Received: 27 September 2019 | Revised: 29 April 2020 | Accepted: 10 June 2020

DOI: 10.1002/edn3.116

ORIGINAL ARTICLE

Environmental DNA
Open Access
Dedicated to the study and use of environmental DNA for basic and applied sciences

WILEY

Metabarcoding unsorted kick-samples facilitates macroinvertebrate-based biomonitoring with increased taxonomic resolution, while outperforming environmental DNA

Lyndall Pereira-da-Conceicoa[1,2,3] | Vasco Elbrecht[4,5] | Andie Hall[2] | Andrew Briscoe[2] | Helen Barber-James[3,6] | Benjamin Price[2]

[1]Department of Natural Sciences, National Museums Scotland, Edinburgh, UK

[2]Life Sciences Department, Natural History Museum, London, UK

[3]Department of Freshwater Invertebrates, Albany Museum, Makhanda, South Africa

[4]Centre for Biodiversity Genomics, University of Guelph, Guelph, ON, Canada

[5]Centre for Molecular Biodiversity Research, Zoological Research Museum Alexander Koenig, Bonn, Germany

[6]Department of Zoology and Entomology, Rhodes University, Makhanda, South Africa

Correspondence
Lyndall Pereira-da-Conceicoa, Department of Natural Sciences, National Museums Scotland, Edinburgh, UK.
Email: lyndall.pereira@gmail.com

Funding information
Department of Business, Energy & Industrial Strategy Rutherford Fund

Abstract

While previous studies have highlighted the potential of DNA-based methods for the biomonitoring of freshwater macroinvertebrates, a limited number have investigated homogenization of bulk samples that include debris, in order to reduce sample-processing costs. This study explores the use of several DNA-based survey methods for water quality and biodiversity assessment in South Africa, comparing morphological and molecular-based identification of freshwater macroinvertebrates at the family level and the level of molecular operational taxonomic units (mOTUs). Seven sites were studied across three rivers with four different sample types collected per site: a standard SASS biomonitoring sample split into a picked sample (also used for morphological identification) and a leftover debris sample; a more intensive-search comprehensive sample; and a filtered water eDNA sample. DNA-based methods recovered higher diversity than morphology, but did not always recover the same taxa, even at the family level. Regardless of the differences in SASS taxon scores, most DNA-based methods, except a few eDNA samples, returned the same water quality assessment category as the standard morphology-based assessment. Homogenized comprehensive samples recovered more freshwater invertebrate diversity than all other methods, suggesting the standardized SASS method overlooks taxa. The eDNA samples recovered more diversity than any other method; however, 90% of the reads were nontarget and as a result eDNA recovered the lowest target (macroinvertebrate) diversity. However, eDNA did find some target taxa that all other methods failed to detect. This study shows that unsorted bulk samples have the potential to be used for water quality biomonitoring, providing higher diversity estimates for macroinvertebrates than either SASS picked or eDNA samples. These results also show the value of incorporating DNA-based approaches into existing South African metrics, providing additional taxonomic resolution to develop more refined metrics for biodiversity management.

KEYWORDS
eDNA, freshwater, macroinvertebrates, next-generation sequencing, South Africa

1 | INTRODUCTION

While large proportions of global biodiversity remain unknown (Stork, 2018), freshwater ecosystems are among the most threatened, due to global climate change (Bates, Kundzewicz, Wu, & Palutikof, 2008) and severe pressure from other anthropogenic impacts (Dallas & Rivers-Moore, 2014). Freshwater resources in Africa are highly impacted, and in South Africa, 57% of river and 75% of wetland ecosystems are highly threatened (Dallas & Rivers-Moore, 2014; Darwall, Smith, Tweddle, & Skelton, 2009; Driver et al., 2012; Nel et al., 2011). A more in-depth knowledge of their current biodiversity would greatly improve our ability to inform management decisions (Hamer, 2013) and would help facilitate freshwater conservation in Africa (Barber-James & Pereira-da-Conceicoa, 2016).

The development of DNA-based identification methods for freshwater macroinvertebrates and their incorporation into biological monitoring programs is rapidly advancing, particularly due to the potential reductions in processing time, greater taxonomic resolution, and reduction in errors compared with current morphological monitoring methods (Hering et al., 2018). While morphological methods often have limited taxonomic resolution due to cryptic species, sex, and life-stages, DNA-based methods can overcome these issues (Ekrem, Stur, & Hebert, 2010; Hebert, Ratnasingham, & de Waard, 2003; Park, Foottit, Maw, & Hebert, 2011; Venter & Bezuidenhout, 2016) when paired with suitable DNA reference libraries. For bulk-collected samples (i.e., unsorted kick/pond-net samples taken from a water body), sample sorting and subsequent morphological identification are time-consuming, the same samples can produce different results depending on the taxonomic expertise and available identification resources (Dickens & Graham, 2002; Haase et al., 2006) and up to a third of specimens and a fifth of taxa can be missed during the sorting stage (Bongard, 2011; Haase, Pauls, Schindehütte, & Sundermann, 2010).

Within the last decade, a growing number of case studies have highlighted the potential applications of DNA-based methods for the bioassessment of freshwater macroinvertebrates (Baird & Hajibabaei, 2012; Blackman et al., 2017; Carew, Pettigrove, Metzeling, & Hoffmann, 2013; Elbrecht, Vamos, Meissner, Aroviita, & Leese, 2017; Elbrecht, Vamos, Steinke, & Leese, 2017; Hajibabaei, Shokralla, Zhou, Singer, & Baird, 2011; Packer, Gibbs, Sheffield, & Hanner, 2009), the majority of which pick out the invertebrates from the sample as the first step. Only a few freshwater macroinvertebrate studies have investigated homogenization in order to remove the specimen sorting stage and reduce sample-processing time, either of cleaned bulk samples where debris has been completely removed (e.g., Andújar et al., 2018; Dowle, Pochon, Banks, Shearer, & Wood, 2016; Gardham, Hose, Stephenson, & Chariton, 2014); bulk samples which have been cleaned, but still include some debris (Majaneva, Diserud, Eagle, Hajibabaei, & Ekrem, 2018) or where debris was added to a mock community to examine inhibition of target species detection (Nichols et al., 2020) .

In parallel, the use of environmental DNA (eDNA) is being explored to detect macroinvertebrates in freshwater environments. However, as eDNA is free of the macroinvertebrate organism and can be transported in flowing waters (Deiner & Altermatt, 2014), these data are spatially complex (Deiner, Fronhofer, Mächler, Walser, & Altermatt, 2016). While eDNA-based studies have been shown to successfully detect macroinvertebrate community richness (Deiner et al., 2017; Deiner, Walser, Mächler, & Altermatt, 2016; Li et al., 2018; Mächler, Little, Wüthrich, Alther, & Fronhofer, 2019), some studies have shown a much lower detection level of invertebrates when compared to morphological methods (Hajibabaei et al., 2019; Macher et al., 2018). These differences can be a result of several factors including DNA shedding rates, persistence and movement in the environment (Barnes & Turner, 2016), sampling, laboratory, and bioinformatic biases as outlined in Blackman et al. (2019).

In addition to the sample processing and bioinformatic biases, DNA-based methods are often limited by the lack of DNA reference libraries (Carew et al., 2017; Porter & Hajibabaei, 2018; Weigand et al., 2019), particularly in regions with high biodiversity and endemism (e.g., South Africa: Venter & Bezuidenhout, 2016). Thus, the comparison of morphological and molecular assessments can be biased due to unrecognized cryptic species. Further compounding this issue, a lack of taxonomic expertise results in a lack of identification keys and hinders the subsequent generation of suitable DNA reference libraries.

While biological monitoring using freshwater macroinvertebrates is used worldwide, the indices are used at various taxon levels, with many indices only using family (or broader) levels of identification (e.g., South Africa: Dickens & Graham, 2002; Tanzania: Kaaya, Day, & Dallas, 2015; Namibia: Palmer & Taylor, 2004). However, such coarse taxonomic resolution has been shown to overlook the varying environmental tolerances that occur at generic or species levels, which can be used to make more well-informed water management decisions (e.g., Barber-James & Pereira-da-Conceicoa, 2016; Macher et al., 2016). Within South Africa, rivers are currently monitored using the South African Scoring System (SASS version 5) protocol (Dickens & Graham, 2002), which uses freshwater macroinvertebrates as a measure of stream ecosystem health. While identification at the SASS level (i.e., mostly at family level) is too coarse for biodiversity and ecological impact assessment, SASS is frequently misused in southern Africa due to the lack of alternative methods (Barber-James & Pereira-da-Conceicoa, 2016) highlighting the need for more suitable approaches to biodiversity and ecological assessment.

In order to explore the use of DNA-based methods of assessment in South Africa, this study compares morphological and a

molecular-based identification of freshwater macroinvertebrates at the SASS family level. DNA-based methods are also compared at the level of the molecular operational taxonomic unit "mOTU." Four DNA-based methods were examined: (a) picked SASS samples, (b) SASS leftover debris (after picking), (c) intensive-search "comprehensive" samples, and (d) filtered water (eDNA) samples. Four main aims were investigated: (1) How do DNA-based methods (picked SASS, leftover SASS, comprehensive sample or eDNA) compare in taxon recovery at both family and mOTU levels; (2) how does molecular identification of SASS samples (either picked or leftovers) compare with traditional morphological identification; (3) how do DNA-based methods compare to traditional SASS water quality assessment; and (4) how does DNA-based mOTU recovery reflect known species recorded from the region.

2 | MATERIALS AND METHODS

2.1 | Study sites and sampling strategy

Samples were collected from a total of seven sites, spread across three Eastern Cape rivers: three sites along each of the Elandsbos ("E") and Tyume tributary ("H") rivers and one site on the Berg ("CD") River (Figure S2, Table S1). River sites occurred across three ecoregions (Dallas, 2005, 2007) namely: Southeastern Coastal Belt (E = Elandsbos), upper Southeastern Uplands (H = Tyume tributary), and Southern Folded (CD = Berg). All sites sampled were selected based on the expected condition of good water quality in a natural or unmodified river condition.

At each site, environmental DNA (eDNA) was collected from surface water using a sterile 1L bucket from five locations within a 5m radius of the sample site and mixed in a sterile 10L bucket. Water was then filtered through three 0.22μm polyethersulfone filters (Sterivex-GP) using a 50ml syringe. Weather conditions were favorable for eDNA sampling, that is, no heavy rainfall or flooding; however, the streams sampled at Hogsback (H) sites had higher turbidity than the other rivers sampled, affecting the amount of water that could be filtered. At each site, water was filtered until the filter blocked (with sediment) or 2 L was reached. The total volume filtered per filter ranged between 0.2 and 2 L, and therefore, total volume for three filters for each Hogsback (H) site was 0.6 L, each Craigdoone (C) site was 1.8 L, and each Tsitsikamma site was 6 L filtered. Each filter was preserved onsite with 96% ethanol and kept refrigerated where possible.

Following eDNA sampling, a SASS sample was taken by an accredited river health practitioner, following the SASS protocol using a 30 × 30 cm framed standard kick-net with a 1mm mesh size: a 2-min kick-sample of stones-in-current across the river; 1-min kick-sample of stones-out-of-current; 2 m total marginal vegetation sweep; and a 1-min stir and sweep of gravel, sand, and mud biotope (Dickens & Graham, 2002). SASS samples, including all debris (i.e., the entire kick-net contents), were preserved in 96% ethanol while in the field. As SASS samples were identified in the laboratory, rather than live in the field as is standard, the results are not directly comparable with other SASS data from these sites. In order to provide a more comprehensive assessment of the biodiversity at each site, whole community samples (here termed "comprehensive") were collected following the SASS sample for 30 min using a 30 × 30 cm framed kick-net with a 250μm mesh. The whole sample, including all debris, was preserved in 96% ethanol while in the field. New kick-nets were used for each river; between-sites within a river kick-nets were washed with consumer dishwashing liquid and then rinsed with 10% bleach.

2.2 | Sample processing

SASS samples were scored by a single individual and identified according to SASS protocol (i.e., a total time of 45 min per sample for identification to family level). Invertebrates from each sample were picked, morphosorted after scoring, and then identified further if possible and counted (here termed "SASS picked"). Morphotaxa were identified mostly to family and genus level using the Guide to the Freshwater Invertebrates of Southern Africa Series (Water Research Commission). A representative individual of each morphospecies from the SASS picked samples was removed as a voucher specimen. Vouchers were imaged before being placed in ATL buffer and proK overnight at 56°C, and DNA was extracted using the Qiagen BioSprint 96 DNA Blood Kit or the DNeasy Blood & Tissue Kit.

