



HAL
open science

Monitoring and methods for estimating the biodiversity of a watercourse in an alpine ecosystem by metabarcoding analyses

Amélie Bernard, Augustin Maniouloux, Clémentine Capitaine, Florian Petrilli, Ibtissam Dich, Jehanne Cordin, Jolan Soyer, Manon Petit, Perrine Duquet, Valentin Garreau, et al.

► **To cite this version:**

Amélie Bernard, Augustin Maniouloux, Clémentine Capitaine, Florian Petrilli, Ibtissam Dich, et al.. Monitoring and methods for estimating the biodiversity of a watercourse in an alpine ecosystem by metabarcoding analyses. 2026. <hal-05495981>

HAL Id: hal-05495981

<https://hal.science/hal-05495981v1>

Preprint submitted on 5 Feb 2026

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

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



Distributed under a Creative Commons CC BY-NC-ND 4.0 - Attribution - Non-commercial use - No Derivative Works - International License

Monitoring and methods for estimating the biodiversity of a watercourse in an alpine ecosystem by metabarcoding analyses

Amélie Bernard¹, Augustin Maniouloux¹, Clémentine Capitaine¹, Florian Petrilli¹, Ibtissam Dich¹, Jehanne Cordin¹, Jolan Soyer¹, Manon Petit¹, Perrine Duquet¹, Valentin Garreau¹, Clémentine François^{1,2}, Johan Pansu^{1,2}, Perrine Cruaud^{1,2}, Tristan Lefebure^{1,2}, Annabelle Haudry^{1,3}.

¹ Université Claude Bernard Lyon I, 69100 Villeurbanne, France

² Laboratoire d'Ecologie des Hydrosystèmes Naturels et Anthropisés, CNRS UMR 5023, Université Claude Bernard Lyon I, 69622 Villeurbanne Cedex, France

³ Laboratoire de Biométrie et Biologie Evolutive, CNRS UMR 5558, Université Claude Bernard Lyon I, 69622 Villeurbanne Cedex, France

Abstract :

Alpine ecosystems, like many other environments, are impacted by climate change. These alpine areas are characterized by significant biodiversity, which is highly vulnerable to climatic changes. To preserve these fragile habitats, it is essential to characterize and monitor their diversity. To this end, we aimed to identify plant and mammal species present in a watershed in Savoie, in the French Alps, using eDNA and metabarcoding. In order to verify the completeness of our inventory, we compared our results with those of the organization responsible for cataloguing fauna and flora (CREA). The protocol was also tested with and without human blocking primers to determine whether this improved the results. Our findings show that, for plants, eDNA approaches and traditional inventories conducted by CREA are complementary, as they identify different species. However, for mammals, CREA was more effective in identifying the species present. The use of human blocking primers proved beneficial for sediment samples, reducing human contamination, but had no significant impact on water samples. Estimates of richness and species composition revealed a differentiation between samples collected at high altitudes and those from lower elevations. Taxonomic assignments further highlighted the superior resolution of the water matrix for both plants and mammals, although sediments provided complementary information by detecting unique taxa.

Keywords : Metabarcoding, Alpine Biodiversity, Bioinformatic, Illumina Sequencing, Blocking Primer, Centre de Recherche des Écosystèmes d'Altitudes (CREA)



Introduction

Global warming, primarily caused by greenhouse gas emissions from human activities, has numerous impacts on the environment and biodiversity; it plays a key role in altering landscapes and driving the rapid, widespread disruption of biodiversity we witness today (Moinard 2023). Some habitats are more affected than others, leading to their classification based on their level of vulnerability (Dudgeon et al. 2006, Vautier et al. 2024). This includes alpine habitats, which are particularly fragile ecosystems (Lamouille-Hébert et al. 2024).

Among these vulnerable habitats, mountainous ecosystems such as the Alps stand out for their exceptional biodiversity and sensitivity to environmental changes. The Alps host approximately 4,500 plant species, including 400 endemics, and up to 30,000 animal species, making this region a major biodiversity hotspot in Europe (Theurillat and Guisan 2001). In addition to climatic pressures, land-use changes, such as the abandonment of traditional agricultural practices or the intensification of tourism, have significantly altered Alpine ecosystems over the past decades (Cernusca et al. 1996) (Merler et al. 1996). These transformations affect species interactions, increase the vulnerability of biological communities, and facilitate the spread of parasites and diseases (Hudson et al. 2002) (Chemini et al. 1997).

To enhance conservation efforts in these areas, it is essential to better understand and evaluate the status of the species occupying them (Moinard 2023). Various techniques are employed for this purpose, such as biodiversity inventories. While essential, these inventories are labour-intensive and time consuming (Moinard 2023). However, other techniques, such as environmental DNA (eDNA), have emerged in recent years to support these efforts.

In our study, numerous traditional inventories have already been conducted, and since 2022 the potential use of eDNA to enhance these

inventories is being explored. eDNA offers several advantages over traditional methods : notably, it can detect species that are difficult to observe, such as rare species or non-native species, at an early stage of their presence in a given environment. Furthermore, this method is non-invasive, fast, and increasingly cost-effective; it also significantly reduces the time and difficulty of fieldwork, thereby minimising risks for observers (Giraud et al. 2024).

Metabarcoding approaches theoretically allow the identification of all species present in an environment through simple sampling of the studied matrix. The use of generalist or universal primers facilitates the detection of the desired species.

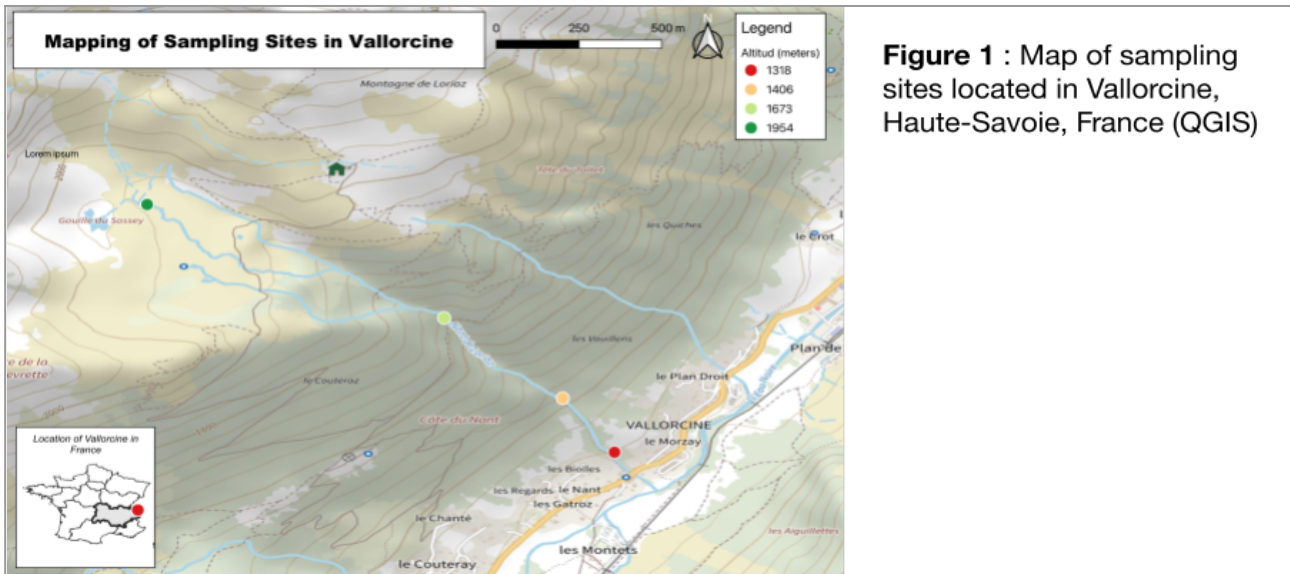
We have defined several objectives that this project aims to address: (1) Compare biodiversity characteristics obtained using eDNA data and inventory data from the *Centre de Recherche des Écosystèmes d'Altitude* (CREA Mont Blanc), (2) assess the complementarity of eDNA signals from different matrices (water and sediment) and (3) evaluate the integration of the eDNA signal along an upstream-to-downstream gradient.

Additionally, we will explore methodological questions such as the effectiveness of human-blocking probes and the sequencing depth required to detect the maximum number of species, to further refine subsequent analysis.

Material and methods

Location

To carry out this project, we travelled to the Haute-Savoie department (74), specifically to the commune of Vallorcine, located in the French Alps. We sampled the Nant de Loriaz river, which is 2.18 km long. It originates at approximately 1,984 meters above sea level and flows into the Eau Noire at 1,294 meters above sea level (Sandre n.d.).



Sampling

We chose four sampling sites along the river (Figure 1), from upstream (station 1) to downstream (station 4). On each station, we sampled water and sediment matrices in triplicates, taking into account site heterogeneity to maximise representativity. We collected 26 samples on the field and added a water filtration control, bringing the total to 27 samples (Annex 1).

Water sampling was conducted by pumping water through a Waterra capsule with a 0.45 µm filter. We filtered 20L for each sample except the 3rd replicate of the 4th station (sample E4C) where only 13L were filtered due to a time constraint (Figure 3). Once the water was filtered, 50mL of Longmire's lysis buffer were added and the capsule was shaken for 10 min manually to suspend the particles from the filter. The resulting mix was placed in a 50mL falcon and in a freezer at -20C for conservation

until extraction. The filtration control was done by filtering sterilized water.

