

Enhancing confidence in eDNA monitoring for freshwater invertebrates

**in Aotearoa–New Zealand through a genetic
and morphological evaluation**

Prepared for Otago Regional Council

September 2026

Prepared by:

Brian J. Smith

Stephen R. Pohe, Pohe Environmental

For any information regarding this report please contact:

Brian Smith

Freshwater Ecologist

Freshwater Ecology

+64 7 856 1756

brian.smith@earthsciences.nz

New Zealand Institute for Earth Science Limited

PO Box 11115

Hamilton 3251

Phone +64 7 856 7026

Client Report No: 2026181HN

Report date: September 2026

Project No: ELF26202

Quality Assurance Statement		
	Reviewed by:	Elizabeth Graham
	Formatting checked by:	Melissa Greentree/Jo Downey
	Approved for release by:	Michael Bruce

© New Zealand Institute for Earth Science Limited ("Earth Sciences New Zealand") 2026. All rights reserved.

This report has been prepared by the New Zealand Institute for Earth Science Limited (Earth Sciences New Zealand) exclusively for Otago Regional Council and for the specific purpose stated ("Purpose"). It is based on available information and professional judgment at the time of preparation, and is subject to any limitations or assumptions set out in the report. Unless otherwise agreed in writing by Earth Sciences New Zealand, Earth Sciences New Zealand provides no warranty as to the use of the report other than by the Otago Regional Council for the Purpose and Earth Sciences New Zealand will not be liable for any loss, damage or expense (on any ground and whether direct, indirect or consequential) arising from use of or reliance on this report or its contents by any person other than to Otago Regional Council for the Purpose.

Contents

Executive summary.....	5
1 Purpose and context.....	7
2 Background.....	9
3 Scope and approach.....	12
4 Findings and insights.....	15
4.1 Reference sequence database coverage	15
4.2 Sequence quality analysis by taxonomic groups.....	18
4.3 Threatened species coverage.....	21
4.4 Misidentified reference sequences – a case study.....	25
4.5 Metadata quality.....	29
4.6 Geographical coverage of sequences.....	32
4.7 Investigation of additional markers	36
5 Where to from here?.....	42
6 Conclusions	46
7 Acknowledgements.....	47
8 References	48

Tables

Table 4-1:	Public COI sequence availability for 675 species of freshwater invertebrates listed in Grainger et al. (2018).	15
Table 4-2:	Private sequence holdings for the 675 freshwater invertebrate species listed in Grainger et al. (2018), illustrating the potential contribution of unpublished records toward improving national reference library coverage and eDNA confidence.	17
Table 4-3:	Public sequence coverage and barcode completeness among major freshwater invertebrate taxonomic groups (675 species) listed in Grainger et al. (2018), highlighting differences in readiness for species-level eDNA interpretation.	18
Table 4-4:	Sequence coverage of conservation-priority freshwater invertebrates across major taxonomic groups (675 species), showing how reference library completeness varies among taxa of management and biodiversity significance.	22

Table 4-5:	Distribution of public sequences for conservation-priority freshwater invertebrates among threat-status categories and major taxonomic groups (140 species).	23
Table 4-6:	Metadata completeness for public sequences, summarised by major taxonomic group (395 species), showing the extent to which supporting information enhances confidence in reference library records.	30
Table 4-7:	Geographic representativeness of public reference sequences among major freshwater invertebrate groups (395 species), indicating the extent to which available records capture species distributions and genetic variation.	33
Table 4-8:	Distribution of publicly available sequences of major freshwater invertebrate taxonomic groups for different gene regions other than COI.	37

Figures

Figure 4-1:	Examples of species misidentification and unnecessary taxonomic ambiguity within public COI reference records for Leptophlebiidae mayflies, illustrating how reference library errors can reduce confidence in eDNA species assignments.	27
Figure 4-2:	Example of a family-level taxonomic misidentification in a public reference sequence dataset.	27
Figure 4-3:	Visualisation of a fragment of mayfly DNA showing the positions of the AQdb-16S primers (Takenaka et al. 2023) that amplify longer 16S DNA fragments useful for reference sequences, and the MiDeca primers (Komai et al. 2019) that amplify a shorter fragment designed for eDNA barcoding of decapods (Crustacea).	39

Executive summary

Regional councils are increasingly using environmental DNA (eDNA) for freshwater biomonitoring, biodiversity assessment, threatened-species and biosecurity surveillance, and development of higher-resolution tools such as an eDNA-based Macroinvertebrate Community Index (eMCI). The effectiveness of DNA barcoding relies on verified reference libraries of known taxa, because accurate species identification requires both comprehensive taxonomic coverage and rigorous quality assurance of reference sequences. However, the reliability of eDNA results depends on matching DNA reads to accurate, comprehensive and well-curated reference sequences.

This report assesses the status, strengths and limitations of DNA reference data for New Zealand freshwater invertebrates, with emphasis on biomonitoring taxa, especially Ephemeroptera (mayflies), Plecoptera (stoneflies) and Trichoptera (caddisflies), collectively known as EPT.

The assessment used the freshwater invertebrate species list from the Department of Conservation's New Zealand Threat Classification System (Grainger et al. 2018) as the working taxonomic framework. All 675 species were reviewed against public GenBank and BOLD sequence database records. During the process, evaluation was also made on additional markers, unpublished private sequence holdings, metadata quality, geographic representation, and taxonomic voucher-specimen issues, including museum-curated material linked to earlier barcode projects.

Publicly available DNA sequences in the GenBank and BOLD repositories are useful but incomplete. Of the 675 New Zealand taxa assessed, nearly 60% had some public GenBank sequence coverage, but only 40% had at least one full-length barcode (651–658 bp) and 40% with shorter sequences (≤ 650 bp), while about 42% had no public GenBank coverage. Full-length and shorter sequence categories overlap and are not mutually exclusive. Freshwater invertebrate coverage is strongest for some EPT taxa, but remains weak for many non-EPT groups, including crustaceans, molluscs, beetles and worm-like taxa. Conservation-priority species are also underrepresented, limiting the ability of eDNA programmes to support threatened-species monitoring without further curation, verification and targeted sequencing.

Sequence availability alone is not a sufficient measure of confidence. Species-level interpretation is constrained by sequence length, taxonomic reliability, voucher support, metadata completeness and geographic representativeness. The review identified specimen misidentifications, incomplete or incorrect metadata, overseas-only records, poor spatial representation and many taxa supported by only one or a few records. These issues can produce false positives, false negatives or over-confident species-level assignments if not quality-flagged.

Therefore, the immediate priority is not simply to generate more sequence data, but to create an expert-verified, quality-coded, New Zealand-specific reference resource (an authenticated sequence library) that councils can rely on for defensible ecological interpretation and future environmental Macroinvertebrate Community Index (eMCI)

development. The lowest-cost, highest-return actions are to publish verified private and institutional sequences, correct known taxonomic and metadata anomalies, apply transparent quality codes to existing records, and require providers to report reference sequence confidence alongside eDNA outputs. New taxa sequencing should be targeted toward habitats and regions where it will most improve management confidence.

This advice document has been prepared primarily for regional councils, at their request, but the messages are equally relevant to the Department of Conservation (DOC), Public Research Organisations (PROs), museums, taxonomic specialists, freshwater researchers, eDNA service providers, and other national partners. Strengthening freshwater invertebrate reference libraries is a shared national priority, and the benefits will extend well beyond council monitoring, including supporting threatened-species recovery, biodiversity reporting, regulatory confidence, and development and use of future tools such as the eMCI.

The practical implementation pathway recommended below reflects this wider context. While councils are key beneficiaries and contributors, many actions cannot be led by councils alone. Delivering a nationally consistent and defensible reference resource will require coordinated effort across freshwater managers and data providers. The pathway proposed below represents a nationally aligned direction rather than a council-only work programme, with councils playing a central role in prioritisation, commissioning, and ensuring that provider reporting meets confidence requirements.

A staged implementation pathway is recommended as follows:

1. triage existing records into confidence categories,
2. release low-risk verified private sequences,
3. resolve known misidentifications and metadata defects,
4. build a curated national freshwater invertebrate reference set,
5. target new collections of threatened, endemic, groundwater, alpine, island, headwater, and poorly sampled taxa, and
6. create a verified New Zealand-specific reference library.

These tasks are achievable but require adequate resourcing and a collective commitment. Better-curated reference libraries will increase confidence in eDNA tools, allowing freshwater managers to make more robust and defensible ecological assessments and management decisions

The detailed species-level results underpinning this assessment are provided in an accompanying Microsoft Excel file ("Freshwater invertebrate eDNA reference sequence assessment dataset"), which contains the complete assessment of all 675 freshwater invertebrate species reviewed in this report.

1 Purpose and context

This report was prepared to provide regional councils with practical, evidence-based advice on the current utility, limitations and improvement priorities for DNA reference data used in freshwater invertebrate environmental DNA (eDNA) monitoring in Aotearoa–New Zealand (hereafter New Zealand). It responds directly to the Envirolink Advice Grant (2624–ORC22 Enhancing confidence in eDNA monitoring for freshwater invertebrates), which identified an urgent need to improve confidence in eDNA-based freshwater biomonitoring by assessing the completeness, reliability and usability of molecular reference databases used to assign DNA sequences to freshwater taxa.

The need for this work arises because regional councils are increasingly investing in and adopting eDNA methods for freshwater monitoring, biodiversity and biosecurity assessment, restoration tracking and future development of higher-resolution biological indices. eDNA provides a powerful way to detect freshwater communities from water samples, including taxa and life stages that are difficult to identify using routine morphology.

The effectiveness of DNA barcoding relies on verified reference libraries of known taxa, because accurate species identification requires both comprehensive taxonomic coverage and rigorous quality assurance of reference sequences (Weigand et al. 2019). However, the reliability of eDNA results depends on matching DNA reads to accurate, comprehensive and well-curated reference sequences. Where reference databases contain missing taxa, mislabelled sequences, taxonomic inconsistencies, poor metadata or geographically unrepresentative records, eDNA outputs can be uncertain, inconsistent or open to challenge.

New Zealand's environmental information system is affected by fragmented holdings, data gaps, inconsistent standards and limited accessibility, which impede effective environmental management and inflate costs (PCE 2026). These systemic issues apply equally to molecular reference data, and without coordinated, standardised and discoverable reference sequences, councils cannot rely on eDNA outputs for regulatory or biodiversity-management decisions. Strengthening freshwater invertebrate reference data therefore contributes directly to national priorities for improving environmental information quality and stewardship.

Current limitations in DNA reference libraries and ecological knowledge remain significant barriers to the interpretation of eDNA-derived ecological indicators, providing a strong rationale for targeted investment in expanding and validating reference sequence resources (e.g., Suren et al. 2024). In advice provided to the Ministry for the Environment (MFE) on incorporating eDNA indicators into the National Objectives Framework (NOF) of the National Policy Statement for Freshwater Management (NPS–FM), Graham & Baker (2025, in White et al. 2025) identified improvement of DNA reference libraries as the highest priority for enhancing the application of eDNA-based macroinvertebrate metrics. This recommendation aligns directly with the purpose of the present project, which seeks to strengthen the taxonomic coverage and reliability of reference libraries to support robust freshwater biomonitoring.

The work has strong regional-council relevance, and the proposal was championed by Otago Regional Council and supported by freshwater practitioners Surface Water Integrated Management (SWIM), and its subgroup the Fish Focus Group, at Te Uru Kahika (Regional and Unitary Councils Aotearoa). Supporting councils included Northland Regional Council, Horizons Regional Council, Marlborough District Council, Tasman District Council, Otago Regional Council, Environment Southland, Auckland Council, Greater Wellington Regional Council, Waikato Regional Council and Environment Canterbury.

The advice also builds on previous Envirolink investment, specifically the Waikato Regional Council-led tools project Stocking New Zealand's molecular library for eDNA monitoring, and related national work on DNA reference data, groundwater fauna molecular sequences, the New Zealand eDNA Development Group and SWIM.

2 Background

New Zealand freshwater ecosystems support a rich and highly endemic invertebrate fauna. Aquatic insects such as mayflies, stoneflies and caddisflies are key indicators of stream health, while many non-insect groups also contribute important ecological and conservation information. Traditional biomonitoring has long relied on morphological identification of benthic macroinvertebrates, but many species are difficult to identify to species level, particularly juveniles, damaged specimens, cryptic taxa, undescribed species and groups that require specialist taxonomic expertise. Routine surveys therefore often report taxa at genus level or higher, which can obscure biodiversity, regional endemism, species-specific ecological traits and sensitivity to environmental change.

Murray et al. (2024) identified detection and outcome monitoring as one of the highest conservation research priorities in New Zealand, noting that freshwater fishes and freshwater invertebrates have among the greatest unmet needs for monitoring tools and biodiversity inventory methods. Their national research gap analysis showed that 90% of assessed threatened taxa require new or improved detection or monitoring methods, with freshwater invertebrates among the groups with the highest research need. This reinforces the importance of developing comprehensive and defensible molecular reference libraries to support emerging eDNA-based monitoring approaches.

eDNA metabarcoding offers an important opportunity to improve freshwater monitoring because it can detect multiple taxa from water samples, reduce some sampling and identification constraints, and incorporate information from life-history stages that are underrepresented in conventional sampling, including juveniles, pupae and emerging adults. These advantages are particularly relevant as councils move toward more sensitive, scalable and cost-effective biomonitoring tools and seek to develop next-generation indicators such as an environmental DNA-based Macroinvertebrate Community Index (eMCI). However, eDNA does not remove the need for taxonomy. Instead, it increases the importance of reliable taxonomic foundations because molecular assignments are only as good as the reference sequences that support them.

