Freshwater Biological Traits Database

Global Change Research Program
National Center for Environmental Assessment
Office of Research and Development
U.S. Environmental Protection Agency
Washington, DC 20460

DISCLAIMER

This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

ABSTRACT

The Freshwater Biological Traits Database (http://www.epa.gov/ncea/global/traits) currently contains traits data for 3,857 North American macroinvertebrate taxa and includes habitat, life history, mobility, morphology, and ecological trait data. Species traits are the characteristics that explain an organisms' relationship to the surrounding environment, including how it grows, feeds and moves. The traits data were compiled for a project on climate change effects on river and stream ecosystems that was conducted by the Global Change Research Program in the National Center for Environmental Assessment in the U.S. EPA Office of Research and Development. The traits data were gathered from multiple sources. Data gathering efforts focused on data that were published or well-documented, available, appropriate for the regions being studied, in a standardized format that could be analyzed or easily converted to a format that could be analyzed, and ecologically relevant to the gradients being considered. The database has been posted online to facilitate further research. This is intended to be a 'living' database, and researchers are encouraged to contribute data and provide suggestions or feedback on how the database can be expanded and improved upon in the future.

TABLE OF CONTENTS

LIST OF TABLES

PREFACE

The report and database were prepared by Tetra Tech, Inc. and the Global Change Research Program (GCRP) in the National Center for Environmental Assessment (NCEA) of the Office of Research and Development at the U.S. Environmental Protection Agency (EPA). They are intended for resource managers and scientists working in freshwater ecosystems who are interested in species traits, biological indicators, bioassessment, biomonitoring, and climate change. The database is intended to be modified and augmented by scientists and resource managers with data and research results.

AUTHORS, CONTRIBUTORS, AND REVIEWERS

AUTHORS

Center for Ecological Sciences, Tetra Tech, Inc., Owings Mills, MD
Jen Stamp, MS, Anna Hamilton, MS, Liejun Wu, MS, Jeffrey White, MS

U.S. EPA
Britta G. Bierwagen, PhD

REVIEWERS

U.S. EPA Reviewers
Wayne Davis, MS (OEI), Lilian Herger, MS (R10), Rachael Novak, MS (OW/OST), Lester Yuan, PhD (ORD/NCEA, now OW/OST)

Other Reviewers
Daren Carlisle, PhD (USGS), M. Siobhan Fennessy, PhD (Kenyon College), Eric P. Smith, PhD (Virginia Polytechnic Institute and State University), R. Jan Stevenson, PhD (Michigan State University), N. Scott Urquhart, PhD (Statistical Consultant)

1. INTRODUCTION

The Freshwater Biological Traits Database (http://www.epa.gov/ncea/global/traits) was compiled as part of a project conducted by the Global Change Research Program (GCRP) in the National Center for Environmental Assessment in the U.S. Environmental Protection Agency (EPA) Office of Research and Development on climate change effects on river and stream ecosystems (U.S. EPA, 2011). For this project, long-term trend analyses were performed on biomonitoring data from Maine, North Carolina, Ohio, and Utah to examine whether biological responses to changes in temperature and hydrology could be detected. One component of these analyses involved compiling and analyzing traits data for North American macroinvertebrate taxa found in lotic systems. Species traits are the characteristics that explain an organisms' relationship to the surrounding environment, including growth, feeding habits and dispersal. Advantages of using traits data for these types of analyses are that they are less susceptible to taxonomic ambiguities or inconsistencies in long-term data sets; they can detect changes in functional community characteristics; and they vary less across geographical areas, which allows for larger-scale trend analyses across regional species pools. Because it took substantial effort to gather the traits data into one place, and because we would like to save other researchers from having to undergo similar efforts, we have integrated the traits data that were gathered for this project into one database and have made it available online.

