Accuracy, limitations and cost-efficiency of eDNA-based community survey in tropical frogs

Miklós Bálint[1,2], Carsten Nowak[1,2], Orsolya Márton[1,3], Steffen U. Pauls[1,2], Claudia Wittwer[1], Jose Luis Aramayo B.[4], Arne Schulze[5], Thierry Chambert[6], Berardino Cocchiararo[1,2], Martin Jansen[1]

[1]Senckenberg Research Institute and Natural History Museum Frankfurt, Senckenberganlage 25, 60325 Frankfurt, Germany

[2]LOEWE Centre for Translational Biodiversity Genomics (LOEWE-TBG), Senckenberganlage 25, 60325 Frankfurt, Germany

[3]Institute for Soil Sciences and Agricultural Chemistry, Centre for Agricultural Research, Hungarian Academy of Sciences, Herman Otto str. 15, H-1022 Budapest, Hungary

[4]Museo Historia Natural Noel Kempff Mercado (Universidad Autónoma Gabriel Rene Moreno), Av. Irala 565, Santa Cruz de la Sierra, Bolivia

[5]Hessisches Landesmuseum Darmstadt (HLMD), Friedensplatz 1, 64283 Darmstadt, Germany

[6]Pennsylvania State University, Department of Ecosystem Science and Management, University Park, PA 16802, USA

Abstract

Rapid environmental change in highly biodiverse tropical regions demands efficient biomonitoring programs. While existing metrics of species diversity and community composition rely on encounter-based survey data, eDNA recently emerged as alternative approach. Costs and ecological value of eDNA-based methods have rarely been evaluated in tropical regions, where high species richness is accompanied by high functional diversity (e.g. the use of different microhabitats by different species and life-stages). We first tested whether estimation of tropical frogs' community structure derived from eDNA data is compatible with expert field assessments. Next we evaluated whether eDNA is a financially viable solution for biodiversity monitoring in tropical regions. We applied eDNA metabarcoding to investigate frog species occurrence in five ponds in the Chiquitano dry forest region in Bolivia and compared our data with a simultaneous visual and audio encounter survey (VAES). We found that taxon lists and community structure generated with eDNA and VAES correspond closely, and most deviations are attributable to different species' life histories. Cost efficiency of eDNA surveys was mostly influenced by the richness of local fauna and the number of surveyed sites: VAES may be less costly in low-diversity regions, but eDNA quickly becomes more cost-efficient in high-diversity regions with many sites sampled. The results highlight that eDNA is suitable for large-scale biodiversity surveys in high-diversity areas if life history is considered, and certain precautions in sampling, genetic analyses and data interpretation are taken. We anticipate that spatially extensive, standardized eDNA biodiversity surveys will quickly emerge in the future.

Keywords

Amphibians, metabarcoding, tropical biodiversity, conservation, community ecology, cost comparison

Introduction

Improvements on most biodiversity loss indicators lag behind the 20 “Aichi Biodiversity Targets” (UNEP 2016) that aim to reduce the decline of biodiversity by 2020 (Tittensor *et al.* 2014). An important component of the biodiversity crisis is the extinction of species. Based on current trends in mammals, birds, reptiles and amphibians, it has been projected that the biodiversity crisis may lead to the 6th Mass Extinction over the next three centuries if all threatened species go extinct (Barnosky *et al.* 2011). Current rates of extinction may even be much higher if one considers the extinction that likely occurred during the last few decades-centuries, but went unnoticed because the now-extinct species had small ranges, were never described or only described on the eve of their extinction (Pimm *et al.* 2014; Lees & Pimm 2015). It is however difficult to assess which species are endangered and to what extent. First, most taxa remain poorly described: in some highly diverse regions many species will likely go extinct before they are discovered (Costello *et al.* 2013; Lees & Pimm 2015). Second, cryptic genetic diversity is common within morphospecies (Pfenninger & Schwenk 2007; Pauls *et al.* 2013), and global change may impact cryptic diversity more severely than morphospecies (Bálint *et al.* 2011). Third, data on population-level trends is scarce, even for well-known species (Butchart *et al.* 2010). Better population-level biodiversity data is thus urgently needed to 1) understand biodiversity patterns and extinction threats, 2) improve forecasting abilities about future biodiversity, and 3) improve humanity’s responses to the challenges of biodiversity loss. This data is crucial in times when conservation action is increasingly demanded by society (Tittensor *et al.* 2014).

The importance of internationally coordinated, standardized biodiversity data collection is long recognized both in science and in conservation (Henry *et al.* 2008). This is particularly true for the most biodiverse areas. The tropics are generally underrepresented in ecological studies

(Clarke *et al.* 2017; Stroud & Feeley 2017). In addition, encounter-based data collection is logistically challenging since it needs considerable funds to bring enough specialists for different organismic groups to remote areas, thus insufficient funds and expertise often limit comprehensive surveys. Indirect species records through environmental DNA are increasingly heralded as an alternative to encounter-based surveys (Thomsen & Willerslev 2015; Pedersen *et al.* 2015). eDNA also facilitates the standardization of biodiversity surveys at both regional and global scales since community composition data can be obtained by high-throughput-sequencing of standardized, taxonomically informative marker genes (metabarcoding) (Taberlet *et al.* 2012a; Cristescu 2014). Aquatic or semiaquatic vertebrates such as frogs and other amphibians or fish have been early targets of eDNA based studies, either focusing on the detection of single species (e.g. Ficetola *et al.* 2008; Goldberg et *al.* 2011; Jerde *et al.* 2011; Thomsen *et al.* 2012a) or entire communities (e.g. Thomsen *et al.* 2012b; Valentini *et al.* 2016; Shaw *et al.* 2016; Yamamoto *et al.* 2017). The use of next-generation sequencing approaches led to a boost in data acquisition (Taberlet *et al.* 2012b) and is considered to make important contributions to biodiversity research (Rees *et al.* 2014; Bohmann *et al.* 2014). eDNA-based metabarcoding may present one of several tools needed to globally coordinate initiatives for ecosystem monitoring and sustainable management (Bush *et al.* 2017; Schmeller *et al.* 2017).

In this study we evaluate whether eDNA metabarcoding is suitable for inventories of frogs, a group with particular high species-diversity in tropical regions. Frogs and other amphibians are sentinel victims of the biodiversity crisis: more than one-third of the approximately 7500 described species are endangered (Stuart *et al.* 2004; Bishop *et al.* 2012; Whittaker *et al.* 2013). Frogs are also known for being a highly diverse, but incompletely described taxon, especially in the tropics (Ferrão *et al.* 2016; Caminer *et al.* 2017). Many "widespread" morphospecies harbor considerable cryptic genetic diversity and are better considered complexes of closely related species with much smaller ranges (Fouquet *et al.* 2007;

92Gehara *et al.* 2013, 2014; Ortega-Andrade *et al.* 2015). Efficient implementation of amphibian
93conservation measures (e.g. the prioritization of areas for conservation, or informing society and
94stakeholders about conservation needs) are only possible with geographically broad-scaled
95fine-grain, taxonomically well-resolved faunistic data, but our current understanding of present
96and future amphibian biodiversity is often based on rare, spatially and temporally scattered
97observations of phenotypically defined taxa.

98 First efforts have been made to test the suitability of eDNA for the survey of tropical frog
99biodiversity (Lopes *et al.* 2016), but important practical aspects remain unaddressed. First, it is
100not clear which fraction of the local species pool is represented by amphibian eDNA in tropical
101water bodies. Existing comparisons of encounter-based surveys and eDNA either do not include
102non-adult life stages, or they use already compiled fauna lists for the evaluation of eDNA
103performance without consideration of life history traits or behavioral aspects at the moment of
104sampling. However, temporal and spatial patterns of microhabitat use by frogs is largely
105species-specific and influenced by phenology and reproduction modes (e.g. Duellmann & Pyles
1061983, Haddad & Prado 2005, Wells 2010), which can induce strong biases in biodiversity data
107(Petitot et al. 2014). Second, the degree to which aquatic eDNA can provide accurate
108assessments of community structure remains largely unevaluated. Most studies to date have
109only investigated the correspondence between encounter-based and metabarcoded taxon lists
110(Miya *et al.* 2015; Valentini *et al.* 2016), although community structure assessments themselves
111may be of higher importance for many applications (Ji *et al.* 2013; Elbrecht *et al.* 2017). Third, it
112is not clear whether, and under what conditions eDNA is financially efficient since comparisons
113are lacking (Lopes *et al.* 2016), although these comparisons are essential for deciding on data
114collection strategies.

115 Here we address whether 1) detectability of tropical amphibians with eDNA is linked to
116species' life history, and 2) eDNA sampling provides accurate data for the characterization of

117community assembly. Finally, we present a framework for cost comparisons between
118encounter- and eDNA-based biodiversity survey that may be adapted to other systems beyond
119amphibians. We compare the results of long- and short-term encounter-based field surveys, and
120an eDNA survey of tropical amphibians in a well-characterized high-diversity area in Bolivia.

