Exploring the NEON Biorepository data portal with Symbiota

Kelsey Yule & Nico Franz
Biocollections & Biodiversity Knowledge Integration Center
Arizona State University

Ecological Society of America
14 August 2019
Louisville, KY (Osage, Shawnee, & Miami lands)
The National Ecological Observatory Network (NEON) is a continental scale program to monitor and forecast ecological change.
The National Ecological Observatory Network (NEON) is a continental scale program to monitor and forecast ecological change. Standardized organismal and environmental data collection for the next 30 years.
The National Ecological Observatory Network (NEON) is a continental scale program to monitor and forecast ecological change. Standardized organismal and environmental data collection for the next 30 years. 81 terrestrial and freshwater sites within 20 ecoclimatic domains across the US, including AK, HI, and PR.
The National Ecological Observatory Network (NEON) is a continental scale program to monitor and forecast ecological change. 81 terrestrial and freshwater sites within 20 ecoclimatic domains across the US, including AK, HI, and PR. Standardized organismal and environmental data collection for the next 30 years. 180 data products, openly available at https://data.neonscience.org.
100k+ samples / year

40+ sample classes

**Aquatic**
- algae, macrophytes, mosses, lichens, liverworts, microbes, fish, zooplankton, macroinvertebrates

**Terrestrial**
- herptiles, invertebrates, plants, small mammals, soil microbes

**Environmental**
- soils, particulate mass filters, wet deposition

**Genomic**
- beetles, fish, macroinvertebrates, small mammals, soil microbes, zooplankton
Physical samples and specimens are collected at all 81 NEON sites.
Domain facilities process samples and send them to the Biorepository at ASU
Samples are arriving and being curated daily.
<table>
<thead>
<tr>
<th>Sample Type</th>
<th>Volume/Year</th>
<th>Sample Type</th>
<th>Volume/Year</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Plant</strong></td>
<td></td>
<td><strong>Vertebrate</strong></td>
<td></td>
</tr>
<tr>
<td>DNA extractions (whole) ▲</td>
<td>500</td>
<td>Fish – DNA extractions ▲</td>
<td>1,100</td>
</tr>
<tr>
<td>Slides ■</td>
<td>700</td>
<td>Fish - fin clip ▲</td>
<td>275</td>
</tr>
<tr>
<td>Whole sample ■</td>
<td>700</td>
<td>Fish – voucher ▲</td>
<td>275</td>
</tr>
<tr>
<td>Macroalgae - whole sample ▲</td>
<td>250</td>
<td>Herptiles – pitfall bycatch ▲</td>
<td>200</td>
</tr>
<tr>
<td><strong>Aquatic Lichens, Mosses &amp; Liverworts</strong></td>
<td></td>
<td>Small mammal - blood ▲</td>
<td>2,000</td>
</tr>
<tr>
<td>Vouchers ▲</td>
<td>55</td>
<td>Small mammal – fecal ▲</td>
<td>6,000</td>
</tr>
<tr>
<td><strong>Invertebrate</strong></td>
<td></td>
<td>Small mammal – hair/whisker ▲</td>
<td>4,000</td>
</tr>
<tr>
<td>Carabids – DNA extractions ▲</td>
<td>4,500</td>
<td>Small mammal – ear punch ▲</td>
<td>4,000</td>
</tr>
<tr>
<td>Carabids – pinned/pointed ■</td>
<td>9,000</td>
<td>Small mammal - voucher ▲</td>
<td>750</td>
</tr>
<tr>
<td>Carabids – pooled ▲</td>
<td>2,500</td>
<td><strong>Environmental</strong></td>
<td></td>
</tr>
<tr>
<td>Macroinvertebrates - pooled ▲</td>
<td>800</td>
<td>Litterfall ▲</td>
<td>125</td>
</tr>
<tr>
<td>Macroinvertebrates – DNA extractions ▲</td>
<td>300</td>
<td>Soil – frozen ▲</td>
<td>21,000</td>
</tr>
<tr>
<td>Pitfall – pooled ▲</td>
<td>3,750</td>
<td>Soil – dry ▲</td>
<td>375</td>
</tr>
<tr>
<td>Mosquitoes – DNA extractions ▲</td>
<td>1,750</td>
<td>Particulate mass filters (PM10) ▲</td>
<td>150</td>
</tr>
<tr>
<td>Mosquitoes – disease pools ▲</td>
<td>21,000</td>
<td>Wet deposition ▲</td>
<td>1,100</td>
</tr>
<tr>
<td>Mosquitoes – pinned/pointed ■</td>
<td>3,000</td>
<td><strong>Microbial</strong></td>
<td></td>
</tr>
<tr>
<td>Mosquitoes – pooled ▲</td>
<td>3,500</td>
<td>Aquatic microbes - Sterivex filters ▲</td>
<td>850</td>
</tr>
<tr>
<td>Tick – disease pools ▲</td>
<td>4,500</td>
<td>Soil microbes – DNA extractions ▲</td>
<td>2,500</td>
</tr>
<tr>
<td>Zooplankton – DNA extractions ▲</td>
<td>75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zooplankton – pooled ▲</td>
<td>75</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total &gt;104,000 samples per year</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Wet
- Dry / 4°C
- LN₂
- -80 / -20°C
Samples are already being used for research

