Institute for Biodiversity and Freshwater Conservation | Institiùd Bith-iomadachd is Glèidhteachais Fìor-uisg'

Allt Mor eDNA Baseline Assessment (2025)

Final report: January 2026

Prepared by: **Dr Nathan Griffiths**

Other UHI project staff:
Dasha Svobodova
Charley Clark
Prof. Bernd Hänfling

Project Leads:
Dr Nathan Griffiths (UHI Inverness)
Ffion Robb (Spey Catchment Initiative)

Correspondence:

Dr. Nathan Griffiths
UHI Inverness
1 Inverness Campus
Inverness
IV2 5NA
Email: nathan.griffiths.ic@uhi.ac.uk

1. Background

Environmental DNA (eDNA) metabarcoding is increasingly used for biodiversity monitoring of aquatic habitats. Metabarcoding combines the collection of eDNA with high-throughput sequencing, and allows the simultaneous characterisation of entire biological communities. Previous research has shown that eDNA metabarcoding is more effective at detecting elusive fish species than established invasive surveying techniques such as electric fishing or netting, and in both lentic and lotic habitats (Griffiths *et al.*, 2020; Handley *et al.*, 2019; Hänfling *et al.*, 2016; Li *et al.*, 2019; Pont *et al.*, 2018). eDNA metabarcoding approaches have also been successfully used to characterise aquatic invertebrate communities (Deiner *et al.*, 2016; Harper *et al.*, 2021) and to detect cryptic non-native invertebrate species (Blackman *et al.*, 2018). DNA is extracted from an environmental sample and a targeted region of the mitochondrial genome, termed a 'barcode', is amplified by polymerase chain reaction (PCR) and sequenced. As different mitochondrial regions have different levels of sequence variation in different taxonomic groups, the specific barcode selected for a study will depend on the taxa of interest. The different barcode sequences identified in the sample are compared against a species sequence database to determine the species present where the sample was collected.

For this biodiversity baseline assessment, 10 sampling sites were surveyed within the Allt Mor restoration site, and analysed using vertebrate (Kelly *et al.*, 2014; Riaz *et al.*, 2011) and invertebrate (Leese *et al.*, 2021) eDNA metabarcoding workflows to determine the species present within the catchment. These sites were specifically selected in order to provide 'baseline' data on the current biodiversity directly at the restoration site and also in a control region directly upstream. Two temporal visits are carried out here, in order to maximise baseline data obtained. While the control samples will enable us to determine the species which are already present in the upstream catchment.

2. Methods summary

All water samples (n = 20) were collected by UHI and SCI staff on 17th July 2025 and 17th September 2025. Each individual sample contained 1.5l of surface water, and the same sampling points were visited on each occasion (Figure 1). Samples were collected over a ~10m stretch, each including 5 x 400ml sub-samples taken from flowing river water, stored in a cool box on ice, and filtered within 24 hours of collection. DNA was extracted following the Mu-DNA water extraction protocol (Sellers *et al.*, 2018). Three PCR replicates were carried out for each sample prior to being pooled. Samples were further processed and sequenced following metabarcoding protocols established at UHI Inverness using a vertebrate specific 12S marker (Riaz et al. 2011) and then an invertebrate specific CO1 marker (Leese *et al.*, 2021).

