

CDL's Path Towards Data Publishing Adoption: Community Infrastructure

John Chodacki,
UC Curation Center (UC3)



University of California

CDL

California Digital Library

Looking Inward: Self Assessment



University of California

CDL

California Digital Library

Last year...

Where's the adoption? Shifting the Focus of Data Publishing in 2018



UC Curation Center [Follow](#)

Dec 18, 2017 · 6 min read

By [Daniella Lowenberg](#)

Institutions build or buy general data repositories (a.k.a. data IRs, non-domain data repositories, campus data solutions)

What we learned



Tools are not researcher centric



Focus on tools has distracted community from our adoption efforts



Researchers do not think at the institutional level



We have to meet researchers where they are at within their workflows



Evaluating Our Success

- × # of Deposits
- × Awareness in the researcher community
- × Integration into workflows
- × Simplified story of value proposition



University of California

CDL

California Digital Library

Meanwhile....

1. Commercialization of this space continues
2. Research data services/investments are unaffordable with library budgets
3. Researchers have continued to deposit where publishers refer them
4. Our focus, rightly so, has been on community infrastructure for articles

Community Infrastructure



University of California

CDL

California Digital Library

Principles for Open Scholarly Infrastructures

Governance



Sustainability

Insurance



Supporter Community Values

6. Respect multiple solutions
7. Stick to your scope
8. Leverage communal wisdom to move quickly
9. Encourage healthy skepticism
10. Collaborate and be stronger together

Full book available: supporters.guide - [@supportersguide](https://twitter.com/supportersguide)



**Effective support to the research
community does not always mean
building new or more tools**



University of California

CDL

California Digital Library

**Looking
Outward:
Community
Success**



University of California

CDL

California Digital Library

Dryad

- Researcher supported and adopted
- Curation for FAIR data & compliance on every dataset
- Embedded in researchers workflows
- Open source & non-profit



University of California

CDL

California Digital Library

Dryad: leader in open data publishing

90,000 researchers

24,500 data submissions

2100 international institutions

700+ academic journals represented

Curation checks on every submission



University of California

CDL

California Digital Library

PUBLIC RELEASE: 30-MAY-2018

Advancing data publishing: California Digital Library & Dryad announce partnership

CALIFORNIA DIGITAL LIBRARY (UNIVERSITY OF CALIFORNIA, OFFICE OF THE PRESIDENT)



SHARE



PRINT



E-MAIL

OAKLAND/May 30, 2018 - University of California's California Digital Library (CDL) and Dryad Digital Repository are formally partnering to address researcher needs and lead an open, community-supported initiative in research data curation and publishing. This partnership is aimed at driving adoption of curated, accessible data publishing in the research community and will leverage the capabilities of both institutions and publishers to better align data publishing within researcher workflows.

We Are Launching a New Dryad Service!

- ★ New product development team
- ★ Migrating classic Dryad onto open-source, nimble CDL technology
- ★ Transparent reporting and curation with administrative layer
- ★ Enhanced submission features: publisher integrations



We are launching a new community model

- Dryad is building an **institutional membership** community
- Dryad will be supporting a new set of institutional features
- **Institutions can support curated research data publishing where researchers are already depositing**



Our Values



We are researcher centered



University of California
CDL
California Digital Library

We are adoption focused



University of California

CDL

California Digital Library

Call to Action



University of California
CDL
California Digital Library

We are an **institutional community** who support data curation, data publishing, and data preservation adoption. We should effectively **band together** and support **community-owned** research data infrastructure!



University of California
CDL
California Digital Library

It is **not** about technology. It is about meeting researchers where they are at and building **global, sustainable, centralized** approaches to achieve **adoption.**



University of California
CDL
California Digital Library



**Values translated
to features**



New Dryad Platform

- ❑ Leverages a Core Trust Seal Certified Preservation Repository
- ❑ Technology is standards-based:
 - ❑ SWORD, OAI-PMH, Schema.org
 - ❑ DataCite schema, ORCID login/co-author ORCIDs, Funder Registry, Versioning
- ❑ Large datasets accepted via cloud manifest
- ❑ Submission and Download APIs
- ❑ Administration and curation layer
- ❑ Standardized data usage and citation metrics (Make Data Count)



Dryad Curation

- Metadata and file checking on every submission
- Team of expert curation staff
- Plugging into institutional data curators
- Modified checks and policies enforced for publishers



Seamless Deposits



University of California

CDL

California Digital Library

Publisher Integrations

- Deposit of data at same time as article
- Researcher never has to leave article submission system
- Privacy during peer review



Transparency & Reporting



University of California

CDL

California Digital Library

Admin Dashboard

At a glance

38,469 Users
87 Datasets

Activity in the last 7 days

0 users added
0 datasets started
6 datasets submitted

Datasets

[Clear Search](#)

Filter by:




[Reset all filters](#)

Title	Status	Author	DOI	Last Modified	Last modified by	Size	Publication Date
Data from: Demographic inferences after a range expansion can be biased: the test case of the blacktip reef shark (<i>Carcharhinus melanopterus</i>)	Unsubmitted	Daisie Huang	10.5061/dryad.553cm8g	12/04/2018 09:47:24			My last modified
Data from: Distinct activity-gated pathways mediate attraction and aversion to CO2 in <i>Drosophila</i>	Published	Daisie Huang	10.5061/dryad.2s8422f	12/03/2018 14:39:58			My last modified