This assessment recognises that the reliability of eDNA-derived taxonomic identifications depends heavily on reference-database completeness, and that gaps in barcode coverage represent a significant limitation for ecological metabarcoding studies (Melchior & Baker 2023). Furthermore, known specimen misidentifications, sequencing errors, conflicting sequences, taxonomic inconsistencies, low taxonomic resolution and incomplete metadata can lead to false positives, false negatives or over-confident species-level assignments. These risks can compromise ecological assessments and ultimately affect regulatory, restoration and conservation decisions. Species that are not represented in reference databases may only be assigned to higher taxonomic levels such as genus or family, reducing taxonomic resolution and limiting the ecological information that can be derived from monitoring data (Melchior & Baker 2023). Melchior & Baker (2023) further noted that unidentified sequences generated during eDNA analyses may represent species that are not yet described scientifically or are absent from current reference databases. As reference libraries expand and new species are

characterised, these previously unassigned sequences may subsequently be resolved to species or other taxonomic levels, highlighting the value of continued investment in reference library development.

Regional councils requested this advice because the effectiveness of eDNA biomonitoring is currently constrained by the limited coverage and variable accuracy of freshwater invertebrate molecular reference sequence data. Public databases such as GenBank and BOLD provide essential infrastructure, but they do not yet comprehensively or accurately represent New Zealand's freshwater invertebrate fauna.

The PCE (2026) report emphasised that environmental information in New Zealand is collected by numerous organisations with different mandates, priorities and methods, resulting in "complexity and fragmentation" and "duplication and overlaps" across datasets. Molecular reference data for freshwater invertebrates exhibit the same structural weaknesses whereby sequences are distributed across GenBank, BOLD, private holdings, museum collections and unpublished institutional datasets. Currently, there is no single custodian or unified system of quality control of data input and, equally, no quality assurance of data produced. This fragmentation reduces confidence in species-level assignments and reinforces the need for a coordinated national reference library framework.

The analysis by Murray et al. (2024) also highlighted that freshwater invertebrates are one of the most poorly understood threatened groups, with limited taxonomic resolution, sparse distribution data and inadequate monitoring methods. Their findings emphasise that many freshwater invertebrates remain Data Deficient or Data Poor under the New Zealand Threat Classification System (NZTCS), and that improved detection methods, including molecular approaches, are essential for resolving conservation status, understanding pressures, and informing management.

The advice provided in this report aligns with regional council strategies and statutory responsibilities by improving the quality and consistency of biodiversity information used for freshwater management. Councils have responsibilities under legislation including the Resource Management Act 1991 and Biosecurity Act 1993 to maintain and enhance indigenous biodiversity and to support effective environmental management.

Macroinvertebrate monitoring is also embedded in freshwater assessment under the National Policy Statement for Freshwater Management (NPS-FM 2020 [updated 2025]) to improve the health of the nation's freshwaters (MFE 2025). By improving reference library confidence, this work supports more robust state of environment reporting, freshwater planning, consent monitoring, restoration assessment, threatened species and biosecurity surveillance and long-term biodiversity management. It also aligns with national priorities advanced through SWIM and the Regional Council Research for Resource Management Strategy by contributing to a more nationally consistent and scientifically defensible molecular reference framework.

The environmental benefits of the work programme are therefore practical and strategic. Better-curated reference libraries will increase confidence in eDNA tools, improve species-level detection, reveal hidden or under-recorded biodiversity, including biosecurity threats, help identify rare, threatened and locally endemic taxa, and reduce the risk that database errors are carried into council monitoring outputs. This will allow freshwater managers to make more robust and defensible ecological assessments and improve the evidence base for decisions that affect water quality, habitat protection, biodiversity management and restoration outcomes across New Zealand.

3 Scope and approach

The New Zealand Threat Classification System (NZTCS) is the national framework used to assess the conservation status of indigenous taxa in New Zealand. Administered by the Department of Conservation (DOC), the system provides a consistent and transparent process for evaluating extinction risk and conservation priority for DOC, but is widely used by government agencies, researchers and environmental practitioners. NZTCS assessments underpin species recovery planning, biodiversity management, regional and national reporting, and statutory decision-making processes, including those associated with the Resource Management Act 1991 and regional planning frameworks. Assessments are undertaken by expert panels, convened by DOC, and comprise specialists drawn from New Zealand and international scientific communities.

This report reviews the current status of DNA reference sequences publicly available for New Zealand freshwater invertebrates using the 675 freshwater taxa listed by Grainger et al. (2018) as the working taxonomic framework. At the time of writing, a revised 2026 freshwater invertebrate conservation-status assessment was ‘in progress’, but was not yet publicly available. Consequently, the Grainger et al. (2018) assessment was used as the most recent accessible and nationally recognised taxon list available during report preparation. The conservation classifications and taxonomic treatment presented in this report are therefore assumed to be current and correct for the purposes of analysis and interpretation, recognising that some threat classifications, nomenclature and species totals may change following publication of the 2026 assessment.

The use of the 2018 assessment provides a consistent baseline against which molecular reference sequence coverage, metadata quality, geographic representation and conservation-priority gaps can be evaluated. Where future NZTCS assessments revise species rankings or taxonomic concepts, the patterns identified in this report are still expected to remain broadly relevant because the principal focus is on the completeness, quality and fitness-for-purpose of freshwater invertebrate DNA reference sequences rather than the conservation status of individual species.

The taxonomic coverage spans multiple freshwater invertebrate groups, including: Ephemeroptera (mayflies), Plecoptera (stoneflies) and Trichoptera (caddisflies), collectively known as EPT, Coleoptera (beetles), “other” aquatic insects (e.g., true-bugs, dobsonflies, dragon-and damselflies), crustaceans (amphipods, crayfish, isopods, shrimps, zooplankton), molluscs (bivalves, limpets, snails) and “worm-like” taxa (i.e., flat worms, Gordian worms, nematodes, round worms). Greater interpretive effort is given to EPT because these groups are sensitive to environmental degradation, central to freshwater biomonitoring and supported by stronger available taxonomic expertise and reference material. This focus is consistent with the advice proposal, which recognised EPT as high-information indicator groups while also identifying the need to assess broader freshwater invertebrate reference library gaps.

Diptera are notably absent from this coverage, having been excluded from Grainger et al. (2018) NZTCS freshwater invertebrate lists on the grounds that they fell under a separate taxon-specific assessment (see Andrew et al. 2012).

However, that assessment covered only a small number of species, so the vast majority of freshwater Diptera (across many families) remain unassessed; representing a gap in taxonomic coverage (and knowledge) rather than in reference sequence quality.

Each species was cross-referenced against GenBank and BOLD, the major public molecular databases, and assessed for the availability and utility of COI barcode data and other potentially useful genetic markers. The assessment considers sequence availability, sequence type, sequence length, metadata completeness, taxonomic reliability, geographic representation, availability of private or institutional sequence holdings, and known or suspected database anomalies. For the purposes of this assessment, a full-length COI barcode is defined as 651–658 bp and all sequences ≤650 bp were classified as ‘shorter’ COI sequences. The review also identifies key issues affecting confidence in current reference sequences, including mislabelled species, sequencing errors, conflicting sequences, unresolved taxonomy, low taxonomic resolution, missing taxa, incomplete metadata and potentially high intraspecific genetic variation.

To support transparency and future reference-library development, the working dataset underpinning this assessment is provided as a companion Microsoft Excel file. The dataset contains the 675 freshwater invertebrate species assessed in this review, together with taxonomic hierarchy, NZTCS conservation status and unique taxon identification number, public and private sequence availability, sequence-length categories, metadata assessments, geographic coverage assessments and associated comments. The spreadsheet is intended as a searchable resource that allows councils, researchers and practitioners to identify taxonomic, geographic and data-quality gaps and to prioritise future sequencing, curation and reference-library development activities.

The approach was designed as a quality control and quality assurance assessment rather than a database correction exercise. Its purpose is to identify where current reference data are fit for council interpretation, where results should be treated cautiously, and where future investment would most improve confidence in eDNA monitoring. The report documents the current state of reference-sequence coverage and quality, highlights priority gaps, and provides advice to support provider reporting requirements, prioritisation of sequencing and curation effort, development of nationally consistent reference resources, and staged implementation toward future eMCI development. Within the Findings and insights section, each key finding is accompanied by an assessment of its implications for regional councils, followed by practical recommendations to improve confidence in eDNA-based freshwater monitoring and interpretation.

This approach aligns with broader national recommendations for environmental information stewardship. The PCE (2026) report argued that environmental data must be “accessible, discoverable and usable” and that data integration is essential to generate joined-up insights. Applying these principles to molecular reference datasets means ensuring that freshwater invertebrate sequences are curated, standardised, linked to metadata, and interoperable across platforms used by councils, researchers, and service providers.

Implementing these recommendations will require coordinated action among councils, research organisations, museums, taxonomic experts, eDNA providers, and national biodiversity–data infrastructure. While this report does not seek to correct public database anomalies, publish private sequence holdings, or undertake new sequencing of priority taxa, it provides a framework to help guide and prioritise those future activities.

4 Findings and insights

4.1 Reference sequence database coverage

Rather than presenting the complete 675-species dataset as a report appendix, the results are provided as a searchable Microsoft Excel file accompanying this report. The spreadsheet contains the detailed species-level assessments used to generate the summaries presented below.

4.1.1 Publicly available sequences

Full (651–658 bp) or shorter (≤ 650 bp) COI barcode sequences were publicly available in GenBank for less than 59% (395) of the 675 freshwater invertebrate species assessed, indicating that more than two-fifths of the fauna currently lack publicly accessible data (Table 4–1).

Of the 395 species assessed with an available sequence, around 40% were represented by at least one full-length COI barcode, providing a potential basis for species-level identification. Breaking this down further, 188 species (data not presented) were represented by one to four full-length sequences, which may be sufficient for preliminary identification, but provide limited insight into intraspecific variability and geographic representation. The remaining 80 species were represented by five or more full-length sequences, providing relatively robust barcode coverage and improving confidence in species-level assignments, subject to verification, metadata quality (see Section 4.5), and geographic representation (see Section 4.6).

A similar 39% of species were only represented by shorter sequences. While shorter sequences may assist taxonomic assignment, they generally provide lower confidence for species-level identification in ecological monitoring and biodiversity surveys.

Table 4–1: Public COI sequence availability for 675 species of freshwater invertebrates listed in Grainger et al. (2018). The proportion of taxa with full-length, shorter, or no publicly available barcode coverage and the implications for species-level eDNA identification. Full length and shorter sequence proportions are independent: a priority taxon can be counted twice when both sequence types are available.

Public sequence availability	No. species	Coverage (%)	Implications
Available	395	58.5	Full or shorter sequences available, but no indication of quality or correctness.
Full-length (651–658 bp)	268	39.7	Good species-level matching, subject to verification, metadata quality and geographic coverage.
Shorter-length (≤ 650 bp)	264	39.1	Potentially useful but subject to the shorter fragment position matching the small eDNA fragment position.
No sequence	280	41.5	Species-level eDNA detection is not currently supported by public GenBank records.

Over 40% or 280 species listed in Grainger et al. (2018) lacked any publicly available sequences, a similar proportion to that found by Weigand et al. (2019) for European aquatic invertebrates. This represents a significant weakness in currently accessible reference sequence data. In addition to highlighting a need to focus on a number of taxonomic groups, a key priority for the future is to continue adding new species with high quality-sequences to publicly accessible reference databases, in order to progressively fill the remaining taxonomic and geographic gaps. Achieving this goal will require sustained investment in sequencing programmes, ongoing funding for database management and support for taxonomic expertise.

4.1.2 Privately-held sequences

Even when biodiversity and environmental information is known to exist, access to the underlying reference data can remain limited as not all sequences are publicly accessible through GenBank or BOLD. Many DNA barcode sequences, specimen records, and eDNA datasets are held by research institutions, consultancies or private databases and may not be publicly accessible, thus limiting the ability of eDNA analyses to reliably identify species. As a result, local and central government agencies, or even community groups, may be unable to fully utilise these data for biodiversity monitoring, species detection, and environmental management.

Although this assessment focuses on publicly available datasets, privately-held sequences are reported here for context because they represent a substantial but currently under-used contribution to national reference sequence coverage. A number of New Zealand projects have resulted in the generation of collections of aquatic insect sequences over the years through BOLD, with some eventually being made public (e.g., Hogg et al. 2009), while others remain as private repositories and therefore inaccessible to most users. However, more recent national data-sharing work shows that this situation can change where data owners, laboratories and biodiversity-data infrastructure agree on permissions, standards and metadata requirements. For example, the Ministry for the Environment MfE-commissioned eDNABridge pilot (Hards & Muellner 2026) successfully published public metabarcoding records from DOC, Hawke's Bay Regional Council and Earth Sciences New Zealand to Global Biodiversity Information Facility (GBIF), demonstrating that large-scale eDNA data sharing is feasible.

Known private holdings include full-length COI sequences for 188 New Zealand freshwater invertebrate species (Table 4-2). Even though some of these records overlap with public data, private datasets include 60 species with full-length sequences that are not represented in GenBank, including 44 EPT species, and a further 85 species with shorter sequences. Importantly, 32 species of freshwater invertebrates are solely represented by full-length sequences in private databases. Several of these privately held records relate to species listed by Grainger et al. (2018) as Nationally Critical, Nationally Endangered or Nationally Vulnerable, highlighting their potential value for improving threatened-species detection and increasing confidence in council eDNA interpretation. If private records could be purchased or made accessible through appropriate permissions, agreements or other access arrangements, this would significantly increase overall coverage of freshwater invertebrates to 63%, an increase of over 20%.

Table 4-2: Private sequence holdings for the 675 freshwater invertebrate species listed in Grainger et al. (2018), illustrating the potential contribution of unpublished records toward improving national reference library coverage and eDNA confidence.

Private sequence holdings	No. of species	% Coverage
Total number	188	27.8
Full-length (651–658 bp)	60	8.8
Shorter-length (≤ 650 bp)	85	12.6
Unique (not in public databases)	32	4.7

4.1.3 Outcome for regional councils

Public COI coverage is sufficient to support useful eDNA interpretation for some freshwater invertebrates, but not yet sufficient for species-level reporting across the full fauna. The major implication for councils is that reference sequence availability should be separated from reference sequence reliability. A taxon may have a public sequence, but that sequence may be shorter, poorly documented, geographically narrow, taxonomically uncertain or supported by too few records to justify high-confidence ecological interpretation. Councils (and other recipients of eDNA data) should therefore require eDNA outputs to report the confidence basis for each taxonomic call and interpret results at the most defensible taxonomic level.