Dr. Michael Weiser (University of Oklahoma)
Bulk invertebrates – Macroecology, species discovery, community dynamics

Identified mosquitos - Phylogenetomics

DNA extracts from soil - Microbial ecology
NEON Biorepository Team

Dr. Nico Franz
Principal Investigator

Laura Steger
Environmental & Zoological Collections Manager

Dr. Laura Rocha Prado
Bioinformatician (II)

Ed Gilbert
Bioinformatician (I)

Dr. Kelsey Yule
Project Manager

Azhar Husain
Cryo Collections Manager

Dr. Andrew Johnston
Invertebrate Collections Manager

Biodiversity Knowledge Integration Center
Arizona State University
How do I access NEON sample and specimen data?
How do I access NEON sample and specimen data?

Biorepository Data Portal
https://biorepo.neonscience.org/
How do I access NEON sample and specimen data?

biorepo@asu.edu

Biorepository Data Portal

https://biorepo.neonscience.org/
Discover and access sample-based data
Discover and access sample data through our Symbiota portal like SEINet or SCAN.
Symbiota - A virtual platform for creating voucher-based biodiversity information communities

Corinna Gries†, Edward E. Gilbert†, Nico M. Franz‡
† University of Wisconsin, Madison, Madison, United States of America
‡ Arizona State University, Tempe, United States of America

Most frequently used software in North America for managing Natural History Collections records

Based on Darwin Core standards

Mid-level data aggregation platform and user-driven content management system

View and contribute directly to sample data and publish in this portal

Abstract

We review the Symbiota software platform for creating voucher-based biodiversity information portals and communities. Symbiota was originally conceived to promote small- to medium-sized, regionally and/or taxonomically themed collaborations of natural history
Explore available samples
Sample search | Map search | Checklists

> 62,000 samples | > 400 taxa

Distribution of samples by collection type.

Distribution of samples by top 5 determined taxa.

Data
Visit the Data Usage Policy page for information on how to cite data obtained from the NEON Biorepository Data Portal.

Specimens
Please consult the Archival Sample Request information page to initiate inquiries about sample accessibility and loans.

Contact
Join the portal as a regular visitor or contributor, and send direct feedback or inquiries to BioRepo@asu.edu.
Mosquito Voucher Collection (NEON-MOSC)
Enter Search Parameters

Fill in one or more of the following query criteria and click "Search" to view your results.

Taxonomic Criteria:

- Include Synonyms
- Scientific Name

Locality Criteria:

Country:
Latitude and Longitude

Bounding box
Northern Latitude: [ ]
Southern Latitude: [ ]
Western Longitude: [ ]
Eastern Longitude: [ ]

Polygon (WKT footprint)

Point-Radius
Latitude: [ ] N
Longitude: [ ] W
Radius: [ ] Kilometers

Collector Criteria

Collector's Last Name: [ ]
Collector's Number: [ ]
Collection Date: [ ]

Specimen Criteria

Catalog Number: [ ]
Include other catalog numbers and GUIDs

Limit to Type Specimens
Limit to Specimens with Images
Limit to Specimens with Genetic Data
Include cultivated/captive occurrences
Include other catalog numbers and GUIDs
Firefox
Enter Search Parameters

Fill in one or more of the following query criteria and click "Search" to view your results.