121

122

5

Materials and methods

The study area is located the Chiquitano region of Bolivia, a forest-savanna ecotone between Amazon, Cerrado and Gran Chaco in a transition zone among humid and dry forests that are special in regard to their taxonomic and functional diversity (Castro *et al.* 1999). The region contains the largest intact, old-growth block of seasonally dry tropical forests in South America (Miles *et al.* 2006; Power *et al.* 2016). Samples were collected from ponds in the vicinity of the Biological Station “Centro de Investigaciones Ecológicas Chiquitos” on the San Sebastián cattle ranch (S16.3622, W62.00225, 500 m a.s.l.), 24 km south of the town of Concepción, Province of Ñuflo de Chávez, Santa Cruz Department, Bolivia. A description of the area is given by Schulze *et al.* (2009). This area is well characterized with respect to amphibians, including both larvae and adults (Schulze *et al.* 2015). Intensive long-term assessments have resulted in the detection of 45 frog species in the area (e.g. Jansen 2009; Schulze *et al.* 2009, 2015; Jansen *et al.* 2011), as well as the discovery of cryptic intraspecific diversity (e.g. Jansen *et al*. 2011; Jansen *et al.* 2016).

We sampled five ponds (T1 – T5) for this study between 19 and 23 November 2014 (see also section “Site and ponds” in the Supplemental information). At these ponds, 35 species have been recorded in previous long-term surveys. Water samples were obtained from three sampling points at each pond (Fig. 1): These water samples consisted of four liters of water collected into two two-liter silicon bottles. One sample of 2 x 2 liters was taken at each sampling point of ponds T1 - T4. Three samples of 2 x 2 liters were taken on each sampling points of pond T5 to check for sampling variation, i.e. whether replication in the sampling records more variation in community composition compared to replication in PCR steps. We also sampled and filtered water from two field controls: Tap water from the station from a covered well, and water

6

from an aquarium with tadpoles of *Leptodactylus vastus* at the station. The water samples were filtered at the station immediately after collection.

Samples were filtrated with a vacuum filtration unit (Duran Group, Wertheim, Germany), which was connected to a vacuum membrane pump (Laboport, both Carl Roth GmbH + Co. KG, Karlsruhe, Germany). Filtration was performed until complete clogging of the filters. For each sample, the water from the two bottles was filtered separately on two glass-fiber filters (GFF, pore size 2 µm). One of the filters was preserved in CTAB (A), the other filter was dried immediately (B). Our intention here was to evaluate if filter preservation influences species detection, since the transportation of dry filters may be considerably simpler compared to wet filters. The flow-through from the two GFF filtrations was combined and filtered with a nylon filter until clogging (NF, pore size 0.2 µm, Millipore, Merck, Darmstadt, Germany). This filter was immediately dried after filtration (C). We intended to evaluate if the larger pore-size filters sufficiently capture anuran eDNA, or finer pore filters allow for additional species detection. To avoid cross-contamination we changed gloves with every new sample and disinfected hands with ethanol (96%), as well we cleaned filtration unit with ethanol (96%) after each pond or site.

Combined visual and audio encounter surveys (VAES, Zimmerman 1994), and a tadpole survey (TS, Schulze *et al.* 2015) were conducted by experienced observers (MJ and AS) to generate a presence/absence matrix of species for each pond. Frogs surveys were performed at night (between 21:00 and 00:30 hours) during 0.5 - 1 hrs transect walks along the ponds to detect and identify specimens in vegetation, in water and on the ground around ponds. Detections occurred from either visual (using flashlights and headlamps) or auditory species identification. Tadpoles were sampled with dip nets once during daytime and once at night in each pond as well as the riparian vegetation.

eDNA samples were processed in Germany in a laboratory dedicated to the pre-PCR handling of environmental DNA samples. Several working routines have been implemented in the laboratory to avoid contamination of samples and reactions, including spatial separation from the other DNA facilities (separate room), strict decontamination protocols using UV light and bleach, physical separation of extraction and PCR work spaces, automated extraction, and restricted access to staff trained in the handling and analysis of forensic and environmental samples.

DNA was extracted from GFF samples (A, B, see above) with a CTAB chloroform extraction method according to Strand *et al.* (2014) and Wittwer et al. (2017). Dried nylon filters (C) were extracted with DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) following Thomsen *et al.* (2012b). Negative controls were included in all extraction sessions (see also DNA extraction in the Supplemental information). The barcode amplification targeted mitochondrial 16S ribosomal DNA. We designed primers (see Primer design in Supplemental information for details on design and primer tests) for a 150 bp fragment with the reference database of local species (sequences from 159 specimens, Jansen *et al.* 2011; Schulze *et al.* 2015). The primers were tested on a subset of 12 species. PCRs were run on 96-well plates in 15-µl reaction volumes with a touchdown protocol, with four technical PCR replicates per sample (see PCR setup in the Supplemental information for PCR protocol and cycling conditions). We included four negative PCR controls (ultra-sterile water), four extraction blanks, and two positive controls on each plate. Positive controls consisted of 1) equimolar positives where DNA of the twelve frog species (primer development) was pooled in equimolar concentration (5 ng/µl), and 2) non-equimolar positives where a concentration series was used by diluting the 5 ng/µl templates 1x to 512x dilution. PCR replicates were individually labeled with multiplexing indices designed by Kozarewa and Turner (2011). PCR products were purified with Agencourt AMPure XP SPRI magnetic beads (Beckman Coulter, Brea, CA), pooled in

equal volumes, and paired-end-sequenced (2 x 150 bp) on an Illumina NextSeq 500 sequencer at StarSEQ GmbH (Mainz, Germany).

The bioinformatic pipeline for sequence processing mostly relied on OBITools v1.2.0 (Boyer *et al.* 2016, see Sequence processing in the Supplemental information for details). We performed two reference-based taxonomic assignments, first with a custom database of 16S sequences of 159 specimens from the regional amphibian fauna (16S_custom), and second with all 16S sequences found in the EMBL (release 125, 16S_EMBL). Further filtering was done in R (R Core Team 2017) with a script supplied on GitHub: https://github.com/MikiBalint/amphibian-eDNA. This includes an ordination of all PCR replicates, negative and positive controls. This ordination was used to identify outlier replicates of samples from different ponds (Supplemental Fig. S1).

We followed recommendations for the statistical analyses of marker gene community data (Bálint *et al.* 2016). eDNA faunistic differences among the five ponds were visualized with a latent variable model-based ordination in R ('boral', Hui *et al.* 2015), and tested with multispecies generalized linear models (GLMs in 'mvabund', Wang *et al.* 2012). We used read abundances as input data for the boral ordination and the multispecies GLMs since presence-absences inferred from read abundances may overestimate the importance of rare sequence variants on community structure (Bálint *et al.* 2016). Different filtrates were included as distinct samples into the analyses since we wanted to see the variation among closely related samples (e.g. the three filtrates taken at the same point of a pond) *versus* the variation among distinct samples (e.g. originating from different ponds). We assumed a negative binomial distribution for both boral and mvabund analyses since read abundances are overdispersed counts, with a strong uniform prior on the dispersion parameter in boral (0-3). The surveyed ponds were markedly different in size, vegetation, water depth, etc, thus we considered pond identity as a good predictor of community composition. The pond effect was evaluated with analysis of

219variance (likelihood-ratio test, PIT-trap resampling, 1 000 bootstrap replicates). We performed a
220model-based ordination also for the VAES presence/absence data (since no abundance data
221was collected during the VAES), and then used a Procrustean superimposition (Peres-Neto &
222Jackson 2001) to evaluate how the VAES-based ordination of ponds matches the ordination of
223centroids of eDNA samples. We compared the efficiency of filter preservation (CTAB or dried)
224on the successful detection of species with a site occupancy model (MacKenzie *et al.* 2002;
225Bailey *et al.* 2014), implemented in the R package 'unmarked' (Fiske *et al.* 2011). We used the
226single season false-positive occupancy model developed by Miller *et al.* (2011, for details see
227Site occupancy models in Supplemental information).

228 For cost comparisons, we considered a typical eDNA survey scenario: samples are
229collected in the field and later processed in a dedicated laboratory. For the encounter-based
230survey we considered a scenario with a similar separation of the fieldwork and species
231identification: species records (audio or visual) are collected by a field biologist, and later
232identified or verified by an expert in the office/laboratory. The parameters in our cost models are
233an approximation of the variables involved in the present study and involve a learning effect in
234the efficiency of the taxonomic expert (see Cost comparison in the Supplemental information for
235details). The sampling and identification costs of VAES are dependent both on the number of
236sampling sites, and the number of species since each species needs to be recorded. During the
237eDNA survey samples are collected by a field biologist, analyzed in a lab and sequenced by an
238external provider. The costs of eDNA sampling and identification depend on the number of sites,
239but not on the number of species. We kept some cost factors constant: the costs for training the
240frog taxonomic expert and the VAES observers, the costs for building up the eDNA
241metabarcoding facilities (clean rooms and equipment to perform DNA manipulations, except
242sequencing), and the databases necessary for the sequence assignment. We assume that
243travel costs are the same for the two survey types, and that the time necessary to walk between

244frog observations and eDNA sampling points is the same. All model parameters and

245calculations are accessible on FigShare (Bálint *et al.* 2017).

