Increasing the efficiency and cost effectiveness of the MDBA-funded Murray and Mitta river biological monitoring programs through the development of molecular tools

Prepared by: Michael Shackleton, Gavin Rees, Garth Watson, & Chris Davey

June 2014

Final report

MDFRC Publication 25/2014

Increasing the efficiency and cost effectiveness of the MDBA-funded Murray and Mitta river biological monitoring programs through the development of molecular tools.

Draft Report prepared for the Murray-Darling Basin Authority by The Murray-Darling Freshwater Research Centre.

Murray-Darling Basin Authority
Level 4, 51 Allara Street | GPO Box 1801
Canberra City ACT 2601

Ph: (02) 6279 0100; Fax: (02) 6248 8053

This report was prepared by The Murray-Darling Freshwater Research Centre (MDFRC). The aim of the MDFRC is to provide the scientific knowledge necessary for the management and sustained utilisation of the Murray-Darling Basin water resources. The MDFRC is a joint venture between the Murray-Darling Basin Authority, La Trobe University and CSIRO.

For further information contact:

Michael Shackleton

The Murray-Darling Freshwater Research Centre
PO Box 991
Wodonga VIC 3689
Ph: (02) 6024 9650; Fax: (02) 6059 7531

Email: m.shackleton@latrobe.edu.au
Web: www.mdfrc.org.au
Enquiries: info@mdfrc.org.au

Report Citation: Shackleton M, Rees G, Watson G, Davey C (2013) Increasing the efficiency and cost effectiveness of the MDBA-funded Murray and Mitta river biological monitoring programs through the development of molecular tools. Draft Report prepared for the Murray Darling Basin Authority by The Murray-Darling Freshwater Research Centre, MDFRC Publication 25/2014, June, 32pp.

Cover Image: Dorsal view of an adult Elmidae of the genus Austrolimnius collected by James Anderson from the Alpine high plains.

Photographer: Chris Davey

Contact us

Inquiries regarding the licence and any use of the document are welcome at:

The Murray-Darling Freshwater Research Centre
PO Box 991,
Wodonga Vic 3689
Ph: 02 6024 9650

Document history and status

Version	Date Issued	Reviewed by	Approved by	Revision type
Draft	17/6/14	Lucy Alderton		External Copy Edit

Distribution of copies

Version	Quantity	Issued to
Final	1	Tapas Biswas, M-DBA

Filename and path: Projects/MDBA/484-Barcoding II/administration/Draft Report/Draft Report (June 2014)

Author(s): Shackleton M, Rees G, Watson G, Davey C

Author affiliation(s): Murray-Darling Freshwater Research Centre

Project Manager: M. Shackleton

Client: Murray Darling Basin Authority

Project Title: Increasing the efficiency and cost effectiveness of the MDBA-funded Murray and Mitta river biological monitoring programs through the development of molecular tools

Document Version: Draft

Project Number: M/BUS/484

Contract Number: MD2616

Acknowledgements:

The work contained herein was funded by the Murray Darling Basin Authority (MDBA). Many of the specimens used in this research were made available through the MDBA funded Mitta Mitta River and Murray River biological macroinvertebrate monitoring programs. Part of the research was undertaken by James Anderson as part of a La Trobe Industry Cadetship program, including collection and genetic analyses of specimens.

Contents

List of tables

List of figures

Executive summary

This report summarises three genetically based projects funded by the Murray Darling Basin Authority (MDBA) and undertaken by the Murray-Darling Freshwater Research Centre (MDFRC). The project was designed to aid in the development of tools for increasing the efficiency and cost effectiveness of the MDBA-funded River Murray and Mitta Mitta River biological monitoring programs.

Genetic tools are currently being developed for large-scale assessment of biodiversity. While some techniques look promising and are being increasingly used for identification of aquatic macroinvertebrates, there is still much work required to resolve some basic issues. Foremost is the need for a comprehensive library of DNA barcodes matched to vouchered specimens so that genetic data gathered during future research and monitoring programs can be assigned species identities. This will ensure that such data are biologically meaningful. Secondly, while various techniques exist for detecting DNA, most focus on single genes with limited capacity to discern between species. Such techniques can also be costly to implement at the multi-gene scale. A technique for rapidly assessing multiple gene fragments, increasing the likelihood of species detection, from bulk samples is desirable. Lastly, the behaviour of DNA, either dissolved or as free cells within aquatic systems, is largely unknown. Understanding how DNA degrades within aquatic systems under various environmental parameters will be important if biological monitoring programs are to take advantage of these cost-effective and data rich methods for assessing biodiversity.

To address the abovementioned issues the current project aimed to:

- increase the macroinvertebrate DNA database by adding new specimens from the Murray-Darling Basin to the macroinvertebrate DNA library database
- investigate the potential use of Hybrid Sequence Capture (HSC) to more accurately detect the presence of species from environmental samples
- investigate how long DNA remains detectable in aquatic ecosystems, thereby determining the extent to which environmental DNA can provide site specific data.

The project increased the existing macroinvertebrate DNA database by 271 specimens. This included DNA data from species for which no sequences are currently described. Analysis of the DNA data revealed a number of taxonomic issues. In particular, the data revealed some larvae that are currently recognised as separate taxa actually belong to a single species; some larvae recognised as a single taxon actually belong to two species; associations between seven undescribed larvae and seven adults; and a possible 12 new species.

The 271 specimens are a valuable addition to the macroinvertebrate database and will provide important information for taxonomists and researchers working with these taxa. The data provided here, along with that already available in databases such as the Barcode of Life Database (BOLD) and GenBank, will form the backbone to answer future DNA based ecological questions, to describe new species, and to investigate future methods employed in biological monitoring programs.

The project identified a protocol that allows amplification of DNA from mitochondria of test organisms. This was the first step in the HSC approach. Actual DNA sequencing (phase two of the HSC approach) is still in progress and should supply relevant information for the further development of the HSC protocol.

Results from the DNA decay component of the project were not clear, although results did indicate that DNA could exist in our experimental systems for eight to 10 days before experimental bottle effects began to develop, leading to unexpected results. Low levels of turbidity occurred in the positive controls, where DNA amplification steadily declined, but large additions of clay (terrestrially

derived suspended solid) inhibited DNA detection. The mechanisms of this inhibition could not be resolved during this project and requires further consideration if DNA is to be examined in aquatic systems where suspended solids may be high.

1 Introduction

Increasingly, genetic techniques are offering new ways in which to answer biological questions and DNA analysis looks to become an important tool in the areas of biological research and monitoring. DNA is becoming widely used as an identification tool for aquatic macroinvertebrates, enabling species level identities to be placed on otherwise unresolvable specimens, such as small or damaged organisms. Recent advances in research have shown how the fauna of bulk samples can be rapidly characterised using Next Generation Sequencing (NGS) techniques. Through capturing freely available DNA, diversity within sampling sites can be measured in environmental samples, such as water or soil. The ultimate goal for the use of these techniques is to enable quantification of abundance data, rather than simple presence/absence data, and some work is progressing in this area.

Before the full potential of genetic techniques for biological monitoring can be realised, some basic investigations and data gathering must be undertaken. Foremost is the establishment of a comprehensive DNA data library, where gene sequences are matched to vouchered specimens. Once such a library is developed, identifying specimens can be a simple matter of comparing their DNA sequences to those housed in the DNA database. This has the potential to revolutionise monitoring programs, allowing greater numbers of samples to be collected and analysed at a reduced cost and with greater taxonomic accuracy. Some work has progressed towards developing such a library. In 2013, the MDBA funded a small project that saw the establishment of a Barcode Of Life Database (BOLD) for Aquatic Macroinvertebrates of the Murray-Darling Basin (AAMDB). The project captured 236 specimens from 12 orders within the database. This project has already raised interest among others working with aquatic invertebrates, with offers to incorporate privately held data into the AAMDB database and a push for the database to increase its bounds to incorporate data from all of Australia. Indeed, since its inception, a further 142 specimens of Trichoptera have been added to the database from a private collection. Complementary to this database are a number of private databases that include sequence data from Ephemeroptera, Trichoptera, and Plecoptera. Over time the data from these databases will be made public, greatly adding to the data pool already available.

If DNA is to be used to determine species presence from environmental samples, the development of better techniques to extract and analyse environmental DNA (eDNA) is required, along with knowledge of how eDNA behaves in aquatic systems. A number of techniques have already been developed, each designed to extract genetic data in order to answer specific questions and each with their own advantages and limitations. For instance, classical single specimen DNA amplification can inform species identification of a single organism and can be scaled up to be used in quantifying abundance data, but is costly to do so. Next Generation Sequencing (NGS) offers a means to identify organisms in bulk by amplifying a specific gene sequence from all available material in a sample, but is not useful for quantitative analysis. Lastly, a technique known as 'shotgun sequencing' looks promising to offer some quantification, albeit relative abundance, and has the potential to characterise multiple gene fragments simultaneously from a given sample. However, its non-targeted nature leaves it prone to not detecting gene fragments of interest and makes it cost ineffective for use with large Eukaryotic genomes.

The persistence of DNA in aquatic environments will play a large role in the utility of genetic tools for analysing eDNA. Due to the flowing nature of streams and rivers, it will be important to quantify how

likely any given eDNA signal has originated from a sampling site at which it was collected or from an area further upstream. River monitoring programs rely on characterisation of fauna at a given site over time. If genetic tools are to be employed, they will need to take into account the fact that eDNA collected at a site may have drifted in from a site many kilometres upstream. As yet, information surrounding the longevity and travelling distances of eDNA is sparse. Some work has been done on DNA persistence (Deere *et al.* 1996; Dejean *et al.* 2011; England *et al.* 2005). However, no study has specifically looked into the persistence of the regularly used COI barcoding gene. The COI fragment is larger than previously studied fragments and is more readily broken apart or degraded through physical pressures.

This report is divided into three components, each component describes a sub-project aimed at addressing one of the issues stated above. Component 1 describes the addition of 271 new specimens to the AAMDB database, together with locality and photographic data. Component 2 describes a feasibility study looking into a new method to capture and analyse eDNA, namely, HSC. Component 3 describes a lab-based DNA longevity study designed to assess the timeframe in which COI gene fragments become undetectable in stable aquatic environments.

2 Component 1: New specimens to the macroinvertebrate DNA library

2.1 Introduction

In 2013, the MDBA funded barcoding work that saw 236 specimens from 12 orders of aquatic macroinvertebrates barcoded and vouchered (Shackleton *et al.* 2013). This information was made available through the Barcode Of Life Database, to be publically accessible for use in research and monitoring programs. This work highlighted many areas in which DNA barcoding could help resolve taxonomic and other identification issues for routine monitoring and research programs. A major outcome of the 2013 work was that the specimens collected and data gathered are now being used to describe new species, associate the adults and larvae of species, and resolve identification issues in laboratories. Through presenting the findings of the 2013 study to others, such as the Australian Society for Limnology Conference 2013, a dialogue has begun between interested parties about ways in which these data can be used, how we can increase macroinvertebrate genetic data, and how such data will be managed.

The present study focused on increasing the macroinvertebrate database for two major taxonomic groups; caddisflies (Trichoptera) and riffle beetles (Elmidae). These taxa were chosen partly due to the taxonomic issues that exist within both groups and partly because they are a current research focus in other areas. For instance, a number of experts in the taxonomy of caddisflies are currently working to describe new species and match unknown larvae to described adults in this order. The data gathered during the present study will be available to such people to be utilised in their research and so increase the outputs from a small research project.

Specimens chosen for the study were either sourced from the MDBA funded Mitta Mitta River and River Murray biological monitoring programs, currently undertaken by MDFRC, or were collected especially for the current project. Those specimens collected specifically for this project were either taken from sites used in the two monitoring programs or sites nearby. Through this project, a further 271 specimens of Trichoptera and Elmidae were added to the macroinvertebrate database (project code AMMDB on the BOLD website), complete with their gene sequences, photographs, and location information. This data will prove invaluable to researchers using aquatic macroinvertebrates.

2.2 Methods

A total of 168 larval caddisfly specimens were obtained from the Mitta Mitta River and River Murray biological monitoring programs. A further 127 adult caddisfly specimens were collected during a one day field trip to Snowy Creek, near the Lightning Creek camp ground. This site was chosen as it is close to a site that is part of the Mitta Mitta River Biological Monitoring Program (Mount Wills Creek 506). Identification of caddisflies were undertaken using the keys of Cartwright (1997; 1998), Dean (1997; 1999; 2000), Jackson (1998), Neboiss (1986; 1992), St Clair (1997; 2000), and Wells (1997).

Elmidae larvae and adults were collected from six sites in the Bogong high plains and in Mount Buffalo National Park as part of a MDFRC and La Trobe University Summer Industry Cadetship Program. The Bogong High plains sites included two sites that form part of the Mitta Mitta River Biological Monitoring Program. A total of 97 adult and 43 larval specimens were collected. Identification of Elmidae was undertaken using the keys of Hinton (1964), Glaister (1999), and Lawrence (1992).

Genetic material was extracted from specimens using the non-destructive extraction fluid XytXtract Insect (ANDE) (Castalanelli *et al.* 2010). A one-in-five dilution of the DNA extract was made and stored at -20°c. The non-diluted extract was stored at -80°c. Further analyses were conducted on the diluted DNA extract only, in order to reduce the number of freeze/thaw cycles the original genetic product was exposed to, which would degrade the quality of the genetic material.

A 658 base pair (bp) fragment of cytochrome oxidase subunit 1 (COI) was amplified through Polymerase Chain Reaction (PCR) using the primers HCO2198 and LCO1490 (Folmer *et al.* 1994) and following the protocols outlined in Webb and Suter (2010). Amplification was attempted at least once on every specimen using 1µl of DNA extract. Specimens that yielded PCR products were sequenced at Macrogen Inc. Korea.

The Barcode of Life Database (BOLD) is a web based database for storing and analysing genetic data. Data within this database can be either private or publicly accessible. From the BOLD database sequence, data can be directly submitted to GenBank, which is a public repository for genetic sequences. The sequence data generated for the current study has been uploaded onto the BOLD under the public folder 'Aquatic Macroinvertebrates of the Murray–Darling Basin' (AMMDB). At the time of this report the sequences are awaiting submission onto the GenBank database.

Some sequences that were not considered of high enough quality for publication have not been made public. For instance, where sequences were not assembled through analysis of both strands of the DNA (that is, they were represented by only a forward or a reverse strand) they have not been uploaded due to the likely errors contained within them. Similarly, where contamination issues were suspected, these have not been uploaded to the database as this would create false identifications attached to sequences.

Sequences were compared to other sequences previously obtained through the MDBA funded MDFRC barcoding program and to those available on GenBank and through BOLD. One of the databases for which the BOLD comparisons were done is currently a private database of caddisfly data to which one of the authors has access. As this data is not yet publicly available, the findings surrounding the comparison with this database are given without reference to actual data points.

Gene trees were created through MEGA6.2 using neighbour joining and the Kimara-2 parameter. Sequence data were interrogated to highlight areas where taxonomic issues could be resolved, such as discovery of new species or association between the adult and larvae of a species.

2.3 Results and discussion

A total of 202 caddisfly and 73 Elmidae specimens yielded sequences. Of the caddisflies, this included 95 adults, 112 larvae and one pupa. Six caddisfly specimens yielded sequences that were likely contaminants from other species, including contamination from other caddisfly species and from mites. Two caddisfly and two Elmidae specimens yielded sequences of low quality and were omitted from the data, bringing the total number of sequences to 271. The sequences for Elmidae obtained as part of this project, and that of the previous MDBA funded project, are the only Australian representatives of this family currently on BOLD and GenBank. In contrast, a relatively large number of caddisflies were available through GenBank and BOLD.