**Taxonomic Criteria:**

- **Scientific Name**: Aedes

**Locality Criteria**

- Country:
- State:
- County:
- City:
- Zip Code:
Mosquito Voucher Collection

**Aedes communis** (De Geer, 1776)
2016-05-13
United States, Massachusetts, Worcester, Harvard Forest Site, CORE, Plot HARV_082 (plot dimensions: 0m), 42.482279 -72.272148, 172m
Full Record Details

**Aedes fulvus pallens**
2017-07-18
United States, Alabama, Greene, Dead Lake Site, Plot DELA_032 (plot dimensions: 0m), 32.530107 -87.815715, 27m
Full Record Details

**Aedes punctor** (Kirby, 1837)
Map samples
Mosquito Voucher Collection

**Aedes communis** (De Geer, 1776)
- 2016-05-13
- United States, Massachusetts, Worcester, Harvard Forest Site, CORE, Plot HARV_082 (plot dimensions: 0m), 42.48279 -72.272148, 172m
- [Full Record Details](#)

**Aedes fulvus pallans**
- 2017-07-18
- United States, Alabama, Greene, Dead Lake Site, Plot DELA_032 (plot dimensions: 0m), 32.530107 -87.815715, 27m
- [Full Record Details](#)

**Aedes punctor** (Kirby, 1837)
Google Map

DISPLAY COORDINATES IN GOOGLE MAP

Google Maps is a web mapping service provided by Google that features a map that users can pan (by dragging the mouse) and zoom (by using the mouse wheel). Collection points are displayed as colored markers that when clicked on, displays the full information for that collection. When multiple species are queried (separated by semi-colons), different colored markers denote each individual species.

Google Earth (KML)

This creates an KML file that can be opened in the Google Earth mapping application. Note that you must have Google Earth installed on your computer to make use of this option.

CREATE KML
Google Map

Google Maps is a web mapping service provided by Google that features a map that users can pan (by dragging the mouse) and zoom (by using the mouse wheel). Collection points are displayed as colored markers that when clicked on, displays the full information for that collection. When multiple species are queried (separated by semi-colons), different colored markers denote each individual species.

Google Earth (KML)

This creates a KML file that can be opened in the Google Earth mapping application. Note that you must have Google Earth installed on your computer to make use of this option.

CREATE KML
View sample record details
Mosquito Voucher Collection

- **Aedes communis** (De Geer, 1776)
  - 2016-05-13
  - United States, Massachusetts, Worcester, Harvard Forest Site, CORE, Plot HARV_082 (plot dimensions: 0m), 42.4B279 - 72.272148, 172m
  - Full Record Details

- **Aedes fulvus pallens**
  - 2017-07-18
  - United States, Alabama, Greene, Dead Lake Site, Plot DELA_032 (plot dimensions: 0m), 32.530107 - 87.815715, 27m
  - Full Record Details

- **Aedes punctor** (Kirby, 1837)
Mosquito Voucher Collection

Secondary Catalog #: MOS.D01.000432
Taxon: Aedes communis (De Geer, 1776)
Family: Culicidae
Collector:
Date: 2016-05-13
Locality: United States, Massachusetts, Worcester, Harvard Forest Site, CORE, Plot HARV_082 (plot dimensions: 0m)
        42.482279 -72.272148 +10m. WGS84
Elevation: 172 meters (564ft)
Habitat: woodyWetlands; slope aspect: 159.51; slope gradient: 4.3; soil type order: Inceptisols
Description: List of individual mosquitoes preserved as vouchers

Usage Rights: CC0 1.0 (Public-domain)

Record Id: 7015c1d0-8ea8-49bf-801a-7e794752a8db

For additional information on this specimen, please contact: NEON Biorepository (biorepo@asu.edu)
Researchers encouraged to become portal managers to link all sample-associated data

Mosquito Voucher Collection

Secondary Catalog #: MOS.D01.000432
Taxon: Aedes communis (De Geer, 1776)
Family: Culicidae
Collector:
Date: 2016-05-13
Locality: United States, Massachusetts, Worcester, Harvard Forest Site, CORE, Plot HARV_082 (plot dimensions: 0m)
  42.482279 -72.272148 +10m. WGS84
Elevation: 172 meters (564ft)
Habitat: woodyWetlands; slope aspect: 159.51; slope gradient: 4.3; soil type order: Inceptisols
Description: List of individual mosquitoes preserved as vouchers

Usage Rights: CC0 1.0 (Public-domain)
Record Id: 7015c1d0-8ea8-49bf-801a-7e794752a8db

For additional information on this specimen, please contact: NEON Biorepository (biorepo@asu.edu)
Download sample data
Mosquito Voucher Collection