11

Results

The sequencing resulted in 12 742 273 read pairs (deposited in ENA as PRJEB22113**)**, of which 9 479 299 were identified as complete 16S amplicons. De-replication of the reads (see Supplemental Information: Sequence processing for more information on the steps) resulted in 631 003 unique sequences variants. Only 22 706 of these variants were represented by at least 10 reads and retained for further processing. The sequence cleanup resulted in 14 442 high quality sequence variants, 13 497 coming from the present experiment. Of these, 4 805 sequence variants were taxonomically assigned by the *ecotag* command with the 16S_custom database, and 8 692 with the 16S_EMBL database. The *ecotag* outputs are accessible on FigShare (https://doi.org/10.6084/m9.figshare.5099842.v5), and can be combined and simplified with the R script provided on GitHub (https://github.com/MikiBalint/amphibian-eDNA). The assigned sequence variants represented 8 011 631 sequences. After bionformatic processing with OBITools 561 of these sequence variants were identified as 'head' sequences, i.e. sequences that have no variants more abundant than a predefined percentage of their own count, here 5% (see also Sequence processing in the Supplemental information) in at least one sample (6 965 866 reads). The original read numbers were re-assigned to these head sequence variants, and the read numbers were used in downstream analyses. Several of these 561 head sequence variants were found also in the negative controls. After removing the reads assigned to these, the final sample - sequence variant abundance matrix contained 5 815 014 sequences. These belonged to several groups: frogs (2 158 534), fish (1 692 613), insects (304 059), mammals (14 006), birds (967), and bacteria and other groups (1 063 156).

Read numbers assigned to frog species varied considerably among the spatial replicates of the eDNA samples (see species_abundance_matrix.csv – zipped - on FigShare: https://figshare.com/articles/_/5099842). Numbers of frog species detected by both methods,

eDNA and VAES, varied between 3 and 18 per pond (Table1). Several of the species present at the ponds had no aquatic life phase during the time of survey (i.e. no larvae or adults in the water), and this likely resulted in lack of detection. To better evaluate the performance of the eDNA approach we thus assembled "reduced" VAES data sets containing only those species that are known to have an aquatic life phase during the survey period. For example, we excluded *Leptodactylus fuscus*, a species that was observed calling from the shore of the pond during the survey, but for which no tadpoles were recorded at that time.

In summary, in the five ponds 31 frog species were detected in total with both VAES and eDNA. Each of these methods detected 25 species, while the TS detected 4 species; 19 species were detected by both eDNA survey and VAES (Fig. 2A; Table 1). Six species were detected only by eDNA, and six species were detected only by VAES. If we consider only species with aquatic life phase during the sampling ("reduced" VAES data set; N=20 for all ponds), of these eDNA detected 19 (Fig. 2B, Table 1, Supplemental Table S1).

We detected 11 species (of the twelve) from the equimolar DNA concentration positive controls, and 6 in the non-equimolar DNA concentration controls (Supplemental Fig. S2). Read numbers in the equimolar PCR controls were highly variable, but strongly correlated among the controls ($R > 0.7$ for each pair of positive controls, Supplemental Fig. S2A). The read numbers in the non-equimolar positive control were strongly linked to the DNA template concentrations of the PCRs (Pearson correlation coefficient $R = 0.97$, Supplemental Fig. S2B). Only *Leptodactylus vastus* was recovered in the field positive control (water from an aquarium with *L. vastus* tadpoles).

We found no difference in detection probability between the filter preservation methods (CTAB buffer - A: detection probability p_1=0.278 [0.242; 0.317]; dry - B: p_1=0.243 [0.209; 0.281]),

294Supplemental Fig. S3). Interestingly, subsequent filtering of water filtered through glass fiber
295filters with a 0.2 µm nylon filter - C, appeared to show an increased detection probability
296(p_1=0.327, 95% C.I. = [0.285; 0.372]. The three approaches were equivalent with respect to
297false positive probabilities (Supplemental Fig. S3).

298 Regarding the community structure assessment based on the eDNA data, replicate
299samples of each pond grouped relatively tightly on the latent variable model ordination (Fig. 3A).
300The pond identity was a statistically significant predictor of frog communities in the five ponds
301(ANOVA, df = 6, dev = 534.99, $p < 0.01$). This is reflected in the 95% confidence intervals of the
302group centroids on the ordination which clearly separates all ponds except T1 and T3 (Fig. 3A).
303The ordination of the eDNA pond centroids closely corresponds with the ordination of
304observations from the five ponds with VAES (Procrustes permutation test, R = 0.8, p = 0.03, Fig.
3053B).

306 The cost model of VAES and eDNA showed that the starting costs (i.e. with few
307sampling sites) for VAES are relatively low, but these costs rapidly increase until the taxonomic
308expert becomes familiar with the regional frog fauna (Fig. 4). The VAES price is dependent on
309the species richness: first, the VAES observer needs to record each species on the field, and
310then the taxonomic expert needs to listen to each recording (Fig. 4). eDNA metabarcoding has a
311relatively high entry price since consumables and sequencing are costly, regardless of the
312number of sites. eDNA survey prices are then a linear function of the number of sampling sites,
313and an increase in the site numbers simply adds to sampling and consumable costs, but does
314not influence neither the time spent in the laboratory, nor the sequencing costs (Fig. 4).

315

Discussion

The one-time eDNA survey detected 25 species in the studied ponds and confirmed the presence of (i) almost all species that were in contact with water during the survey time, (ii) all species that had tadpoles in the ponds, (iii) more than half of all 45 species ever recorded from the area, and (iv) about 65% of the 35 species ever recorded in the five surveyed ponds. The eDNA dataset recorded clear differences among the frog communities in the surveyed ponds, a result also in accordance with the VAES data. The comparison of VAES and eDNA cost models show that eDNA biodiversity surveys may be a cost-efficient alternative to VAES in species-rich areas, but not necessarily in areas with low species numbers. Specific recommendations and technical remarks are presented in the Supplemental information.

eDNA approaches for routine use in tropical regions will benefit from the implementation of straightforward and robust sampling techniques. The comparison of the filtration-preservation approaches shows that filters can be dried in the field before being sent to a lab. Currently only few comparative studies exist regarding the preservation of eDNA filtrates on filters (e.g., Hinlo et al., 2017; Spens *et al.*, 2017). Hinlo *et al.* (2017) showed that the simple refrigeration of filters may be preferred to frozen storage. Here we show that filters can be dried in the field: this simplifies transportation and storage since no special precautions are needed, unlike for liquid handling. Interestingly, small pore-size nylon filters seemed to recover a high proportion of target DNA that had previously passed through a 2 µm mesh sized glass fibre filter. While this result may be caused by the different extraction methods used, we consider it likely that a large proportion of DNA occurred in extracellular form. This result suggests that alternative filtering methods based on fine-scale mesh sized filtration with limited filtering volumes may further increase success rates in eDNA studies on amphibian communities. By limited volumes we do not suggest sampling only a few hundred ml of water, but rather to sample enough from a

waterbody to ensure representativeness, and then filter from this well-mixed sample until the filter clogs.

During the one-time surveys, six species were only detected with eDNA, and not with VAES (Fig. 2A). Four of these species (*Dermatonotus muelleri*, *Leptodactylus elenae*, *Osteocephalus taurinus*, *Rhinella schneideri*) are quite common in the area. The detection of these four species by eDNA but not by VAES may result from the low abundance of adults or tadpoles in the ponds, or a lack in acoustic activity. In some cases eDNA detections are even likely based on single tadpoles or single adults only. For example, one adult *L. syphax* was heard calling from a crevice of a rocky outcrop, its usual habitat (de Sá *et al.* 2014), approximately 30 m away from T2, but detection by eDNA most probably was due to the presence of tadpoles in pond T2. All other species that had tadpoles in the ponds were also detected by eDNA (Table 1). Furthermore, our results suggest that detection of rare or solitary species by eDNA is possible, since the presence of *L. vastus* in the eDNA most possibly can be traced back to single, territorial adult males that were present during VAES of T4 and T5. These examples show the great potential of eDNA for detection of elusive life stages like tadpoles and less abundant species. The remaining two species detected only with eDNA are not known to occur in the area, thus they are candidates for false assignments. The low read numbers assigned to these species also supports this (Supplemental Table S1). Both of these were assigned with sequence data from EMBL (*L. latinasus*: KM091595, *L. laevis*: AY843696). *Lysapsus laevis* would be the first record of this genus in the study area, but the identification of *Lysapsus* populations only using short 16S eDNA sequences is questionable, especially when considering the unclear taxonomy of the group in Bolivia. Nevertheless, based on their distribution and ecology all of the four known species of *Lysapsus* (*L. boliviana*, *L. caraya*, *L. laevis*, *L. limellum*) may occur at the site (De la Riva *et al.* 2000; Lavilla *et al.* 2004; Reichle 2004a; b; Angulo 2008; Jansen *et al.* 2011; Frost 2016). They may have remained undetected

thus far simply because of their small size and rather inconspicuous advertisement call. *Leptodactylus latinasus* is another species possibly occurring in the area that had so far never been recorded (de Sá *et al.* 2014) and we cannot exclude the possibility that the corresponding sequence variant is actually an erroneous variant of one of the six local *Leptodactylus* species.