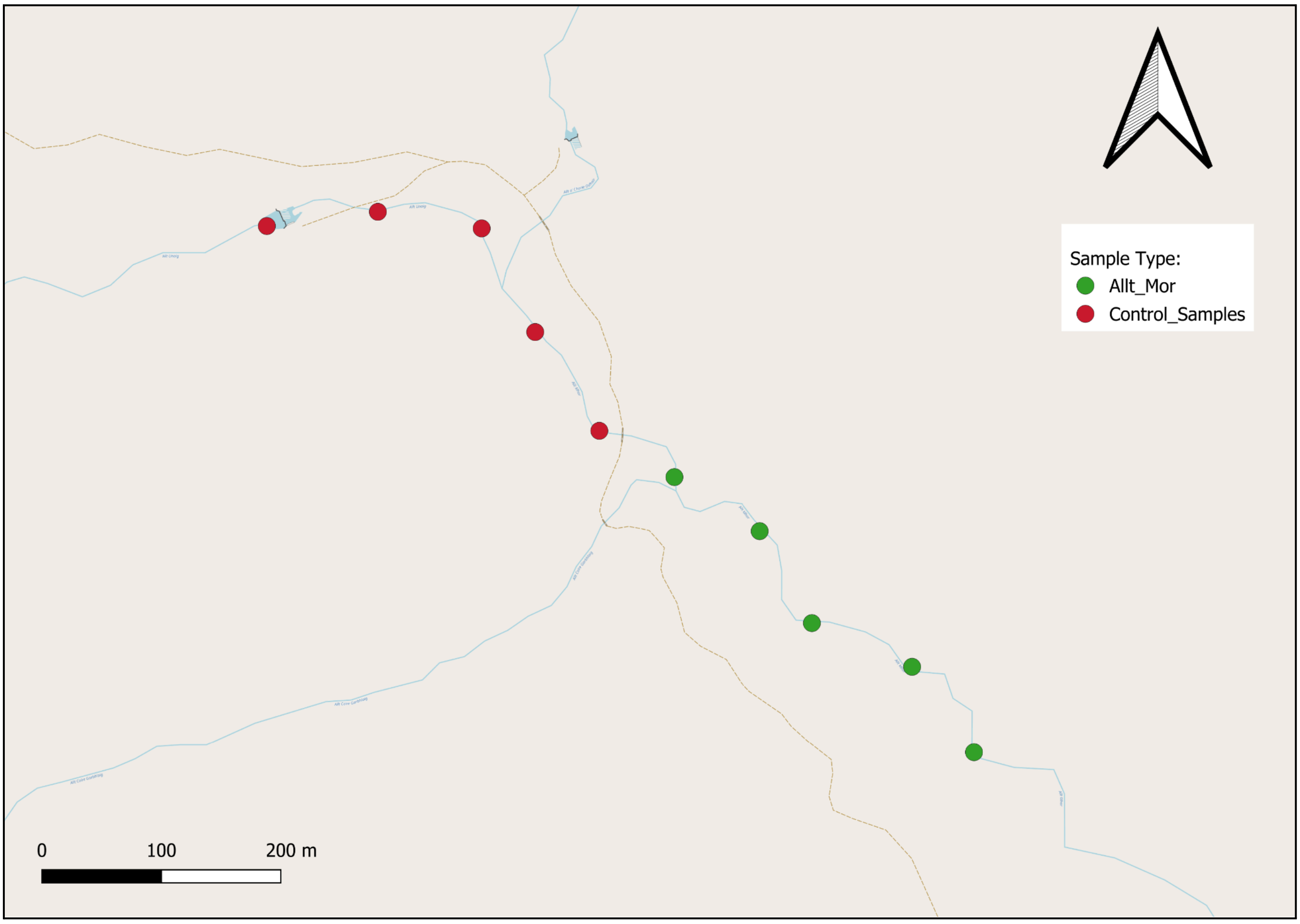

Figure 1. Map of the 10x sampling points taken for this baseline survey at Allt Mor, showing the restoration site in Green (🟢) and the upstream control samples in Red (🔴).

Raw sequencing data were analysed using a reproducible metabarcoding bioinformatic workflow, *Tapirs* (https://github.com/EvoHull/Tapirs). Sequencing reads underwent a *BLAST* (Zhang *et al.*, 2000) taxonomic assignment against curated reference databases for UK vertebrates and invertebrates respectively. A criterion of 98% (vertebrates) and 95% (invertebrates) identity with the reference sequences was applied. Following taxonomic assignment, a noise threshold of 0.1% of total reads per sample was applied when handling fish data to remove low-frequency reads (Hänfling *et al.*, 2016), while a minimum reads threshold was applied to invertebrate data, removing any species with <20 reads assigned. Reads assigned to positive controls, reads which could not be assigned to any taxon and samples with no taxonomically assignable reads were also excluded from the analysis.

3. Results

DNA extraction yielded sufficient quality and quantity for all samples, both 12S and COI metabarcoding were performed successfully.

3.1 Vertebrate Metabarcoding

The average number of accepted 12S DNA sequence reads per sample was 65,467. Metabarcoding workflows included quality controls to determine PCR success and to monitor for potential contamination. Within the vertebrate workflow, only low levels of contamination from human DNA and taxonomically unassignable reads were observed across lab process blanks (Figure 2). However, these low levels of human contamination are of no concern, and thus all post-filtering threshold detections from environmental samples are regarded as being present at site. Note, the positive control *Maylandia zebra* consistently amplified across this sequencing run.