2. METHODS

Data gathering efforts focused on data that were published or well-documented, available, appropriate for the regions being studied, in a standardized format that could be analyzed or easily converted to a format that could be analyzed, and ecologically relevant to the gradients being considered. The data search revealed that traits data compilations in North America have been at smaller scales and are less comprehensive than the European efforts (i.e., Euro-limpacs Consortium: www.freshwaterecology.info—The Taxa and Autecology Database for Freshwater Organisms), but nevertheless show promise. In 2006, the U.S. Geological Survey (USGS) published a database of lotic invertebrate traits for North America (Vieira et al., 2006). This database represented the first comprehensive summary of traits for North American invertebrate taxa and the first effort to compile this traits information in a Web-based database.

The traits information was gathered from over 3,000 keys, texts, peer-reviewed publications, and reports on North American aquatic invertebrates.

Another important source of traits information for North American lotic insect taxa is the Traits Matrix that was published in Poff et al. (2006). The Traits Matrix provides information on 20 traits (in 59 trait states) that span four broad categories of trait groups (life history, morphological, mobility, and ecological) for 311 taxa from 75 families. The traits information in the Traits Matrix was cross-referenced with the USGS (i.e., Vieira et al., 2006) traits database described above. An older series of publications was also included in the traits database: the EPA series on environmental requirements and pollution tolerance of Ephemeroptera, Plecoptera, Trichoptera, and Common Freshwater Chironomidae (Surdick et al., 1978; Beck et al., 1977; Harris et al., 1978; Hubbard et al., 1978). Traits information in these publications was compiled from general literature searches. The database created for this project contains information on 362 Plecoptera taxa, 240 Trichoptera taxa, 218 Chironomidae taxa, and 396 Ephemeroptera taxa from this older series of EPA publications.

Also included in the database are thermal optima and tolerance data that were generated from weighted average or generalized linear model calculations that were performed on biomonitoring data from Maine, North Carolina, Ohio, and Utah (U.S. EPA, 2011), as well as from Oregon (Yuan, 2006), Idaho (Brandt, 2001), and the Lahontan/Sierra Nevada region of California (Herbst and Silldorff, 2007). Weighted-average inference is a simple, robust approach for estimating the central tendencies of different taxa, for example, optima and tolerance values (i.e., ter Braak and Looman, 1986). For the climate change pilot study analyses in Maine, North Carolina, and Utah, the guidelines of Yuan (2006) were used to calculate optima values based on instantaneous water-temperature measurements and occurrences of organisms. Optima values for Maine and Utah were derived from weighted-average inferences. The lists for Utah were supplemented with weighted-average inferences derived from data sets from Idaho (i.e., Brandt, 2001) and Oregon (i.e., Yuan, 2006). Maximum-likelihood inferences were used in North Carolina because North Carolina Department of Environment and Natural Resources abundance data are categorical (1 = rare: 1–2 species; 3 = common: 3–9 species; 10 = abundant: ≥10 species). To improve model performance, optima values were calculated only for taxa occurring in >9 sites or samples.

These tolerance data were used to derive lists of cold- and warm-water-preference taxa in Maine, North Carolina, and Utah. Because the methods used to derive the thermal optima values and the specific characteristics of the data sets (e.g., range of collection dates, station locations, elevation) varied, an arbitrary ranking scheme was developed to make results more comparable across data sets. Taxa in each state were assigned rankings ranging from 1 to 7 based on percentiles within each data set. Initially, taxa with rankings ≤3 (<40th percentile) were designated as cold-water taxa and taxa with rankings ≥5 (>60th percentile) as warm-water taxa. Thermal optima values were not available for all taxa, so literature—primarily the traits matrix in Poff et al. (2006) and the USGS traits database (Vieira et al., 2006)—were used as a basis for making some additional initial designations.

