Marine and Freshwater Research
https://doi.org/10.1071/MF20269_AC

Supplementary material

Comparison of an extracellular *v.* total DNA extraction approach for environmental DNA-based monitoring of sediment biota

Johan Pansu[A,B,C,D], *Michelle B. Chapman*[C], *Grant C. Hose*[C] *and Anthony A. Chariton*[C]

[A]Station Biologique de Roscoff, UMR 7144, CNRS, Sorbonne Université,
Place Georges Teissier, F-29680 Roscoff, France.

[B]CSIRO Ocean & Atmosphere, New Illawarra Road, Lucas Heights, NSW 2234, Australia.

[C]Department of Biological Sciences, Macquarie University, Balaclava Road,
Macquarie Park, NSW 2109, Australia.

[D]Corresponding author. Present address: Institut des Sciences de l'Évolution de Montpellier (ISEM), CNRS, IPHE, IRD, Université de Montpellier, Place Eugène Bataillon, F-34000 Montpellier, France. Email: johan.pansu@gmail.com

Table S1. Statistics regarding metabarcoding data processing

	Bacteria (16S)			Eukaryote (18S)			Metazoan (COI)		
	Number of reads	Number of OTUs	Mean number of reads per sample	Number of reads	Number of OTUs	Mean number of reads per sample	Number of reads	Number of OTUs	Mean number of reads per sample
After demultiplexing	3366075	-	-	7813456	-	-	3051786	-	-
Post-GHAP pipeline	2620834	27836	36913 +/- 1524	6582313	12663	88950 +/- 1794	2611810	12818	40182 +/- 2201
Post-Filtering (before removing rare OTUs)	2555517	27093	35993 +/- 1499	6443202	10806	87070 +/- 1753	1674650	5367	26166 +/- 1506
Final dataset	1708175	1362	24059 +/- 1026	5672529	998	76655 +/- 1533	1585648	1414	25575 +/- 1358

Table S2. Pairwise comparisons in pre-PCR (polymerase chain reaction) concentrations among the different protocol variants

Z-values from Dunn's multiple-comparison tests are reported here. Significance: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$. Over the entire dataset (column 'All'), only pairwise comparisons involving the totDNA extraction protocol with 1 g of sediment are significant

Pairwise comparison	Sediment type			
	All	Pond	Estuary	Marine
extDNA 1 g - extDNA 10 g	-0.80	-0.34	-0.52	-1.46
extDNA 1 g - extDNA 200 g	-0.40	-0.30	1.42	-3.05*
extDNA 10 g - extDNA 200 g	0.39	0.04	1.93	-1.59
extDNA 1 g - totDNA 1 g	2.50*	2.23·	2.49*	0.73
extDNA 10 g - totDNA 1 g	3.29**	2.58*	3.00**	2.19
extDNA 200 g - totDNA 1 g	2.90*	2.54*	1.07	3.78**
extDNA 1 g - totDNA 10 g	-1.30	-0.73	-1.68	-1.37
extDNA 10 g - totDNA 10 g	-0.50	-0.39	-1.16	0.09
extDNA 200 g - totDNA 10 g	-0.90	-0.43	-3.09**	1.68
totDNA 1 g - totDNA 10 g	-3.79**	-2.96*	-4.17***	-2.11

Table S3. Pairwise comparisons in operational taxonomic unit (OTU) richness among the different protocol variants for each primer pair

Z-value from the Dunn's multiple-comparison tests are reported here. None of the *Z* values was significant ($P > 0.05$), although marked differences were observed for eukaryotes

Pairwise comparison	Bacteria (16S)		Eukaryote (18S)		Metazoan (COI)	
	***Z* value**	**Adjusted *P* value**	***Z* value**	**Adjusted *P* value**	***Z* value**	**Adjusted *P* value**
extDNA 1 g - extDNA 10 g	-0.446	1.000	-1.155	0.414	-0.544	0.977
extDNA 1 g - extDNA 200 g	0.773	1.000	-1.936	0.132	-0.316	0.940
extDNA 10 g - extDNA 200 g	1.265	1.000	-0.781	0.543	0.228	0.819
extDNA 1 g - totDNA 1 g	-0.397	0.988	0.822	0.587	-0.802	0.845
extDNA 10 g - totDNA 1 g	0.043	0.965	1.957	0.168	-0.243	0.898
extDNA 200 g - totDNA 1 g	-1.199	1.000	2.724	0.064	-0.477	0.904
extDNA 1 g - totDNA 10 g	-0.046	1.000	-1.774	0.152	-2.265	0.235
extDNA 10 g - totDNA 10 g	0.407	1.000	-0.620	0.595	-1.721	0.284
extDNA 200 g - totDNA 10 g	-0.836	1.000	0.161	0.872	-1.949	0.256
totDNA 1 g - totDNA 10 g	0.357	0.901	-2.566	0.051	-1.522	0.320

