

The background of the slide features a repeating pattern of small, green seedlings growing in circular containers, possibly trays or pots. The seedlings are in various stages of growth, with some showing distinct root systems and others with small leaves. The overall color palette is a vibrant green, creating a fresh and agricultural atmosphere.

Companion planting

Generalist and specialist
repositories working together
to promote agricultural data
sharing and reuse

About me

- Head of Community Engagement at [Dryad](#)
- A trained librarian with an interest in advancing open and equitable scholarly communication



DRYAD

Not-for-profit and online since
2008. Learn more: datadryad.org

About Dryad

An open data publishing platform
and community committed to the
open availability and **routine
re-use** of all research data

50,00+ data publications
200,000+ researchers
70,000+ institutions
1,270+ academic journals



Data and metadata are:

- **Curated by our expert team to ensure suitability for publishing and to facilitate discovery and reuse**
- **Permanently stored in our CoreTrustSeal repository**
- **Published under a Creative Commons Public Domain (CCo) license**
- **Accessible via our open API**



DRYAD

Not-for-profit and online since
2008. Learn more: datadryad.org

What to expect

- A brief introduction to Dryad
- A perspective on the value of generalist and specialist data publishers
- A case study of data sharing in multiple platforms
- Collaborative strategizing on enhancing connections



DRYAD

Not-for-profit and online since
2008. Learn more: datadryad.org

Science is a social process

www.science.org/doi/10.1126/science.adi0333

Discoveries do not become knowledge until the findings are shared with the scientific community, to be vetted, challenged, and expanded on.

Data sharing policies

“Whenever possible, researchers should deposit data in discipline-specific or data-type specific repositories.

Otherwise, they should use a trusted generalist repository.”

What happens when a project has both?



DRYAD

Not-for-profit and online since
2008. Learn more: datadryad.org

Generalist repositories

Accommodate heterogeneous data

Provide a home for data that might not have a place elsewhere

Facilitate broad and serendipitous discovery

Expands potential audience for data

May support custom metadata and linking to related content managed elsewhere

May be free to use (within limits)

Specialist repositories

Optimized for disciplinary needs

Enhance discovery and reuse by using discipline-specific metadata

Support niche file formats

May support embedded visualization or analysis

Serve as “community hubs”



DRYAD

Not-for-profit and online since
2008. Learn more: datadryad.org

Data sharing in multiple platforms

- **When producing heterogeneous data types or formats**
 - Are there subsets of data that belong in an accepted specialist repository due to their file type or subject matter?
 - Are there funder requirements governing where data should be deposited?
- **When there are security and confidentiality considerations**
 - If raw data is too sensitive for broad sharing, are there subsets or processed datasets that could be openly distributed?

Focus on reuse

Reproduce, replicate, repurpose,
build-upon, perform meta-analysis,
seed machine learning algorithms




















DRYAD

Not-for-profit and online since
2008. Learn more: datadryad.org



Reproducibility in ecology and evolution: Minimum standards for data and code

Gareth B. Jenkins¹  | Andrew P. Beckerman²  | Céline Bellard³  |
Ana Benítez-López⁴  | Aaron M. Ellison^{5,6}  | Christopher G. Foote¹  |
Andrew L. Hufton⁷  | Marcus A. Lashley⁸  | Christopher J. Lortie⁹  | Zha
Allen J. Moore¹¹  | Shawn R. Narum¹²  | Johan Nilsson¹³  | Bridget O'Be
Diogo B. Provete^{14,15}  | Orly Razgour¹⁶  | Loren Rieseberg¹⁷  | Cynthia
Luca Santini¹⁹  | Benjamin Sibbett¹ | Pedro R. Peres-Neto²⁰ 

¹John Wiley & Sons Ltd, Oxford, UK

²School of Biosciences, Ecology and Evolutionary Biology, University of Sheffield, Sheffield, UK

³CNRS, AgroParisTech, Ecologie Systématique Evolution, Université Paris-Saclay, Orsay, France

⁴Department of Zoology, Faculty of Sciences, University of Granada, Granada, Spain

⁵Harvard University Herbaria, Cambridge, Massachusetts, USA

⁶Sound Solutions for Sustainable Science, Boston, Massachusetts, USA

⁷Wiley-VCH GmbH, Weinheim, Germany

⁸Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, Florida, USA

⁹York University, Toronto, Ontario, Canada

¹⁰John Wiley & Sons Ltd, Shanghai, China

¹¹Department of Entomology, University of Georgia, Athens, Georgia, USA

¹²University of Idaho, Hagerman, Idaho, USA

¹³Nordic Society Oikos, Lund University, Lund, Sweden

“Too often open data are uploaded piecemeal, with no accompanying metadata or missing context on processing that happened before data deposition.