The lowest-hanging fruit to improve reference sequence reliability is the mobilisation of existing verified sequence data. Privately-held sequences and institutional holdings already include full-length COI sequences for taxa not represented in public databases, including some EPT and conservation-priority species. If these records can be verified, linked to voucher material, populated with minimum metadata and made accessible through an agreed national framework, councils would gain immediate improvements in detection confidence.

4.1.4 Recommendations

Actions should be prioritised as follows:

1. publish low-risk, voucher-backed private and institutional records where identifications and metadata are already robust,
2. quality-code all public and private records using transparent criteria for sequence length, taxonomic verification, voucher status, metadata completeness and geographic coverage,
3. correct or annotate known anomalies, including misidentified records, coordinate errors, overseas-only records and unresolved names,
4. establish a curated, national freshwater invertebrate reference library, and

5. target new collecting and sequencing toward taxa with no public coverage, conservation-priority species, regionally endemic taxa and underrepresented regions and habitats.

This staged approach delivers immediate gains while building a defensible reference resource for long-term monitoring.

4.2 Sequence quality analysis by taxonomic groups

Direct comparison of sequence coverage estimates among studies should be made with caution because different analytical units and assessment methods can produce different results. Barcode Index Numbers (BINs) are increasingly being used as a proxy for species richness (e.g., Buckley et al. 2026). This can create a potential for misinterpretation when estimating biodiversity or reference-library completeness because they do not always correspond to taxonomically validated species. Consequently, they can overstate reference-library completeness, particularly when multiple BINs exist for a single species or when BINs include records of uncertain identity (Emerson 2025). In contrast, the present study evaluated sequence availability at the recognised species level.

Public COI sequence coverage was uneven across taxonomic groups (Table 4-3). EPT were better represented than most other freshwater invertebrate groups. This likely reflects their long-standing use in biomonitoring, the availability of autecological information and taxonomic resources, and access to specialist expertise and well-curated collections. Nonetheless, substantial gaps remain even within the EPT.

Table 4-3: Public sequence coverage and barcode completeness among major freshwater invertebrate taxonomic groups (675 species) listed in Grainger et al. (2018), highlighting differences in readiness for species-level eDNA interpretation. * = EPT; "other" insects = Hemiptera, Lepidoptera, Mecoptera, Megaloptera, Odonata; Crustacea = amphipods, crayfish, isopods, shrimps, zooplankton; Mollusca = bivalves, limpets, snails; "worm-like" = flat worms, Gordian worms, nematodes, round worms, note, full length and shorter sequence proportions are independent: a species can be counted twice when both sequence types are present and not necessarily sum to the value in "Yes".

Major taxonomic groups	No. species	Public sequence availability (%)		Sequence length based on availability (%)	
		No	Yes	Full-length (651–658 bp)	Shorter-length (≤650 bp)
Ephemeroptera*	59	40.7	59.3	44.1	28.8
Plecoptera*	107	24.3	75.7	32.7	72.0
Trichoptera*	242	21.5	78.5	71.5	41.3
Coleoptera	62	75.8	24.2	24.2	4.8
"other" insects	46	54.3	45.7	30.4	32.6

Major taxonomic groups	No. species	Public sequence availability (%)		Sequence length based on availability (%)	
		No	Yes	Full-length (651–658 bp)	Shorter-length (≤ 650 bp)
Crustacea	57	75.4	24.6	5.3	24.6
Mollusca	79	60.8	39.2	2.5	38.0
“worm-like”	23	65.2	34.8	0	34.8
All groups	675	41.5	58.5	39.7	39.1

Full-length COI barcode coverage was highest for Trichoptera, with approximately 72% of sequenced species represented by full-length records, followed by Ephemeroptera at 44% and Plecoptera at 33%. These EPT patterns indicate that the best-developed parts of current reference sequence datasets are concentrated in taxa already used routinely for freshwater biomonitoring. This improves the likelihood that some mayfly, stonefly and caddisfly detections can be resolved to species level, but it does not remove the need for taxonomic verification, metadata review and geographic assessment. A species represented by a single full-length sequence, or by sequences from only one geographic region, should still be treated differently from a species represented by multiple expert-verified, voucher-backed records across its distribution.

In contrast, several non-EPT groups remain poorly supported for species-level eDNA interpretation. Coleoptera and other aquatic insect groups had lower public sequence availability overall, while crustaceans, molluscs and worm-like taxa were especially weakly represented by full-length public COI records. For some molluscs and crustaceans, COI alone is not always sufficient for species resolution, and a second mitochondrial marker (commonly 16S) may be used (e.g., Asmyhr & Cooper 2012; Chee 2015), which may, in part, explain the comparatively fewer public COI records available for these groups (see Section 4.7).

This does not mean these groups are ecologically unimportant, but rather, it shows that the foundation of reference sequences are less developed for many taxa that may be locally endemic, habitat-specialist, conservation-significant or difficult to identify morphologically. For these groups, eDNA outputs are more likely to be reported at genus, family or higher taxonomic levels, and apparent species-level detections should be treated cautiously unless supported by verified reference material.

Shorter sequences are generally most useful for coarser taxonomic identification, such as genus level or above. They may support species-level assignment in some cases, particularly where closely related taxa are adequately represented, but they generally provide lower confidence than full-length barcodes for environmental reporting and biomonitoring. This distinction is especially important for non-insect groups, where few full-length sequences are available and interpretation may depend disproportionately on shorter records.

Gaps in molecular reference data are closely linked to gaps in taxonomic expertise. Consistent with the observations of Winterbourn (2014), freshwater invertebrate groups that have attracted sustained taxonomic attention generally possess more comprehensive reference-sequence resources, whereas poorly studied groups remain underrepresented. The resulting uneven barcode coverage reflects historical differences in research effort, biomonitoring demand, specimen availability and specialist expertise. The “taxonomic impediment” described by Taylor (1983) forms part of a broader, long-documented decline in taxonomic capacity and species-description effort globally (e.g., Löbl et al. 2023). New Zealand is not immune to this problem (Lester et al. 2014; Leschen et al. 2016); limited investment in local and national taxonomic expertise and training constrains both the generation of newly-validated reference sequences and the verification and quality assurance of existing records.

EPT taxa have benefited from decades of ecological use, specialist attention and curated collections, while less-charismatic invertebrate taxa (e.g., mites, nematodes, worms) are increasingly reliant on legacy material collected during an earlier era when taxonomic effort and museum and university-based expertise were more strongly supported (e.g., Páll-Gergely et al. 2024). As expertise has declined, so too has the capacity to generate new, well-vouchered reference sequences for these groups, leaving many reliant on older curated material (Csabai et al. 2023). Since DNA degrades over time, this material is more likely to produce poor quality, shorter sequence reads (Strutzenberger et al. 2012), so new collecting and modern preservation remain important for building high-confidence reference libraries.

For regional councils, the key message is that sequence coverage is not evenly transferable across taxonomic groups. eDNA may already provide relatively strong species-level evidence for selected EPT taxa, but it should not be assumed to provide equivalent resolution for all freshwater invertebrates. eDNA provider reports should therefore distinguish between groups where species-level assignments are well supported and groups where the underlying reference data are incomplete, sequences are shorter, geographically narrow or taxonomically uncertain. This distinction is essential for avoiding over-confident interpretation of biodiversity inventories, threatened-species searches, consent-monitoring outputs and future eMCI development.

4.2.1 Outcome for regional councils

The utility of eDNA will not extend equally across the freshwater invertebrate fauna without targeted reference library investment. Councils can place greater confidence in species-level calls for taxa supported by full-length, expert-verified and geographically relevant COI records, but should be cautious for groups dominated by shorter, absent or poorly documented sequences. This is particularly important for crustaceans, molluscs, beetles and worm-like taxa, where public full-length COI coverage is low and where ecological or conservation significance may be high despite poor reference sequence representation.

4.2.2 Recommendations

Priorities for improving sequence quality should be: (1) require full-length COI barcodes where practical for species-level interpretation; (2) use recently collected, well-preserved and voucher-backed specimens wherever possible; (3) involve taxonomic specialists in specimen identification, database vetting and quality control; (4) review type, voucher or authoritatively identified material for problem taxa; (5) further evaluate additional markers such as 16S and/or 12S in the development of specific assays for freshwater invertebrates, to improve species detections where COI provides poor amplification or has limited reference sequence coverage (particularly for taxa such as Crustacea and Mollusca).

These actions should focus first on high-use EPT taxa, threatened or locally significant species, and non-EPT groups where current public coverage is weakest.

4.3 Threatened species coverage

Conservation-status classifications used in this assessment follow Grainger et al. (2018), which was the most recent publicly available NZTCS freshwater invertebrate assessment available during report preparation. Although an updated assessment had been completed in 2026, it was not available for incorporation at the time of writing. Consequently, some species classifications and threat rankings presented here may differ from those adopted in the forthcoming assessment.

In this report, the term conservation-priority species refers to a threatened or declining species identified as needing high-priority research to support its conservation (*sensu* Murray et al. 2024). For these species, as well as other rare freshwater invertebrates, an apparent absence in an eDNA result should not automatically be interpreted as ecological absence. Where a species lacks a reliable reference sequence, or is represented only by short, poorly documented or geographically distant sequences, metabarcoding may fail to assign reads to that species even when its DNA is present in the sample. Gaps in public reference sequences can therefore generate false negatives or low-resolution assignments for the taxa that councils most need to detect for biodiversity protection, freshwater planning and conservation management.

The gaps found in reference sequence datasets align with national-scale findings from Murray et al. (2024), who reported that freshwater invertebrates have some of the highest research needs of any threatened group in New Zealand. Their analysis showed that 97% of threatened freshwater invertebrates require research to understand causes of decline, 85% require research to develop management plans, and 84% require improved detection or outcome monitoring methods. Weak barcode coverage therefore directly limits councils' ability to detect, monitor and manage conservation-priority freshwater invertebrates using eDNA.

Table 4-4: Sequence coverage of conservation-priority freshwater invertebrates across major taxonomic groups (675 species), showing how reference library completeness varies among taxa of management and biodiversity significance. Conservation-priority refers to a threatened or declining taxon identified as needing high-priority research to support its conservation; * = EPT; “other” insects = Hemiptera, Lepidoptera, Mecoptera, Megaloptera, Odonata; Crustacea = amphipods, crayfish, isopods, shrimps, zooplankton; Mollusca = bivalves, limpets, snails; “worm-like” = flat worms, Gordian worms, nematodes, round worms; note, full length and shorter sequence proportions are independent, and a priority taxon can be counted twice when both sequence types are present and not necessarily sum to the value in “Priority”.

Major taxonomic groups	No. species	Conservation-priority status (%)			Priority taxa sequence length (%)	
		Non-priority	Priority	Priority but no sequence	Full-length (651–658 bp)	Short (≤650 bp)
Ephemeroptera*	59	59.3	40.7	23.7	29.2	16.7
Plecoptera*	107	53.3	46.7	17.8	18.0	56.0
Trichoptera*	242	58.3	41.7	19.4	42.6	22.8
Coleoptera	62	25.8	74.2	61.3	17.4	2.2
“other” insects	46	80.4	19.6	17.4	0	11.1
Crustacea	57	19.3	80.7	66.7	4.3	17.4
Mollusca	79	15.2	84.8	53.2	1.5	37.3
“worm-like”	23	47.8	52.2	39.1	0	25.0
All groups	675	47.4	52.6	31.9	19.7	26.2

This risk is greatest for species that are naturally rare, range-restricted, cryptic, recently described, taxonomically unresolved or associated with habitats that are poorly represented in routine monitoring. Seepages, waterfalls, groundwater habitats, alpine streams and small headwater streams support a diverse range of conservation-interest freshwater invertebrates, including, but not limited to EPT (e.g., Collier & Smith 2006; Graham et al. 2022; Pohe et al. 2020; Pohe et al. 2023; Smith et al. 2024), crustaceans (e.g., Fenwick 2001) and molluscs (e.g., Haase 2008). Without targeted collection, sequencing and expert verification, current reference databases are likely to underrepresent these threatened and locally significant species. Table 4-4 summarises sequence coverage for taxa groups assigned conservation-priority status in Grainger et al. (2018). Priority species are categorised as Threatened (including Nationally Critical, Nationally Endangered, and Nationally Vulnerable statuses), At Risk (including Declining and Naturally Uncommon statuses), and Data Deficient taxa with uncertain risk of extinction. Non-priority species are classified as Not Threatened, Vagrant, Coloniser, or Introduced and Naturalised.

The results show that molecular coverage for conservation-priority species is particularly weak. Over half (355 species, 52.6%) of the 675 species listed by Grainger et al. (2018) are of conservation priority (Table 4-4). Of these (355 species), only 140 species (39%) had any form of barcode representation, and only 70 species (20%) were represented by full-length COI barcode sequences. More concerning, no public barcodes were available for 215 priority species; 61% of species listed as Threatened, At Risk, or Data Deficient) and 32% of all species listed by Grainger et al. (2018). As previously noted elsewhere in this report, EPT were better represented than most other freshwater invertebrate groups, but substantial gaps remain even within these routinely monitored and ecologically informative orders.

Table 4-5: Distribution of public sequences for conservation-priority freshwater invertebrates among threat-status categories and major taxonomic groups (140 species). * = EPT; “other” insects = Hemiptera, Lepidoptera, Mecoptera, Megaloptera, Odonata; Crustacea = amphipods, crayfish, isopods, shrimps, zooplankton; Mollusca = bivalves, limpets, snails; “worm-like” = flat worms, Gordian worms, nematodes, round worms; Data Deficient species cannot be assigned a threat classification due to a lack of current information on their distribution and abundance.

