Data papers about vectors of diseases

Vectors of human disease series

Scott Edmunds

ID 0000-0001-6444-1436
scott@gigasciencejournal.com
30th March 2023
Why we need to fill biodiversity data gaps

Expert predictions of species richness

Completeness of biodiversity records

https://www.nature.com/articles/ncomms9221
Why we need to fill biodiversity data gaps

Vector-borne diseases account for ~\( \frac{1}{4} \) of all infectious diseases

Global distribution of malaria, lymphatic filariasis, leishmaniasis, dengue, Japanese, encephalitis, yellow fever and Chagas disease

What are disease vectors?

 виды векторных заболеваний
Rewarding open data: *GigaScience*

Launched July 2012, now partnering with OUP. Publishes “Data Notes” for CC0 data.

http://gigasciencejournal.com/

Published by:

- **BioMed Central**
  - 2011-2016
- **Oxford University Press**
  - 2016-date

http://gigasciencejournal.com/
Rewarding open data: GigaScience

APC covers curation and 1TB of storage in our GigaDB repository

Since 2011, and working with

http://gigadb.org/
Signed on the 23rd November 2021 at the 41st session of UNESCO by 193 Members States

The Recommendation affirms the importance of open science as a vital tool to improve the quality and accessibility of both scientific outputs and scientific process, to bridge the science, technology and innovation gaps between and within countries and to fulfill the human right of access to science.

Recommends that Member States apply the provisions of this by taking whatever legislative or other measures may be required to give effect within their jurisdictions to the principles of this Recommendation.

https://unesdoc.unesco.org/ark:/48223/pf0000379949
Addressing this with a new journal: (GIGA)byte

Main advantage of workflow is XML from start to end

Traditional workflow: API spaghetti

New: straightforward, single platform

https://gigabytejournal.com/
Inclusive features: the cost barrier

Huge time + cost savings from XML-first workflow

DOCX/.latex to XML to paper. Automated production (only needing humans for pagination)

End-to-end analytics

Production times in hours not days

Current APC = $400 USD covers these publication costs (with 10% markup)

Real stats from DARVi (Data Analyser, Reporter and Visualiser) https://rivervalley.io/darvi/
Thinking about users: authors, reviewers, readers

What does a GigaByte data paper look like?

Data Release: a short, updatable, description of a research dataset

Discoverability & credit: Highlights and help to contextualize openly available datasets to encourage reuse.

Sharing: All data can be linked to the Data Release via GBIF, GigaDB or other data DOIs or accessions.

Data, not analysis: Incentivizes and allows more rapid releases of data before subsequent detailed analysis has been carried out. Or in coordination with publication of an analysis paper.

Simple: Structure = Context, Methods, Data Validation and QC, Reuse Potential, Data Availability

Submit via:

Integrates with preprints:

https://gigabytejournal.com/data-release-description
Access to data on vectors and vector-borne diseases is improved through the release of a special issue publication of a series of data papers

Publication of a series of 11 papers with data on vectors that transmit vector-borne diseases is a significant advance in the availability of such data in an easy and open access format. Published by the journal GigaByte by GigaScience Press, in partnership with the Global Biodiversity Information Facility and supported by TDR, the papers improve knowledge on vectors and pave the way forward in data sharing.

Currently affecting about half of the world’s population, vector-borne diseases are transmitted by arthropod (insect and tick) vectors such as mosquitoes and flies; they include diseases such as malaria, dengue and leishmaniasis. While some diseases such as malaria show a decline in numbers of fatalities, others, such as those caused by arboviruses like dengue, chikungunya and Zika viruses, show a worrying increase in number of cases (although improved case management has led to a decrease in fatalities). Factors such as climate change and urbanization have the potential to further affect the impact of these diseases, so it is vital that research is ongoing to understand more about the vectors and to ensure that data are shared widely with researchers and policy-makers in an open access, freely available and discoverable format.

Rewarding open data: TDR/GBIF

First thematic help desk to support mobilization of biodiversity data related to disease vectors

Efforts to boost mobilization of data on wild host, vector and reservoir species connected to human health

health@gbif.org

Additional content of submissions

Papers presenting sampling methods and protocols, diverse data types, examples working with citizen scientists and indigenous communities in the Amazon, and papers showcasing many novel technical features...

https://doi.org/10.46471/GIGABYTE_SERIES_0002
End product: Open Science for Public Health

Data published from the GigaByte Vectors of human disease series includes:

- **Vectors**
  - Mosquitos
  - Sand flies
  - Ticks
  - Triatominae

- **GBIF mediated data**
  - >500,000 Occurrence Records
  - Specimens >675,000

- **More than 50 Countries**

- **Data Openness**
  - CC0 Public Domain

- **Data Types**
  - Observations
  - Specimens
  - Imaging Data
  - DNA Barcodes
  - Citizen Science

PHASE II IS OPEN FOR SUBMISSIONS...submit by 30 April 2023.

https://doi.org/10.46471/GIGABYTE_SERIES_0002
End product: Open Science for Public Health
Multilingualism for better accessibility, understandability and trust

https://doi.org/10.46471/gigabyte.61
End product: Open Science for Public Health

Multilingualism for better accessibility, understandability and trust

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9876538/
Notes on the vectors series

The data papers submitted should describe datasets with the following criteria:

- Data has clear relevance for research on vectors of human vector-borne diseases
- Dataset contains more than 5,000 records that are new to GBIF.org in 2022/23 with high-quality data and metadata
- Data is dedicated to the public domain under an open CC0 designation
- Deadline for submission is 30 April 2023
Notes on the vectors series

Data deposition is key, and supported by GBIF helpdesk and GigaDB curators

• Authors should start by preparing the dataset and publishing it through GBIF.org before writing
• Support from health@gbif.org for questions on publishing data through GBIF, data standards, etc.
• GigaDB team (database@gigasciencejournal.com) on hand to help with additional supporting data
• GigaDB curators will also help review process by providing a data audit for each submission
Publishing data to support the fight against human vector-borne diseases


A GigaScience Press, BGI Hong Kong Tech Co Ltd, Rm 26F Kings Wing Plaza 2, 1 On Kwun Street, Shek Mun, Sha Tin, NT, Hong Kong SAR

1Special Programme for Research & Training in Tropical Diseases (ISP), World Health Organization, Avenue Appia 20, 1211 Geneva 27, Switzerland
2GIGF Secretariat Universitetsparken 15, DK-2100 Copenhagen Ø, Denmark
3Coleção de Fitoterímeds (FOCHRIO/COLFEP), Instituto René Rachou, FIOCRUZ Minas Avenida Augusto de Lima, 1715 - Baro Preto, 30190909, Belo Horizonte, Brazil
4Group of Estudos em Leishmanioses, Instituto René Rachou, FIOCRUZ Minas Avenida Augusto de Lima, 1715 - Baro Preto, 30190909, Belo Horizonte, Brazil, Brazil
5Group of Entomology, Instituto Nacional de Salud, Bogota, Colombia
6Secretaría de Salud Departamental del Cauca, Popayan, Colombia
7Universidad de Ciencias Aplicadas y Ambientales, Bogota, Colombia
8Programa de Estudios de Medio Ambiente Tropicales (PIMENT), Medellin, Colombia
9Centro de Estudios Avanzados de Biología (CEAB-CIC), Ce蔥/a la Ca St Francisco 14, 17000 Blanes, Girona, Spain; Centre d’Estudis Avançats de Blanes (CEAB-CIC), Ce蔥/a la Ca St Francisco 14, 17000 Blanes, Girona, Spain
10Centre de Recerca Ecológica i Aplicacions Forestals (CREAF), Edificio C Campus de Bellaterra, Barcelona, Spain
11Department of Geosciences and Social Studies, Universitat Pompeu Fabra, Placa de la Meridiana, 11-13, 08002 Barcelona, Spain
12Institut Català de Recerca i Estudis Avançats (ICREA), 25 Passage de Lluis Companys, 08010 Barcelona, Spain
13Istituto Cavalli di Recerca e Studi Avanzati (ICRC), 25/27/29 Piazza del Popolo, Roma, Italy
14Research Institute for Nature and Forest (INBO), Havenlaan 45 b/5, 1000, Brussels, Belgium
15Research Institute for Nature and Forests (INBO) - Boek-Sciences Heritage Service, Havenlaanstraat 29, 1000, Brussels, Belgium
16Arctic-GIS INC, Koksarvesti 33, 9000, Ivalo, Finland
17Instituut voor Bioloogisch-Medische Wetenschappen, Universiteit Gent, De Pintelaan 185, 9000, Gent, Belgium
18Department of Tropical Medicine (STO), Nationalleibnitz 153, 2000, Antwerp, Belgium
19Research Museum of Central Africa (RMCA - Boek), Leuvensesteenweg 17, 3080 ten Boom, Belgium
20Agency for Nature and Forests (ANB), Havenlaan 45 b/5, 1000, Brussels, Belgium
21Independent Researcher, Belgium
22Universidad Federal do Acre, Departamento de Ciências da Saúde e Educação Física. Universidade Federal do Acre, Distrito Industrial, Rio Branco, 6900900, Brazil
23Centro de Estudios Paracifolíticos y de Vientos (CEPV-CCV) La Palma (CONCET-CVEN), La Palma, Buenos Aires 190, Argentina
24Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentina
25College of Veterinary Medicine and Biobio Sciences, Texas A&M University, College Station, Texas, USA
26Museo del Instituto de Zoología Agüito (MIZA), Maracay, Venezuela
27Corresponding author. Scott Edmunds, BGI Hong Kong Tech Co Ltd., Hong Kong, NT HONG KONG. E-mail: scott@gigasciencejournal.com

Abstract

Vector-borne diseases are responsible for more than 17% of human cases of infectious diseases. In most situations, effective control of debilitating and deadly vector-borne diseases (VBDs), such as malaria, dengue, chikungunya, yellow fever, Zika and Chagas requires up-to-date, robust and comprehensive information on the presence, diversity, ecology, biomics and geographic spread of the organisms

Received: October 21, 2022. Accepted: October 24, 2022. © The Author(s) 2022. Published by Oxford University Press Gigascience. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by-4.0), which permits unrestricted reuse, redistribution and reproduction in any medium, provided the original work is properly cited.

http://dx.doi.org/10.1093/gigascience/giac114
Thanks to TDR/WHO for support of this datasets on vectors of human diseases series

Due to this very generous sponsorship the article processing fee (normally $400 USD) will be waived for the first 15 papers that are accepted and meet the series criteria.
Many thanks to our partners

For further questions contact: editorial@gigabytejournal.com

Submit by 30th April 2023:
https://gigabytejournal.com/