

## An initial snapshot of the dipteran superfamily Chironomoidea at Holy Loch Nature Reserve, Argyll, Scotland, as revealed by automated DNA barcoding

N. Hammatt<sup>1</sup> & Wellcome Sanger Institute BIOSCAN Collective<sup>2</sup>

<sup>1</sup>Holy Loch Nature Reserve, Sandbank, Argyll PA23 8PD

<sup>2</sup>Wellcome Genome Campus, Hinxton, Cambridge CB10 1SA

<sup>1</sup>E-mail: [holylochnaturereserve@gmail.com](mailto:holylochnaturereserve@gmail.com)

### ABSTRACT

This study utilises the UK BIOSCAN molecular pipeline at the Wellcome Sanger Institute to establish the first high-resolution, DNA-based taxonomic baseline for the large dipteran superfamily Chironomoidea at Holy Loch Nature Reserve, Argyll, Scotland. During 2024, sampling was conducted via Malaise trapping in carr woodland and a supralittoral saltmarsh habitat mosaic. The methodology employs DNA barcoding, specifically targeting the cytochrome *c* oxidase subunit I (COI) mitochondrial gene. This genetic marker serves as a diagnostic tool, allowing for species-level identification by comparing DNA sequences/barcodes against a global reference library. The audit revealed at least 134 species, including (when compared with records on the Global Biodiversity Information Facility, GBIF) ten new species (or near species) for Scotland and/or the U.K. The study extended standard UK BIOSCAN procedures to larvae and to marine specimens. Despite representing only a snapshot of chironomid diversity at the Holy Loch, this study provides a foundational metadata anchor and reference library, essential for future biodiversity monitoring at the site.

### INTRODUCTION

The world is currently experiencing an accelerating rate of species extinctions driven by human economic activity (IPBES, 2019). Biodiversity monitoring is central to both global conservation efforts and small community-led projects such as that involving the undisturbed, protected area of the Holy Loch Nature Reserve (HLNR) on the Cowal Peninsula, Argyll, Scotland. Such monitoring requires a comprehensive baseline audit of all resident taxa (Blaxter *et al.*, 2005). Given the central role of insects in providing ecosystem services for human well-being (Scudder, 2017), in a time of finite research resources it is clear that novel methods are required that are both at scale (to be relevant) and automated (to reduce costs).

For over two centuries, traditional taxonomy has relied on specialists working methodically on individual groups to publish morphological identification guides and dichotomous keys. However, applying these methods to insects requires an extensive and costly investment in human expertise. Due to a chronic lack of

such a global strategic investment in taxonomists - the so-called "taxonomic impediment" (Engel *et al.*, 2021; Wheeler *et al.*, 2004) - humanity remains a long way from formally documenting every species via a Latin binomial, even at a single global site. Because many insect families remain difficult for non-specialists (and often experts) to identify during field-based surveys, compiling a comprehensive multi-class baseline at a community-led reserve like HLNR has historically been impossible. Consequently, many species may be lost to climate change and other environmental factors (Urban, 2024; Tedesco *et al.*, 2014) before they are ever described.

DNA barcoding (sequencing short DNA fragments from specific genes; Hebert *et al.*, 2003; Porter & Hajibabaei, 2018) offers a major step forward that enables rapid, species identification (Antil *et al.*, 2023). The Barcode of Life Database (BOLD), based in Guelph, Canada, maintains a growing library of DNA sequences from the cytochrome *c* oxidase subunit I (COI) mitochondrial gene, which have been deposited by taxonomists over recent decades. Additionally, specimens of insects and other organisms deposited in museum collections by taxonomists can also be sequenced and added to BOLD where they are clustered into statistically distinct sequence groups known as Biodiversity Index Numbers (BINs; Ratnasingham & Hebert, 2007, 2013). These BINs have a similarity threshold of  $\geq 98\%$  when they also tend to equate to species. Thus, even BINs without names, but which have unique DNA sequences, can be used as "proxy" species while taxonomic studies are carried out.

While earlier barcoding methods required the destruction of the specimen, the UK BIOSCAN project (coordinated by the Wellcome Sanger Institute, Hinxton, Cambridgeshire) utilises non-destructive DNA extraction methods (Park *et al.*, 2023). This is a critical advancement: where a barcode does not match an existing record on BOLD, the physical specimen remains intact, allowing taxonomists to conduct subsequent morphological studies to assign an existing name or describe a new species.

The UK BIOSCAN project aims to characterise the genetic diversity of flying insects from across the U.K.

(for an outline of the project, see Wellcome Sanger Institute (2026)). It employs robotic automation to generate high volumes of COI barcodes from invertebrates captured by standard Malaise traps deployed across multiple U.K. sites. Never before has such a vast dataset of barcodes been available for U.K. biodiversity and ecological research. This paper outlines a foundational method for monitoring taxonomic changes over time. The dipteran superfamily Chironomoidea (comprising four families) was selected for this study because available taxonomic resources for the group vary greatly: from the relatively recent identification keys for non-biting midges (Chironomidae) (Langton & Pinder, 2007), to web-based keys for biting midges (Ceratopogonidae), and somewhat dated resources for blackflies (Simuliidae) (Davies, 1966, 1968).

## METHODS

### Location

This study was carried out in 2024 and 2025 using insects trapped at the Holy Loch Nature Reserve (HLNR). This is located at the head of a sea loch, an inlet of the Firth of Clyde in Argyll, Scotland. This site has been relatively undisturbed for many decades in the ownership of Argyll and Bute Council and leased by Sandbank Community Development Trust. Light sheep grazing occurs in a small area of the saltmarsh from January to March annually. Otherwise, there is no other significant human activity.

### Field sampling and specimen acquisition

The aim of this project was to provide a first snapshot of chironomid diversity over the course of a year. Monthly invertebrate samples were obtained from standard two-metre Malaise traps deployed in two distinct ecological zones: (1) a complex, supralittoral habitat mosaic (Figs. 1 and 2; hereafter referred to as “marsh”), which comprises a beach with decomposing seaweeds and other organic debris, herbs (Fig. 3), gorse scrub and woodland on a sea-facing gravel ridge, *Festuca rubra*-dominated, herb-rich middle saltmarsh with water-containing depressions, and channels (Fig. 4), all within 40 m of the trap site; and (2) a stable carr woodland (Grid ref. 55.991205 N, -4.962823; what3words: soggy.fidgeting.squaring; Fig. 5) dominated by *Alnus glutinosa* (alder), *Betula pubescens* (downy birch) and *Salix caprea* (goat willow). There is evidence of an underground freshwater seep to the north of the marsh site which significantly alters the vegetation in this freshwater “lens”, and fills a sinuous channel. The marsh trap was deployed at the locations shown in Fig. 2.