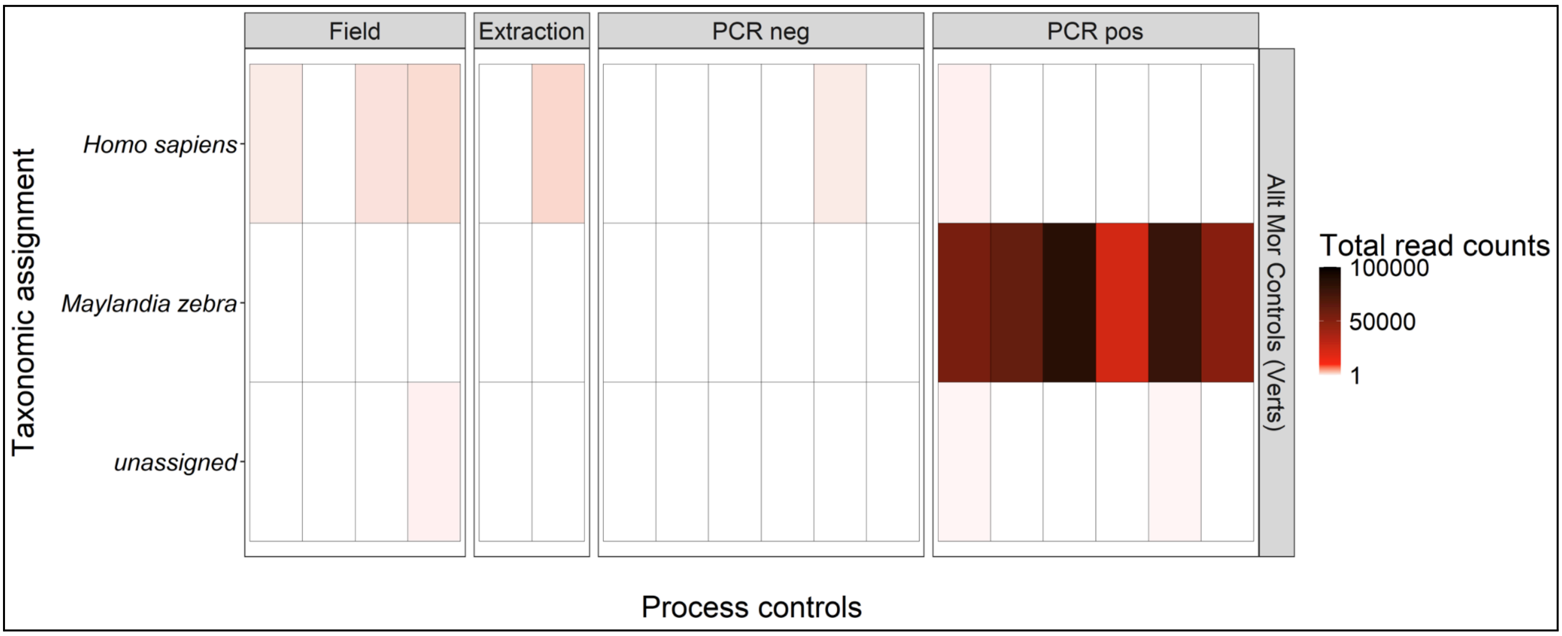

Figure 2. A proportional contamination heatmap to visualise contamination with vertebrate DNA across controls during different lab and field processes.

Following the 0.1% low-frequency reads threshold for fish, only brown trout (*Salmo trutta*) were detected across this baseline survey (Figure 3). Post-threshold, brown trout made up 100% of fish reads in every sample, in both July and September. However, prior to the fish threshold being applied, European eel (*Anguilla anguilla*) DNA was detected at very low signal in one sample taken in July (Figure 4). While this is unlikely to have arisen from lab contamination, standard practice with fish is to take a conservative approach and omit such low frequencies from metabarcoding datasets (Griffiths *et al.*, 2020; Hänfling *et al.*, 2016). In summary, this catchment is dominated by brown trout, although future work should not rule out the possibility of eel presence within this catchment.

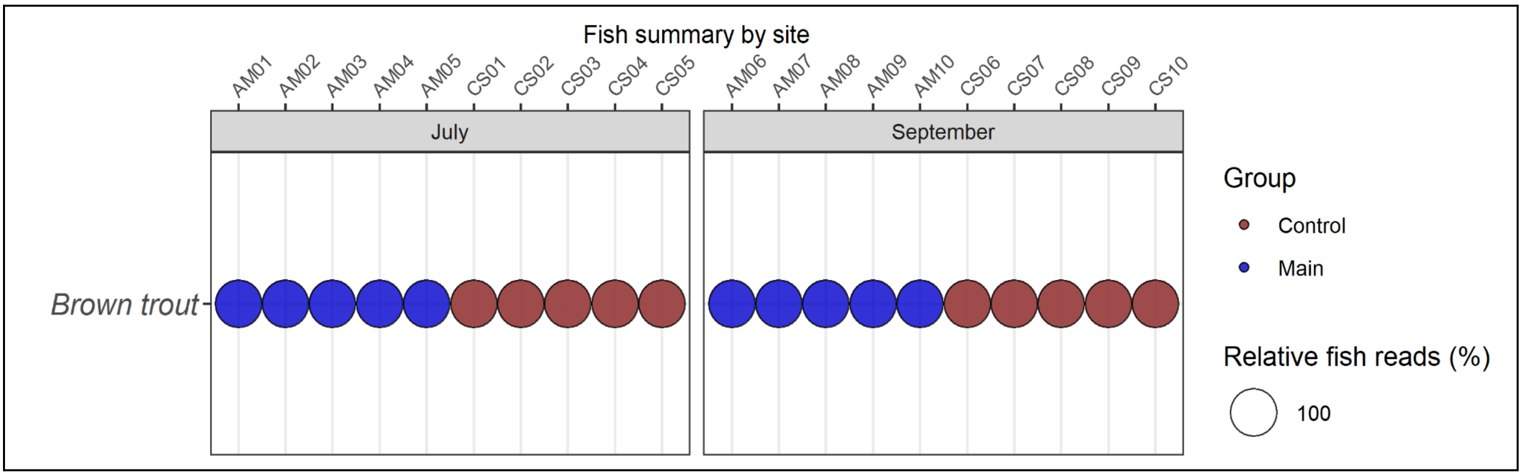

Figure 3. The fish community and relative abundance/reads, presented by sampling site as a bubble chart (common names are used here for ease of interpretation).