Species' ecology might explain why some species were detected only with VAES, but not by eDNA (Fig. 2A, Supplemental Table S1). None of these species are strictly bound to ponds in the life stages occurring during our sampling: *Dendropsophus arndti* and *D. salli*, usually call from plants on the pond shores (Schulze *et al.* 2009) and have only sporadic contact with water. *Leptodactylus fuscus* and *Pseudopaludicola* sp. also do not enter the water but usually call from nearby muddy grounds or grasslands. We did not detect tadpoles of these species in the ponds: the tadpoles of these species develop during the rainy season which triggers reproduction, but our sampling slightly preceded the rainy season. Some of the other undetected species have terrestrial or quasi-terrestrial life histories: *L. fuscus* deposits eggs within foam nests in underground burrows at some distance from ponds and these are washed into nearby water bodies by floods that follow heavy rains (Heyer 1978; Lucas *et al.* 2008). *Boana geographica* was likely missed as a result of PCR bias: we could never recover it from positive controls when DNA from other species was also present (Supplemental Fig. S2), although single-template PCR reactions worked. *H. geographicus* was also not detected by VAES, so it is possible that the species was not present during the survey. If we consider the specific life histories, the one-time eDNA survey only missed a single species (*Sphaenorhynchus lacteus*) in the area which was found calling from the water during the VAES (Fig. 2B, Supplemental Table S1). Regarding the whole local species pool (45 species in 10 years), some of the species (10) may not be detectable by eDNA since they reproduce outside of water (e.g. *Leptodactylus mystacinus*), or are forest dwellers (e.g. *Leptodactylus* cf. *didymus*, Fig. 2D, Table S2). Overall, the results suggests that eDNA performed similarly well in detecting

390species as VAES, but it was not successful to identify all species occurring at the specific
391ponds. Differences among individual ponds may result from the differential observational biases
392of eDNA and VAES approaches, none of which is providing a “true” list of species (hence the
393long-standing need to model species occupancy). eDNA methods seem to be a powerful tool for
394the detection of elusive life stages and less abundant species in tropical communities: with a
395single sampling eDNA detected more than half of the 45 species known to be present in the
396area, and about 65% of the detectable species from the area (23 out of 35).

397 The sampling was done at the beginning of the rainy season when only few species
398reproduced in the ponds. Repeating the sampling during the rainy season could have increased
399species detections by eDNA. The results show the importance of life history in the
400metabarcoding-based survey of tropical frogs and emphasize that sampling at multiple time
401points may be essential for more complete species pools also with eDNA (see
402Recommendations and technical remarks in the Supplemental information). Comprehensive
403and taxonomically sound sequence databases are essential for eDNA metabarcoding studies:
404we had a database that contained all 45 species that were ever recorded in the area. This
405database was essential for both initial primer development, and taxonomic assignments: indeed,
406both eDNA-recorded taxa that were not present on the complete local species list of the present
407study (Fig. 2, Supplemental Table S2) were identified in the EMBL-based assignment that
408followed the assignment with the custom local 16S database. The importance of sequence
409assignment databases is long recognized, with considerable efforts underway to establish them
410(Ratnasingham & Hebert 2007; Coissac *et al.* 2016). There is an urgent need of further DNA
411sampling to create regional reference data banks, but only a few studies aim at completing DNA
412sampling of local anurofaunas in South America (e.g. Jansen *et al.* 2011; Guarnizo *et al.* 2015;
413Moraes *et al.* 2017). However, a preferably complete geographical and taxonomic coverage is
414necessary for continent-wide eDNA-based frog inventories. Increased DNA sampling will also

increase knowledge about actual distributions and taxonomy of Neotropical frogs (including cryptic lineages). This will improve IUCN evaluations, which currently clearly lack distributional information (Supplemental Table S1).

There was considerable variation in the species recorded with the spatial replicates of the eDNA samples, and this underlines that eDNA sampling should be replicated for a good representation of community composition. These samples may be pooled before DNA extraction to optimize costs if site-level variation is not of interest. Interestingly, community structures across ponds inferred from both eDNA and VAES datasets where highly similar (Fig. 3). The eDNA data did not distinguish the communities from ponds T1 and T3, and these communities also grouped closely in the VAES results. Although the comparison of frog community structures among the assessed ponds may be confounded by the different sample sizes (with the eDNA ordination based on many spatial replicates and the VAES ordination on a single observation event per pond), and the eDNA ordinations are based on read-abundances in contrast to the VAES presence-absence ordinations, the correspondence of the results is still striking. Similar results were found in insect (Ji *et al.* 2013; Elbrecht *et al.* 2017) and plant communities (Niemeyer *et al.* 2017) where eDNA-based and morphology-based identifications result in similar conclusions about community structure, despite markedly different species lists. Our results provide further evidence that eDNA-based biodiversity surveys are highly sensitive to differences among ecological communities. These inferences are comparable to those derived with encounter-based observations and are informative about processes that underlay community assembly.

Although central to deciding on a method for biodiversity surveys, cost comparisons are not straightforward since they need to be based on expert knowledge both in VAES and eDNA. Cost comparisons were performed for single species eDNA detection (Huver *et al.* 2015; Davy *et al.* 2015; Smart *et al.* 2016), but we are not aware of frameworks suitable for eDNA

metabarcoding. Here the VAES cost estimation is informed by over a decade of field and integrative taxonomic work with tropical frogs (Jansen *et al.* 2007, 2016; Schulze *et al.* 2009; Brusquetti *et al.* 2014), while the eDNA part is only informed by a few years of eDNA biodiversity surveys (Bálint *et al.* 2016; Vörös *et al.* 2017). Such comparisons are urgently needed due to stakeholder demands in eDNA (e.g. governmental agencies, conservation NGOs, fisheries, etc.).

The two species richness scenarios we defined for cost comparisons (Fig. 4) shows that VAES costs become high in regions with high frog richness since they are a function of both the number of sites and the number of species. Biodiversity surveys with eDNA are not necessarily cheaper in low-richness regions since entry costs for the eDNA work are high: lab consumables and sequencing are costly. However, eDNA survey costs are not dependent on the local biodiversity since metabarcoding can consider thousands of species simultaneously in a sample (Taberlet *et al.* 2012b). Consequently, eDNA costs are function of sample numbers, which influence the collection time spent on the field and consumables. eDNA metabarcoding operations are easily scaled up in a sense that hundreds of samples can be simultaneously processed (Ficetola *et al.* 2015).

Several aspects of our cost models are contentious. One issue is whether the relatively untrained VAES observers, or taxonomic experts perform the fieldwork, since taxonomic experts may identify many of the targeted frog species immediately on the field. Currently, most surveys of high-diversity areas are directly done by experts interested in the local fauna, but we argue that this will not work for continental - global biodiversity surveys simply because there are not enough taxonomic experts (Buyck 1999; Haas & Häuser 2003). We also did not consider a scenario when VAES surveys are performed with automated recording devices (ARDs), and sounds are automatically identified by algorithms (see Recommendations and technical remarks in the Supplemental Information). The sound complexity in tropical environments currently

prohibits the use of automated sound identifications (Campos-Cerqueira & Aide 2016). It is also difficult to compare fundamental infrastructure and training costs (i.e. the establishment of an eDNA laboratory versus the training of taxonomic experts). Discussions about cost models are timely since they will play major roles in devising much needed continental and global biodiversity surveys (Tittensor *et al.* 2014).

21

Conclusions

eDNA seems to be suitable to standardized biodiversity surveys of frogs even in species-rich areas, but it may be overly costly for smaller studies in low-richness regions. Differences between eDNA and traditional surveys seem to result largely from different observational biases. Consideration of life histories promises to improve comprehensiveness of both types of surveys and thus also their correspondence. The eDNA data is certainly suitable to characterize not only community composition, but also the factors that shape communities: this gives an unprecedented opportunity to incorporate eDNA as a standard toolkit for community ecology and macroecology. Assessing community structure with eDNA-based community data foresees global biodiversity surveys and monitoring that will support both biodiversity research, and informed decisions on sustainable use of biological diversity.

22

Acknowledgements

The authors thank Jenny Wertheimer for technical assistance during eDNA extraction, José Ribeiro for a helpful discussion on passive acoustic monitoring, and Simon Vitecek for helpfuls suggestions on the manuscript. MJ, AS and JLA thank the owners of Hacienda San Sebastián (family Werding) for their invitation to conduct herpetological surveys on their properties and for logistic support. TC was supported by the USGS Amphibian Research and Monitoring Initiative. MJ was supported by the Erika and Walter Datz-Stiftung, Bad Homburg v. d. H., equipment was sponsored by Globetrotter, Frankfurt. This research presents an outcome of the Centre for Translational Biodiversity Genomics (LOEWE-TBG) and was supported by the research funding programme “LOEWE – Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz" of Hesse's Ministry of Higher Education, Research, and the Arts.

23

References

Angulo A (2008) *Lysapsus boliviana*. *The IUCN Red List of Threatened Species*.