Traps were deployed for 24 h concluding on the following dates in 2024: 27th January (carr only), 27th February, 31st March, 27th April, 22nd May, 30th June, 23rd July, 22nd August, 15th September, 21st October, and 26th November, and for seven days from 18th to 24th December. Calms days, ideally with sunshine, were chosen in order to sample during maximum invertebrate activity. The Malaise trapping bottle contained 100% ethanol. Individual specimens were then sorted (Fig. 6)

and distributed into the wells of 96-well microtitre plate which are arranged in an eight-row by twelve-column grid format, with alphanumeric identification for easy sample tracking (e.g. A1, H12). Prior to that, Darwin Tree of Life metadata were collected for each specimen (Lawniczak *et al.*, 2022).

Supplemental sampling was performed from 1st to 15th September 2024 to completely fill the most recent plate prior to dispatch. This involved manual collection of larvae from submerged timber within a saltmarsh drainage ditch (Grid ref. 55.961783 N, -4.961925 E; what3words: notch.january.similar) and the deployment for 15 days of a moth pheromone trap baited with a solution of cider vinegar and *Rubus fruticosus* agg. (bramble) fruits in a hedge at the back of the saltmarsh (Grid ref: 55.963345 N, -4.961988 E; what3words: animator.lame.deputy). To prevent possible polymerase chain reaction (PCR: see below) inhibition by saline residues or acetic acid from the vinegar, all specimens retrieved from the saltmarsh or baited trap were washed in a phosphate buffer (pH 7.0) prior to preservation in 100% ethanol in the 96-well plates. Plates were stored in a freezer at at least -20°C prior to dispatch to the Wellcome Sanger Institute by courier.

### Molecular pipeline and DNA barcoding

Specimens were processed through the UK BIOSCAN high-throughput pipeline at the Wellcome Sanger Institute as described by Park *et al.* (2023). Following evaporation of the ethanol from each well, specimens were submerged in a lysis buffer, which allowed seepage of DNA-containing body fluids through the insects' exoskeleton and preserved the structure of the "voucher" specimen for any subsequent taxonomic examination. The collection is curated at the Natural History Museum, London.

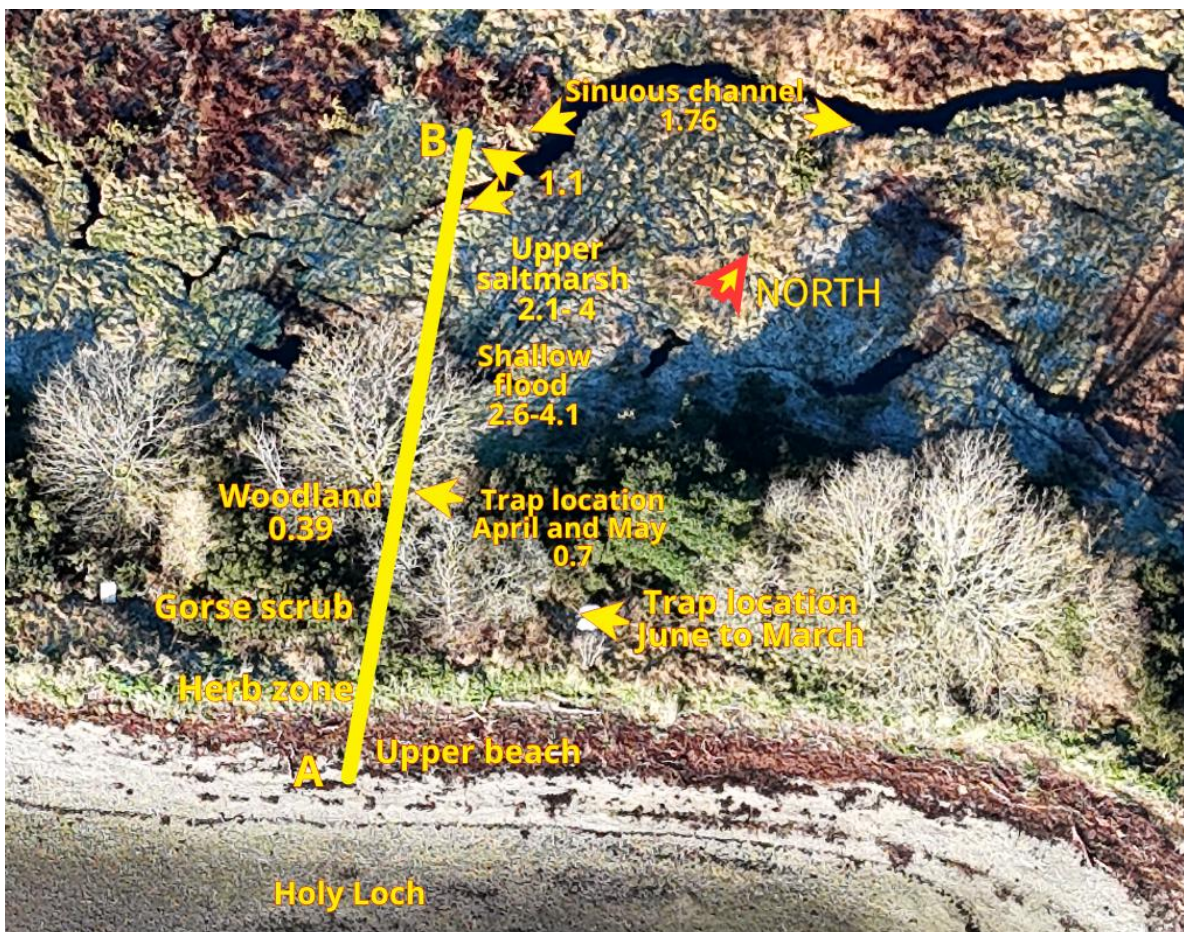
The PCR technique was then employed to produce multiple copies of a *ca.* 658 base pair fragment of the mitochondrial cytochrome *c* oxidase subunit I (COI) gene present in the DNA lysates (Mullis *et al.*, 1986). The amplified DNA strands were sequenced using the Illumina NovaSeq platform. The resulting raw sequence data were processed through a bioinformatics pipeline (mBRAVE; Ratnasingham, 2019) to isolate high-quality barcodes, with around 658 nucleotides, which were then uploaded to the Barcode of Life Data System (BOLD) under project HLNR.