The complete gene trees for each caddisfly family and the total of the Elmidae, including data from specimens obtained in the previous MDBA funded MDFRC barcoding project, are given in Appendix A. Below we discuss the major findings from the current project within these two orders.

2.3.1 Elmidae

The barcoding gene fragment separated the genera *Kingolus* and *Austrolimnius* into well supported clades (see Appendix B). *Coxelmis* also occurred as a separate clade but consisted of only two species. In contrast the genera *Notriolus* and *Simsonia* did not separate out, due to the placement of one specimen, CO0007, identified as *Simsonia tasmanica* occurring within a clade of *Notriolus* species.

Austrolimnius

Larvae of *Austrolimnius* are difficult to distinguish, with many of the discriminating features being small differences in the shape of the terminal abdominal segment or the distribution of tubercles on the abdominal segments. For this reason, subjective discretion in assessing the extent of slight defences between the diagnostic characters is highly likely to be leading to erroneous data. The genetic data compiled for this project supports this assertion, with multiple species designations for larval specimens occurring within clades that represent single species (Figure 1).

In Figure 1, clade 1 includes larval specimens identified as *A. metasternalis*, *A. diamensis*, and species L10E along with adult specimens identified as *A. victoriensis*. Given the lack of confidence in the characters used to identify larval specimens and, conversely, the strong confidence in identifying adults of *A. victoriensis*, it is likely that these larvae belong to this species and that the variation in morphological features that exists between the specimens does not reflect actual species separations.

In clade 2, species L36E and L65E occur together, suggesting they belong to the same species. As yet, the species identity of the adults represented within this clade have not been ascertained and so no association between the two larval forms and a described adult can be made at this time.

Clade 3 is composed mostly of adult specimens, some of which have been identified as *A. anytus*. However, included in this clade are two larval forms identified as *A. diamensis* and species L44E. It is interesting to note that specimens identified as *A. diamensis* occurred in both clades 1 and 3, indicating that what constitutes the defining characters of the larva of this species may actually define two different species.

Lawrence (1992) places all the ambiguous larval forms mentioned here (i.e. those with the format of L##E) into the same sub-group within *Austrolimnius*. It is, therefore, not surprising that the characters used to discern between the two forms may not reflect actual species separations.

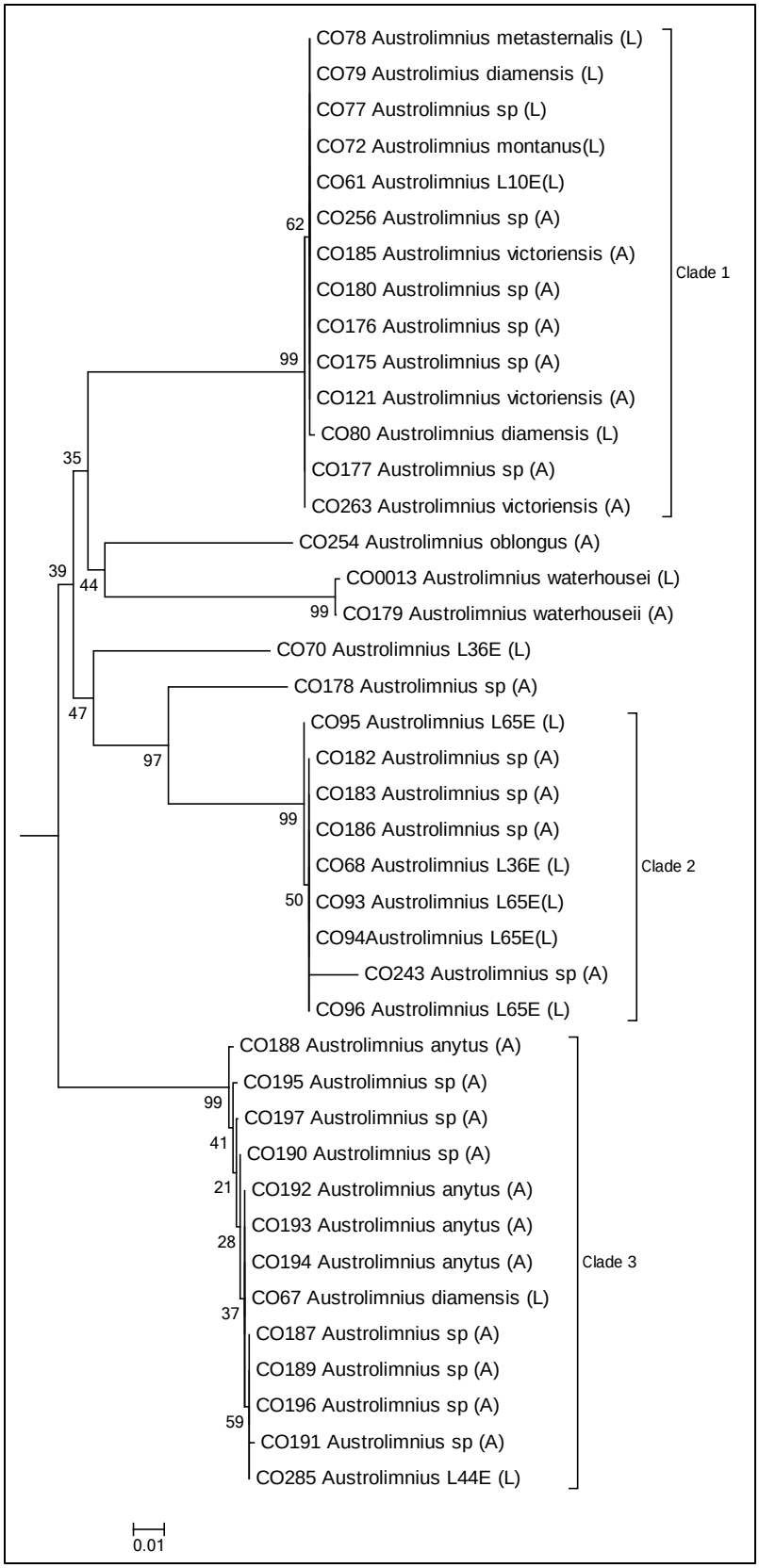


Figure 1. The COI gene tree of *Austrolimnius* specimens collected as part of both MDBA funded barcoding projects

Kingolus

Kingolus specimens grouped into two clades are shown in Figure 2. Adult and larval specimens identified as *K. yarrensis* occur within a single clade, indicating high confidence in the identification of this species. One larval specimen, CO0024, identified as *K. aeratus* occurred close to the *K. metallicus* clade and shared highly similar gene sequences with these specimens. It is possible that our current understanding of the characters that delineate larvae of *K. metallicus* and *K. aeratus* is erroneous.

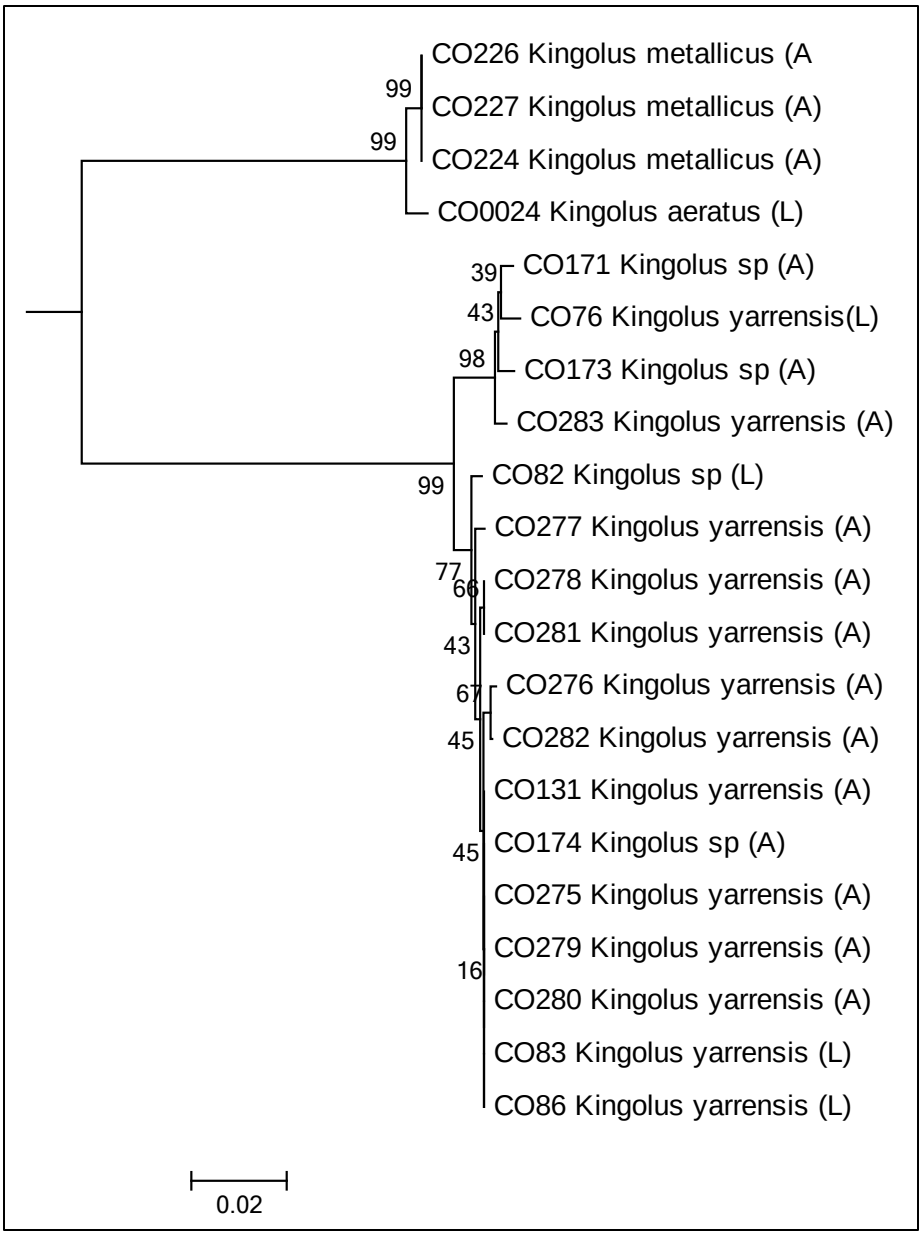


Figure 2. The COI gene tree of *Kingolus* specimens collected as part of both MDBA funded barcoding projects

Notriolus and Simsonia

As mentioned above, the placement of TR0007, *Simsonia tasmanica*, resulted in the Notriolus and Simsonia specimens not forming separate clades. Apart from this, of note is the similarity between the gene sequences of *Simsonia angusta* and L12E, suggesting that L12E is a variant belonging to this species. A larva identified as *Simsonia wilsoni* occurred with those of *S. longipes* as well as with an undescribed adult. This placement may be due to identifier error, given that the character that separates these two species is the colouration of the dorsal humps as viewed from above. Alternatively, it may indicate that this character does not reliably distinguish the two species. Confirming the actual identity of the larvae may rely on identifying the adults to which they group with.

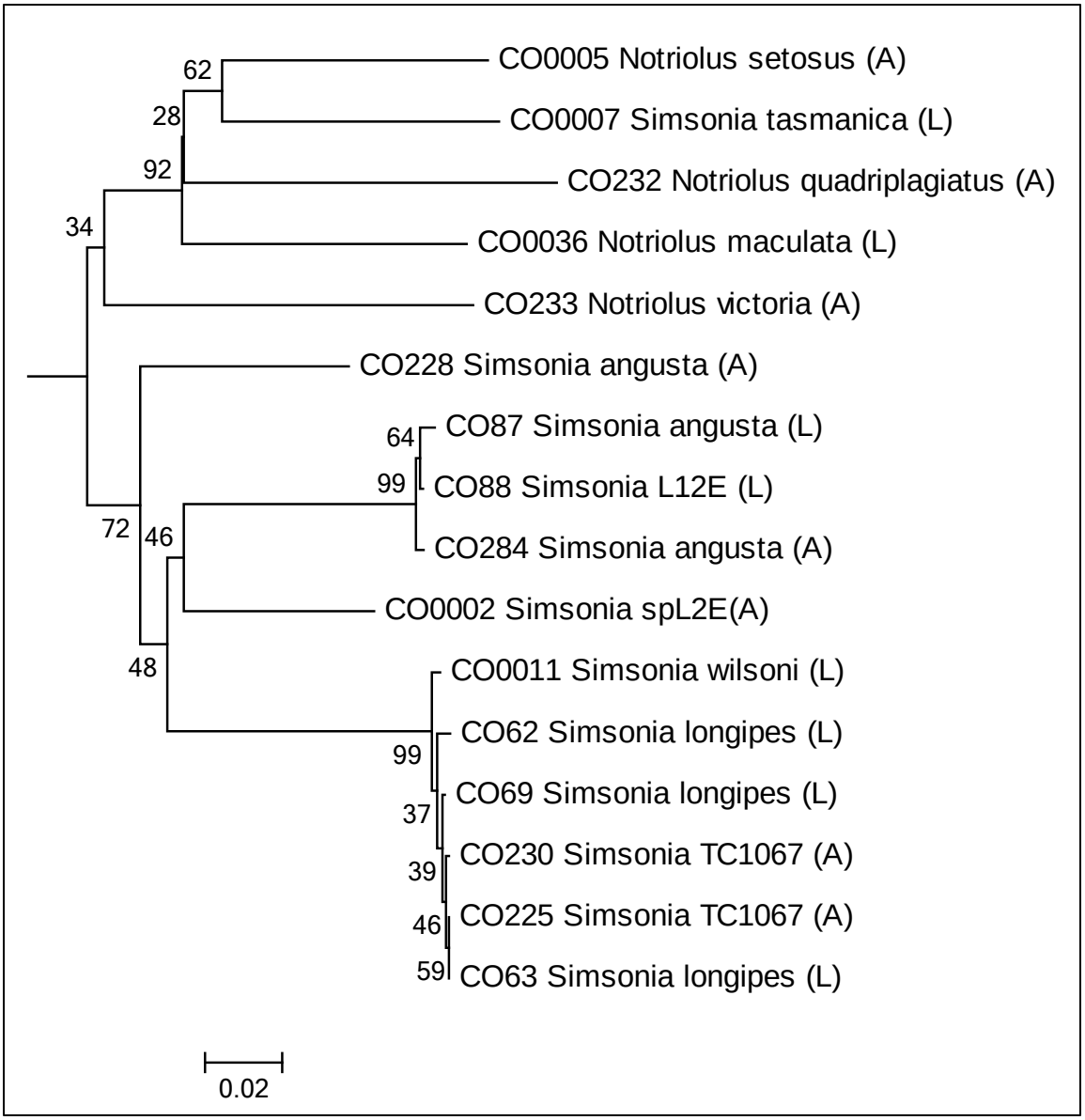


Figure 3. The COI gene tree of *Notriolus* and *Simsonia* specimens collected as part of both MDBA funded barcoding projects

2.3.2 Trichoptera

The COI data organised all families of caddisflies collected into discrete clades on the gene tree. Due to the large size of the complete gene tree, it is given in Appendix A. Below we illustrate and discuss only those sections of the gene tree that provided the most interesting information.