Sediment sampling was conducted with a spoon directly from the riverbed into a 50mL falcon. After decantation, any additional water was poured out of the tube, and the tubes were also placed in a freezer at -20C.

DNA extraction and sequencing

DNA extraction : We extracted DNA from water samples, including an extraction control (ETEX, see Annex 2), using the NucleoSpin eDNA Water kit (Macherey-Nagel ref : 740402); and DNA from sediment samples including two extraction controls (STEX1 and STEX2, see Annex 2) using the NucleoSpin Soil kit (Macherey-Nagel ref : 740780). We measured total DNA concentration using a high sensitive (HS) fluorometry kit with Qubit.

Amplicon region : For the plant analysis, we targeted the P6 loop region of the trnI intron using the forward and reverse trnI primers (Taberlet et al. 2007) (Table 1). The size of the amplicon can range from 10 to 145 bp, with an average size of 45 bp. For mammals, we targeted the 16S mitochondrial gene with the MamP007 primer pair (Giguet-Covex et al. 2014) (Table 1). The resulting amplicon can vary between 60 and 84 pb. We also used a blocking primer (MamP007_B_Hum1) (Giguet-Covex et al. 2014) (Table 1) in the mammal DNA extraction to avoid human DNA amplification and contamination. Replicates A

Table 1 : Primer sequence used for plant and mammals, and for blocking primer

Primers	illumina adapter + primer sequence (5'-3')	Reference
trnI-g-F1	CTTTCCCTACACGACGCTCTTCCG ATCTGGGCAATCTCGAGCCAA	(Taberlet et al. 2007)
trnI-g-R1	GGAGTTCAGACGTGTGCTCTTCCG ATCTCCATTGAGTCTCGCACCTAT C	(Taberlet et al. 2007)
MamP007_F1	CTTTCCCTACACGACGCTCTTCCG ATCTCGAGAAGACCCTATGGAGCT	(Giguet-Covex et al. 2014)
MamP007_R1	GGAGTTCAGACGTGTGCTCTTCCG ATCTCCGAGGTCRCCCAACC	(Giguet-Covex et al. 2014)
MamP007_B_Hum1	GGAGCTTTAATTIATTAATGCAAAACA GTACC-C3-	(Giguet-Covex et al. 2014)

from each sampling site were also processed without blocking primers as a control. Universal Illumina adapter tails were added to the trnL and MamP007 primers to enable sequencing (Table 1).

PCR : PCR was performed in triplicates using the Qiagen Master Mix Multiplex PCR Plus (ref : 206152). The PCR cycling parameters were as follows: initial denaturation for 15 minutes at 95°C, followed by 45 cycles of denaturation for 30 seconds at 95°C, annealing for 30 seconds at 50°C, extension for 30 seconds at 72°C, followed by a final extension for 10 minutes at 68°C. All triplicates were pooled in a new PCR plate. Some samples were pooled in duplicates only due to pipetting errors during the PCR mix distribution. PCR products quality was checked by gel electrophoresis, using five 50ml gels at 1% with 0.5X TAE and the 1kb DNA ladder (Jena Bioscience) (Annex 3).

Sequencing : For Illumina sequencing, the second PCR and the sequencing process were outsourced to the sequencing platform in Toulouse (France Génomique). The amplicons were sequenced using the Illumina MiSeq platform (150 bp paired end).

Bioinformatic analysis

The quality of the sequenced reads was assessed using FastQC (v0.11.9), and primers were removed with Cutadapt (v3.5). We filtered the reads based on quality, merged R1 and R2, removed chimeras, and generated ASVs using DADA2 (v1.32.0)

Clustering : ASVs were clustered into OTUs using Swarm (v3.1.5), followed by denoising with Lulu (v0.1.0).

Taxonomic Assignment : Taxonomic assignment was performed using VSEARCH (v2.21.1) with different parameters to assess potential improvements in the results. First, the Lowest Common Ancestor (LCA) threshold was set to 0.95, and the sequence similarity (ID) was set to 0.98. Then, the LCA threshold was adjusted to 0.98, and the ID was set to 0.90. Three databases were used for taxonomic assignment: a global database for mammals based on Mamm02 sequences obtained from

NCBI, and two databases for plants—one local database provided by CREA and a global database based on Sper01 sequences, also retrieved from NCBI. As no differences were observed in taxonomic assignments across the tested parameters, subsequent data processing and analyses were performed using an LCA threshold of 0.95 and a sequence similarity of 0.98.

Decontamination : The cleaning of contaminants was carried out in parallel using Metabar (v.1.0.0) and Decontam (v.24.0). True contaminants within samples have been identified and discarded using the `subset_metabarlist` function. External contaminants were excluded using `decontam`. The two OTU tables obtained were then merged into a single table.

Monitoring and biological question

The dataset was processed and analysed using R libraries (R version 4.4.2, 2024-10-31).

CREA data comparison : To evaluate the complementarity between the metabarcoding approach and the traditional inventory methods used by CREA for plants and mammals, a comparison was conducted between the local database (corresponding to the taxonomic assignments obtained via metabarcoding) and the CREA database. This analysis aimed to quantify the proportion of shared genera between the two datasets. In addition, for CREA data, floristic surveys were carried out over the years 2016 and 2020. In order to limit the loss of information, we decided to take into account the two-year records as the CREA database.

Matrix complementarity : To statistically analyse the complementarity between the two matrices sampled (water and sediment), the datasets presenting the samples with OTUs and the associated taxonomic assignments were rarefied to the lowest number of OTUs in the samples for each matrix using a Phyloseq object (Package R Phyloseq) for mammals and plants. Firstly, complementarity in the number of OTUs assigned to a species according to the water and sediment matrix was measured for

mammals and plants using Venn diagrams. Similar Jaccard's presence/absence matrices were then calculated for plants in order to compare samples from the different matrices, taking only those OTUs with a taxonomic assignment to species. A similar analysis could not be carried out for mammal samples because the quantitative differences in terms of OTUs were too large between the water and sediment matrices to obtain statistically robust results. The datasets with the local and global plant reference databases were merged to keep the assignments unique. If two assignments were discordant, the taxonomic assignment from the local reference database took priority over the global reference database. The similarity matrices of the plant matrices were then visualised using dimensionality reduction (NDMS). PERMANOVAs (n=999) were also performed on the similarity matrices to statistically test the difference between the water and sediment samples. The comparison of taxonomic richness between samples and matrices for mammals and plants was carried out by taking into account the abundance of OTUs assigned by plants to a family level and mammals to a genus one. As for the comparison of the alpha diversity of the samples and matrices for the two taxa, Shannon indices were established, and the statistical difference was calculated using Wilcoxon tests.

Altitudinal question : Analyses were performed on mammal and plant OTU tables, utilizing metadata and taxonomic assignments. Data from water and sediment samples were extracted and cleaned to remove missing values. OTU tables and taxonomic annotations were integrated with environmental data for each matrix using the phyloseq library (v1.48.0). OTU data were normalized by rarefaction to standardize sequencing depth across samples, with sample size set to the minimum observed value. The Shannon diversity index was used to assess species richness in the samples. Variations across sites and matrices (water/sediment) were tested using Kruskal-Wallis tests. Bray-Curtis distance matrices were computed to quantify compositional differences among communities.

Non-Metric Multidimensional Scaling (NMDS) analyses were employed to visualise the data, and multi-site dissimilarities were evaluated using the betapart package (v1.6). It should be noted that no mammal-specific diversity analysis was conducted using the sediment matrix, as no DNA was sequenced from it.

Methodological questions : Statistical differences between samples with or without bhum were tested with a Wilcoxon test. To verify whether the sequencing depth in our samples was sufficient, we used the rarecurve function from the vegan package (v 2.6-8).

Results and Discussion

Data comparison with existing CREA data

Comparison between the CREA database and the local database : A comparative analysis of plant and mammal genera detected using the DNA-based database (Taxons_DataBase_ADN) and the CREA database (Taxons_DataBase_CRE) reveals a notable difference in the diversity detected. The DNA-based database identified 538 unique genera, compared to 30 genera identified using the CREA database (Figure 2,(a)). An overlap of 44 genera was observed between the two databases (Figure 2,(a)), representing approximately 57% of the total genera identified by the CREA database. These findings validate the metabarcoding approach while emphasizing its ability to uncover a substantially greater diversity of genera. The observed disparity underscores the complementary nature of these methods, with metabarcoding serving as a powerful tool for comprehensive biodiversity assessments by capturing a broader range of taxa than traditional inventories alone.