Across the baseline dataset, all vertebrate detections were assigned to a total of 28 taxonomic units. These comprised 2 fish species, 16 mammal species, 3 amphibian species, and 7 avian taxonomic groups. The taxonomic resolution of this 12S marker is variable for birds with several key groups only identifiable to family level; bird data were therefore included at a higher taxonomic level where appropriate. Note, that for this figure the 0.1% fish threshold has not been applied in order to allow visualisation of raw baseline data (Figure 4). Interestingly, more mammals were detected in September compared to July. This is most likely due to increased run-off and rainfall comparatively, leading to more accumulation / downstream transportation of DNA in the system of species which are only intermittently in contact with the water. This particular result highlights the importance of multiple visits when obtaining overall biodiversity baseline data.

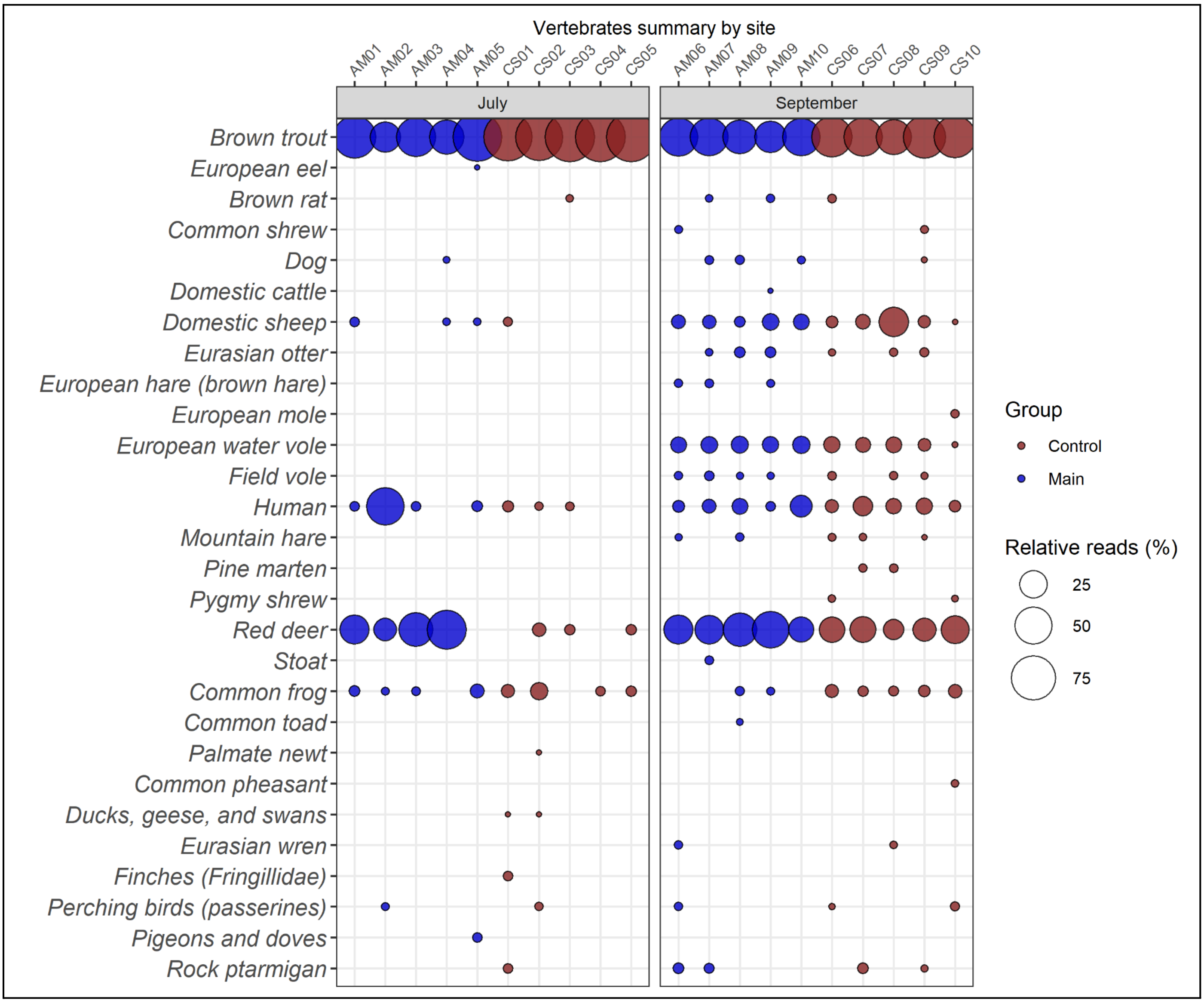

Figure 4. The overall vertebrate community and relative abundance/reads, presented by sampling site as a bubble chart (common names are used here for ease of interpretation).