Bailey LL, MacKenzie DI, Nichols JD (2014) Advances and applications of occupancy models. *Methods in ecology and evolution / British Ecological Society*, **5**, 1269–1279.

Bálint M, Bahram M, Eren AM *et al.* (2016) Millions of reads, thousands of taxa: microbial community structure and associations analyzed via marker genes. *FEMS microbiology reviews*, **40**, 686–700.

Bálint M, Domisch S, Engelhardt CHM *et al.* (2011) Cryptic biodiversity loss linked to global climate change. *Nature climate change*, **1**, 313–318.

Bálint M, Nowak C, Márton O *et al.* (2017) Input files for the comparison of visual-audio encounter surveys and eDNA metabarcoding of species-rich tropical frog communities. https://doi.org/10.6084/m9.figshare.5099842

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

Bishop PJ, Angulo A, Lewis JP (2012) The amphibian extinction crisis-what will it take to put the action into the amphibian conservation action plan? *SAPI EN. S. Surveys*.

Bista I, Carvalho GR, Walsh K *et al.* (2017) Annual time-series analysis of aqueous eDNA reveals ecologically relevant dynamics of lake ecosystem biodiversity. *Nature communications*, **8**, 14087.

Bohmann K, Evans A, Gilbert M *et al.* (2014) Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in ecology & evolution*, **29**, 358–367.

Boyer F, Mercier C, Bonin A *et al.* (2016) obitools: a unix-inspired software package for DNA metabarcoding. *Molecular ecology resources*, **16**, 176–182.

Brusquetti F, Jansen M, Barrio-Amorós et al. (2014) Taxonomic review of *Scinax*

518 *fuscomarginatus* (Lutz, 1925) and related species (Anura; Hylidae). *Zoological journal of*
519 *the Linnean Society*, **171**, 783–821.

520Bush A, Sollmann R, Wilting A *et al.* (2017) Connecting Earth observation to high-throughput
521 biodiversity data. *Nature Ecology & Evolution*, **1**, 176.

522Butchart SHM, Walpole M, Collen B *et al.* (2010) Global biodiversity: indicators of recent
523 declines. *Science*, **328**, 1164–1168.

524Buyck B (1999) Taxonomists are an endangered species in Europe. *Nature*, **401**, 321.

525Caminer MA, Milá B, Jansen M *et al.* (2017) Systematics of the *Dendropsophus leucophyllatus*
526 species complex (Anura: Hylidae): Cryptic diversity and the description of two new species.
527 *PloS one*, **12**, e0171785.

528Campos-Cerqueira M, Aide TM (2016) Improving distribution data of threatened species by
529 combining acoustic monitoring and occupancy modelling. *Methods in ecology and evolution*
530 */ British Ecological Society*, **7**, 1340–1348.

531Castro AAJF, A A J, Martins FR, Tamashiro JY, Shepherd GJ (1999) How rich is the flora of
532 Brazilian Cerrados? *Annals of the Missouri Botanical Garden. Missouri Botanical Garden*,
533 **86**, 192.

534Clarke DA, York PH, Rasheed MA, Northfield TD (2017) Does biodiversity–ecosystem function
535 literature neglect tropical ecosystems? *Trends in ecology & evolution*, **32**, 320–323.

536Coissac E, Hollingsworth PM, Lavergne S, Taberlet P (2016) From barcodes to genomes:
537 extending the concept of DNA barcoding. *Molecular ecology*, **25**, 1423–1428.

538Costello MJ, May RM, Stork NE (2013) Can we name Earth's species before they go extinct?
539 *Science*, **339**, 413–416.

540Cristescu ME (2014) From barcoding single individuals to metabarcoding biological
541 communities: towards an integrative approach to the study of global biodiversity. *Trends in*
542 *ecology & evolution*, **29**, 566–571.

543Davy CM, Kidd AG, Wilson CC (2015) Development and validation of environmental DNA

544 (eDNA) markers for detection of freshwater turtles. *PloS one*, **10**, e0130965.

545De la Riva I, Köhler J, Lötters S, Reichle S (2000) Ten years of research on Bolivian
546 amphibians: updated checklist, distribution, taxonomic problems, literature and
547 iconography. *Revista espanola de herpetologia*, **14**, 19–164.

548de Sá RO, Grant T, Camargo A *et al.* (2014) Systematics of the neotropical genus
549 *Leptodactylus* Fitzinger, 1826 (Anura: Leptodactylidae): phylogeny, the relevance of non-
550 molecular evidence, and species accounts. *South american journal of herpetology /*
551 *Sociedade Brasileira de Herpetologia*, **9**, S1–S100.

552Duellman W, Pyles R (1983) Acoustic resource partitioning in anuran communities. *Copeia,* **1983**,
553 639–649.

554Elbrecht V, Vamos EE, Meissner K, Aroviita J, Leese F (2017) Assessing strengths and
555 weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine
556 stream monitoring. *Methods in ecology and evolution / British Ecological Society*, **8**, 1265–
557 1275.

558Ferrão M, Colatreli O, de Fraga R *et al.* (2016) High species richness of *Scinax* treefrogs
559 (Hylidae) in a threatened Amazonian landscape revealed by an integrative approach. *PloS*
560 *one*, **11**, e0165679.

561Ficetola, GF, Miaud C, Pompanon F, Taberlet P (2008) Species detection using environmental
562 DNA from water samples. *Biology letters*, **4**, 423–425.

563Ficetola, GF, Miaud C, Pompanon F, Taberlet P (2008) Species detection using environmental
564 DNA from water samples. *Biology letters*, **4**, 423–425.

565Fiske I, Chandler R, Others (2011) unmarked: An R package for fitting hierarchical models of
566 wildlife occurrence and abundance. *Journal of statistical software*, **43**, 1–23.

567Fouquet A, Gilles A, Vences M *et al.* (2007) Underestimation of species richness in neotropical
568 frogs revealed by mtDNA analyses. *PloS one*, **2**, e1109.

569Frost DR (2016) Amphibian Species of the World: an Online Reference. Version 6.0.

570Gehara M, Canedo C, Haddad CFB, Vences M (2013) From widespread to microendemic:
571 molecular and acoustic analyses show that *Ischnocnema guentheri* (Amphibia:
572 Brachycephalidae) is endemic to Rio de Janeiro, Brazil. *Conservation genetics* , **14**, 973–
573 982.
574Gehara M, Crawford AJ, Orrico VGD *et al.* (2014) High levels of diversity uncovered in a
575 widespread nominal taxon: continental phylogeography of the Neotropical tree frog
576 *Dendropsophus minutus*. *PloS one*, **9**, e103958.
577Goldberg, CS, Pilliod DS, Arkle RS, Waits LP (2011) Molecular detection of vertebrates in
578 stream water: A demonstration using Rocky Mountain Tailed Frogs and Idaho Giant
579 Salamanders. *Plos* One, **6**, e22746.
580Guarnizo CE, Paz A, Muñoz-Ortiz A, *et al.* (2015) DNA barcoding survey of anurans across the
581 eastern Cordillera of Colombia and the impact of the Andes on cryptic diversity. *PLoS one*,
582 **10**, e0127312.
583Haas F, Häuser CL (2003) Taxonomists: an endangered species? In: *Success stories in*
584 *implementation of the programmes of work on dry and sub-humid lands and the global*
585 *taxonomy initiative*, p. 87. Montreal: Secretariat of the Convention on Biological Diversity
586Haddad, CFB, Prado CPA (2005); Reproductive modes in frogs and their unexpected diversity
587 in the Atlantic Forest of Brazil. *BioScience*, **55**, 207–217.
588Henry P-Y, Lengyel S, Nowicki P *et al.* (2008) Integrating ongoing biodiversity monitoring:
589 potential benefits and methods. *Biodiversity and conservation*, **17**, 3357–3382.
590Heyer WR (1978) Systematics of the fuscus group of the frog genus *Leptodactylus* (Amphibia,
591 Leptodactylidae). Natural History Museum of Los Angeles County.
592Hinlo R, Gleeson D, Lintermans M, Furlan E (2017) Methods to maximise recovery of
593 environmental DNA from water samples. *PloS one*, **12**, e0179251.
594Hui FKC, Taskinen S, Pledger S, Foster SD, Warton DI (2015) Model-based approaches to
595 unconstrained ordination. *Methods in ecology and evolution / British Ecological Society*, **6**,

596 399–411.

597Huver JR, Koprivnikar J, Johnson PTJ, Whyard S (2015) Development and application of an

598 eDNA method to detect and quantify a pathogenic parasite in aquatic ecosystems.

599 *Ecological applications: a publication of the Ecological Society of America*, **25**, 991–1002.

600Jansen M (2009) Measuring temporal variation in calling intensity of a frog chorus with a data

601 logging sound level meter: results from a pilot study in Bolivia. *Herpetology notes*, **2**, 143–

602 149.

603Jansen M, Álvarez LG, Köhler G (2007) New species of *Hydrolaetare* (Anura, Leptodactylidae)

604 from Bolivia with some notes on Its natural history. *Journal of Herpetology*, **41**, 724–732.

605Jansen M, Bloch R, Schulze A, Pfenninger M (2011) Integrative inventory of Bolivia's lowland

606 anurans reveals hidden diversity. *Zoologica scripta*, **40**, 567–583.

