The Canterbury Museum mayfly collection data resource

Johnathon D. Ridden,* Terry R. Hitchings, Tim R. Hitchings

Canterbury Museum, 11 Rolleston Avenue, Christchurch Central City, Christchurch 8013, New Zealand

ABSTRACT

A nationally significant collection of mayflies that has been amassed and curated at Canterbury Museum, Christchurch, New Zealand is described. A project to formally catalogue the backlog of this collection was completed in 2018. This collection has been primarily worked on, added to, and curated by Terry Hitchings since the early 1990s, with his son Tim Hitchings assisting this work since the late 2000s. This paper outlines this process involved in cataloguing the collection and preparing the data for publication to online biorepositories. The dataset was published to the Atlas of Living Australia (ALA) and the Global Biodiversity Information Facility (GBIF) in late 2021. This dataset contains just under 49,000 published specimen records with high quality field collection information. It represents nearly all currently described mayfly species in New Zealand. Areas of collecting focus include most of the South Island of New Zealand, with collecting gaps in South Westland and Marlborough. There are large collecting gaps throughout the North Island of New Zealand. An overview of the trends shown in the dataset is provided. Future work is identified and recommended to enhance and improve this dataset to highlight and promote freshwater ecosystems in New Zealand.

INTRODUCTION

Mayflies (Ephemeroptera) are a group of aquatic insects which are important for monitoring the health of freshwater ecosystems and environments globally. Water quality can be assessed by documenting mayfly abundance and diversity along with other freshwater insects such as caddisfly (Trichoptera) and stonefly (Plecoptera), commonly referred to as Ephemeroptera, Plecoptera and Trichoptera (EPT) (Lenat, 1988). The state of New Zealand freshwater is sensitive due to various impacts, primarily climate change and human mediated land use change (Ministry for the Environment and Stats NZ, 2020). Information on where species are found and how abundant they are is vital for tracking the state of freshwater ecosystems. In New Zealand, the mayfly fauna currently consists of 58 described mayfly species (Pohe, 2018, Hitchings and Hitchings, 2018, Hitchings and Hitchings, 2021), which is a small proportion of the currently described global fauna of just over 3,800 species (Barber-James et al., 2013). There are more species, predominantly in *Deleatidium*, *Nesameletus* and *Zephlebia,* that are yet to be discovered and formally described in New Zealand (Pohe, 2018). Providing information on species distribution and abundance of mayflies is beneficial for monitoring the status of New Zealand's freshwater environment over time.

Large datasets, including from natural history collections, are being made available on online biorepositories, such as the Global Biodiversity Information Facility (GBIF) https://www.gbif.org/, or more locally to New Zealand and Australasia *via* the Atlas of Living Australia (ALA) https://www.ala.org.au/. Publishing data to these platforms will continue to make biodiversity data more accessible and usable for research and decision-making purposes. To make this information consistent and usable, data must conform to the Darwin Core standard for biodiversity data (Wieczorek *et al*., 2012).

The Canterbury Museum mayfly collection, which is nationally significant, was published to the ALA and GBIF in late 2021, providing valuable distribution data on New Zealand mayflies (Atlas of Living Australia, 2021; Canterbury Museum, 2021). These data can inform efforts to document the conservation status of New Zealand freshwater insects, and fill knowledge gaps about

Corresponding author: JRidden@canterburymuseum.com

Key words: Occurence; New Zealand; Ephemeroptera; Atlas of Living Australia; Global Biodiversity Information Facility; Data Paper.

Citation: Ridden JD, Hitchings TR, Hitchings TR. The Canterbury Museum mayfly collection data resource. *J. Limnol. 2023;82:2097.*

Edited by: Diego Fontaneto, *National Research Council, Water Research Institute (CNR-IRSA), Verbania Pallanza, Italy.*

Received: 20 October 2022.
Accepted: 27 January 2023.

J. Limnol., 2023; 82(s1):2097
DOI: 10.4081/jlimnol.2023.2097

the mayfly fauna (Grainger *et al*., 2018; Drinan *et al*., 2020). Another important purpose of these data is providing baseline information on New Zealand mayflies, so trends in abundance and diversity can be tracked over time. It is important to contextualise the history of this collection and the metadata associated with it, to improve its usefulness.

This paper will outline the history and development of the collection, which is primarily due to work by Terry Hitchings since the early 1990s. The data publication process is described, and a broad overview of the data is given. The aim is to provide context around this collection to improve and utilise its usage by a range of stakeholders, such as researchers and policy makers.

METHODS

General collection description

This dataset provides information on freshwater mayfly specimens held at Canterbury Museum. Data from this collection have accurately recorded locality and georeference data. All specimens have well curated collection labels recording vital field collection data (Fig. 1). Nearly all specimens are identified to species, with a small proportion only identified to genus. There are currently 48,992 specimens catalogued and published from the collection, representing 56 unique species (see Taxonomic Coverage), compromising over 10,000 collecting event lots which continue to be added to this dataset, and published online in future.

History and development of the collection

An overview of the Canterbury Museum mayfly collection was given in Hitchings (2001). Hitchings (2019) gives an account of the history and development of the Canterbury Museum mayfly collection and its digital status as recorded on a Microsoft Access database. Since the early 2000s, Terry Hitchings, currently a Research Fellow at Canterbury Museum, has significantly enriched the mayfly collection, by collecting and identifying specimens from throughout New Zealand. Hitchings (2001) provides important context for the history of the Canterbury Museum mayfly collection for interpreting vial label data, with information on collectors, localities, grid references and other abbreviated codes. The collection predominately consists of specimens preserved in 80% ethanol and a small number of slide mounted dissections from specimens. Some additional information is recorded in notebooks, with several folders of anatomical drawings for species descriptions.