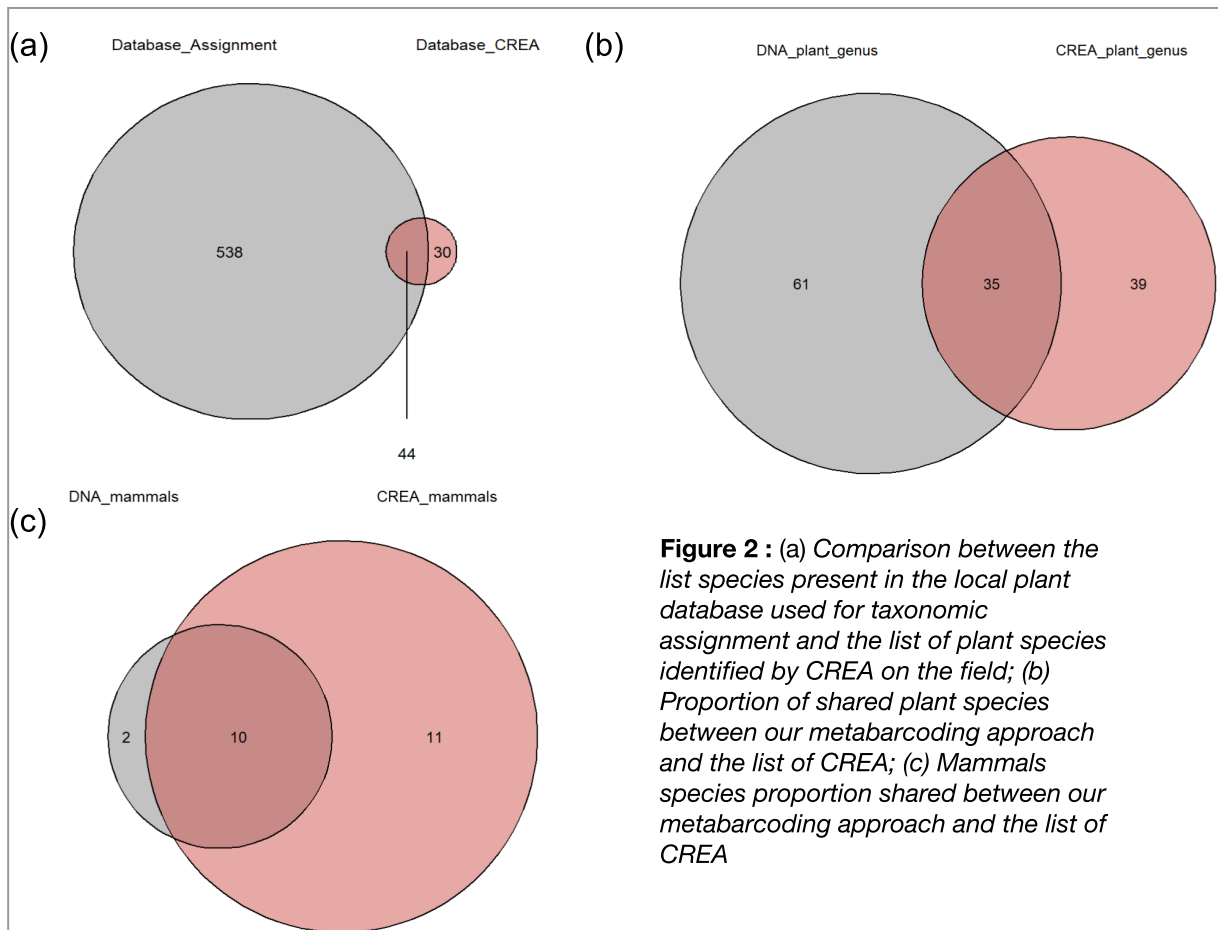


Figure 2 : (a) Comparison between the list species present in the local plant database used for taxonomic assignment and the list of plant species identified by CREA on the field; (b) Proportion of shared plant species between our metabarcoding approach and the list of CREA; (c) Mammals species proportion shared between our metabarcoding approach and the list of CREA

Comparison between the data of this study and the CREA database : A comparative analysis of plant genera based on our analyses using metabarcoding (DNA_plant_genus) and CREA inventories (CREA_plant_genus) revealed an overlap of 35 genera out of a total of 96 detected by metabarcoding and 74 by CREA (Figure 2(b)). Our results showed that metabarcoding identified 61 unique genera, while CREA inventories identified 39 unique genera (Figure 2(b)). This underscores the complementarity of the two methods, with metabarcoding detecting potentially cryptic diversity and CREA inventories capturing taxa that may not be amplified by PCR. These results highlight the superiority of the metabarcoding approach for a comprehensive biodiversity assessment. This method offers a much finer resolution than traditional inventories, in line with previous studies that have demonstrated its effectiveness in capturing a greater diversity of taxa in various ecological contexts (Taberlet et al. 2012, Deiner et al. 2017).

The overlap of only 57% between the two databases also shows that traditional methods, although limited in terms of diversity detected, can provide additional information. These methods can sometimes better detect rare species or taxa whose sequences are not yet included in the metabarcoding reference databases (Pawlowski et al. 2021). This reinforces the idea that DNA and traditional approaches are complementary, rather than mutually exclusive, to biodiversity studies.

Finally, the results highlight the importance of the quality of reference databases for the success of DNA analyses. Better taxonomic coverage of databases can increase the ability to identify rare or cryptic taxa, a challenge identified by several studies (Cristescu 2014). Integrating these databases with traditional tools could therefore be a promising way to achieve more robust and comprehensive ecosystem assessments. These results pave the way for future applications, particularly in conservation, where

metabarcoding could play a key role in providing accurate information on biodiversity in diverse environments.

A comparative analysis of mammal communities based on our results using metabarcoding (DNA_mammals) and traditional CREA inventories (camera trap) (CREA_mammals) showed both similarities and differences between the two methods. Metabarcoding identified 12 species, 10 of which were also detected by CREA, while CREA detected an additional 11 species not identified by metabarcoding (Figure 2(c)). Additionally, metabarcoding identified 2 species absent from the traditional inventories (Figure 2(c)). These findings highlight the complementary strengths of both approaches: traditional inventories provide a broader species count, while metabarcoding can uncover cryptic diversity and enable quicker, less intrusive detection.

The differences observed between plant and mammal results could potentially be explained by the lack of specificity of the metabarcoding marker for mammals. This might suggest that marker specificity plays a role in the relative effectiveness of metabarcoding across taxonomic groups, and highlights an area for methodological refinement in future studies.

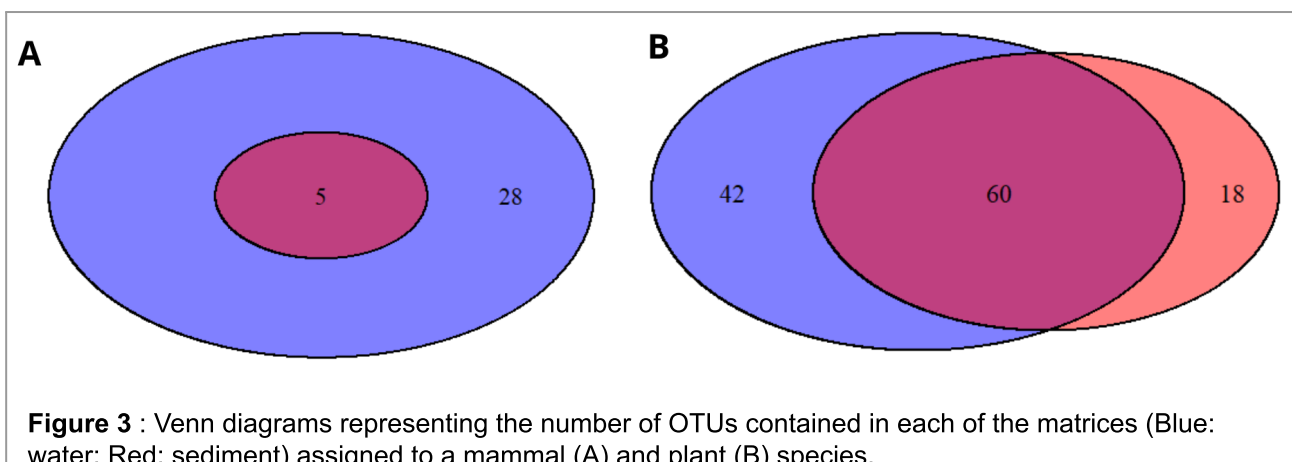
Most abundant genus : The CREA survey revealed that 90% of the plant species identified were represented by just 20 genera. While metabarcoding confirmed 14 of these, the remaining, most abundant genera were classified as Poaceae. The inability of the

primers to efficiently amplify Poaceae sequences may account for their absence in the metabarcoding results, despite their prevalence in the samples.

Matrix complementarity

After taxonomic assignment, a large portion of OTUs remains unassigned. Among the 364 OTUs identified for mammals, only 9.1% of OTUs in the water matrix and 1.4% in the sediment matrix are assigned to this taxon. The Venn diagram (Figure 3, A) shows that out of the 33 determined OTUs, 5 are shared between the two matrices, while 28 OTUs are specific to the water matrix. The taxonomic richness bar plots (Figure 4, B) also indicate a much lower diversity in sediments, with some samples, such as the S3 replicates, completely lacking any assignment. Furthermore, numerous human contaminations were observed, sometimes affecting entire samples (e.g., S4C and S2C). For plants, among the 311 identified OTUs, a relatively higher percentage is assigned in the water matrix (32.7%) compared to the sediment matrix (25%).

