial standardization of the method.

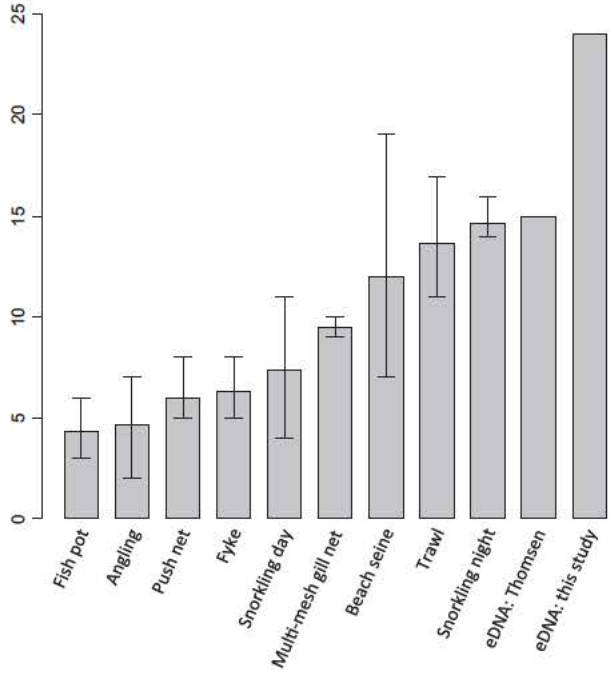

Fig. 4 Mean number of fish species in the marine coastal ecosystem recorded by nine different conventional survey methods in 2009, 2010 and 2011 (Thomsen *et al.* 2012b) and two environmental DNA metabarcoding methods (Thomsen *et al.* 2012b and this study). Error bars represent the standard deviation. Figure modified from Thomsen *et al.* (2012b).

Importance of the choice of eDNA metabarcoding markers

The importance of marker choice in eDNA metabarcoding has recently been emphasized (Coissac *et al.* 2012; Deagle *et al.* 2014). Because there is no ideal universal metabarcode (Riaz *et al.* 2011), marker choice should be specific to the target taxonomic group, and validation is required before application of the metabarcoding analysis in situ (Deagle *et al.* 2014). Mitochondrial rRNA genes have been recommended for animal identification because they have a similar taxonomic resolution as the COI marker and they present conserved regions that flank variable regions, which allows the design of primers with high-resolution power for the target taxonomic group (Deagle *et al.* 2014).

The comparison of results obtained using different metabarcodes (this study) also underlines the importance of selecting an appropriate marker. For example, the primers proposed by Kelly *et al.* (2014b) or Thomsen *et al.* (2012b) show a lower taxonomic coverage and resolution compared to the 'teleo' primer pair (Table S5, Supporting information). However, it should be noted that the Thomsen primers were specifically designed to estimate local biodiversity in Danish coastal waters. Recently, Miya *et al.* (2015) proposed an alternative universal primer pair for the amplification of fish species. The amplified fragment, located in the 12S gene, is nearly twice the size of the amplified fragment using 'teleo' primers for only a slightly higher taxonomic resolution. We also identified a nearly identical universal primer pair using the ECOPRIMERS software during the course of this study. However, to optimize the robustness of the amplification without losing significant taxonomic resolution, we opted for the shorter 'teleo' marker, and not for primers similar to those published by Miya *et al.* (2015).

The analysis performed using the same DNA extracts from Danish coastal marine fishes as used in Thomsen *et al.* (2012b) detected more fish species than the previous study (24 vs. 15). These results are not simply because of the primers used but more generally because of a better performance of the metabarcoding approach, including the number of PCR replicates (12 vs. 8), the sequencing technology used (Illumina Miseq vs. Roche 454), the sequencing depth (144 851 vs. 20 315 reads) and sequence data analysis.

Limitations of the eDNA metabarcoding approach

Population structure and size. As with every biodiversity assessment technique, the eDNA metabarcoding approach proposed here presents some limitations. First, the use of this approach does not permit information of size, developmental stage and sex of the target organisms to be obtained. Second, considering that mitochondrial DNA is maternally inherited in most cases (Giles *et al.* 1980), it is not possible to distinguish hybrids from their maternal species when using this eDNA approach. Third, the eDNA approach does not easily provide quantitative estimates for the surveyed species, whereas density and/or biomass information are often required to comply with legal statutes (e.g. European Water Framework Directive; European Council 2000). Some studies have demonstrated a relationship between the amount of eDNA and the biomass for some species (e.g. Takahara *et al.* 2012; Thomsen *et al.* 2012a), but further studies are required to evaluate this relationship in various environments, with a range of species densities on different taxonomic groups and especially when different life stages excrete DNA to the water column simultaneously (juveniles may produce more eDNA per biomass than adults because of increased metabolism during growth; Klymus *et al.* 2015). Therefore, the number of sequences obtained per taxon may not be interpreted as quantitative but rather as semi-quantitative and can differentiate between abundant and rare species (e.g. Pompanon *et al.* 2012; Kelly *et al.* 2014b). It should be noted, however, that a reliable quantification is also difficult using traditional

methods depending on the studied ecosystem characteristics because of sampling and identification bias for certain species (Miranda & Schramm 2000).