John Ward, who was a Research Fellow at Canterbury Museum studying caddisflies, mentored Hitchings, and helped him develop his skills as an amateur mayfly taxonomist (Patrick, 2016). Ward was a prolific field collector

Fig. 1. Style of vial label found in the collection. The information shown is as follows: Crosby Code recorded in the top left by the two-letter system, locality information, grid references or coordinates, elevation, collection date, collector, species identification, life stage, date identified and identifier. The tripartite accession number is placed in the vial also for each individual specimen.

and during his many field trips collecting caddisfly, often alongside his colleague and friend Brian Patrick, collected many mayflies, which were given to Hitchings. Throughout his tenure at Canterbury Museum, Hitchings has described several new species, some along with his son Tim Hitchings, primarily in the genus *Deleatidium* (Hitchings, 2006, 2009, 2010; Hitchings and Hitchings, 2016, 2018, 2021) but also species in other genera with colleague Arnold Staniczek, including *Nesameletus* (Hitchings and Staniczek, 2003) and *Rallidens* (Staniczek and Hitchings, 2013). Richard Hitchings (Tim Hitching's son and Terry Hitchings's grandson) also assists in the curation and identification of mayfly material alongside his father and grandfather, under their tutelage. Mayfly material had not been formerly catalogued and documented in the Canterbury Museum database, the Vernon Collection Management System (CMS) before 2014. A project from 2014-2018 occurred to transfer the data held in the Microsoft Access database to Vernon CMS and catalogue the mayfly collection. The data held on the Microsoft Access database are recorded at the collection event/specimen tube level, not at an individual specimen level.

Canterbury Museum has a collection policy, which means every individual object and specimen in the museum is assigned an accession number. For vials with multiple specimens, an accession number is assigned to each individual specimen in the vial and the range of accession numbers is recorded in the vial. Data were captured onto a Microsoft Excel spreadsheet template formatted to Canterbury Museum internal data standards, by transcribing information from the Microsoft Access database and checking this against vial labels during data entry. This internal standard does not directly map to Darwin Core data standards (Wieczorek, 2012). Some specimen data were not captured in the cataloguing process, including dwc:dateIdentified, dwc:verbatimLocality and only some records have data for dwc:verbatimCoordinates recorded. All specimens were assigned coordinates from map grid references or GPS readings and follow the protocol outlined in Hitchings *et al*. (2015). Each given collector is assigned an associated dwc:coordinateUncertaintyInMeters, based on the information they provided with specimens. All dwc:coordinateUncertaintyInMeters data were verified by (JDR) by comparing locality information against given coordinates.

Specimen identifications were primarily made by Terry Hitchings, up until 2009, when his son Tim started at Canterbury Museum as a Research Associate, and subsequently became a Research Fellow in 2016. Both Terry and Tim Hitchings are recorded as identifiers for material collected since 2009, as the specific specimen identifier information was not included in the cataloguing project. Specimens catalogued after the initial cataloguing project was completed in 2018 have data recorded in some of the Darwin Core fields mentioned above e.g., dwc:dateIdentified, dwc:verbatimCoordinates. All mayfly data were downloaded from the Canterbury Museum Vernon CMS in 2021 and prepared for data publication.

Project description

A project to catalogue and document the backlog of mayflies in Canterbury Museum commenced in 2014, finishing in 2018. This involved assigning unique catalogue numbers to each specimen and recording collection data onto spreadsheets. The cataloguing process involved the following steps:

i. Information from the Microsoft Access database was copied and entered in a new Microsoft Excel spreadsheet, with each specimen getting a row of data recorded with its unique accession number.
ii. Around 1000 specimen records were added to a given spreadsheet by a technician, then the data were checked by a curator for errors and passed to the Registration team. One more check of the data was completed by a member of the Registration team and the spreadsheet was imported into the Canterbury Museum collection database, Vernon CMS.
iii. A Registration team member verified the accession numbers matched the associated specimens on the database record.
iv. This process was repeated until all backlog mayfly specimens in the collection were processed.

Mayfly records were exported from Canterbury Museum Vernon CMS as a comma separated value (.csv) file. Each specimen has georeferenced coordinates, with the dwc:coordinateUncertaintyInMeters determined primarily per Hitchings *et al*. (2015) or for specimens catalogued after 2018, by the Georeferencing Quick Reference Guide (Zermoglio, 2020). Georeference data was checked and verified by a curator (JDR) against the locality information associated with the coordinates. Errors of map grid reference or coordinates were identified by this process, such as incorrect data entry or a data lapsus during transcription. These errors were corrected and checked to confirm the given locality information corresponded to the coordinates. The data were then cleaned using the OpenRefine application version 3.2 (OpenRefine, 2022) and mapped to the Darwin Core data standard for biodiversity (Wieczorek *et al*., 2012).

The cleaned data as a (.csv) file were then run through the GBIF Data Validator tool. Issues identified by this process were addressed. The data were then uploaded in the ALA Sandbox tool. After the data were parsed by the ALA Sandbox, issues identified were rectified. The clean data were uploaded into the ALA Sandbox again and once they passed all the data quality checks; a Darwin Core Archive was created using the GBIF Test Integrated Publishing Toolkit (IPT). This archive was then submitted to

ALA for publishing *via* the Sandbox tool. Data on ALA are shared with GBIF for publication online and are now available to download from both biorepositories (Atlas of Living Australia, 2021, Canterbury Museum, 2021). This collection will continue to be updated as new specimen records are catalogued and digitised and if specimen identifications are updated.