3.2 Macroinvertebrate Metabarcoding

The average number of accepted COI DNA sequence reads per sample was 67,132. Metabarcoding workflows included quality controls to determine PCR success and to monitor for potential contamination. Within the invertebrate workflow, only low levels of contamination from 4 taxa were observed across lab process blanks (Figure 5). The most prevalent species found in blanks / negative controls was *Baetis rhodani*. Nonetheless, these detections were not consistent and were present in low frequencies. This low-level contamination is not of concern, and therefore our standard invertebrate <20 reads threshold is applied to all samples. If specifically considering the aforementioned species however, low-frequency reads should be interpreted with caution. Note, the positive control *Eudasyphora cyanicolor* consistently amplified across this sequencing run.

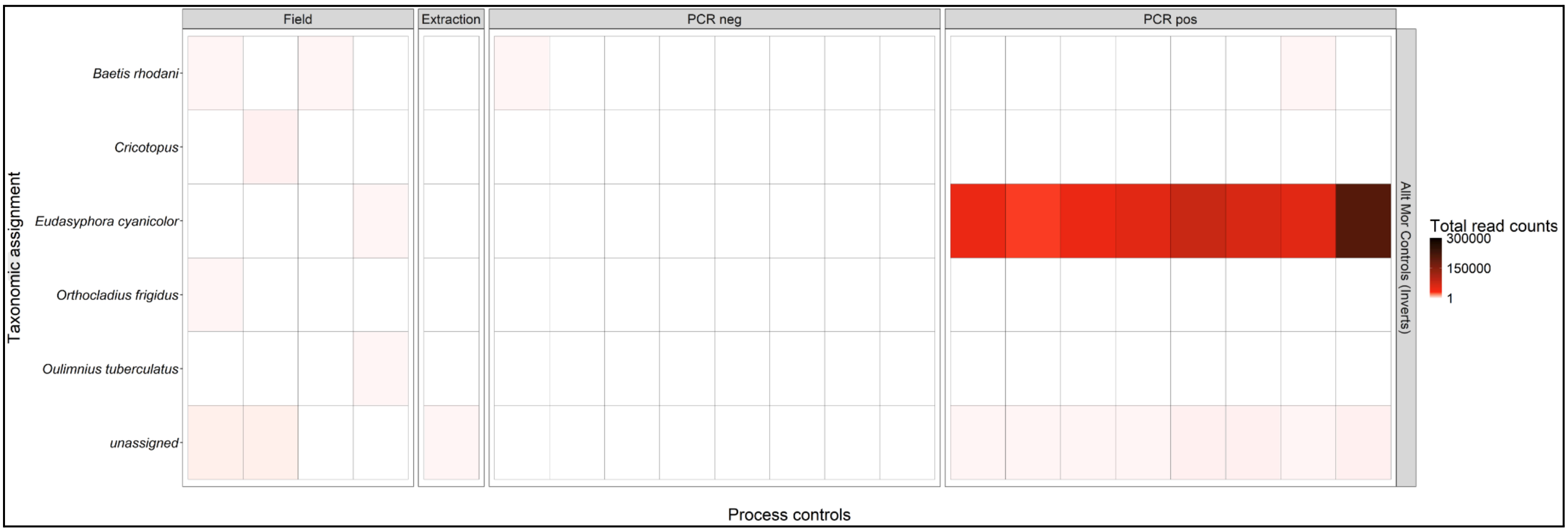

Figure 5. A proportional contamination heatmap to visualise contamination with invertebrate DNA across controls during different lab and field processes.

In total, the COI reads from this baseline survey were assigned to a total of 94 macroinvertebrate species, belonging to 10 different orders (Figure 6). An overview of the number of species detected within these different groups is highlighted below (Figure 6).