After making initial cold- and warm-water designations, the lists in each state were refined based on case studies and best professional judgment from regional advisory groups. Thermal tolerance values, which were calculated using the methods described above (i.e., Yuan, 2006), were also taken into consideration. We thought these additional considerations were necessary because some taxa occurred with greater frequency in warm- or cold-water habitats but were not present exclusively in one or the other. For example, some taxa initially designated as cold-water taxa also were present at sites that had the hottest recorded water temperatures. During the refinement process, these taxa were removed from the cold-water list. In some cases, taxa were removed from the lists because regional taxonomists did not think that the literature-based designations were appropriate for their region. The cold-water-preference lists in Maine, North Carolina, and Utah consisted of 39, 32, and 33 taxa, respectively. The warm-water-preference lists in Maine, North Carolina, and Utah consisted of 40, 27, and 16 taxa, respectively. Lists of the cold and warm water taxa can be found in Appendix A. The relatively low number of taxa on the Utah warm-water-preference list was partially a consequence of the need to use a family-level operational taxonomic unit (OTU) for Chironomidae because of inconsistencies in the long-term data set that arose from a change in taxonomic laboratories.

These lists of cold- and warm-water taxa are included in region-specific traits tables that were compiled for the Maine, North Carolina, and Utah climate change pilot study analyses (U.S. EPA, 2011). Also included in these tables are information on traits related to life-cycle features (i.e., life-cycle duration, reproductive cycles per year, aquatic stages), resilience or resistance potentials (i.e., dispersal, locomotion, resistance forms), physiology and morphology

(i.e., respiration, maximum size), and reproduction and feeding behavior (i.e., reproduction, food, and feeding habits). Table 1 contains a list of the traits that were included the climate

Table 1. Summary of the traits and trait states in the Maine, North Carolina, and Utah climate change traits tables (modified from Poff et al., 2006)

Trait Category	Trait	Trait States
Life history	Voltinism	Semivoltine (<1 generation/yr), univoltine (1 generation/yr), bi- or multivoltine (>1 generation/yr)
	Development	Fast seasonal, slow seasonal, nonseasonal
	Synchronization of emergence	Poorly synchronized (wk), well synchronized (d)
	Adult life span	Very short (<1 wk), short (<1 mo), long (>1 mo)
	Adult ability to exit	Absent (not including emergence), present
	Ability to survive desiccation	Absent, present
Mobility	Dispersal (adult)	Low (<1 km flight before laying eggs), high (>1 km flight before laying eggs)
	Adult flying strength	Weak (e.g., cannot fly into light breeze), strong
	Occurrence in drift	Rare (catastrophic only), common (typically observed), abundant (dominant in drift samples)
	Maximum crawling rate	Very low (<10 cm/h), low (<100 cm/h), high (>100 cm/h)
	Swimming ability	None, weak, strong
Morphology	Attachment	None (free-ranging), some (sessile, sedentary)
	Armoring	None (soft-bodied forms), poor (heavily or partly sclerotized), good (i.e., some cased caddisflies, hard-shelled organisms)
	Shape	Streamlined (flat, fusiform), not streamlined (cylindrical, round or bluff)
	Respiration	Tegument, gills, plastron or spiracle (aerial)
	Size at maturity	Small (<9 mm), medium (9−16 mm), large (>16 mm)
Resource acquisition/preference	Rheophily	Depositional, depositional and erosional, erosional
	Habit (primary)	Burrower, climber, sprawler, swimmer, clinger, diver, skater
	Functional feeding group (primary)	Collector-filterer, collector-gatherer, predator, shredder, scraper, piercer, herbivore, parasite

Table 1. continued…

Trait Category	Trait	Trait States
Temperature	Temperature optimum	Numeric value derived from weighted average calculation
	Temperature tolerance	Numeric value derived from weighted average calculation
	Rank of temperature optimum	Scores range from 1 (lowest optima values) to 7 (highest optima values), based on percentile of optimum value
	Rank of temperature tolerance	Scores range from 1 (narrowest tolerance ranges) to 7 (widest temperature ranges), based on percentile of tolerance value
	Rank of temperature optimum-tolerance	Combination of the optimum and tolerance ranks. Values range from 1–1 to 7–7
	Temperature indicator	Cold or warm. Designations were made by Jen Stamp of Tetra Tech, Inc., based on weighted average or maximum likelihood calculations, literature, best professional judgment, and case studies
Enrichment tolerance	Tolerance	Values range from 0 (most intolerant) to 10 (most tolerant)

change traits tables, which were modeled after the Poff et al. (2006) Traits Matrix. These traits were selected for their relevance to the climate change pilot studies, which focused on biological responses to changes in temperature and hydrology.