### Taxonomic assignment and the BIN system

Sequences were assigned to Biodiversity Index Numbers (BINs) using the BOLD Refined Single Linkage (RESL) algorithm (Ratnasingham & Hebert, 2013). This system clusters sequences into statistically distinct operational taxonomic units (OTUs) which serve as high-resolution species proxies. BIN sequences are then compared with named sequences in BOLD's growing reference library.



**Fig. 1.** Aerial photograph of the “marsh” sampling site at Holy Loch Nature Reserve, Sandbank, Argyll, Scotland. (Photo: Phill Rawlins)



**Fig. 2.** Closer view of the boxed area in Fig. 1. A-B indicates the location of the 40 m linear transect. A: 55.9945 N, -4.9583; what3words: standard.sunflower.beginning. B: 55.9941 N, -4.9580; what3words: phantom.overgrown.junction. The values are soil conductivity measurements ( $\text{mS cm}^{-1}$ ). (Photo: Phill Rawlins)



**Fig. 3.** The sinuous channel labelled in Fig. 2, located in flower-rich grassland (June 2024). (Photo: Neil Hammatt)



**Fig. 4.** A malaise trap at the June to March site (see Fig. 2). (Photo: Neil Hammatt)



**Fig. 5.** General view of the location of the carr woodland malaise trap, showing the dense, herbaceous understory. (Photo: Neil Hammatt)



**Fig. 6.** The “marsh” catch from June 2024. (Photo: Neil Hammatt)

At HLNR, taxonomic assignment of these BINs then followed a four-tier taxonomic confidence hierarchy (Table 1), depending on the amount of DNA and other supporting evidence.

In addition to the standard BOLD identification engine, all sequences that returned only a genus-level assignment were checked manually using BLAST (Basic Alignment Research Tool) searches on BOLD (<https://id.boldsystems.org>) and/or NCBI

([https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE\\_TYPE=BlastSearch&LINK\\_LOC=blast\\_home](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blast_home)). This step is necessary because the automatic BOLD pipeline is intentionally conservative and may stop at genus even when multiple high-quality species-level matches exist. Manual BLAST bypasses BIN-based restrictions and compares the sequence directly with all public reference records, allowing clear species-level identifications to be recognised when the evidence is unambiguous. Because BOLD is based on arthropod COI sequences and GenBank used by BLAST can contain any sequence from any taxon, results from BOLD were given more credence, but if BLAST added evidence, then that was also taken into consideration.

#### **Primary data source**

All sequences were downloaded to the publicly-available database on the Bioscan Report Card. This paper allots these sequences to the various trapping sites.

#### **Abiotic Data Integration**

Substrate pH and electrical conductivity (a proxy for total salinity) were recorded on 3rd November 2025, to provide a quantitative understanding of habitat heterogeneity at the two trapping sites within a relatively short flying distance between emerged adults and the traps. At the woodland site, four 5 m cardinal transects were established for substrate pH testing. Additionally, a 40 m longitudinal transect was established from the beach through successive “marsh” vegetation zones (Fig. 2). Measurements were taken at 1 m intervals along all transects using calibrated Hanna Instruments soil probes. These parameters are essential for mapping a single snapshot of physiological constraints of resident larvae, particularly in the dynamic saline conditions of the saltmarsh caused by spring tides of varying heights.

Considered evidence	Holy Loch Nature Reserve BIN taxonomic category			
	Confirmed identity	Rare biosphere	Emerging identity	Barcode silence
Molecular Sequence similarity to BOLD references	Close to 100% match to BOLD references; unambiguous species-level assignment	≥98% match to BOLD references; usually species-level, may require manual verification on BOLD or	Genus-level match only; multiple species in BOLD may exist; no confident species-level	No reliable match in BOLD; single or very few BLAST hits only; assignment not possible beyond
Number of reference sequences	≥3	<3	Any	No limit
Geographical range of reference BINs	≥2 countries	≥2 countries	Any	Any
Collector diversity for reference BINs	≥2 independent collectors	≥2 independent collectors	Any	Any
Longevity of BINs	>5 years	Can be newly formed	Can be newly formed	Can be newly formed
Recurrence (number of sequences per BIN at HLNR)	≥2 months	At least one month	Any	0 or very few (BLAST-only)
Ecological plausibility (if niche known)	Supporting evidence	Supporting evidence		
Entry to the UK Species Inventory (UKSI)	Supporting evidence	Supporting evidence		

**Table 1.** Confidence criteria used for BIN taxonomic conclusions at Holy Loch Nature Reserve, Sandbank, Argyll, Scotland.

## RESULTS

### Abiotic data

The substrate pH in the carr woodland remained relatively uniform throughout the study period, with values recorded between 6.6 and 6.9 (Table 2). In contrast, the saltmarsh habitat mosaic displayed a significant environmental gradient; pH values shifted from highly acidic (4.5) within the mature woodland atop the gravel ridge to 6.8 in the upper saltmarsh. Electrical conductivity on the saltmarsh largely reflected elevation, which determines the frequency and duration of tidal inundation during spring tides (Fig. 2).

### Barcoding

Chironomoidea comprised the greatest number of individuals and BINs recorded at HLNR in 2024. Within the family Chironomidae, the genus *Limnophyes* was the most abundant, producing 842 barcodes, of which 348 were assigned to *L. natalensis*. Successful naming of BINs depended on the quality and completeness of the reference library; in several cases, such as BOLD:ACF9576, BINs may contain more than one species, specifically *Culicoides obsoletus* or *C. scoticus*, where morphological differences are not reflected in a 2% DNA sequence divergence. Conversely, some species, such as *Dasyhelea turficola*, spanned multiple BINs. This is a reminder that intraspecific genetic variation can be larger than the BIN algorithm can accommodate.

Within the Chironomidae, 23 BINs were assigned to “Confirmed identity” (Table 3), while thirty-two BINs occurred only once or a few times and were categorised as “Rare biosphere” (Table 4). Among 25 BINs with “Emerging identities” assignable only to genus (Table 5), manual BLAST analysis enabled the assignment of up to three possible species. Two relatively frequent BINs assigned to “Barcode silence” were identifiable only to the family level (Table 6). For Ceratopogonidae, at least seven species were identified

confidently (Table 7), and fifteen rare BINs also had solid identifications. Among the 12 with emerging identities, BLAST analysis revealed that several could not be deciphered between two near-identical species.