Agapetus

The larval form of *Agapetus* designated as species AV1 is a known species complex (Cartwright 1998). The genetic data resolved three distinct clades that are likely to represent three species (Figure 4). For two of the putative species (species 1 and 2) adult males were collected. These are to be examined by experts on this group to determine if they belong to a new species, and if so, to be described. As yet, no morphological characters reliably separate the larval type AV1. The ability to examine larval types that have been designated to a species based on DNA data may facilitate establishing morphologically based identification tools. The data collected as part of this study will aid in such an endeavour. The distributions of these three species overlap to some extent. Species 1 and 2 exhibited the largest distributions. While the northern most distribution of Species 2 was the River Murray at Jingellic (just above Lake Hume), comparison with a private database of caddisflies suggests that species 2 occurs as far north as the Allyn River in the Barrington Tops National Park, northern NSW. Similarly, a specimen collected near the Avon Wilderness Area suggests that species 1 occurs as far west as that area. At this stage, Species 3 appears to be largely confined to the upper reaches of the Mitta Mitta River catchment.

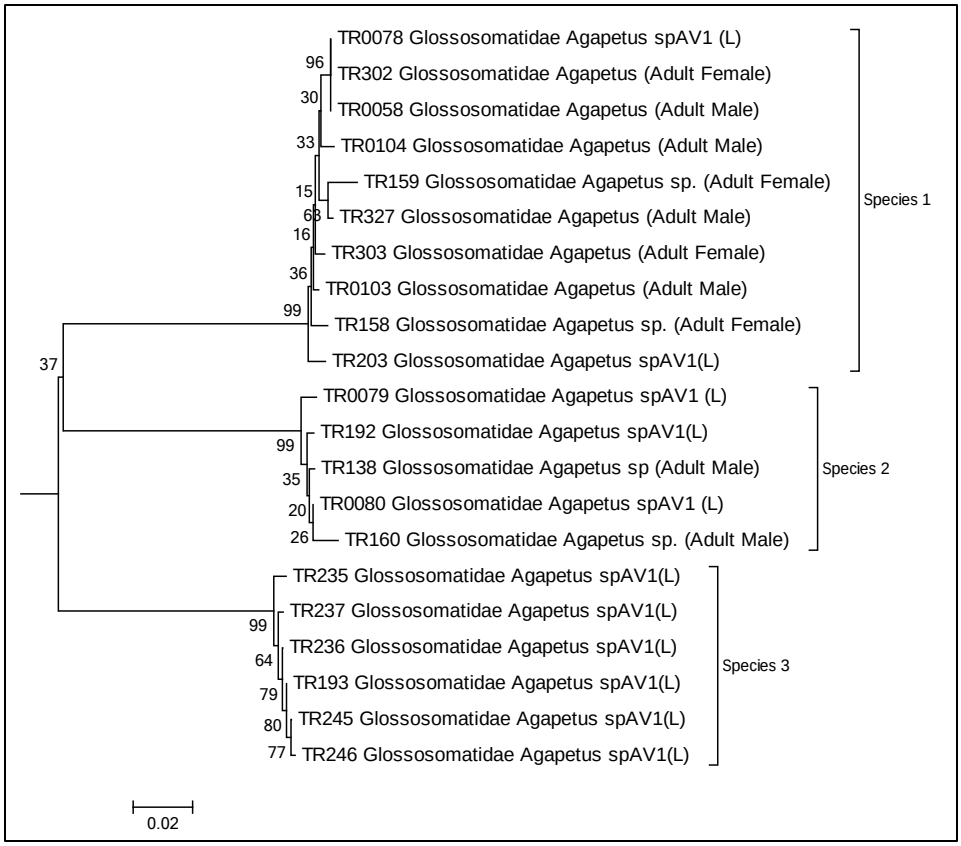


Figure 4. The COI gene tree of the genus *Agapetus* showing three distinct clades representing three putative species, collected as part of both MDBA funded barcoding projects

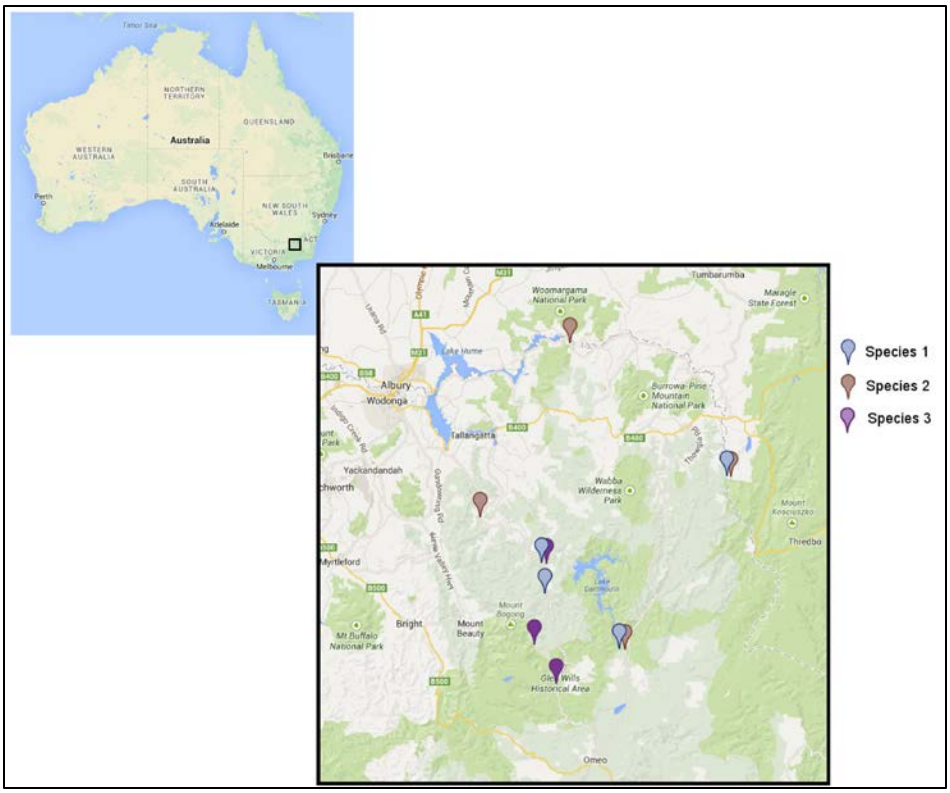


Figure 5. Map of sites where the three *Agapetus* species were collected during the present study

Hydropsychidae

Two species of Hydropsychidae were collected in the present study, bringing the total from both MDBA funded studies to three species. *Asmicridea* species AV1 was suggested as being either the larva of *A. edwardsii* or belonging to a species complex described by Dean (1999). The data presented here (Figure 6) suggests that the latter is correct. Only one species is described from mainland Australia (*A. edwardsii*). The presence of three distinct clades in the gene tree for this species suggests that at least three cryptic species exist in the current concept of *A. edwadsii*. Comparison with data form the private caddisfly database suggests that the number of species is more likely closer to 10. The data and specimens collected for the present study will be made available to John Dean, who is currently reviewing the taxonomy of this genus.

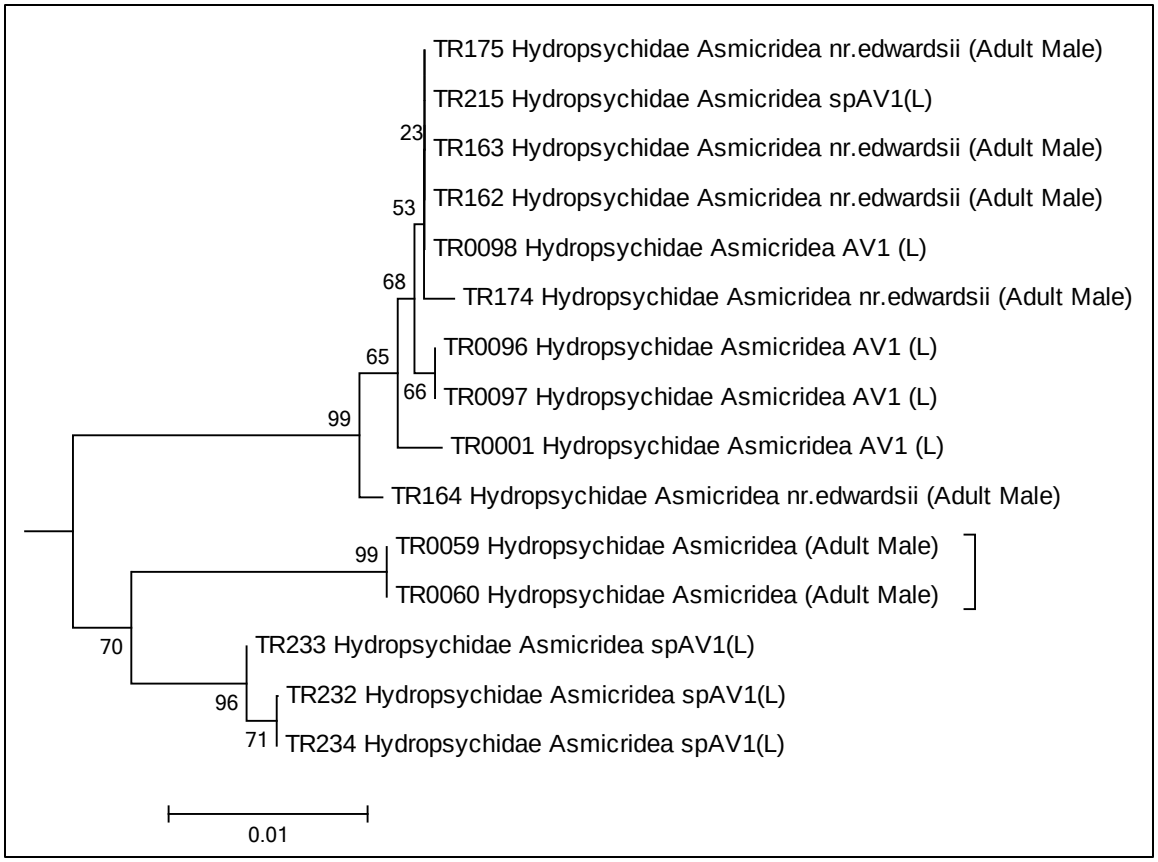


Figure 6. The COI gene tree of Hydropsychidae specimens collected as part of both MDBA funded barcoding projects

Calamoceratidae

Three species of Calamoceratidae are known to occur in the south-east of Australia. In the previous MDBA funded barcode program, three species were collected. However, species names could not be assigned to the specimens collected as they were all larvae and the taxonomy around the larval types is, as yet, unresolved. As a consequence of the current study, we were able to identify the larvae occurring in one clade (top-most in Figure 7) as being *Anisocentropus vulgas* through the collection and DNA barcoding of adult males of this species. Comparison with the private caddisfly database further confirmed specimens TR274 and MDFRC77 as being *A. bicoloratus*. The data and specimens collected as part of this project will be used in resolving morphological identification tools for the larvae of this genus.

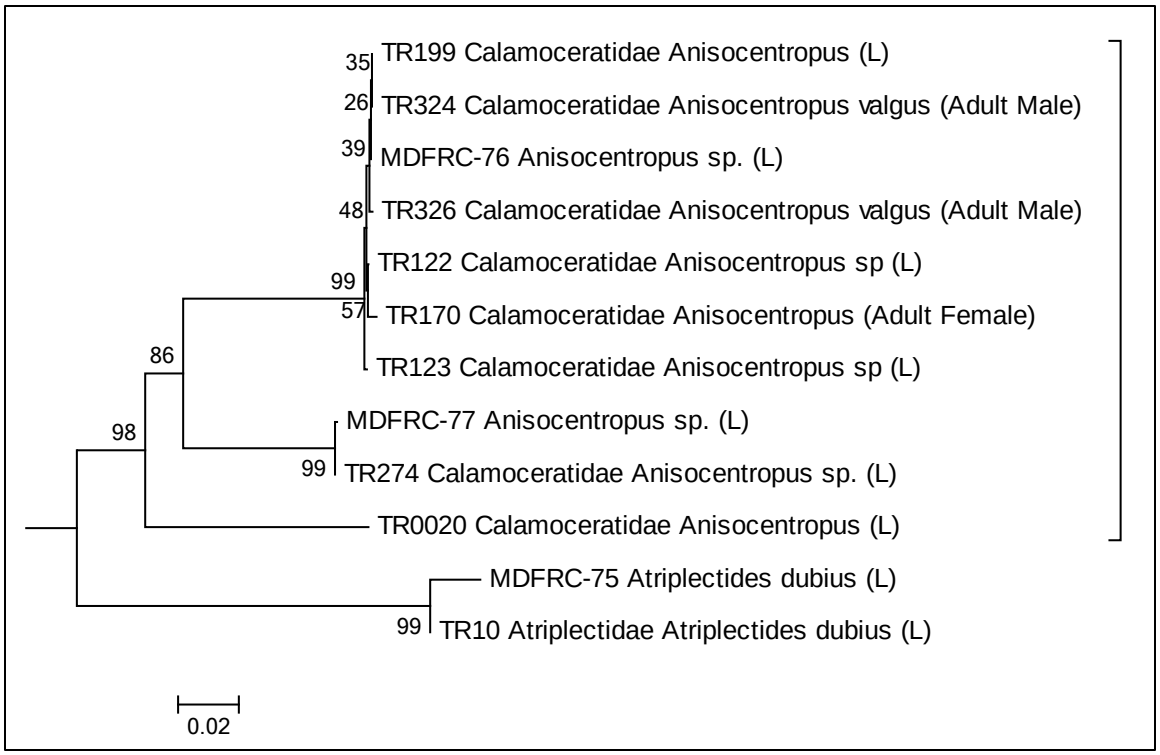


Figure 7. The COI gene tree for specimens of Calamoceratidae with *Atriplectides* as an out-group, collected as part of both MDBA funded barcoding projects

Philopotamidae

During the previous MDBA funded barcoding project, ambiguous larvae, identified as being either *Chimarra australica* or *Chimarra monticola*, were assigned to the species *C. australica* based on the similarities in gene sequences between adult and larval specimens. The present study provided further support that these larvae are *C. australica* and not *C. monticola* through the collection and comparison of the sequence of a specimen of *C. monticola*. Data from the present study also suggests that a larval type identified as species AV3 is also a variant of *C. australica* (Figure 8). To firmly establish this association further examination and specimens will be required.

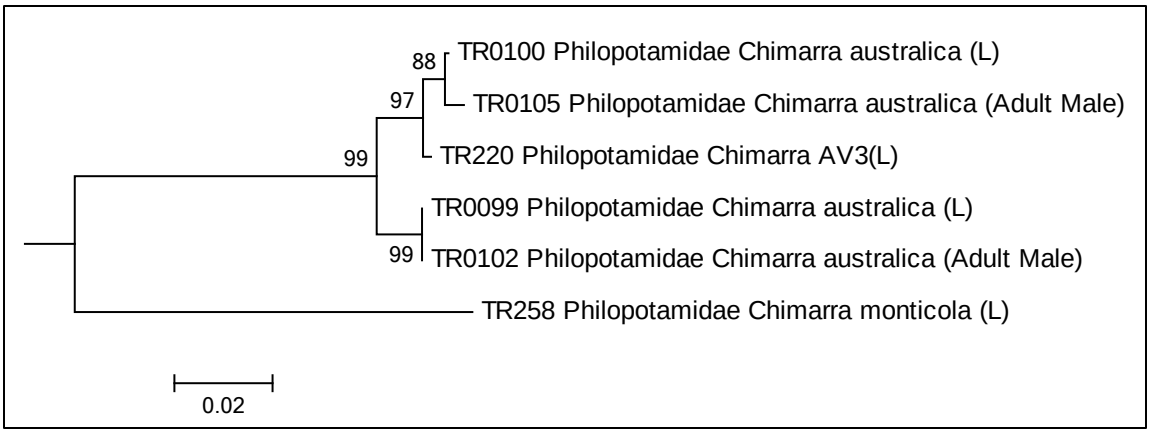


Figure 8. The COI gene tree for specimens of Philopotamidae collected as part of both MDBA funded barcoding projects

Conoesucidae

The present study has resulted in the collection of two new species of Conoesucidae and the association of two ambiguous larvae. As yet, the larvae Coenoria species AV5 and Conoesucus species AV5 have not been associated to any known species. The data suggest that all three larvae belong to an, as yet, undescribed species.