Data from multiple sources were incorporated into the Maine, North Carolina, and Utah climate change traits tables. Main sources were the USGS traits database (Vieira et al., 2006) and the Poff et al. trait matrix (2006), which were available in an electronic format and were imported directly into the database. The EPA's 1970s publications had to be hand-entered. Quality assurance procedures were performed on 10% of these entries, and the data entry error rate was less than 5%. To maintain consistency and standardization across the multiple data sources, data integration rules were developed. These rules are described in detail in the 'Data Integration Rules' documents (see Appendix B). Efforts were also made to identify gaps in each traits data set. Results of these 'traits gap' analyses can be found in the 'Traits Gap Analysis' documents (see Appendix C).

Although species-level data were available in each of the state databases, genus-level or higher OTUs were used in the Maine, North Carolina, and Utah climate change traits tables. This was due to taxonomic ambiguities in the long-term data that had resulted from factors such as changes in taxonomic keys and changes in taxonomic labs. Previous research has shown that traits analyses utilizing genus and family levels have been successful at characterizing aquatic communities for bioassessment purposes (i.e., Vieira et al. [2006] cites Dolédec et al. [1998, 2000] and Gayraud et al. [2003]) and that congeneric species typically have similar functional trait niches (Poff et al., 2006). Species-level identification is typically not necessary for traits-based analytical approaches used in biomonitoring programs, is more costly and error prone, and may result in taxonomic ambiguities because individuals are not identifiable to the same taxonomic level (Vieira et al. [2006] who also cites Moulton et al. [2000]).

3. RESULTS

The Freshwater Biological Traits Database is available online at http://www.epa.gov/ncea/global/traits. The database currently has 11,912 unique records for 3,857 different taxa and includes location, habitat, life history, mobility, morphology, and ecological traits data, along with tolerance calculations for temperature and flow. A list of traits

and metadata can be found in Appendix D. Levels of taxonomic resolution vary, as do data types (i.e., binary, categorical, text notes entries). Instructions on how to conduct data searches can be found in Appendix E.

Listed below are brief descriptions of the 14 data sources that have been integrated into the database at this time. These data sources are available for download online on the Data Source page.

- **Vieira et al., 2006**
  Description: In 2006, the U.S. Geological Survey (USGS) published a database of lotic invertebrate traits for North America. This was a collaborative effort between the USGS National Water-Quality Assessment Program and Colorado State University. This database represented the first comprehensive summary of traits for North American invertebrate taxa and the first effort to compile this traits information in a Web-based database. The traits information was gathered from over 3,000 keys, texts, peer-reviewed publications, and reports on North American aquatic invertebrates. Traits were grouped into four general categories: ecology, morphology, behavior, or physiology. Trait states were established based on the types of information available in the literature and were expressed in categorical, binary, and quantitative terms. The traits could be mutually exclusive (only one or the other) or co-occurring (more than one trait state is appropriate and is, therefore, listed). Species-level resolution was used, but the focus and quality assurance efforts were concentrated on genus and family-level trait summaries.

- **Poff et al., 2006**
  Description: The Traits Matrix in the Appendix of this journal article provides information on 2 traits (in 59 trait states) that span four broad categories of trait groups (i.e., life history, morphological, mobility, and ecological) for 311 taxa from 75 families. Each trait has anywhere from 2 to 6 trait states. Each taxonomic unit is assigned to only one trait state (based on literature information and expert opinion). The traits information in the Traits Matrix was cross-referenced with the USGS (i.e., Vieira et al., 2006) traits database. This database is in a format that can be readily analyzed.

- **U.S. EPA, 2011**

  Description: These tables were compiled for the Maine, North Carolina, and Utah climate change pilot study analyses. The focus of these analyses was to look for biological responses to changes in temperature and hydrology. Data from multiple sources are incorporated into these data sets. Main sources include the USGS traits database (2006) and the Poff et al. trait matrix (2006).

