4 MORE Ways to Make Papers a Little More Open and Reproducible

Back in 2015 I wrote a post on Ways to Make Papers a Little More Open and Reproducible – Roberts Lab. I thought it would be a fun exercise to revisit this 5 years later and see if any improvements have been made.

Spencer, LH, Venkataraman, YR, Crim R, Ryan S, Horwith MJ, and Roberts SB (2020) Carryover effects of temperature and pCO2 across multiple Olympia oyster populations. Ecological

Case 1

Applications 00(00):e02060. doi:10.1002/eap.2060 First off I am proud to say there is a PrePrint in bioRxiv. **Open Research**

Data are available on Figshare: https://doi.org/10.6084/m9.figshare.8872646

Data Availability

```
Supporting Information
 Filename
                                                                               Description
  eap2060-sup-0001-AppendixS1.pdf PDF document, 2.4 MB
Please note: The publisher is not responsible for the content or functionality of any supporting
the corresponding author for the article.
```

information supplied by the authors. Any queries (other than missing content) should be directed to The Supporting Information is rich with recipes and detailed larval rearing info. On FigShare there is an archived version of the associated GitHub Repo. Download (1.55 GB) Cite + Collect (you need to log in first) GitHub repo for Spencer et al. 2019, Oly carryover effects 141 downloads views Version 3 ✓ Dataset posted on 16.07.2019, 14:12 by Laura H. Spencer, Yaamini Venkataraman, Steven Roberts, Micah J Horwith, Ryan Crim, Stuart Ryan CATEGORIES

analyses data images Inotebooks <u>ipynb_checkpoints</u> 5.1 kB Locus_ids.ipynb VCF_readcounts.ipynb 23.1 kB ipyrad_assembly.ipynb 47.4 kB ipyrad_log.txt 0 Bytes ipyrad_script.ipynb 5.5 kB locus_id.txt 0 Bytes readme.md 364 Bytes readme.md 4.7 kB scripts Rapp.history 0 Bytes readme.md 38 Bytes Zenodo links us out to GitHub https://github.com/jldimond/P.ast-transplant/tree/v2.0.0 and OpenAire SEARCH DEPOSIT LINK CONTENT PROVIDERS **XPLORE** jldimond/P.ast-transplant: Porites astreoides 🔑 LINK THIS SOFTWARE TO... transplant experiment

SOFTWARE

Jay Dimond;

Publisher: Zenodo

(2019)

Subject:

data 🖿

images

notebooks

notebooks"

Build and launch a repository

https://github.com/jldimond/P.ast-transplant

https://mybinder.org/v2/gh/jldimond/P.ast-transplant/master

Copy the URL below and share your Binder with others:

GitHub repository name or URL

Git branch, tag, or commit

Git branch, tag, or commit

GitHub →

or could use nbviewer.

jupyter

OPEN SOURCE

Related identifiers: doi: 10.5281/zenodo.3576526

mesheuropmc: fungi | social sciences | technology, industry, and

This is a repository detailing an experiment in which colonies of the coral

Updated analysis and R script

https://github.com/jldimond/P.ast-transplant/blob/v2.0.0/readme.md

GitHub has a great Readme with software needed and directory structure explained.

agriculture | population characteristics | geographic locations

Update directory

Update readmes

Porites furcata Porites divaricata

The goal is to evaluate (1) if there is any genetic structuring of these individuals using the ddRAD data and (2) if there is

What is neat here is you can go to mybinder.org and "Turn a Git repo into a collection of interactive

8 binder

Turn a Git repo into a collection of interactive

notebooks

Have a repository full of Jupyter notebooks? With Binder, open those notebooks in an executable

environment, making your code immediately reproducible by anyone, anywhere.

Path to a notebook file (optional)

Path to a notebook file (optional)

File ▼

FAQ

JUPYTER

SIGN IN 🔍

6

?

\$

3 years ago

Share - Bookmark

Download from

Cite this software

select a citation style

2019)

☑ Zenodo

Zenodo via Zenodo (Software,

form of "Base Counts (CATG)". This information is used in the analysis of the EpiRADseq data. !head data2.vcf In [1]: ##fileformat=VCFv4.0 ##fileDate=2017/04/03 ##source=ipyrad v.0.5.15 ##reference=GCA_001939145.1_ASM193914v1_genomic.fna ##phasing=unphased ##INFO=<ID=DP,Number=1,Type=Integer,Description="Total Depth"> ##FORMAT=<ID=GT, Number=1, Type=String, Description="Genotype"> ##FORMAT=<ID=DP, Number=1, Type=Integer, Description="Read Depth"> ##FORMAT=<ID=CATG, Number=1, Type=String, Description="Base Counts (CATG)"> In [12]: !vcftools --vcf data2.vcf --extract-FORMAT-info DP VCFtools - 0.1.15 (C) Adam Auton and Anthony Marcketta 2009 Parameters as interpreted: --vcf data2.vcf --extract-FORMAT-info DP

After filtering, kept 96 out of 96 Individuals

After filtering, kept 51977 out of a possible 51977 Sites

Actions

Watch

▼

ılı Insights

Blame

Security

Wiki

Projects 0

This notebook performs BLAST searches of ddRADseq reads against a custom

esearch -db nuccore -query "Symbiodinium [ORGN] AND Symbiodinium sp. clade A [ORGN]" | efetch

Symbiodinium database. The goal of this analysis is to identify symbiont

| esearch -db nuccore -query "Symbiodinium [ORGN] AND Symbiodinium sp. clade A [ORGN] efetch -format fasta > /Volumes/toaster/jdimond/Porites-astreoides-RAD/data/SymCladeA.fasta

lesearch -db nuccore -query "Symbiodinium [ORGN] AND Symbiodinium sp. clade B [ORGN]" efetch -format fasta > /Volumes/toaster/jdimond/Porites-astreoides-RAD/data/SymCladeB.fasta

esearch -db nuccore -query "Symbiodinium [ORGN] AND Symbiodinium sp. clade A [ORGN]" |

NCBI EDirect installed according to instructions (https://www.ncbi.nlm.nih.gov/books/NBK179288/#chapter6.Getting_Started)

★ Star

Find file

History

9c87c2f on Nov 12, 2019

Y Fork

Copy path

Outputting FORMAT information for DP

Run Time = 3.00 seconds

Pull requests 0

Branch: master - P.ast-transplant / notebooks / SymbiodiniumBLAST.ipynb

populations in each coral at the clade level.