Presently there is only one described species of *Coenoria*. The data collected as part of this study suggests that at least two species occur in the region sampled (Figure 9). Further data from the private caddisfly database suggests at least five species exist. Descriptions of the species discovered here will require collection of more adult males.

Five species of *Conoesucus* from Tasmania and one from New South Wales are currently described. Examination of the male specimen TR310, collected during the present study, indicated the presence of a further species occurring in Victoria (Figure 10). Furthermore, similarities between the gene sequence of this species and that of *Conoesucus* AV5 suggest an association between the two. However, two specimens collected during the present study (TR259 and TR260), identified as *Conoesucus* species AV5, occur as a distinct clade on the gene tree, separate from the new species. This suggests that our current understanding of *Conoesucus* species AV5 includes two cryptic species. Forming a description of the new species will require examination of more adult males, and associating the adult with the larval form will require discerning morphological differences between the two larval AV5 types.

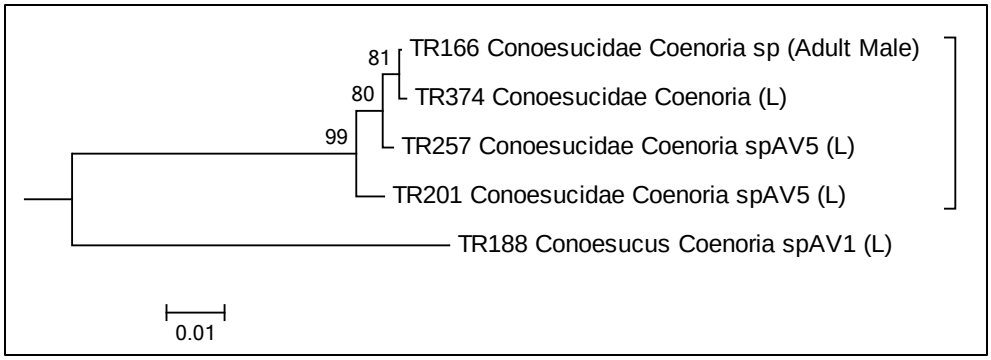


Figure 9. the COI gene tree for specimens of *Coenoria* collected as part of both MDBA funded barcoding projects

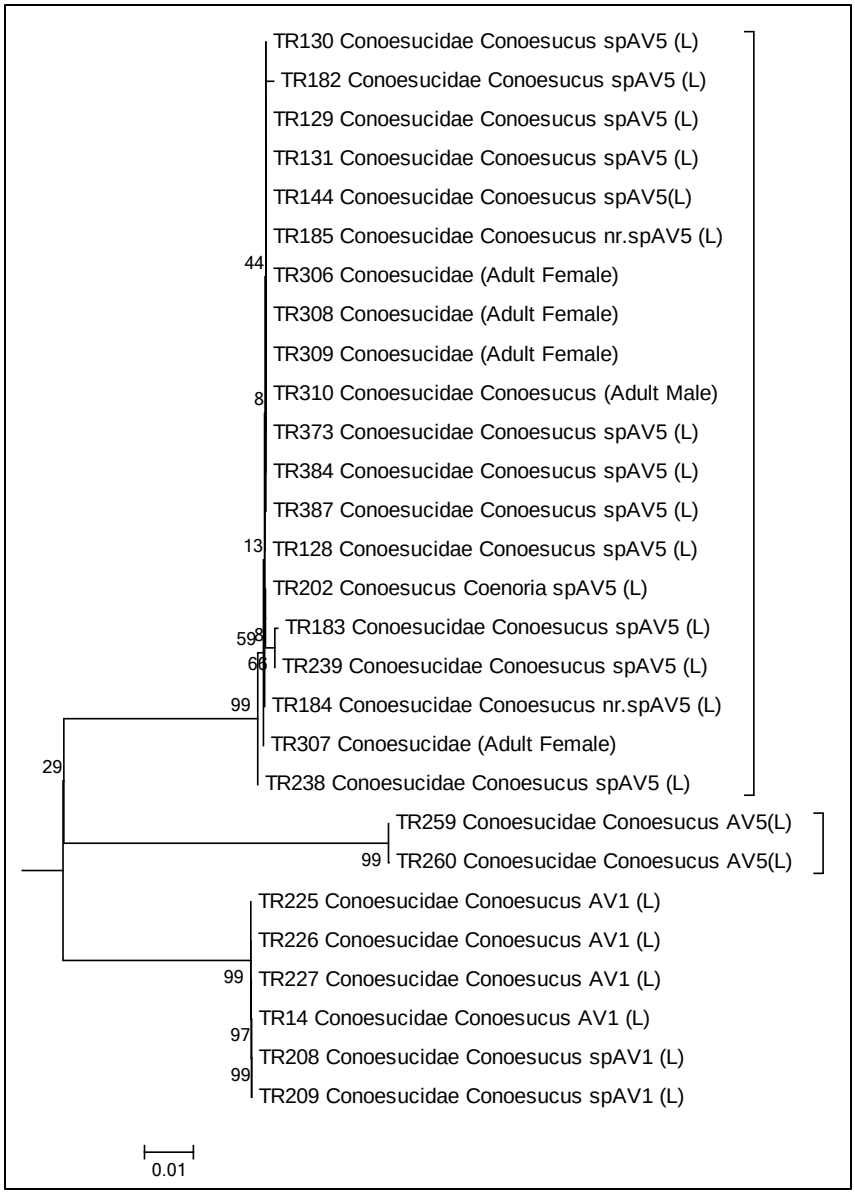


Figure 10. The COI gene tree for specimens of *Conoesucus* collected as part of both MDBA funded barcoding projects

Hydroptilidae

Hydroptilidae are often referred to as micro-caddisflies due to their small size, and are difficult to identify. For some genera in the family, keys to the larvae are lacking and as a result identification of these animals is generally only to genus level. Through the MDBA funded barcoding projects, we have successfully amplified COI sequences from nine species (Figure 11). *Hellyethira simplex* is one species for which an association has been made and the identification of larvae can be taken to species level. However, our genetic data suggests that the current concept of *H. simplex*, in as far as the larvae are concerned, includes two species. Specimens TR117 and TR178 occur as a distinct taxon on the COI gene tree. The data also confirmed the species identity of larvae identified as *Hellyethira scamandra*. As more data is obtained on these taxa, it may be possible to resolve many of the taxonomic issues around the placement of ambiguous larvae.

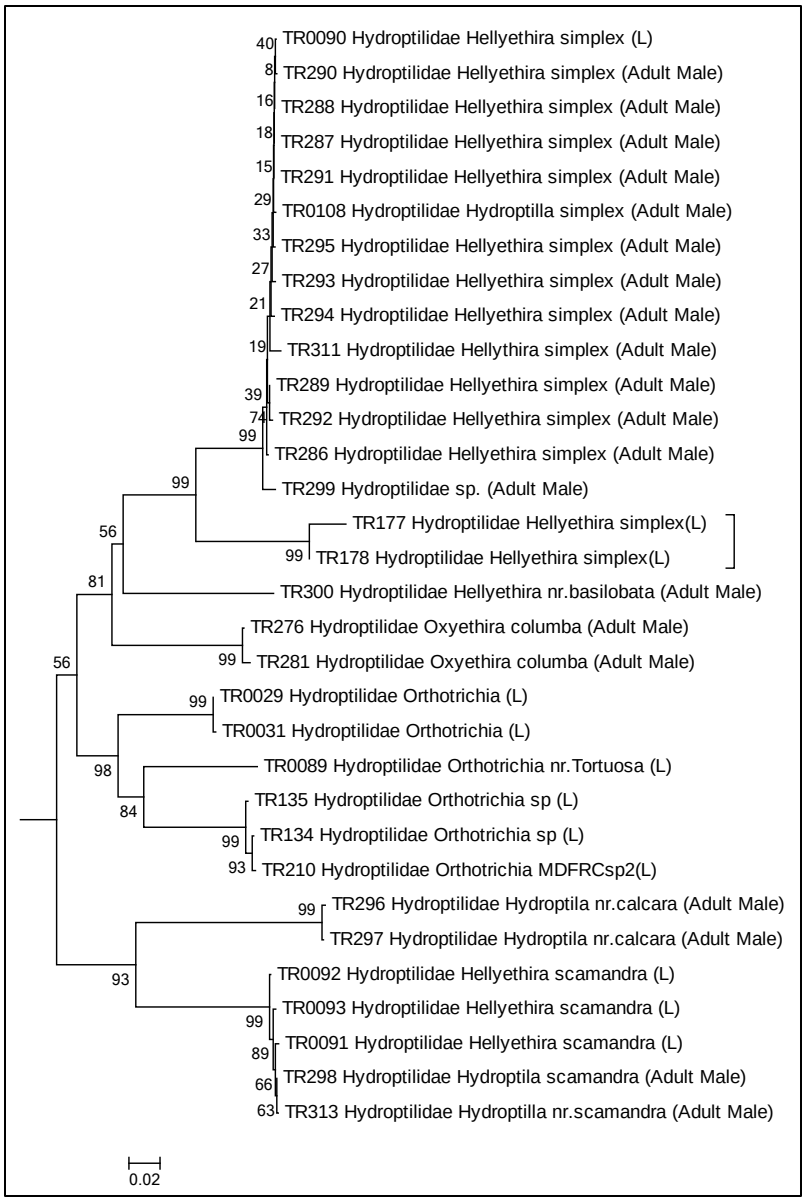


Figure 11. The COI gene tree for specimens of Hydroptilidae collected as part of both MDBA funded barcoding projects

Hydrobiosidae

Five species of *Ulmerochorema* were collected during the current study (Figure 12). Three of these species are described (*U.* membrum, *U.* rubiconum and *U. lentum*) with a fourth species identified as being similar to *U. tasmanica*, a species known only to occur in Tasmania. It is possible that this specimen belongs to an, as yet, undescribed species that shares some morphological similarity to *U. tasmanica,* or this may be the first record of this species on the mainland. A fourth species, occurring as a clade at the bottom of the gene tree in Figure 12, is likely a new species. Confirmation of this will need to be established and if found to be true, these specimens will be used towards describing the new species.

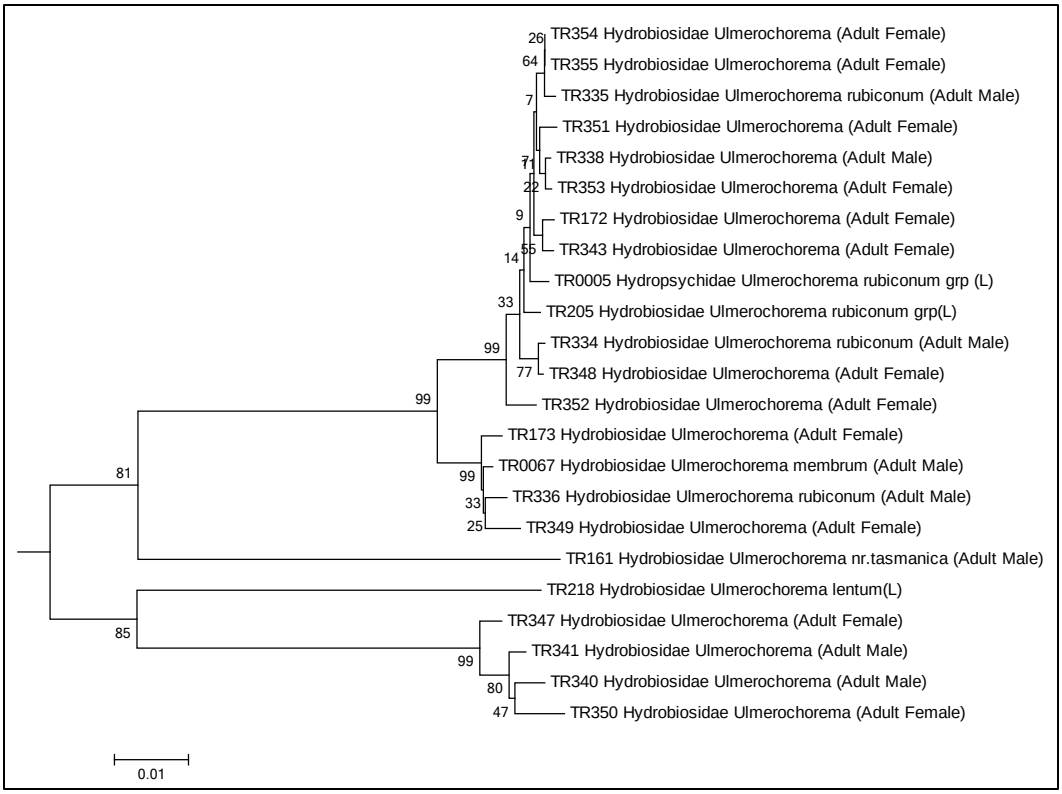


Figure 12. The COI gene tree for specimens of Ulmerochorema collected as part of both MDBA funded barcoding projects

2.4 Concluding remarks

The 271 specimens identified as part of this study are a valuable addition to the macroinvertebrate database and will provide important information for taxonomists and researches working with these taxa. From the Trichoptera data alone we have discovered a possible nine new species, based on adult material, and a further three species from larval material. These will be made available to appropriate taxonomists. If they are found truly to be unique species, they will be described, adding to our understanding of the diversity of aquatic biota.

This work is of a type that attracts value-adding, as the data that it has generated can be readily used in a variety of research areas. The data provided here, along with that already available in databases such as BOLD and GenBank, will form the backbone to answer future DNA based ecological questions, to describe new species, and to investigate future methods employed in biological monitoring programs.

3 Component 2: Hybrid sequence capture

3.1 Introduction

Recently, high-throughput DNA sequencing has been recognised as having the capacity to be applied to the area of rapid biodiversity assessment (RBA). The goal of RBA is to characterise the biodiversity of a sample or site in a rapid and cost effective manner. Through applying genetic techniques it may be possible to characterise biodiversity through detecting the presence of species specific gene sequences in samples. This method has the potential to reduce the extensive man-hours currently employed in biological studies, just in identifying individual specimens.

Current DNA techniques for detecting the presence of species from environmental samples generally rely on detecting single indicator genes. Such analyses are limited in the strength of data they produce and may provide false negative results by failing to detect the indicator gene. A better approach is to analyse multiple indicator genes in order to increase the chance of detection; however, this is currently costly and time consuming.

HSC offers a cost effective way of simultaneously detecting multiple indicator genes from environmental samples. In theory, the method uses gene fragments of known species to create probes that can then be used to capture multiple indicator genes from a greater fauna. The simultaneous creation of multiple probes greatly reduces the cost and increases the likelihood of detecting a species from a sample. Current techniques employ single probes that are synthetically created. Using multiples of these probes quickly becomes expensive. Furthermore, the use of single probes, as is currently used, results in a trade-off between detecting a breadth of biodiversity and taxonomic resolution. Probes that are conserved enough to amplify DNA fragments that are shared between a wide range of fauna tend to be found in more conserved areas of DNA and so lack the power to discriminate between species of the same genus or family. In contrast, more specific, less conserved primers work well at delineating species, but are likely not to work over a broader taxonomy, such as between orders. An advantage of using the HSC method is that the probes do not need to be identical to their targeted DNA, allowing for highly divergent regions, with the capability of discriminating between species, to be captured if they are flanked by more conserved regions where the probes can bind.