Taxonomic resolution and target group detectability. Taxonomic resolution does not always reach the species level. To improve taxonomic resolution, one approach is to complement the group-specific primers with one or several additional primer pairs specifically designed to amplify more discriminant genetic regions for families with many closely related species (e.g. Cyprinidae or Gadidae).

Additionally, the primers may amplify nontarget groups (Fig. 1). To overcome this issue, the use of blocking primers for abundant nontarget species (e.g. human) is recommended and a high sequencing depth is needed to detect rare species of the target group.

Dealing with errors. One of the main challenges associated with eDNA methods is the risk of false-positive and false-negative detections (Darling & Mahon 2011). In the present study, precautions were taken in the field and in the laboratory to prevent the occurrence of such errors (e.g. Dejean *et al.* 2012; Biggs *et al.* 2015), and no false positives were encountered at the six control sites. The reliability of the proposed method was demonstrated using in silico, in vitro and in situ validation tests. Both the set-up of the laboratory, which should employ similar rigorous standards as those described for ancient DNA laboratories (Cooper & Poinar 2000), and the experience of laboratory personnel working with rare and degraded DNA are key factors that strongly influence the reliability of eDNA results.

Another challenge is to properly address errors caused by DNA degradation or that are produced during the PCR and sequencing steps that can lead to taxonomic misidentification. The choice of reliable bioinformatics tools (Coissac *et al.* 2012) is crucial in order to deal with this kind of errors. The use of a carefully produced local and exhaustive reference database for the target group allows avoiding a high level of sequencing errors (Harris 2003), mislabelled species (Santos & Branco 2012) and a lack of sequences for the target genetic region in public databases. Additionally, the use of a local reference database allows increasing the taxonomic resolution of the assigned eDNA sequences (e.g. Taberlet *et al.* 2007).

To avoid PCR errors, the amplification step could be omitted, and the entire eDNA retrieved from the sample could be sequenced directly using a shotgun approach, as proposed by Mahon *et al.* (2014). However, in natural aquatic ecosystems, the proportion of target group DNA (e.g. fish, amphibians) is most often minimal compared with nontarget DNA (bacteria, phytoplankton, etc.). Therefore, to detect rare species in environmental samples, several millions of reads per sample would be required. However, even if NGS becomes more affordable with an ever-increasing number of produced sequences, the cost of shotgun sequencing eDNA currently remains too high to be used routinely in biodiversity monitoring.

Conclusion

This study demonstrated the performance and reliability of the novel eDNA metabarcoding approach proposed for ecological studies and for aquatic biodiversity. It also underlined that eDNA metabarcoding and traditional approaches can be considered complementary. If the objective is to obtain a list of species present in an aquatic ecosystem, including rare or secretive species, then eDNA metabarcoding is the most efficient tool. If additional data are required (development stage, sex, etc.), traditional survey methods should be used in parallel. In the present study, the eDNA metabarcoding method was validated for fish and amphibian surveys, and similar approaches could be developed for other taxonomic groups. In a time of global biodiversity loss and substantial environmental change, we firmly believe that the proposed approach represents a next-generation tool for efficient, precise and standardized monitoring of aquatic biodiversity in various ecosystems. This approach can therefore deliver key data for addressing many fundamental and applied research questions in ecology.

Acknowledgements

This work was financially supported by SPYGEN, ONEMA, EDF, Parc National du Mercantour, STOWA, Waterboard Bra bantse Delta, Waterboard Rijn en Ijssel, Waterboard Vallei en Veluwe, Waterboard de Dommel and Network Ecological Monitoring (NEM). We thank all persons and organisms who participated in the field work, development of the method and in the construction of the reference databases, in particu lar the Museum National d'Histoire Naturelle of Paris, the Aquarium of Aix les Bains, the CNRS Moulis, Biotope, Franco Andreone, Dick Bekker, Nicolas Bergher, Patrick Berrebi, Lau rence Blanc, Jeroen Bosveld, Pierre André Crochet, Arthur de Bruin, Jean François Dagand, Karin Didderen, Guillame Evanno, Christine Gazeau, Elsa Grand, Joran Janse, Stéphane Lefebvre, Jean Paul Lena, Jonathan Lereaux, Sébastien Manne, Jean Muratet, Sandro Parussatti, Martijn Schiphouwer, Jean Secondi, William Sremski, Alexandre Teynié and Wim van den Heuvel. We would like to thank Dr. Kristy Deiner and two anonymous reviewers for very useful and constructive comments.