Table S4. Results of perMANOVA analyses conducted per sediment type to test for the effect of the extraction type (extDNA and totDNA, regardless of the initial mass of sediment) on Bray–Curtis distances between samples

	Bacteria (16S)			Eukaryote (18S)			Metazoan (COI)		
	F	r^2	*P*	F	r^2	*P*	F	r^2	*P*
Estuary	33.86	0.61	0.001	20.30	0.47	0.001	10.73	0.33	0.001
Marine	7.011	0.25	0.001	4.55	0.17	0.002	-	-	-
Pond	12.29	0.36	0.001	8.30	0.27	0.001	3.67	0.16	0.001

Table S5. Mean (± s.e.m.) relative read abundance (RRA) of phyla obtained from each protocol variant and averaged fold change (AFC) ratio between totDNA and extDNA extractions

Only pairs of protocols with similar amount of starting material were compared using AFC ratios. For each clade, AFC values reflects the positive bias (i.e. enrichment) by either of the extraction method (extDNA or totDNA). Significant differences between the two approaches were assessed using Wilcoxon test. Significance: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$. Only the 10 most abundant phyla for each marker are represented. Significant differences were observed for eukaryotes and metazoans, but not bacteria

Bacteria (16S)

extDNA 1 g vs totDNA 1 g

Phylum	Mean (± SEM) RRA extDNA 1 g	Mean (± SEM) RRA totDNA 1 g	AFC	Enriching method	Significance
Acidobacteria	6.16% ± 1.25	8.02% ± 0.87	1.30	totDNA 1 g	
Actinobacteria	2.52% ± 0.72	2.58% ± 0.30	1.02	totDNA 1 g	
Bacteroidetes	8.88% ± 1.42	10.04% ± 1.89	1.13	totDNA 1 g	
Chloroflexi	2.72% ± 0.52	3.51% ± 0.82	1.29	totDNA 1 g	
Euryarchaeota	1.16% ± 0.46	1.09% ± 0.39	1.07	extDNA 1 g	
Firmicutes	0.47% ± 0.06	0.65% ± 0.08	1.39	totDNA 1 g	
Planctomycetes	1.09% ± 0.39	0.93% ± 0.32	1.18	extDNA 1 g	
Proteobacteria	63.52% ± 1.83	59.51% ± 1.44	1.07	extDNA 1 g	
Thaumarchaeota	2.73% ± 0.93	3.14% ± 1.20	1.15	totDNA 1 g	
Verrucomicrobia	2.74% ± 0.58	2.48% ± 0.53	1.10	extDNA 1 g	

extDNA 10 g vs totDNA 10 g

Phylum	Mean (± SEM) RRA extDNA 10 g	Mean (± SEM) RRA totDNA 10 g	AFC	Enriching method	Significance
Acidobacteria	5.93% ± 1.08	7.58% ± 1.21	1.28	totDNA 10 g	
Actinobacteria	2.55% ± 0.63	2.25% ± 0.36	1.13	extDNA 10 g	
Bacteroidetes	8.70% ± 1.35	11.57% ± 1.91	1.33	totDNA 10 g	
Chloroflexi	2.54% ± 0.46	2.99% ± 0.47	1.18	totDNA 10 g	
Euryarchaeota	1.32% ± 0.45	0.95% ± 0.35	1.39	extDNA 10 g	
Firmicutes	0.49% ± 0.07	0.77% ± 0.13	1.56	totDNA 10 g	
Planctomycetes	1.18% ± 0.39	0.66% ± 0.22	1.80	extDNA 10 g	
Proteobacteria	62.95% ± 1.60	61.74% ± 1.05	1.02	extDNA 10 g	
Thaumarchaeota	3.52% ± 1.15	1.77% ± 0.78	1.98	extDNA 10 g	
Verrucomicrobia	2.27% ± 0.39	2.12% ± 0.36	1.07	extDNA 10 g	

extDNA 200 g

Phylum	Mean (± SEM) RRA extDNA 200 g
Acidobacteria	6.31% ± 1.22
Actinobacteria	1.02% ± 0.28
Bacteroidetes	8.80% ± 1.38
Chloroflexi	1.16% ± 0.21
Euryarchaeota	1.56% ± 0.58
Firmicutes	0.52% ± 0.05
Planctomycetes	1.02% ± 0.33
Proteobacteria	64.95% ± 1.93
Thaumarchaeota	3.35% ± 0.98
Verrucomicrobia	2.32% ± 0.36