The similarity analysis based on the Jaccard matrix (presence/absence) and its visualization via NMDS reveals marked differences between water and sediment samples for plants. Water samples (blue points) form a well-defined group, reflecting relative homogeneity, while sediment samples (red points) show greater dispersion, indicating higher heterogeneity in this matrix. Samples from site 1 (S1A, B, C for sediments and E1A, B, C for water) show



notable similarity between the two matrices, which contrasts with other sites, where the matrices do not exhibit such correspondence. The Venn diagram (Figure 3, B) illustrates that out of 120 identified OTUs, 60 are common to both matrices, while 42 OTUs are specific to the water matrix and 18 are specific to the sediment matrix. This distinction between matrices is confirmed by a PERMANOVA and by the Shannon diversity index, which shows significantly higher diversity in water samples ($p < 0.05$). The bar plots (Figure 4, C) indicate a higher proportion of Pinaceae in water samples, particularly at site 4. Samples from site 1, for both water and sediment matrices, also stand out from other sites with their specific taxonomic proportions.

For mammals, the low assignment in sediments prevents an in-depth comparison between the matrices. However, the available results show a significant disparity: the water matrix dominates in terms of assigned taxonomic diversity, while some sediment samples completely lack any assignment. Similar to plants, samples from site 1 exhibit distinct

taxonomic richness compared to other sites, for both matrices.

It seems that the water matrix is more resolutive for plants than the sediment matrix. Although more OTUs identified for plants are detected in sediments, most of the assigned OTUs are found in water samples, suggesting better specificity of the primers for this taxon. This hypothesis is supported by the presence of a greater number of OTUs found specifically in the water matrix, using the same marker for both matrices. However, a large proportion of OTUs remains unassigned, likely due to the incompleteness of reference databases of plants, notably the absence of some groups (e.g. algae). Even with the combined use of a global database and a local one, the limitations of these references affect the quality of the assignments. Furthermore, the Shannon diversity index, which indicates higher diversity in water, aligns with this hypothesis. However, it is not necessary to rely solely on the water matrix. Indeed, the NMDS analysis and PERMANOVA, which reveal a significant difference between samples from the two

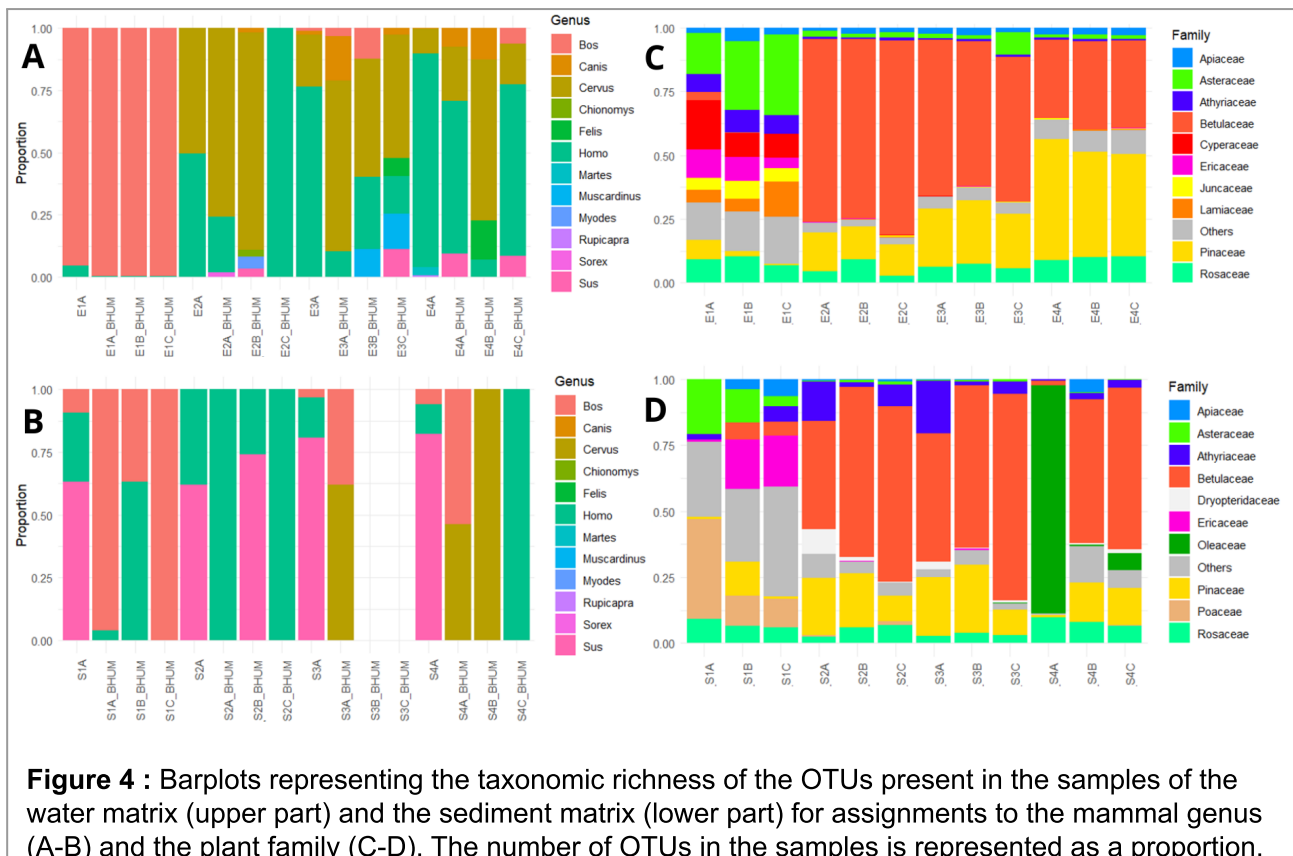


Figure 4 : Barplots representing the taxonomic richness of the OTUs present in the samples of the water matrix (upper part) and the sediment matrix (lower part) for assignments to the mammal genus (A-B) and the plant family (C-D). The number of OTUs in the samples is represented as a proportion.

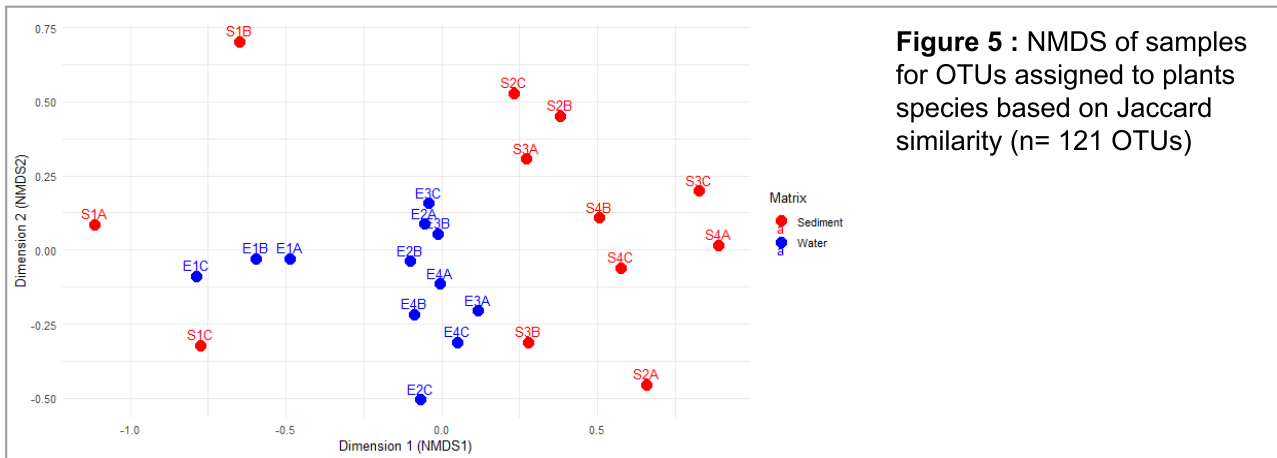


Figure 5 : NMDS of samples for OTUs assigned to plants species based on Jaccard similarity (n= 121 OTUs)

matrices, encourage the use of both methods, despite the lower resolution in sediments. Additionally, the bar plots confirm the complementary nature of the two matrices. Despite a lower number of OTUs, the sediment matrix allows for the detection of plant families such as Dryopteridaceae and Poaceae, which do not appear in the analysis of water samples. Moreover, the fact that the first sample site is particularly distinct from the others within each matrix may bias our results. This site exhibits greater diversity than the other sites in water samples, which could overestimate the overall diversity of water samples—a methodological bias noted in other studies on riverine plant eDNA (Espinosa Prieto et al. 2024).

For mammals, the comparison between the two matrices is more challenging due to the low number of OTUs in sediments. At first glance, the Venn diagram and bar plots suggest that the water matrix is sufficient and that sediment sampling is not relevant, as all OTUs are covered by the water matrix, and some sediment samples detect only a single mammal genus. Some studies (Sales et al. 2020) present the water matrix as more resolutive for mammals, arguing that taxa detected in sediments come from established populations, requiring a minimum threshold of individuals for detection. However, we do not recommend considering our results as reliable. Firstly, because the sequencing depth is too low to compare the matrices; and secondly, because the marker used is likely too broad to specifically target mammals, thus diluting the target sequences among non-target taxa.