As a result, **potential for their reuse in either replicated studies, or in metaanalyses ... or their use in generating novel results ... is reduced.**”

Make and sustain connections

Publishers, researchers, institutions, and data repositories globally are collaborating to advance open science



DRYAD

Not-for-profit and online since 2008. Learn more: datadryad.org



Photo by Erwan Hesry on Unsplash

A data sharing case study


As a researcher, my study uses **genomic data** in combination with **landscape, dispersal, and occupancy data**, to inform [conservation unit] CU delineation in Nevada populations of the Great Basin Distinct Population Segment of the Columbia spotted frog (*Rana luteiventris*).

Forester, Brenna et al. (2022), Genomics-informed delineation of conservation units in a desert amphibian, Dryad, Dataset, <https://doi.org/10.5061/dryad.w6m905qqn>



Not-for-profit and online since
2008. Learn more: datadryad.org

Raw sequencing
data



National Library of Medicine
National Center for Biotechnology Information

BioProject

A BioProject is a collection of biological data related to a single initiative, originating from a single organization or from a consortium. A BioProject record provides users a single place to find links to the diverse data types generated for that project.

Filtered data
(Variant Call
Format [VCF] and
metadata (text,
TSV)



DRYAD

Not-for-profit and online since
2008. Learn more: datadryad.org

Build connections and provide context

- Indicate use of multiple repositories in DMSP
- Add persistent identifiers (PIDs) for related datasets (and other outputs)
- Describe how and why data has been divided between repositories
- Strategically apply metadata

Indicate use of multiple repositories in DMSP

Element 4: Data Preservation, Access, and Associated Timelines

A. Repository where scientific data and metadata will be archived:

Raw, demultiplexed sequencing data will be made available on the NCBI Sequence Read Archive.

Filtered VCFs and metadata will be made available via Dryad.

Add persistent identifiers (PIDs)

Genomics-informed delineation of conservation units in a desert amphibian

Forester, Brenna, Colorado State University,  <https://orcid.org/0000-0002-1608-1904>

Murphy, Melanie, University of Wyoming

Mellison, Chad, United States Fish and Wildlife Service

Petersen, Jeffrey, Nevada Department of Wildlife

Pilliod, David, United States Geological Survey

Van Horne, Rachel, US Forest Service

Harvey, Jim, US Forest Service

Funk, W. Chris, Colorado State University,  <https://orcid.org/0000-0002-6466-3618>

brenna.forester@colostate.edu, chris.funk@colostate.edu


Publication date: August 25, 2022

Publisher: Dryad

<https://doi.org/10.5061/dryad.w6m905qqn>



Data files

 Download dataset

> August 25, 2022

** changes not displayed to the public*

Related works

Article
<https://doi.org/10.1111/mec.16660>

Dataset
<https://www.ncbi.nlm....oproject/PRJNA869693>



ORIGINAL ARTICLE |  Open Access |  

Genomics-informed delineation of conservation units in a desert amphibian

Brenna R. Forester  Melanie Murphy, Chad Mellison, Jeffrey Petersen, David S. Pilliod, Rachel Van Horne, Jim Harvey, W. Chris Funk