The remainder of the picked morphospecies ("SASS picked") samples were placed in 2ml tubes and dried using a DNA SpeedVac DNA120 at high heat (65°C) for one large sample (H4) and medium heat (43°C) for all other samples. The SASS picked samples were then homogenized with three glass beads per 2ml tube using the Qiagen TissueLyser II (30 Hz for 1 min) and digested overnight in 10:1 solution of ATL buffer and proK. Following digestion, DNA was extracted using three columns of the DNeasy Blood & Tissue kit to a final elution of 900 μl. Two subsamples from each extraction were then used for PCR amplification (i.e., PCR replicates).

For each site, the remainder of the SASS sample, consisting of debris and some macroinvertebrates (here termed "SASS leftovers"), were homogenized (after the ethanol was drained off) for up to 10 s using a consumer blender (Breville VBL062, 300 W, 600 ml; Figure 1).

Between samples, the blender blades and container were sterilized using 12% industrial bleach. Each homogenized sample was then subsampled in two extraction replicates, each comprising 10 g homogenate (Figure 2). Subsamples were dried overnight at room temperature, and DNA was extracted using concentrated proK (20 mg/ml) with overnight digestion following Ransome et al. (2017) and then the Qiagen DNeasy PowerMax Soil kit with a final DNA elution of 5 ml. The comprehensive samples were processed in the same way as the SASS leftovers.

For the eDNA samples, the Sterivex filters (three replicates per site, Figure 2) were dried and DNA extracted using the Qiagen DNeasy PowerWater Sterivex kit, following the alternative method

FIGURE 1 An example of a comprehensive, unsorted kick-sample from this study. All debris was included in the sample when processed using a blender

(using a microcentrifuge and not a vacuum manifold), and excluding the 90°C incubation and powerbead steps.

All DNA concentrations were quantified using a NanoDrop 8000 (Thermo Fisher Scientific), and the concentration of DNA was adjusted to 25 ng in the downstream PCRs.

2.3 | Voucher barcoding

Voucher barcoding followed standard in-house protocols using LCO1490 and HCO2198 primers (Folmer, Black, Hoeh, Lutz, & Vrijenhoek, 1994). Each reaction consisted of 1 mM total dNTPs, 3 mM $MgCl_2$, 1.25 μl Bio-Taq DNA polymerase (Bioline), 0.1 μM each primer, and 1× reaction buffer. Cycling conditions were as follows: initial denaturation 94°C for 1min followed by 35 cycles of 94°C for 30 s, 50°C for 30 s, and 72°C for 30 s, with a final elongation of 10 min at 72°C. PCR products were visualized using gel electrophoresis, purified using Agencourt AMPure XP beads, and then sequenced bidirectionally using BigDye terminator reaction mix v3.1 in a 3730xl DNA analyzer (Applied Biosystems) at the NHM sequencing facility. Voucher data and corresponding sequences were uploaded to the Barcode of Life Data System (BOLD) website (Ratnasingham & Herbert, 2007) and can be found under the project name "BEISA" in the public data portal. Uploaded sequences correspond to 115 taxa, comprising 85 identified to genus and 30 identified to species, from five classes and 55 families. Specimens were identified to genus and species levels whenever respective taxonomic keys were available.

2.4 | DNA metabarcoding

DNA metabarcoding followed the 2-step PCR approach outlined in Elbrecht and Steinke (2018) using the BF2/ BR2 freshwater macroinvertebrate fusion primer sets developed for the cytochrome c oxidase subunit I (COI) gene (Elbrecht & Leese, 2017). The single extraction from each SASS picked sample was used in two PCRs (i.e., two PCR replicates) whereas one PCR replicate was performed for each of the two extraction replicates of both comprehensive and SASS leftover samples and for each eDNA filter extraction (Figure 2).

Two-step PCRs were performed on each extraction and PCR replicate using the Qiagen Multiplex PCR Plus Kit with 0.5 μM of each primer in a final volume of 25ul. PCRs were run on Techne™ Prime Elite Thermal Cyclers (Thermo Fisher Scientific) with the following conditions, PCR 1:94°C for 5 min; 25 cycles of 94°C for 30 s; 50°C for 30 s; 65°C for 50 s; and final extension at 65°C for 5 min; the eDNA samples used 27 cycles. Untailed BF2/ BR2 primers were used for PCR 1. Then, 1 μl of amplicon from PCR 1 was used as a template for PCR 2 under the following conditions: 94°C for 5 min; 13 cycles of 94°C for 30 s; 50°C for 30 s; 65°C for 2 min; and final extension at 65°C for 5 min using the tailed BF2/ BR2 fusion primers (Elbrecht & Leese, 2017). One negative control was used in PCR 1 and subsequently used as template for PCR 2. PCR products were visualized using gel electrophoresis.

Following the second PCR, the tagged amplicons and negative control were purified using Agencourt AMPure XP beads at 0.8x ratio. The eDNA sample from site E3 was gel cut using the QIAquick

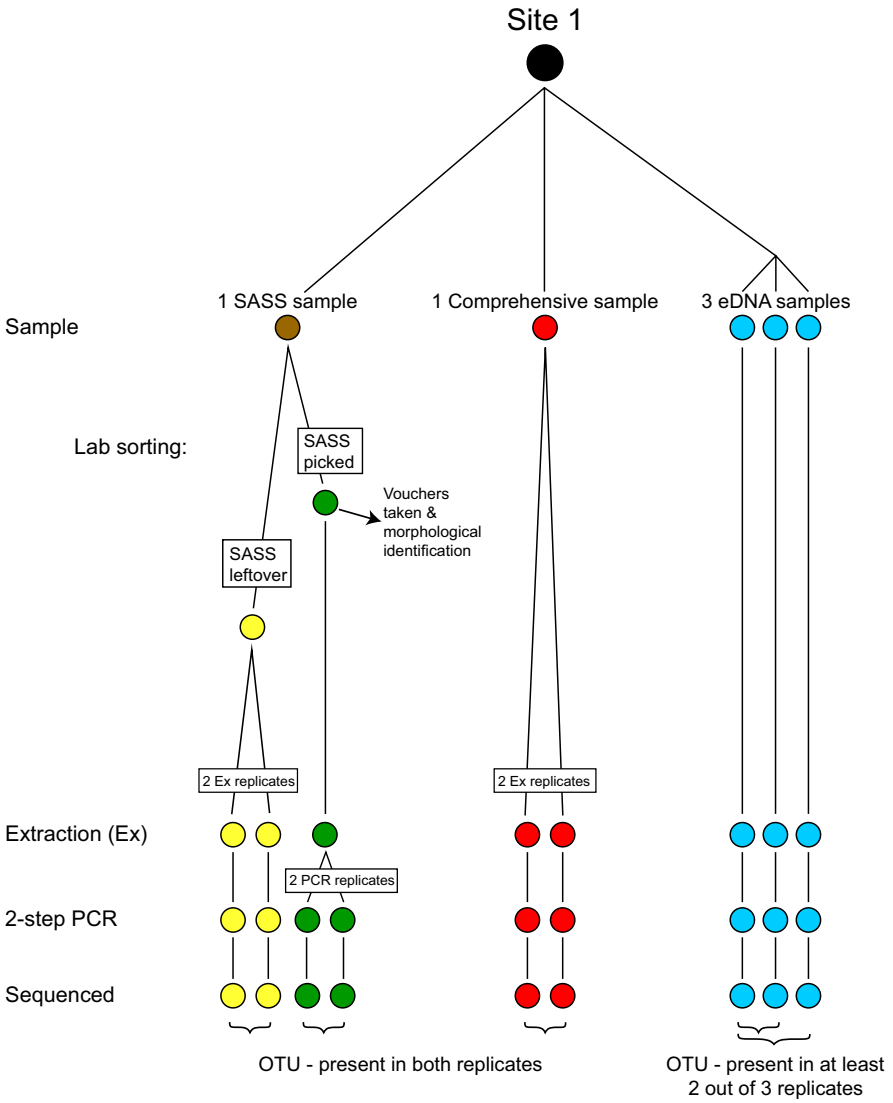

FIGURE 2 Sampling strategy per river site showing laboratory procedures and the corresponding technical extraction and PCR replicates used for the study

Gel Extraction Kit to remove a shorter nontarget band, although this can result in a loss of target DNA, thus introducing biases into the data, this sample was not shown to be different from other eDNA samples from the same site. Following clean-up, the DNA concentration of each individual library was measured with a SPECTROstar Nano (BMG Labtech) and then equimolar pooled with the negative control added at the maximum volume added for any single library (15 μl).

The size-corrected concentration of the pooled libraries was determined following analysis with an Agilent 2200 Tapestation system and Qubit 2.0 Fluorometer (Invitrogen). The pool was loaded onto an Illumina MiSeq at 9 pM, with 5% Phi-X, using a 600 cycle V3 kit with 300-bp paired-end sequencing (index read steps skipped).

2.5 | Bioinformatics

Bioinformatics processing was performed using JAMP v0.66 (https://github.com/VascoElbrecht/JAMP) with detailed scripts available as supporting information (Scripts S1). Reads were demultiplexed using JAMP, and paired-end reads were merged using Usearch v8.1.1861 (Edgar, 2013) with relaxed settings to maximize the number of reads merged (allowing up to 99 mismatches in the overlapping region). Where necessary, the reverse complement of the reads was generated, to ensure all sequences are present in the same orientation. BF2 and BR2 primers were then removed using cutadapt 1.18 with default settings (Martin, 2011), discarding reads where the primer sequences remained undetected. Only sequences of 411 to 431 bp were used for further analysis (filtered with cutadapt). Low-quality sequences were then filtered from all samples, using fastq_filter with maxee = 1 (Edgar & Flyvbjerg, 2015). Sequences from all samples were then pooled, dereplicated (minuniquesize = 2), and clustered into molecular operational taxonomic units (mOTUs), using cluster_otus with a 97% identity threshold (Edgar, 2013) which includes chimera removal. Prefiltered reads for all samples were dereplicated again, but singletons were included to maximize the information extracted from the sequence data. Sequences from each sample were matched against the mOTUs with a minimum match of 97% using usearch_global.

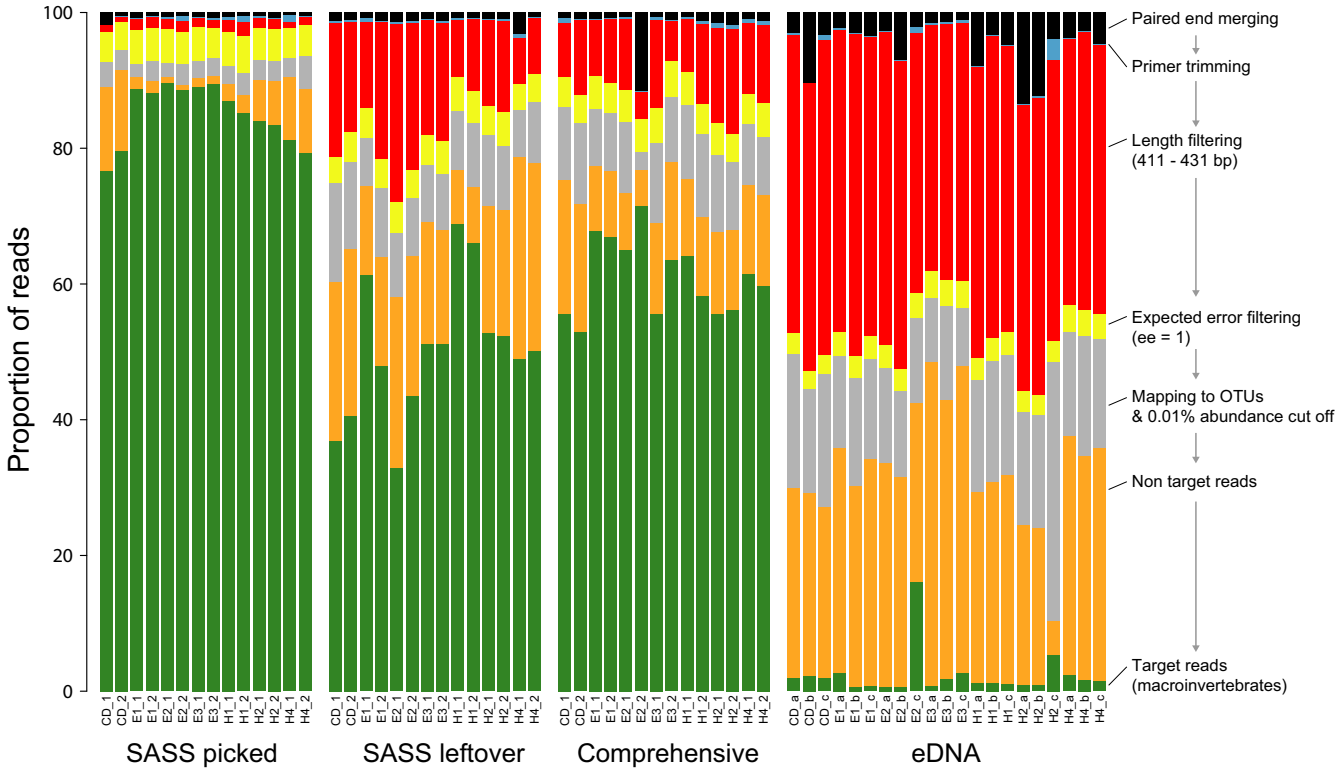

FIGURE 3 Proportion of reads lost during each processing step, an overview of sequences discarded from raw data, bioinformatics processing, and nontarget hits (e.g., bacteria) compared to the target macroinvertebrate taxa (green)

For each sample, only mOTUs with a read abundance above 0.01% in at least two technical replicates were considered for downstream analyses using a 3% divergence threshold which is consistent with other studies and observations of real invertebrate communities (e.g., Elbrecht, Peinert, & Leese, 2016; Elbrecht, Vamos, Steinke, et al., 2017; Hajibabaei, Janzen, Burns, Hallwachs, & Hebert, 2006).