Conflicts of interest

PT and EC are co-inventors of a patent on 'batra' and 'teleo' primers and on the use of the amplified fragment for identifying amphibians and fish species from environmental samples. This patent only restricts commercial applications and has no impact on the use of this method by academic researchers. AV, EB, CG, PJ and TD are research scientists in a private company, specialized on the use of eDNA for species detection.

References

Allard L, Grenouillet G, Khazraie K *et al.* (2014) Electrofishing efficiency in low conductivity neotropical streams: towards a non destructive fish sampling method. *Fisheries Management and Ecology*, **21**, 234 243.

Argillier C, Caussé S, Gevrey M *et al.* (2013) Development of a fish based index to assess the eutrophication status of Euro pean lakes. *Hydrobiologia*, **704**, 193 211.

Barnosky AD, Matzke N, Tomiya S *et al.* (2011) Has the Earth's sixth mass extinction already arrived? *Nature*, **471**, 51 57.

Bellemain E, Carlsen T, Brochmann C *et al.* (2010) ITS as an environmental DNA barcode for fungi: an in silico approach reveals potential PCR biases. *BMC Microbiology*, **10**, 189.

Beumer C, Martens P (2013) IUCN and perspectives on biodi versity conservation in a changing world. *Biodiversity and Conservation*, **22**, 3105 3120.

Biggs J, Ewald N, Valentini A *et al.* (2015) Using eDNA to develop a national citizen science based monitoring pro gramme for the great crested newt (*Triturus cristatus*). *Biologi cal Conservation*, **183**, 19 28.

Boyer F, Mercier C, Bonin A *et al.* (2015) obitools: a unix inspired software package for DNA metabarcoding. *Molecu lar Ecology Resources*, doi:10.1111/1755 0998.12428.

Burnham KP, Anderson DR (2004) Multimodel inference understanding AIC and BIC in model selection. *Sociological Methods & Research*, **33**, 261 304.

Coissac E, Riaz T, Puillandre N (2012) Bioinformatic challenges for DNA metabarcoding of plants and animals. *Molecular Ecology*, **21**, 1834 1847.

Cooper A, Poinar HN (2000) Ancient DNA: do it right or not at all. *Science*, **289**, 1139.

Copp GH, Oliver JM, Peňáz M, Roux AL (1991) Juvenile fishes as functional describers of fluvial ecosystem dynamics: appli cations on the river Rhône, France. *Regulated Rivers: Research & Management*, **6**, 135 145.

Crooks GE, Hon G, Chandonia J M, Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Research*, **14**, 1188 1190.

Daan N (2001) The IBTS database: a plea for quality control. *ICES CM*, **3**, 1 5.

Darling JA, Mahon AR (2011) From molecules to management: adopting DNA based methods for monitoring biological invasions in aquatic environments. *Environmental Research*, **111**, 978 988.

De Barba M, Miquel C, Boyer F *et al.* (2014) DNA metabarcod ing multiplexing and validation of data accuracy for diet assessment: application to omnivorous diet. *Molecular Ecology Resources*, **14**, 306 323.

Deagle BE, Jarman SN, Coissac E, Pompanon F, Taberlet P (2014) DNA metabarcoding and the cytochrome c oxidase subunit I marker: not a perfect match. *Biology Letters*, **10**, 20140562.

Deiner K, Walser J C, Machler E, Altermatt F (2015) Choice of capture and extraction methods affect detection of freshwater biodiversity from environmental DNA. *Biological Conserva tion*, **183**, 53 63.

Dejean T, Valentini A, Miquel C *et al.* (2012) Improved detec tion of an alien invasive species through environmental DNA barcoding: the example of the American bullfrog *Litho bates catesbeianus*. *Journal of Applied Ecology*, **49**, 953 959.

European Council (2000) *Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 Establishing a Framework for Community Action in the Field of Water Policy*. Office for official publications of the European Communities, Brussels.

Evans NT, Olds BP, Turner CR *et al.* (2015) Quantification of mesocosm fish and amphibian species diversity via eDNA metabarcoding. *Molecular Ecology Resources*, doi:10.1111/ 1755 0998.12433.