Activity Log for Data from: Post-Cretaceous bursts of evolution along the benthic-pelagic axis in marine fishes

doi:10.5061/dryad.9bq5n41

Add Note

Timestamp ↓ ¹ / ₂	Status	Action taken by	Notes	Keywords
11/19/2018 16:00:00	Submitted		Submitted by Ricardo Betancur-R. (betanri@gmail.com) on 2018-11-20T17:25:31Z workflow start=Step: requiresReviewStep - action:noUserSelectionAction No. of bitstreams: 0	
11/20/2018 16:00:00	Curation		Approved by ApproveRejectReviewItem based on metadata for RSPB-2018-2010 on 2018-11-21T09:20:02Z (GMT) article title was updated from "Data from: Post-Cretaceous bursts of evolution along the benthic-pelagic axis in marine fishes".	
11/20/2018 16:00:00	Published		Made available in DSpace on 2018-11-21T21:27:12Z (GMT). No. of bitstreams: 0	
12/03/2018 14:17:54	Status Unchanged		Step: dryadAcceptEditReject - action:dryadAcceptEditRejectAction Approved for entry into archive by Rich Yaxley (ryaxley@datadryad.org) on 2018-11-21T21:27:09Z (GMT)	

Discoverability & Usability



University of California
CDL
California Digital Library

Data from: Neurospora and the dead-end hypothesis: genomic consequences of selfing in the model genus.

Gioti, Anastasia

Stajich, Jason E.

Johannesson, Hanna

Publication date: October 24, 2018

Publisher: Dryad

<https://doi.org/10.5061/dryad.4n9b4>

Citation

Gioti, Anastasia; Stajich, Jason E.; Johannesson, Hanna (2018), Data from: Neurospora and the dead-end hypothesis: genomic consequences of selfing in the model genus., Dryad Dash, Dataset, <https://doi.org/10.5061/dryad.4n9b4>



Download the dataset ~ 0 B



Download Data Publication (PDF)

Metrics



0 views



0 downloads



[1 citations](#)

Keywords

Data from: Neurospora and the dead-end hypothesis: genomic consequences of selfing in the model genus.



Download the dataset ~ 0 B



Download Data Publication (PDF)

Gioti, Anastasia

Stajich, Jason E.

Johann

Public

Public

https:

Cita

Gioti, Anastasia; Stajich, Jason E.; Johannesson, Hanna (2018), Data from: Neurospora and the dead-end hypothesis: genomic consequences of selfing in the model genus., Dryad Dash, Dataset, <https://doi.org/10.5061/dryad.4n9b4>

Abstract

It is becoming increasingly evident that adoption of different reproductive strategies, such as sexual selfing and asexuality, greatly impacts genome evolution. In this study, we test theoretical predictions on genomic maladaptation of selfing lineages using empirical data from the model fungus *Neurospora*. We sequenced the genomes of four species representing distinct transitions to selfing within the history of the genus, as well as the transcriptome of one of these, and compared with available data from three outcrossing species. Our results provide evidence for a relaxation of purifying selection in protein-coding genes and for a reduced efficiency of transposable element silencing by Repeat

Metrics



Works Referencing This Dataset

Gioti, Anastasia; Stajich, Jason E.; Johannesson, Hanna (2013), NEUROSPORA AND THE DEAD-END HYPOTHESIS: GENOMIC CONSEQUENCES OF SELFING IN THE MODEL GENUS, Wiley, Article-journal, <https://doi.org/10.1111/evo.12206>

Keywords

Mating Systems

molecular evolution

Mutations

natural selection

Citation

Gioti, Anastasia; Stajich, Jason E.; Johannesson, Hanna (2018), Data from: Neurospora and the dead-end hypothesis: genomic consequences of selfing in the model genus., Dryad Dash, Dataset, <https://doi.org/10.5061/dryad.4n9b4>

Abstract

It is becoming increasingly evident that adoption of different reproductive strategies, such as sexual selfing and asexuality, greatly impacts genome evolution. In this study, we test theoretical predictions on genomic maladaptation of selfing lineages using empirical data from the model fungus *Neurospora*. We sequenced the genomes of four species representing distinct transitions to selfing within the history of the genus, as well as the transcriptome of one of these, and compared with available data from three outcrossing species. Our results provide evidence for a relaxation of purifying selection in protein-coding genes and for a reduced efficiency of transposable element silencing by Repeat Induced Point mutation. A reduction in adaptive evolution was also identified in the form of reduced codon usage bias in highly expressed genes of selfing *Neurospora*, but this result may be confounded by mutational bias. Potentially counteracting these negative effects, the nucleotide substitution rate and the spread of transposons is reduced in selfing species. We suggest that differences in substitution rate relate to the absence of the asexual pathway producing conidia in selfing *Neurospora*. Our results support the dead-end theory and show that *Neurospora* genomes bear signatures of both sexual and asexual reproductive mode.

References

This dataset is cited by <https://doi.org/10.1111/evo.12206>



[1 citations](#)

Keywords

Mating Systems

molecular evolution

Mutations

natural selection

sequence alignments

selfing

outcrossing

License

This work is licensed under a [CC0 1.0 Universal \(CC0 1.0\) Public Domain Dedication](#) license.

Compliance



University of California
CDL
California Digital Library

Compliance

- FAIR Data: Curation
- Funder Requirements:
 - CoreTrustSeal, FAIR Data, Preservation
- Publisher Policies:
 - Private during peer review, published alongside article, citations
- **Institutional Values: [Your Input Here!]**



What sets Dryad apart?

- Focus: ease of deposit and user centered
- Goal: drive adoption of **curated, compliant**, research data publishing
- Features: innovations to uphold our goals & values