Voucher sequences were uploaded to BOLD before taxonomy assignment on the DNA data commenced. Taxonomy was assigned to remaining mOTUs using R scripts to search against both BOLD (https://github.com/VascoElbrecht/JAMP/blob/master/JAMP/R/BOLD_web_hack.R) and NCBI (https://github.com/VascoElbrecht/JAMP) databases. The taxonomy assigned mOTU table was then filtered by Phylum to include targeted taxa (Arthropoda, Annelida, Porifera, Coelenterata, Turbellaria, Hydracarina, Gastropoda and Bivalvia), taxonomy was further validated and checked, removing any terrestrial invertebrates, and any conflicting assignments between BOLD and NCBI were handled individually (unresolved cases were removed). A rough guideline for taxonomy assignment using percent similarity was used, where a hit with 98% similarity was used for species, 95% for genus, 90% for family, and 85% for order levels. Potential overestimation of taxonomic diversity from the clustering algorithm due to numts or possible amplification errors was checked by aligning target mOTUs using MAFFT (Katoh & Standley, 2013) and examined for stop codons and indels, then analyzed using maximum likelihood in RAxML v8 (Stamatakis, 2014) and visualized using FigTree v.1.4.3 (Rambaut, 2009) to identify any long branches. These errors were found to be negligible as they affected a very low proportion of the mOTUs (0.01%). UpSetR (Conway, Lex, & Gehlenborg, 2017; Lex et al., 2014) was used to visualize the number of shared mOTUs/ families between each metabarcoding sampling method across all sites.

The family level taxa detected by DNA metabarcoding of the SASS picked and SASS leftover samples were then compared with those found using the standard morphological approach. Within the SASS picked samples, singletons were not included in the comparison as they were used for voucher sequencing (i.e., not available for metabarcoding in the SASS picked sample). The percentage of taxon overlap (at SASS level) of DNA methods with morphology was then calculated using R (version 3.5.3; R Core Team, 2019).

2.6 | Water quality assessment

Within South Africa, rivers are currently monitored using the South African Scoring System (SASS version 5) protocol (Dickens & Graham, 2002), which uses freshwater macroinvertebrates as a measure of stream ecosystem health. Organisms are identified to a mixed taxon (typically family) level (Figure S1) and are assigned a quality score based on pollution sensitivity (Dickens & Graham, 2002). The abundance of organisms at family level per live sample is roughly estimated into categories (where 1 = 1 individual, A = 2–10, B = 10–100, C = 100–1,000, D > 1,000) and recorded on the scoring sheet; however, it is not used to calculate the SASS scores. SASS returns three principal indices: (a) the Taxon Score, the total of quality scores

for all taxa found in a sample; (b) the Number of Taxa; and (c) the Average Score Per Taxon (ASPT) which is the Taxon Score divided by the Number of Taxa (Dickens & Graham, 2002). The SASS score and ASPT are then standardized across South Africa using biological bands that ranges from "A" to "E/F," indicating unmodified (A) to seriously modified (E/F) states which are calibrated for each ecoregion (Dallas, 2007; Dallas & Day, 2007; Kleynhans, Thirion, & Moolman, 2005; Omernik, 1987). As some natural variation will occur throughout the year, seasonal data (i.e., multiple samples) are required to capture a holistic view of the taxa occurring at a site (Dickens & Graham, 2002). For each site, the ecoregion classification was determined based on Kleynhans et al. (2005), and the total Taxon Score and Average Score Per Taxon (ASPT) values calculated for each method (morphology, SASS picked, SASS leftover, comprehensive and eDNA) and plotted against the biological bands for that region. The SASS picked and SASS leftover taxon lists for each sample were also combined and deduplicated to simulate a full SASS sample being processed as a single unit ("SASS combined").

2.7 | Historical species records

The taxa previously recorded for each river/sampling region were extracted from the database of the Albany Museum, Makhanda (previously called Grahamstown), which houses the largest freshwater invertebrate collection in Africa, including from the sites sampled in this study, and compared to the target mOTUs recovered by molecular methods. Only historical records identified at the genus or species level were used for comparison. Available data on the number of described species for the broader region (Eastern Cape/South Africa/southern Africa as applicable) were added to the analyses.

3 | RESULTS

3.1 | Sequencing statistics

The MiSeq run yielded 16 million reads from the 68 tagged samples (raw data available from https://doi.org/10.5281/zenodo.3462633; Sequence Read Archive (SRA) data code: PRJNA629361). After library demultiplexing, an average of 221,630 (*SD* = 69,340) read pairs was retained. After bioinformatic processing, a total of 17,660 mOTUs were detected which was then reduced to 5,117 mOTUs that were present in more than one technical replicate, of these a total of 404 "target" mOTUs (i.e., freshwater macroinvertebrates) were found across all samples (Table S2).

The proportion of reads lost from each processing step to the 404 target taxa found across samples and sites are shown in Figure 3. SASS picked samples represented the "cleanest" samples, with the lowest proportion of reads discarded (mostly under 20% discarded), and comprehensive and SASS leftovers showed similar results (ca. 40%–50% discarded), while over 90% of reads were discarded for the eDNA samples (Figure 3).

3.2 | Comparison of DNA-based methods

3.2.1 | mOTU level

When considering the number of mOTUs shared between methods across all sites (Figure 4a–c), only 57 of the 5,117 mOTUs (1%) found were shared between all four methods (Figure 4a). As expected, SASS picked returned mostly arthropods (237 mOTUs), 76% of which were also present in the SASS leftover samples. Across all sites, 84% of SASS picked mOTUs were shared with the comprehensive sample (Figure 4a). The eDNA samples showed the least overlap with other methods, but returned the highest number of mOTUs (3,446 mOTUs) most of which were unique (84%). The comprehensive samples (comprising 1,249 mOTUs) and SASS leftovers (comprising 1,780 mOTUs) shared a total of 841 mOTUs (Figure 4a).

For targeted SASS mOTUs (i.e., freshwater macroinvertebrates), only 42 of 404 mOTUs (10%) were shared between all four methods (Figure 4b). SASS leftover (227 mOTUs) and comprehensive samples (331 mOTUs) overlapped with SASS picked by 78% and 93% respectively, and shared 196 mOTUs. The eDNA samples recovered 80 target mOTUs, 19% of which were unique and not found by any other method. The comprehensive sample had the highest number of unique mOTUs (103) corresponding to 26% of all target mOTUs found (Figure 4b).

When comparing the number of operational taxonomic units (OTUs), either morphology-based or DNA-based, for the target taxa across the seven sites (Figure 5), the comprehensive method recovered the highest number of OTUs (69–116 OTUs across sites) compared to other methods. Morphology, SASS picked, and SASS leftover methods found similar numbers of OTUs per sample; however, there was only partial overlap in the SASS picked and SASS leftover samples, and thus, the total number of OTUs recovered per sample increased when these were artificially combined as SASS combined (Figure 5). The eDNA samples recovered the lowest number of target mOTUs (11–31 OTUs) (Figure 5).

3.2.2 | Family level

At the SASS (mostly family) level, 19 out of 57 taxa were shared between all four methods (Figure 4c). The three kick-net-based methods (SASS picked, SASS leftover and comprehensive) shared a further 14 taxa. The eDNA samples shared another 2 taxa with SASS leftovers, while each method, except SASS picked, found 3–4 unique family-level taxa (Figure 4c).

Comprehensive sample DNA metabarcoding generally found more target families than all the other methods (25.71 ± 4.75, mean ± *SD*) (Figure 6). SASS leftover recovered between 14 and 26 target families (19.14 ± 3.72), and SASS picked recovered between 15 and 25 target families (20.86 ± 3.63) while taxon recovery from the eDNA samples was consistently lower than other methods, between 5 and 16 target families (10.29 ± 3.82; Figure 6).

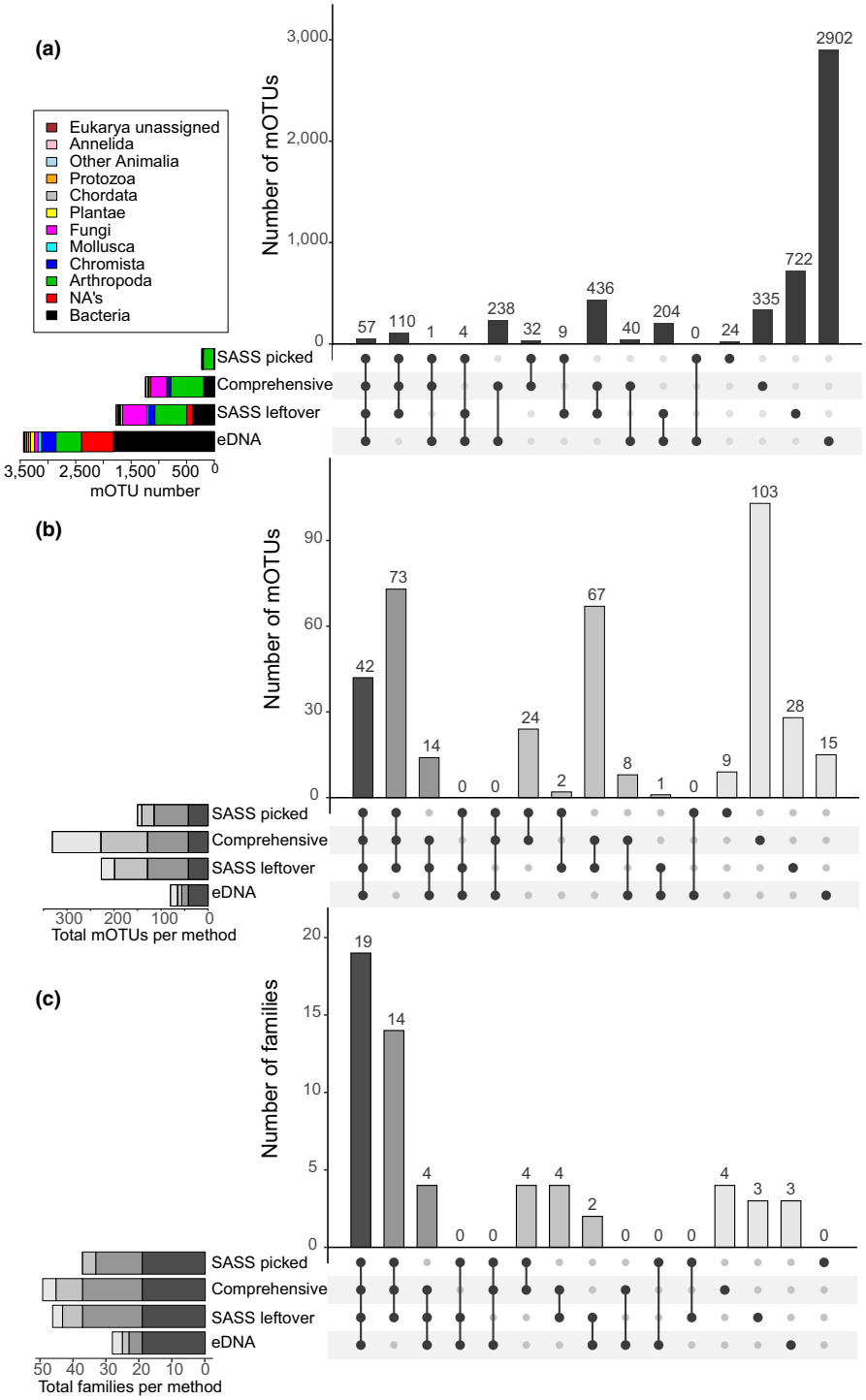

FIGURE 4 *UpSet bar plot showing shared mOTUs between each sampling method across all sites. The last four bars in each plot reflect the remaining mOTUs unique in each method. (a) For the 5,117 mOTUs and (b) of the 404 mOTUs representing target freshwater invertebrates listed in the SASS protocol. (c) Overlap of family level taxa between DNA-based methods*

3.3 | Comparison of morphological and molecular id of SASS samples

3.3.1 | Morphospecies/mOTU level

A total of 319 morphospecies were picked from the SASS samples, identified to mixed taxon level and Sanger sequenced to be used as vouchers. Of these, duplicate species as identified by COI sequences and nontarget taxa (i.e., terrestrial invertebrates) were removed, leaving 254 unique freshwater macroinvertebrate morphotaxa which were subsequently uploaded to BOLD. Species-level assignment was poor as the reference library for South African macroinvertebrate fauna is limited, and only 42 out of 254 morphospecies and 57 out of 404 mOTUs could be assigned to species.