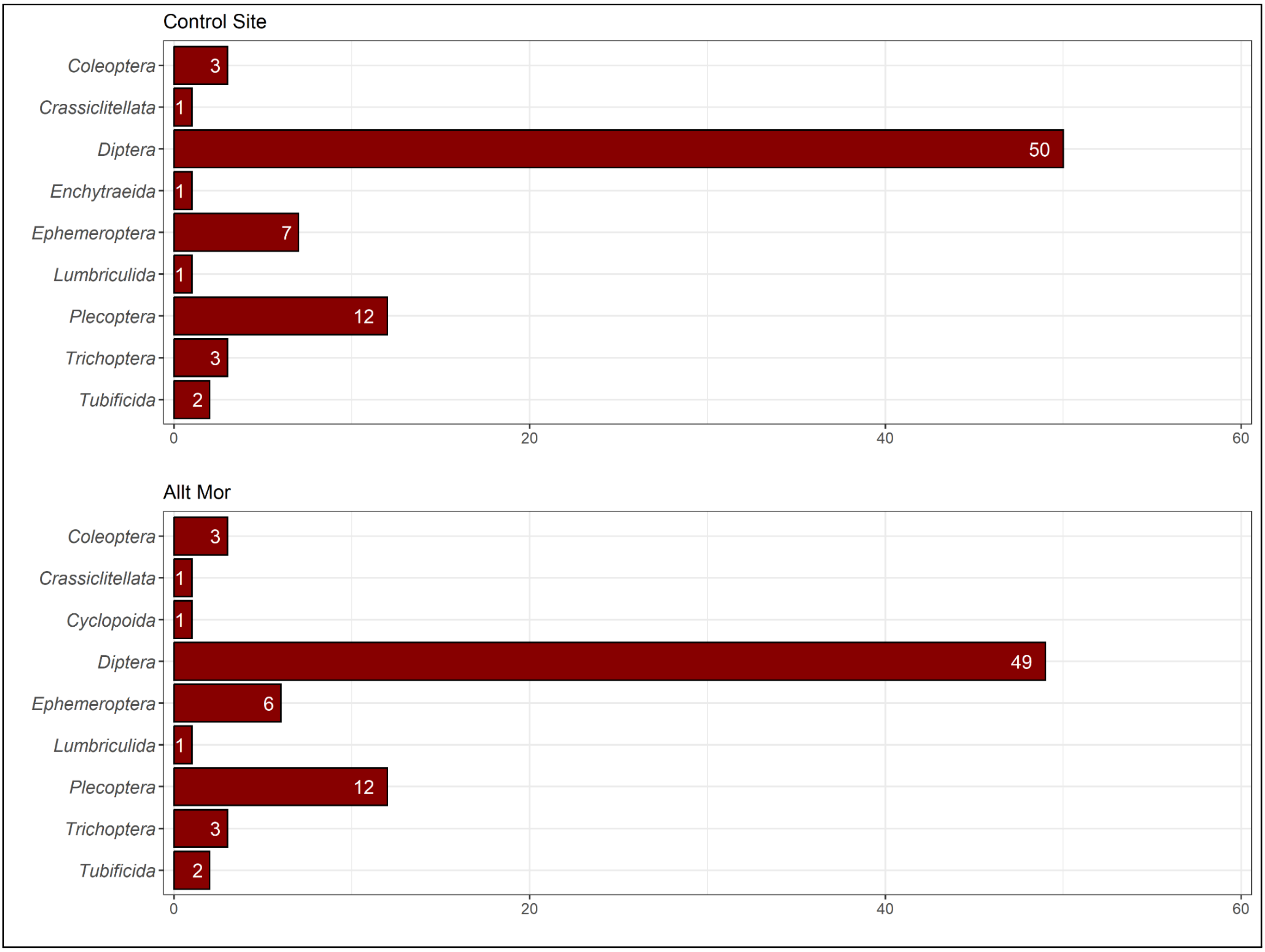

Figure 6. A visual overview of the number of invertebrate species detected within each group (Order) at the upstream control region and the Allt Mor restoration site.

Three of the insect families - Baetidae (mayflies), Leuctridae (stoneflies), and Chironomidae (Non-biting midges) - contributed the largest number of assigned reads. However several other families were also consistently detected across all sites and timeframes (Figure 7).