Table S5. (*Cont.*)

Eukaryote (18S)

extDNA 1 g vs totDNA 1 g

Phylum	Mean (± SEM) RRA extDNA 1 g	Mean (± SEM) RRA totDNA 1 g	AFC	Enriching method	Significance
Annelida	16.07% ± 4.84	6.78% ± 2.15	2.37	extDNA 1 g	
Arthropoda	10.93% ± 3.92	21.47% ± 4.38	1.96	totDNA 1 g	
Bacillariophyta	10.56% ± 2.21	5.64% ± 1.10	1.87	extDNA 1 g	
Chlorophyta	0.77% ± 0.11	7.94% ± 2.86	10.25	totDNA 1 g	
Ciliophora	4.70% ± 1.00	2.70% ± 0.35	1.74	extDNA 1 g	
Gastrotricha	5.77% ± 1.16	4.30% ± 2.45	1.34	extDNA 1 g	**
Myzozoa	5.15% ± 1.80	4.81% ± 1.06	1.07	extDNA 1 g	
Nematoda	3.62% ± 0.97	8.20% ± 1.19	2.27	totDNA 1 g	**
Platyhelminthes	21.72% ± 5.20	15.02% ± 3.58	1.45	extDNA 1 g	
Xenacoelomorpha	4.58% ± 1.68	6.38% ± 2.99	1.40	totDNA 1 g	

extDNA 10 g vs totDNA 10 g

Phylum	Mean (± SEM) RRA extDNA 10 g	Mean (± SEM) RRA totDNA 10 g	AFC	Enriching method	Significance
Annelida	15.62% ± 4.22	12.28% ± 4.48	1.27	extDNA 10 g	
Arthropoda	11.64% ± 3.30	17.19% ± 3.40	1.48	totDNA 10 g	
Bacillariophyta	9.31% ± 1.82	4.82% ± 0.95	1.93	extDNA 10 g	
Chlorophyta	0.79% ± 0.17	5.81% ± 2.14	7.31	totDNA 10 g	
Ciliophora	4.96% ± 1.29	2.93% ± 0.56	1.69	extDNA 10 g	
Gastrotricha	6.32% ± 1.45	4.76% ± 1.88	1.33	extDNA 10 g	
Myzozoa	4.90% ± 1.39	4.18% ± 1.07	1.17	extDNA 10 g	
Nematoda	1.78% ± 0.37	5.56% ± 0.71	3.12	totDNA 10 g	***
Platyhelminthes	23.57% ± 5.65	17.75% ± 4.31	1.33	extDNA 10 g	
Xenacoelomorpha	4.84% ± 1.45	7.98% ± 3.01	1.65	totDNA 10 g	

extDNA 200 g

Phylum	Mean (± SEM) RRA extDNA 200 g
Annelida	18.20% ± 5.74
Arthropoda	9.30% ± 2.92
Bacillariophyta	11.68% ± 2.65
Chlorophyta	0.74% ± 0.15
Ciliophora	4.73% ± 1.27
Gastrotricha	5.43% ± 1.40
Myzozoa	4.69% ± 1.15
Nematoda	0.69% ± 0.11
Platyhelminthes	23.80% ± 6.04
Xenacoelomorpha	5.47% ± 1.68

Table S5. (*Cont.*)