3.3.2 | Family level

At the SASS family level, morphological-based identification resulted in 51 taxa (Table S3) with a mean of 21.86 ± 2.97 (mean ± *SD*) taxa

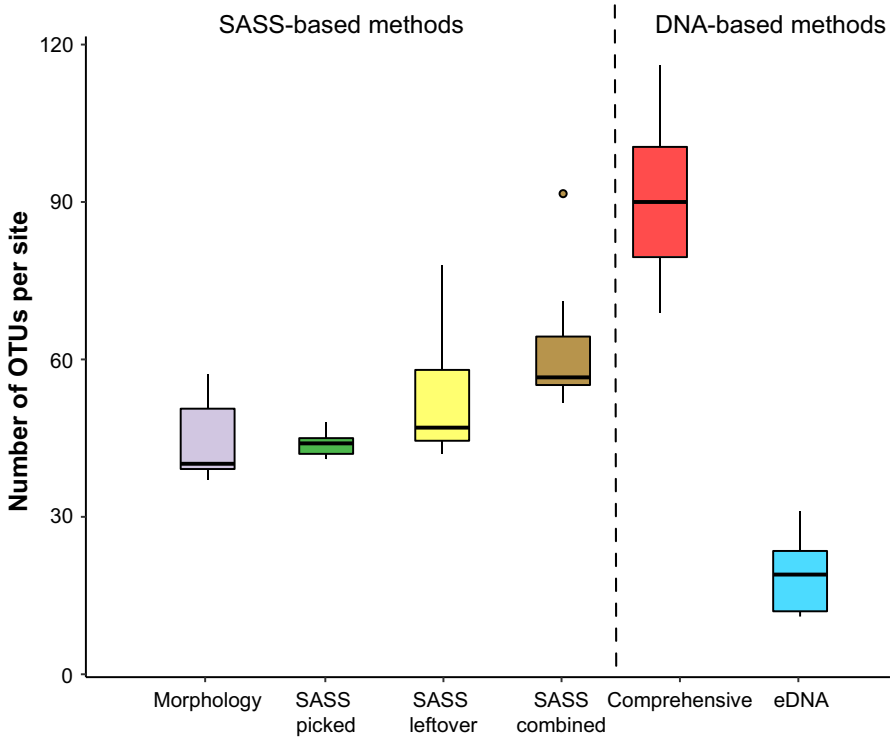

FIGURE 5 Box and whisker plot showing the median and interquartile ranges for SASS-based and DNA-based methods at the out level over the seven sites

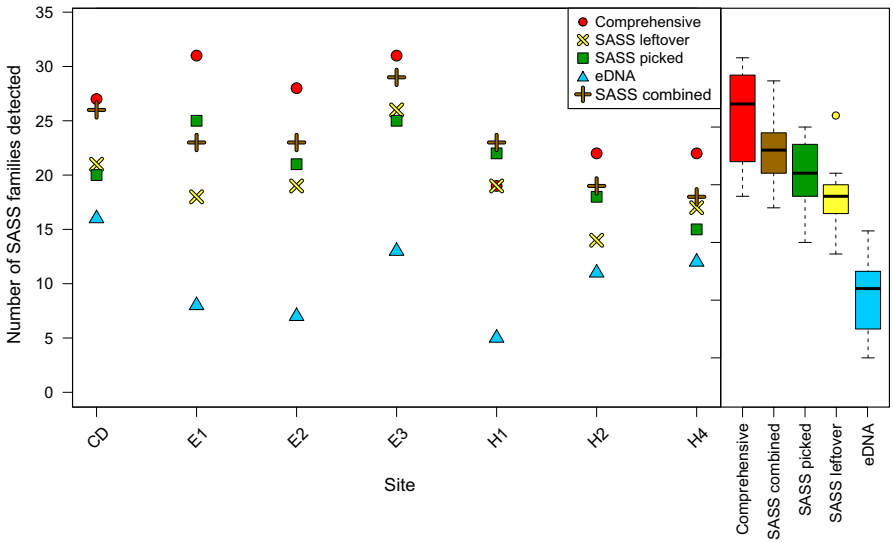

FIGURE 6 Number of family-level taxa detected by DNA-based methods (comprehensive, SASS combined, SASS picked, SASS leftover, and eDNA samples) across sites

at each site (Figure 7a). SASS leftover and SASS picked samples recovered similar results at the family level (as mentioned above), and SASS picked was most similar to morphology (mean = 20.86 ± 3.63), as expected given the morphotaxa came directly from these samples. Four SASS taxa were not recovered by DNA metabarcoding across the SASS picked samples: Ancylidae (Gastropoda) and Hydroptilidae (Trichoptera) were missing from 2 of the 7 samples, while Gyrinidae and Chironomidae were each missing from one sample. Seven SASS families that were not recorded by morphology in some SASS picked samples were recovered by DNA metabarcoding: Caenidae, Notonemouridae, Philopotamidae, Pisuliidae, Teloganodidae, and Tipulidae.

The overlap of SASS-based molecular methods with morphology (Figure 7b) showed that SASS picked samples had the highest overlap with morphology at 74%–100%, failing to detect between 0 and 5 family-level taxa (mean = 2.29 ± 1.80) across all samples, whereas SASS leftover samples failed to detect an average of 34.1% taxa (6–11 taxa, mean = 7.43 ± 1.81).

3.4 | Comparison of water quality assessments

Water quality assessment metrics calculated for SASS-level morphology and DNA metabarcoding data were generally similar across

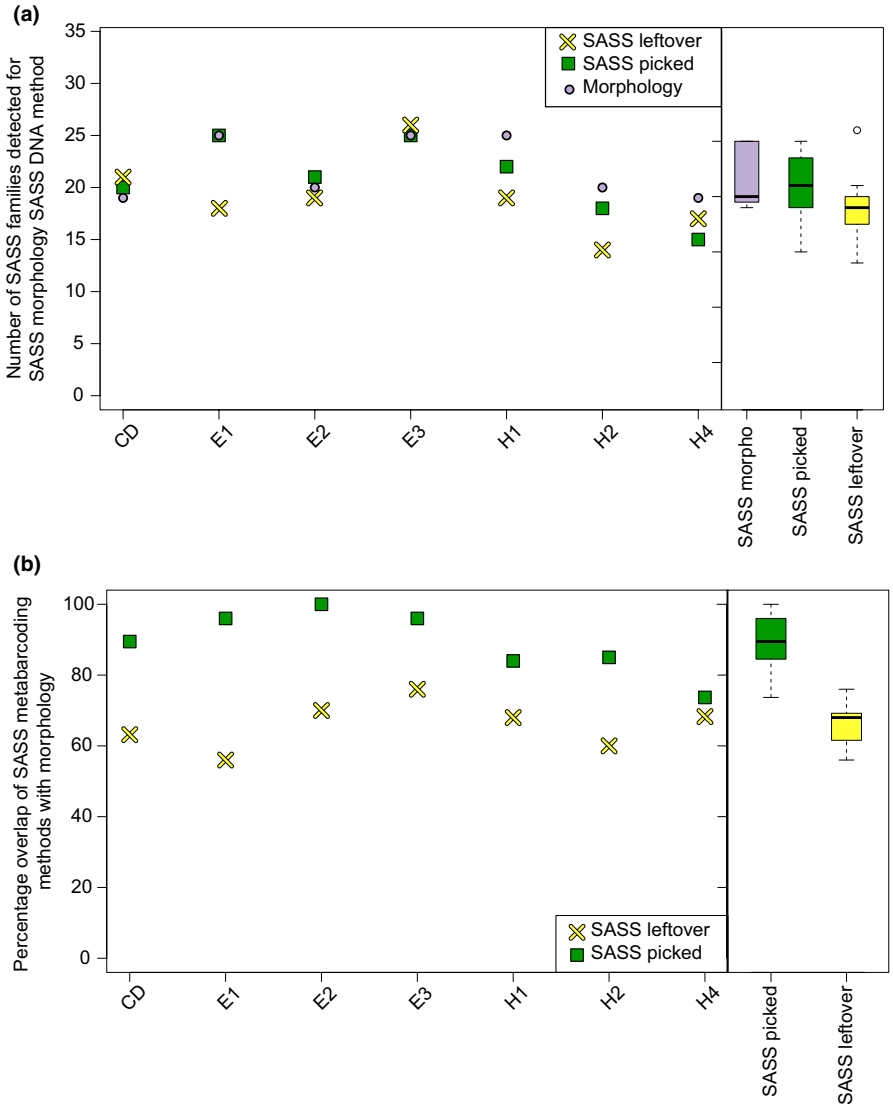

FIGURE 7 (a) Number of SASS mixed taxon-level taxa detected by morphological SASS and DNA-based SASS methods (SASS picked and SASS leftover) across sites. (b) Percentage overlap of the taxa detected by DNA-based SASS methods (SASS picked and SASS leftover) with morphology

methods when interpreted using Ecoregion Level 1 classified biological bands (Dallas, 2007), except for some eDNA samples (Figure 8a-c). Although SASS morphology was scored in the laboratory instead of in the field, it is noteworthy that the SASS results obtained from this study were consistent with SASS results obtained from the same river sites during previous monitoring events (Helen Barber-James, SASS Accreditor, pers. comm.). The total taxon scores varied considerably across sites sampled (Figure 8a–c), even based on morphology. However, ASPT was less variable and is recommended as the common SASS index of choice (Dickens & Graham, 2002). The variation is particularly evident in the Craigdoone site (Figure 8a), where the total taxon scores differed by over 75 units while ASPT remained very similar between methods.

3.5 | MOTU recovery compared to historical records

Comparing mOTUs in these three rivers against historical records of species from the region highlighted several groups with potentially high levels of cryptic diversity (Table 1). This was especially apparent within the true flies (Diptera) where molecular methods found a total of 101 mOTUs for Chironomidae (comprehensive = 88; SASS picked = 15; SASS leftover = 65; eDNA = 21; see Figure S3) whereas current museum records from these rivers include two species, and an estimated 30 species from the wider Eastern Cape province. Within the Simuliidae (Diptera), molecular methods found 34 mOTUs (comprehensive = 31; SASS picked = 23; SASS leftover = 29; eDNA = 11; see Figure S3) whereas current museum records from these rivers include 12 species, and 23 species known from the wider Eastern Cape province. Within the mayflies (Ephemeroptera), two families showed high levels of cryptic diversity; in the Baetidae (Ephemeroptera), molecular methods found 24 mOTUs (comprehensive = 24; SASS picked = 23; SASS leftover = 19; eDNA = 8; see Figure S3) whereas museum records from these rivers include 17 species, and 37 species known from the entire region of southern Africa. Within the Leptophlebiidae (Ephemeroptera), molecular methods found 9 mOTUs (comprehensive = 8; SASS picked = 9; SASS leftover = 8; eDNA = 4; see Figure S3) whereas museum records from these rivers include 5 species, and 20 species known from the entire region of southern Africa. Within