Major taxonomic groups	No. species	Threatened			At Risk		Insufficient data
		Nationally Critical (%)	Nationally Endangered (%)	Nationally Vulnerable (%)	Declining (%)	Naturally Uncommon (%)	Data Deficient (%)
Ephemeroptera*	10	0	22.2	0	33.3	9.3	3.8
Plecoptera*	31	45.0	11.1	0	0	16.3	26.4
Trichoptera*	54	25.0	55.6	77.8	16.7	55.8	22.6
Coleoptera	8	0	0	0	0	0	15.1
“other” insects	1	0	0	0	0	2.3	0
Crustacea	8	0	11.1	0	33.3	2.3	7.5
Mollusca	25	25.0	0	22.2	16.7	14.0	20.8
“worm-like”	3	5.0	0	0	0	0	3.8
All groups	140	14.3	6.4	6.4	4.3	30.7	37.9

A breakdown of the 140 priority species by conservation status qualifier is presented in Table 4-5 by risk of extinction. EPT collectively accounted for 95 (68%) of all priority species represented in public sequence databases and dominated most conservation status categories. Their contribution was particularly pronounced among species classified as Data Deficient and Naturally Uncommon, where they comprised most species with available sequence data.

The remaining 45 priority species are drawn from the non-EPT taxonomic groups, although the greatest shortfall occurs among Nationally Critical species. This highlights a substantial gap in barcode reference data for many of New Zealand's most conservation-dependent freshwater invertebrate species.

4.3.1 Outcome for regional councils

eDNA provides valuable evidence for threatened-species surveillance, but current species-level detections and non-detections should be interpreted in relation to reference sequence confidence and the conservation context of each species. DOC's threatened-species research gap analysis found that most assessed Threatened and At Risk taxa require further research to support management, with freshwater invertebrates among the groups with especially high needs for research on causes of decline, management planning, and detection or outcome monitoring methods (Murray et al. 2024).

For regional councils, this means that weak barcode coverage for conservation-priority freshwater invertebrates directly limits the ability of eDNA programmes to support freshwater planning, consent monitoring, restoration assessment, biodiversity reporting and threatened-species management. A non-detection may reflect true absence, but it may also reflect insufficient sampling, low DNA shedding or transport, poor amplification, incomplete reference coverage, or unresolved taxonomy and, on some level, a combination of them all. The most defensible approach is therefore to report the highest taxonomic resolution supported by the reference data and to clearly flag records affected by absent, shorter, geographically distant, poorly documented or taxonomically uncertain sequences.

Murray et al. (2024) emphasised that many threatened freshwater invertebrates cannot currently be managed effectively because their distributions, population trends and pressures remain poorly understood. They found that only one freshwater invertebrate taxon in their assessment had sufficient knowledge to be fully managed, and that most require substantial investment in detection and monitoring methods before management can be effective. This reinforces the need for improved molecular reference coverage to support defensible eDNA-based detection of threatened species.

4.3.2 Recommendations

Improving reference-sequence coverage for threatened freshwater invertebrates should be treated as a shared priority across organisations involved in biodiversity monitoring, conservation research, environmental management and eDNA development, rather than as a standalone technical issue. Achieving this will require coordinated effort among regional councils, DOC, PROs, museums, taxonomic specialists, eDNA providers, freshwater managers, and other biodiversity-data partners to ensure that threatened and poorly represented taxa are prioritised for collection, sequencing, curation and validation.

Actions should be prioritised as follows:

1. target sequencing toward Threatened, At Risk, Data Deficient and locally significant freshwater invertebrates with no public coverage,
2. prioritise taxa and taxonomic groups where improved detection or monitoring methods would benefit multiple threatened species,
3. use recent, well-preserved and voucher-backed specimens or, if not available, support further targeted collection and curation of fresh voucher specimens,
4. require expert taxonomic verification before sequences are used for reporting,
5. publish verified private and institutional holdings where appropriate, and
6. apply minimum metadata standards covering locality, coordinates, collection date, collector, identifier and voucher details.

For high-consequence applications (e.g., threatened-species management, consent monitoring, freshwater planning and restoration assessment), councils should use eDNA alongside ecological context, conventional survey data, DOC threat-status information and specialist taxonomic review, and should avoid treating unsupported species-level non-detections as evidence that a threatened species is absent. Councils should also seek opportunities to align new sequencing and monitoring investments with DOC research priorities, particularly for taxa where improved detection and outcome monitoring would help resolve management-relevant knowledge gaps.

4.4 Misidentified reference sequences – a case study

The accuracy of DNA barcode reference libraries depends on robust specimen identifications, as errors can be perpetuated through subsequent bioinformatic sequence assignments to unverified specimens (Buckley et al. 2026). As part of this evaluation, the Canterbury Museum in Christchurch was visited during May 2026 to examine voucher specimens linked to earlier BOLD sequence barcode projects for freshwater insects.

These projects generated 1,243 EPT barcodes of varying quality between 2011 and 2016, with approximately 48% of the barcoded specimens held at Canterbury Museum. The museum holdings were particularly important for Ephemeroptera and Plecoptera, with 92% of BOLD-barcoded mayfly specimens and 90% of stonefly specimens held there, compared with 31% of caddisfly specimens. This made the Canterbury Museum collection a practical and high-value starting point for assessing the taxonomic reliability of existing sequence records, particularly because the authors were also able to provide expert taxonomic validation for these orders.

DNA barcode reference libraries are strongest when sequences are linked to vouchered specimens that can be re-examined by taxonomic specialists (Santos et al. 2023).

Even in well-curated national collections, specimen identifications reflect the expertise available at the time, the state of the taxonomy, the condition of the material, and the period since the specimens were last examined (McGinley 1993, *in* Pentinsaari et al. 2020). However, taxonomic errors do occur, and even experts can make mistakes in species identification (Yakimenko et al. 2025).

Within EPT, mayflies were selected as the primary focus for museum review because Canterbury Museum held the largest proportion of BOLD-barcoded material, and staff and research associates were available to assist with specimen access and metadata checks. This voucher specimen review provided an essential quality-control step, demonstrating that sequence availability alone does not guarantee the reliability of a DNA barcode reference library.

Of the 171 BOLD-barcoded mayfly specimens held at Canterbury Museum, 138 (81%) were examined during the visit. Taxonomic issues were common: 76 specimens (55%) required correction, mostly at species level. Most corrections related to individual specimens rather than whole Operational Taxonomic Units (OTUs), but several OTU-level problems were also detected. This demonstrates that even reference records associated with museum-held voucher material can require expert reassessment before they are used to support species-level eDNA interpretation.

Misidentifications often involved closely related congeners, undescribed taxa and cryptic species that were unlikely to have been recognised when the original barcoding was undertaken. The mayfly example in Figure 4-1 illustrates the issue where *Deleatidium* cf. *fumosum* (KX038358) is recorded as *Atalophlebioides cromwelli*. In addition, subsequent quality assurance should have identified that *Atalophlebioides* sp. (KX038081) refers to a monospecific genus, making the “sp.” designation redundant and unnecessarily ambiguous. This is not simply a historical or taxonomic detail. If an incorrect name is attached to a reference sequence, subsequent eDNA assignments can generate a plausible-looking but erroneous species detection. Conversely, unresolved or ambiguous names can force results to be reported at a coarser taxonomic level, reducing their value for biodiversity reporting, restoration assessment and threatened-species surveillance.

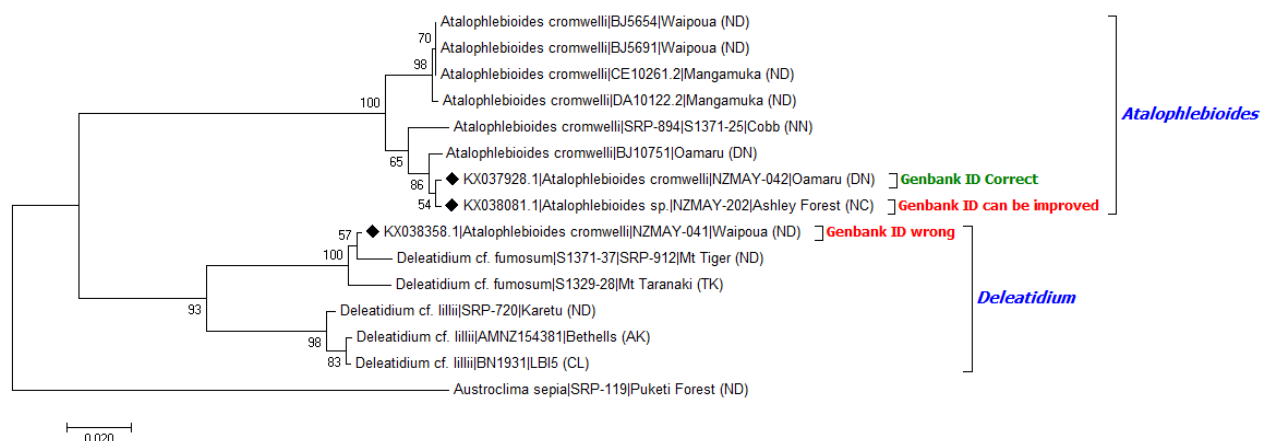


Figure 4-1: Examples of species misidentification and unnecessary taxonomic ambiguity within public COI reference records for Leptophlebiidae mayflies, illustrating how reference library errors can reduce confidence in eDNA species assignments.

Examples found during the review included errors at species, genus and family level. Notable cases included *Ameletopsis* (Ameletopsidae) misidentified as *Rallidens* (Rallidentidae), *Deleatidium* (Leptophlebiidae) misidentified as *Nesameletus* (Nesameletidae), and, as shown in Figure 4-2, *Austroclima* (Leptophlebiidae) was misidentified as *Siphlaenigma* (Siphlaenigmatidae).

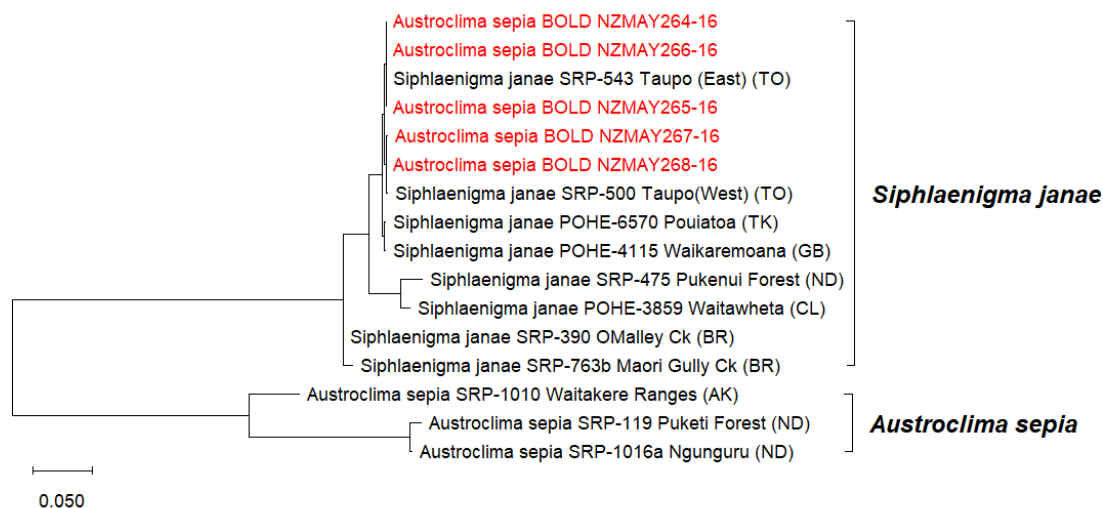


Figure 4-2: Example of a family-level taxonomic misidentification in a public reference sequence dataset. The leptophlebiid mayfly *Austroclima sepia* has been misidentified as *Siphlaenigma janae* from the family Siphlaenigmatidae. The misidentifications are shown in red.

These examples show that errors can occur even among ecologically important and well-studied taxa, and that the consequences may be substantial where eDNA data are used to infer species presence, distribution, conservation status or ecological condition.

Given the proportion of corrections required, the remaining unreviewed voucher material should be reassessed, and affected reference records should be corrected, annotated or quality-flagged before being used uncritically in council reporting.

Misidentified reference sequences are particularly problematic because they can generate false species detections, rather than simply failing to detect a species. For this reason, taxonomic verification remains indispensable for interpreting eDNA data. As emphasised by Buckley (2025), the future of entomological taxonomy in New Zealand depends on sustained investment in expert identification, curation, and revisionary work, capacities that cannot be replaced by molecular tools alone. Similarly, Löbl et al. (2023) argue that the global decline in taxonomic expertise represents a “silent extinction” that threatens biodiversity science itself. Both papers caution against the assumption that eDNA and automated pipelines can substitute for taxonomists, but instead, molecular approaches increase the need for expert-vetted identifications, curated voucher collections, and specialist oversight.

4.4.1 Outcome for regional councils

The mayfly voucher review showed that even comparatively well-studied, well-vouchered and taxonomically informative groups can contain substantial identification and metadata problems in public reference databases. The examples in Figure 4–1 and Figure 4–2 illustrates the practical risk where a sequence labelled as one mayfly taxon, or left unnecessarily unresolved, can propagate through eDNA pipelines as either a plausible but erroneous species detection or larger scale family errors with potentially far-reaching consequences. Councils should therefore treat species-level eDNA detections linked to unverified, disputed or taxonomically unresolved reference records as provisional, particularly where results are being used for freshwater planning, consent monitoring, biodiversity reporting or threatened-species management. This reinforces the need for stronger engagement with regional taxonomic expertise and highlights the ongoing importance of taxonomic verification in molecular biodiversity research.

4.4.2 Recommendations

Regional councils should use the mayfly example as a model for reference library quality assurance. Database records should not be accepted as reliable solely because they are public, barcoded or associated with a voucher. Practical actions should be prioritised as follows:

1. require eDNA providers to state whether species-level calls are supported by verified, voucher-backed and taxonomically current reference sequences,
2. undertake targeted re-examination of voucher specimens for high-use or high-consequence taxa,
3. correct and annotate misidentified records before they are used in council interpretation,
4. apply quality flags to uncertain, disputed or taxonomically unresolved names,

5. strengthen engagement with regional taxonomic expertise; and
6. avoid treating unsupported assignments as definitive evidence of species presence or absence in high-consequence applications.