Our greater aim is to investigate the feasibility of developing this tool through investigating its performance in detecting target DNA within samples that include: 1) single species, 2) mock environmental samples, and 3) real environmental samples that contain some known species. However, it was recognised that much of the downstream investigations into this tool relied on the successful amplification and sequencing of large fragments of DNA, specifically whole mitochondria, and so this project was set up as a feasibility study. At the time of preparing this report, we have succeeded in amplifying the whole mitochondrion of five species. Sequencing of these species has not yet been undertaken, but is in process. While the study has not been completed, the first step in HSC has been achieved, demonstrating that this method may represent a feasible technique to examine communities from simple environmental samples.

3.2 Methods and results

Five specimens from five separate species were selected for DNA extraction and amplification based on their taxonomy. Three specimens were selected from among the Decapoda to be used to test the application of the method within an order. These were a yabby (*Cherax destructor*), a shrimp (*Paratya australiensis*), and a prawn (*Macrobrachium australiense*). One whirly-gig beetle (*Macrogyrus oblongus*) and one stone fly (*Cosmioperla kuna*) were selected as members from distantly related taxa.

3.2.1 DNA extraction

Two methods were used for DNA extraction. First, a non-destructive method was employed in order to retain whole specimens following the extraction process. For this method DNA was extracted from a whole animal using the commercial extraction fluid XYTXtract (www.ande.com.au) and following the protocols outlined in Castalanelli *et al.* (2010) except that samples were heated for two minutes rather than 20 minutes. This method does not remove impurities, such as RNA and proteins, from the final product. To reduce the effect these impurities may have on the amplification success, a one in five dilution of the extracted product was made by diluting 5 µl of product in 20 µl of sterile water.

The second method employed the use of the DNeasy Blood and Tissue Kit (www.Qiagen.com). Extractions were performed on part of a specimen, usually a leg, following the manufacturer's instructions. This method is destructive, so the material use was destroyed in the process. The final product is a clean genomic sample relatively free from impurities.

3.2.2 DNA amplification

The whole mitochondrion was targeted using the primers HPK16Saa and HPK16Sbb, developed by Hwang *et al.* (2001). The Invitrogen amplification kit Platinum Taq DNA Polymerase High Fidelity was used to amplify the mitochondrion following the manufacturer's instructions. Five microliters of amplified material was run on a 2% Agarose gel, containing 50 g Agarose and 5 µl SybrSafe, for 60 minutes at 80 volts. Successful amplifications were those that fluoresced when viewed over a black light. The strength of fluorescence indicated the amount of amplified material and was graded from 0 for a negative result to 3 for the strongest result. Results were recorded.

First attempts were conducted on the 1:5 dilutions and the raw extraction products of the XytXtract material. These failed to produce results under a range of thermal regimes including setting the annealing temperature to decrease by one degree with each cycle from 65 to 55 degrees. Each amplification attempt took around eight hours or more.

The second attempt was conducted on the material extracted using the DNeasy Blood and Tissue kit. These amplifications produced some success. An attempt at amplification was conducted on each specimen four times. Each specimen yielded results but not always consistently (see Table 1). The only specimen to produce positive results four times was the yabby.

Table 1 Results of the success rate of amplifying whole mitochondria of five specimens

Specimen	Replicate	Result (score from 0-3)
Yabby	1	3
Yabby	2	3
Yabby	3	3
Yabby	4	3
Prawn	1	0
Prawn	2	0
Prawn	3	2
Prawn	4	3
Shrimp	1	3
Shrimp	2	0
Shrimp	3	3
Shrimp	4	1
Beetle	1	3
Beetle	2	0
Beetle	3	3
Beetle	4	0
Stone Fly	1	1
Stone Fly	2	3
Stone Fly	3	1
Stone Fly	4	0

3.2.3 DNA sequencing

The whole mitochondrion PCR products were mailed to the CSIRO in dry ice for purification and sequencing. At the time of this report, sequencing is being undertaken. Replicates have been pooled so as to provide 50 µl of 10 ng/µl solution. The shotgun sequence method is being employed where the complete mitochondrion is fragmented into short random and overlapping fragments, which are cloned and sequenced. The resultant sequences are then assembled, in-silico, to reproduce the whole mitochondrial genome sequence.

3.3 Discussion

The successful amplification and sequencing of specimens from a range of taxa was an integral component to the testing of the HSC method. Before HSC probes can be developed and tested, it is important to know the sequence of the organism used to make the probes. This aids an understanding of where in the mitochondrion the probes came from and confirms that it is the mitochondrion of this particular species that has been amplified and not a contaminant. Furthermore, in order to assess if the probes do capture DNA from test species, it is important to be able to compare the captured DNA to that of the previously sequenced mitochondrion of the test specimen. This component of the overall project has provided the integral first step to testing the applicability of the HSC method.

One of the main time constraints with this component was in the amplification of genetic material. Amplification protocols need refining to attain results, and this involves somewhat blindly testing thermal regimes. Furthermore, each amplification attempt requires large amounts of time to run. The study successfully amplified DNA and this is now in the process of being sequenced. It is our intent to continue this aspect of work once the sequence information becomes available.

4 Component 3: DNA longevity experiment

4.1 Introduction

The use of DNA barcodes in identifying aquatic invertebrates is a relatively new but now well accepted technique. However, the current practice of sequencing DNA from single specimens to gain identification is not feasible to apply to large-scale biological monitoring programs. For this reason, recent developments utilising high-throughput next generation sequencing (NGS) technology, such as pyrosequencing and shotgun sequencing, have been investigated as ways to address this issue. These methods have the potential to rapidly characterise the fauna of given sites simply by analysing environmental samples (e.g. soil or water samples) for the presence of species specific DNA fragments. Research investigating the potential use of NGS has found the method useful. For instance, Carew *et al.* (2013) used 454 pyrosequencing to investigate Chironomidae fauna and found that 96% of species could be identified with a single gene fragment (COI) and that this increased to 99% with the inclusion of a second fragment (CytB). Ficetola *et al.* (2008) was able to demonstrate that the presence of bullfrogs in a water body can be determined through detecting gene sequence specific to this species simply from water samples where bullfrogs had been present.

It is clear that the current technology can be utilised for the detection of species from environmental samples. However, while testing the presence or absence of species under controlled laboratory conditions has offered promising results, questions arise as to how DNA behaves under natural conditions, such as within flowing rivers. Before these methods can be put to use in rapid bio-monitoring for riverine systems, it is important to understand how long DNA persists in the environment, how easily it can be collected, and how far it is likely to have travelled. The distance DNA travels will likely be a product of 1) how long DNA persists within water before it degrades due to shearing forces, 2) how readily it is decomposed or assimilated back into the food web, 3) how water chemistry might affect the detection success of DNA, and 4) the characteristics of the river's flow i.e. slow and meandering, or fast and turbulent. Few investigations have looked into DNA longevity in aquatic environments. England *et al.* (2005) found that a 530 base pair (bp) gene fragment persisted, in detectable amounts, in a field microcosm for about 24 hours. In contrast, (Deere *et al.* 1996) found that a 423 bp fragment of extracellular DNA was detectable in loamy substrate water samples for around four weeks and in sandy substrate water samples for about seven weeks, under laboratory conditions. Dejean *et al.* (2011) found that a 98 bp fragment of Sturgeon DNA and a 79 bp fragment of bullfrog DNA remained detectable for less than one month.

The study hypothesised that the chemical and biotic properties of the water in which DNA is being detected may play a role in the longevity of detectable DNA. In particular, the study investigated if DNA persistence was influenced by:

1) The presence of nutrients
2) The presence of carbon
3) The presence of both carbon and nutrients
4) The presence of clays

A simple method was chosen for amplifying the DNA COI barcoding region, a 658 bp region regularly used for species detection. Due to unintended PCR inhibiting agents in some samples, comparisons involving the carbon and clay treatments could not be made. Furthermore, the unexpected contamination issues from non-targeted microorganisms hampered attempts at measuring the final decay of DNA in the samples. However, the data suggested that the COI barcode probably remained within the system for somewhere between six and 14 days. The study also produced some inadvertent findings such as the possibility that a bacterium or bacteria occurring in freshwater

samples may be readily detectable using classical COI barcoding techniques. This finding raises some concerns for the use of barcoding environmental samples.

4.2 Methods

Water from fish tanks, used for housing crustaceans and fish at the MDFRC, was passed through 50 µm mesh netting to remove micro and macroinvertebrates. The water was collected into a 40 L tub and stirred to homogenise particulate matter. Eighteen sterilised, 2 L Schott bottles were filled with 2 L of this water. To each of these, one of six treatments was added, thereby replicating each treatment three times. A further three Schott bottles were filled with 2 L of sterilised water, to be used as negative controls. A diagram of the treatments and the sampling regime is given in Figure 13.

Schott bottles that were treated with nutrients were done so by adding 0.002 g of commercial Thrive. This was achieved by dissolving 0.02 g of commercial Thrive in 10 ml of sterile water and adding one millilitre of this stock solution to each Schott bottle.

A stock solution of leaf leachate was created to be used as a source of dissolved organic carbon (DOC) for experimental treatments. Dried leaf material weighing a total of 5 grams was added to 100 ml of water and left to leach for 24 hours. 20 ml of this stock solution was added to each of six Schott bottles to give a concentration of 30 mg DOC/L. Three of these bottles contained only tank water, while the remaining three were treated with Thrive. Due to problems with PCR inhibition present in the DOC treatments, a further six 1 L Schott bottles were dosed with 1 g/L of glucose instead of the leachate derived carbon, with three of these containing 0.002 g of Thrive

To investigate the effect of turbidity on the PCR ability of DNA, clay was added to six Schott bottles at two concentrations. Three Schott bottles were treated with clay to 60.8 Nephelometric Turbidity Units (NTUs) and three to 310 NTUs, as measured using a Quanta Hydrolab multimeter. The untreated water measured 6.4 NTUs.

Three bottles contained autoclaved, sterile water and were used as negative controls in order to assess the likelihood or extent of DNA contamination from external sources.

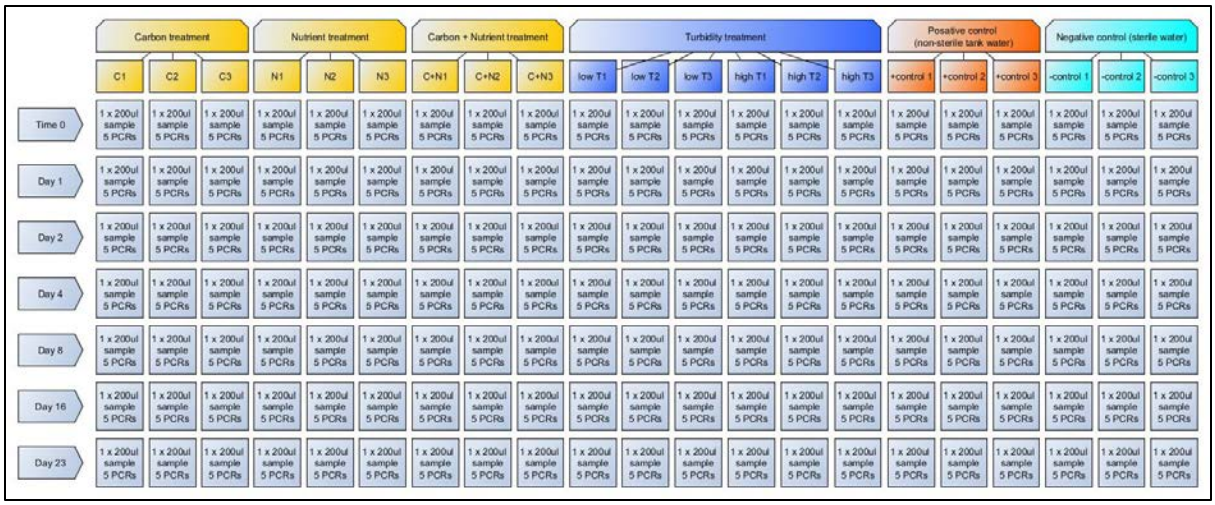


Figure 13. Diagram of the study design. Each treatment is replicated three times and each replicate is subsampled for PCR 5 times on each of the sampling days. Sampling days increment exponentially except for the final sample which was taken at day 23 to further investigate anomalies in the data on day 16.

One 200 µl sample was taken from each Schott bottle on days 0,1,2,4,8,16, and 23 (see Figure 13). From each of these samples, five Polymerase Chain Reactions (PCR) were conducted using 11.5 µl of sample water as the DNA template. A 658 base pair (bp) fragment of the Cytochrome Oxidase subunit 1 (COI) gene was amplified. The PCRs were undertaken using the *Taq* PCR Master Mix from QIAGEN. The PCR cocktail consisted of 12.5 µl of *taq* PCR Master Mix, 0.5 µl of each of the primers LCO1490 and HCO2198 (Folmer *et al.* 1994). Positive and negative controls contained 10.5 µl of sterile water, with the positive also containing 1 µl of a previously tested and amplifiable genomic DNA template. The PCR thermal regime followed that of Webb and Suter (2010). Agarose gel electrophoresis was used to test for successful amplification of template DNA. PCR products that provided a discernable band on the agarose gel were considered to be positive results, and those where no band was present were considered to be negative results.

4.3 Results

Tank water (positive control) and nutrient dosed tank water treatments provided positive results from day one of the experiment. The results of these two treatments are graphed in Figure 14. All negative controls did not return any positive results. Overall, percent success decreased steadily for the first eight days and then increased, with the decrease over the first eight days more rapid than the increase during the following 15 days.

At day zero the control samples yielded only an 86.6% success rate, due to two PCRs failing to amplify products. For this reason, the control data peaks at day one. The percent success of amplification then decreased rapidly until day four, where success was only 26.6%. By day eight the percent success began to rise again.

In contrast, the nutrient dosed samples began with 100% success on day zero and declined steadily until day eight. While the percent success matched closely with that of the control for the first two days, on day four the nutrient samples had a 60% chance of success as opposed to 26.6% for the control samples.

Both leachate and turbidity treatments failed to yield expected positive results on days zero and one, and were not analysed further. In a modified experiment, where glucose was used as a carbon source, removing any inhibitory effect from the leachate, amplification tests provided positive results on day zero. However, on day four, while amplification in the control treatments declined, the glucose treated samples provided increasingly strong positive results. This experiment was not continued as microscopic examination of the samples indicated that the strong signals were a result of the PCR amplifying DNA from organisms growing in the samples. The glucose was probably supplying food for microinvertebrates to grow, increasing their number, rather than DNA being degraded as anticipated.