Metazoan (COI)

extDNA 1 g vs totDNA 1 g

Phylum	Mean (± SEM) RRA extDNA 1 g	Mean (± SEM) RRA totDNA 1 g	AFC	Enriching method	Significance
Annelida	35.21% ± 7.31	22.62% ± 5.25	1.56	extDNA 1 g	
Arthropoda	15.34% ± 5.73	26.64% ± 3.25	1.74	totDNA 1 g	*
Cnidaria	2.21% ± 1.03	2.81% ± 0.89	1.27	totDNA 1 g	
Gastrotricha	0.39% ± 0.07	1.65% ± 0.56	4.24	totDNA 1 g	
Mollusca	31.41% ± 10.50	22.41% ± 6.63	1.40	extDNA 1 g	
Nematoda	0.25% ± 0.13	3.16% ± 1.44	12.73	totDNA 1 g	***
Platyhelminthes	0.87% ± 0.27	0.93% ± 0.22	1.06	totDNA 1 g	
Porifera	0.64% ± 0.14	0.76% ± 0.19	1.18	totDNA 1 g	
Rotifera	0.10% ± 0.06	1.21% ± 0.50	11.64	totDNA 1 g	
Xenacoelomorpha	3.00% ± 1.1	1.83% ± 0.66	1.64	extDNA 1 g	

extDNA 10 g vs totDNA 10 g

Phylum	Mean (± SEM) RRA extDNA 10 g	Mean (± SEM) RRA totDNA 10 g	AFC	Enriching method	Significance
Annelida	45.39% ± 7.70	29.44% ± 6.79	1.54	extDNA 10 g	
Arthropoda	11.50% ± 3.37	22.41% ± 2.84	1.95	totDNA 10 g	*
Cnidaria	2.29% ± 0.69	1.49% ± 0.41	1.53	extDNA 10 g	
Gastrotricha	0.57% ± 0.16	1.69% ± 0.63	2.98	totDNA 10 g	
Mollusca	23.86% ± 9.60	28.48% ± 8.90	1.19	totDNA 10 g	
Nematoda	0.28% ± 0.12	0.76% ± 0.25	2.75	totDNA 10 g	*
Platyhelminthes	0.70% ± 0.26	0.79% ± 0.19	1.13	totDNA 10 g	
Porifera	0.72% ± 0.21	0.54% ± 0.21	1.32	extDNA 10 g	
Rotifera	0.14% ± 0.07	0.69% ± 0.28	5.00	totDNA 10 g	
Xenacoelomorpha	2.77% ± 1.23	1.90% ± 0.66	1.46	extDNA 10 g	

extDNA 200 g

Phylum	Mean (± SEM) RRA extDNA 200 g
Annelida	41.25% ± 8.56
Arthropoda	9.44% ± 3.12
Cnidaria	2.29% ± 0.91
Gastrotricha	0.44% ± 0.16
Mollusca	31.90% ± 10.01
Nematoda	0.16% ± 0.10
Platyhelminthes	0.93% ± 0.33
Porifera	0.73% ± 0.21
Rotifera	0.09% ± 0.06
Xenacoelomorpha	3.63% ± 1.32