Distance (m) from trap location 2	pH	Conductivity (mS cm <sup>-1</sup> )	Habitat
-18	6.6		Beach
-16	6		<b>Trap location 1</b>
-11	6.5		Scrub ecotone
-8	4.4	0.39	Woodland
0	6.9	0.23	<b>Trap location 2</b>
1	6.6	0.47	Downslope scrub
2	6.5	0.7	Downslope scrub
3	6.6	0.93	Upper saltmarsh
4	6.8	1.48	Upper saltmarsh
5	6	1.54	Shallow flood edge
6	6.1	4	Shallow flood
7	6.5	2.6	Shallow flood
8	6.5	2.6	Shallow flood
9	5.3	2.6	Shallow flood
10	5.7	3.9	Shallow flood edge
11	6.1	4	Upper saltmarsh
12	6.3	2.8	Upper saltmarsh
13	6.2	2.6	Upper saltmarsh
14	6.5	3.5	Upper saltmarsh
15	6.1	2.4	Upper saltmarsh
16	6.6	2.1	Upper saltmarsh
17	6	2.7	Channel edge
18	5.7		Rushy shallow edge
19	6.4	1.76	Sinuuous channel
20	5.8	2.6	Sinuuous channel edge
21	6	3.5	Upper saltmarsh
22	6.7	1	Upper saltmarsh

**Table 2.** Variation in pH and soil conductivity and habitat description along a 40 m linear transect on vegetated gravel at Holy Loch Nature Reserve, Sandbank, Argyll, Scotland, on 2nd November 2025 (see Fig. 2). Both trap locations were in gorse and herb-rich vegetated gravel.

BIN	Assigned name	Wood		Marsh	
		Total	Months	Total	Months
BOLD:AAD7724	<i>Brillia bifida</i>	3	Feb, Dec	2	Jan
BOLD:ACF6903	<i>Chaetocladius perennis</i>	2	Feb	6	Jan, Jun, Aug, Sep
BOLD:AAF3277	<i>Chironomus apralinus</i>	1	Aug	13	Apr, May
BOLD:AAN5317	<i>Chironomus salinarius</i>			15	Jun, Jul, Aug
BOLD:ACD4501	<i>Gymnometriocnemus brumalis</i>	5	Sep, Oct	14	Sep, Oct
BOLD:AAB8751	<i>Halocladius variabilis</i>			21	Mar, May, Jun, Jul, Aug
BOLD:AEF6833	<i>Halocladius varians</i>			6	Jul, Oct
BOLD:AAB9700	<i>Limnophyes edwardsi</i>	1	Dec	40	Mar, Apr, Jun, Aug
BOLD:AAJ2130	<i>Limnophyes habilis</i>	2	Feb, Jun	11	Jun, Jul, Aug, Sep
BOLD:AAB7359	<i>Limnophyes natalensis</i>	1	Jul	351	Apr, May, Jun, Jul, Aug
BOLD:AAE6392	<i>Limnophyes pentaplastus</i>	7	Feb, Mar, Jun, Dec	12	Jan, Jun, Jul, Aug
BOLD:ABU5525	<i>Limnophyes</i> sp. 14ES-unnamed	9	Feb, Mar, Jun, Jul, Dec	17	Jan, Apr, Jun, Aug, Dec
BOLD:ABW2734	<i>Limnophyes</i> sp. 14ES-unnamed	1	Jul	11	Jan, Feb, Apr, May, Jun, Jul, Dec
BOLD:AAB1806	<i>Limnophyes</i> sp. 3ES-unnamed	17	Aug, Nov, Dec	179	Jun, Jul, Aug, Oct, Nov, Dec
BOLD:AAC6643	<i>Metriocnemus albolineatus</i>			5	Jun, Aug
BOLD:AAI1573	<i>Metriocnemus fuscipes</i>	3	Feb, Oct	27	Feb, May, Jun, Jul, Aug
BOLD:AAI1572	<i>Metriocnemus picipes</i>	2	Apr, Aug	25	Apr, Jun, Aug, Sep
BOLD:AAC4200	<i>Paraphaenocladius impensus</i>	4	Jun, Oct, Dec	2	Apr
BOLD:ACN9514	<i>Paraphaenocladius impensus</i>			21	Mar, Apr, Jun, Jul, Aug
BOLD:ACG6805	<i>Pseudorthocladius</i> cf. <i>curtistylus</i>	3	Jun, Jul	10	Jun, Jul, Aug
BOLD:AAD9251	<i>Pseudorthocladius filiformis</i>			5	Aug
BOLD:ACP4407	<i>Pseudosmittia trilobata</i>			5	Apr, May
BOLD:ABW7981	<i>Smittia leucopogon</i>			5	Jun, Aug

**Table 3.** BOLD BINs with “Confirmed identity” from the family Chironomidae captured at the Holy Loch Nature Reserve, Sandbank, Argyll, Scotland, in 2024 from carr woodland (Wood) and the saltmarsh-based habitat mosaic (Marsh). This includes the total number caught as adults in 2024 and the months in which they were trapped.

Specific attention was given to the Simuliidae, where three larvae collected on 15th September from the saltmarsh field drain were assigned to two *Simulium* BINs, while a total of 12 adults collected in 2024 were allocated to three BINs (Table 8). Notably, BOLD:ACX9864 was captured both as a larva in September and as an adult on October 21st.

Comparison of the HLNLR species found by UK BIOSCAN with the GBIF (Global Biodiversity Information Facility) distribution maps for Scotland and the U.K. on 9th December 2025, revealed many species not previously recorded in these jurisdictions. In Chironomidae, these include *Limnophyes pentaplastus*, *Pseudosmittia trilobata*, *Pseudorthocladius* cf. *curtistylus*, *Micropsectra apposita*, *M. klinki*, *Tanytarsus ejuncidus*, and *Xenopelopia nigricans*. In Ceratopogonidae, *Culicoides comosioculatus* appears to be new to the U.K., alongside *Dasyhelea arenivaga*, *Forcipomyia monilicornis*, and *Atrichopogon* cf. *hirtidorsum*.

## DISCUSSION

In view of the trapping methodology employed – the use of a Malaise trap for short (mainly 24 h) monthly periods over the course of one year – the data generated in this investigation can provide only a snapshot of the chironomid assemblage at HLNLR. Also, taxonomic interpretations may be compromised by the fact that there are quite a few chironomid BINs with unclear or provisional taxonomy in BOLD, due to lack of revision or undescribed species. Despite these caveats, the

molecular biodiversity audit presented here represents a transformative frontier for biological recording in the West of Scotland. Given the coastal setting and the diversity of aquatic and semi-aquatic environments present, including flooded carr, saltmarsh pans, brackish creeks, freshwater burns, and freshwater seeps, it is unsurprising that Chironomoidea, particularly Chironomidae (Armitage *et al.*, 1995), comprised the greatest number of individuals and BINs recorded at HLNLR. DNA barcoding has similarly revealed substantial hidden chironomid diversity in other regions (Namayandeh *et al.*, 2024), highlighting the value of genomic approaches in complex wetland systems.