RESULTS AND DISCUSSION

Data resource

Data package title: Canterbury Museum Mayfly Collection
Resource link: https://doi.org/10.15468/5ksxu8
Alternative identifiers: https://doi.ala.org.au/doi/33adcafe-168c-4c8e-a126-468c599b71da, https://collections.ala.org.au/public/show/dr17655, CMNZmayfly
Number of data sets: 1
Data set name: Canterbury Museum Mayfly Collection
Data format: Darwin Core Archive
Data licence: CC BY 4.0
Description: This is a collection amassed primarily by Terry Hitchings at Canterbury Museum, New Zealand. Hitchings began collecting, curating, identifying, and describing new species from the collection in the early 1990s, incorporating the existing Canterbury Museum mayfly collections into this work. From the late 2000s, Terry Hitchings's son, Tim Hitchings, began working with his father on the collection. Their work on the group is ongoing. A project to fully catalogue this collection was completed in 2018. Further additions are made to the collection in lots of material, are identified and curated. Explanations of the Darwin Core terms used for publishing this dataset are given (Tab. 1).

Tab. 1. Summary of interpretation of Darwin Core terms used in published dataset.

Column label	Column description
occurrenceID	The globally unique identifier for the record, which is a combination of insititionCode and the Vernon CMS unique record ID number.
type	The nature or genre of the resource, i.e. "PhysicalObject".
modified	The most recent date on which the resource was changed.
language	The language of the resource, i.e. English.
licence	The legal document giving official permission to do something with the resource. i.e. "https://creativecommons.org/licenses/by/4.0/".
institutionID	An identifier from the Global Registry for Biodiversity Repositories (https://grbio.org/) for the institution (CMNZ) which has custody of the object(s) or information referred to in the record.
institutionCode	The name of the institution having custody of the object(s) or information referred to in the record, i.e. Canterbury Museum, New Zealand.
collectionCode	The name, acronym, coden, or initialism identifying the collection or dataset from which the record was derived, i.e., the Canterbury Museum Mayfly Collection (CMNZmayfly).
basisOfRecord	The specific nature of the data record, i.e. "PreservedSpecimen".
catalogNumber	An unique tripartite internal CMNZ identifier for the record within the dataset or collection. The first number represents the year of processing e.g. 2020, the second number represents when the lot was processed in that year e.g. 12 is the twelfth object lot to be processed and the last number is the unique identifier for each specimen in the lot starting from 1.
recordedBy	The primary collector or collectors of the specimen(s), e.g. Hitchings, Terry R.
individualCount	The number of individuals represented in the data record.
organismQuantity	A number or enumeration value for the quantity of organisms per record in the collection.
organismQuantityType	The type of quantification system used for the quantity of organisms, i.e. every specimen is an individual.
lifeStage	The age class or life stage of the biological individual(s), i.e. "adult imago", "adult subimago", "nymph" or "exuvia".
establishmentMeans	Statement about whether an organism or organisms have been introduced to a given place and time through the direct or indirect activity of modern humans, e.g. native.
occurenceStatus	A statement about the presence or absence of a Taxon at a Location. This dataset only records "present" individuals.
preparations	A list (concatenated and separated) of preparations and preservation methods for a specimen, e.g. whole animal (ethanol).
eventDate	The date-time or interval during which an Event occurred. For occurrences, this is the date-time when the event was recorded. Dates conform to ISO 8601-1:2019 as best practice.
year	The four-digit integer year in which the Event occurred, according to the Common Era Calendar.
month	The integer month in which the Event occurred.

To be continued on next page

Tab. 1. Continued from previous page.

Column label	Column description
day	The integer day of the month on which the Event occurred.
higherGeography	A list (concatenated and separated) of geographic names less specific than the information captured in the locality term. This is based on broad, well established geographic places, e.g. Oceania | New Zealand | South Island | Canterbury.
continent	The name of the continent on which the collection occurred, i.e. Oceania.
island	The name of the island on or near which the Location occurs, e.g. South Island of New Zealand.
country	The name of the country or major administrative unit in which the Location occurs, i.e. New Zealand.
countryCode	The standard code for the country in which the Location occurs, using the ISO 3166-1-alpha-2 country code as best practice, i.e. NZ.
county	The full, unabbreviated name of the next smaller administrative region than stateProvince (county, shire, department, etc.) in which the Location occurs. For invertebrate collections in New Zealand the Crosby Code system is used here Crosby et al. (1998). These are given as the full region name, not the two letter code, in this dataset.
locality	The specific description of the place. This term may contain information modified from the original to correct perceived errors or standardise the description, e.g. Maruia R, SH7, would be interpreted as, Maruia River, State Highway 7.
verbatimElevation	The original description of the elevation (altitude, usually above sea level) of the Location, in metres.
verbatimCoordinates	The verbatim original spatial coordinates of the Location. Coordinates given as map grid references from either New Zealand Transverse Mercator (NZTM) or New Zealand Map Grid (NZMG), e.g. 1639842 6067836 (NZTM) or 23696 57050 (NZMG).
decimalLatitude	The geographic latitude (in decimal degrees, using the spatial reference system given in geodeticDatum) of the geographic center of a Location. Positive values are north of the Equator, negative values are south of it. Legal values lie between -90 and 90, inclusive.
decimalLongitude	The geographic longitude (in decimal degrees, using the spatial reference system given in geodeticDatum) of the geographic center of a Location. Positive values are east of the Greenwich Meridian, negative values are west of it. Legal values lie between -180 and 180, inclusive.
geodeticDatum	The ellipsoid, geodetic datum, or spatial reference system (SRS) upon which the geographic coordinates given in decimalLatitude and decimalLongitude as based. This is primarily WGS84.
coordinateUncertaintyInMeters	The horizontal distance (in meters) from the given decimalLatitude and decimalLongitude describing the smallest circle containing the whole of the Location.
georeferencedBy	A list (concatenated and separated) of names of people, groups, or organizations who determined the georeference (spatial representation) for the Location.
georeferenceDate	The date on which the Location was georeferenced. Dates conform to ISO 8601-1:2019 as best practice.
georeferenceProtocol	A description or reference to the methods used to determine the spatial footprint, coordinates, and uncertainties. The majority of the collection was georeferenced based on the protocol of Hitchings et al. (2015), or GeoLOCATE (Rios and Bart, 2010) following the Georeferencing Quick Reference Guide (Zermoglio, et al. 2020, https://doi.org/10.35035/e09p-h128).
georeferenceSources	A list (concatenated and separated) of maps, gazetteers, or other resources used to georeference the Location, described specifically enough to allow anyone in the future to use the same resources. The primary source used was www.freshmap.co.nz, and subsequently www.topomap.co.nz from 2010. Some records were georeferenced using GeoLOCATE (Rios and Bart, 2010) web application.
georeferenceVerificationStatus	A categorical description of the extent to which the georeference has been verified to represent the best possible spatial description for the Location of the Occurrence. The author (JDR) checked and verified the majority of unique collecting localities against the decimalLatitude and decimalLongitude.
georeferenceRemarks	Notes or comments about the spatial description determination, explaining assumptions made in addition or opposition to the those formalized in the method referred to in georeferenceProtocol. Remarks given based on protocol in Hitchings et al. (2015) or GeoLOCATE (Rios and Bart, 2010) following the Georeferencing Quick Reference Guide (Zermoglio et al. 2020, https://doi.org/10.35035/e09p-h128).
identifiedBy	A list (concatenated and separated) of names of people, groups, or organizations who assigned the Taxon to the subject.
dateIdentified	The date on which the subject was determined as representing the Taxon. Dates conform to ISO 8601-1:2019 as best practice.
typeStatus	A list of nomenclatural types with type status applied to the subject.
scientificName	The full scientific name, with authorship and date information if known. When forming part of an Identification, this should be the name in lowest level taxonomic rank that can be determined. This term should not contain identification qualifications, which should instead be supplied in the IdentificationQualifier term.
higherClassification	A list (concatenated and separated) of taxa names terminating at the rank immediately superior to the taxon referenced in the taxon record, e.g., Animalia | Arthropoda | Insecta | Ephemeroptera | Leptophlebiidae | Deleatidium.