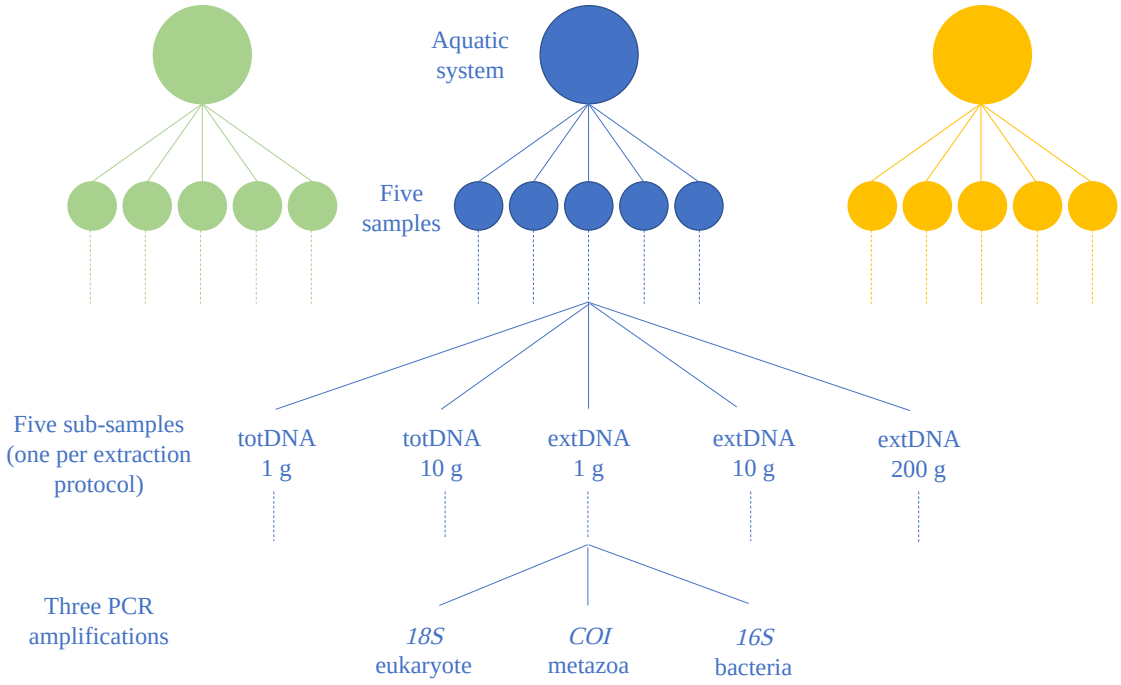

Fig. S1. Study design. Three aquatic systems (pond, estuarine and marine) were studied; five samples (~500 g) were collected in each (sample *n* = 15). From each sample, five subsamples of different volume were used for assessing the efficiency of two DNA extraction approaches: 1 g, 10 g and 200 g for the extracellular DNA method, 1 g and 10 g for the total DNA extraction method (extraction *n* = 75). Each DNA extract was then amplified with three different primer pairs to characterise bacterial, eukaryote and metazoan communities (PCR *n* = 225).