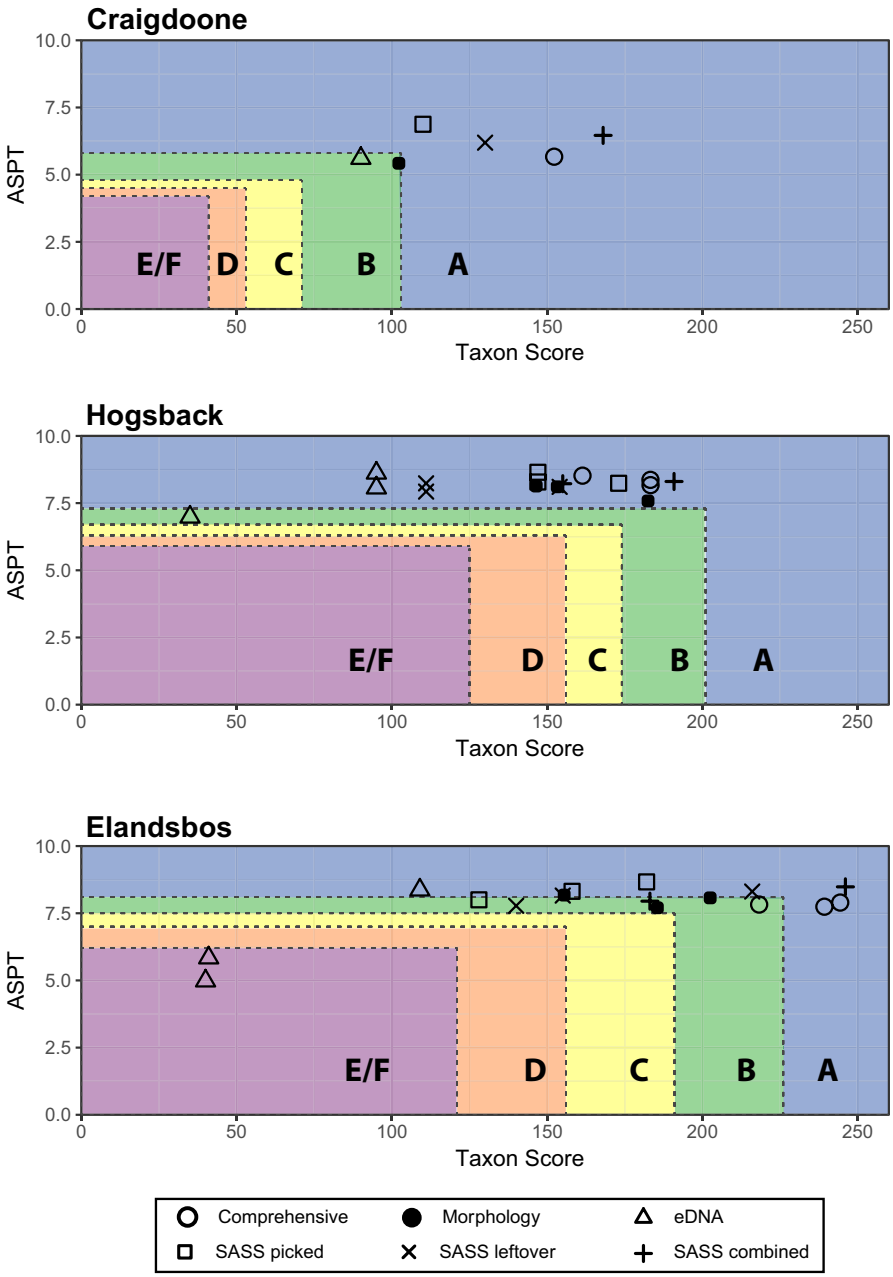

FIGURE 8 Plot of Taxon Score and ASPT scores, showing results for each site (Craigdoone [C], Hogsback [H] and Elandsbos [E]) according to method. Biological bands are shown for each region, calculated through the intersection of total score and ASPT: A = natural (blue), B = good (green), C = fair (yellow), D = poor (red), and E/F = seriously modified (purple)

the stoneflies (Plecoptera), the Notonemouridae exhibited high levels of potential cryptic diversity with the molecular methods finding 13 mOTUs (comprehensive = 13; SASS picked = 7; SASS leftover = 7; eDNA = 1; see Figure S3) whereas current museum records from these rivers include 2 species, and 14 species known from all of South Africa (Table 1).

Within the dragonflies and damselflies (Odonata), the family Platycnemididae exhibited very high levels of potential cryptic diversity with the molecular methods finding 10 mOTUs (comprehensive = 5; SASS picked = 6; SASS leftover = 2; eDNA = 0; see Figure S3) whereas museum records from these rivers include 1 species, with only 3 species known from the wider Eastern Cape province. Within the Oligochaeta which are typically not identified further in the SASS protocol, molecular methods found 42 mOTUs (comprehensive = 34; SASS picked = 8; SASS leftover = 28; eDNA = 2; see Figure S3) whereas current museum records from these rivers include 1 species, with 29 species known from the wider Eastern Cape province. Within the Potamonautidae (Crustacea), molecular methods found 23 mOTUs (comprehensive = 22; SASS picked = 4; SASS leftover = 3; eDNA = 2; see Figure S3) whereas museum records from these rivers include 2 species, with 22 species known from all of South Africa (Table 1).

4 | DISCUSSION

4.1 | Taxon recovery by different DNA-based methods

In this study, we compared four DNA-based methods for recovering biodiversity at the family and the mOTU level in the context of water

TABLE 1 Number of mOTUs recovered for each family with DNA-based methods, compared to Albany Museum historical specimen records of species found at the rivers sampled in this study

Taxon	# mOTUs	# Historical records	# Species known from region	Region
Chironomidae	101	2	30	EC[1]
Oligochaeta	42	1	29	EC[2]
Simuliidae	34	12	23	EC[3]
Baetidae	24	17	37	Southern Africa[4]
Potamonautidae	23	2	22	SA[5]
Notonemouridae	13	4	8	EC[6]
Hydracarina	10	8	23	EC[7]
Platycnemididae	10	1	3	EC[8,9]
Coenagrionidae	9	10	15	EC[8,9]
Gyrinidae	9	4	24	EC[10]
Leptophlebiidae	9	5	20	Southern Africa[4]
Leptoceridae	8	6	36	WC/EC[11]
Ceratopogonidae	5	2	142	Southern Africa[12]
Dytiscidae	5	3	27	EC[13]
Elmidae	5	-	5	EC[14]
Scirtidae	5	1	?	Unknown[15]
Gomphidae	4	5	5	EC[8,9]
Hydropsychidae	4	5	9	WC/EC[11]
Synlestidae	2	5	6	EC[8,9]
Aeshnidae	1	4	6	EC[8,9]
Libellulidae	1	8	22	EC[8,9]

Note: Regional records for families with number of mOTUs or historical records >5 and for those with regional species information readily available. Regional information available includes Eastern Cape (EC), Western Cape (WC), South Africa (SA), and southern Africa.

Sources: [1]Harrison, 2004; [2]van Hoven & Day, 2002; [3]de Moor, 2002b; [4]Barber-James & Lugo-Ortiz, 2003; [5]Daniels, Busschau, & Cumberlidge, 2019; [6]Stevens & Picker, 2003; [7]Jansen van Rensburg & Day, 2002; [8]Samways & Wilmot, 2003; [9]Simaika & Samways, 2009; [10]Stals, 2008; [11]de Moor & Scott, 2003; [12]de Meillon & Wirth, 2002; [13]Biström, 2008; [14]Nelson, 2008; [15]Endrödy-Younga & Stals, 2008; [16]Harrison, Prins, & Day, 2002

quality biomonitoring and rapid biodiversity assessment. We found that the comprehensive sampling method, including whole sample homogenization, recovered the highest number of target taxa and the highest number of unique target taxa at both the mOTU and family levels, while our eDNA sampling and processing method found the most mOTUs overall but the fewest target taxa (Figures 4 and 6).

Similarly, whole sample homogenization of the SASS leftover samples recovered more target taxa at both the mOTU and family levels than the picked SASS samples, suggesting taxa are missed in the sample sorting process (Haase et al., 2006, 2010) or present as gut contents or free DNA. Although not experimentally validated in our study, there appeared to be no evidence of sample inhibition due to the substantial debris in both the comprehensive and SASS leftover samples, likely due to our use of a soil DNA extraction kit which includes an inhibitor removal step. Thus, the dual benefits of whole sample homogenization are that sample processing time is dramatically reduced, and fewer taxa are missed; indeed, more freshwater macroinvertebrate studies are focusing on whole-sample homogenization (Andújar et al., 2018; Dowle et al., 2016; Gardham et al., 2014; Hajibabaei et al., 2011; Majaneva, Diserud, Eagle, Boström, et al., 2018); however, these studies have partially or completely removed debris before homogenization.

The greater taxon recovery in the comprehensive sample over the combined SASS samples suggests that a longer sampling time with a smaller size mesh net captures more of the local macroinvertebrate diversity. This is likely due to the combination of greater sampling effort and the inclusion of taxa that would typically be considered as "out of season" or undetectable due to their life stage (e.g., eggs/ small larvae). Thus, the comprehensive sampling approach has the potential to reduce the number of sampling visits to a site than currently required by the SASS protocol; however, this result needs further testing.

While our eDNA sampling and processing method found up to 10 times more mOTUs than the other methods, a remarkable proportion of reads (over 90%, Figure 3) were nontarget and discarded during bioinformatics processing. These nontarget reads are likely due to the highly degenerate primers used, which may

be unsuitable for eDNA-based studies of freshwater macroinvertebrates (Macher et al., 2018; Smith et al., 2012). As a result, the number of target taxa detected with eDNA was lower than for all other methods. Amplicon size may also explain the lower number of target taxa in the eDNA samples than the other methods, and amplifying a shorter region of COI may have resulted in obtaining more freshwater macroinvertebrates in the eDNA samples from degraded eDNA. Although the eDNA samples recovered fewer taxa than the kick-samples, eDNA found more mOTUs for the Gastropoda, Amphipoda, Porifera (sponges), Coelenterata, and Platyhelminthes than any other method (Figure S3), suggesting eDNA may complement traditional kick-sampling methods for these groups.

While these results show that eDNA as currently sampled and analyzed is an unlikely alternative to sampling whole organisms for detection of whole macroinvertebrate communities in the near term (Barnes & Turner, 2016; Blackman et al., 2019; Hajibabaei et al., 2019; Macher et al., 2018), the field is advancing at a rapid rate (Blackman et al., 2019; Deiner et al., 2017; Deiner, Fronhofer, et al., 2016; Deiner, Walser, et al., 2016; Li et al., 2018; Mächler et al., 2019; Majaneva, Diserud, Eagle, Boström, et al., 2018). While our stringent two-out-of-three replicate approach should substantially reduce OTUs remaining which are a result of sequencing error, it is likely to have resulted in target taxa being lost from the eDNA samples at this sequencing depth.

4.2 | Water quality assessment through DNA-based and morphological methods

Despite differences in taxon recovery between all methods at the family level, results for water quality assessment metrics were very similar across morphology-based and DNA-based methods. All rivers sampled are considered to be in a "good" to "natural" state (biological bands A and B), which was reflected in both the morphology-based and DNA-based results. While these results suggest that metabarcoding (with the exception of eDNA in this case) can produce usable data for current assessment techniques, the DNA methods tested here are able to detect more families than are be detected with morphology, and thus, their inclusion in the current morphology-based SASS indices could lead to inflated results. While these results are promising for DNA-based water quality assessment in South Africa, it is clear that further research at sites which represent the full spectrum of water quality bands (poor to good quality) is required to facilitate intercalibration of these new DNA methods.

For most biodiverse countries, DNA reference libraries are still a major limiting factor (e.g., Venter & Bezuidenhout, 2016); however, most taxa can at least be assigned confidently to family level or below, using current data on BOLD, as was the case for the target taxa in this study, suggesting that reference libraries may not hinder the incorporation of DNA metabarcoding at family level into the current SASS protocol.

When comparing morphospecies with the SASS picked mOTUs (essentially the same sample), some of the discrepancies may be due to several factors, including primer bias, rare taxa, the amplification of gut contents, the presence of DNA or free cells attached to the body of the organisms or contamination from DNA transfer during identification steps—although the latter is unlikely due to careful handling and the use of decontaminated equipment.

Our results show that while some DNA-based metabarcoding methods detect more diversity than morphology, even at family level, they do not always find the same taxa, as has been reported in other studies (Carew et al., 2013; Clarke, Beard, Swadling, & Deagle, 2017; Elbrecht, Vamos, Meissner, et al., 2017; Gibson et al., 2015; Lejzerowicz et al., 2015; Zimmermann, Glöckner, Jahn, Enke, & Gemeinholzer, 2015).

Our samples differed considerably in both the number of taxa and the total taxon score (Figure 8), yet when compared at the SASS level of biological bands, these differences became minimal. A number of studies have warned that family-level identification, being a compromise between accuracy and the level of taxonomic expertise required, is not precise enough to assess water quality changes, primarily due to the different habitat requirements and pollution tolerances found between species within the same family (Barber-James & Pereira-da-Conceicoa, 2016; de Moor, 2002a; Odume, Muller, Arimoro, & Palmer, 2012; Odume, Palmer, Arimoro, & Mensah, 2015). The current lack of data on the habitat requirements and pollution tolerances of species within some groups would prove problematic in a water quality monitoring context, especially in biodiverse but water-stressed areas such as South Africa. However, a DNA-based approach may rapidly provide these data over large taxonomic and geographic scales when combined with relevant abiotic information from the sites sampled and the inclusion of reference sites for calibration.

In the context of SASS water quality biomonitoring in South Africa, our results highlight that eDNA has the potential to be incorporated as it found similar water quality scores to the morphological-based methods in 5 out of 7 sites; however, additional research is required to explore macroinvertebrate eDNA metabarcoding across the different water quality bands used in South Africa and the use of more specific primer sets and/or a greater sequencing depth. While sampling for eDNA may be influenced by weather conditions (high precipitation can result in large amounts of sediments that may clog filters) and natural variation, including pH and turbidity, our results show the potential use of eDNA in assisting with ecosystem management by offering a unique method of biomonitoring, and once established, could provide a much safer and quicker way of monitoring rivers, especially those that harbor dangerous animals.