*Aedes communis* (De Geer, 1776)
2016-05-13
United States, Massachusetts, Worcester, Harvard Forest Site, CORE, Plot HARV_082 (plot dimensions: 0m), 42.482279 -72.272148, 172m
Full Record Details

*Aedes fulvus pallens*
2017-07-18
United States, Alabama, Greene, Dead Lake Site, Plot DELA_032 (plot dimensions: 0m), 32.530107 -87.815715, 27m
Full Record Details

*Aedes punc tor* (Kirby, 1837)
Data Usage Guidelines

By downloading data, the user confirms that he/she has read and agrees with the general data usage terms. Note that additional terms of use specific to the individual collections may be distributed with the data download. When present, the terms supplied by the owning institution should take precedence over the general terms posted on the website.

Download Georeference Data

File Format:  
- Comma Delimited (CSV)
- Tab Delimited

Character Set:  
- ISO-8859-1 (western)
- UTF-8 (unicode)

Compression:  
- Compressed ZIP file

[Download Data]
Map search
.csv file

Darwin Core record
Request samples
Discover and access sample-based data.
Sample Use Policy

NEON, the National Ecological Observatory Network, aims to provide "open data to understand how our aquatic and terrestrial ecosystems are changing". Therefore, the NEON Sample Use Policy reflects the need to provide researchers with access to NEON samples for a wide-variety of purposes while preserving the future research potential of those samples.

- **Contact us** well in advance of any grant proposal deadlines in order to optimally integrate sample uses into project narratives, data management plans, and budgets.
- The NEON Biorepository data portal is capable of hosting or linking to many forms of sample-associated data. Researchers using samples are strongly encouraged to become portal managers in order to disseminate their data to the public.
- The NEON Biorepository Advisory Committee, consisting of Biorepository and NEON staff as well as the external Biorepository Technical Working Group, may be consulted prior to approval of any sample use request.
- Sample uses can be non-invasive, invasive, consumptive, or destructive. Uses that reduce the future research potential of a sample will likely require stronger justification and a plan to disseminate all sample-associated data.
- Researchers are responsible for proper handling of all samples, adhering to all aspects of the sample use agreement, and following all NEON and National Science Foundation data reporting and citation policies.

Please read the full policy for more details and contact us for more information.
Sample Use Policy

Acceptable sample uses, sample use approval process

The National Ecological Observatory Network, aims to provide "open data to show our aquatic and terrestrial ecosystems are changing". Therefore, the NEON policy reflects the need to provide researchers with access to NEON samples for a variety of purposes while preserving the future research potential of those samples.

Contact us well in advance of any grant proposal deadlines in order to optimally integrate sample uses into project narratives, data management plans, and budgets.

The NEON Biorepository data portal is capable of hosting or linking to many forms of sample-associated data. Researchers using samples are strongly encouraged to consult with the portal managers in order to disseminate their data to the public.

The NEON Biorepository Advisory Committee, consisting of Biorepository and NEON staff as well as the external Biorepository Technical Working Group, may be consulted prior to approval of any sample use request.

- Sample uses can be non-invasive, invasive, consumptive, or destructive. Uses that reduce the future research potential of a sample will likely require stronger justification and a plan to disseminate all sample-associated data.
- Researchers are responsible for proper handling of all samples, adhering to all aspects of the sample use agreement, and following all NEON and National Science Foundation data reporting and citation policies.

Please read the full policy for more details and contact us for more information.
Sample Use Policy

NEON, the National Ecological Observatory Network, aims to provide "open data to understand how our aquatic and terrestrial ecosystems are changing". Therefore, the NEON Sample Use Policy reflects the need to provide researchers with access to NEON samples for a wide-variety of purposes while preserving the future research potential of those samples.

- **Contact us** well in advance of any grant proposal deadlines in order to optimally integrate sample uses into project narratives, data management plans, and budgets.
- The NEON Biorepository data portal is capable of hosting or linking to many forms of sample-associated data. Researchers using samples are strongly encouraged to become portal managers in order to disseminate their data to the public.
- The NEON Biorepository Advisory Committee, consisting of Biorepository and
Submit this form for Letters of Support or to receive samples for research

Contact us at biorepo@asu.edu with inquiries

Sample Use Request

NEON Biorepository at Arizona State University - Sample Use Request Form

Version 01.2 - August 8, 2019

NEON, the National Ecological Observatory Network, aims to provide “open data to understand how our aquatic and terrestrial ecosystems are changing”. NEON samples are meant to be used very frequently; however, immediate usage must be balanced against the potential future research value of those same samples. Additionally, data associated with these samples should be openly available. In keeping with these needs, the NEON Biorepository at Arizona State University (ASU) requires information in order to process sample use requests.