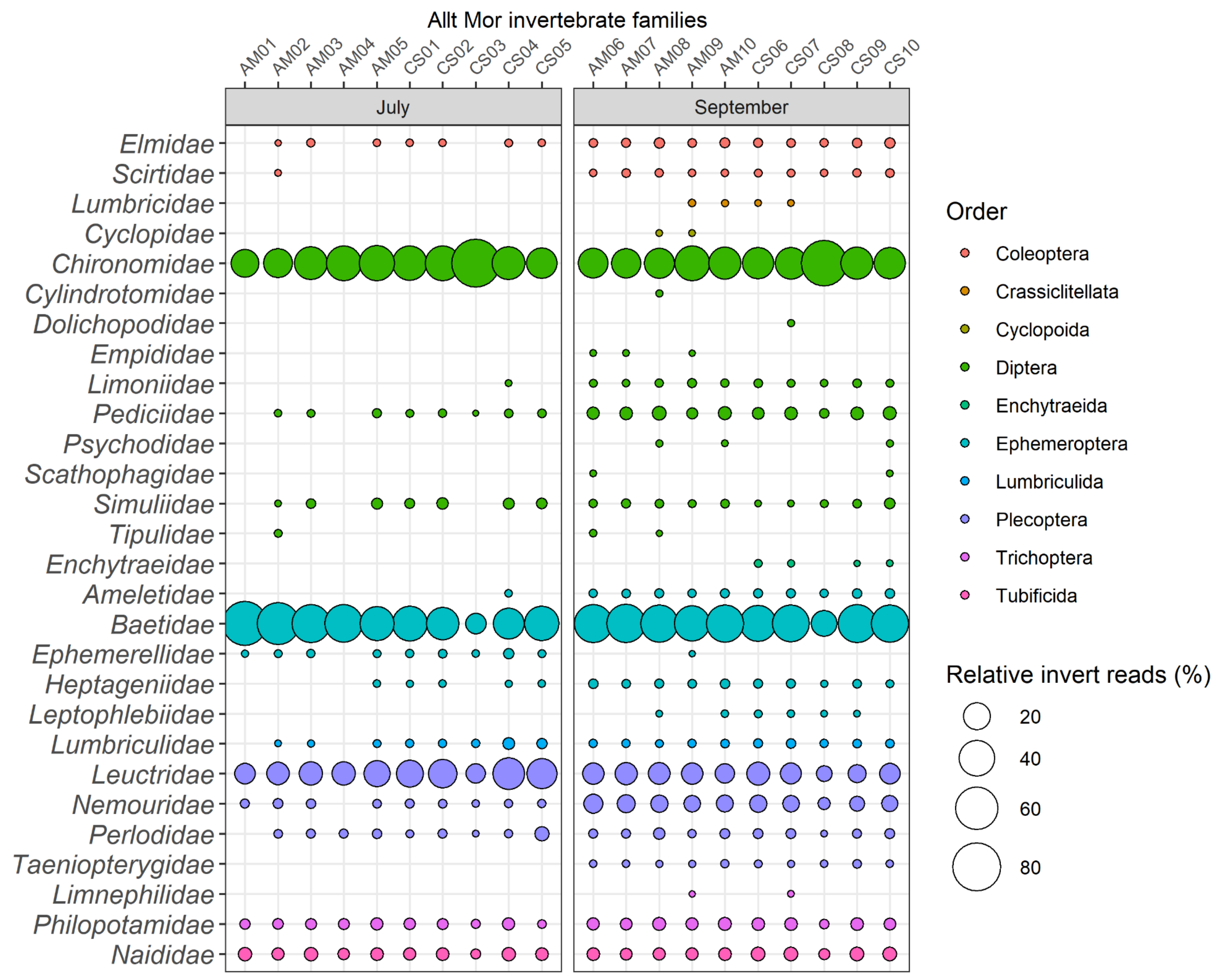

Figure 7. A visual overview of the invertebrate families detected, and their relative reads.

Read counts for all identified taxa with >1000 total reads in the dataset are shown in Figure 8. These show that the high read counts for Baetidae in particular are largely driven by a single species: the large dark olive mayfly *B. rhodani* (Figure 8). Not all invertebrate families and/or species were consistently identified at all sites with >1000 reads. In particular, several temporal shifts can be observed between the two different sampling months. However, within months, detections are mostly consistent when looking at species with >1000 reads.