4.3 | Taxon recovery at finer taxonomic resolution & mOTU versus historical records

Our results indicate that DNA-based methods offer a finer taxonomic resolution compared to morphology-based identification, in agreement with previous studies (Baird & Sweeney, 2011; Elbrecht,

Vamos, Meissner, et al., 2017; Gibson et al., 2015; Stein et al., 2014; Sweeney, Battle, Jackson, & Dapkey, 2011). In particular groups which are harder to identify, like larval Diptera (Chironomidae and Simuliidae in particular), mites (Hydracarina) and Oligochaeta are often not included at a fine taxonomic resolution in biomonitoring protocols; however, DNA metabarcoding is a powerful tool in this respect and is able to uncover remarkable mOTU diversity for these groups. Across the sites sampled in our study, 101 chironomid mOTUs were found using the comprehensive method, while less than a third of that was picked from the SASS samples as morphospecies (Table 1), and this diversity is usually recorded as a single family and potentially misrepresented by a relatively low SASS taxon score. While the high level of cryptic diversity in chironomids has been found in temporary wetlands (Theissinger et al., 2018), a similar pattern was seen in nearly all target groups in this study (Table 1), which is surprising given our limited sampling in only three rivers.

While it is argued that the mOTU approach overestimates recognizable species (Clare, Chain, Littlefair, & Cristescu, 2016; Flynn, Brown, Chain, Macisaac, & Cristescu, 2015), access to morphology-based identification expertise has always limited the level to which taxa (and therefore patterns of biodiversity) can be identified.

Although there is concern with the conceptual and practical difficulties of translating mOTUs to species (Brown, Chain, Crease, Macisaac, & Cristescu, 2015), the strength of a mOTU approach is the potential to estimate diversity and ensure comparability in the absence of described species (Blaxter et al., 2005; Clare et al., 2016) and is not a species concept (Floyd, Abebe, Papert, & Blaxter, 2002). This is useful in cases of cryptic diversity (Delić, Trontelj, Rendoš, & Fišer, 2017), where community composition data are grossly underestimated using morphological approaches because the largest and most diverse groups (e.g., Chironomidae) are grouped together or omitted from analyses.

It is unrealistic to wait for these reference databases to be built first, and it is important to note that developing a metabarcoding tool despite limited availability of reference libraries is that the data can be "time stamped." These data can then be reanalyzed as reference libraries are developed in parallel. By following a standardized approach, the analysis of the effects of single or multiple stressors will be possible, even for cryptic and undescribed species (Beermann et al., 2018; Macher et al., 2016). For those genera and species that can be identified using DNA, assessments can be further optimized by including and integrating species functional diversity, ecological preferences, and the effects of ecosystem stressors, especially for indicator taxa (Macher et al., 2016). This detailed level of information could then be used for biomonitoring, biodiversity, and ecological impact assessments, providing detailed evidence that can be used to inform managers, governments, and policy.

4.4 | Limitations and future considerations

Throughout this study, a number of limitations were identified. Firstly, molecular OTU diversity as defined by the clustering algorithm can overestimate taxonomic diversity (Clare et al., 2016); however, more aggressive filtering runs the risk of losing true signal, and so more emphasis is made here on the mOTUs as estimates, which should be used with caution. While mOTU estimates are unreliable and depend on sequencing processing, sequencing depth and bioinformatics (Holovachov, Haenel, Bourlat, & Jondelius, 2017) rather than actual diversity in the sample, the use of reference databases improves their reliability. This approach can be problematic with incomplete reference databases, such as the case with South African macroinvertebrates, and so this was partly overcome by building a reference database in parallel. Prior alignment screening for stop codons and indels showed negligible influence on the mOTU results of this study; however, putative species inferences could be further improved by using more robust procedures such as phylogenetic tree inferences or gap analysis (Flynn et al., 2015).

The differing physical nature of samples required different laboratory pipelines, which may have biased the results, and while these biases are unlikely to affect the main conclusions of this study, the comparisons need to be considered with caution. Furthermore to reduce the influence of PCR stochasticity, especially with eDNA samples, more PCR replicates should be included in future (Beentjes, Speksnijder, Schilthuizen, Hoogeveen, & Van Der Hoorn, 2019) and the efficiency of species detection can be further improved by using a multispecies occupancy modeling framework to determine the appropriate replicate numbers (Doi et al., 2019).

5 | CONCLUSIONS

DNA-based techniques, especially those using whole sample homogenization, are not only able to recover the same water quality assessments as morphological SASS, but are able to identify taxa to a much finer resolution. The intensive-search "comprehensive" sampling out-performed all other methods for the number of focal taxa detected, highlighting that the standardized SASS sampling method does miss relevant target taxa present in the environment, potentially due to the larger mesh size or the limited SASS sampling time. Comprehensive sampling with whole-sample homogenization not only offers the potential to drastically reduce sample processing time, but also potentially collects out-of-season taxa reducing the seasonal variation in SASS scores at a site. While comprehensive sample homogenization is a method that has demonstrated promising results, other techniques using fixatives from bulk samples are being developed (Blackman et al., 2019; Carew, Coleman, & Hoffmann, 2018; Erdozain et al., 2019; Martins et al., 2019; Zizka, Leese, Peinert, & Geiger, 2019) and would be advantageous in that the sample remains intact for morphological identification if needed.

This study focussed on the official SASS species list of macroinvertebrates; however, the strength of eDNA methods for wider catchment-based assessments and detection of ecological condition is promising (Blackman et al., 2019) and has the potential to benefit biomonitoring protocols by detecting different organisms, such as highly diverse microbial communities and taxa of priority

conservation concern, including vertebrates. New indices will need to be developed to take advantage of these eDNA-based approaches.

This study shows that unsorted samples, including those with significant debris, recover the same water quality scores as a morphology-based assessment and recover much higher diversity scores than both picked and eDNA samples. This approach has significant advantages in that it provides both higher taxon-based water quality assessment and species-level data at the same time. As a result it is possible to integrate DNA-based approaches into existing metrics while providing much more information for the development of more refined metrics and, as reference libraries are developed, provides data that can be reused for conservation and biodiversity management.

ACKNOWLEDGMENTS

This research was funded by the Department of Business, Energy & Industrial Strategy Rutherford Fund. Florian Leese, Edith Vamos, Cristina Hartmann, and Chris Hempel are thanked for providing the indexed BF2/BR2 primers. Steve Russell (NHM) is thanked for barcoding the voucher specimens. Many thanks to Albany Museum staff who helped us in the field, Alex Holland, Musa Mlambo, Nonkazimulo Mdidimba, Ina Ferreira, and Zezethu Mnqeta. Rhodes University students, Lauren James and Jacqui James. Juan Tedder (GroundTruth), Kirsty Venter, and James Pereira da Conceicoa. Special thanks to Dirk Roux and Nerina Kruger (SANPARKS) for assisting with permits and accommodation and Melanie de Morne (SANPARKS) for assisting in the field. Research material was collected under a once off sampling permit for the Tsitsikamma Section—Garden Route National Park issued on 2017/11/10 by Nerina Kruger, permit number 0056_AAA041-0053 and in the Eastern Cape under CRO 69/17CR and CRO 70/17CR. We thank the two reviewers for their constructive and knowledgeable input, which has added great value to the manuscript.

CONFLICT OF INTEREST

None declared.

AUTHOR CONTRIBUTIONS

LP and BP designed the study. LP, HB-J, and BP carried out sample collection. LP, BP, HB-J, AH, and AB generated the data. LP, VE, and AB analyzed the data. LP, VE, and BP wrote the original manuscript, and all authors contributed to revisions and accepted the final version.

DATA AVAILABILITY STATEMENT

Raw MiSeq data can be found here: https://doi.org/10.5281/zenodo.3462633

ORCID

Lyndall Pereira-da-Conceicoa https://orcid.org/0000-0001-9192-3051
Vasco Elbrecht https://orcid.org/0000-0003-4672-7099
Andie Hall https://orcid.org/0000-0001-5546-7255
Andrew Briscoe https://orcid.org/0000-0002-9422-6056
Helen Barber-James https://orcid.org/0000-0002-1029-9275
Benjamin Price https://orcid.org/0000-0001-5497-4087

REFERENCES

Andújar, C., Arribas, P., Gray, C., Bruce, C., Woodward, G., Yu, D. W., & Vogler, A. P. (2018). Metabarcoding of freshwater invertebrates to detect the effects of a pesticide spill. *Molecular Ecology*, *27*(1), 146–166. https://doi.org/10.1111/mec.1441026

Baird, D. J., & Hajibabaei, M. (2012). Biomonitoring 2.0: A new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. *Molecular Ecology*, *21*, 2039–2044. https://doi.org/10.1111/j.1365-294X.2012.05519.x

Baird, D. J., & Sweeney, B. W. (2011). Applying DNA barcoding in benthology: The state of the science. *Journal of the North American Benthological Society*, *30*(1), 122–124. https://doi.org/10.1899/10-154.1

Barber-James, H. M., & Lugo-Ortiz, C. (2003). Ephemeroptera. In I. de Moor, J. Day, & F. de Moor (Eds.), *Guides to the freshwater invertebrates of Southern Africa Volume 7: Insecta I: Ephemeroptera, Odonata & Plecoptera* (pp. 16–159). University of Cape Town. Water Research Commission Report No: TT 207/03.

Barber-James, H., & Pereira-da-Conceicoa, L. (2016). Efficacy and deficiencies of rapid biomonitoring in biodiversity conservation: A case study in South Africa. *African Journal of Aquatic Science*, *41*(3), 337–343. https://doi.org/10.2989/16085914.2016.1192019

Barnes, M. A., & Turner, C. R. (2016). The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics*, *17*(1), 1–17. https://doi.org/10.1007/s10592-015-0775-4

Bates, B., Kundzewicz, Z. W., Wu, S., & Palutikof, J. P. (Eds.) (2008). *Climate change and water. Technical Paper of the Intergovernmental Panel on Climate Change* (pp. 1–210). Geneva, Switzerland: IPCC Secretariat.

Beentjes, K. K., Speksnijder, A. G. C. L., Schilthuizen, M., Hoogeveen, M., & Van Der Hoorn, B. B. (2019). The effects of spatial and temporal replicate sampling on eDNA metabarcoding. *PeerJ*, *2019*(7), 1–18. https://doi.org/10.7717/peerj.7335

Beermann, A. J., Elbrecht, V., Karnatz, S., Ma, L., Matthaei, C. D., Piggott, J. J., & Leese, F. (2018). Multiple-stressor effects on stream macroinvertebrate communities: A mesocosm experiment manipulating salinity, fine sediment and flow velocity. *Science of the Total Environment*, *610–611*, 961–971. https://doi.org/10.1016/j.scitotenv.2017.08.084

Biström, O. (2008). Dytiscidae. In R. Stals, & I. de Moor (Eds.), *Guides to the freshwater invertebrates of Southern Africa. Volume 10, Coleoptera* (pp. 69–84). Pretoria, South Africa: Water Research Commission.

Blackman, R. C., Constable, D., Hahn, C., Sheard, A. M., Durkota, J., Hänfling, B., & Handley, L. L. (2017). Detection of a new non-native freshwater species by DNA metabarcoding of environmental samples – first record of *Gammarus fossarum* in the UK. *Aquatic Invasions*, *12*(2), 177–189. https://doi.org/10.3391/ai.2017.12.2.06

Blackman, R., Mächler, E., Altermatt, F., Arnold, A., Beja, P., Boets, P., … Deiner, K. (2019). Advancing the use of molecular methods for routine freshwater macroinvertebrate biomonitoring – the need for calibration experiments. *Metabarcoding and Metagenomics*, *3*, 49–57. https://doi.org/10.3897/mbmg.3.34735

Blaxter, M., Mann, J., Chapman, T., Thomas, F., Whitton, C., Floyd, R., & Abebe, E. (2005). Defining operational taxonomic units using DNA barcode data. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1462), 1935–1943. https://doi.org/10.1098/rstb.2005.1725

Bongard, T. (2011). Detecting invertebrate species change in running waters: an approach based on the sufficient sample size principle. *Open Environmental & Biological Monitoring Journal*, *4*(1), 72–82. https://doi.org/10.2174/1875040001104010072

Brown, E. A., Chain, F. J. J., Crease, T. J., Macisaac, H. J., & Cristescu, M. E. (2015). Divergence thresholds and divergent biodiversity estimates: Can metabarcoding reliably describe zooplankton communities? *Ecology and Evolution*, *5*(11), 2234–2251. https://doi.org/10.1002/ece3.1485

Carew, M. E., Coleman, R. A., & Hoffmann, A. A. (2018). Can non-destructive DNA extraction of bulk invertebrate samples be used

for metabarcoding? *PeerJ*, *6*, e4980. https://doi.org/10.7717/peerj.4980

Carew, M. E., Nichols, S. J., Batovska, J., St Clair, R., Murphy, N. P., Blacket, M. J., & Shackleton, M. E. (2017). A DNA barcode database of Australia's freshwater macroinvertebrate fauna. *Marine and Freshwater Research*, *68*(10), 1788–1802. https://doi.org/10.1071/MF16304

Carew, M. E., Pettigrove, V. J., Metzeling, L., & Hoffmann, A. A. (2013). Environmental monitoring using next generation sequencing: Rapid identification of macroinvertebrate bioindicator species. *Frontiers in Zoology*, *10*(1), 45. https://doi.org/10.1186/1742-9994-10-45

Clare, E. L., Chain, F. J. J., Littlefair, J. E., & Cristescu, M. E. (2016). The effects of parameter choice on defining molecular operational taxonomic units and resulting ecological analyses of metabarcoding data1. *Genome*, *59*(11), 981–990. https://doi.org/10.1139/gen-2015-0184

Clarke, L. J., Beard, J. M., Swadling, K. M., & Deagle, B. E. (2017). Effect of marker choice and thermal cycling protocol on zooplankton DNA metabarcoding studies. *Ecology and Evolution*, *7*(3), 873–883. https://doi.org/10.1002/ece3.2667

Conway, J. R., Lex, A., & Gehlenborg, N. (2017). UpSetR: An R package for the visualization of intersecting sets and their properties. *Bioinformatics*, *33*(18), 2938–2940. https://doi.org/10.1093/bioinformatics/btx364

Dallas, H. F. (2005). *Inventory of national river health programme monitoring sites: Volume 2*. Cape Town, South Africa: Freshwater Consulting Group/Freshwater Research Unit, University of Cape Town.