607Jansen M, Plath M, Brusquetti F, Ryan MJ (2016) Asymmetric frequency shift in advertisement

608 calls of sympatric frogs. *Amphibia-reptilia: publication of the Societas Europaea*

609 *Herpetologica*, **37**, 137–152.

610Jerde CL, Mahon AR, Chadderton WL, Lodge DM (2011) "Sight-unseen" detection of rare

611 aquatic species using environmental DNA. Conservation letters, **4**, 150–157.

612Ji Y, Ashton L, Pedley SM *et al.* (2013) Reliable, verifiable and efficient monitoring of

613 biodiversity via metabarcoding. *Ecology letters*, **16**, 1245–1257.

614Jungfer K-H, Reichle S, Piskurek O (2010) Description of a new cryptic southwestern

615 Amazonian species of leaf-gluing treefrog, genus *Dendropsophus* (Amphibia: Anura:

616 Hylidae). *Salamandra*, **46**, 204–213.

617Kozarewa I, Turner DJ (2011) 96-plex molecular barcoding for the Illumina Genome Analyzer.

618 In: *High-Throughput Next Generation Sequencing* (eds Kwon YM, Ricke CR), pp. 279–298.

619 Humana Press, Totowa, NJ.

620Lavilla E, Reichle S, Lajmanovich R, Faivovich, Julian (2004) *Lysapsus limellum*. *The IUCN*

621 *Red List of Threatened Species*.

622Lees AC, Pimm SL (2015) Species, extinct before we know them? *Current biology: CB*, **25**,
623 R177–80.
624Lopes CM, Sasso T, Valentini A *et al.* (2016) eDNA metabarcoding: a promising method for
625 anuran surveys in highly diverse tropical forests. *Molecular ecology resources*, **17**, 904–
626 914.
627Lucas EM, Brasileiro CA, Oyamaguchi HM, Martins M (2008) The reproductive ecology of
628 *Leptodactylus fuscus* (Anura, Leptodactylidae): new data from natural temporary ponds in
629 the Brazilian Cerrado and a review throughout its distribution. *Journal of natural history*, **42**,
630 2305–2320.
631MacKenzie DI, Nichols JD, Lachman GB *et al.* (2002) Estimating site occupancy rates when
632 detection probabilities are less than one. *Ecology*, **83**, 2248–2255.
633Miles L, Newton AC, DeFries RS *et al.* (2006) A global overview of the conservation status of
634 tropical dry forests. *Journal of biogeography*, **33**, 491–505.
635Miller DA, Nichols JD, McClintock BT *et al.* (2011) Improving occupancy estimation when two
636 types of observational error occur: non-detection and species misidentification. *Ecology*, **92**,
637 1422–1428.
638Miya M, Sato Y, Fukunaga T *et al.* (2015) MiFish, a set of universal PCR primers for
639 metabarcoding environmental DNA from fishes: detection of more than 230 subtropical
640 marine species. *Royal Society open science*, **2**, 150088.
641Moraes LJCL, Almeida AP, Fraga R *et al.* (2017) Integrative overview of the herpetofauna from
642 Serra da Mocidade, a granitic mountain range in Northern Brazil . *ZooKeys*, **715**, 103-159.
643Møller AP (2010) When climate change affects where birds sing. *Behavioral ecology: official*
644 *journal of the International Society for Behavioral Ecology*, **22**, 212–217.
645Niemeyer B, Epp LS, Stoof-Leichsenring KR, Pestryakova LA, Herzschuh U (2017) A
646 comparison of sedimentary DNA and pollen from lake sediments in recording vegetation
647 composition at the Siberian treeline. *Molecular ecology resources*, **17**, e46–e62.

Ortega-Andrade HM, Rojas-Soto OR, Valencia JH *et al.* (2015) Insights from integrative systematics reveal cryptic diversity in *Pristimantis* frogs (Anura: Craugastoridae) from the Upper Amazon Basin. *PloS one*, **10**, e0143392.

Pauls SU, Nowak C, Bálint M, Pfenninger M (2013) The impact of global climate change on genetic diversity within populations and species. *Molecular ecology*, **22**, 925–946.

Pedersen MW, Overballe-Petersen S, Ermini L *et al.* (2015) Ancient and modern environmental DNA. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **370**, 20130383.

Peres-Neto PR, Jackson DA (2001) How well do multivariate data sets match? The advantages of a pProcrustean superimposition approach over the Mantel test. *Oecologia*, **129**, 169–178.

Petitot M, Manceau N, Geniez P, Besnard A (2014) Optimizing occupancy surveys by maximizing detection probability: application to amphibian monitoring in the Mediterranean region. *Ecology and evolution*, **4**, 3538–3549.

Pfenninger M, Schwenk K (2007) Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. *BMC evolutionary biology*, **7**, 121.

Pimm SL, Jenkins CN, Abell R *et al.* (2014) The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, **344**, 1246752.

Power MJ, Whitney BS, Mayle FE *et al.* (2016) Fire, climate and vegetation linkages in the Bolivian Chiquitano seasonally dry tropical forest. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **371**, 20150165.

Ratnasingham S, Hebert PDN (2007) BARCODING: bold: The barcode of life data system (http://www.barcodinglife.org). *Molecular ecology notes*, **7**, 355–364.

R Core Team (2017) *R: A Language and Environment for Statistical Computing*.

Rees HC, Maddison BC, Middleditch DJ *et al.* (2014) The detection of aquatic animal species using environmental DNA - a review of eDNA as a survey tool in ecology. Journal of applied

674 ecology, **51**,1450–1459.

675Reichle S (2004a) *Lysapsus laevis*. *The IUCN Red List of Threatened Species*.

676Reichle S (2004b) *Lysapsus caraya*. *The IUCN Red List of Threatened Species*.

677Schmeller DS, Böhm M, Arvanitidis C *et al.* (2017) Building capacity in biodiversity monitoring at

678 the global scale. *Biodiversity and conservation*, 26, 2765–2790.

679Schulze A, Jansen M, Köhler G (2009) Diversity and ecology of anuran communities in San

680 Sebastián (Chiquitano region, Bolivia). *Salamandra*, **45**, 75–90.

681Schulze A, Jansen M, Köhler G (2015) Tadpole diversity of Bolivia's lowland anuran

682 communities: molecular identification, morphological characterisation, and ecological

683 assignment. *Zootaxa*, **4016**, 1–111.

684Shaw JLA, Clarke LJ, Wedderburn SD, *et al*. (2016) Comparison of environmental DNA

685 metabarcoding and conventional fish survey methods in a river system. *Biological*

686 *Conservation*, **197**, 131–138.

687Smart AS, Weeks AR, van Rooyen AR *et al.* (2016) Assessing the cost-efficiency of

688 environmental DNA sampling. *Methods in Ecology and Evolution*, **7**, 1291–1298.

689Spens J, Evans AR, Halfmaerten D, *et al.* (2017) Comparison of capture and storage methods

690 for aqueous macrobial eDNA using an optimized extraction protocol: advantage of enclosed

691 filter. *Methods in Ecology and Evolution*, **8**, 635-645.

692Strand DA, Jussila J, Johnsen SI *et al.* (2014) Detection of crayfish plague spores in large

693 freshwater systems. *The Journal of applied ecology*, **51**, 544–553.

694Stroud JT, Feeley KJ (2017) Neglect of the Tropics Is Widespread in Ecology and Evolution: A

695 Comment on Clarke et al. *Trends in ecology & evolution*, **32**, 626–628.

696Stuart SN, Chanson JS, Cox NA *et al.* (2004) Status and trends of amphibian declines and

697 extinctions worldwide. *Science*, **306**, 1783–1786.

698Taberlet P, Coissac E, Hajibabaei M, Rieseberg LH (2012a) Environmental DNA. *Molecular*

699 *ecology*, **21**, 1789–1793.

700Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (2012b) Towards next-
701 generation biodiversity assessment using DNA metabarcoding. *Molecular ecology*, **21**,
702 2045–2050.
703Thomsen, PF, Kielgast J, Iversen LL *et al*. (2012a) Monitoring endangered freshwater
704 biodiversity using environmental DNA. Molecular Ecology, **21**, 2565–2573.
705Thomsen PF, Kielgast J, Iversen LL *et al.* (2012b) Detection of a diverse marine fish fauna
706 using environmental DNA from seawater samples. *PloS one*, **7**, e41732. Thomsen PF,
707 Willerslev E (2015) Environmental DNA – An emerging tool in conservation for monitoring
708 past and present biodiversity. *Biological conservation*, **183**, 4–18.
709Tittensor DP, Walpole M, Hill SLL *et al.* (2014) A mid-term analysis of progress toward
710 international biodiversity targets. *Science*, **346**, 241–244.
711UNEP (2016) The strategic plan for biodiversity 2011-2020, the Aichi biodiversity targets and
712 NBSAPs. In: *Sourcebook of Opportunities for Enhancing Cooperation among the*
713 *Biodiversity-Related Conventions at National and Regional Levels*, pp. 108–135.
714Valentini A, Taberlet P, Miaud C *et al.* (2016) Next-generation monitoring of aquatic biodiversity
715 using environmental DNA metabarcoding. *Molecular ecology*, **25**, 929–942.
716Vörös J, Márton O, Schmidt BR, Gál JT, Jelić D (2017) Surveying Europe's only cave-dwelling
717 chordate species (*Proteus anguinus*) using environmental DNA. *PloS one*, **12**, e0170945.
718Wang Y, Naumann U, Wright ST, Warton DI (2012) mvabund- an Rpackage for model-based
719 analysis of multivariate abundance data. *Methods in ecology and evolution / British*
720 *Ecological Society*, **3**, 471–474.
721Wells KD (2010) *The ecology and behavior of amphibians*. University of Chicago Press.
722Whittaker K, Koo MS, Wake DB, Vredenburg VT (2013) Global declines of amphibians. In:
723 *Encyclopedia of Biodiversity (Second Edition)* (ed Levin SA), pp. 691–699. Academic
724 Press, Waltham.
725Wilson DM, Bart J (1985) Reliability of singing bird surveys: effects of song phenology during

726 the breeding season. *The Condor*, **87**, 69–73.