Substrate pH (Carstensen & Duarte, 2019) and salinity are the principal environmental factors structuring saltmarsh invertebrate assemblages. Tolerance to these variables, particularly soil salinity, determines the viability of species whose life cycles depend on wet substrates and decomposing organic matter. Both trap sites at HLNLR were acidic: saltmarsh pH ranged from 4.5 to 6.8, whereas the carr woodland showed relatively uniform pH values, reflecting its homogeneous vegetation structure. As expected, conductivity on the saltmarsh largely reflected elevation, which determines the frequency and duration of tidal inundation during spring tides. These hydrological gradients influence plant community composition (Hammatt, unpublished data) and, consequently, the types of detrital resources available to invertebrates.

BIN	Assigned name	Wood		Marsh	
		Total	Months	Total	Months
BOLD:AAW1344	<i>Bryophaenocladus flavoscutellatus</i>	1	Jul	1	Jul
BOLD:AER6168	<i>Bryophaenocladus flexidens</i>			2	Oct
BOLD:AAM6273	<i>Bryophaenocladus ictericus</i>	1	Nov	1	Aug
BOLD:AAD4034	<i>Chaetocladus melaleucus</i>	1	Nov		
BOLD:AAJ4282	<i>Chironomus luridus</i>	1	Nov		
BOLD:AAD1162	<i>Corynoneura lobata</i>			1	Aug
BOLD:ACG3525	<i>Corynoneura</i> sp. 24ES			2	Jun
BOLD:AAP5927	<i>Cricotopus pulchripes</i>			1	Jun
BOLD:AAE4568	<i>Eukiefferiella claripennis</i>			1	Mar
BOLD:AAX3566	<i>Macropelopia nebulosa</i> **			1**	Sep
BOLD:AAB8862	<i>Metriocnemus eurynotus</i>			1	Jun
BOLD:ACB9910	<i>Metriocnemus</i> sp. 1SW-unnamed			1	Aug
BOLD:AEJ0479	<i>Micropsectra apposita</i>	1	Feb		
BOLD:AAW0057	<i>Micropsectra klinki</i>	1	Feb	1	Jun
BOLD:AAD1527	<i>Micropsectra lindrothi</i>			1	Aug
BOLD:AEE6875	<i>Micropsectra pallidula</i>	1	Dec		
BOLD:AAC4200	<i>Paraphaenocladus impensus</i>			1	Mar
BOLD:ACF8295	<i>Paraphaenocladus pseudirritus</i>			1	Apr
BOLD:ACM5899	<i>Paratanytarsus austriacus</i>	1	Jul		
BOLD:AAF0341	<i>Polypedilum uncinatum</i>	1	Apr		
BOLD:AAG6458	<i>Pseudosmittia albipennis</i>	1	May	1	Jun
BOLD:AAD5902	<i>Rheocricotopus atripes</i>			1	Jan
BOLD:AEE5575	<i>Rheocricotopus</i> sp. 3ES			1***	Aug
BOLD:ACE0351	<i>Smittia</i> cf. <i>stercoraria</i>	1	Dec		
BOLD:AGO4959	<i>Smittia</i> cf. <i>stercoraria</i>			2	Aug, Dec
BOLD:ACK3746	<i>Smittia paranudipennis</i>			1	Jun
BOLD:ACM5388	<i>Smittia pratorum</i>			3	Jun, Oct
BOLD:AAP5616	<i>Tanytarsus brundini</i>			1	Jun
BOLD:AAW1102	<i>Tanytarsus ejuncidus</i>			1	Jun
BOLD:ACG9673	<i>Tanytarsus signatus</i>			1	Aug
BOLD:AAD5253	<i>Thienemanniella xena</i> *			1	Aug
BOLD:ADA6313	<i>Xenopelopia nigricans</i>	1	Apr		
BOLD:AEV9540	<i>Limnophyes asquamatus</i>	2	Apr		

**Table 4.** BOLD BINs from the “Rare biosphere” of the family Chironomidae captured at the Holy Loch Nature Reserve, Sandbank, Argyll, Scotland, in 2024 from carr woodland (Wood) and the saltmarsh-based habitat mosaic (Marsh). This includes the total number caught as adults (and a single larva) in 2024 and the months in which they were trapped. \*Identified by manual BLAST and the BOLD identification engine. \*\*Larva from a seaweed pile on a beach. \*\*\*Swept.

BIN	Genus	BCY	Wood		Marsh	
			Total	Months	Total	Months
BOLD:ACR0319	<i>Bryophaenocladus</i>		1	Oct	1	Oct
BOLD:AAI7060	<i>Chaetocladus</i>				3	Dec
BOLD:AEQ8805	<i>Chaetocladus</i>				1	Jan
BOLD:ACD9509	<i>Georthocladus</i>				1	Mar
BOLD:AAI4193	<i>Gymnometriocnemus</i>				1	Apr
BOLD:AAU6603	<i>Hydrosmittia</i>				1	Aug
BOLD:AGL0475	<i>Krenopelopia</i> (nr. <i>binotata</i> )*	2025			1	Jun
BOLD:AGO4493	<i>Krenopelopia</i> (nr. <i>binotata</i> )*	2025			1	Aug
BOLD:AAI2213	<i>Krenopelopia</i> (nr. <i>binotata</i> )*	2010			1	Aug
BOLD:ACO9143	<i>Limnophyes</i>				7	Feb, Jun, Aug
BOLD:ACR0830	<i>Limnophyes</i>		13	Dec	7	Feb, Mar, Apr, Jul
BOLD:ACT5780	<i>Limnophyes</i>				2	Apr, Oct
BOLD:AGX5543	<i>Limnophyes</i>				116	Apr, May, Jun, Aug, Sep
BOLD:AAX4301	<i>Metriocnemus tristellus</i> **	2011			2	Apr, Oct
BOLD:ACC1758	<i>Metriocnemus</i>	2012			1	Aug
BOLD:AEC4629	<i>Procladius</i>				1	Aug
BOLD:ACT1181	<i>Procladius</i>				1	Aug
BOLD:ACF9730	<i>Pseudorthocladus</i>				1	Apr
BOLD:AAM6263	<i>Pseudosmittia</i>		2	May		
BOLD:ACG7714	<i>Pseudosmittia</i>				4	Jul
BOLD:ADM2886	<i>Smittia</i>				1	Jun
BOLD:AEH1919	<i>Smittia</i>		1	May		
BOLD:ACF7218	<i>Smittia</i> ***				14	Jun, Aug
BOLD:ACU9413	<i>Smittia</i>		2	May, Jun		
BOLD:ACQ9002	<i>Thienemanniella</i> (nr. <i>vittata</i> )****				1	Jun