4.5 Metadata quality

The PCE (2026) report highlighted that environmental information often lacks standardisation, with inconsistent formats, missing fields and incompatible structures across organisations. These same issues are evident in molecular reference datasets, where metadata such as voucher information, locality, collector, sequencing method and taxonomic authority are frequently incomplete or absent. As the Commissioner noted, poor standardisation reduces usability and increases the risk of misinterpretation. Improving metadata completeness is therefore essential for defensible eDNA interpretation.

A genetic sequence is only as useful as the information attached to it. For eDNA monitoring, a COI barcode can confirm that a species has been sequenced, but without reliable information on where, when, from what specimen, and by whom that material was collected and identified, its value for freshwater management is greatly reduced. While poorly documented records may still support species discrimination, they provide limited evidence for local occurrence, regional population structure, range change, conservation status, or the effectiveness of restoration and management actions.

The quality and completeness of the metadata associated with COI sequences was assessed for each of the 395 freshwater invertebrate species publicly databased and summarised by major taxonomic group in Table 4–6. For a sequence to be recognised as acceptable quality it needs to have, at a minimum: a location, coordinates and date recorded. The following qualifiers were assigned to each species based on metadata completeness; Low 0–40%, Moderate 41–70%, High 71–100%.

4.5.1 Metadata completeness

Sequence metadata completeness varied substantially across the 395 species, revealing strong biases in data curation among major groups. Ephemeroptera exhibited the highest metadata quality, with 80% of species possessing fully complete records, while Trichoptera showed moderate completeness overall. In contrast, Plecoptera, Coleoptera, and Crustacea were dominated by low-quality metadata, with Mollusca almost approaching complete absence of key information.

Across EPT taxa, fewer than 47% of assessed records had high-quality metadata, mainly because Plecoptera records were consistently poorly documented despite stronger metadata completeness for Ephemeroptera. Overall, only 42% of assessed freshwater invertebrate species had high-quality metadata, leaving substantial gaps for the remaining 58%, particularly among non-insect taxa.

Of the 138 mayfly voucher specimens examined at the Canterbury Museum, which are linked to BOLD (see Section 4.4), 107 specimens (76%), comprising 255 unique metadata qualifiers across the reference dataset needed updating or correcting.

Despite identifying metadata gaps in the material examined, we also report substantial metadata gaps in publicly available EPT databases in general. For example, 42% of mayfly records, 67% of stonefly records and 51% of caddisfly records contain incomplete collection metadata. These missing data we report now need to be incorporated into the BOLD system, and many hundreds, if not thousands of specimen entries still contain empty metadata fields requiring completion.

Table 4–6: Metadata completeness for public sequences, summarised by major taxonomic group (395 species), showing the extent to which supporting information enhances confidence in reference library records. EPT = Ephemeroptera, Plecoptera, Trichoptera combined; “other” insects = Hemiptera, Lepidoptera, Mecoptera, Megaloptera, Odonata; Crustacea = amphipods, crayfish, isopods, shrimps, zooplankton; Mollusca = bivalves, limpets, snails; “worm-like” = flat worms, Gordian worms, nematodes, round worms.

Major taxonomic groups	Metadata completeness (%)		
	Low (0–40%)	Moderate (41–70%)	High (71–100%)
EPT combined	27.5	25.8	46.7
Ephemeroptera	11.4	8.6	80
Plecoptera	59.3	19.8	21.0
Trichoptera	16.8	31.6	51.6
Coleoptera	46.7	20	33.3
“other” insects	38.1	23.8	38.1
Crustacea	78.6	0	21.4
Mollusca	96.8	0	3.2
“worm-like”	37.5	0	62.5
All groups	36.2	22.0	41.8

International reviews show that reference library reliability depends on more than the presence of a DNA sequence. Weigand et al. (2019) emphasised that barcode reference libraries used for aquatic biomonitoring require explicit quality assurance and quality control, including traceable specimen identifications, collection data, geographic context and linkage to voucher material. Buckner et al. (2021) similarly argued that vouchers provide the physical evidence required to verify taxonomic identity, support reproducibility, document legality and enable future taxonomic revision. For New Zealand freshwater invertebrates, this means metadata, vouchers and specimen–sequence linkage should be treated as core components of eDNA reference library fitness.

A record that lacks locality information, collection date, identifier, voucher number or a clear link between the specimen and sequence may still be technically useful, but it should not be treated as equivalent to a well-documented, voucher-backed record for council reporting or high-consequence management decisions. Essentially, this means metadata quality assurance is simply not about filling in the blank fields. Quality control of metadata completeness determines whether eDNA data can be validated, reused, published, reconciled across systems, and trusted for regional council decisions.

Recent New Zealand analyses reinforce that metadata limitations are a practical constraint on eDNA interpretation, not simply a database management issue. Buckley et al. (2026) found that New Zealand insect barcode data are taxonomically, geographically and elevationally biased, with many records lacking species-level identification, explicit identification-method information, or reliable georeferencing. In addition, they also recommended that barcoded specimens be accessioned into public biological collections with accessible digitised (and complete) metadata. Similarly, the eDNABridge pilot demonstrated that large-scale publication of New Zealand eDNA data to GBIF is technically feasible, but depends on validation, consistent Darwin Core mapping, clear data ownership, project-level metadata, and reconciliation between laboratory taxonomies and global biodiversity infrastructure (Hards and Muellner 2026). Together, these studies show that for regional council eDNA programmes, metadata quality must extend beyond locality and date fields to include identification method, voucher status, data provenance, taxonomic authority, ownership, permissions, and where relevant, local context and data-sovereignty considerations.

These metadata deficiencies also affect interpretation of the patterns summarised above. Incomplete or incorrect locality data can mislead range interpretation, create false assumptions about regional occurrence, and make it difficult to determine whether an eDNA detection is biologically plausible or a database artefact. Future reference library development should therefore require minimum metadata standards, including verified locality, coordinates, collection date, collector, identifier, voucher details and a clear connection between each sequence and the specimen examined. Existing records with unresolved metadata issues should be corrected where possible or retained with explicit quality flags so they are not used uncritically in regulatory interpretation, biodiversity reporting or conservation management.

4.5.2 Outcome for regional councils

Metadata quality should be treated as quality assurance component of eDNA interpretation, not as a secondary database attribute 'to (never) be filled out later'. Where records lack complete locality, coordinates, collection date, collector, identifier, voucher details or specimen to sequence linkage, councils have reduced ability to assess whether a species-level assignment is taxonomically reliable, geographically relevant or appropriate for management use. This is particularly important where eDNA outputs are used for freshwater planning, consent monitoring, restoration assessment, biodiversity reporting or threatened-species management. This essentially means metadata gaps affect both taxonomic confidence and data reuse.

Councils may receive eDNA results that are technically valid within a provider's pipeline but difficult to compare across regions, upload to national or international repositories, reconcile with GBIF taxonomy, or defend in regulatory settings if specimen, sequence, project and ownership metadata are incomplete.

4.5.3 Recommendations

Regional councils should require eDNA providers to supply both sequence-level and dataset-level metadata, including identification method, voucher status, collection locality, coordinates, collection date, taxonomic authority, reference database, data owner, permissions and intended data-use conditions.

Where datasets are intended for publication through GBIF or similar infrastructure, councils should ensure Darwin Core-compatible metadata are captured at the point of sampling rather than reconstructed later. For existing New Zealand freshwater invertebrate records, councils should support efforts to resolve metadata gaps where records are relevant to their monitoring programmes and ensure that metadata associated with council-funded or council-collected datasets are complete and fit for purpose. Councils should also support publication of verified institutional and private holdings with complete metadata and encourage the use of quality flags for records with missing or uncertain provenance.

For high-consequence detections (e.g., Nationally Critical species or biosecurity detections) or non-detections, councils should request supporting information from eDNA service providers, data generators, or the custodians of the reference datasets used for taxonomic assignment. This information should demonstrate that the underlying reference sequences are voucher-backed, taxonomically verified and geographically appropriate before results are used as evidence for regulatory, biodiversity or threatened-species decisions.

4.6 Geographical coverage of sequences

The completeness and quality of molecular reference databases are known to be geographically biased or not spatially representative of a species' distribution (Keck et al. 2023). Although a sequence may be technically correct, this does not necessarily imply there is adequate geographical coverage of genetic diversity.

Geographic coverage of publicly available COI sequences was assessed for each of the 395 freshwater invertebrate species with public data and is summarised by major taxonomic group in Table 4–7. Coverage was evaluated using published distribution records, private or unpublished database records and the authors' expert knowledge of regions known or likely to contain distinct genetic lineages or regional haplotypes. Each species was assigned to one of three categories; Low, Moderate or High, to provide the best current estimate of COI representation across its known range.

For example, the widespread North Island endemic mayfly *Acanthophlebia cruentata* was assigned 'High' coverage because reference sequences are available from across a broad geographic range (Trewick et al. 2022).

In contrast, the caddisfly *Zelolessica cheira* was assigned ‘Low’ coverage because only a single public sequence is available despite the species being widespread throughout the North and South Islands (Henderson & Ward 2007). Some species were classified as ‘Low’ because the locality metadata were too sparse or incomplete to determine whether the available COI records suitably captured their geographic range. In other cases, species with published COI data were still assigned ‘Low’ because the sequences themselves were too short or otherwise unsuitable for assessing geographic representation or supporting reliable eDNA interpretation. Together, these issues mean that “Low” can reflect either poor metadata or insufficient sequence quality, even when sequences exist.

Table 4–7: Geographic representativeness of public reference sequences among major freshwater invertebrate groups (395 species), indicating the extent to which available records capture species distributions and genetic variation. * = EPT; “other” insects = Hemiptera, Lepidoptera, Mecoptera, Megaloptera, Odonata; Crustacea = amphipods, crayfish, isopods, shrimps, zooplankton; Mollusca = bivalves, limpets, snails; “worm-like” = flat worms, Gordian worms, nematodes, round worms. Geographic coverage was evaluated using published distribution records, private or unpublished database records and the authors’ expert knowledge.

Major taxonomic groups	Geographical coverage of public sequences (%)		
	Low	Moderate	High
Ephemeroptera*	71.4	14.3	14.3
Plecoptera*	43.2	42.0	14.8
Trichoptera*	63.2	31.1	5.8
Coleoptera	86.7	6.7	6.7
“other” insects	85.7	4.8	9.5
Crustacea	57.1	21.4	21.4
Mollusca	71.0	29.0	0.0
“worm-like”	87.5	12.5	0.0
All groups	62.8	28.6	8.6

Table 4–7 shows that geographic representation is one of the weakest components of the current public COI reference dataset. Across all groups, 62.8% of species with public sequences were assessed as having Low geographic coverage, while only 8.6% had High coverage. This means that the majority of public records provide evidence that a species has been sequenced, but do not provide strong evidence that the available sequence data represent the genetic variation likely to occur across that species’ full distribution.

The result is especially important because geographic under-representation can produce the same practical outcome as poor taxonomic or metadata quality, whereby species-level assignments become less defensible, apparent non-detections become harder to interpret, and local or regional lineages may remain invisible to routine eDNA analysis.

The pattern is not uniform among groups. Plecoptera had the strongest relative representation, with 42.0% of species in the Moderate category and 14.8% in the High category, suggesting that some stonefly lineages are comparatively well represented across their known distributions. However, this still leaves more than 40% of sequenced Plecoptera in the Low category. Ephemeroptera and Trichoptera, despite being central to biomonitoring and better represented in public barcode datasets than many non-EPT groups, were dominated by Low coverage records (71.4% and 63.2%, respectively). This indicates that even where species-level COI barcodes exist, many EPT records are likely to be geographically shallow and may not capture regional haplotypes, island differences, upland-lowland variation, or disjunct distributions as commonly occurs across many freshwater invertebrates (e.g., McCulloch et al. 2010).

The non-EPT groups show an even stronger spatial-bias signal. Coleoptera, “other” aquatic insects and “worm-like” taxa all had more than 85% of assessed species in the Low coverage category, while Mollusca had no species classed as High and 71.0% classed as Low. Crustacea was the only non-EPT group with a relatively high proportion of High coverage records (21.4%), but more than half of its sequenced species still had Low coverage. These results suggest that current public reference libraries are not only taxonomically incomplete, but also spatially uneven within the taxa that are present. For groups with naturally fragmented distributions, low dispersal capacity or strong habitat specialisation, this lack of spatial representation can limit confidence in species-level eDNA interpretation.

Of the data available, 16 species with publicly available COI data were represented only by sequences (full and/or short) from outside New Zealand, including Australia, Japan and Indonesia. Notably, the Nationally Endangered tadpole shrimp *Lepidurus apus viridis* and four Data Deficient species (three beetles and one snail) were represented solely by Australian sequences. Similarly, six dragonfly species considered vagrants or recent colonisers of New Zealand were represented by sequences from Australia and more distant locations.

The dominance of Low coverage records also means that species occurrence cannot be validated confidently in many regional contexts. A sequence from one part of New Zealand, or from overseas, may be adequate for recognising a broad taxonomic match, but may not be adequate for confirming the occurrence of a local population, distinguishing close congeners, resolving cryptic diversity, or excluding the possibility of an unsampled regional lineage. This problem is most acute for regionally restricted, disjunct, high-elevation, groundwater, spring-fed, island and headwater populations, where local haplotypes may be poorly represented or absent from public databases. In these circumstances, eDNA outputs should be interpreted as reference-limited biological evidence rather than as definitive species-level confirmation.

This highlights a broader limitation of relying on geographically distant reference sequences for species identification, as such sequences may not adequately capture the genetic diversity present within local populations. This consideration is particularly important for highly dispersive or migratory taxa, including many odonate species. While long-distance movement can result in widespread haplotypes across large geographic areas, isolated populations may also retain distinct mitochondrial lineages. The globally-distributed dragonfly *Pantala flavescens*, which is an occasional visitor to New Zealand, provides a useful example. Studies have demonstrated extensive gene flow among migratory populations, while also identifying geographically isolated, non-migratory populations with greater genetic differentiation (Low et al. 2017).

Over half of the freshwater invertebrates listed in Grainger et al. (2018) are of conservation interest, and only 20% of these are represented by full-length COI barcode sequences (see Section 4.3), despite many having restricted distributions and habitat preferences. This suggests unique haplotypes would be majorly underrepresented, potentially missing the ability to detect cryptic, rare or threatened species.