- **Rankin and Yoder, 2009**

  Description: This report was prepared by the Midwest Biodiversity Institute for the USEPA GCRP Climate Change Pilot Project (U.S. EPA, 2011). Appendix Table 2 of the report contains thermal optima and current optima data (referred to as Weighted Stressor Values [WSVs] in this document) for macroinvertebrates in headwater and wadeable streams and were calculated using Ohio EPA data. In addition to weighted average values, general tolerance and functional feeding group assignments specific to Ohio were included in the database entries. Fish data are also available in Appendix Table 2 but have not yet been incorporated into the Freshwater Biological Traits Database.

- **Brandt, 2001**

  Description: Thermal optima and tolerance data for were obtained from Idaho Department of Environmental Quality (DEQ). Data were derived from Idaho DEQ bioassessment program samples collected from water bodies throughout Idaho. Included in this report is a list of cold water obligate taxa, which are based on Idaho's water quality criterion for cold water taxa (which is not to exceed a daily average stream temperature of 19ºC).

- **Herbst and Silldorff, 2007**

  Description: Thermal optima data for 99 taxa were provided by David Herbst and Erik Silldorff of the Sierra Nevada Aquatic Research Laboratory—University of California (see pages 9–11 of report). Data were derived from summer sampling events in the eastern Sierra Nevadas. Taxa were designated as 'thermal sensitive' if the optima values were ≤13ºC and 'thermal tolerant' if the optima values were ≥17ºC.

- **Huff et al., 2008**

  Description: Thermal optima and tolerance data for 234 taxa were provided by Shannon Hubler of Oregon DEQ. These data were derived from Oregon DEQ data from a wide range of wadeable stream types and span all of the major ecoregions in Oregon.

- **Yuan, 2006**

  Description: Thermal optima values from Table C-1 in Appendix C of this report were entered into the database. These data were derived from EMAP-West samples that were collected in 2000-2001.

- **EPA 1970s series on environmental requirements and pollution tolerance of aquatic macroinvertebrates**

  Description: Traits information for this series was compiled from general literature searches (it does not include exhaustive surveys of the literature, only major sources). Data are grouped into broad categories such as general habitat, specific habitat, turbidity, current, temperature, pH, dissolved oxygen, seasonal distribution, timing of emergence, and geographical distribution (by EPA region). Each page has a species profile that summarizes the range of environmental conditions under which the species has been found (values and ranges reflect the experimental and observational bias of each study), along with the sources from which the information was gathered. These publications were intended to provide a baseline to which further information could be added as further research was conducted and more information became available. Some might consider the information in these publications to be outdated. However, there have been very few comprehensive efforts to gather this information (especially that compile and publish it in one place and in a consistent format), and the comprehensive bibliographies and documentation are very valuable. Electronic copies of this publication are not available, and hard copies are difficult and expensive to obtain. To obtain lists of citations for the primary literature that was reviewed for these publications, one needs to reference the hard copies. This series is composed of four publications:

- **Beck, 1977**

  Description: Information on 216 Chironomidae taxa was taken from this
  publication and included in the online database.

- **Harris and Lawrence, 1978**

  Description: Information on 240 Trichoptera taxa was taken from this publication
  and included in the online database.

- **Hubbard and Peters, 1978**

  Description: Information on 396 Ephemeroptera taxa was taken from this
  publication and included in the online database.

- **Surdick and Gaufin, 1978**

  Description: Information on 362 Plecoptera taxa was taken from this publication
  and included in the online database.

4. FUTURE DIRECTIONS

Currently, there are no plans to further develop this database, although there are several
possible directions. Next steps could include adding fish and periphyton data, along with more
functionality (e.g., new queries, automated import function, interactive map). The automated
import function in particular is important because in order for this database to reach its full
potential, researchers will need to actively contribute to it. Further development of this database
would also benefit from collaborations with other agencies, institutions, and researchers,
domestically and internationally, interested in freshwater species traits.