Dallas, H. F. (2007). *Data interpretation guidelines*. Pretoria, South Africa: Institute of Natural Resources and Department of Water Affairs and Forestry. Retrieved from https://www.dwa.gov.za/iwqs/rhp/methods/SASSInterpretationGuidelines.pdf

Dallas, H. F., & Day, J. A. (2007). Natural variation in macroinvertebrate assemblages and the development of a biological banding system for interpreting bioassessment data – A preliminary evaluation using data from upland sites in the south-western Cape, South Africa. *Hydrobiologia*, *575*, 231–244. https://doi.org/10.1007/s10750-006-0374-y

Dallas, H. F., & Rivers-Moore, N. (2014). Ecological consequences of global climate change for freshwater ecosystems in South Africa. *South African Journal of Science*, *110*(5–6), 1–11. https://doi.org/10.1590/sajs.2014/20130274

Daniels, S. R., Busschau, T., & Cumberlidge, N. (2019). Two new species of freshwater crabs of the genus Potamonautes MacLeay, 1838 (Decapoda: Brachyura: Potamonautidae) from the forests of KwaZulu-Natal, South Africa. *Journal of Crustacean Biology*, *39*(4), 426–435. https://doi.org/10.1093/jcbiol/ruz024

Darwall, W., Smith, K., Tweddle, D., & Skelton, P. (2009). *The status and distribution of freshwater biodiversity in southern Africa. The IUCN Red List of Threatened Species – Regional Assessment*. Gland, Switzerland: IUCN.

de Meillon, B., & Wirth, W. (2002). Ceratopogonidae. In J. Day, A. Harrison, & I. de Moor (Eds.), *Guides to the freshwater invertebrates of Southern Africa. Vol. 9: Diptera* (pp. 50–56). Water Research Commission Report TT201/02. Pretoria, South Africa: Water Research Commission.

de Moor, F. C. (2002a). Shortcomings and advantages of using rapid biological assessment techniques for determining the health of rivers in South Africa. *Verhandlungen Des Internationalen Verein Limnologie*, *28*, 651–662.

de Moor, F. C. (2002b). Chapter 5: Simuliidae. In J. A. Day, A. D. Harrison, & I. de Moor (Eds.), *Guides to the freshwater invertebrates of Southern Africa. Vol. 9: Diptera* (pp. 1–200). African Water Research Commission Report TT201/02. Pretoria, South Africa: Water Research Commission.

de Moor, F., & Scott, K. (2003). Chapter 5: Trichoptera. In I. de Moor, J. Day, & F. de Moor (Eds.), *Guides to the freshwater invertebrates of Southern Africa. Volume 8: Insecta II* (pp. 84–169). Pretoria, South Africa: Water Research Commission.

Deiner, K., & Altermatt, F. (2014). Transport distance of invertebrate environmental DNA in a natural river. *PLoS One*, *9*(2), e88786. https://doi.org/10.1371/journal.pone.0088786

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., … Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, *26*(21), 5872–5895. https://doi.org/10.1111/mec.14350

Deiner, K., Fronhofer, E. A., Mächler, E., Walser, J. C., & Altermatt, F. (2016). Environmental DNA reveals that rivers are conveyer belts of biodiversity information. *Nature Communications*, *7*, 12544. https://doi.org/10.1038/ncomms12544

Deiner, K., Walser, J.-C., Mächler, E., & Altermatt, F. (2016). Choice of capture and extraction methods affect detection of freshwater biodiversity from environmental DNA. *Biological Conservation*, *183*, 53–63. https://doi.org/10.1016/j.biocon.2014.11.018

Delić, T., Trontelj, P., Rendoš, M., & Fišer, C. (2017). The importance of naming cryptic species and the conservation of endemic subterranean amphipods. *Scientific Reports*, *7*, 3391. https://doi.org/10.1038/s41598-017-02938-z

Dickens, C. W., & Graham, P. M. (2002). The South African Scoring System (SASS) version 5 rapid bioassessment method for rivers. *African Journal of Aquatic Science*, *27*(1), 1–10. https://doi.org/10.2989/16085914.2002.9626569

Doi, H., Fukaya, K., Oka, S.-I., Sato, K., Kondoh, M., & Miya, M. (2019). Evaluation of detection probabilities at the water-filtering and initial PCR steps in environmental DNA metabarcoding using a multispecies site occupancy model. *Scientific Reports*, *9*(1), 1–8. https://doi.org/10.1038/s41598-019-40233-1

Dowle, E. J., Pochon, X., Banks, C. J., Shearer, K., & Wood, S. A. (2016). Targeted gene enrichment and high-throughput sequencing for environmental biomonitoring: A case study using freshwater macroinvertebrates. *Molecular Ecology Resources*, *16*(5), 1240–1254. https://doi.org/10.1111/1755-0998.12488

Driver, A., Sink, K. J., Nel, J. N., Holness, S., Van Niekerk, L., Daniels, F., … Maze, K. (2012). *National Biodiversity Assessment 2011: An assessment of South Africa's biodiversity and ecosystems. Synthesis Report*. Pretoria, South Africa: South African National Biodiversity Institute and Department of Environmental Affairs.

Edgar, R. C. (2013). UPARSE: Highly accurate OTU sequences from microbial amplicon reads. *Nature Methods*, *10*(10), 996–998. https://doi.org/10.1038/nmeth.2604

Edgar, R. C., & Flyvbjerg, H. (2015). Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics*, *31*(21), 3476–3482. https://doi.org/10.1093/bioinformatics/btv401

Ekrem, T., Stur, E., & Hebert, P. D. N. (2010). Females do count: Documenting Chironomidae (Diptera) species diversity using DNA barcoding. *Organisms Diversity and Evolution*, *10*(5), 397–408. https://doi.org/10.1007/s13127-010-0034-y

Elbrecht, V., & Leese, F. (2017). Validation and development of COI metabarcoding primers for freshwater macroinvertebrate bioassessment. *Frontiers in Environmental Science*, *5*, 1–11. https://doi.org/10.3389/fenvs.2017.00011

Elbrecht, V., Peinert, B., & Leese, F. (2016). Sorting things out – Assessing effects of unequal specimen biomass on DNA metabarcoding. *PeerJ*, *5*, e2561v3. https://doi.org/10.7287/PEERJ.PREPRINTS.2561V1

Elbrecht, V., & Steinke, D. (2018). Scaling up DNA metabarcoding for freshwater macrozoobenthos monitoring. *Freshwater Biology*, *64*, 380–387. https://doi.org/10.1111/fwb.13220

Elbrecht, V., Vamos, E. E., Meissner, K., Aroviita, J., & Leese, F. (2017). Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream

monitoring. *Methods in Ecology and Evolution*, *8*, 1265–1275. https://doi.org/10.1111/2041-210X.12789

Elbrecht, V., Vamos, E. E., Steinke, D., & Leese, F. (2017). Assessing intraspecific genetic diversity from community DNA metabarcoding data. *PeerJ Preprints*, *5*, e3269v1. https://doi.org/10.7287/peerj.preprints.3269v1

Endrödy-Younga, S., & Stals, R. (2008). Chapter 14: Scirtidae. In R. Stals, & I. de Moor (Eds.), *Guides to the Freshwater Invertebrates of Southern Africa. Volume 10, Coleoptera* (pp. 133–138). Pretoria, South Africa: Water Research Commission.

Erdozain, M., Thompson, D. G., Porter, T. M., Kidd, K. A., Kreutzweiser, D. P., Sibley, P. K., … Hajibabaei, M. (2019). Metabarcoding of storage ethanol vs. conventional morphometric identification in relation to the use of stream macroinvertebrates as ecological indicators in forest management. *Ecological Indicators*, *101*(September), 173–184. https://doi.org/10.1016/j.ecolind.2019.01.014

Floyd, R., Abebe, E., Papert, A., & Blaxter, M. (2002). Molecular barcodes for soil nematode identification. *Molecular Ecology*, *11*(4), 839–850. https://doi.org/10.1046/j.1365-294X.2002.01485.x

Flynn, J. M., Brown, E. A., Chain, F. J. J., Macisaac, H. J., & Cristescu, M. E. (2015). Toward accurate molecular identification of species in complex environmental samples: Testing the performance of sequence filtering and clustering methods. *Ecology and Evolution*, *5*(11), 2252–2266. https://doi.org/10.1002/ece3.1497

Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, *3*(5), 294–299.

Gardham, S., Hose, G. C., Stephenson, S., & Chariton, A. A. (2014). DNA metabarcoding meets experimental ecotoxicology: Advancing knowledge on the ecological effects of copper in freshwater ecosystems. In *Advances in ecological research* (1st ed., Vol. *51*). Cambridge, MA: Academic Press.

Gibson, J. F., Shokralla, S., Curry, C., Baird, D. J., Monk, W. A., King, I., & Hajibabaei, M. (2015). Large-scale biomonitoring of remote and threatened ecosystems via high-throughput sequencing. *PLoS One*, *10*(10), 1–15. https://doi.org/10.1371/journal.pone.0138432

Haase, P., Murray-Bligh, J., Lohse, S., Pauls, S., Sundermann, A., Gunn, R., & Clarke, R. (2006). Assessing the impact of errors in sorting and identifying macroinvertebrate samples. *Hydrobiologia*, *566*(1), 505–521. https://doi.org/10.1007/s10750-006-0075-6

Haase, P., Pauls, S. U., Schindehütte, K., & Sundermann, A. (2010). First audit of macroinvertebrate samples from an EU Water Framework Directive monitoring program: Human error greatly lowers precision of assessment results. *Journal of the North American Benthological Society*, *29*(4), 1279–1291. https://doi.org/10.1899/09-183.1

Hajibabaei, M., Janzen, D. H., Burns, J. M., Hallwachs, W., & Hebert, P. D. N. (2006). DNA barcodes distinguish species of tropical Lepidoptera. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(4), 968–971. https://doi.org/10.1073/pnas.0510466103

Hajibabaei, M., Porter, T. M., Robinson, C., Baird, D. J., Shokralla, S., & Wright, M. (2019). Watered-down biodiversity? A comparison of metabarcoding results from DNA extracted from matched water and bulk tissue biomonitoring samples. *BioRxiv*, *45*(15), 37. https://doi.org/10.1101/575928

Hajibabaei, M., Shokralla, S., Zhou, X., Singer, G. A. C., & Baird, D. J. (2011). Environmental barcoding: A next-generation sequencing approach for biomonitoring applications using river benthos. *PLoS One*, *6*(4), e17497. https://doi.org/10.1371/journal.pone.0017497

Hamer, M. (2013). *A national strategy for zoological taxonomy (2013–2020)* (pp. 1–53). Pretoria, South Africa: South African National Biodiversity Institute.

Harrison, A. D. (2004). *The non-biting midges (Diptera: Chironomidae) of South Africa*. Rhodes University online checklist. Retrieved from https://www.ru.ac.za/media/rhodesuniversity/resources/martin/chironomidae.html

Harrison, A., Prins, A., & Day, J. A. (2002). Chapter 2: Lesser-known Nematocera. In J. Day, A. Harrison, & F. de Moor (Eds.), *Guides to the freshwater invertebrates of Southern Africa. Vol. 9: Diptera* (pp. 1–200). Water Research Commission Report TT201/02. Pretoria, South Africa: Water Research Commission.