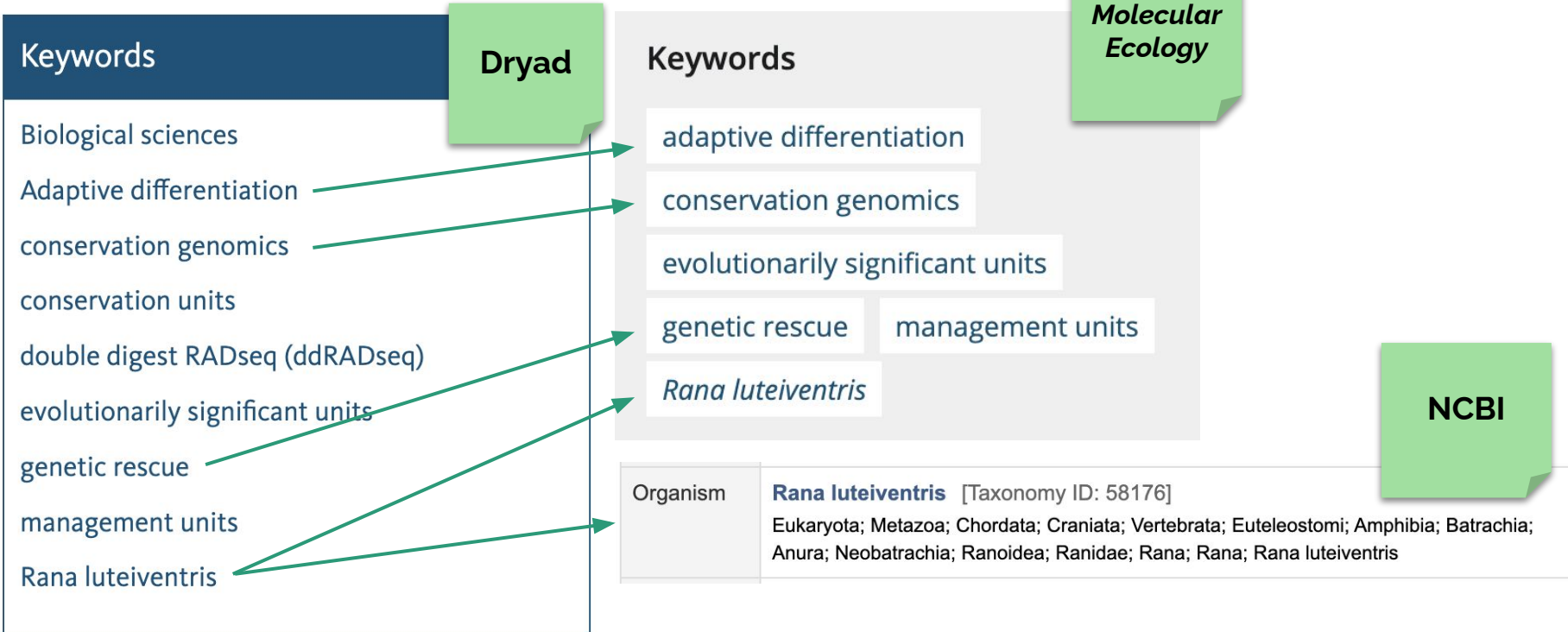
First published: 17 August 2022 | <https://doi.org/10.1111/mec.16660> | Citations: 1

Open Research

DATA AVAILABILITY STATEMENT

Raw, demultiplexed sequencing data are available on the NCBI Sequence Read Archive under BioProject PRJNA869693: <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA869693> . Filtered VCFs and metadata are available on Dryad (Forester et al., [2022](#)): <https://doi.org/10.5061/dryad.w6m905qqn> .

Strategically apply metadata





ASHRAE global database of thermal comfort field measurements

Parkinson, Thomas, University of Sydney, <https://orcid.org/0000-0002-0088-8754>
Tartarini, Federico, Berkeley Education Alliance for Research in Singapore

Data files

Download dataset



Characterizing and through microRNA

Nanayakkara, Jina, Queen's Unive
 Yang, Xiaojing, Queen's University
 Tyryshkin, Kathrin, Queen's Unive
 Wong, Justin J.M., Queen's Univer
 Vanderbeck, Kaitlin, Queen's Univ
 Ginter, Paula S., Weill Cornell Med
 Scognamiglio, Theresa, Weill Corn

Works Referencing This Dataset

- Panarelli, Nicole et al. (2019), Evaluating gastroenteropancreatic neuroendocrine tumors through microRNA sequencing, Endocrine-Related Cancer, Journal-article, <https://doi.org/10.1530/erc-18-0244>
- Renwick, Neil et al. (2013), Multicolor microRNA FISH effectively differentiates tumor types, Journal of Clinical Investigation, Journal-article, <https://doi.org/10.1172/jci68760>
- Cheung, Irene Y. et al. (2014), Deep MicroRNA sequencing reveals downregulation of miR-29a in neuroblastoma central nervous system metastasis, Genes, Chromosomes and Cancer, Journal-article, <https://doi.org/10.1002/gcc.22189>



Search

help | My Datasets | Logout

Download dataset

2020

not displayed to the public

works

DATA COUNT



Citation

Parkinson, Thomas et al. (2022), ASHRAE global database of thermal comfort field measurements, Dryad, Dataset, <https://doi.org/10.6078/D1F671>

9916 views

2108 downloads

3 citations



These findings also have some implications for expression programs, many retaining morphologic and functional similarities. Using machine learning approaches, we identified 17 miRNAs to discriminate 15 NEN pathological types and subsequently constructed a multi-layer classifier, correctly identifying 217 (98%) of 221 samples and overturning one histologic diagnosis. Through our research, we have identified common and type-specific miRNA tissue markers and constructed an accurate miRNA-based classifier, advancing our understanding of NEN diversity.