To be continued on next page

Geographic coverage

The geographic coverage of the collection covers all three main islands of New Zealand (North Island, South Island and Stewart Island) as well a record from the Auckland Islands (Fig. 2). The altitudinal gradient of collecting is from alpine waterways all the way to sea level. Most specimens were collected from the South Island, with a focus on Banks Peninsula and the Canterbury foothills. The Lewis Pass and Arthur Pass in the South Island be-

Tab. 1. Continued from previous page.

Column label	Column description
kingdom	The full scientific name of the kingdom in which the taxon is classified, i.e., “Animalia”.
phylum	The full scientific name of the phylum or division in which the taxon is classified, e.g. “Arthopoda”.
class	The full scientific name of the class in which the taxon is classified, e.g., “Insecta”.
order	The full scientific name of the order in which the taxon is classified, e.g., “Ephemeroptera”.
family	The full scientific name of the family in which the taxon is classified, e.g., “Leptophlebiidae”.
genus	The full scientific name of the genus in which the taxon is classified, e.g., “Deleatidium”.
specificEpithet	The name of the first or species epithet of the scientificName, e.g., “myzobranchia”.
taxonRank	The taxonomic rank of the most specific name in the scientificName, i.e., “Family”, “Genus” or “Species”.
nomenclaturalCode	The code of nomenclature that governs the scientificName, i.e., “ICZN”, the International Code of Zoological Nomenclature.