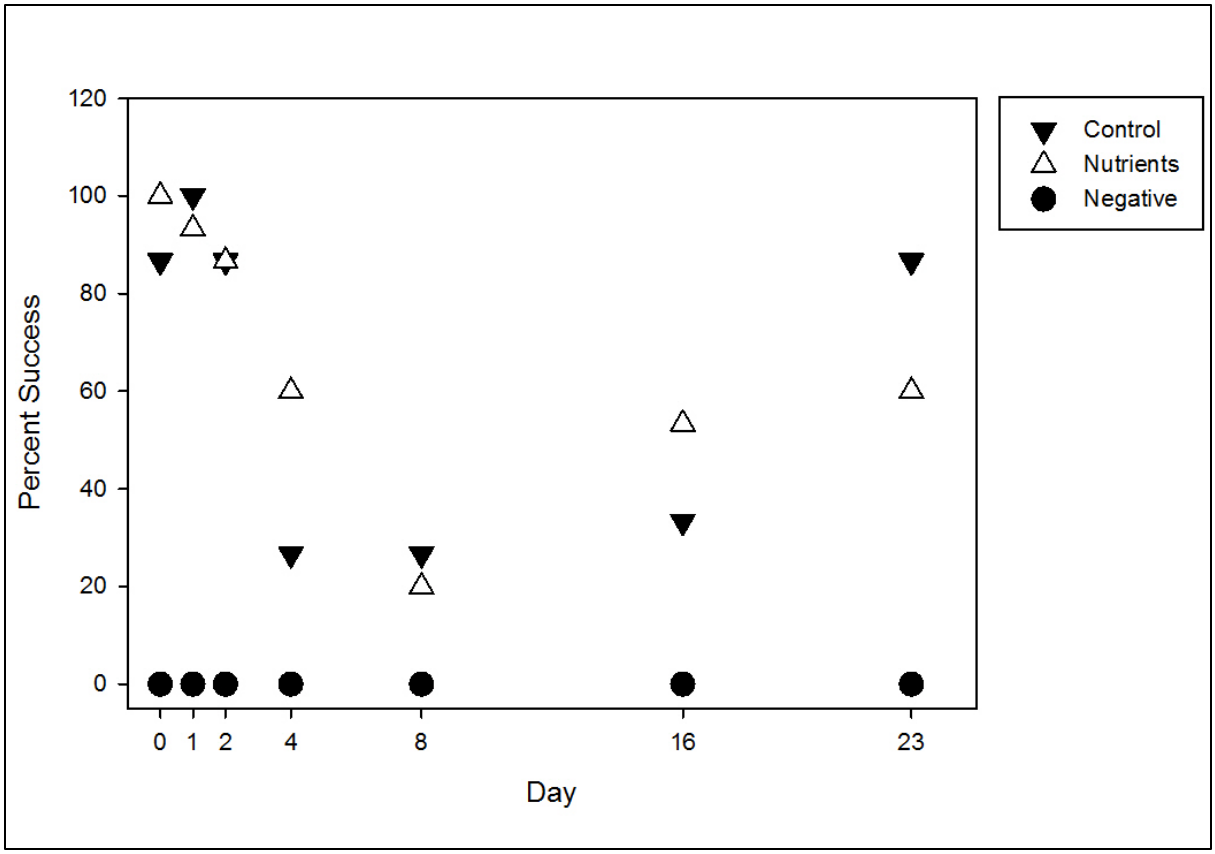


Figure 14. Line graph showing the percent success of amplifying from the template DNA as calculated as a percentage of the five PCRs performed per sampling day

4.4 Discussion

Determining the residence of DNA in water is hampered by choosing an appropriate experimental system. An ideal scenario would be to carry out the experiment in a small stream, but the scale of water flow and dilutions means it is difficult to add or measure a specific type of DNA in the water. Lab systems alleviate the latter problem, but the small containment of water samples then creates the opportunity for contaminant organisms to grow, obscuring results. Our data suggests that the COI fragment remained detectable in the water until somewhere between day four and 14. However, due to the increase in DNA following day eight, we are unable to determine if a zero percent success rate exists. For the nutrient treated samples, a straight line drawn from the data point at day zero to day eight suggests that the DNA should have become undetectable by around day 10. Increases in DNA amplification success following day eight are most likely due to organisms growing within the samples. Our original assumption was that the COI fragment would be quickly consumed by bacteria and other organisms. We had also assumed that the primers used would not amplify these micro-organisms. Our data suggests that the COI barcoding fragment can be successfully amplified from at least one of these micro-organisms. A strong candidate would be from among the bacteria that were present in large numbers in the glucose dosed samples. Further culturing out of various strains of these bacteria in the future will test this hypothesis. If correct, it will be particularly important to choose an appropriate gene so bacteria are not included in the assay system. The leachate treatments did not yield results, possibly due to the presence of humic matter, which is known to inhibit PCRs. The study design could have been improved by filtering out the humic matter, however, we were interested in examining free DNA in solution in the presence of

dissolved organic carbon and there was no surety that filtering would remove the dissolved contaminants. Given that the pilot of our method worked for straight tank water, which contained some amount of humic matter, we did not expect this experiment to subsequently fail.

It is difficult to interpret the results from our first turbidity experiment. It is possible that all the DNA was immediately bound to the clay particles and that once bound, DNA amplification could not take place. The binding to clay may also cause reactions with other materials also bound to the clay (e.g. DOC) and prevent amplification. Interestingly, the tank water contained low levels of turbidity, yet yielded results. This may be because the tank turbidity was not largely derived from a clay-based source, or that there is a lower threshold where DNA detection can still occur.

While answers to our hypotheses were clear, the component did suggest a longevity for the barcoding fragment of about 10 days in the presence of elevated nutrients. This may not be too far removed from the decay under natural conditions. The failure to PCR straight water with elevated clay and leachate levels suggests that this will need to be taken into account when collecting DNA samples from rivers. The presence of a possible bacterium that the standard LCO/HCO barcoding primers are able to detect is an interesting result and will be followed up in further investigations. Such bacteria are known from marine samples (Siddall *et al.* 2009). However, to our knowledge, this is the first account in a freshwater sample.

5 Conclusions and recommendations

The three components here provide important steps towards developing tools for cost effective and accurate identification of aquatic macroinvertebrates. Component 1 delivered 271 new barcode sequences for aquatic invertebrates and deposited these in a public database. The development of COI barcode databases provides the backbone for future advancements in genetics-based approaches to species identification. The importance of developing a database of the Australian aquatic invertebrate fauna has been largely recognised and in October 2014 a consortium of researchers and business partners agreed to collaboratively develop such a database. The sequence information gathered as part of Component 1 will contribute to this database and so will aid in developing an important national repository for DNA sequences.

Component 1 has also provided insights into our current understanding of the taxonomy of aquatic insects. In particular, it highlighted areas within the *Elmidae* where taxonomic delineations were not supported by the molecular evidence. Through the collection and genetic sequencing of Trichoptera, a possible nine new species have been discovered. These will be described and will further aid in our understanding of the diversity within this order.

Component 2 provided the first steps towards developing a HSC method that would have the power to detect genetically diverse species within a sample at a comparatively low cost. Through developing and testing PCR protocols we successfully amplified the whole mitochondrions of five species of macroinvertebrate. The sequence information provided by this project will directly feed into the development of the HSC method. Further development of this method will provide tools for more accurately detecting species from environmental samples in a cost effective manner.

Component 3 investigated the retention time of free DNA in water under varying conditions. In particular, it focussed on how long the COI barcoding gene fragment could be detected in a water sample. This information will be integral to understanding how DNA is transported through river systems. Our study suggested a retention time between 4 and 14 days. However, it also revealed a possible contamination factor that will need to be accounted for with the future development of eDNA detection. The signal from the COI barcoding gene fragment increased after day 8 and this was likely due to the amplification of bacteria that were rapidly growing in the water media. The use of eDNA to characterise the fauna of rivers will need to take such contaminants into account. By

determining the DNA sequence of bacterial contaminants it will be possible to exclude these sequences from eDNA analyses.

5.1.1 Recommendations for priority research

Future utilisation of COI barcodes for identification purposes requires the development of a reference COI barcode database. To some extent this has progressed and recent collaborative efforts look set to develop such a database on a national scale. However, data for much of the fauna remains missing. Some taxa are better represented than others and will require less effort to produce comprehensive databases for. Other taxa are missing altogether. Nationally, there are notable gaps in COI barcode data for Hemiptera, Odonata, and Coleoptera. We recommend that future research focus on filling these gaps.

As next generation sequencing technologies evolve, obtaining diagnostic barcodes (possibly from multiple gene fragments) will become more rapid and cost effective. Currently, developing the COI barcode database is costly and time consuming. Recent advances in NGS technology suggest that COI barcode sequences from multiple specimens could be obtained simultaneously. These methods will greatly reduce the cost of developing the COI barcode database. To our knowledge, these methods have not yet been developed and we recommend that future research should focus in this direction.

References

Carew ME, Pettigrove VJ, Metzeling L, Hoffmann AA (2013) Environmental monitoring using next generation sequencing: rapid identification of macroinvertebrate bioindicator species. *Front Zool* **10**, 1-15.

Cartwright DI (1997) *Preliminary guide to the identification of late instar larvae of Australian Ecnomidae, Philopotamidae and Tasimiidae (Insecta: Trichoptera). Identification guide number 10* Cooperative Research Centre for Freshwater Ecology, Thurgoona, NSW.

Cartwright DI (1998) *Preliminary guide to the identification of late instar larvae of Australian Polycentropodidae, Glossosomatidae, Dipsuedopsidae and Psychomyiidae (Insecta: Trichoptera). Identification guide number 15* Cooperative Research Centre for Freshwater Ecology, Thurgoona, NSW.

Castalanelli MA, Severtson DL, Brumley CJ, Szito A, Foottit RG, Grimm M, Munyard K, Groth DM (2010) A rapid non-destructive DNA extraction method for insects and other arthropods. *JOURNAL OF ASIA PACIFIC ENTOMOLOGY* **13**, 243-248.

Dean JC (1997) *Larvae of the Australian Hydrobiosidae (Insecta: Trichoptera). Identification guide number 11* Cooperative Research Centre for Freshwater Ecology, Thurgoona, NSW.

Dean JC (1999) *Preliminary keys for the identification of Australian Trichoptera larvae of the family Hydropsychidae. Identification guide number 22* Cooperative Research Centre for Freshwater Ecology, Thurgoona, NSW.

Dean JC (2000) *Preliminary keys for the identification of Australian caddisfly larvae of the families Antipodoeciidae, Atriplectididae, Limnephilidae and Plectrotarsidae* Cooperative Research Centre for Freshwater Ecology Albury, NSW.

Deere D, Porter J, Pickup R, Edwards C (1996) Survival of cells and DNA of Aeromonas salmonicida released into aquatic microcosms. *Journal of Applied Microbiology* **81**, 309-318.

Dejean T, Valentini A, Duparc A, Pellier-Curit S, Pompanon F, Taberlet P, Miaud C (2011) Persistence of Environmental DNA in Freshwater Ecosystems. *PLoS ONE* **6**, e23398.

England L, Pollok J, Vincent M, Kreutzweiser D, Fick W, Trvors JT, Holmes SB (2005) Persistence of extracellular baculoviral DNA in aquatic microcosms: extraction, purification, and amplification by the polymerase chain reaction (PCR). *Molecular and cellular probes* **19**, 75-80.

Ficetola GF, Miaud C, Pompanon F, Taberlet P (2008) Species detection using environmental DNA from water samples. *Biology Letters* **4**, 423-425.

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. *Mol Mar Biol Biotechnol* **3**, 294-299.

Glaister A (1999) *Guide to the identification of Australian Elmidae larvae, (Insecta: Coleoptera). Identification guide number 21* Cooperative Research Centre for Freshwater Ecology, Thurgoona, NSW.

Hinton HE (1964) A revision of the Australian species of Austrolimnius (Coleoptera: Elmidae). *Australian journal of zoology* **13**, 97-172.

Hwang UW, Park CJ, Yong TS, Kim W (2001) One-step PCR amplification of complete arthropod mitochondrial genomes. *Molecular phylogenetics and evolution* **19**, 345-352.

Jackson J (1998) *Preliminary guide to the identification of late instar larvae of Australian Calocidae, Helicophidae and Conoesucidae (insecta: Trichoptera). Identification guide number 16* Cooperative Research Centre for Freshwater Ecology, Thurgoona, NSW.

Lawrence JF (1992) *Coleoptera Workshop.* Murray-Darling Freshwater Research Centre, Albury, NSW.

Neboiss A (1986) *Atlas of Trichoptera of the SW Pacific-Australian region* W. Junk, Dordrecht, The Netherlans and Boston and Hingham, MA, USA.

Neboiss A (1992) *Illustrated keys to the families and genera of Australian Trichoptera I. Adults. ASL special publication number 9.* Australian Society for Limnology.

Shackleton M, Rees GN, Watson GO, Davey C (2013) Enhancing the quality and value of the MDFRC macroinvertebrate collection through application of genetic bar coding. Final Report prepared for the Murray-Darling Basin Authority by The Murray-Darling Freshwater Research Centre, MDFRC Publication, p. 29pp.

Siddall ME, Fontanella FM, Watson SC, Kvist S, Erséus C (2009) Barcoding bamboozled by bacteria: convergence to metazoan mitochondrial primer targets by marine microbes. *Systematic biology* **58**, 445-451.

St Clair RM (1997) *Preliminary guide to the identification of late instar larvae of Australian Philorheithridae, Calamoceratidae and Helicopsychidae (Insecta: Trichoptera). Identification guide number 12* Cooperative Research Centre for Freshwater Ecology, Thurgoona, NSW.

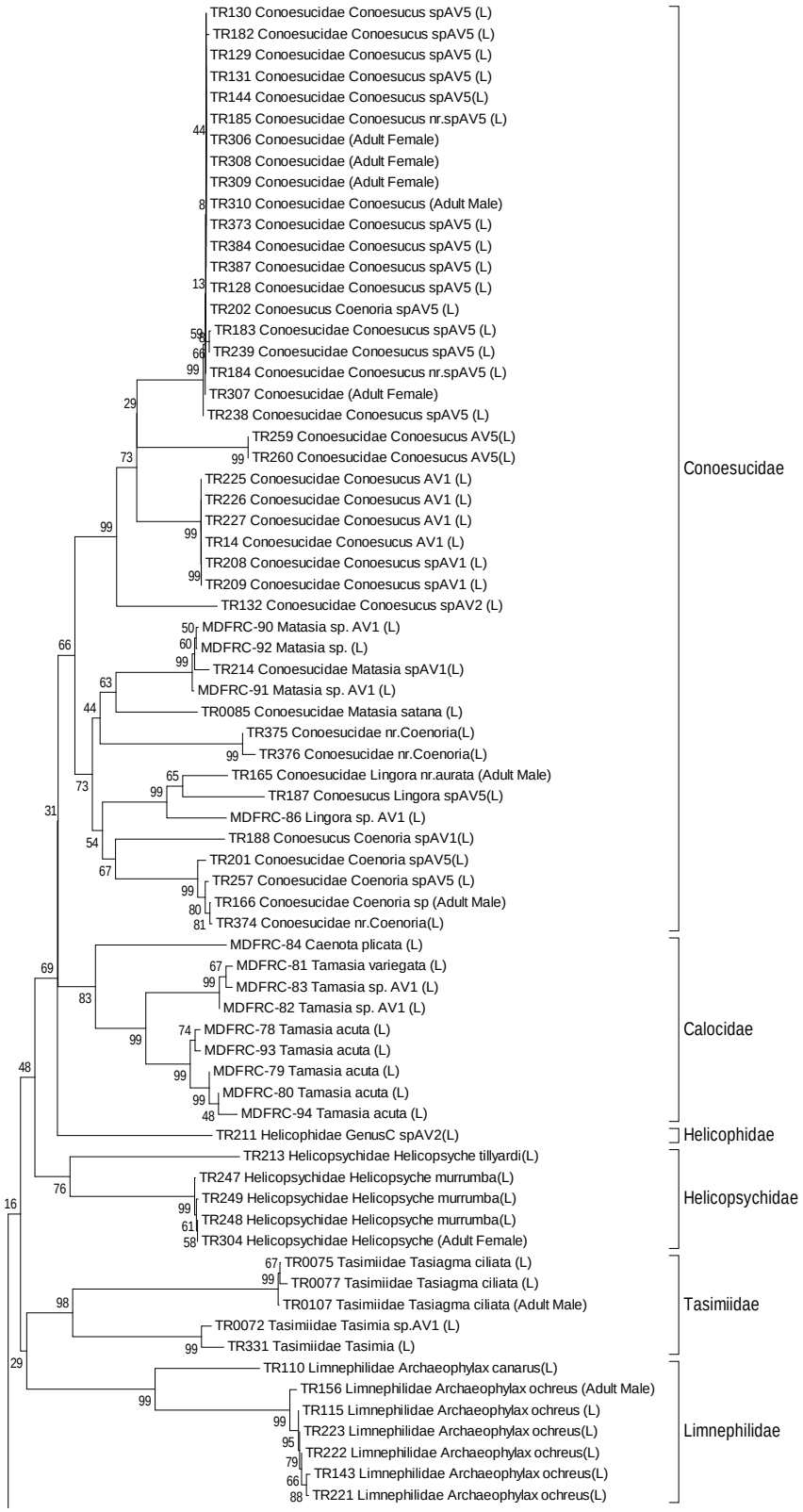
St Clair RM (2000) *Preliminary keys for the identification of Australian larvae of the family Leptoceridae. Identification guide number 27* Cooperative Research Centre for Freshwater Ecology, Thurgoona, NSW.