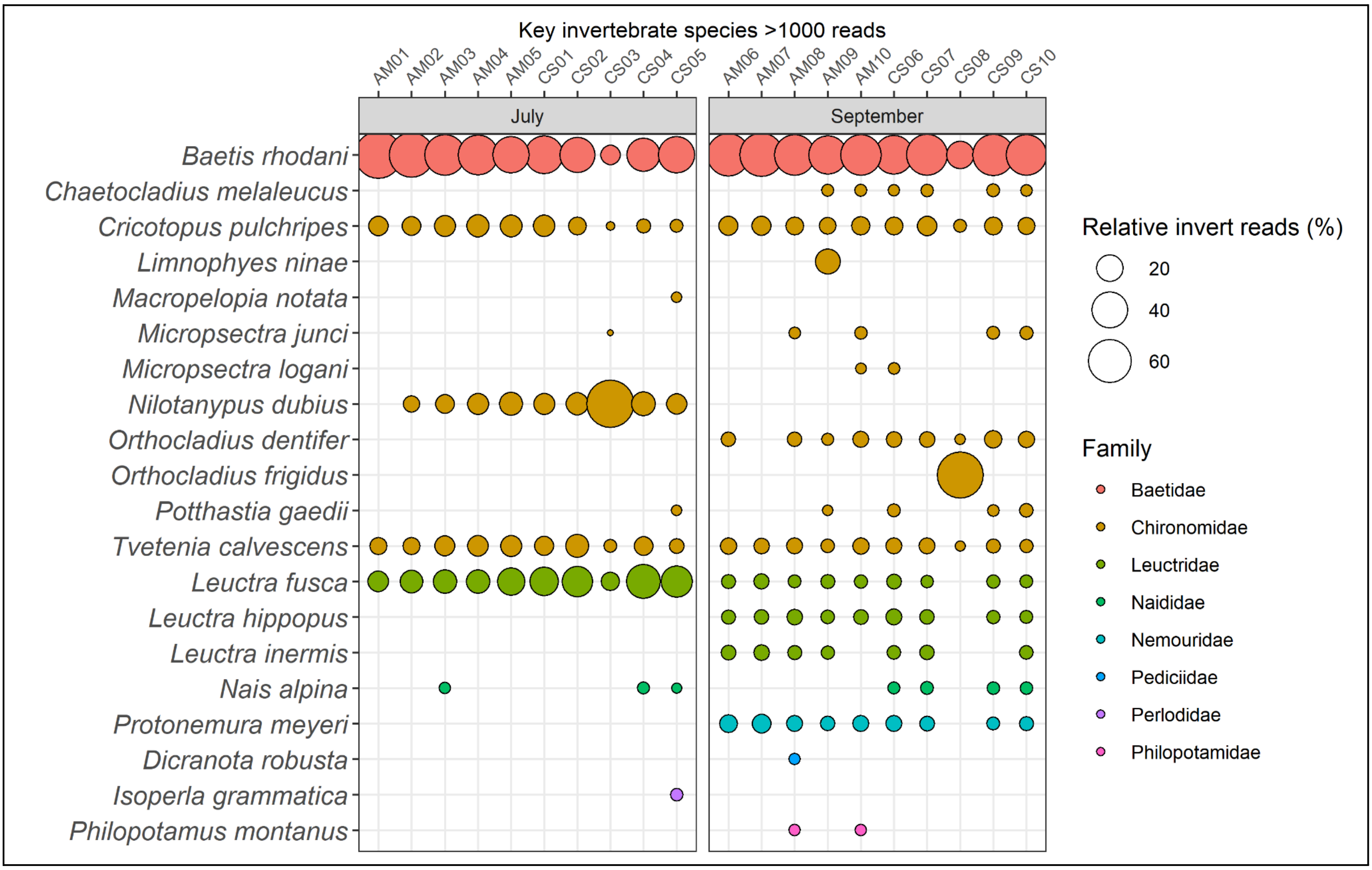

Figure 8. An overview of the invertebrate species detected (>1000 reads), and their relative reads.

References

Blackman, R.C., Benucci, M., Donnelly, R., Hänfling, B., Harper, L.R., Kimbell, H., Sellers, G.S., *et al.* (2018), *Targeting the Invaders – Targeted Detection of Four Priority Freshwater Invasive Non-Native Species Using Environmental DNA*, No. e27284v1, PeerJ Preprints, doi: 10.7287/peerj.preprints.27284v1.

Deiner, K., Fronhofer, E.A., Mächler, E., Walser, J.-C. and Altermatt, F. (2016), "Environmental DNA reveals that rivers are conveyer belts of biodiversity information", *Nature Communications*, Vol. 7, p. 12544.

Griffiths, N.P., Bolland, J.D., Wright, R.M., Murphy, L.A., Donnelly, R.K., Watson, H.V. and Hänfling, B. (2020), "Environmental DNA metabarcoding provides enhanced detection of the European eel Anguilla anguilla and fish community structure in pumped river catchments", *Journal of Fish Biology*, Blackwell Publishing Ltd Oxford, UK, Vol. 97 No. 5, pp. 1375–1384.

Handley, L.L., Read, D.S., Winfield, I.J., Kimbell, H., Johnson, H., Li, J., Hahn, C., *et al.* (2019), "Temporal and spatial variation in distribution of fish environmental DNA in England's largest lake", *Environmental DNA*, doi: 10.1002/edn3.5.

Hänfling, B., Lawson Handley, L., Read, D.S., Hahn, C., Li, J., Nichols, P., Blackman, R.C., *et al.* (2016), "Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods", *Molecular Ecology*, Vol. 25 No. 13, pp. 3101–3119.

Harper, L.R., Lawson Handley, L., Sayer, C.D., Read, D.S., Benucci, M., Blackman, R.C., Hill, M.J., *et al.* (2021), "Assessing the impact of the threatened crucian carp (Carassius carassius) on pond invertebrate diversity: A comparison of conventional and molecular tools", *Molecular Ecology*, Vol. 30 No. 13, pp. 3252–3269.

Kelly, R.P., Port, J.A., Yamahara, K.M. and Crowder, L.B. (2014), "Using environmental DNA to census marine fishes in a large mesocosm", *PloS One*, Vol. 9 No. 1, p. e86175.

Leese, F., Sander, M., Buchner, D., Elbrecht, V., Haase, P. and Zizka, V.M.A. (2021), "Improved freshwater macroinvertebrate detection from environmental DNA through minimized nontarget amplification", *Environmental DNA (Hoboken, N.J.)*, Wiley, Vol. 3 No. 1, pp. 261–276.

Li, J., Lawson Handley, L.J., Harper, L.R., Brys, R., Watson, H.V., Di Muri, C., Zhang, X., *et al.* (2019), "Limited dispersion and quick degradation of environmental DNA in fish ponds inferred by metabarcoding", *Environmental DNA*, Vol. 5, p. 133.

Pont, D., Rocle, M., Valentini, A., Civade, R., Jean, P., Maire, A., Roset, N., *et al.* (2018), "Environmental DNA reveals quantitative patterns of fish biodiversity in large rivers despite its downstream transportation", *Scientific Reports*, Vol. 8 No. 1, p. 10361.

Riaz, T., Shehzad, W., Viari, A., Pompanon, F., Taberlet, P. and Coissac, E. (2011), "ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis", *Nucleic Acids Research*, Vol. 39 No. 21, p. e145.

Sellers, G.S., Di Muri, C., Gómez, A. and Hänfling, B. (2018), "Mu-DNA: a modular universal DNA extraction method adaptable for a wide range of sample types", *Metabarcoding and Metagenomics*, Pensoft Publishers, Vol. 2, p. e24556.

Zhang, Z., Schwartz, S., Wagner, L. and Miller, W. (2000), "A greedy algorithm for aligning DNA sequences", *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology*, Vol. 7 No. 1-2, pp. 203–214.