**Table 5.** BOLD BINs with “Emerging identity” from the family Chironomidae captured at the Holy Loch Nature Reserve, Sandbank, Argyll, Scotland, in 2024 from carr woodland (Wood) and the saltmarsh-based habitat mosaic (Marsh). This includes the total number caught as adults in 2024 and the months in which they were trapped. BCY = BIN creation year. \*Each BIN has a few matches to sequences from *Krenopelopia binotata* in a BOLD manual algorithm run. These are all over 97% match. \*\*Single match to *Metriocnemus tristellus* in a BLAST search. \*\*\*Single match on BLAST. \*\*\*\*Matched on BLAST. Nineteen barcodes were assigned to *Limnophyes* without first including them in a BIN.

BIN	Date BIN created	No. of specimens
BOLD:ACY5749	06/05/2019	4
BOLD:AEE3886	21/09/2020	2
BOLD:AEH2536	19/02/2021	27
BOLD:AGK6424	07/01/2025	1
BOLD:AGN3217	07/01/2025	13
BOLD:AGM4472	07/01/2025	3
BOLD:AGO3668	07/01/2025	2

**Table 6.** BOLD BINs with “Barcode silence” from the family Chironomidae captured at the Holy Loch Nature Reserve, Sandbank, Argyll, Scotland in 2024.

With basic descriptions of the trapping locations, habitat composition, and substrate variables, this paper serves as a metadata anchor for all subsequent UK BIOSCAN publications from HLNLR. Given the potential number of distinct physiological niches, at a time when the concept of ecological niche itself is being critically re-examined (Ferrante, 2025), the substantially higher abundance of individuals and BINs on the saltmarsh compared with the structurally more uniform carr woodland is expected (Buche *et al.*, 2022). Both pH and salinity also influence dissolved oxygen concentrations, another fundamental determinant of invertebrate physiology (Banerjee *et al.*, 2019). Systematic measurements of dissolved oxygen across HLNLR are scheduled to begin shortly to further refine these niche models.

At present, understanding of the specific, and potentially flexible, niches occupied by individual chironomid taxa remains limited. Consequently, devising targeted conservation measures for declining species, or even detecting climate-driven declines against a background of naturally variable population dynamics, remains challenging. However, the statistically robust volumes of BIN occurrence data now generated by UK BIOSCAN enable ecological modelling based on identifiable taxa rather than hypothetical species. Such models can incorporate variables such as habitat preference and flight period. In the present study, for example, *Limnophyes natalensis* was almost entirely restricted to the saltmarsh between April and August, whereas *Limnophyes* sp. 3ES was recorded more evenly throughout the year. While some taxa may be woodland specialists, it is notable that most BINs were recorded at both sites, suggesting a degree of landscape-scale connectivity.

Where chironomid niches have been investigated in detail, complex ecological relationships have emerged. These include the dependence of *Halocladus variabilis*, recorded at HLNLR, on seaweeds (Tarakhovskaya & Garbary, 2009). In the context of projected climate change for western Scotland (Werrity & Sugden, 2012), it is notable that within the genus *Brillia*, temperature and moisture gradients are reported as key drivers of population divergence, with evidence of adaptive differentiation among species (Xu *et al.*, 2025). Food type has also been identified as a determinant of presence or absence in *B. bifida* (García & Pardo, 2015). Research on Ceratopogonidae has historically focused on their nuisance value to humans, particularly the

Highland midge, *Culicoides impunctatus*, and their role as vectors of pathogens such as blue tongue virus (Blackwell *et al.*, 1992; Mullens *et al.*, 2015; King *et al.*, 2025).

By establishing a high-resolution genomic baseline at HLNLR, investigations such as this bypass the “taxonomic impediment” that has historically stalled community-led conservation efforts (Engel *et al.*, 2021; Wheeler *et al.*, 2024). For over two centuries, the requirement for hyper-specialised morphological expertise has acted as a barrier, preventing small sites, such as Local Nature Reserves, from fully documenting their taxa. DNA barcoding democratises this process, allowing non-specialists to generate data of a calibre that matches or exceeds that of institutional surveys (Roslin *et al.*, 2022).

The identification of at least ten species new to the Scottish and/or U.K. lists – including *Culicoides comosoculatus*, *Limnophyes pentaplastus*, and *Xenopelopia nigricans* – signals that a significant volume of “hidden” biodiversity has been overlooked. Were any present, traditional morphological keys, which rely heavily on male genitalia (Langton & Pinder, 2007), could not resolve endemic genetic lineages. The use of non-destructive DNA extraction by the Sanger Institute is a critical technological advance (Park *et al.*, 2023); it ensures that the physical “voucher” remains intact, allowing specialists to revisit the exact specimens that produced a “new to science” barcode. This creates a vital bridge between 19th century morphological traditions and 21st century genomics.

This study validates HLNLR as a vital site for long-term monitoring and provides the “genetic library” required for future environmental DNA (eDNA) metabarcoding. With the future integration of barcodes collected in 2025, 2026, and beyond, the resolution of our ecological map will continuously increase. This multi-year data stream will progress beyond simple inventorying into the analysis of seasonal shifts and population dynamics.

Ultimately, the phenological and distributional data captured in this inaugural year will serve as the permanent control against which all future ecological shifts will be measured.

#### ACKNOWLEDGEMENTS

This research was supported by Seachangers, William Grant Foundation, Flora and Fauna International, and the people of Cowal. We are grateful to Phill Rawlins for aerial photography and to two reviewers for their detailed feedback. COI barcode data used here were generated by the BIOSCAN project led by Mara Lawniczak’s team at the Wellcome Sanger Institute. The BIOSCAN project and many of the people named in the Wellcome Sanger Institute BIOSCAN Collective author list are funded by the Wellcome Trust award 220540/Z/20/A.