Integration of DNA signal along the ecological gradient

Mammal diversity along gradient in water

matrix data : The alpha diversity obtained from mammal-specific primers shows notable differences among the four sampling sites along the altitudinal gradient. Sites at higher altitudes exhibit slightly lower species richness compared to downstream sites, reflecting a reduction in available niches for mammals at higher elevations. Beta diversity analyses indicate significant turnover between sites, particularly through the balanced variation component of abundance, suggesting a substantial replacement of mammal species across different altitudes. No notable unidirectional abundance gradient was detected (Table 2), reinforcing the idea that variations between sites are primarily driven by species replacement rather than overall abundance differences.

Plant diversity along altitudinal gradient in water and sediment matrix :

The plant diversity data obtained from samples in the two matrices reveal that the alpha diversity of plants at high altitudes is significantly higher than at the three other sites located at lower altitudes, which show no significant differences among themselves (Figure 6c and 6e). These observations suggest that sites at the highest altitude harbour species-rich communities, while at lower altitudes, a homogeneity of environmental conditions could explain the similarity in their alpha diversity. This data highlights a possible niche differentiation

starting at the second site in mid-altitude. Beta diversity analyses conducted using abundance matrices show a strong structuring of communities along the altitudinal gradient. The partitioning of beta diversity indicates that the total dissimilarity between sites is entirely attributable to the turnover component, with no

unidirectional abundance gradient detected (Table 2). This reflects a significant replacement of species between sites, without a progressive increase or decrease in overall abundances along the gradient. These results are corroborated by the NMDS analysis, where data from the water matrices reveal a clear

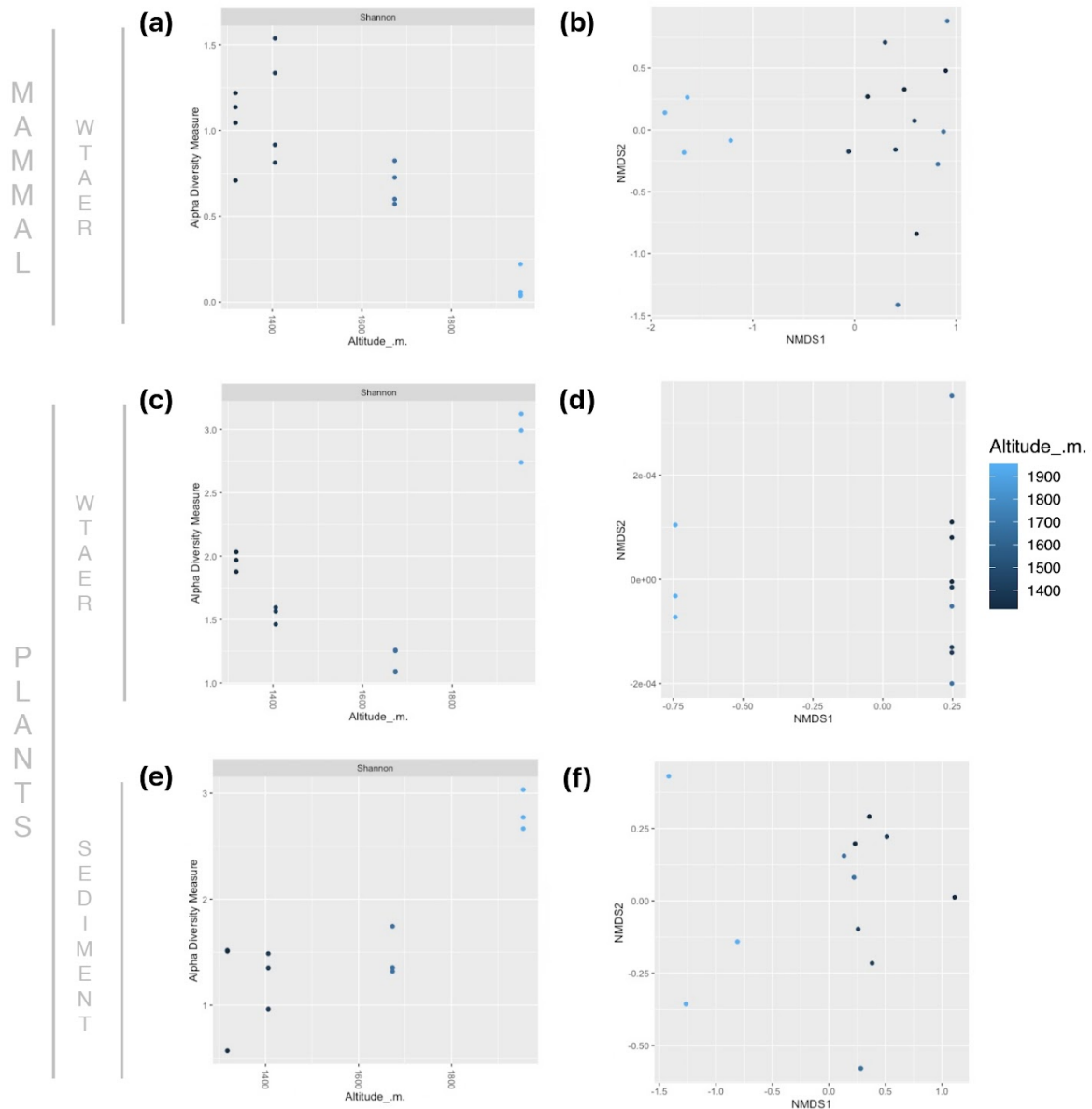


Figure 6: alpha and beta diversity plots of mammal and plant in water matrix for mammal and water and sediment matrix for plants : (a) alpha mammal diversity in water matrix, (b) beta mammal diversity in water matrix, (c) alpha plant diversity in water matrix, (d) beta plant diversity in water matrix, (e) alpha plant diversity in sediment matrix, (f) beta plant diversity in sediment matrix

Table 2: beta partition index calculate by betapart package on beta diversity of mammals and plants

Beta partition	“betapart” index	Mammal water	Plant water	Plant sediment
Turnover	\$beta.BRAY.BAL	0.69	0.67	0.68
Abundance gradient	\$beta.BRAY.GRA	0	0	0
Total dissimilarity	\$beta.BRAY	0.69	0.67	0.68

separation of high-altitude sites from those at intermediate and lower altitudes. Intermediate and lower-altitude sites form a relatively homogeneous group, with similar coordinates along the main axis of variation. In contrast, high-altitude sites display strongly negative coordinates, indicating a distinct community composition. The minimal variation observed along the secondary axis (MDS2) suggests that community structuring is primarily driven by altitude (Figure 6d). Analyses from the sediment matrices similarly reveal distinct structuring of the high-altitude site, which is separated from

the three other sites at mid and low altitudes. However, sediment data also show greater variation along the main NMDS axis and the notable presence of one high-altitude replicate that strongly deviates from the rest (Figure 6f).

These results emphasize that altitude acts as a major structuring factor, influencing community composition through strong species turnover between sites. Communities at high-altitude sites are characterized by specialized species adapted to extreme environmental conditions, while downstream sites share more