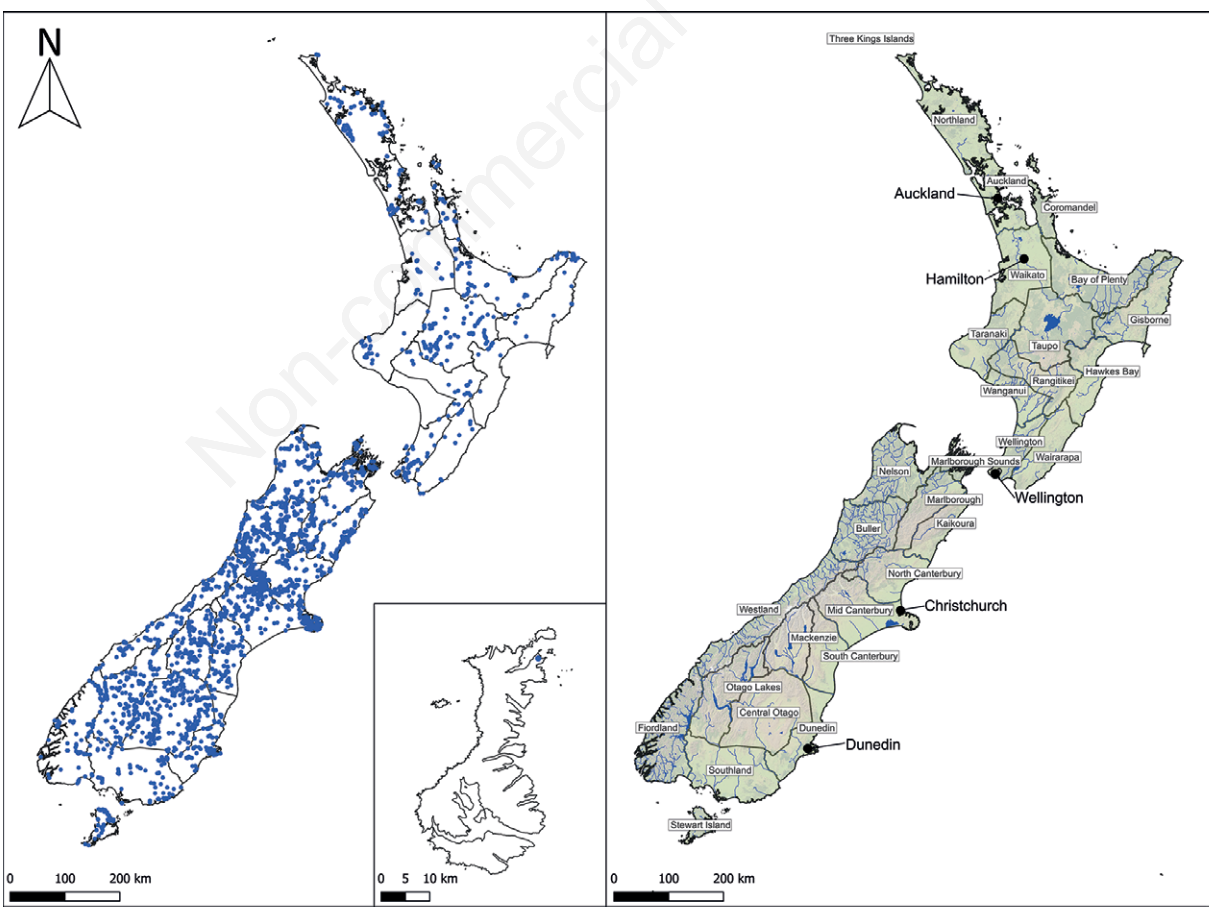

Fig. 2. Distribution map (left) of mayfly specimens in the Canterbury Museum collection with Crosby Code (Crosby *et al.*, 1998) collecting regions shown on the map. Inset map shows specimen record on the Auckland Islands. New Zealand geographic context map (left) shows the name of Crosby Code regions, significant lakes/rivers and the five main city centres.

tween the West Coast and the Canterbury Plains are well collected areas. In the southern area of New Zealand, areas that are not as well represented include remote parts of Fiordland, Southland, and the south-eastern coast of Stewart Island. Parts of Marlborough (the Richmond Ranges) and Nelson have collection gaps. There are also several large collection gaps in the North Island, including Whanganui, Wairarapa, Hawkes Bay, Gisborne, Bay of Plenty and Waikato. Parts of Northland (Kaipara and Whangarei), inland Taranaki and areas in the Central Plateau also present gaps in the collection. Collecting efforts by Canterbury Museum staff and associates are focussed primarily on the Canterbury region and generally the South Island. This is the primary reason there are collection gaps listed above occur, as they are located farther from Canterbury. Also, most collection events were completed by researchers based in Christchurch, who collected more frequently near Christchurch.

Geographic extent as bounding coordinates: -65.513 and -28.613 Latitude; 157.852 and -153.984 Longitude.

Taxonomic coverage

This collection represents all eight families of Ephemeroptera recorded in New Zealand (Pohe, 2018). All species and genera represented in the collection and published in the dataset are presented (Tab. 2). It should

Tab. 2. Taxonomic representation and number of specimens from the Canterbury Museum Mayfly Collection.

Rank	Scientific name	Taxonomic authority	Number of specimens
species	*Acanthophlebia cruentata*	(Hudson, 1904)	244
species	*Ameletopsis perscitus*	(Eaton, 1899)	335
species	*Arachnocolus phillipsi*	Towns and Peters, 1979	20
species	*Atalophlebioides cromwelli*	(Phillips, 1930)	403
genus	*Atalophlebioides*	Phillips, 1930	1
species	*Auporiella pohei*	Winterbourn, 2009	1
genus	*Austroclima*	Towns and Peters, 1979	1
species	*Austroclima jollyae*	Towns and Peters, 1979	516
species	*Austroclima sepia*	(Phillips, 1930)	614
species	*Austronella planulata*	(Towns, 1983)	78
species	*Coloburiscus humeralis*	(Walker, 1853)	3876
species	*Cryophlebia aucklandensis*	(Peters, 1971)	1
genus	*Deleatidium*	Eaton, 1899	84
species	*Deleatidium acerbum*	Hitchings and Hitchings, 2016	109
species	*Deleatidium angustum*	Towns and Peters, 1996	678
species	*Deleatidium atricolor*	Hitchings, 2009	1122
species	*Deleatidium autumnale*	Phillips, 1930	4599
species	*Deleatidium branchiola*	Hitchings, 2009	110
species	*Deleatidium cerinum*	Phillips, 1930	881
species	*Deleatidium cornutum*	Towns and Peters, 1996	773
species	*Deleatidium crawfordi*	Hitchings and Hitchings, 2018	12
species	*Deleatidium fumosum*	Phillips, 1930	6430
species	*Deleatidium insolitum*	(Towns & Peters, 1979)	9
species	*Deleatidium kawatiri*	Hitchings and Hitchings, 2016	75
species	*Deleatidium kiwa*	Hitchings, 2010	261
species	*Deleatidium lillii*	Eaton, 1899	4467
species	*Deleatidium magnum*	Towns & Peters, 1996	110
species	*Deleatidium myzobranchia*	Phillips, 1930	6378
species	*Deleatidium patricki*	Hitchings, 2008	474
species	*Deleatidium rapidum*	Hitchings and Hitchings, 2018	18
species	*Deleatidium townsi*	Hitchings, 2009	261
species	*Deleatidium vernale*	Phillips, 1930	4718
species	*Deleatidium wardorum*	Hitchings, 2010	721

To be continued on next page

be noted that early identifications of specimens may need to be revisited as they could be inaccurate due to subsequent descriptions and elaborations of new species, for example, species complexes to be uncovered in the genus *Deleatidium* (Hitchings, 2019). Some specimens could only be described to genus level as they were not preserved well or are early instar nymphs, and taxonomic characters required for species identification were not present. Genera such as *Deleatidium* have several species that are very morphologically similar and cannot confidently be assigned to species level. Molecular analysis would assist in species identification for these specimens. Most specimens are nymphs (33975) as they are easier to collect than subimago (4505) or imagoes (4118), which require night-time light trapping equipment. A portion of specimens did not have life stage recorded (6394). Information on the number of males or females was not recorded when catalogued for imagoes. The records without life stage or sex recorded should be captured and uploaded to the dataset in future.