These findings illustrate that patterns of population connectivity can vary markedly across a species' range and caution against assuming that a geographically distant COI sequence is representative of a local population. Consequently, the presence of a species in a reference database such as BOLD does not necessarily mean that the available sequence accurately reflects the genetic variation present in New Zealand populations, potentially reducing confidence in species-level identifications based on a single geographically distant reference sequence (Buckley et al. 2026).

4.6.1 Outcome for regional councils

Geographic coverage should be treated as a core component of reference library confidence. The results in Table 4–7 indicate that many species-level eDNA detections are constrained not only by whether a species has a sequence, but by whether that sequence is spatially relevant to the catchment, region or habitat being monitored. Current New Zealand reference datasets contain significant spatial biases, which can limit species-level identification and reduce confidence in regional biodiversity assessments. For regional councils, this means that detections and non-detections for taxa with Low geographic coverage should be interpreted cautiously, especially where they relate to biodiversity reporting, consent monitoring, freshwater planning, restoration outcomes, threatened species, or locally endemic fauna. Councils will obtain the greatest long-term benefit from eDNA if routine monitoring is linked to targeted reference library improvement that identifies spatial gaps, directs new collecting effort to under-sampled regions and habitats, and increases geographic representation across priority taxa.

4.6.2 Recommendations

Regional councils should support a coordinated programme of targeted collection, expert identification, vouchering and sequencing to fill spatial gaps in the freshwater invertebrate reference library alongside routine eDNA monitoring.

Priority should be given to taxa with public sequences from few localities, taxa represented only by overseas or geographically distant records, conservation-priority species, regionally endemic species, and habitats likely to contain distinct genetic lineages, including groundwater, springs, wetlands, alpine streams, islands, headwaters and poorly sampled catchments. Expanding reference collections across under-sampled regions, habitats and locally important taxa will improve the accuracy, consistency and management value of eDNA-based monitoring programmes. New reference material should be collected with complete locality, coordinate, date, habitat, collector, identifier and voucher metadata, and should be published or curated in a way that allows future quality assurance. This spatially explicit approach would move the reference library from a simple list of sequenced taxa toward a defensible, robust and regionally representative database capable of supporting high-confidence eDNA interpretation across New Zealand.

4.7 Investigation of additional markers

Environmental DNA (eDNA) monitoring in New Zealand has so far relied almost entirely on the COI gene region, which has become the standard marker for metabarcoding because of its broad taxonomic utility and widespread representation in global reference databases (Folmer et al. 1994; Hebert et al. 2003; Leese et al. 2021). COI has therefore been the focus of this assessment. However, COI does not amplify consistently across all freshwater invertebrate groups (Elbrecht & Leese 2017; Takenaka et al. 2024), and several taxa of ecological or conservation importance remain difficult to detect or resolve using COI alone. This was particularly evident for the 23.5% of non-insect species assessed that are poorly represented in current COI public reference datasets.

These limitations have prompted increasing interest in whether additional genetic markers could strengthen detection performance, improve taxonomic resolution, or increase confidence in species-level assignments. The objective of this assessment was not to identify a replacement for COI, but to evaluate whether supplementary markers, such as 12S, 16S and CytB, could enhance eDNA monitoring when used alongside COI.

To address this, in addition to COI, publicly available reference sequences of other genes commonly associated with the 675 freshwater invertebrate species listed by Grainger et al. (2018) were reviewed. The evaluation considered the availability, length and taxonomic coverage of the alternative markers, and examined recent developments in primer design. This approach allowed assessment of whether additional markers could help address known COI amplification gaps and support more reliable eDNA interpretation across the freshwater fauna.

The use of eDNA for freshwater monitoring in New Zealand has gained considerable momentum since around 2020, as freshwater managers increasingly recognised its potential and broad range of applications. Nevertheless, the adoption of commercial eDNA tools for freshwater invertebrate monitoring may have outpaced scientific validation, with low detection rates reported and limited formal validation or quality assurance of the online reference data used to support species detections. Early development of eDNA in New Zealand focused on fish detection using multiple gene regions and primer sets, with a single “universal” COI primer added to provide broader

taxonomic coverage for invertebrates. Commercial assays have since expanded, with “basic” and “comprehensive” testing options now available, with the latter having specifically-designed markers for challenging fish (i.e., bullies), high-profile invasive species (i.e., gold clams) and native mussels. Despite these developments, methods continue to rely on a single COI marker across both test options for insects, and independent validation of freshwater invertebrate detections remains limited.

4.7.1 Reference sequence coverage of alternative gene regions

Six gene regions were regularly encountered on GenBank associated with publicly available freshwater invertebrate sequences, representing the full range of taxonomic groups assessed (Table 4–8). These included slow to moderately-evolving nuclear genes (H3, 18S, 28S) and fast-evolving mitochondrial genes (12S, 16S, CytB). Although the nuclear markers are unlikely to be useful for species-level eDNA identification, they were included for completeness and to provide an overview of existing sequence resources.

Coverage of the 675 species assessed was generally low for all six gene regions, ranging from 3.6% to 16.3% (Table 4–8). Despite this, several notable patterns emerged. Plecoptera were particularly well-represented for H3 and 18S, with 43% and 51% of stonefly species having publicly available sequences, reflecting widespread use of these markers in phylogenetic research (e.g., McCulloch et al. 2017). Some insect groups with relatively few species (“other insects”) had comparatively high sequence coverage for 16S and 28S (37% and 35%, respectively). Molluscs also exhibited strong representation for 16S, with sequence data available for around 40% of species. “Worm-like” taxa (annelids, flatworms, Gordian worms, nematodes) had reasonable coverage of 18S and 28S (approximately 40% and 30%, respectively).

Table 4–8: Distribution of publicly available sequences of major freshwater invertebrate taxonomic groups for different gene regions other than COI. * = EPT; “other” insects = Hemiptera, Lepidoptera, Mecoptera, Megaloptera, Odonata; Crustacea = amphipods, crayfish, isopods, shrimps, zooplankton; Mollusca = bivalves, limpets, snails; “worm-like” = flat worms, Gordian worms, nematodes, round worms; conservation interest species are those recognised under the NZTCS as Threatened, At Risk or Data Deficient.

Major taxonomic groups	No. species	H3 (%)	12S (%)	16S (%)	18S (%)	28S (%)	CytB (%)
Ephemeroptera*	59	22.0	10.2	11.9	11.9	18.6	3.4
Plecoptera*	107	43.0	9.3	4.7	51.4	11.2	0.9
Trichoptera*	242	1.2	1.2	7.0	1.7	2.9	1.2
Coleoptera	62	16.1	9.7	11.3	12.9	9.7	8.1
“other” insects	46	15.2	13.0	37.0	17.4	34.8	10.9
Crustacea	57	5.3	7.0	12.3	7.0	10.5	1.8
Mollusca	79	16.5	6.3	39.2	19.0	16.5	5.1

Major taxonomic groups	No. species	H3 (%)	12S (%)	16S (%)	18S (%)	28S (%)	CytB (%)
"worm-like"	23	4.3	4.3	13.0	39.1	30.4	13.0
Conservation interest	355	10.1	2.5	9.9	10.4	5.6	1.4
All groups	675	14.2	6.1	13.9	16.3	11.6	3.6

Although overall coverage was low, these patterns indicate that several taxonomic groups already have comparatively strong representation for particular gene regions. This suggests opportunities already exist to utilise some sequences for other gene markers, particularly molluscs, crustaceans and some non-EPT taxa, if additional marker development is adopted. As Melchior & Baker (2023) noted, unidentified sequences generated during eDNA analyses often reflect gaps in reference libraries or marker performance rather than true ecological absence, reinforcing the value of expanding marker coverage where promising foundations already exist.

4.7.2 Implications for marker choice and multi-assay eDNA approaches

The uneven performance of COI and limited availability of alternative markers have several implications for freshwater biomonitoring, invasive species surveillance and threatened-species detection. Systematic under-detection may occur for taxa with poor COI amplification.

Marker choice directly influences taxonomic resolution, with some groups likely to remain identifiable only to genus or family until additional markers are sequenced. False negatives may arise when taxa are present but not amplified by COI, particularly in mixed communities dominated by strongly amplifying species. Conversely, false positives or over-confident assignments may occur if alternative markers are used without adequate reference coverage or quality control.

Globally, considerable focus has been placed on developing robust genus- and species-specific markers across many branches of biodiversity (Dawkins et al. 2025; Khan et al. 2026; Prabowo et al. 2026), including New Zealand, where a range-restricted stonefly has been targeted for assay development (McCulloch et al. 2025). There has also been growing interest in assay panels capable of detecting multiple taxa within broader freshwater communities (Chen et al. 2025; Takenaka et al. 2026; Zhang et al. 2026). Of particular relevance, assay panels using novel primer sets, including MiDeca-16S, MtlInsects-16S, MtlInsects-12S and AQdb-16S, have been developed to improve DNA barcoding of aquatic invertebrate communities (Komai et al. 2019; Takenaka et al. 2023) by targeting a range of fragment regions and fragment lengths (Figure 4–3).

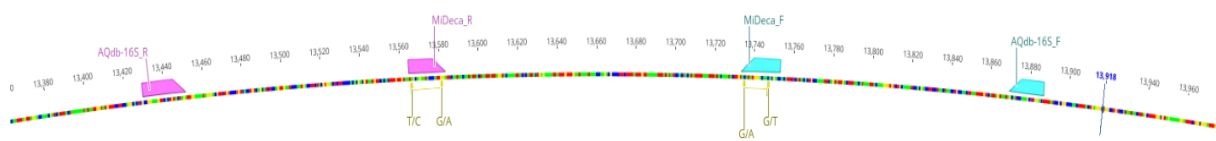


Figure 4-3: Visualisation of a fragment of mayfly DNA showing the positions of the AQdb-16S primers (Takenaka et al. 2023) that amplify longer 16S DNA fragments useful for reference sequences, and the MiDeca primers (Komai et al. 2019) that amplify a shorter fragment designed for eDNA barcoding of decapods (Crustacea). Source: Jackson Treece, University of Otago.

Multi-assay panels have the potential to enhance the accuracy and taxonomic breadth of eDNA-based biodiversity assessments by improving detection across multiple taxonomic levels. However, enhancing eDNA accuracy with assay panels comes with sizeable trade-offs. Achieving reliable species-level resolution across diverse freshwater invertebrate communities typically requires multiple markers and more complex analytical workflows, increasing laboratory effort, bioinformatic processing requirements and overall costs. Consequently, assay design must balance taxonomic coverage, detection accuracy, operational complexity and cost-effectiveness.

Collectively, these findings indicate that although COI remains the most practical and widely supported marker for freshwater invertebrate eDNA monitoring in New Zealand, its performance is unlikely to be optimal across all taxonomic groups. Several alternative gene regions already have useful reference-sequence resources for particular taxa, especially 16S for molluscs and some crustaceans, and H3 and 18S for selected insect groups. However, no alternative marker currently possesses sufficient taxonomic coverage or reference-library support to replace COI as the primary marker for freshwater invertebrate monitoring. Instead, future improvements in detection and taxonomic resolution are more likely to come from targeted multi-marker approaches that combine the strengths of different gene regions.

4.7.3 Outcome for regional councils

For councils, the main outcome of this assessment is that current eDNA results must be interpreted with explicit awareness of marker limitations. Reliance on a single universal COI marker provides strong coverage for many biomonitoring taxa, but its weaknesses for crustaceans, molluscs, beetles and worm-like groups mean that absence of detection cannot be assumed to indicate ecological absence. Councils should expect variable detection performance across taxa and require providers to report marker-specific confidence, including known amplification biases and taxa with poor COI performance.

The evaluation of additional gene regions showed that although overall sequence coverage remains limited across all alternative markers, several taxonomic groups have comparatively strong representation for 16S, 18S, 28S and H3.

These patterns indicate that a single-marker approach is unlikely to provide optimal performance across the full freshwater invertebrate fauna. In particular, taxa that are poorly served by COI, such as molluscs, crustaceans and some non-EPT insect groups, already have partial foundations for targeted marker development. This suggests that future improvements in freshwater invertebrate eDNA monitoring are likely to come from multi-marker assay panels designed specifically for New Zealand taxa, rather than further reliance on COI alone.

Developing multi-marker approaches offers the potential to improve species detection, increase confidence in taxonomic assignments, and support future applications such as an environmental DNA based Macroinvertebrate Community Index (eMCI). An opportunity already exists to accelerate this work through the authors' research, client projects and the Envirolink Tools Grant (2024) *Stocking New Zealand's molecular library for eDNA monitoring*, where DNA extracts have been generated for hundreds of New Zealand freshwater invertebrate species. These extracts are currently held in storage and could be used to fast-track the expansion of an invertebrate sequence library for new gene regions.

However, councils should recognise that increased confidence will involve trade-offs. Multi-marker approaches require additional laboratory processing, more complex bioinformatic analysis, expanded reference libraries and greater investment in validation and quality assurance. Maximising taxonomic coverage, species-level resolution and affordability may not be simultaneously achievable. In some situations, assays targeting indicator taxa lists or ecologically informative groups (such as EPT) may provide a more practical and cost-effective monitoring solution than attempting comprehensive species-level detection across entire freshwater invertebrate communities.