Methods

Sequencing-based miRNA expression profiles from 378 clinical samples, comprising 239 neuroendocrine neoplasm (NEN) cases and 139 site-matched non-NEN controls, were used in this study. Expression profiles were either compiled from published studies (n=149) or generated through small RNA sequencing (n=229). Prior to sequencing, total RNA was isolated from formalin-fixed paraffin-embedded (FFPE) tissue blocks or fresh-frozen (FF) tissue samples. Small RNA cDNA libraries were sequenced on HiSeq 2500 Illumina platforms using an established small RNA sequencing (Hafner et al., 2012 *Methods*) and sequence annotation pipeline (Brown et al., 2013 *Front Genet*) to generate miRNA expression profiles. Scaling our existing approach to miRNA-based NEN classification (Panarelli et al., 2019 *Endocr Relat Cancer*; Ren et al., 2017 *Oncotarget*), we constructed and cross-validated a multi-layer classifier for discriminating NEN pathological types based on selected miRNAs.

Usage Notes

Diagnostic histopathology and small RNA cDNA library preparation information for all samples are presented in Table S1 of the associated manuscript.

Funding

Academic Health Sciences Center Alternative Funding Plan Innovation Fund

Canada Foundation for Innovation John R Evans Leaders Fund

Carcinoid and Neuroendocrine Tumor Society Canada

Ontario Research Fund-Research Infrastructure

Southeastern Ontario Academic Medical Organization

Robertson Therapeutic Development*

Rockefeller University, Award: UL1TR001866

Ontario Institute for Cancer Research

Robertson Therapeutic Development

License

This work is licensed under a [CC0 1.0 Universal \(CC0 1.0\) Public Domain Dedication license](https://creativecommons.org/licenses/by/4.0/).

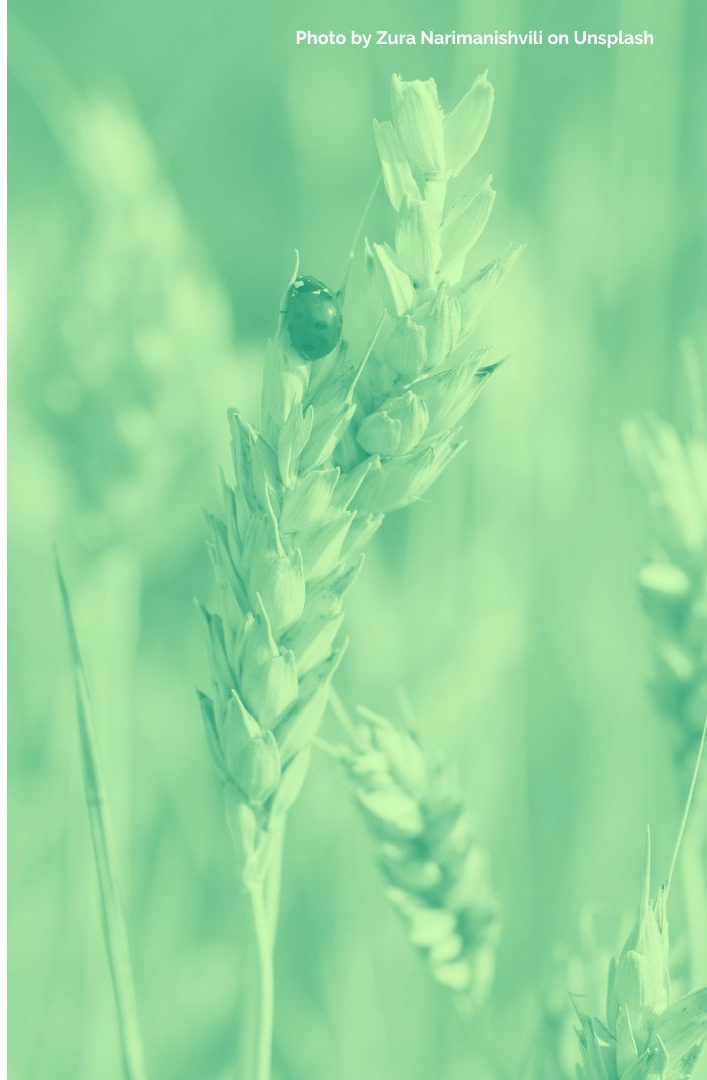


Room to grow

- Support discipline-specific metadata
- Strengthen connections with discipline-specific repositories
- Support author “wayfinding”
- What else?



Not-for-profit and online since
2008. Learn more: datadryad.org



Thank you!

hello@datadryad.org

sarah@datadryad.org

www.datadryad.org