Over 75% of records are from the family Leptophlebiidae, a globally distributed family (Fig. 3). The swimming mayflies in the family Nesameletidae are the second most represented, followed by the Coloburiscidae, which is represented by one taxonomically valid species *Coloburiscus humeralis* commonly found throughout New Zealand. Fewer number of specimens are represented for the five other families recorded in New Zealand. This collection contains the majority of described mayfly species in New Zealand. Two species are not represented in the collection, *Oniscigaster intermedius* and *Coloburiscus tonnoiri* which are considered *nomina dubia* (Pohe, 2018).

Temporal coverage

This dataset represents over a century of collecting from 1914 - 2019. It is still actively worked on and added

Tab. 2. Continued from previous page.

Rank	Scientific name	Taxonomic authority	Number of specimens
species	*Ichthybotus bicolor*	Tillyard, 1923	103
species	*Ichthybotus hudsoni*	(McLachlan, 1894)	89
species	*Isothraulus abditus*	Towns & Peters, 1979	8
species	*Mauiulus aquilus*	Towns & Peters, 1996	35
species	*Mauiulus luma*	Towns & Peters, 1979	96
species	*Neozephlebia scita*	(Walker, 1853)	1820
genus	*Nesameletus*	Tillyard, 1933	2
species	*Nesameletus austrinus*	Hitchings & Staniczek, 2003	1752
species	*Nesameletus flavitinctus*	(Tillyard, 1923)	837
species	*Nesameletus murihiku*	Hitchings & Staniczek, 2003	143
species	*Nesameletus ornatus*	(Eaton, 1883)	2523
species	*Nesameletus vulcanus*	Hitchings & Staniczek, 2003	171
species	*Oniscigaster distans*	Eaton, 1899	346
species	*Oniscigaster wakefieldi*	McLachlan, 1873	100
species	*Rallidens mcfarlanei*	Penniket, 1966	145
species	*Rallidens platydontis*	Staniczek & Hitchings, 2014	154
species	*Siphlaenigma janae*	Penniket, 1962	139
genus	*Tepakia*	Towns & Peters, 1996	2
species	*Tepakia caligata*	Towns & Peters, 1996	10
genus	*Zephlebia*	Penniket, 1961	4
species	*Zephlebia borealis*	(Phillips, 1930)	101
species	*Zephlebia dentata*	(Eaton, 1871)	510
species	*Zephlebia inconspicua*	Towns, 1983	114
species	*Zephlebia nebulosa*	Towns & Peters, 1996	236
species	*Zephlebia pirongia*	Towns & Peters, 1996	110
species	*Zephlebia spectabilis*	Towns, 1983	303
species	*Zephlebia tuberculata*	Towns & Peters, 1996	142
species	*Zephlebia versicolor*	(Eaton, 1899)	589

to. There are two periods when the collection has been significantly added to, between the 1950s-1970s and from the late 1980s until the present day (Fig. 4).

The period of activity between the 1950s-1970s was driven by John G Penniket and Vida Stout (Hitchings, 2019). Alexander Grant McFarlane, a trichopteran (caddisfly) Research Fellow based at Canterbury Museum contributed to a lesser extent during this period with by

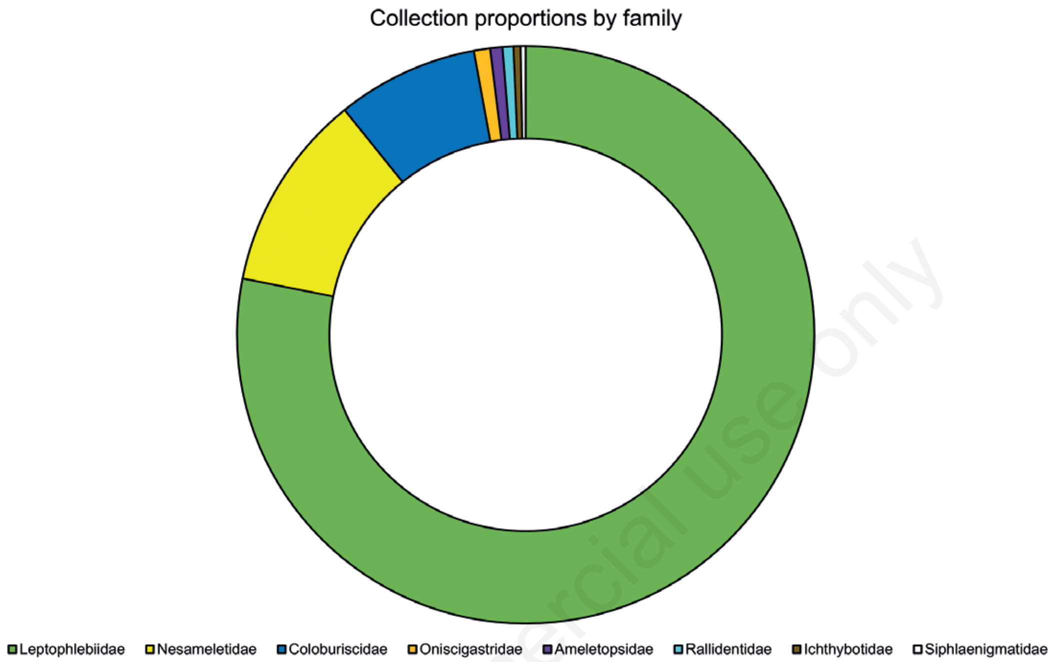

Fig. 3. Proportion of specimens in collection represented in each family.