4.7.4 Recommendations

To improve confidence in freshwater invertebrate eDNA monitoring in New Zealand, the following actions are recommended:

1. investigate and formally validate additional markers for freshwater invertebrate monitoring, prioritising 16S and CytB, which show promising taxonomic coverage for groups poorly served by COI (e.g., molluscs, crustaceans, selected insect orders),
2. evaluate the suitability of multi-marker assay panels for New Zealand freshwater invertebrate communities, drawing on emerging international developments such as MiDeca-16S, MtInsects-16S, MtInsects-12S and AQdb-16S, recognising their potential to improve amplification consistency and taxonomic resolution,
3. develop and test freshwater invertebrate-specific assay panels for the New Zealand fauna, with priority given to taxa of biomonitoring, biodiversity and conservation significance,
4. strengthen collaboration across ecological, taxonomic, molecular and commercial sectors to ensure that new markers and assay panels are

robustly designed, independently validated and fit for council monitoring programmes, addressing current gaps in external quality assurance,

5. expand national reference sequence resources for alternative gene regions, and prioritise taxa with promising partial coverage (e.g., molluscs, crustaceans, “other insects”, worm-like taxa) to support future multi-marker assay development and reduce false negatives,
6. working with molecular specialists, commence generation of additional sequences with new COI primers or for new gene regions (i.e., 16S) using existing DNA extract generated through previous projects and the 2024 Envirolink Tools Grant,
7. evaluate targeted monitoring approaches for indicator taxa, recognising that focused detection of ecologically informative groups (e.g., EPT) may provide greater value and affordability than attempting comprehensive species-level detection across entire freshwater invertebrate communities, and
8. require marker-specific confidence reporting from providers, to ensure that councils receive clear information on amplification biases, reference coverage and taxonomic limitations for each marker used, improving defensibility of eDNA-based decisions.

5 Where to from here?

National environmental-information reforms reinforce the need for coordinated molecular reference resources. The Government's recent investment in environmental data infrastructure, described by the PCE (2026), aims to support federated, interoperable systems that reduce duplication and improve data accessibility. A curated national freshwater invertebrate molecular reference library would align with these reforms by providing a standardised, authoritative dataset that can be integrated into emerging digital platforms and used consistently across councils.

This should be treated as a shared priority across organisations involved in biodiversity monitoring, conservation research, environmental management and eDNA development, including regional councils, DOC, PROs, museums, taxonomic specialists, eDNA providers and other biodiversity-data partners. In practice, a federated model would allow these organisations to retain stewardship of their own datasets while contributing to a shared framework of standards, metadata requirements, quality codes and identifiers, enabling information from multiple sources to be discovered, linked and interpreted consistently across New Zealand.

The critical gaps identified in this assessment fall into three overlapping categories: taxonomic gaps, data-quality gaps and implementation gaps. This structure is useful because it separates the absence of reference data from the reliability of existing records, and from the way results are reported or used by regional councils. A taxon may be represented in GenBank, BOLD, GEOME or a private database, but still be unsuitable for high-confidence interpretation if the sequence is shorter, taxonomically uncertain, geographically distant, poorly documented, unsupported by a voucher specimen or affected by unresolved nomenclature. The practical priority is therefore to distinguish sequence presence from sequence utility.

Taxonomic gaps include the 280 freshwater invertebrate species with no public GenBank sequence coverage, conservation-priority taxa without usable barcodes, and poorly represented groups such as molluscs, crustaceans, aquatic beetles, odonates, zooplankton and worm-like taxa. Within EPT, priority should be given to nationally threatened, At Risk, Data Deficient, range-restricted, undescribed and locally significant species that remain absent from public reference libraries or have inadequate barcode coverage. Particular attention is needed for taxa associated with under-sampled habitats, including headwaters, alpine streams, seepages, springs, waterfalls, groundwater systems, islands and poorly sampled catchments.

Data-quality gaps include shorter sequences, suspected misidentifications, unresolved names, incomplete metadata, weak voucher support, overseas-only records and low geographic representation. At a minimum, quality-control categories should record mitochondrial marker used, sequence length, taxonomic identification confidence, voucher status, metadata completeness, database source, geographic representativeness and whether the record has been expert reviewed.

High-confidence records should be full-length or otherwise demonstrably diagnostic, linked to a voucher, taxonomically current, supported by complete metadata and geographically relevant to the region or habitat being assessed. Lower-confidence records may still be useful but should be flagged so that results are reported at the most defensible taxonomic level.

Implementation gaps arise when reference library limitations are not communicated clearly to councils or incorporated into procurement, reporting and interpretation. eDNA suppliers should identify the database used, mitochondrial marker used, assignment threshold, confidence basis for each taxonomic call and any known limitations affecting interpretation. Council reports should clearly separate high-confidence detections from provisional assignments, identify where non-detections are reference-limited, and avoid treating unsupported species-level results as definitive evidence of presence or absence. High-consequence uses, including invasive species detection, threatened-species management, consent monitoring, freshwater planning and restoration assessment, should combine eDNA results with ecological context, conventional survey data and expert taxonomic review.

Priority implementation actions are in perceived order of importance:

1. publish and curate verified, New Zealand, voucher-backed sequences,
2. correct or annotate problematic public records,
3. contribute to a shared, versioned and quality-coded national freshwater invertebrate reference set,
4. prioritise sequencing of threatened, regionally endemic and poorly represented taxa,
5. capture complete metadata at the point of collection, and
6. interpret results at the taxonomic level supported by the evidence.

This staged approach delivers immediate gains from existing data while directing new collecting and sequencing toward the taxa, habitats and regions where improved reference coverage will most increase management confidence.

Murray et al. (2024) highlighted the importance of developing eDNA reference libraries to support regional management of conservation-priority species, as required under the National Policy Statement for Freshwater Management (NPS-FM). Their analysis highlights key research themes such as “Detection and outcome monitoring” and “Determining the full range of biodiversity requiring management,” both of which rely on robust molecular reference data to enable accurate species identification, address data deficiencies, and enhance conservation assessments. eDNA can also support assessment of threatened species and help councils delineate threatened-species habitats, as relevant to section 3.8(3)(c) of the NPS-FM (2020). Developing comprehensive freshwater invertebrate reference libraries therefore directly supports regional councils in fulfilling their monitoring and management responsibilities under the NPS-FM.

Aligning council-led efforts to improve reference libraries with these priorities can help ensure efficient use of resources, reduce duplication, and contribute to more effective conservation outcomes and invasive species surveillance at the regional level. Essentially, strengthening freshwater invertebrate reference libraries therefore directly supports regional councils in fulfilling their monitoring and management responsibilities under the NPS-FM.

While regional councils are key beneficiaries and contributors, many of the actions outlined below cannot be led by councils alone. Delivering a nationally consistent and defensible freshwater invertebrate reference library will require coordinated effort across DOC, Public Research Organisations, museums, taxonomic specialists, eDNA service providers, and existing national groups such as SWIM and the New Zealand eDNA Development Group. The pathway therefore represents a shared national direction rather than a council-only work programme, with councils playing a central role in prioritisation, commissioning, and ensuring that provider reporting meets confidence requirements.

A practical implementation pathway, numbered in perceived order of importance, is recommended:

1. **Immediate quality triage:** classify existing public, private and institutional records by sequence length, taxonomic verification, voucher support, metadata completeness and geographic coverage so councils can distinguish high-confidence records from provisional ones.
2. **Publish low-risk verified sequences:** campaign for the release of voucher-backed private and institutional COI records where identifications and metadata are already robust. This is the clearest low-hanging fruit because it improves coverage without waiting for new field collections.
3. **Resolve high-value anomalies:** audit and correct records with known or suspected misidentifications, coordinate errors, overseas-only sequences, unresolved names and taxonomic conflicts before they are used in council interpretation.
4. **Develop a New Zealand curated reference set:** building on useful existing information, including freshwater invertebrate sequence data generated through the successful Envirolink Tool Grant (2024) *Stocking New Zealand's molecular library for eDNA monitoring*, establish a versioned, quality-coded freshwater invertebrate reference library linked to voucher specimens and transparent metadata standards. The immediate priority is not simply to generate more sequence data, but establish best practices by consolidating, curating and creating an expert-verified New Zealand-specific reference library. This will provide quality assured data to councils enabling defensible ecological interpretation and future development of tools such as an Environmental DNA-based Macroinvertebrate Community Index (eMCI).

5. **Target new collections and sequencing:** direct new effort to taxa with no public coverage, conservation-priority species, regionally endemic taxa, groundwater and spring fauna, alpine and headwater taxa, island populations, poorly sampled catchments and non-EPT groups where coverage is weakest.
6. **Increase geographic representation:** identify spatial gaps within each taxon and prioritise collections that capture regional haplotypes, disjunct populations and habitat-specific lineages rather than relying on single locality or overseas-only records.
7. **Evaluate additional markers:** assess markers such as 16S and 12S for taxa or genera where COI provides poor amplification, weak discrimination or limited ecological utility.
8. **Embed QA/QC in procurement:** require eDNA suppliers to report reference library confidence, taxonomic assignment thresholds, database source, marker, quality flags and the most defensible reporting level with all council monitoring outputs.

6 Conclusions

Environmental DNA monitoring has substantial potential for freshwater biodiversity assessment, but its performance depends on the marker and taxonomic group being assessed. Public COI reference coverage remains incomplete and uneven, and COI amplification or species-level discrimination may be inadequate for some non-insect groups.

The New Zealand freshwater invertebrate barcode reference resource is not a single unified library, but a dispersed collection of sequences held across multiple repositories, including BOLD, GenBank, GEOME and various private or institutional datasets. Collectively, these sources provide adequate COI coverage for some groups, particularly caddisflies and mayflies, but the overall reference base remains far from comprehensive. Major limitations include the absence of public COI sequences for 41.5% of assessed taxa, insufficient sequence depth for many represented species, documented taxonomic misidentifications, metadata inaccuracies, spatially biased geographic coverage, and poor representation of threatened and non-insect taxa.

The immediate priority is not simply to generate more sequence data, but to consolidate these dispersed holdings into an expert-verified, quality-coded, New Zealand-specific reference resource, a sequence library, that councils can rely on for defensible ecological interpretation. The most efficient pathway is a staged programme that first makes better use of existing verified data, then corrects known errors, then fills priority taxonomic and spatial gaps. Publishing verified private and institutional sequences, improving metadata records, correcting misidentifications, increasing geographic representation and targeted strategic sequencing will provide councils with a stronger evidence base for biodiversity reporting, threatened-species surveillance, consent and biosecurity monitoring, restoration assessment and future development of tools such as the eMCI.

7 Acknowledgements

This work was supported and funded through the Envirolink Advice Grant programme (2624–ORC22 Enhancing confidence in eDNA monitoring for freshwater invertebrates). We wish to thank project champions Ciaran Campbell and Scott Jarvie (Otago Regional Council), and all the organisations and regional councils that supported this work. We are especially grateful to Canterbury Museum for their generous support, particularly to Johnathon Ridden (former Assistant Curator) and Timothy Hitchings (Research Associate) for continual access to specimens and data, and a safe space to discuss all things taxonomy. Figure 4–3 was kindly supplied by, and used with permission from, Jackson Treece and Joanne Gillum, University of Otago.

8 References

- Andrew, I.G., Macfarlane, R.P., Johns, P.M., Hitchmough, R.A., Stringer, I.A.N. (2012) The conservation status of New Zealand Diptera. *New Zealand Entomologist*, 35: 99–102. <https://doi.org/10.1080/00779962.2012.686312>
- Asmyhr, M.G., Cooper, S.J.B. (2012) Difficulties barcoding in the dark: the case of crustacean stygofauna. *Invertebrate Systematics*, 26(5–6): 583–591. <https://doi.org/10.1071/IS12032>
- Buckley, T.R. (2025) Charting a future for entomological taxonomy in New Zealand. *New Zealand Entomologist*, 48(1): 31–47. <https://doi.org/10.1080/00779962.2024.2407230>
- Buckley, T.R., Dopheide, A., Ward, D., Dhami, M. (2026) Prospects for complete DNA barcode coverage of the New Zealand insect fauna. *Journal of the Royal Society of New Zealand*, 56(4): e70078. <https://doi.org/10.1002/snz2.70078>
- Buckner, J.C., Sanders, R.C., Faircloth, B.C., Chakrabarty, P. (2021) Science Forum: The critical importance of vouchers in genomics. *eLife*, 10: e68264. <https://doi.org/10.7554/eLife.68264>
- Chee, S.Y. (2015) Limitations of cytochrome oxidase I for the barcoding of Neritidae (Mollusca: Gastropoda) as revealed by Bayesian analysis. *Genetics and Molecular Research*, 14(2): 5677–5684. <http://dx.doi.org/10.4238/2015.May.25.20>
- Chen, X., Yan, Z., Li, S., Yao, M. (2025) Advancing aquatic biodiversity assessments of invertebrates using eDNA metabarcoding: A systematic evaluation of primers for marine and freshwater communities. *Methods in Ecology and Evolution*, 16 (10): 2408–2430. <https://doi.org/10.1111/2041-210X.70152>
- Collier, K.J., Smith, B.J. (2006) Distinctive invertebrate assemblages in rockface seepages enhance lotic biodiversity in northern New Zealand. *Biodiversity & Conservation*, 15(11): 3591–3616. <https://doi.org/10.1007/s10531-005-5395-8>
- Csabai, Z., Čiamporová-Zaťovičová, Z., Boda, P., Čiampor Jr, F. (2023) 50%, not great, not terrible: Pan-European gap-analysis shows the real status of the DNA barcode reference libraries in two aquatic invertebrate groups and points the way ahead. *Science of the Total Environment*, 863: e160922. <https://doi.org/10.1016/j.scitotenv.2022.160922>
- Dawkins, K.L., Alexander, J.B., Huey, J., Curran, M., Main, D. Rodman, S. (2025) Using a species-specific assay to improve detection of an elusive stygobitic vertebrate, *Ophisternon candidum*, through environmental DNA. *Conservation*