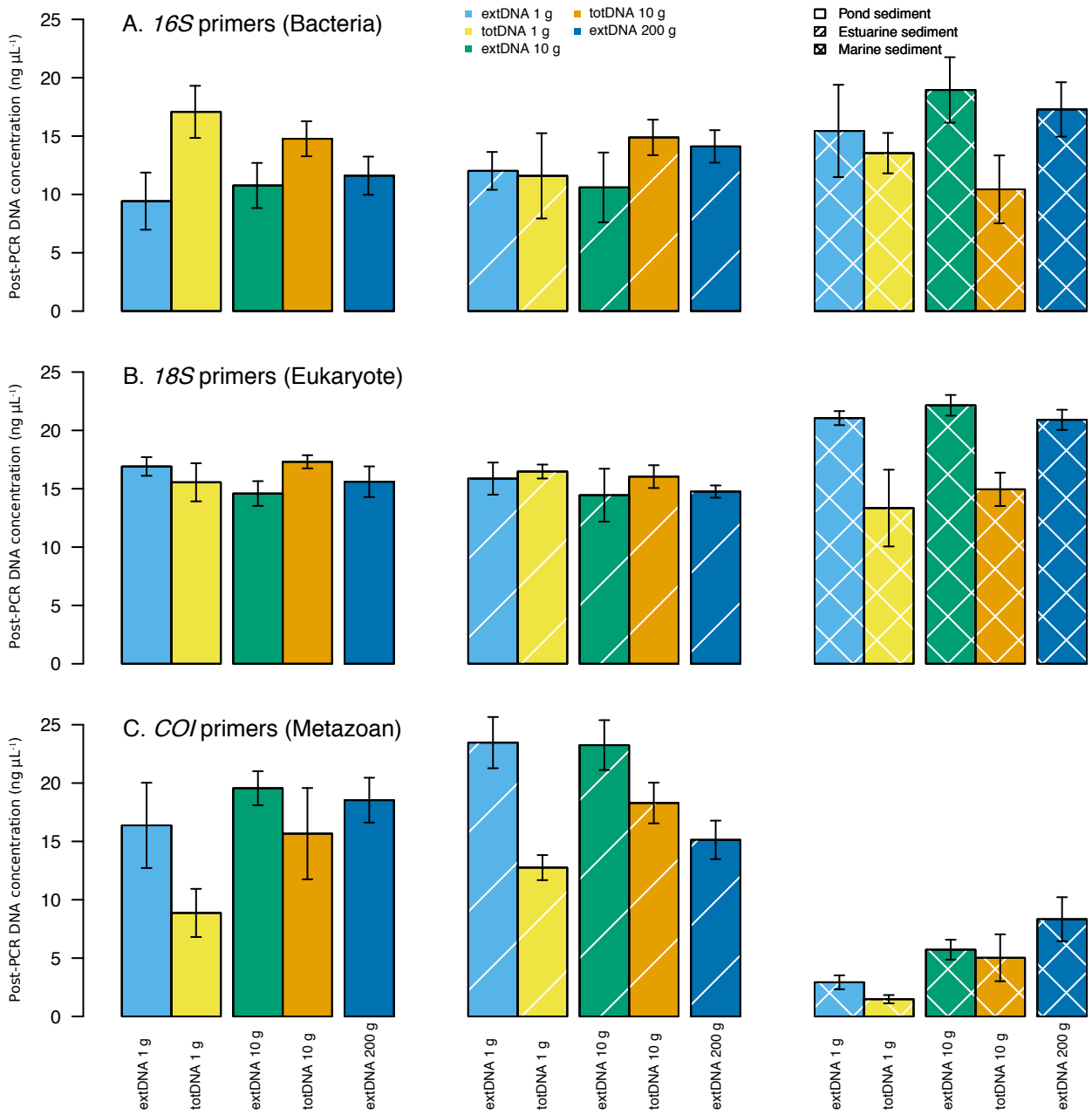

Fig. S2. Mean (± s.e.m.) post-PCR (polymerase chain reaction) concentrations per extraction protocol variant for each sediment type. Each row corresponds to one of the primer pairs used in this study. (*a*) *16S* bacteria, (*b*) *18S* eukaryote, (*c*) *COI* metazoan.