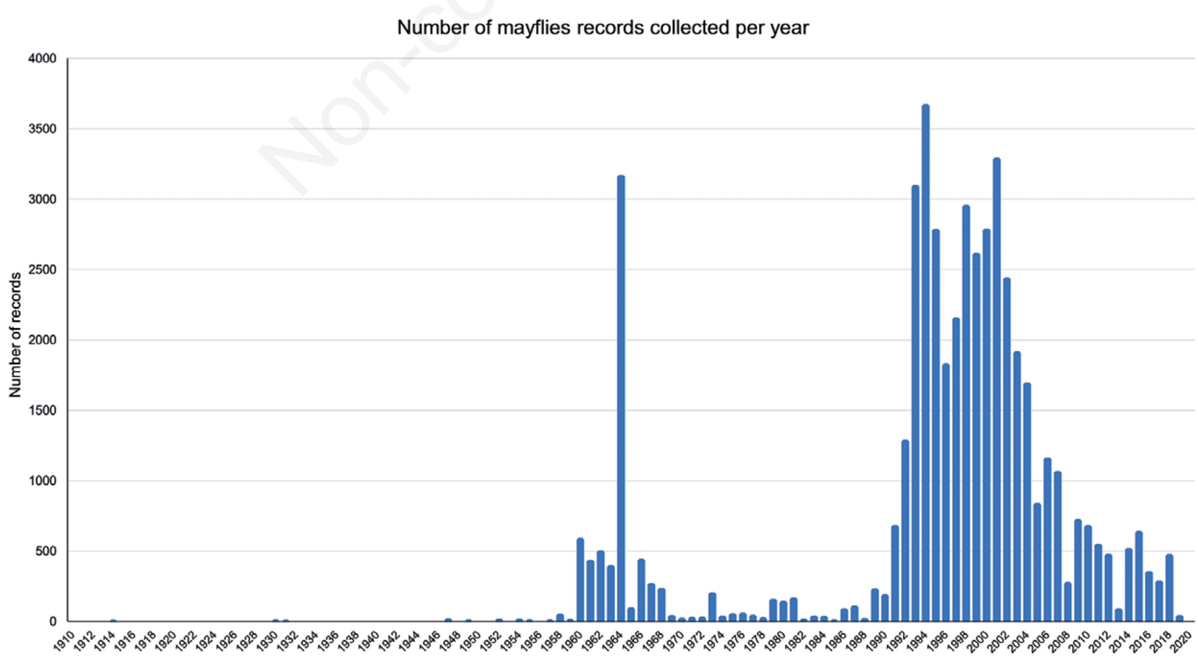

Fig. 4. Number of specimens collected between 1910-2020 that are published to ALA and GBIF from the Canterbury Museum collection.

catch (Savill and Ward, 1993). Lower volumes of specimens were collected from the 1970s-1990s. John Ward joined the Museum as a Research Fellow in 1990, carrying on McFarlane's taxonomic research of Trichoptera (Patrick, 2016). Ward collected extensively from 1990 until the mid-2000s, accumulating large volumes of mayflies, alongside his targeted caddisflies, which were set aside to be worked on by Hitchings.

Terry Hitchings began as a Research Associate at Canterbury Museum in 1991. Hitchings was also actively collecting between the 1990s until about 2010. The large volumes of mayfly material were collected during this period, which was identified and curated by Hitchings. Tim Hitchings developed an interest in his father's mayfly research and joined Canterbury Museum as a Research Associate in 2009. Tim continues to collect material for the collection, to identify, and discover new species to be described. In 2016, Tim Hitchings was made a Research Fellow of Canterbury Museum. Collection of mayflies since 2010 has been predominantly made by Tim Hitchings.

CONCLUSIONS

This work presents an overview of the publication of the Canterbury Museum mayfly collection. It is a nationally significant collection with high quality metadata provided and contextualised in this paper. Making this data available fills gaps on the distribution of New Zealand Ephemeroptera. Many years of identification and curation have improved the taxonomic resolution of the collection and identified areas where future collecting efforts should be focussed.

As new specimens are added to the Canterbury Museum mayfly collection, they should continue to be published to ALA and GBIF. One important aspect of this data is the provision of good quality georeference information for specimens, which promote utilisation for a wide range of purposes, such as species distribution models and informing conservation status of different species. Fine scale specific site information e.g., field notes, images, should be recorded and presented to supplement and enrich published distribution data, as this information can only be recorded when a specimen is collected. Future collection work that could enrich this dataset would involve capturing specific data not initially recorded during the cataloguing project (e.g. dwc:identifiedBy, dwc:verbatimCoordinates and dwc:verbatimLocality). This could be during a specimen and label imaging project. The overseas mayfly collection contains important specimens, including important New Caledonian material, which has not been catalogued. Cataloguing this material and publishing the associated collection data will further increase the value and use of the Canterbury Museum mayfly collection.

ACKNOWLEDGMENTS

Thanks is given to *GBIF* and the *Journal of Limnology* for putting out a call for data paper submissions to highlight species associated with freshwater ecosystems around the world and GBIF specifically, for sponsoring article processing charges. Jan Legind (GBIF) is thanked for providing access to the GBIF Test Integrated Publishing Toolkit to produce a Darwin Core Archive of the data. Thanks is given to the Atlas of Living Australia for efficient support in publishing this dataset, particularly Peggy Newman and Doug Palmer for support in upload data to the ALA sandbox and for assistance with queries to prepare data. Marie Grosjean (GBIF) is thanked for facilitating the uploading of data onto GBIF via ALA and providing access to update information on Canterbury Museum on the GRSciColl registry. Canterbury Museum is thanked for providing resourcing and staff time to catalogue the mayfly collection. The Canterbury Museum Registration team is thanked for database management and data checking and the Canterbury Museum Curatorial team of technicians for cataloguing the specimens. All the collectors of material that have donated specimens to Canterbury Museum are greatly appreciated and are fundamental to this work occurring. We thank both reviewers for comments that greatly improved the manuscript.