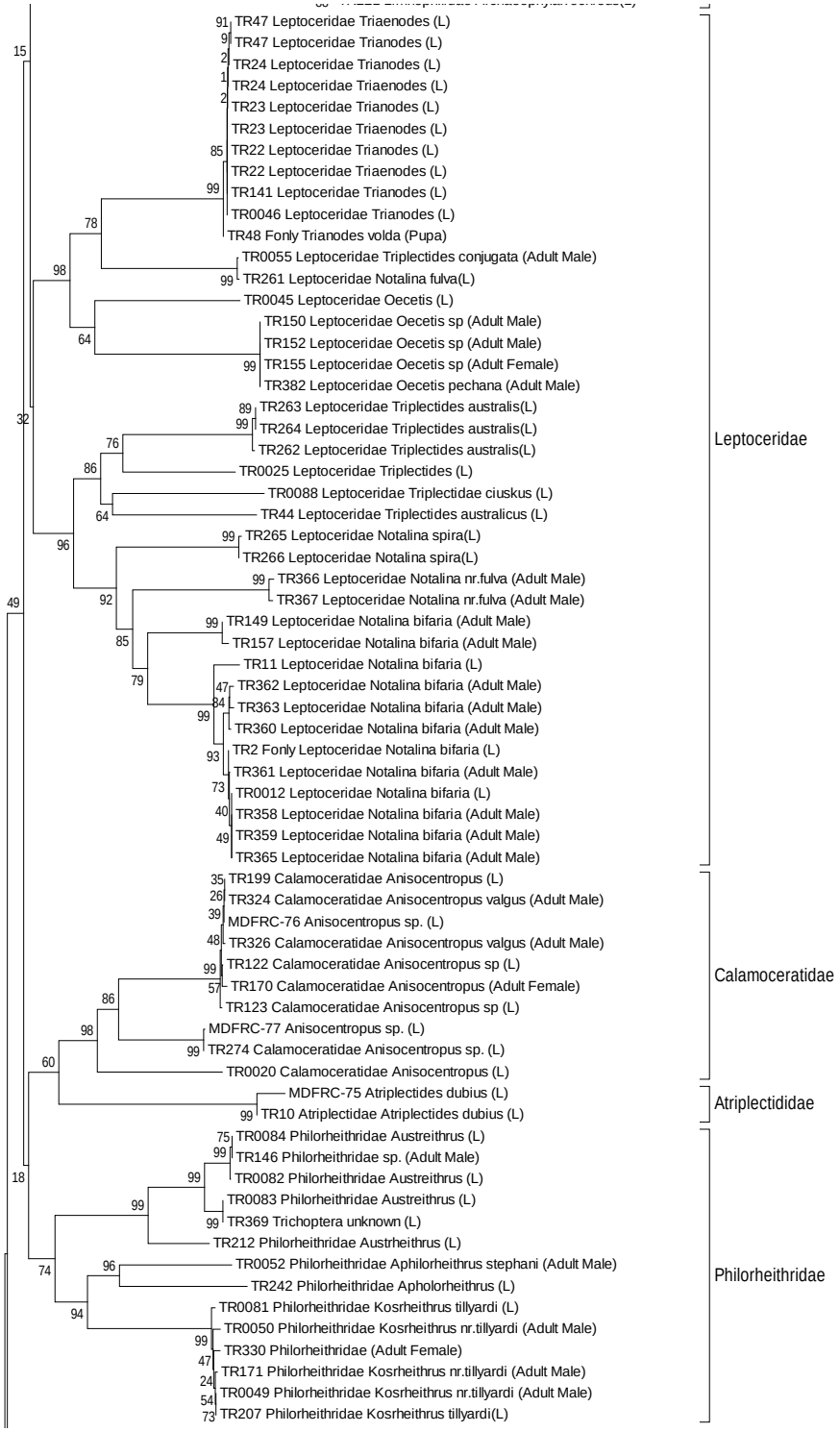
Webb JM, Suter PJ (2010) Revalidation and redescription of *Bungona illiesi* (Lugo-Oritz & McCafferty) (Ephemeroptera: Baetidae) from Australia, based on mitochondrial and morphological evidence. *Zootaxa* **2481**, 37-51.

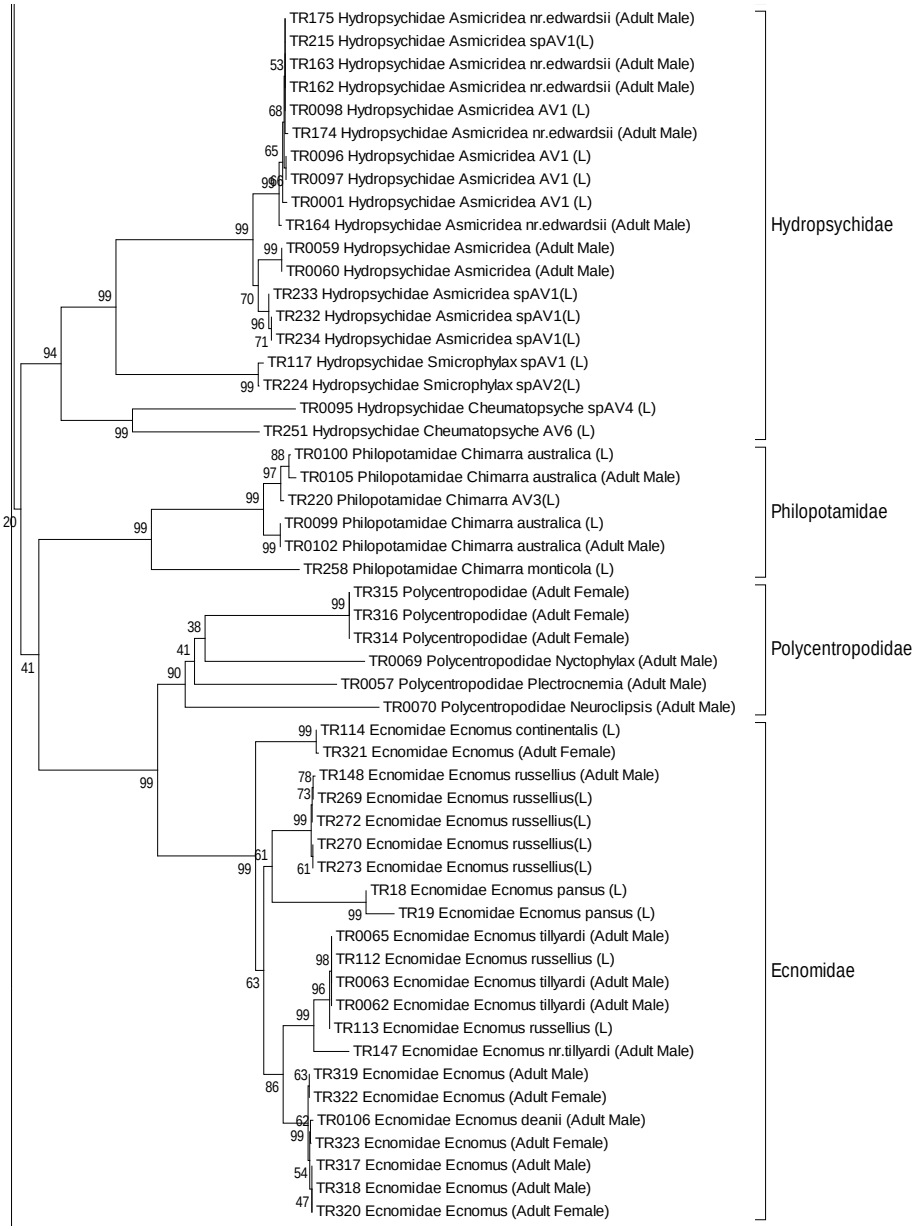
Wells A (1997) *A preliminary guide to the identification of larval hydroptilidae (Insecta: Trichoptera). Identification guide number 13* Cooperative Research Centre for Freshwater Ecology, Thurgoona, NSW.

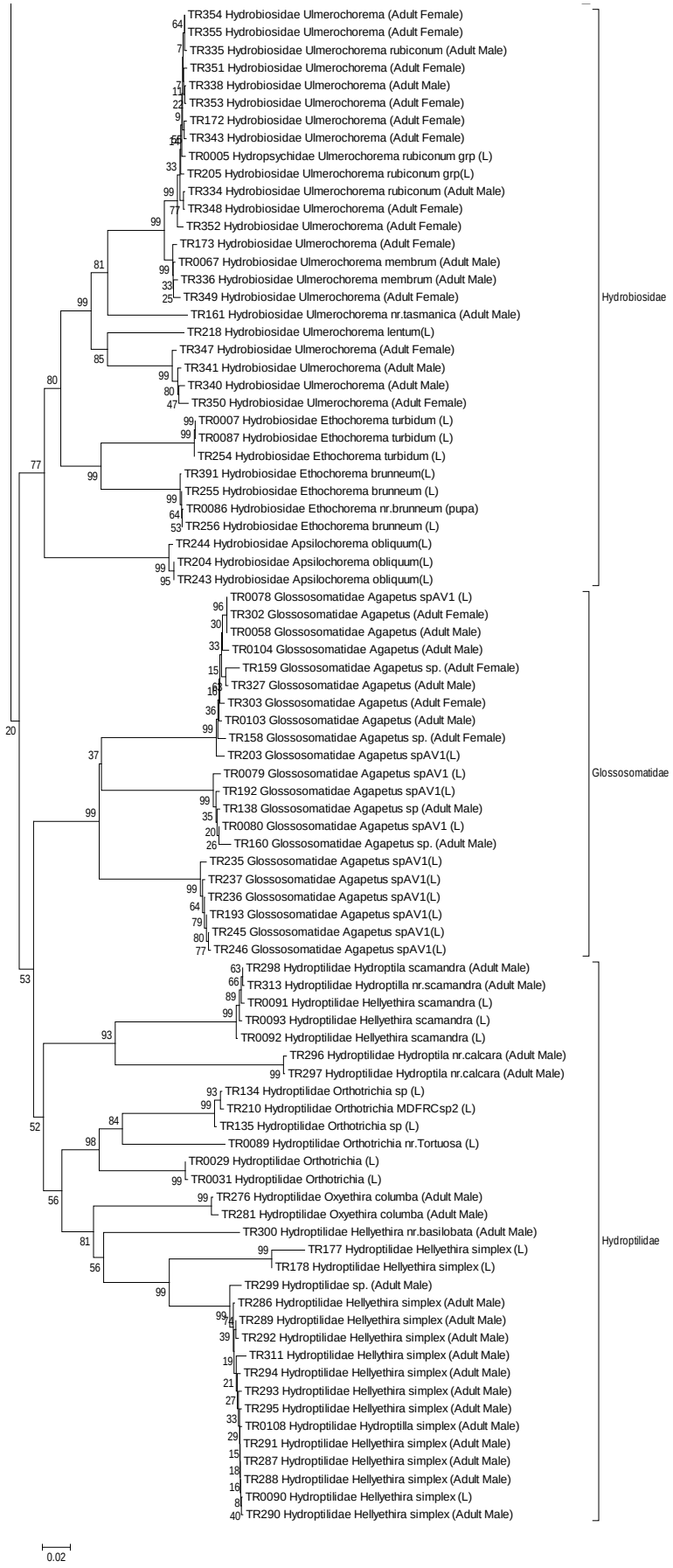
Appendices

Appendix A COI gene tree of Trichoptera specimens added to the DNA barcoding library