BIN	Assigned name	Wood		Marsh	
		Total	Months	Total	Months
<b>Confirmed identity</b>					
BOLD:AAV9748	<i>Culicoides circumscriptus</i>			4	Jun
BOLD:ACG4335	<i>Culicoides impunctatus</i>	1	Jun	23	Jun, Jul
BOLD:ACX2603	<i>Culicoides punctatus</i>			11	Jun, Jul
BOLD:AAV9630	<i>Culicoides segnis</i>	3	May, Jun		
BOLD:ACX2767	<i>Dasyhelea arenivaga</i>			14	Jun, Jul, Aug
BOLD:AEH4464	<i>Dasyhelea turficola</i>			11	Jun, Jul, Aug
BOLD:ABW3987	<i>Forcipomyia tenuis</i>			6	Jun, Jul, Aug
<b>Rare biosphere</b>					
BOLD:ACP4327	<i>Ceratopogon grandiforceps</i>			1	Apr
BOLD:ADF6075	<i>Culicoides comosioculatus</i>			1	Aug
BOLD:AEC6694	<i>Culicoides newsteadi</i>			2	Jul, Aug
BOLD:AGT3278	<i>Dasyhelea modesta</i>			1	Jun
BOLD:AAV5098	<i>Dasyhelea turficola</i>			1	Jul
BOLD:AEH4801	<i>Dasyhelea turficola</i>			1	Jun
BOLD:AGR6137	<i>Dasyhelea turficola</i>			1	Jun
BOLD:AAN5148	<i>Forcipomyia bipunctata</i>			1	Jun
BOLD:ABW3942	<i>Forcipomyia nigra</i>	1	Jun		
BOLD:AEA8665	<i>Palpomyia flavipes</i>	1	Aug		
BOLD:ACX1684	<i>Palpomyia serripes</i> (using manual BLAST)			1	Aug
BOLD:ACL4093	<i>Schizohela leucopeza</i>			1	Jun
BOLD:ACY6482	<i>Serromyia femorata</i>			1	Jun
BOLD:ACS7502	<i>Serromyia ledicola</i>			1	Jul
BOLD:ACG4212	<i>Stilobezzia ochracea</i>	1	Jul		
<b>Emerging identity</b>					
BOLD:AGL8522	<i>Atrichopogon</i>			5	May, Jun
BOLD:ABZ8851	<i>Atrichopogon</i> cf. <i>hirtidorsum</i>			2	Jul, Aug
BOLD:ACP8436	<i>Bezzia ornata</i>			1	Jun
BOLD:ACF9576	<i>Culicoides obsoletus</i> or <i>C. scoticus</i> *			1	Jun
BOLD:ACF7867	<i>Culicoides obsoletus</i> or <i>C. scoticus</i> *	2	Jun		
BOLD:ACS7831	<i>Dasyhelea</i> sp.			14	Jun, Jul, Aug
BOLD:ACX1316	<i>Dasyhelea</i> sp.			1	Jul
BOLD:ACN2557	<i>Forcipomyia monilicornis</i>			1	Jun
BOLD:ABW3952	<i>Forcipomyia sphagnophila</i> or <i>F. nigrans</i> *	4	Mar, Jun	3	Jul
BOLD:ABW3964	<i>Forcipomyia sphagnophila</i> or <i>F. nigrans</i> *			8	Jun, Jul, Aug
BOLD:ACX3286	<i>Palpomyia spinipes</i>	1	Aug		
<b>Barcode silence</b>					
BOLD:AFM9892	Ceratopogonidae	1	Aug	2	May, Jul
BOLD:AGN3401	Ceratopogonidae			1	Jun
BOLD:AGR1629	Ceratopogonidae			1	May
BOLD:AEH2536	Ceratopogonidae			10	Jun
BOLD:ADM5192	Ceratopogonidae			1	Jul
BOLD:AGL8523	Forcipomyiinae			1	May

**Table 7.** BOLD BINs from the family Ceratopogonidae captured at the Holy Loch Nature Reserve, Sandbank, Argyll, Scotland, in 2024 from carr woodland (Wood) and the saltmarsh-based habitat mosaic (Marsh), and assigned to four taxonomy confidence categories. This includes the total number caught as adults in 2024 and the months in which they were trapped. \*These BINs had maximum similarity with at least two species. Identified by manual BLAST and BOLD.

BIN	Location	Catch details
BOLD:AAB8624	Saltmarsh field drain	2 larvae on 15th Sep
BOLD:ACX9864	Saltmarsh field drain	1 larva on 15th Sep
BOLD:ACX9864	Marsh	1 or 2 adults in Jan, Feb, Mar, Jun, Jul, Aug
BOLD:AFJ9886	Wood	1 adult in Mar
BOLD:AGO5113	Wood	1 adult in Jun

**Table 8.** Mixed category BINs from adult and larval *Simulium* species from carr woodland (Wood) and the saltmarsh-based habitat mosaic (Marsh) at Holy Loch Nature Reserve, Sandbank, Argyll, Scotland, in 2024, their trapping locations and dates.