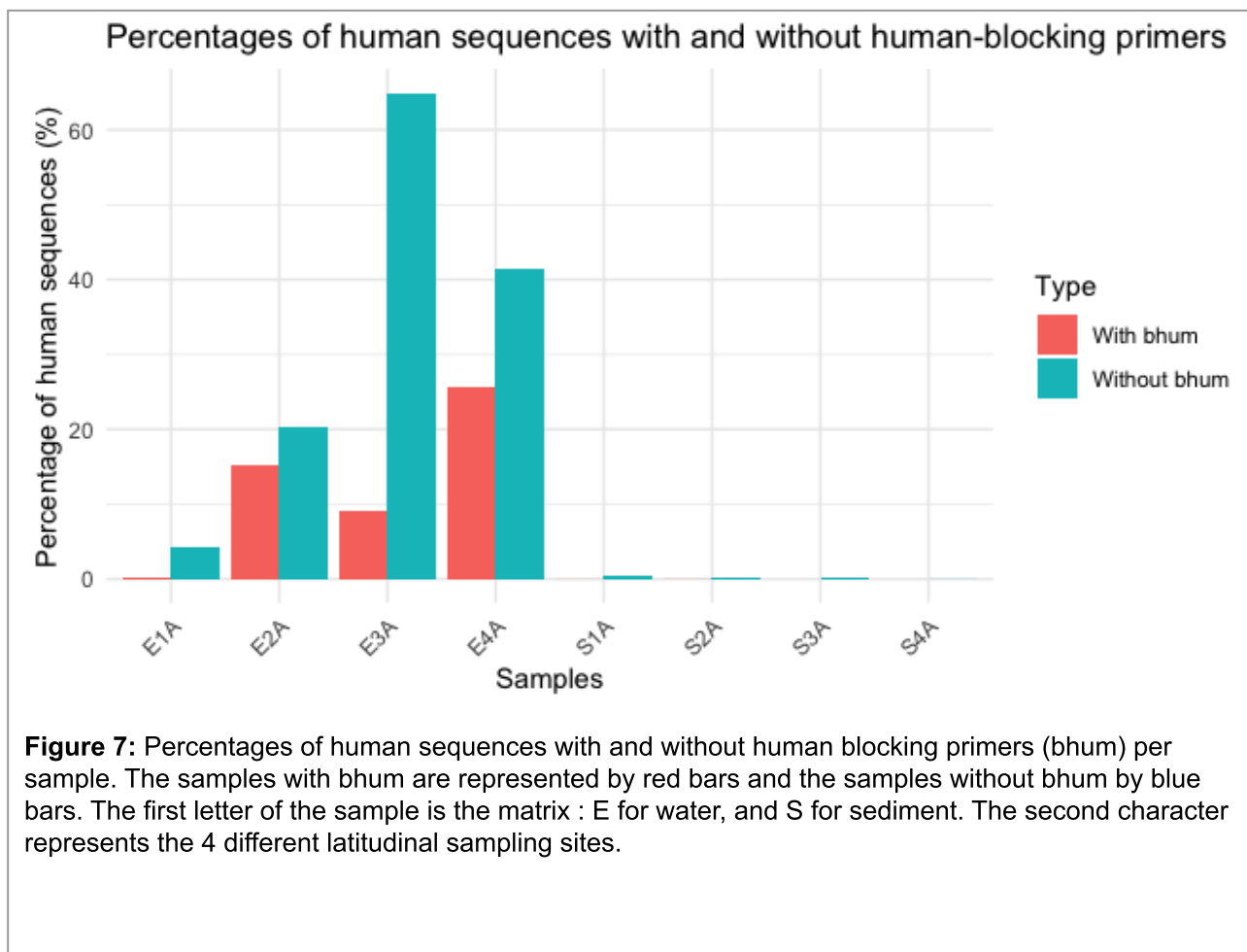


Figure 7: Percentages of human sequences with and without human blocking primers (bhum) per sample. The samples with bhum are represented by red bars and the samples without bhum by blue bars. The first letter of the sample is the matrix : E for water, and S for sediment. The second character represents the 4 different latitudinal sampling sites.

homogeneous communities, shaped by more stable and diverse environmental conditions. These findings align with previous biodiversity studies along altitudinal gradients in the Alps, where species turnover is a dominant driver of alpine species composition, not limited to plants and mammals. Such analyses are increasingly important in the context of climate change, as it remains challenging to identify thresholds or clear patterns of dissimilarity across different altitudinal zones. Addressing these challenges is a crucial step toward implementing conservation strategies tailored to the specific needs of each alpine ecosystem (Viterbi et al. 2013, Lefebvre et al. 2018, Fontana et al. 2020).

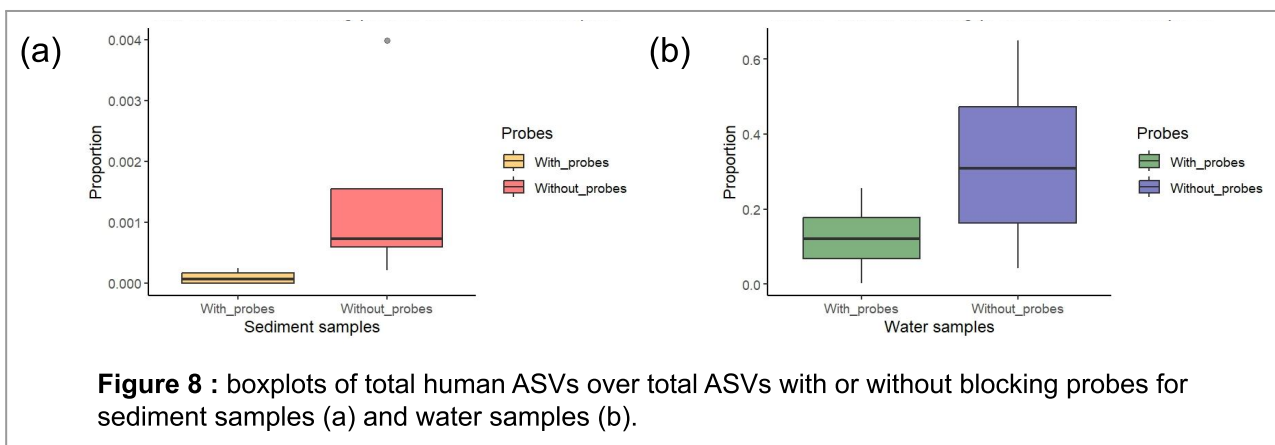
Methodology question

This study is the second of its kind, conducted by master's students under the same leadership team, using samples from the same river. The aim is to continually refine the protocol in order to enhance both efficiency and results in the years to come. This year, two key questions were raised: (1) should subsequent attempts also incorporate blocking primers, and (2) is the current sequencing depth optimal for the analysis?

Blocking primers effectiveness : One of the biggest issues with amplicon sequencing is contaminant amplification. A way to avoid such an amplification is to use blocking primers, provided that the target or contaminant is known (Vestheim et Jarman 2008). In this case, the main suspected contaminant was human DNA, whether it occurred during the sampling process or the extraction steps (Giguet-Covex

et al. 2014). To test the effectiveness of the blocking primers used, we amplified a few samples without blocking primers, while the others were amplified using blocking primers. We then compared the results of the analysis between the two types of samples, with or without primers. To test whether that difference was significant, a paired Wilcoxon test was performed. This test yielded a p-value of 0.125 for the water samples, indicating no significant difference between samples treated with a human blocking probe and those without. However, the test was significant for sediment samples (p-value = 0.007813). The effectiveness of the blocking probe appears to depend on the matrix used. In the case of sediments, the probe could consistently be used to reduce human sequence reads and, consequently, the proportion of reads assigned to humans in future studies.

OTU distribution : Among the 12 mammal species identified across all samples, 9 species were detected in the samples containing the human-blocking primers (bhum), while only 7 species were found in the samples without the blocking primers. The use of the human-blocking primers thus facilitated more taxonomic assignments. However, three species (*Rupicapra rupicapra*, *Bos mutus*, and *Martes martes*) were not detected in the samples with bhum. The five species missing from the samples without the primers were *Muscardinus avellanarius*, *Felis catus*, *Myodes glareolus*, *Chionomys nivalis*, and *Sorex minutus*. The findings suggest that the inclusion of human-blocking primers in the sequencing protocol increases the ability to detect a broader range of mammalian species. This is



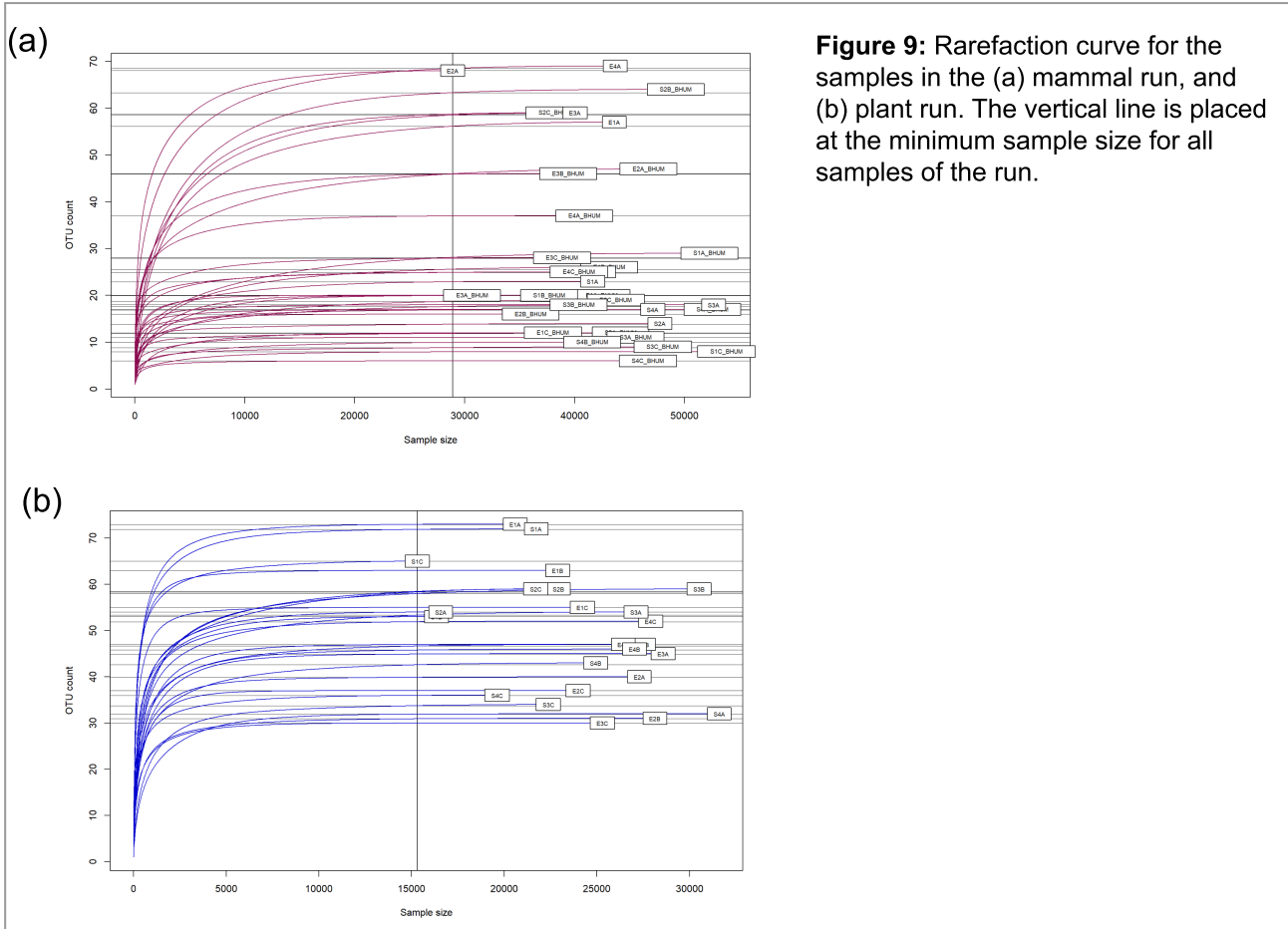


Figure 9: Rarefaction curve for the samples in the (a) mammal run, and (b) plant run. The vertical line is placed at the minimum sample size for all samples of the run.