727Wittwer C, Stoll S, Cocchiararo B, Strand D, Vrålstad T, Nowak C, Thines M (2017) eDNA-

728 based crayfish plague monitoring is superior to conventional trap-based assessments in

729 year-round detection probability. *Hydrobiologia* (pending minor revision).

730Yamamoto S, Masuda R, Sato Y *et al*. (2017) Environmental DNA metabarcoding reveals local

731 fish communities in a species-rich coastal sea. *Scientific reports*, **7**, 40368.

732Zimmerman BL (1994) Audio strip transects. In: *Measuring and monitoring biological*

733 *biodiversity. Standard methods for amphibians* (eds Heyer WR, Donnelly MA, McDiarmid

734 RW, Hayek LC, Foster MS), pp. 99–108. Smithsonian Institution Press, Washington DC.

735

736

737

738

739 Data accessibility

740 DNA sequences: EMBL accession PRJEB22113

741 Cost model parameters and calculations: FigShare,

742https://doi.org/10.6084/m9.figshare.5099842.v5

743 Input data for statistical analyses: FigShare,

744https://doi.org/10.6084/m9.figshare.5099842.v5.

745 R-script for bioinformatic filtering and statistical analyses:

746https://doi.org/10.5281/zenodo.1294092

747

34

Author contributions

M.B., M.J., C.N., C.W. designed research; J.L.A., M.B., M.J., O.M., A.S., C.W., B.C. performed research; B.C., T.C., C.N., S.U.P. contributed reagents or analytic tools; M.B., M.J. analysed data; M.B., T.C., M.J., O.M., C.N., S.U.P., wrote the manuscript.

35

Tables

Table 1.

Numbers of detected species by method per pond (VAES = Visual-audio encounter surveys. Reduced VAES = only considering species that were in contact with water during survey).

Pond	eDNA	VAES	Reduced VAES	Tadpole survey
T1	10	8	6	4
T2	7	4	4	3
T3	3	7	3	0
T4	9	15	13	1
T5	13	18	14	1

36

Figures

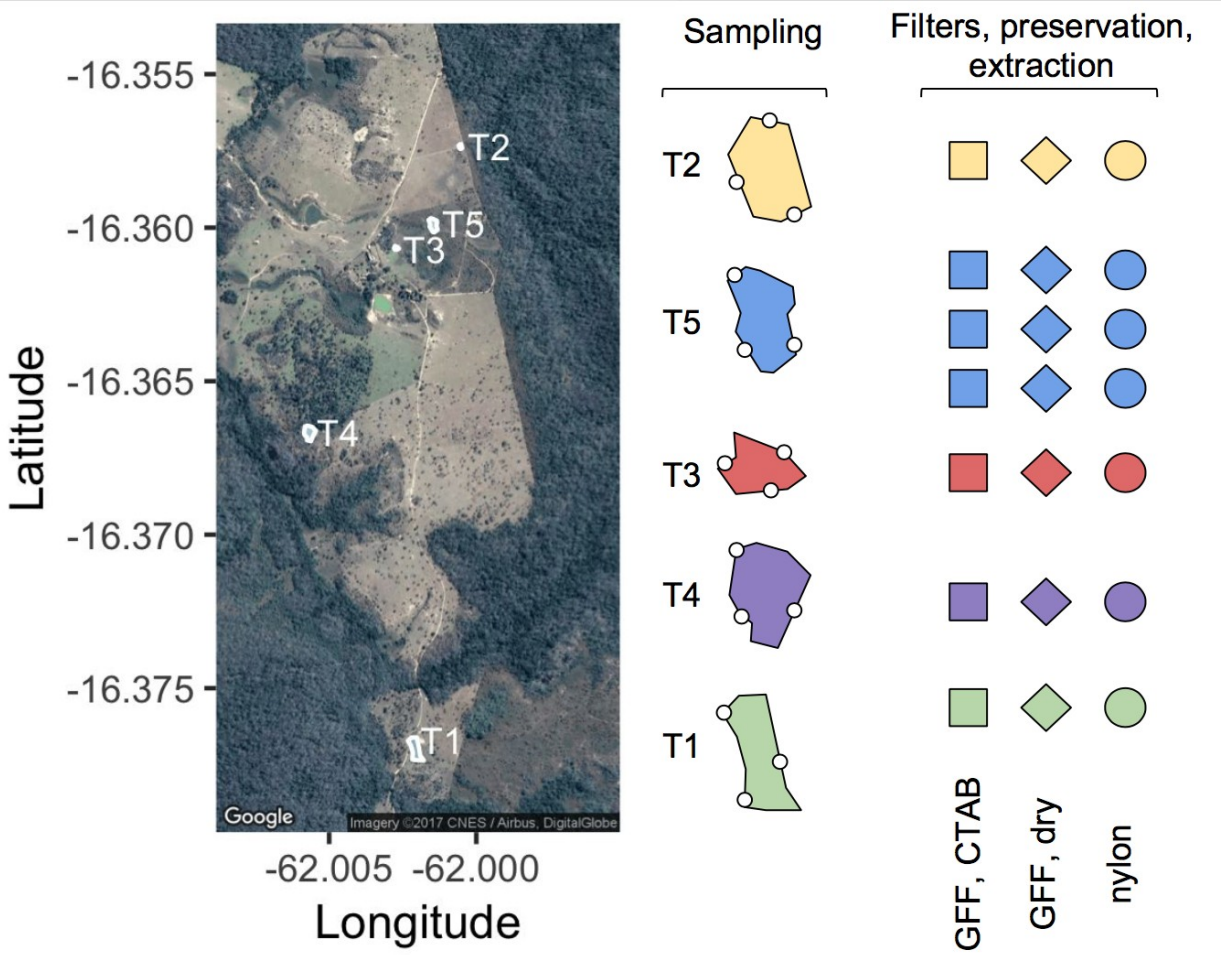


Fig. 1.

Sampled ponds, sampling replication, and filtration - filter preservation strategy. GFF - glass-fiber filter, 2 µm; nylon - nylon filter, 0.2 µm. Each sample (three per pond, marked with small circles) were processed with each three filtration strategies. Multiple samples of pond T5 were three times processed with each filtration strategy.