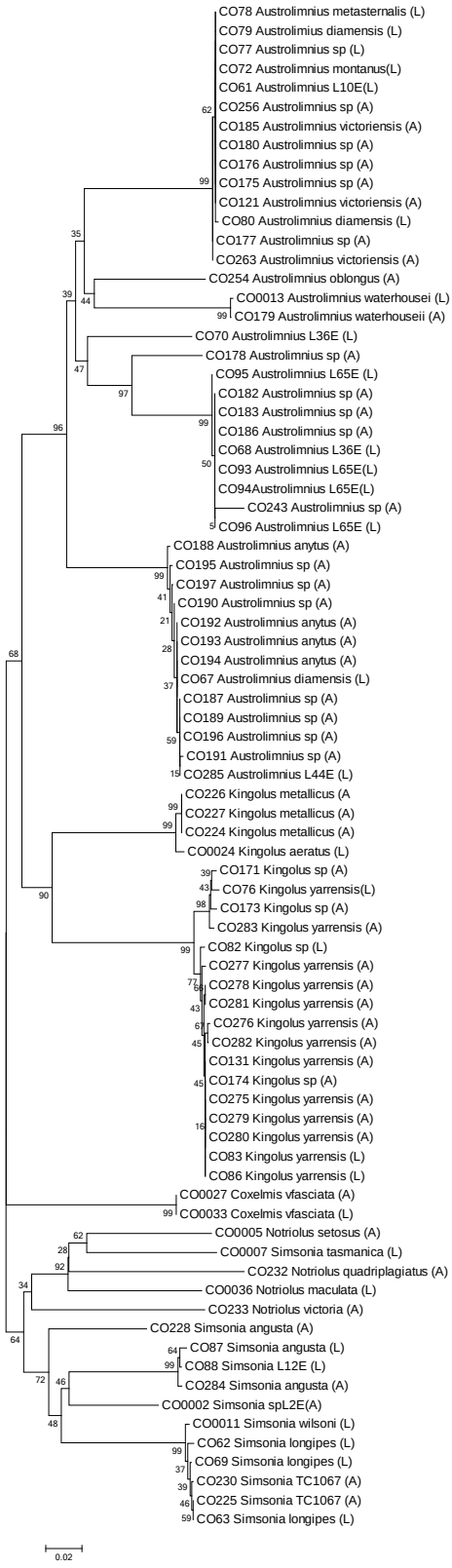



TR47 Leptoceridae Triaenodes (L)
TR47 Leptoceridae Trianodes (L)
TR24 Leptoceridae Trianodes (L)
TR24 Leptoceridae Triaenodes (L)
TR23 Leptoceridae Trianodes (L)
TR23 Leptoceridae Triaenodes (L)
TR22 Leptoceridae Trianodes (L)
TR22 Leptoceridae Triaenodes (L)
TR141 Leptoceridae Trianodes (L)
TR0046 Leptoceridae Trianodes (L)
TR48 Fonly Trianodes volda (Pupa)
TR0055 Leptoceridae Triplectides conjugata (Adult Male)
TR261 Leptoceridae Notalina fulva(L)
TR0045 Leptoceridae Oecetis (L)
TR150 Leptoceridae Oecetis sp (Adult Male)
TR152 Leptoceridae Oecetis sp (Adult Male)
TR155 Leptoceridae Oecetis sp (Adult Female)
TR382 Leptoceridae Oecetis pechana (Adult Male)
TR263 Leptoceridae Triplectides australis(L)
TR264 Leptoceridae Triplectides australis(L)
TR262 Leptoceridae Triplectides australis(L)
TR0025 Leptoceridae Triplectides (L)
TR0088 Leptoceridae Triplectidae ciuskus (L)
TR44 Leptoceridae Triplectides australicus (L)
TR265 Leptoceridae Notalina spira(L)
TR266 Leptoceridae Notalina spira(L)
TR366 Leptoceridae Notalina nr.fulva (Adult Male)
TR367 Leptoceridae Notalina nr.fulva (Adult Male)
TR149 Leptoceridae Notalina bifaria (Adult Male)
TR157 Leptoceridae Notalina bifaria (Adult Male)
TR11 Leptoceridae Notalina bifaria (L)
TR362 Leptoceridae Notalina bifaria (Adult Male)
TR363 Leptoceridae Notalina bifaria (Adult Male)
TR360 Leptoceridae Notalina bifaria (Adult Male)
TR2 Fonly Leptoceridae Notalina bifaria (L)
TR361 Leptoceridae Notalina bifaria (Adult Male)
TR0012 Leptoceridae Notalina bifaria (L)
TR358 Leptoceridae Notalina bifaria (Adult Male)
TR359 Leptoceridae Notalina bifaria (Adult Male)
TR365 Leptoceridae Notalina bifaria (Adult Male)
TR199 Calamoceratidae Anisocentropus (L)
TR324 Calamoceratidae Anisocentropus valgus (Adult Male)
MDFRC-76 Anisocentropus sp. (L)
TR326 Calamoceratidae Anisocentropus valgus (Adult Male)
TR122 Calamoceratidae Anisocentropus sp (L)
TR170 Calamoceratidae Anisocentropus (Adult Female)
TR123 Calamoceratidae Anisocentropus sp (L)
MDFRC-77 Anisocentropus sp. (L)
TR274 Calamoceratidae Anisocentropus sp. (L)
TR0020 Calamoceratidae Anisocentropus (L)
MDFRC-75 Atriplectides dubius (L)
TR10 Atriplectidae Atriplectides dubius (L)
TR0084 Philorheithridae Austreithrus (L)
TR146 Philorheithridae sp. (Adult Male)
TR0082 Philorheithridae Austreithrus (L)
TR0083 Philorheithridae Austreithrus (L)
TR369 Trichoptera unknown (L)
TR212 Philorheithridae Austrheithrus (L)
TR0052 Philorheithridae Aphilorheithrus stephani (Adult Male)
TR242 Philorheithridae Apholorheithrus (L)
TR0081 Philorheithridae Kosrheithrus tillyardi (L)
TR0050 Philorheithridae Kosrheithrus nr.tillyardi (Adult Male)
TR330 Philorheithridae (Adult Female)
TR171 Philorheithridae Kosrheithrus nr.tillyardi (Adult Male)
TR0049 Philorheithridae Kosrheithrus nr.tillyardi (Adult Male)
TR207 Philorheithridae Kosrheithrus tillyardi(L)
Leptoceridae
Calamoceratidae
Atriplectididae
Philorheithridae


TR175 Hydropsychidae Asmicridea nr.edwardsii (Adult Male)
TR215 Hydropsychidae Asmicridea spAV1(L)
TR163 Hydropsychidae Asmicridea nr.edwardsii (Adult Male)
TR162 Hydropsychidae Asmicridea nr.edwardsii (Adult Male)
TR0098 Hydropsychidae Asmicridea AV1 (L)
TR174 Hydropsychidae Asmicridea nr.edwardsii (Adult Male)
TR0096 Hydropsychidae Asmicridea AV1 (L)
TR0097 Hydropsychidae Asmicridea AV1 (L)
TR0001 Hydropsychidae Asmicridea AV1 (L)
TR164 Hydropsychidae Asmicridea nr.edwardsii (Adult Male)
TR0059 Hydropsychidae Asmicridea (Adult Male)
TR0060 Hydropsychidae Asmicridea (Adult Male)
TR233 Hydropsychidae Asmicridea spAV1(L)
TR232 Hydropsychidae Asmicridea spAV1(L)
TR234 Hydropsychidae Asmicridea spAV1(L)
TR117 Hydropsychidae Smicrophylax spAV1 (L)
TR224 Hydropsychidae Smicrophylax spAV2(L)
TR0095 Hydropsychidae Cheumatopsyche spAV4 (L)
TR251 Hydropsychidae Cheumatopsyche AV6 (L)
Hydropsychidae
TR0100 Philopotamidae Chimarra australica (L)
TR0105 Philopotamidae Chimarra australica (Adult Male)
TR220 Philopotamidae Chimarra AV3(L)
TR0099 Philopotamidae Chimarra australica (L)
TR0102 Philopotamidae Chimarra australica (Adult Male)
TR258 Philopotamidae Chimarra monticola (L)
Philopotamidae
TR315 Polycentropodidae (Adult Female)
TR316 Polycentropodidae (Adult Female)
TR314 Polycentropodidae (Adult Female)
TR0069 Polycentropodidae Nyctophylax (Adult Male)
TR0057 Polycentropodidae Plectrocnemia (Adult Male)
TR0070 Polycentropodidae Neuroclipsis (Adult Male)
Polycentropodidae
TR114 Ecnomidae Ecnomus continentalis (L)
TR321 Ecnomidae Ecnomus (Adult Female)
TR148 Ecnomidae Ecnomus russellius (Adult Male)
TR269 Ecnomidae Ecnomus russellius(L)
TR272 Ecnomidae Ecnomus russellius(L)
TR270 Ecnomidae Ecnomus russellius(L)
TR273 Ecnomidae Ecnomus russellius(L)
TR18 Ecnomidae Ecnomus pansus (L)
TR19 Ecnomidae Ecnomus pansus (L)
TR0065 Ecnomidae Ecnomus tillyardi (Adult Male)
TR112 Ecnomidae Ecnomus russellius (L)
TR0063 Ecnomidae Ecnomus tillyardi (Adult Male)
TR0062 Ecnomidae Ecnomus tillyardi (Adult Male)
TR113 Ecnomidae Ecnomus russellius (L)
TR147 Ecnomidae Ecnomus nr.tillyardi (Adult Male)
TR319 Ecnomidae Ecnomus (Adult Male)
TR322 Ecnomidae Ecnomus (Adult Female)
TR0106 Ecnomidae Ecnomus deanii (Adult Male)
TR323 Ecnomidae Ecnomus (Adult Female)
TR317 Ecnomidae Ecnomus (Adult Male)
TR318 Ecnomidae Ecnomus (Adult Male)
TR320 Ecnomidae Ecnomus (Adult Female)
Ecnomidae


TR354 Hydrobiosidae Ulmerochorema (Adult Female)
TR355 Hydrobiosidae Ulmerochorema (Adult Female)
TR335 Hydrobiosidae Ulmerochorema rubiconum (Adult Male)
TR351 Hydrobiosidae Ulmerochorema (Adult Female)
TR338 Hydrobiosidae Ulmerochorema (Adult Male)
TR353 Hydrobiosidae Ulmerochorema (Adult Female)
TR172 Hydrobiosidae Ulmerochorema (Adult Female)
TR343 Hydrobiosidae Ulmerochorema (Adult Female)
TR0005 Hydropsychidae Ulmerochorema rubiconum grp (L)
TR205 Hydrobiosidae Ulmerochorema rubiconum grp(L)
TR334 Hydrobiosidae Ulmerochorema rubiconum (Adult Male)
TR348 Hydrobiosidae Ulmerochorema (Adult Female)
TR352 Hydrobiosidae Ulmerochorema (Adult Female)
TR173 Hydrobiosidae Ulmerochorema (Adult Female)
TR0067 Hydrobiosidae Ulmerochorema membrum (Adult Male)
TR336 Hydrobiosidae Ulmerochorema membrum (Adult Male)
TR349 Hydrobiosidae Ulmerochorema (Adult Female)
TR161 Hydrobiosidae Ulmerochorema nr.tasmanica (Adult Male)
TR218 Hydrobiosidae Ulmerochorema lentum(L)
TR347 Hydrobiosidae Ulmerochorema (Adult Female)
TR341 Hydrobiosidae Ulmerochorema (Adult Male)
TR340 Hydrobiosidae Ulmerochorema (Adult Male)
TR350 Hydrobiosidae Ulmerochorema (Adult Female)
TR0007 Hydrobiosidae Ethochorema turbidum (L)
TR0087 Hydrobiosidae Ethochorema turbidum (L)
TR254 Hydrobiosidae Ethochorema turbidum (L)
TR391 Hydrobiosidae Ethochorema brunneum(L)
TR255 Hydrobiosidae Ethochorema brunneum (L)
TR0086 Hydrobiosidae Ethochorema nr.brunneum (pupa)
TR256 Hydrobiosidae Ethochorema brunneum (L)
TR244 Hydrobiosidae Apsilochorema obliquum(L)
TR204 Hydrobiosidae Apsilochorema obliquum(L)
TR243 Hydrobiosidae Apsilochorema obliquum(L)
Hydrobiosidae
TR0078 Glossosomatidae Agapetus spAV1 (L)
TR302 Glossosomatidae Agapetus (Adult Female)
TR0058 Glossosomatidae Agapetus (Adult Male)
TR0104 Glossosomatidae Agapetus (Adult Male)
TR159 Glossosomatidae Agapetus sp. (Adult Female)
TR327 Glossosomatidae Agapetus (Adult Male)
TR303 Glossosomatidae Agapetus (Adult Female)
TR0103 Glossosomatidae Agapetus (Adult Male)
TR158 Glossosomatidae Agapetus sp. (Adult Female)
TR203 Glossosomatidae Agapetus spAV1(L)
TR0079 Glossosomatidae Agapetus spAV1 (L)
TR192 Glossosomatidae Agapetus spAV1(L)
TR138 Glossosomatidae Agapetus sp (Adult Male)
TR0080 Glossosomatidae Agapetus spAV1 (L)
TR160 Glossosomatidae Agapetus sp. (Adult Male)
TR235 Glossosomatidae Agapetus spAV1(L)
TR237 Glossosomatidae Agapetus spAV1(L)
TR236 Glossosomatidae Agapetus spAV1(L)
TR193 Glossosomatidae Agapetus spAV1(L)
TR245 Glossosomatidae Agapetus spAV1(L)
TR246 Glossosomatidae Agapetus spAV1(L)
Glossosomatidae
TR298 Hydroptilidae Hydroptila scamandra (Adult Male)
TR313 Hydroptilidae Hydroptilla nr.scamandra (Adult Male)
TR0091 Hydroptilidae Hellyethira scamandra (L)
TR0093 Hydroptilidae Hellyethira scamandra (L)
TR0092 Hydroptilidae Hellyethira scamandra (L)
TR296 Hydroptilidae Hydroptila nr.calcara (Adult Male)
TR297 Hydroptilidae Hydroptila nr.calcara (Adult Male)
TR134 Hydroptilidae Orthotrichia sp (L)
TR210 Hydroptilidae Orthotrichia MDFRCsp2 (L)
TR135 Hydroptilidae Orthotrichia sp (L)
TR0089 Hydroptilidae Orthotrichia nr.Tortuosa (L)
TR0029 Hydroptilidae Orthotrichia (L)
TR0031 Hydroptilidae Orthotrichia (L)
TR276 Hydroptilidae Oxyethira columba (Adult Male)
TR281 Hydroptilidae Oxyethira columba (Adult Male)
TR300 Hydroptilidae Hellyethira nr.basilobata (Adult Male)
TR177 Hydroptilidae Hellyethira simplex (L)
TR178 Hydroptilidae Hellyethira simplex (L)
TR299 Hydroptilidae sp. (Adult Male)
TR286 Hydroptilidae Hellyethira simplex (Adult Male)
TR289 Hydroptilidae Hellyethira simplex (Adult Male)
TR292 Hydroptilidae Hellyethira simplex (Adult Male)
TR311 Hydroptilidae Hellythira simplex (Adult Male)
TR294 Hydroptilidae Hellyethira simplex (Adult Male)
TR293 Hydroptilidae Hellyethira simplex (Adult Male)
TR295 Hydroptilidae Hellyethira simplex (Adult Male)
TR0108 Hydroptilidae Hydroptilla simplex (Adult Male)
TR291 Hydroptilidae Hellyethira simplex (Adult Male)
TR287 Hydroptilidae Hellyethira simplex (Adult Male)
TR288 Hydroptilidae Hellyethira simplex (Adult Male)
TR0090 Hydroptilidae Hellyethira simplex (L)
TR290 Hydroptilidae Hellyethira simplex (Adult Male)
Hydroptilidae
0.02

Appendix B COI gene tree of Elmidae specimens added to the DNA barcoding library



Appendix C Table recording the successfulness of detecting DNA in treatments of 23 days

Bottle	Sample treatment	Replicate	Day 0	Day 1	Day 2	Day 4	Day 8	Day 16	Day 23
1	Nutrients	1	1	0	1	1	0	0	1
1	Nutrients	2	1	1	1	1	1	0	1
1	Nutrients	3	1	1	1	1	0	1	0
1	Nutrients	4	1	1	1	0	1	1	1
1	Nutrients	5	1	1	1	1	0	0	0
2	Nutrients	1	1	1	1	1	0	0	1
2	Nutrients	2	1	1	1	1	0	1	1
2	Nutrients	3	1	1	1	0	0	0	1
2	Nutrients	4	1	1	1	0	0	1	0
2	Nutrients	5	1	1	1	0	0	1	0
3	Nutrients	1	1	1	1	1	0	0	1
3	Nutrients	2	1	1	1	1	0	1	0
3	Nutrients	3	1	1	0	0	0	1	0
3	Nutrients	4	1	1	1	1	1	1	1
3	Nutrients	5	1	1	0	0	0	0	1
1	Control	1	1	1	1	1	1	1	1
1	Control	2	0	1	1	1	1	1	1
1	Control	3	0	1	1	0	0	0	0
1	Control	4	1	1	1	0	0	0	1
1	Control	5	1	1	0	1	1	0	1
2	Control	1	1	1	1	0	0	1	0
2	Control	2	1	1	1	0	0	0	1
2	Control	3	1	1	1	0	0	0	1
2	Control	4	1	1	1	0	0	0	1
2	Control	5	1	1	1	0	0	0	1
3	Control	1	1	1	1	0	0	1	1
3	Control	2	1	1	1	1	1	0	1
3	Control	3	1	1	1	0	0	1	1
3	Control	4	1	1	1	0	0	0	1
3	Control	5	1	1	0	0	0	0	1
1	Negative	1	0	0	0	0	0	0	0
1	Negative	2	0	0	0	0	0	0	0
1	Negative	3	0	0	0	0	0	0	0
1	Negative	4	0	0	0	0	0	0	0
1	Negative	5	0	0	0	0	0	0	0
2	Negative	1	0	0	0	0	0	0	0
2	Negative	2	0	0	0	0	0	0	0
2	Negative	3	0	0	0	0	0	0	0
2	Negative	4	0	0	0	0	0	0	0
2	Negative	5	0	0	0	0	0	0	0
3	Negative	1	0	0	0	0	0	0	0
3	Negative	2	0	0	0	0	0	0	0
3	Negative	3	0	0	0	0	0	0	0
3	Negative	4	0	0	0	0	0	0	0
3	Negative	5	0	0	0	0	0	0	0