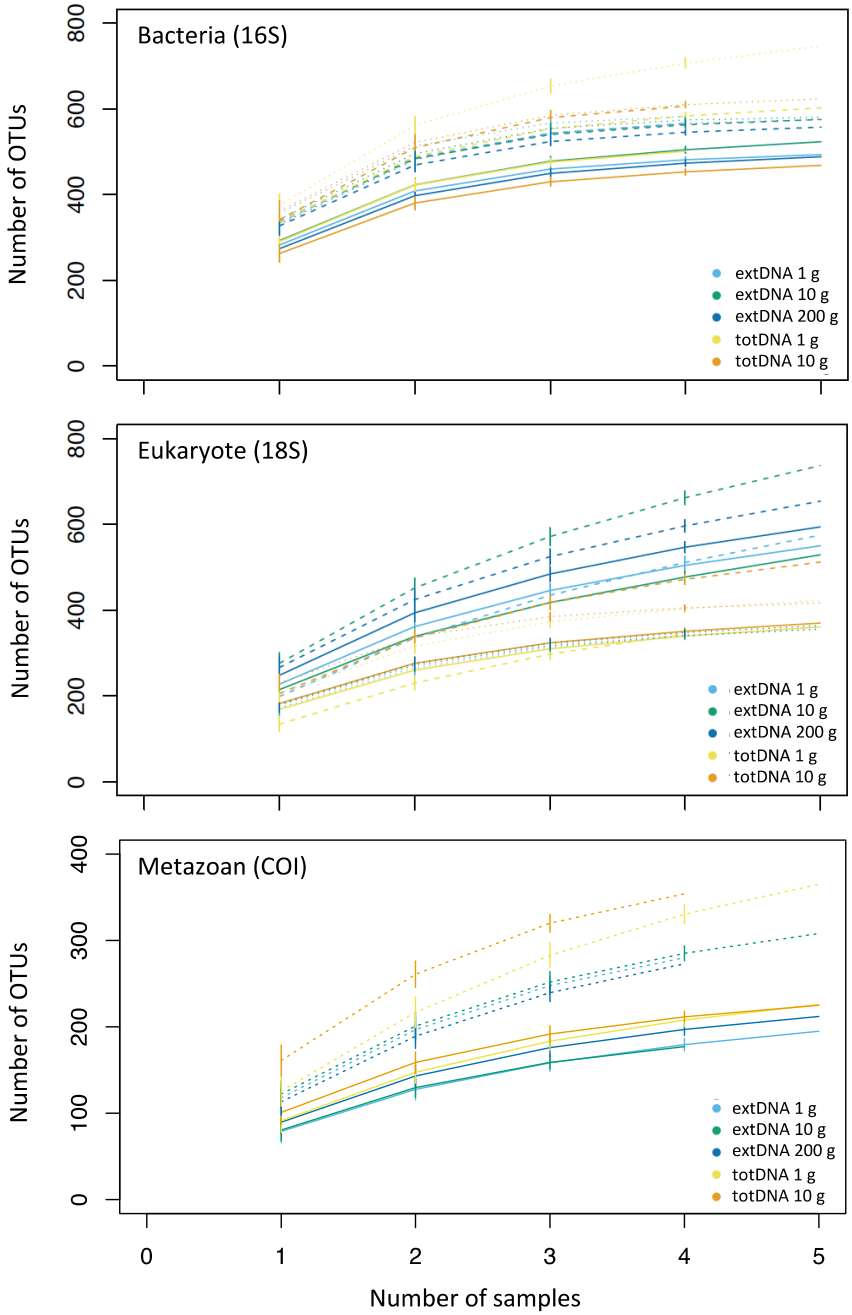

Fig. S3. Operational taxonomic unit (OTU) accumulation curve per sediment type (pond, estuarine or marine) according to the extraction protocol variant for bacteria (top panel), eukaryotes (central panel) and metazoans (bottom panel). Colours correspond to the extraction protocol variant. Estuary samples (plain line); marine samples (dashed line); pond samples (dotted line). Coleman method for species accumulation curve was used to estimate the expected richness. Errors bars correspond to the standard deviation from 100 permutations.

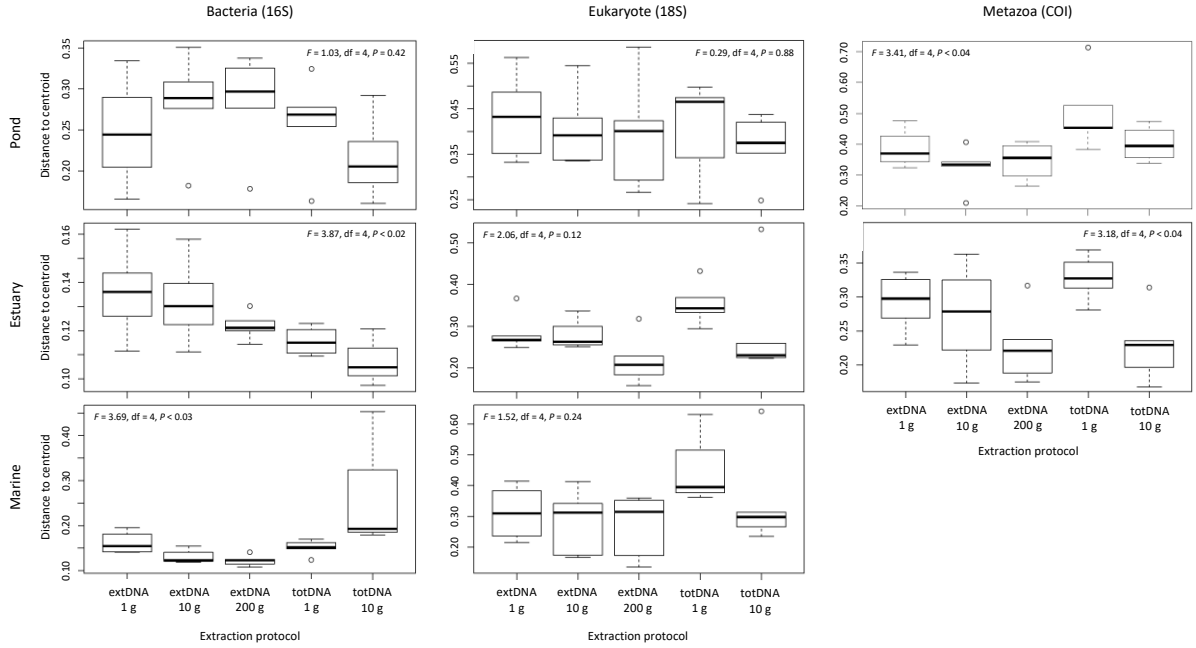

Fig. S4. Homogeneity of samples according to the extraction protocol variant. For each sediment type and taxonomic group, samples are grouped per extraction protocol variant, and Bray–Curtis distances of each sample to the centroid of its group are reported ('*betadisper*' analysis; J. Oksanen, F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, E. Szoecs, and H. Wagner, https://CRAN.R-project.org/package=vegan). Low distances to the centroid indicate high homogeneity among sample replicates. ANOVA was performed for each taxon in each sediment type to test for significant differences between protocols.