likely due to the reduction of human DNA interference, which can otherwise obscure the presence of other species. The absence of these 3 species in the samples with bhum might indicate that the primers may also have unintentionally affected the detection of these species. On the other hand, the 5 species missing from the samples without blocking primers might have been underrepresented due to the higher background noise from human DNA in these samples. These results confirm that the bhum successfully decreased human sequences in the samples tested, aligning with findings in environmental DNA studies that demonstrate the effectiveness of blocking primers in reducing human DNA contamination and enhancing the detection of target species (Boessenkool et al. 2012). This highlights the importance of optimizing primer selection in order to enhance the detection of target species while minimizing bias from human DNA.

Sequencing depth : A major parameter in sequencing is the depth of coverage. It refers to the number of times a specific base (nucleotide) in the DNA is read during the sequencing process, and is often measured as the average number of reads per base in a sample. It is primarily controlled by the number of sequencing cycles run by the sequencing platform, and is directly linked to the cost of the sequencing. A higher sequencing depth allows for more precise data, as it enables error correction, and in metabarcoding studies such as the present one, it is also directly linked to the accurate detection of rare taxa (Sims et al., 2014). Rarefaction curves (Figure 9) show that for all samples, the curves surpass the inflection point where slopes begin to decrease, and reach a plateau (asymptote). This indicates that the sequencing depth is sufficient to detect the biodiversity contained in every sample, and it could even be lowered without much loss of information.

Conclusion

This study aimed to utilise environmental DNA (eDNA) as a complementary approach to traditional inventories for assessing species diversity in fragile ecosystems, with a particular focus on alpine areas. The primary objective was to evaluate the degree of complementarity between the two methods, as well as the complementarity of the two sampling matrices: water and sediment. Additionally, we aimed to determine whether the information accumulated along an upstream to downstream gradient or if the species detected were more reflective of their altitudinal distribution. The study also sought to address several methodological questions to enhance the quality of results in future similar studies, specifically regarding sequencing depth and the use of blocking primers.

We found that traditional inventories and eDNA approaches do not uncover the same diversity, with traditional approaches finding a broader range of species, while eDNA surveys have a greater capacity to detect cryptic species. The water matrix is more resolutive than the sediment matrix, and altitude acts as a major structuring factor in the data collected from an up to downstream gradient. Human blocking primers have a clear positive effect on species detection, and the sequencing depth utilised in this study did not seem to negatively affect species detection.

Acknowledgment

We want to thank our teachers for their help in all steps of this experience. We also want to thank CREA for giving us their inventory of plants and mammals.

Authors contribution

- Sampling : All authors equally
- Molecular biology : All authors equally
- Bioinformatic pipeline :
 - data processing : All authors equally

- Clustering/denoising : Soyer, Petit, Cordin, Duquet
- Taxonomic assignments : Bernard, Maniouloux, Petrilli
- Decontamination : Garreau, Dich, Capitaine
- Statistics :
 - > CREA : Petit, Dich, Duquet
 - > Matrix : Soyer, Maniouloux
 - > Ecological gradient : Petrilli, Garreau
 - > Methodology : Capitaine, Cordin, Bernard
- Writing:
 - Abstract : Petrilli, Petit
 - Introduction : Bernard, Duquet, Capitaine, Petrilli
 - Material and methods : Duquet, Petrilli, Bernard, Cordin, Petit, Maniouloux
 - Results and discussion : All authors
 - Conclusion : Cordin
- Review : Cordin
- Script : Soyer, Maniouloux
- Logo : Duquet
- Page layout: Petrilli, Duquet

References

- Boessenkool, S., L. S. Epp, J. Haile, E. Bellemain, M. Edwards, E. Coissac, E. Willerslev, and C. Brochmann. 2012. Blocking human contaminant DNA during PCR allows amplification of rare mammal species from sedimentary ancient DNA. *Molecular Ecology* 21:1806–1815.
- Cernusca, A., U. Tappeiner, M. Bahn, N. Bayfield, C. Chemini, F. Fillat, W. Graber, M. Rosset, R. Siegwolf, and J. Tenhunen. 1996. Ecological Effects of Land Use Changes on European Terrestrial Mountain Ecosystems. *Pirineos : Revista de Ecología de Montaña* 147–148.
- Chemini, C., A. Rizzoli, S. Merler, C. Furlanello, and C. Genchi. 1997. *Ixodes ricinus* (Acari: Ixodidae) infestation on roe deer (*Capreolus capreolus*) in Trentino, Italian Alps. *Parassitologia* 39:59–63.
- Cristescu, M. E. 2014. From barcoding single individuals to metabarcoding biological communities: towards an integrative approach to the study of global biodiversity. *Trends in Ecology & Evolution* 29:566–571.
- Deiner, K., H. M. Bik, E. Mächler, M. Seymour, A. Lacoursière-Roussel, F. Altermatt, S. Creer, I. Bista, D. M. Lodge, N. de Vere, M. E. Pfrender, and L. Bernatchez. 2017. Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology* 26:5872–5895.
- Dudgeon, D., A. H. Arthington, M. O. Gessner, Z.-I. Kawabata, D. J. Knowler, C. Lévêque, R. J. Naiman, A.-H. Prieur-Richard, D. Soto, M. L. J. Stiassny, and C. A. Sullivan. 2006. Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews of the Cambridge Philosophical Society* 81:163–182.
- Espinosa Prieto, A., L. Hardion, N. Debortoli, T. Bournonville, T. Mathot, J. Marescaux, E. Chanez, C. Staentzel, and J.-N. Beisel. 2024. A comparative analysis of eDNA metabarcoding and field surveys: Exploring freshwater plant communities in rivers. *Science of The Total Environment* 954:176200.
- Fontana, V., E. Guariento, A. Hilpold, G. Niedrist, M. Steinwandter, D. Spitale, J. Nascimbene, U. Tappeiner, and J. Seeber. 2020. Species richness and beta diversity patterns of multiple taxa along an elevational gradient in pastured grasslands in the European Alps. *Scientific Reports* 10:1–11.
- Giguet-Covex, C., J. Pansu, F. Arnaud, P.-J. Rey, C. Griggo, L. Gielly, I. Domaizon, E. Coissac, F. David, P. Choler, J. Poulénard, and P. Taberlet. 2014. Long livestock farming history and human landscape shaping revealed by lake sediment DNA. *Nature Communications* 5:3211.
- Giraud, C., N. Wabete, C. Lemeu, N. Selmaoui-Folcher, D. Pham, V. Boulo, and N. Callac. 2024. Environmental factors and potential probiotic lineages shape the active prokaryotic communities associated with healthy *Penaeus stylirostris* larvae and their rearing water. *FEMS Microbiology Ecology* 100:fiac156.
- Hudson, P., A. Rizzoli, B. Grenfell, J. A. P. Heesterbeek, and A. Dobson. 2002. *The Ecology of Wildlife Diseases*.
- Lamouille-Hébert, M., F. Arthaud, and T. Datry. 2024. Climate change and the biodiversity of alpine ponds: Challenges and perspectives. *Ecology and Evolution* 14:e10883.
- Lefebvre, V., C. Villemant, C. Fontaine, and C. Daugeron. 2018. Altitudinal, temporal and trophic partitioning of flower-visitors in Alpine communities. *Scientific Reports* 8:4706.
- Merler, S., C. Furlanello, C. Chemini, and G. Nicolini. 1996. Classification Tree Methods for Analysis of Mesoscale Distribution of *Ixodes ricinus* (Acari: Ixodidae) in Trentino, Italian Alps. *Journal of Medical Entomology* 33:888–893.
- Moinard, S. 2023, December 13. Estimateurs de biodiversité robustes pour le metabarcoding. phdthesis, Université Grenoble Alpes [2020-....].
- Pawłowski, J., A. Bonin, F. Boyer, T. Cordier, and P. Taberlet. 2021. Environmental DNA for biomonitoring. *Molecular Ecology* 30:2931–2936.
- Sales, N. G., M. B. McKenzie, J. Drake, L. R. Harper, S. S. Browett, I. Coscia, O. S.