37

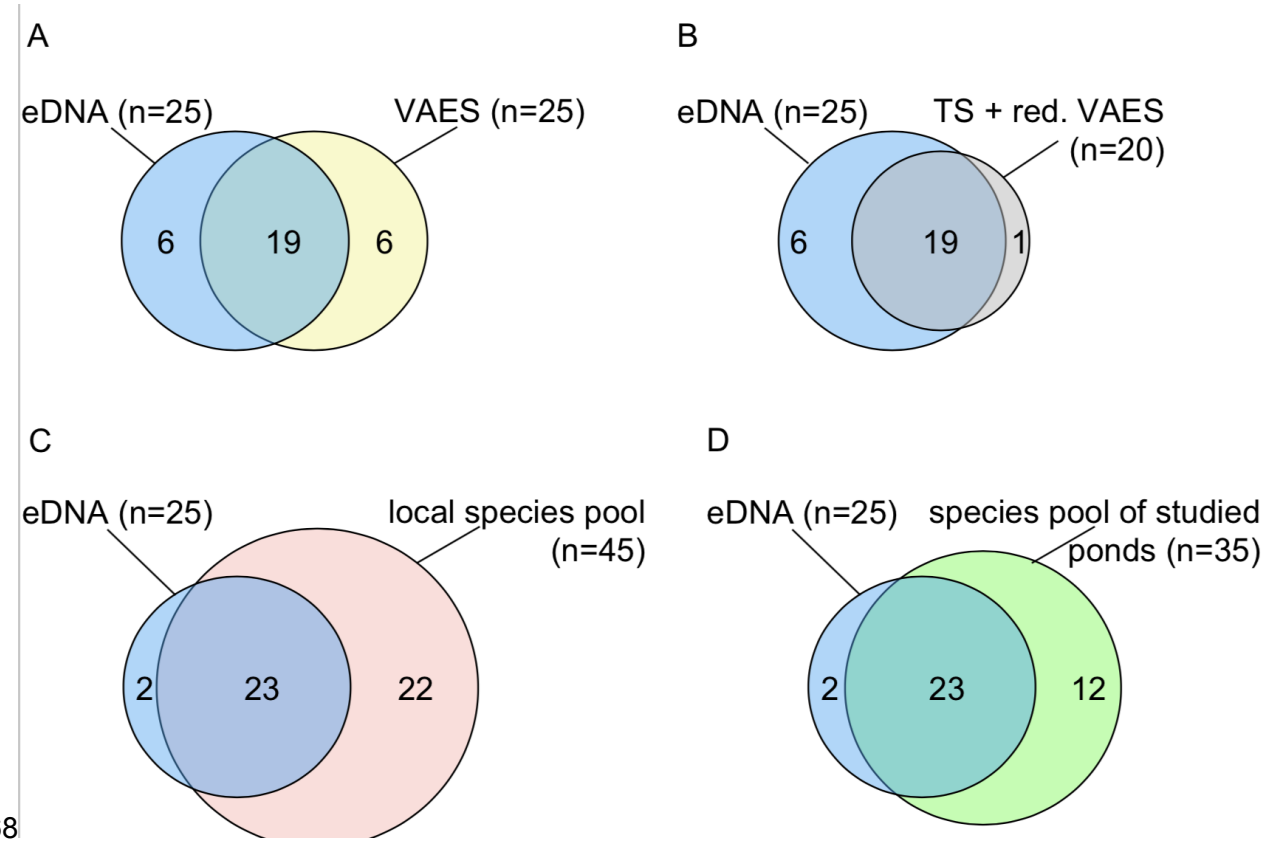
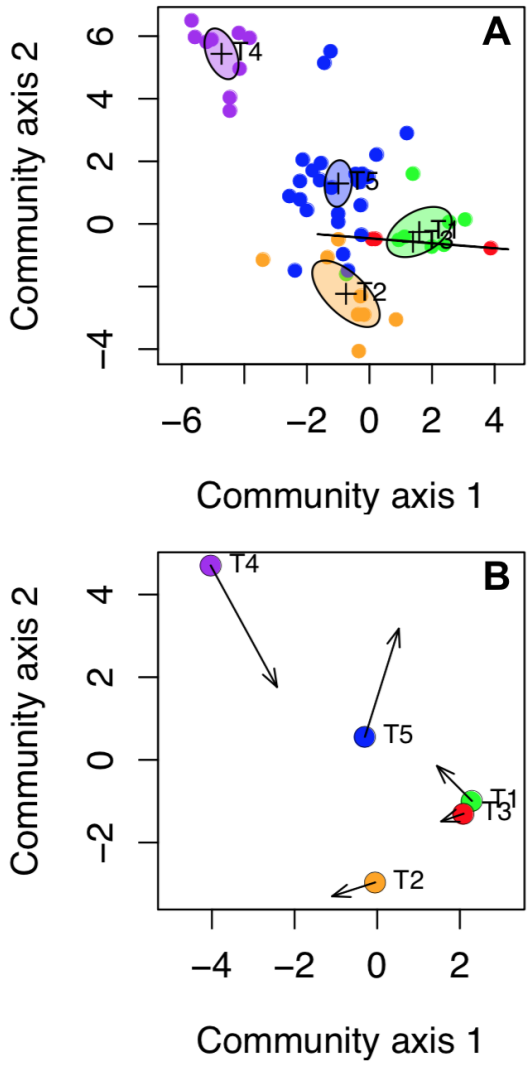


Fig. 2.

Comparison of species lists generated by eDNA (blue) and A) visual and audio encounter surveys (VAES, yellow), B) tadpole survey (TS) plus reduced VAES - only species with high eDNA detectability in ponds (see text for details; grey), C) complete local species pool (red), and D) species pool of studied ponds based on long-term monitoring (green).

38



777

778 Fig. 3.

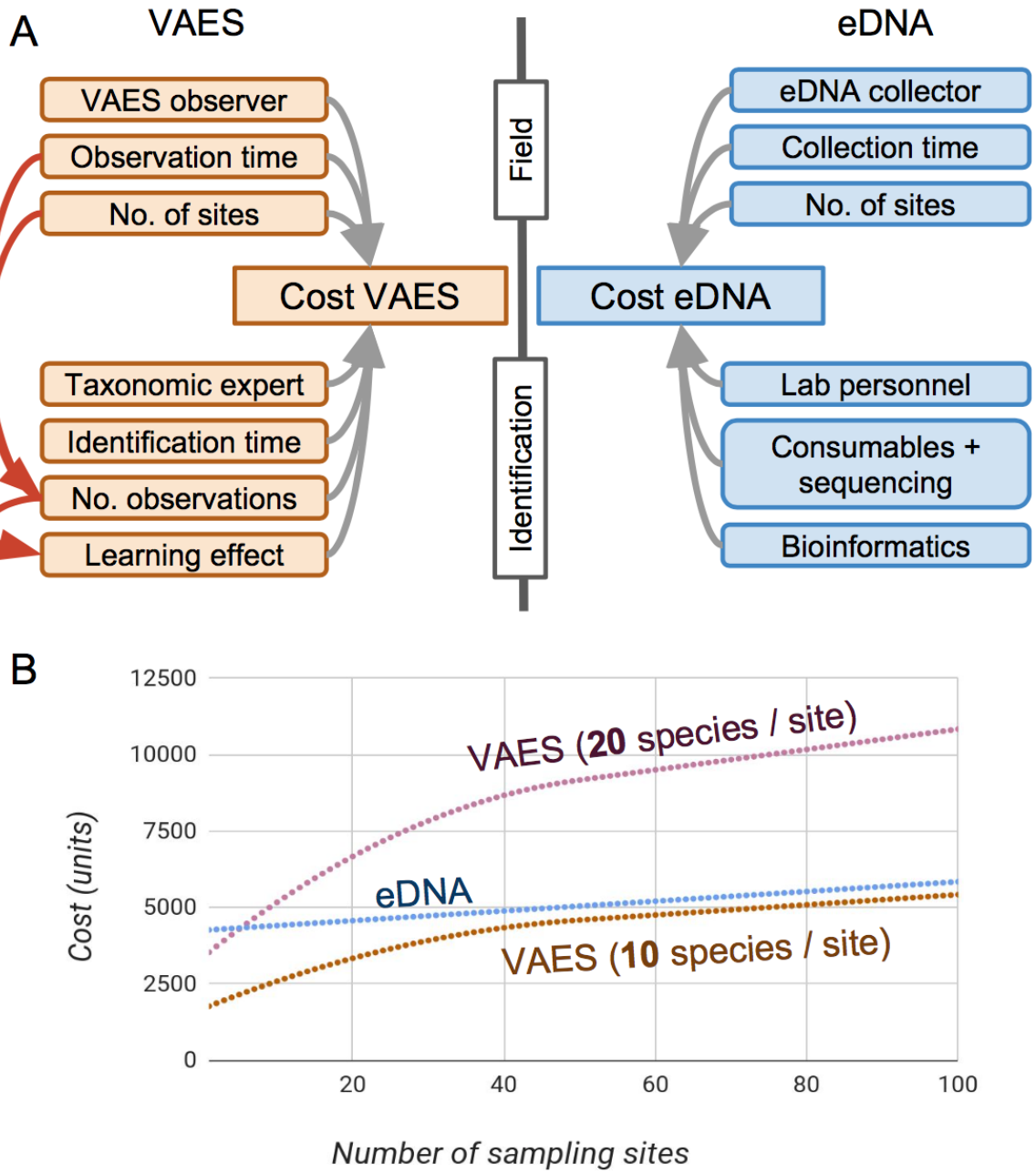
779Comparison of ecological signal between visual and audio encounter surveys (VAES) and
780eDNA metabarcoding surveys. A - Latent variable model ordination of the pond samples
781according to the read numbers of species. Water samples from each pond were taken at three
782locations. For ponds T1-T4, three samples at each location were filtered through two GFF (A,B),
783and one nylon filter (C), resulting in nine samples per pond. For pond T5, each filtration was
784done three times. Not all samples contained reads after the bioinformatic quality filtering and
785these samples are not shown on the ordination (see e.g. T3 - red). The ellipses represent 95%
786confidence intervals for the standard errors of the pond centroids (marked with +). B - Non-

787metric multidimensional scaling plot of the five ponds (Jaccard distance), according to VAES

788species presence-absences. The arrows represent the Procrustes rotation of the VAES pond

789ordination and they target the group centroids of the latent variable model ordination.

40



790

791 Fig. 4.

792 Cost comparisons between visual and audio encounter surveys (VAES) and eDNA

793metabarcoding surveys. A - Schematic overview of the visual and audio encounter survey

794(VAES) and eDNA cost models. Model parameters and calculations are accessible through

795FigShare (Balint_et_al_survey_cost_calculations.xlsx,

796https://doi.org/10.6084/m9.figshare.5099842.v5). B - Cost comparisons of frog diversity surveys

797with eDNA, and two VAES scenarios (a low local species richness scenario - 10 species per

798site, and a medium-high species richness scenario - 20 species per site).

42