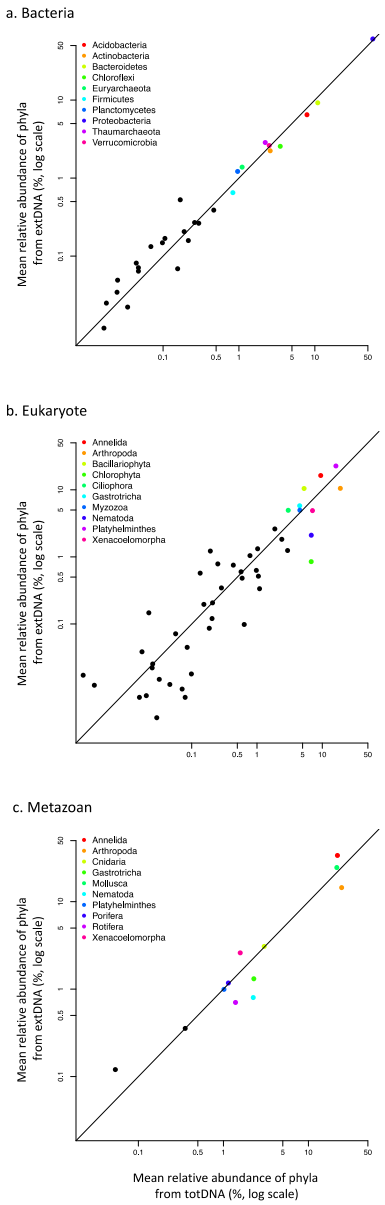

Fig. S5. Comparison of phylum average relative abundances between the totDNA and extDNA extraction approaches for bacteria, eukaryotes and metazoans. The solid line indicates a 1 : 1 relationship. Data presented here are log-transformed, one read have been added to each operational taxonomic unit (OTU) count prior analysis to avoid null values.

Code for the inertia decomposition analysis

The inertia of a group of points represents their dispersion and corresponds to the sum of the squared distances from each point to centroid of the group (i.e. cloud of points). The additive property of the inertia makes that the inertia of a group of points is equal to the sum of all the inertia of subgroups (Prud'Homme 2012). Here, for each sedimentary habitat, five samples were collected, and each of them was extracted using five different protocols. The variation between DNA extracts from the same habitat occurs at two different levels, namely, variation owing to sampling and variation owing to extraction. We used the additive property of the inertia to decompose the total inertia (Itot) between the portion owing to the variation between extractions from a sample (Iext) and the portion owing to the biological variation between samples (Isamp): Itot = Isamp + Iext

For this analysis, we first conducted correspondence analyses for each sediment type, by retaining the highest possible number of dimensions D. We then calculated the total inertia (i.e. the dispersion of the entire cloud of points) by summing the squared distances of each point to the centroid of these points in the D dimensions. Points were then grouped according to their sample of origin, and the centroid of each of these subgroups was calculated. The inertia owing to the variation between extractions corresponds then to the sum of the squared distances of points to their respective subgroups. The fraction of the inertia owing to the variation between extractions from a same sample FIext (i.e. 'extraction effect') was then defined as FIext = Iext / Itot. Finally, because we have only two levels of variation here, the fraction of the inertia owing to the variation between samples is FIsamp is: FIsamp = 1 – FIext.

We thank Eric Coissac and Sophie Prud'Homme for sharing their code, which was originally developed as part of Sophie Prud'Homme's Master of Research thesis (Prud'Homme 2012).

```
###Function 'barycentre' for calculating the centroid of a group of points
barycentre = function(line,grp) {
        aggmean = function(x) aggregate(x,grp,mean)[,2]
grp=as.character(grp)
weight = table(grp)
grp = list(class=grp)
centre = rbind(apply(line,2,aggmean))
names=aggregate(as.vector(line[,1]),grp,mean)[,1]
if (length(names)==1)
        rownames(centre)[1]=names
else
        rownames(centre)=names
attr(centre,"weight")=weight
centre
```

```
}


### Function ‘inertia’ to calculate the inertia of a group of points
inertia = function(line,grp) {
        centre = barycentre(line,grp)
        (line - centre[grp,])**2
}


library(ade4)
###Correspondance analysis
#Tab: mOTUs x sample matrix
coa = dudi.coa(Tab, scan = FALSE, nf = nrow(Tab)-1)


###Total Inertia
#nbSamp: total number of points included in the CA
Itot = sum(inertia(coa$li, rep(1, nbSamp)))


###Inertia extraction
#SubG: vector describing the sub-group (i.e. the original sample) of each point (e.g. S1, S1, S2, S2)
Iext: sum(inertie(pca1$li, SubG))


###Fraction of the inertia due to the variation between extractions
FIext = Iext / Itot


###Fraction of the inertia due to the variation between sample
FIsamp = 1- FIext
```

Reference

Prud’Homme, S. (2012) Environmental DNA metabarcoding for plant biodiversity assessment. M.Res. Thesis, Ecole Normale Supérieure Lyon & University of Lyon, France.