- Wangenstein, C. Baillie, E. Bryce, D. A. Dawson, E. Ochu, B. Hänfling, L. Lawson Handley, S. Mariani, X. Lambin, C. Sutherland, and A. D. McDevitt. 2020. Fishing for mammals: Landscape-level monitoring of terrestrial and semi-aquatic communities using eDNA from riverine systems. *Journal of Applied Ecology* 57:707–716.
- Sandre. (n.d.). Nant de Ioriaz [V0370560] - Cours d'eau selon la version Carthage 2017. Text. https://www.sandre.eaufrance.fr/geo/CoursEau_Carthage2017/V0370560.
- Taberlet, P., E. Coissac, F. Pompanon, C. Brochmann, and E. Willerslev. 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology* 21:2045–2050.
- Taberlet, P., E. Coissac, A. Valentini, T. Vermet, and E. Willerslev. 2007. Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Research* 35.
- Theurillat, J.-P., and A. Guisan. 2001. Potential Impact of Climate Change on Vegetation in the European Alps: A Review. *Climatic Change* 50:77–109.
- Vautier, M., H. Rogissart, J. Guillard, and I. Domaizon. 2024. Development of quantitative eDNA methods for the study of fish phenology in lakes eDNA-based innovations: Deliverable 3 – Advances in eRNA approaches and other eDNA-based innovations.
- Viterbi, R., C. Cerrato, B. Bassano, R. Bionda, A. von Hardenberg, A. Provenzale, and G. Bogliani. 2013. Patterns of biodiversity in the northwestern Italian Alps: a multi-taxa approach. *Community Ecology* 14:18–30.

Supplementary annexes

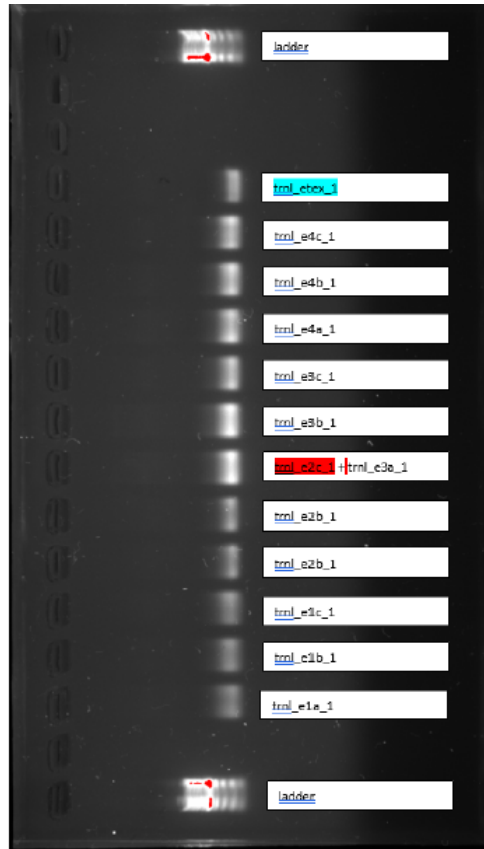
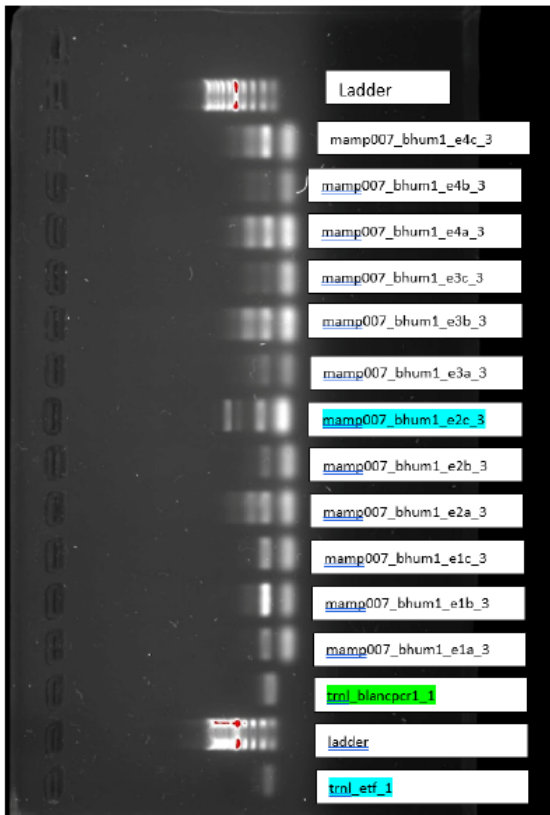
Annex 1: Sampling Site Metadata Table

TABLEAU RECAPITULATIF ECHANTILLONNAGE VALLORCINE 2023													
code_echantillon	surface_echantillonnee	site_echantillonne	echantillon	date	coordonnees_GPS	altitude	heure	type_sediment	prelevement	courant	quantite_eau	quantite_tampon	remarque
S1A	sediment	site 1	A	05/09/2023	46.03675N, 6.91058E ± 5m	1954m	17h	grossier	cuillere	moyen			
S1B	sediment	site 1	B	05/09/2023	46.03675N, 6.91058E ± 5m	1954m	17h	fin	seringue	stagnant			
S1C	sediment	site 1	C	05/09/2023	46.03675N, 6.91058E ± 5m	1954m	17h	grossier	cuillere	rapide			
E1A	eau	site 1	A	05/09/2023	46.03675N, 6.91058E ± 5m	1954m	17h				20L	50mL	
E1B	eau	site 1	B	05/09/2023	46.03675N, 6.91058E ± 5m	1954m	17h				20L	50mL	Pompe debranchee à moitie du temps
E1C	eau	site 1	C	05/09/2023	46.03675N, 6.91058E ± 5m	1954m	17h				20L	40mL	
S2A	sediment	site 2	A	06/09/2023	46.03304N, 6.91739E ± 4m	1673m	11h	grossier + fin	cuillere + seringue	moyen + stagnante			
S2B	sediment	site 2	B	06/09/2023	46.03304N, 6.91739E ± 4m	1673m	11h	MO+++	cuillere	moyen			
S2C	sediment	site 2	C	06/09/2023	46.03304N, 6.91739E ± 4m	1673m	11h	gravier moyen	cuillere	moyen + remourds			
E2A	eau	site 2	A	06/09/2023	46.03304N, 6.91739E ± 4m	1673m	11h				20L	50mL	
E2B	eau	site 2	B	06/09/2023	46.03304N, 6.91739E ± 4m	1673m	11h				20L	50mL	
E2C	eau	site 2	C	06/09/2023	46.03304N, 6.91739E ± 4m	1673m	11h				9L	50mL	colmatage du filtre de la capsule
S3A	sediment	site 3	A	06/09/2023	46.02978N, 6.92202E ± 5m	1406m	12h45	moyen grossier+fin	cuillere	moyen			
S3B	sediment	site 3	B	06/09/2023	46.02978N, 6.92202E ± 5m	1406m	12h45	fin	cuillere	lent			
S3C	sediment	site 3	C	06/09/2023	46.02978N, 6.92202E ± 5m	1406m	12h45	grossier	cuillere	lent			
E3A	eau	site 3	A	06/09/2023	46.02978N, 6.92202E ± 5m	1406m	12h45				20L	50mL	
E3B	eau	site 3	B	06/09/2023	46.02978N, 6.92202E ± 5m	1406m	12h45				20L	50mL	trajet difficile donc conta tres possible
E3C	eau	site 3	C	06/09/2023	46.02978N, 6.92202E ± 5m	1406m	12h45				20L	50mL	
S4A	sediment	site 4	A	06/09/2023	46.02758N, 6.92404E	1318m	14h	grossier	cuillere	moyen			
S4B	sediment	site 4	B	06/09/2023	46.02758N, 6.92404E	1318m	14h	grossier	cuillere	moyen			
S4C	sediment	site 4	C	06/09/2023	46.02758N, 6.92404E	1318m	14h	grossier/fin	cuillere	moyen			
E4A	eau	site 4	A	06/09/2023	46.02758N, 6.92404E	1318m	14h				20L	50mL	
E4B	eau	site 4	B	06/09/2023	46.02758N, 6.92404E	1318m	14h				20L	50mL	
E4C	eau	site 4	C	06/09/2023	46.02758N, 6.92404E	1318m	14h				13L	50mL	limite par le temps
ET-	eau ultra pure	parking campus		06/09/2023								1L	tampon avec impurte donc potentiel contamination

Annex 2: concentration of DNA extracted from samples

Echantillons	concentration Qubit (ng/mL)	concentration tube final (ng/μL)
E1A	16.2	3.24
E1B	8	1.6
E1C	8	1.6
E2A	6.91	1.382
E2B	36.2	7.24
E2C	0.56	0.112
E3A	53	10.6
E3B	24.7	4.94
E3C	8.59	1.718
E4A	38.9	7.78
E4B	9.1	1.82
E4C	2.19	0.438
S1A	70.5	14.1
S1B	173	34.6

S1C	85	17
S2A	59	11.8
S2B	118	23.6
S2C	134	26.8
S3A	124	24.8
S3B	81.9	16.38
S3C	110	22
S4A	77.6	15.52
S4B	61.7	12.34
S4C	38.3	7.66
ETF	<0.5	/
ETEX	<0.5	/
STEX1	<0.5	/
STEX2	0	/



Annex 3: Electrophoresis Gel Results for DNA Extractions