## REFERENCES

- Antil S., Abraham, J.S., Sripoorna, S., Maurya, S., Dagar, J., Makhija, S. *et al.* (2023). DNA barcoding: an effective tool for species identification. *Molecular Biology Reports* 50, 761–775. <https://doi.org/10.1007/s11033-022-08015-7>
- Armitage, P., Cranston, P.S. & Pinder, L.C.V. (1995). *The Chironomidae: Biology and Ecology of Non-Biting Midges*. Springer, Berlin. <https://doi.org/10.1007/978-94-011-0715-0>
- Banerjee, A., Chakrabarty, M., Rakshit, N., Bhowmick, A.R. & Ray, S. (2019). Environmental factors as indicators of dissolved oxygen concentration and zooplankton abundance: deep learning versus traditional regression approach. *Ecological Indicators* 100, 99–117. <https://doi.org/10.1016/j.ecolind.2018.09.051>
- Blackwell, A., Mordue, A.J., Young, M.R. & Mordue, W. (2009). Bivoltinism, survival rates and reproductive characteristics of the Scottish biting midge, *Culicoides impunctatus* (Diptera: Ceratopogonidae) in Scotland. *Bulletin of Entomological Research* 82, 299–306. <https://doi.org/10.1017/S0007485300041055>
- 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* 360, 1935–1943. <https://doi.org/10.1098/rstb.2005.1725>
- Buche, L., Spaak, J.W., Jarillo, J. & De Laender, F. (2022). Niche differences, not fitness differences, explain predicted coexistence across ecological groups. *Journal of Ecology* 110, 2785–2796. <https://doi.org/10.1111/1365-2745.13992>
- Carstensen, J. & Duarte, C.M. (2019). Drivers of pH variability in coastal ecosystems. *Environmental Science & Technology* 53, 4020–4029. <https://doi.org/10.1021/acs.est.8b03655>
- Davies, L. (1966). The taxonomy of British blackflies (Diptera: Simuliidae). *Transactions of the Royal Entomological Society* 118, 413–506. <https://doi.org/10.1111/j.1365-2311.1966.tb00832.x>
- Davies, L. (1968). *A Key to the British Species of Simuliidae (Diptera) in the Larval, Pupal and Adult Stages*. Scientific Publication No 24. Freshwater Biological Association, Milnthorpe.
- Engel, M.S., Ceriaco, L.M.P., Daniel, G.M., Dellapé, P.M., Löbl, I., Marinov, M. *et al.* (2021). The taxonomic impediment: a shortage of taxonomists, not the lack of technical approaches. *Zoological Journal of the Linnean Society* 193, 381–387. <https://doi.org/10.1093/zoolinnean/zlab072>
- Ferrante, M. (2025). The niche concept in a changing world. *Nature Reviews Biodiversity* 1, 419. <https://doi.org/10.1038/s44358-025-00060-x>
- Garcia, L. & Pardo, I. (2015). Food type and temperature constraints on the fitness of a dominant freshwater shredder. *International Journal of Limnology* 51, 227–235. <https://doi.org/10.1051/limn/2015017>
- Hebert, P.D., Cywinska, A., Ball, S.L. & deWaard, J.R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society B Biological Sciences* 270, 313–21. <https://doi.org/10.1098/rspb.2002.2218>
- IPBES (2019). Summary for policymakers of the global assessment report on biodiversity and ecosystem services. Zenodo <https://doi.org/10.5281/zenodo.3553458> Accessed 20th January 2026.
- King, S., Nicholls, M., Scales, J., Gubbins, S., Pearce-Kelly, P., Saverimuttu, S. *et al.* (2025). The efficacy of vector-proof accommodation for the protection of livestock against *Culicoides* biting midges. *Parasites & Vectors* 18, 108. <https://doi.org/10.1186/s13071-025-06736-9>
- Langton, P.H. & Pinder, L.C.V. (2007). *Keys to the Adult Male Chironomidae of Britain and Ireland*. Special Publication 64. Freshwater Biological Association, Ambleside, Cumbria.
- Lawniczak, M.K.N., Davey, R.P., Rajan, J., Pereira da Conceição, L.L., Kiliyas, E., Hollingsworth, P.M. *et al.* (2022). Specimen and sample metadata standards for biodiversity genomics: a proposal from the Darwin Tree of Life project. *Wellcome Open Research* <https://doi.org/10.12688/wellcomeopenres.17605.1>
- Mullens, B.A., McDermott, E.G. & Gerry, A.C. (2015). Progress and knowledge gaps in *Culicoides* ecology and control. *Veterinaria Italiana* 51, 313–323.
- Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. & Erlich, H. (1986). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbor Symposia on Quantitative Biology* 51, 263–73. <https://doi.org/10.1101/SQB.1986.051.01.032>
- Namayandeh, A., Guerra, S., Islam, N. & James, T. (2024). New species and a fascinating diversity of Chironomidae (Diptera, Insecta) in and around an overlooked urban vernal pool. *ZooKeys* 1208, 133–163. <https://doi.org/10.3897/zookeys.1208.124495>
- Park, N., Dawson, E., Thurston, S., Tuameh, A., Mosca, M.M., Pereira da Conceição, L. *et al.* (2023). High-throughput DNA barcoding library construction and sequencing protocol for BIOSCAN using unpurified non-destructively extracted DNA from arthropods. *Protocols.io*. <https://doi.org/10.17504/protocols.io.8epv5jzxd11b/v1>
- Porter, T.M. & Hajibabaei, M. (2018). Scaling up: a guide to high-throughput genomic approaches for biodiversity analysis. *Molecular Ecology* 27, 313–338. <https://doi.org/10.1111/mec.14478>
- Ratnasingham, S. (2019). mBRAVE: the multiplex barcode research and visualization environment. *Biodiversity Information Science and Standards* 3: e37986. <https://doi.org/10.3897/biss.3.37986>
- Ratnasingham, S. & Hebert, P.D.N. (2007). BOLD: The Barcode of Life Data System. *Molecular Ecology Notes* 7, 355–364. <https://doi.org/10.1111/j.1471-8286.2007.01678.x>
- Ratnasingham, S. & Hebert, P.D.N. (2013). A DNA-

based registry for all animal species: the Barcode Index Number (BIN) system *PLoS ONE* 8(7): e66213.

<https://doi.org/10.1371/journal.pone.0066213>

Roslin, T., Somervuo, P., Pentinsaari, M., Hebert, P.D.N., Agda, J., Ahlroth, P. *et al.* (2022). A molecular-based identification resource for the arthropods of Finland. *Molecular Ecology Resources* 22, 803–822.

<https://doi.org/10.1111/1755-0998.13510>

Scudder, G.G.E. (2017). The importance of insects. In: Footitt, R.G. & Adler, P.H. (Editors). *Insect Biodiversity: Science and Society Vol. 1* (2nd edition), pp.9-43. John Wiley & Sons, Hoboken, New Jersey.

<https://doi.org/10.1002/9781118945568.ch2>

Tarakhovskaya, E.R. & Garbary, D.J. (2009). *Halocladus variabilis* (Diptera: Chironomidae): a marine insect symbiotic with seaweeds from the White Sea, Russia. *Journal of the Marine Biological Association of the United Kingdom* 89, 1381-1385.

<https://doi.org/10.1017/S0025315409000071>

Tedesco, P.A., Bigorne, R., Bogan, A.E., Giam, X., Jézéquel, C. & Huguény, B. (2014). Estimating how many undescribed species have gone extinct. *Conservation Biology* 28, 1360-1370.

<https://doi.org/10.1111/cobi.12285>

Urban, M.C. (2024). Climate change extinctions. *Science* 386, 1123-1128.

<https://doi.org/10.1126/science.adp4461>

Wellcome Sanger Institute (2026). BIOSCAN. <https://www.sanger.ac.uk/collaboration/bioscan/> Accessed 20th January 2026.

Werritty, A. & Sugden, D. (2012). Climate change and Scotland: recent trends and impacts. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh* 103, 133-147.

<https://doi.org/10.1017/S1755691013000030>

Wheeler, Q.D., Raven, P.H. & Wilson, E.O. (2004). Taxonomy: impediment or expedient? *Science* 303, 285.

<https://doi.org/10.1126/science.303.5656.285>

Xu, H.F., Lv, M.Y., Zhao, Y., Zhang, Z.C., Z., Liu, Z., & Lin, X.L. (2025). Cryptic diversity and climatic niche divergence of *Brillia* Kieffer (Diptera: Chironomidae): insights from a global DNA barcode dataset. *Insects* 16(7), 675.

<https://doi.org/10.3390/insects16070675>