Ficetola GT, Coissac E, Zundel S *et al.* (2010) An in silico approach for the evaluation of DNA barcodes. *BMC Geno mics*, **11**, 434.

Fiske I, Chandler R (2011) Unmarked: an R package for fitting hierarchical models of wildlife occurrence and Abundance. *Journal of Statistical Software*, **43**, 1 23.

Geniez P, Cheylan M (2012) *Les amphibiens et les reptiles du Languedoc Roussillon et régions limitrophes Atlas biogéographique*. Biotope Editions, Mèze, Paris.

Giles RE, Blanc H, Cann HM, Wallace DC (1980) Maternal inheritance of human mitochondrial DNA. *Proceedings of the National Academy of Sciences, USA*, **77**, 6715 6719.

Harris JD (2003) Can you bank on GenBank? *Trends in Ecology & Evolution*, **18**, 317 319.

Hubert WA, Pope KL, Dettmers JM (2012) Passive capture techniques. In: *Fisheries Techniques* (eds Zale AV, Parrish DL, Sutton TM), pp. 223 265. American Fisheries Society, Bethesda, Maryland.

Hudy M (1985) Rainbow trout and brook trout mortality from high voltage AC electrofishing in a controlled environment. *North American Journal of Fisheries Management*, **5**, 475 479.

Kelly RP, Port JA, Yamahara KM *et al.* (2014a) Harnessing DNA to improve environmental management. *Science*, **344**, 1455 1456.

Kelly RP, Port JA, Yamahara KM, Crowder LB (2014b) Using environmental DNA to census marine fishes in a large meso cosm. *PLoS ONE*, **9**, e86175.

Kéry M, Schmidt BR (2008) Imperfect detection and its conse quences for monitoring for conservation. *Community Ecology*, **9**, 207 216.

Klymus KE, Richter CA, Chapman DC, Paukert C (2015) Quan tification of eDNA shedding rates from invasive bighead carp *Hypophthalmichthys nobilis* and silver carp *Hypoph thalmichthys molitrix*. *Biological Conservation*, **183**, 77 84.

MacKenzie DI, Nichols JD, Lachman GB *et al.* (2002) Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, **83**, 2248 2255.

Mahon AR, Nathan LR, Jerde CL (2014) Meta genomic surveillance of invasive species in the bait trade. *Conservation Genetics Resources*, **6**, 563 567.

Miranda LE, Schramm HL (2000) Selecting gear for monitoring fish assemblages. In: *Management and Ecology of River Fisheries* (ed. Cowx IG), pp. 3 13. Blackwell Science Ltd., Oxford.
Miya M, Sato Y, Fukunaga T *et al.* (2015) MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: detection of more than 230 subtropical marine species. *Royal Society Open Science*, **2**, 150088.
Monastersky R (2014) Biodiversity: life a status report. *Nature*, **516**, 158 161.
Pimm SL, Jenkins CN, Abell R *et al.* (2014) The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, **344**, 1246752.
Plotner J, Uzzell T, Beerli P *et al.* (2008) Widespread unidirectional transfer of mitochondrial DNA: a case in western Palaearctic water frogs. *Journal of Evolutionary Biology*, **21**, 668 681.
Pompanon F, Deagle BE, Symondson WOC *et al.* (2012) Who is eating what: diet assessment using next generation sequencing. *Molecular Ecology*, **21**, 1931 1950.
R Core Team (2014) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
Riaz T, Shehzad W, Viari A *et al.* (2011) ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research*, **39**, e145.
Roset N, Grenouillet G, Goffaux D, Pont D, Kestemont P (2007) A review of existing fish assemblage indicators and methodologies. *Fisheries Management and Ecology*, **14**, 393 405.
Santos AM, Branco M (2012) The quality of name based species records in databases. *Trends in Ecology & Evolution*, **27**, 6 7.
Schnell IB, Bohmann K, Gilbert MTP (2015) Tag jumps illuminated reducing sequence to sample misidentifications in metabarcoding studies. *Molecular Ecology Resources*, **15**, 1289 1303.
Stuart SN, Chanson JS, Cox NA *et al.* (2004) Status and trends of amphibian declines and extinctions worldwide. *Science*, **306**, 1783 1786.
Taberlet P, Coissac E, Pompanon F *et al.* (2007) Power and limitations of the chloroplast *trn*L (UAA) intron for plant DNA barcoding. *Nucleic Acids Research*, **35**, e14.
Taberlet P, Coissac E, Hajibabaei M, Rieseberg LH (2012) Environmental DNA. *Molecular Ecology*, **21**, 1789 1793.
Takahara T, Minamoto T, Yamanaka H, Doi H, Kawabata Z (2012) Estimation of fish biomass using environmental DNA. *PLoS ONE*, **7**, e35868.
Tanadini LG, Schmidt BR (2011) Population size influences amphibian detection probability: implications for biodiversity monitoring programs. *PLoS ONE*, **6**, e28244.
Thomsen PF, Willerslev E (2015) Environmental DNA an emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, **183**, 4 18.
Thomsen PF, Kielgast J, Iversen LL *et al.* (2012a) Monitoring endangered freshwater biodiversity using environmental DNA. *Molecular Ecology*, **21**, 2565 2573.
Thomsen PF, Kielgast J, Iversen LL *et al.* (2012b) Detection of a diverse marine fish fauna using environmental DNA from seawater samples. *PLoS ONE*, **7**, e41732.
Tréguier A, Paillisson J M, Dejean T *et al.* (2014) Environmental DNA surveillance for invertebrate species: advantages and technical limitations to detect invasive crayfish *Procambarus clarkii* in freshwater ponds. *Journal of Applied Ecology*, **51**, 871 879.
Wheeler QD, Raven PH, Wilson EO (2004) Taxonomy: impediment or expedient? *Science*, **303**, 285.
WWF (2014) Living Planet Report 2014: Summary. (eds McLellan R, Iyengar L, Jeffries B, Oerlemans N), WWF, Gland, Switzerland.
Yoccoz NG, Bråthen KA, Gielly L *et al.* (2012) DNA from soil mirrors plant taxonomic and growth form diversity. *Molecular Ecology*, **21**, 3647 3655.