Visit the NEON Biorepository Data Portal (https://biorepo.neonscience.org) to explore sample availability. Samples are being added to the portal as they arrive at ASU. Some sample classes and legacy samples may be located at other facilities. Contact us (biorepo@asu.edu) if you are interested in samples not listed on the portal.

Please consult the latest version of the NEON Sample Use Policy for guidelines. Send an e-mail request to BioRepo@asu.edu for more information.

* Required

Email address *

Your email

NEXT
In development:
Full integration of sample requests into portal
Load sample records into R

Set Your Working Directory

```r
setwd("~/MyFilePath/")
```

Read in Your Data

```r
my.data<-read.csv("MyBiorepositoryData.csv")
```
Explore records in R
Relevant R packages for analyses of Biorepository occurrence record data

vegan (BiodiversityR) – descriptive community ecology
maxent – species distribution modeling
picante – phylogenetic community ecology
phenology – describe species’ phenologies
bdvis – visualize biodiversity data coverage, gaps, biases
spocc – download occurrence records from other biodiversity portals
neonUtilities – download and format NEON data
Main NEON data portal

https://data.neonscience.org/
Welcome to the NEON Data Portal

The National Ecological Observatory Network provides open data to understand changing ecosystems. NEON data are currently construction-grade and provisional - learn more at our Data Quality Program webpage. To learn more about NEON, check out the Resources tab above or visit our main portal by clicking the NEON icon in the upper left corner of this page. Visit the Data Product Catalog for more specific information about individual data products, the Data Availability webpage to learn more about when data will become available after collection, or Data Portal News for occasional updates.
Configure Data Package for Download

Mosquito-borne pathogen status

Product ID: DP1.10041.001

1. Select sites (by site code, state, or domain) and date range:

   Start Date: Jan 2015
   End Date: Aug 2019

2. Do you want to include documentation about this data product?
   - Yes, include relevant documents for this Data Product
     EML files for this Data Product are included in all downloads (More about EML at NEON and KNB.)

3. Select data type:
   - Basic: Includes the data product, summary statistics, expanded uncertainty, and final quality flag
   - File type to be downloaded: ZIP (Compressed Text)
     Estimated size: 0.425 MB

Download Data
.csv files packaged by month and site into zip files within a parent zip file
Use the neonUtilities Package to Access NEON Data

AUTHORS: Claire K. Lunch, Megan A. Jones

This tutorial goes over how to use the neonUtilities R package (formerly the neonDataStackR package).

The package contains several functions:

- `stackByTable()`: Takes zip files downloaded from the Data Portal or downloaded by `zipsByProduct()`, unzips them, and joins the monthly files by data table to create a single file per table.
- `zipsByProduct()`: A wrapper for the NEON API; downloads data based on data product and site criteria. Stores downloaded data in a format that can then be joined by `stackByTable()`.
- `loadByProduct()`: Combines the functionality of `zipsByProduct()` and `stackByTable()`: Downloads the specified data, stacks the files, and loads the files to the R environment.
- `getPackage()`: A wrapper for the NEON API; downloads one site-by-month zip file at a time.
Use the neonUtilities Package to Access NEON Data

AUTHORS: Claire K. Lunch, Megan A. Jones

This tutorial goes over how to use the neonUtilities R package (formerly the neonDataStackR package).

The package contains several functions:

- `stackByTable()`: Takes zip files downloaded from the Data Portal or downloaded by `zipsByProduct()`, unzips them, and joins the monthly files by data table to create a single per table.
- `zipsByProduct()`: A wrapper for the NEON API; downloads data based on data product and site criteria. Stores downloaded data in a format that can then be joined by `stackByTable()`.
- `loadByProduct()`: Combines the functionality of `zipsByProduct()` and `stackByTable()`: Downloads the specified data, stacks the files, and loads the files to the R environment.
- `getPackage()`: A wrapper for the NEON API; downloads one site-by-month zip file at a time.
NEON Biorepository data tutorials:

What would be most useful?
Contact us…

biorepo@asu.edu