- Genetics Resources*, 17(2): 131–140. <https://doi.org/10.1007/s12686-025-01387-5>
- Elbrecht, V., Leese, F. (2017) Validation and development of COI metabarcoding primers for freshwater macroinvertebrate bioassessment. *Frontiers in Environmental Science*, 5: 11. <https://doi.org/10.3389/fenvs.2017.00011>
- Emerson, B.C. (2025) Delimiting species – prospects and challenges for DNA barcoding. *Molecular Ecology*, 34(5): e17677. <https://doi.org/10.1111/mec.17677>
- Fenwick, G.D. (2001) The freshwater Amphipoda (Crustacea) of New Zealand: A review. *Journal of the Royal Society of New Zealand*, 31(2): 341–363. [DOI:10.1080/03014223.2001.9517658](https://doi.org/10.1080/03014223.2001.9517658)
- Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, 3(5): 294–299.
- Graham, E., Baker, C. (2025) Incorporating eDNA into macroinvertebrate metrics. In: White, R., Graham, E., Baker, C. (2025) Environmental DNA indicators for supporting implementation of the National Objectives Framework (NOF) in the NPS-FM. *Earth Sciences Client Report MFE25205*. Prepared for the Ministry for the Environment: 41.
- Graham, S.E., Smith, B.J., Pohe, S.R. (2022) Distributions of alpine freshwater invertebrates: Mt Taranaki field survey. *NIWA Client Report 2022190HN*. Prepared for Department of Conservation: 33.
- Grainger, N., Harding, J., Drinan, T., Collier, K., Smith, B., Death, R., Makan, T., Rolfe, J. (2018) Conservation status of New Zealand freshwater invertebrates, 2018. *New Zealand Threat Classification Series 28*. Department of Conservation, Wellington: 25.
- Haase, M. (2008) The radiation of hydrobiid gastropods in New Zealand: a revision including the description of new species based on morphology and mtDNA sequence information. *Systematics and Biodiversity*, 6(1): 99–159. <https://doi.org/10.1017/S1477200007002630>
- Hards, K., Muellner, P. (2026) *eDNABridge: Streamlining the transfer of New Zealand eDNA data into GBIF (CR703)*. Prepared for the Ministry for the Environment by Epi. Wellington: 16. <https://environment.govt.nz/publications/ednabridge-streamlining-the-transfer-of-new-zealand-edna-data-into-gbif/>
- Hebert, P.D.N., Cywinska, A., Ball, S.L., deWaard, J.R. (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, 270 (1512): 313–321. <https://doi.org/10.1098/rspb.2002.2218>

- Henderson, I.M., Ward, J.B. (2007) Three new species in the endemic New Zealand genus *Alloecentrella* (Trichoptera), and a re-evaluation of its family placement. *Aquatic Insects*, 29(2): 79–96.
<https://doi.org/10.1080/01650420701381155>
- Hogg, I.D., Smith, B.J., Banks, J.C., Dewaard, J.R., Hebert, P.D. (2009) Testing use of mitochondrial COI sequences for the identification and phylogenetic analysis of New Zealand caddisflies (Trichoptera). *New Zealand Journal of Marine and Freshwater Research*, 43(5): 1137–1146.
<https://doi.org/10.1080/00288330.2009.9626536>
- Keck, F., Couton, M., Altermatt, F. (2023) Navigating the seven challenges of taxonomic reference databases in metabarcoding analyses. *Molecular Ecology Resources*, 23(4): 742–755. <https://doi.org/10.1111/1755-0998.13746>
- Khan, A.A., Keber, B.L.M., Harder, K.G., Ogden, J.L., Docker, M.F. (2026) Species-specific environmental DNA (eDNA) assay for monitoring the distribution of the endangered carmine shiner *Notropis percobromus* in Manitoba. *Environmental Biology of Fishes*, 109(2): 78.
<https://doi.org/10.1007/s10641-026-01848-2>
- Komai, T., Gotoh, R.O., Sado, T., Miya, M. (2019) Development of a new set of PCR primers for eDNA metabarcoding decapod crustaceans. *Metabarcoding and Metagenomics*, 3: e33835. <https://doi.org/10.3897/mbmg.3.33835>
- Leese, F., Sander, M., Buchner, D., Elbrecht, V., Haase, P., Zizka, V.M.A. (2021) Improved freshwater macroinvertebrate detection from environmental DNA through minimized nontarget amplification. *Environmental DNA*, 3(1): 261–276. <https://doi.org/10.1002/edn3.177>
- Leschen, R.A.B., Collier, K.J., Death, R., Harding, J.S., Smith, B.J. (2016) Systematics expertise and taxonomic status of New Zealand's freshwater insects. *New Zealand Science Review*, 73: 92–95. <https://doi.org/10.26686/nzsr.v73i3-4.8534>
- Lester, P.J., Brown, S.D.J., Edwards, E.D., Holwell, G.I., Pawson, S.M., Ward, D.F., Watts, C.H. (2014) Critical issues facing New Zealand entomology. *New Zealand Entomologist*, 37(1): 1–13.
<https://doi.org/10.1080/00779962.2014.861789>
- Löbl, I., Klausnitzer, B., Hartmann, M., Krell, F.-T. (2023) The silent extinction of species and taxonomists — An appeal to science policymakers and legislators. *Diversity*, 15(10): 1053. <https://doi.org/10.3390/d15101053>
- Low, V.L., Norma-Rashid, Y., Yusoff, A., Vinnie-Siow, W.Y., Prakash, B.K., Tan, T.K., Noorhidayah, M., Chen, C.D., Sofian-Azirun, M. (2017) Pleistocene demographic expansion and high gene flow in the Globe Skimmer dragonfly *Pantala flavescens* Fabricius (Odonata: Libellulidae) in Peninsular

- Malaysia. *Zoologischer Anzeiger*, 266: 23–27.
<https://doi.org/10.1016/j.jcz.2016.10.002>
- McCulloch, G.A., Pohe, S.R., Wilkinson, S.P., Drinan, T.J., Waters, J.M. (2025) Targeted eDNA metabarcoding reveals new populations of a range-limited stonefly. *Ecology and Evolution*, 15(4): e71244.
<https://doi.org/10.1002/ece3.71244>
- McCulloch, G.A., Wallis, G.P., Waters, J.M. (2010) Onset of glaciation drove simultaneous vicariant isolation of alpine insects in New Zealand. *Evolution*, 64(7): 2033–2043. <https://doi.org/10.1111/j.1558-5646.2010.00980.x>
- McCulloch, G.A., Wallis, G.P., Waters, J.M. (2017) Does wing size shape insect biogeography? Evidence from a diverse regional stonefly assemblage. *Global Ecology and Biogeography*, 26(1): 93–101.
<https://doi.org/10.1111/geb.12529>
- McGinley, R.J. (1993) Where's the management in collections management? Planning for improved care, greater use and growth of collections. *Int. Symp. & First World Congress on Pres. Cons. of Nat. Hist. Col.* 3: 309–338.
- Ministry for the Environment (2025) National Policy Statement for Freshwater Management 2020 [amended December 2025]. ME1948. Wellington: 72.
<https://environment.govt.nz/assets/publications/Freshwater/NPSFM-amended-october-2024.pdf?v=d1f23ad0dde4c67bc165b427c2ec97267a80749c>
- Melchior, M., Baker, C. (2023) Environmental DNA guidelines and field protocols for lotic systems. *NIWA Client Report 2023/279HN*. Prepared for Department of Conservation: 84.
- Murray, T.J., Borkin, K.M., Brumby, A., O'Donnell, C.F.J. (2024) Threatened species research gaps and priorities for the Department of Conservation Te Papa Atawhai. *DOC Research and Development Series 371*. Department of Conservation, Wellington: 40.
- Páll-Gergely, B., Krell, F.T., Ábrahám, L., Bajomi, B., Balog, L.E., Boda, P., Csuzdi, C., Dányi, L., Fehér, Z., Hornok, S., Horváth, A., Kóbor, P., Koczor, S., Kontschán, J., Kovács, P., Kovács, T., Lukátsi, M., Majoros, G., Murányi, D., Németh, T., Perneckner, B., Puskás, G., Rózsa, L., Soltész, Z., Szita, É., Szűts, T., Tóth, B., Tőke, A., Vas, Z., Zsuga, K., Zsupos, V., Csabai, Z., Móra, A. (2024) Identification crisis: a fauna-wide estimate of biodiversity expertise shows massive decline in a Central European country. *Biodiversity and Conservation*, 33(13): 3871–3903. <https://doi.org/10.1007/s10531-024-02934-6>
- Parliamentary Commissioner for the Environment (2026). Unlocking New Zealand's environmental information. Wellington: Parliamentary Commissioner for the Environment: 138.

- Pentinsaari, M., Ratnasingham, S., Miller, S.E., Hebert, P.D.N. (2020) BOLD and GenBank revisited – Do identification errors arise in the lab or in the sequence libraries? *PLoS One*, 15(4): e0231814.
<https://doi.org/10.1371/journal.pone.0231814>
- Pohe, S.R., Smith, B.J., Ball, O.J.-P. (2020) Towards determination of aquatic insect conservation status in New Zealand: EPT survey results from Westland–Brunner, Tasman–Motueka and the Canterbury high country. Unpublished report prepared for the Department of Conservation: 29.
- Pohe, S.R., Smith, B.J., Ball, O.J.-P. (2023) Towards determination of aquatic insect conservation status in New Zealand 2: EPT surveys from Kahurangi National Park and environs, including a summary of all ‘Taxonomically Indeterminate’ EPT taxa. Unpublished technical report for the Department of Conservation, 30.
- Prabowo, D.A., Setyawan, E., Hasan, A.W., Mambrasar, R., Prasetyo, A.P., Sutikno, S., Ewaldo, M.F., Ambafen, O., Saleo, N., Putra, M.I.H., Erdmann, M.V. (2026) Tracking fin-prints on reef flats: A genus-specific eDNA metabarcoding assay for epaulette sharks (*Hemiscyllium* spp.) with field validation in Raja Ampat. *Environmental DNA*, 8(3): e70281. <https://doi.org/10.1002/edn3.70281>
- Santos, B.F., Miller, M.E., Miklasevskaja, M., McKeown, J.T.A., Redmond, N.E., Coddington, J.A., Bird, J., Miller, S.E., Smith, A., Brady, S.G., Buffington, M.L., et al. (2023) Enhancing DNA barcode reference libraries by harvesting terrestrial arthropods at the Smithsonian’s National Museum of Natural History. *Biodiversity Data Journal*, 11: e100904.
<https://doi.org/10.3897/BDJ.11.e100904>
- Smith, B.J., Graham, S.E., Pohe, S.R. (2024) Distributions of alpine freshwater invertebrates 2: Mt Taranaki and Fiordland field study. *NIWA Client Report* 2024180HN. Prepared for Department of Conservation: 65.
- Strutzenberger, P., Brehm, G., Fiedler, K. (2012) DNA barcode sequencing from old type specimens as a tool in taxonomy: a case study in the diverse genus *Eois* (Lepidoptera: Geometridae). *PLoS One*, 7(11): e49710.
<https://doi.org/10.1371/journal.pone.0049710>
- Suren, A.M., Burdon, F.J., Wilkinson, S.P. (2024) eDNA is a useful environmental monitoring tool for assessing stream ecological health. *Environmental DNA*, 6(4): e596. <https://doi.org/10.1002/edn3.596>
- Takenaka, M., Hasebe, Y., Yano, K., Okamoto, S., Tojo, K., Seki, M., Sekiguchi, S., Jitsumasa, T., Morohashi, N., Handa, Y., Sakaba, T. (2024) Environmental DNA metabarcoding on aquatic insects: Comparing the primer sets of MtlInsects-16S based on the mtDNA 16S and general marker based on the mtDNA COI region. *Environmental DNA*, 6(4): e588. <https://doi.org/10.1002/edn3.588>
- Takenaka, M., Uchida, N., Hasebe, Y., Yano, K., Ueki, G. (2026) Environmental DNA metabarcoding for insect biodiversity surveys: Current status and future

- prospects for potential applications in highly diverse insect taxa.
Entomological Science, 29(1): e12616. <https://doi.org/10.1111/ens.12616>
- Takenaka, M., Yano, K., Suzuki, T., Tojo, K. (2023) Development of novel PCR primer sets for DNA barcoding of aquatic insects, and the discovery of some cryptic species. *Limnology*, 24 (2): 121–136.
<https://doi.org/10.1007/s10201-022-00710-5>
- Taylor R.W. (1983) Descriptive taxonomy: past, present, and future. In: E. Highley & R.W. Taylor (Eds) *Australian Systematic Entomology: A Bicentenary Perspective*. CSIRO, Canberra: 93–134.
- Trewick, S.A., Henderson, I.M., Pohe, S.R., Morgan–Richards, M. (2022) Spatial variation of *Acanthophlebia cruentata* (Ephemeroptera), a mayfly endemic to Te Ika-a-Māui—North Island of Aotearoa, New Zealand. *Insects*, 13(7): 567.
<https://doi.org/10.3390/insects13070567>
- Weigand, H., Beermann, A.J., Čiampor, F., Costa, F.O., Csabai, Z., Duarte, S., Geiger, M.F., Grabowski, M., Rimet, F., Rulik, B., Strand, M., Szucsich, N., Weigand, A.M., Willassen, E., Wyler, S.A., Bouchez, A., Borja, A., Čiamporová–Zaťovičová, Z., Ferreira, S., Dijkstra, K.–D.B., Eisendle, U., Freyhof, J., Gadawski, P., Graf, W., Haegerbaeumer, A., van der Hoorn, B.B., Japoshvili, B., Keresztes, L., Keskin, E., Leese, F., Macher, J.N., Mamos, T., Paz, G., Pešić, V., Pfannkuchen, D.M., Pfannkuchen, M.A., Price, B.W., Rinkevich, B., Teixeira, M.A.L., Várбірó, G., Ekrem, T. (2019) DNA barcode reference libraries for the monitoring of aquatic biota in Europe: Gap–analysis and recommendations for future work. *Science of the Total Environment*, 678: 499–524.
<https://doi.org/10.1016/j.scitotenv.2019.04.247>
- Winterbourn, M. (2014) Gentlemen or players: New Zealand's aquatic insect taxonomists. *The Wētā*, 47: 58–61.
<https://weta.ento.org.nz/index.php/weta/article/view/162>
- Yakimenko, E.V., Romanovich, A.E., Lukhtanov, V.A. (2025). Questionable species names for distinct species clusters: An empirical test of the BOLD Molecular Identification Engine. *Insects*, 16(11): 1172.
<https://doi.org/10.3390/insects16111172>
- Zhang, Y., Yang, H., Liang, X., Zhang, M., Zhu, W., Yang, H., Fan, W., Ping, Y., Li, K. (2026) Decoding benthic macroinvertebrate communities in freshwater ecosystems leveraging environmental DNA. *Water Research*, X 32: 100550.
<https://doi.org/10.1016/j.wroa.2026.100550>