REFERENCES

Atlas of Living Australia, 2021. Atlas of Living Australia occurrence. Accessed: 2021-09-14. Available from: https://doi.org/10.26197/ala.33adcafe-168c-4c8e-a126-468c599b71da

Barber-James H, Sartori M, Gattolliat J-L, Webb J, 2013. World checklist of freshwater Ephemeroptera species. In O Bánki, Y Roskov, M Döring, G Ower, L Vandepitte, D Hobern, D Remsen, P Schalk, RE DeWalt, M Keping, J Miller, T Orrell, R Aalbu, R Adlard, E Adriaenssens, C Aedo, E Aescht, N Akkari, MA Alonso-Zarazaga, et al, Catalogue of Life Checklist. Available from: https://doi.org/10.48580/d4tm-3cn

Canterbury Museum, 2021. Canterbury Museum Mayfly Collection. Occurrence dataset Accessed *via* GBIF.org on 2021-12-12. Available from: https://doi.org/10.15468/5ksxu8

Crosby TK, Dugdale JS, Watt JC, 1998. Area codes for recording specimen localities in the New Zealand subregion. New Zeal J Zool 25:175-183.

Drinan T, Grainger NJ, Harding J, Collier K, Smith B, Death R, Makan T, Rolfe J, 2020. Analysis of the conservation status of New Zealand freshwater invertebrates: temporal changes, knowledge gaps, impediments, and management implications. New Zeal J Zool 48:81-96.

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 Zeal Threat Class Ser 28:1-25.

Hitchings TR, Staniczek A, 2003. Nesameletidae (Insecta: Ephemeroptera). Fauna New Zeal 46:1-72.

Hitchings TR, 2006. A new species of *Deleatidium* (*Penniketellum*) and the adult of *D.* (*P.*) *cornutum* Towns and Peters (Ephemeroptera: Leptophlebiidae) from New Zealand. Rec Cant Mus 20:31-43.

Hitchings TR, 2009. Three new species of *Deleatidium* (*Deleatidium*) (Ephemeroptera: Leptophlebiidae) from New Zealand. Rec Cant Mus 23:35-50.

Hitchings TR, 2010. Two new species of *Deleatidium* (*Deleatidium*) (Ephemeroptera: Leptophlebiidae) from the South Island, New Zealand. Rec Cant Mus 24:27-38.

Hitchings TR, 2001. The Canterbury Museum mayfly collection and database (Insecta: Ephemeroptera). Rec Cant Mus 15:11-32.

Hitchings TR, Hitchings TR, Shaw MD, 2015. A revision of the distribution maps and database of New Zealand mayflies (Ephemeroptera) at Canterbury Museum. Rec Cant Mus 29:7-36.

Hitchings TR, Hitchings TR, 2016. Two further species of *Deleatidium* (*Deleatidium*) (Ephemeroptera: Leptophlebiidae) from New Zealand. Rec Cant Mus 30:52-64.

Hitchings TR, Hitchings TR, 2018. Two new species of *Deleatidium* (*Deleatidium*) (Ephemeroptera: Leptophlebiidae) from the central North Island of New Zealand. Rec Cant Mus 32:5-15.

Hitchings TR, 2019. The Canterbury Museum mayfly collection (Insecta: Ephemeroptera) and what it can tell us about changes in species abundance with time. Zoosymposia 16:139-151.

Hitchings TR, Hitchings TR, 2021. *Deleatidium kakahu*, new species of *Deleatidium* (Ephemeroptera: Leptophlebiidae) from New Zealand. Rec Cant Mus 35:247-253.

Lenat DR, 1988. Water quality assessment of streams using a qualitative collection method for benthic macroinvertebrates. J N Am Benthol Soc 7:222-233.

Ministry for the Environment and Stats NZ, 2020. New Zealand's Environmental Reporting Series: Our freshwater 2020. Ministry for the Environment and Stats NZ. Available from: https://environment.govt.nz and www.stats.govt.nz

OpenRefine (Ed), 2022. A free, open source, power tool for working with messy data. v 3.2. Metaweb Technologies, Inc. Available from: https://github.com/OpenRefine

Patrick B, 2016. John Brigham Ward PhD, 10 January 1928–5 April 2016. New Zeal Entomol 39:147-149.

Pohe SR, 2018. An annotated checklist of New Zealand mayflies (Ephemeroptera), 2018. New Zeal Nat Sci 43:1-20.

Rios NE, Bart HL, 2010. GEOLocate (Version 3.22). Computer software. Belle Chasse, LA: Tulane University Museum of Natural History.

Savill A, Ward JB, 1993. ALEXANDER GRANT McFARLANE 1899-1992. New Zeal Entomol 16:102-103.

Staniczek AH, Hitchings TR, 2013. A new species of Rallidens (Ephemeroptera: Rallidentidae) from New Zealand. Rec Cant Mus 27:1-9.

Wieczorek JR, Bloom D, Guralnick R, Blum S, Döring M, Giovanni R, Robertson T, Vieglais D, 2012. Darwin Core: An evolving community-developed biodiversity data standard. PLoS One 7:1-8.

Zermoglio PF, Chapman AD, Wieczorek JR, Luna MC, Bloom DA, 2020. Georeferencing Quick Reference Guide. Copenhagen: GBIF Secretariat. Available from: https://docs.gbif.org/georeferencing-quick-reference-guide/1.0/en/