A.V., P.T., C.M., R.C., J.H., P.F.T., N.P., N.R., D.P., C.A., J.M.B., T.P., A.J.C., A.O., M.A., M.L.B. and T.D. designed the experiment. P.T. and E.C. designed the 'batra' and 'teleo' primers. P.T., E.C. and F.B. performed the in silico validation. A.V., C.M., P.F.T., E.B., N.P., N.R., P.G., T.P., A.J.C., P.R.M. and T.D. built the local sequences database for amphibians and fish. A.V. and T.D. performed the in vitro validation. C.M., R.C., J.H., N.R., T.P., A.O., M.A., P.J. and T.D. made the traditional survey analysis and A.V., R.C., J.H., P.F.T., C.G., P.J., T.D. the eDNA metabarcoding analysis. A.B. performed the occupancy model analysis. A.V., P.T., C.M., J.H., P.F.T., E.B., A.B., N.P., G.H.C., E.W. and T.D. wrote the manuscript. G.H.C., P.G., D.P., C.A., J.M.B., A.J.C., M.L.B., P.R.M. and E.W. provided guidance during the study.

Data accessibility

Sequences for the reference databases are uploaded as online supporting information (Tables S1 and S2, Supporting information) and on Dryad doi: 10.5061/dryad.jm58p. All Illumina raw sequences data and OTU tables are available on Dryad doi:10.5061/dryad.jm58p.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Amphibian tissue samples used for the reference database construction.

Table S2 Fish tissue samples used for the reference database construction.

Table S3 Geographical coordinates, sampling date and results of traditional and eDNA metabarcoding surveys for amphibians.

Table S4 Location, site characteristics and sampling methods used for eDNA metabarcoding and traditional surveys for the fish comparative study.

Table S5 In silico assessment of different primer pairs target ing Batrachia and Teleostei.

Table S6 Number of reads obtained from the NGS runs per sample before and after bioinformatic filtering.

Table S7 Number of detection events of amphibian species using eDNA metabarcoding and traditional surveys; species detection probability with confidence intervals between brack ets for eDNA metabarcoding, traditional surveys and historical data and number of visits required to achieve a 95% chance of species detection.

Table S8 Results from the eDNA metabarcoding survey and historical data for fish in control sites (sites 1 4).

Table S9 Results from eDNA metabarcoding and traditional surveys for fish in pond ecosystems (sites 5 8).

Table S10 Results from eDNA metabarcoding and traditional surveys for fish in ditch ecosystems (sites 9 12).

Table S11 Results from eDNA metabarcoding and traditional surveys for fish in lake ecosystems (site 13).

Table S12 Results from eDNA metabarcoding and traditional surveys for fish in stream ecosystems in the Netherlands (sites 14 17).

Table S13 Results from eDNA metabarcoding and traditional surveys for fish in stream ecosystems in France (sites 18 19).

Table S14 Results from eDNA metabarcoding and traditional surveys for fish in river ecosystems (sites 20 23).

Table S15 Results from eDNA metabarcoding for fish in the marine ecosystem (site 23).