Hebert, P. D. N., Ratnasingham, S., & de Waard, J. R. (2003). Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society B: Biological Sciences*, *270*(Suppl_1), S96–S99. https://doi.org/10.1098/rsbl.2003.0025

Hering, D., Borja, A., Jones, J. I., Pont, D., Boets, P., Bouchez, A., … Kelly, M. (2018). Implementation options for DNA-based identification into ecological status assessment under the European Water Framework Directive. *Water Research*, *138*, 192–205. https://doi.org/10.1016/j.watres.2018.03.003

Holovachov, O., Haenel, Q., Bourlat, S. J., & Jondelius, U. (2017). Taxonomy assignment approach determines the efficiency of identification of OTUs in marine nematodes. *Royal Society Open Science*, *4*(8), 170315. https://doi.org/10.1098/rsos.170315

Jansen van Rensburg, C., & Day, J. (2002). Chapter 2: Water mites. In I. de Moor, J. Day, & F. de Moor (Eds.), *Guides to the freshwater invertebrates of southern Africa. Volume 6: Arachnida and Mollusca* (pp. 23–41). Pretoria, South Africa: Water Research Commission.

Kaaya, L., Day, J., & Dallas, H. (2015). Tanzania River Scoring System (TARISS): A macroinvertebrate-based biotic index for rapid bioassessment of rivers. *African Journal of Aquatic Science*, *40*(2), 109–117. https://doi.org/10.2989/16085914.2015.1051941

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

Kleynhans, C., Thirion, C., & Moolman, J. (2005). *A Level 1 River Ecoregion classification system for South Africa*. Pretoria, South Africa: Lesotho and Swaziland.

Lejzerowicz, F., Esling, P., Pillet, L., Wilding, T. A., Black, K. D., & Pawlowski, J. (2015). High-throughput sequencing and morphology perform equally well for benthic monitoring of marine ecosystems. *Scientific Reports*, *5*, 1–10. https://doi.org/10.1038/srep13932

Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R., Pfister, H., & Manuscript, A. (2014). UpSet: Visualization of intersecting sets Europe PMC Funders Group. *IEEE Transactions on Visualization and Computer Graphics*, *20*(12), 1983–1992. https://doi.org/10.1109/TVCG.2014.2346248

Li, F., Peng, Y., Fang, W., Altermatt, F., Xie, Y., Yang, J., & Zhang, X. (2018). Application of environmental DNA metabarcoding for predicting anthropogenic pollution in rivers. *Environmental Science and Technology*, *52*(20), 11708–11719. https://doi.org/10.1021/acs.est.8b03869

Macher, J. N., Salis, R. K., Blakemore, K. S., Tollrian, R., Matthaei, C. D., & Leese, F. (2016). Multiple-stressor effects on stream invertebrates: DNA barcoding reveals contrasting responses of cryptic mayfly species. *Ecological Indicators*, *61*, 159–169. https://doi.org/10.1016/j.ecolind.2015.08.024

Macher, J. N., Vivancos, A., Piggott, J. J., Centeno, F. C., Matthaei, C. D., & Leese, F. (2018). Comparison of environmental DNA and bulk-sample metabarcoding using highly degenerate cytochrome c oxidase I primers. *Molecular Ecology Resources*, *18*(6), 1456–1468. https://doi.org/10.1111/1755-0998.12940

Mächler, E., Little, C. J., Wüthrich, R., Alther, R., & Fronhofer, E. A. (2019). Assessing different components of biodiversity across a river network using eDNA. *Environmental DNA*, *1*, 290–301.

Majaneva, M., Diserud, O. H., Eagle, S. H. C., Boström, E., Hajibabaei, M., & Ekrem, T. (2018). Environmental DNA filtration techniques affect recovered biodiversity. *Scientific Reports*, *8*, 4682. https://doi.org/10.1038/s41598-018-23052-8

Majaneva, M., Diserud, O. H., Eagle, S. H. C., Hajibabaei, M., & Ekrem, T. (2018). Choice of DNA extraction method affects DNA metabarcoding of unsorted invertebrate bulk samples. *Metabarcoding and Metagenomics*, *2*, e26664. https://doi.org/10.3897/mbmg.2.26664

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *Embnet Journal*, *17*, 10–12. https://doi.org/10.14806/ej.17.1.200

Martins, F. M. S., Galhardo, M., Filipe, A. F., Teixeira, A., Pinheiro, P., Paupério, J., … Beja, P. (2019). Have the cake and eat it: Optimizing nondestructive DNA metabarcoding of macroinvertebrate samples for freshwater biomonitoring. *Molecular Ecology Resources*, *19*(4), 863–876. https://doi.org/10.1111/1755-0998.13012

Nel, J., Driver, A., Strydom, W. F., Maherry, A., Petersen, C., Hill, L., … Smith-Adao, L. B. (2011). *Atlas of freshwater ecosystem priority areas in South Africa*. WRC Report number: K5/1801. Pretoria, South Africa: Water Research Commission.

Nelson, H. (2008). Chapter 15: Elmidae. In R. Stals, & I. de Moor (Eds.), *Guides to the freshwater invertebrates of southern Africa. Volume 10: Coleoptera* (pp. 139–152). Pretoria, South Africa: Water Research Commission.

Nichols S. J., Kefford B. J., Campbell C. D., Bylemans J., Chandler E., Bray J. P., … Furlan E. M. (2020). Towards routine DNA metabarcoding of macroinvertebrates using bulk samples for freshwater bioassessment: Effects of debris and storage conditions on the recovery of target taxa. *Freshwater Biology*, *65*, (4), 607–620. https://doi.org/10.1111/fwb.13443

Odume, O. N., Muller, W. J., Arimoro, F. O., & Palmer, C. G. (2012). The impact of water quality deterioration on macroinvertebrate communities in the Swartkops River, South Africa: A multimetric approach. *African Journal of Aquatic Science*, *37*(2), 191–200. https://doi.org/10.2989/16085914.2012.670613

Odume, O. N., Palmer, C. G., Arimoro, F. O., & Mensah, P. K. (2015). Influence of selected biotopes on chironomid-based bioassessment of the swartkops river, Eastern Cape, South Africa. *Water SA*, *41*(3), 343–358. https://doi.org/10.4314/wsa.v41i3.06

Omernik, J. M. (1987). Ecoregions of the conterminous United States. *Annals of the Association of American Geographers*, *77*(1), 118–125. https://doi.org/10.1111/j.1467-8306.1987.tb00149.x

Packer, L., Gibbs, J., Sheffield, C., & Hanner, R. (2009). DNA barcoding and the mediocrity of morphology. *Molecular Ecology Resources*, *9*, 42–50. https://doi.org/10.1111/j.1755-0998.2009.02631.x

Palmer, R. W., & Taylor, E. D. (2004). The Namibian Scoring System (NASS) version 2 rapid bio-assessment method for rivers. *African Journal of Aquatic Science*, *29*(2), 229–234. https://doi.org/10.2989/16085910409503814

Park, D. S., Foottit, R., Maw, E., & Hebert, P. D. N. (2011). Barcoding bugs: DNA-based identification of the true bugs (Insecta: Hemiptera: Heteroptera). *PLoS One*, *6*(4), e18749. https://doi.org/10.1371/journal.pone.0018749

Porter, T. M., & Hajibabaei, M. (2018). Over 2.5 million COI sequences in GenBank and growing. *PLoS One*, *13*(9), 1–16. https://doi.org/10.1371/journal.pone.0200177

R Core Team (2019). *R: A language and environment for statistical computing*. Retrieved from https://www.r-project.org/

Rambaut, A. (2009). *FigTree version 1.4. 3*. Retrieved from https://tree.bio.ed.ac.uk/software/figtree/

Ratnasingham, S., & Herbert, P. (2007). The Barcode of Life Data System BOLD: (www.barcodinglife.org). *Molecular Ecology Notes*, *7*, 355–364. https://doi.org/10.1111/j.1471-8286.2006.01678.x

Ransome E., Geller J. B., Timmers M., Leray M., Mahardini A., Sembiring A., … Meyer C. P. (2017). The importance of standardization for biodiversity comparisons: A case study using autonomous reef monitoring structures (ARMS) and metabarcoding to measure cryptic diversity on Mo'orea coral reefs, French Polynesia. *PLOS ONE*, *12*, (4), e0175066. https://doi.org/10.1371/journal.pone.0175066

Samways, M., & Wilmot, B. (2003). Chapter 3: Odonata. In I. de Moor, J. Day, & F. de Moor (Eds.), *Guides to the Freshwater Invertebrates of Southern Africa. Volume 7: Insecta I: Ephemeroptera, Odonata & Plecoptera* (pp. 160–200). Pretoria, South Africa: Water Research Commission.

Simaika, J. P., & Samways, M. J. (2009). An easy-to-use index of ecological integrity for prioritizing freshwater sites and for assessing habitat quality. *Biodiversity and Conservation*, *18*(5), 1171–1185. https://doi.org/10.1007/s10531-008-9484-3

Smith, M. A., Bertrand, C., Crosby, K., Eveleigh, E. S., Fernandez-triana, J., Fisher, B. L., … Packer, L. (2012). Wolbachia and DNA barcoding insects: Patterns, potential, and problems. *PLoS One*, *7*(5), e36514. https://doi.org/10.1371/journal.pone.0036514

Stals, R. (2008). Chapter 3: Gyrinidae. In R. Stals, & I. J. de Moor (Eds.), *Guides to the freshwater invertebrates of southern Africa. Volume 10: Coleoptera* (pp. 53–58). Pretoria, South Africa: Water Research Commission.

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stein, E. D., White, B. P., Mazor, R. D., Jackson, J. K., Battle, J. M., Miller, P. E., … Sweeney, B. W. (2014). Does DNA barcoding improve performance of traditional stream bioassessment metrics? *Freshwater Science*, *33*(1), 302–311. https://doi.org/10.1086/674782

Stevens, D., & Picker, M. D. (2003). Chapter 4: Plecoptera. In I. de Moor, J. Day, & F. de Moor (Eds.), *Guides to the Freshwater Invertebrates of Southern Africa. Volume 7: Insecta I: Ephemeroptera, Odonata & Plecoptera* (pp. 213–263). Pretoria, South Africa: Water Research Commission.

Stork, N. E. (2018). How many species of insects and other terrestrial arthropods are there on earth? *Annual Review of Entomology*, *63*(1), 31–45. https://doi.org/10.1146/annurev-ento-020117-043348

Sweeney, B. W., Battle, J. M., Jackson, J. K., & Dapkey, T. (2011). Can DNA barcodes of stream macroinvertebrates improve descriptions of community structure and water quality? *Journal of the North American Benthological Society*, *30*(1), 195–216. https://doi.org/10.1899/10-016.1

Theissinger, K., Kästel, A., Elbrecht, V., Makkonen, J., Michiels, S., Schmidt, S., … Brühl, C. (2018). Using DNA metabarcoding for assessing chironomid diversity and community change in mosquito controlled temporary wetlands. *Metabarcoding and Metagenomics*, *2*(2000), e21060. https://doi.org/10.3897/mbmg.2.21060

Van Hoven, W., & Day, J. (2002). Chapter 13: Oligochaeta. In J. Day, & I. de Moor (Eds.), *Guides to the freshwater invertebrates of Southern Africa, Volume 5: Nonarthropods (Protozoans, Porifera, Cnidaria, Plathyhelminthes, Nemertea, Rotifera, Nematoda, Nematomorpha, Gastrotrichia, Bryozoa, Tardigrada, Polychaeta, Oligochaeta and Hirudinea)* (pp. 203–236). Water Research Commission Report TT167/02. Pretoria, South Africa: Water Research Commission.

Venter, H. J., & Bezuidenhout, C. C. (2016). DNA-based identification of aquatic invertebrates – Useful in the South African context? *South African Journal of Science*, *112*(5–6), 4–7. https://doi.org/10.17159/sajs.2016/20150444

Weigand, H., Beermann, A. J., Čiampor, F., Costa, F. O., Csabai, Z., Duarte, S., … Ekrem, T. (2019). DNA barcode reference libraries for the monitoring of aquatic biota in Europe: Gap-analysis and recommendations for future work. *Science of the Total Environment*, *678*, 499–524. https://doi.org/10.1016/j.scitotenv.2019.04.247

Zimmermann, J., Glöckner, G., Jahn, R., Enke, N., & Gemeinholzer, B. (2015). Metabarcoding vs. morphological identification to assess diatom

diversity in environmental studies. *Molecular Ecology Resources*, *15*(3), 526–542. https://doi.org/10.1111/1755-0998.12336

Zizka, V. M. A., Leese, F., Peinert, B., & Geiger, M. F. (2019). DNA metabarcoding from sample fixative as a quick and voucher-preserving biodiversity assessment method. *Genome*, *62*(3), 122–136. https://doi.org/10.1139/gen-2018-0048

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Pereira-da-Conceicoa L, Elbrecht V, Hall A, Briscoe A, Barber-James H, Price B. Metabarcoding unsorted kick-samples facilitate macroinvertebrate-based biomonitoring with increased taxonomic resolution, while outperforming environmental DNA. *Environmental DNA*. 2020;00:1–19. https://doi.org/10.1002/edn3.116