File "<ipython-input-15-3af491a48041>", line 1

File "<ipython-input-16-012118ea55dc>", line 1

In [16]: #Extract Symbiodinium clade B records from NCBI nucleotide database

I am a phd student at the University of Glasgow, UK, and

my work focuses on genetic and epigenetic variation in

wild populations of salmonids. I am writing you because

In [15]: #Extract Symbiodinium clade A records from NCBI nucleotide database using EDirect

In [13]: !head out.DP.FORMAT

GitHub also renders ipynb files.

jldimond / P.ast-transplant

! Issues 0

595 lines (595 sloc) | 18.3 KB

-format fasta

And people do care.

Dear Dr. Dimond,

SyntaxError: invalid syntax

SyntaxError: invalid syntax

and of course folks could launch locally in Jupyter.

ildimond Updated jupyter notebooks and R analysis

<> Code

1 contributor

interested, and I'm happy to hear any comments you might

This work was funded by Univ. of Glasgow College of Medical, Veterinary and Life Sciences

making their data public so we could explore it for this paper. We thank Glasgow Polyomics

whitefish scaffolds, we thank L Bernatchez, C Rougeux, S Pavey, E Normandeau, S Lien, and

doctoral training programme. We thank JL Dimond, SK Gamblewood, and SB Roberts for

and J Galbraith for sequencing, M Capstick for support in the laboratory, and A Jacobs for

analysis advice and comments on the draft manuscript. For access to unpublished lake

eastern oyster (Crassostrea virginica)

see also https://osf.io/3xf6m/

have.

Acknowledgements

Case 4 Geoduck Genome Paper (Big Files)

Yaamini R. Venkataraman, Alan M. Downey-Wall, Justin Ries, Isaac Westfield, Samuel J. White, Steven B. Roberts, Kathleen E. Lotterhos bioRxiv 2020.01.07.897934; doi: https://doi.org/10.1101/2020.01.07.897934 https://github.com/epigeneticstoocean/paper-gonad-meth My Quick Files My Projects Support 🝘 Steven Roberts 🕶 Search Panopea_generosa_genome Files Wiki Analytics Registrations Contributors Add-ons Make Private Public ľΟ Panopea_generosa_genome Contributors: Steven Roberts, Sam White Affiliated Institutions: University of Washington Date created: 2019-08-22 02:06 PM | Last Updated: 2020-03-03 09:11 PM Create DOI Category: Project Description: Add a brief description to your project License: Add a license Wiki GCitation Add important information, links, or images here to describe your project. Link Projects Components Add Component Files Add components to organize your project. **Z** Click on a storage provider or drag and drop to upload Q Filter i Tags Modified ^ ~ Name ^ v Add a tag to enhance discoverability Panopea_generosa_genome

citations 13 Marine Biology GitHub repository for manuscript, Carry-over effects of temperature and pCO2 across Ecological Physiology multiple Olympia oyster populations. Global Change Biology Invertebrate Biology To view in GitHub: https://github.com/RobertsLab/Paper-oly-adult-oa KEYWORD(S) Ostrea lurida Abstract reproduction phenology Impacts of adult exposure to elevated winter temperature and altered carbonate Intergenerational effects chemistry on reproduction and offspring viability were examined in the Olympia oyster transgenerational exposure (Ostrea lurida). Three distinct populations of adult, hatchery-reared O. lurida, plus an climate change additional cohort spawned from one of the populations, were sequentially exposed to That is citable Spencer, Laura H.; Venkataraman, Yaamini; Roberts, Steven; Horwith, Micah J; Crim, Ryan; Ryan, Stuart (2019): GitHub repo for Spencer et al. 2019, Oly carryover eff ects. figshare. Dataset. https://doi.org/10.6084/m9.figshare.8872646.v3 and in fact has been cited 3 times! 10.6084/m9.figshare.8872646 X Dimensions Support Register **FILTERS PUBLICATIONS** DATASETS ANALYTICAL VIE **FAVORITES** GRANTS PATENTS 0 PUBLICATION YEAR CLINICAL TRIALS POLICY DOCUMENTS FIELDS OF RESEAF O 2020 06 Biological Sciences 2019 Sort by: Relevance > 0602 Ecology Title, Author(s), Bibliographic reference - About the metrics > RESEARCHER Carryover effects of temperature and pCO2 across OVERVIEW multiple Olympia oyster populations FIELDS OF RESEA. Laura H. Spencer, Yaamini R. Venkataraman, Ryan Crim, Stuart Ryan, ... 2020, Ecological Applications - Article Citations Citations (M > PUBLICATION TYPE 0.50 Altmetric 12 ≡+ Add to Library SOURCE TITLE Carryover effects of temperature and pCO2 across > JOURNAL LIST multiple Olympia oyster populations Laura H Spencer, Yaamini R Venkataraman, Ryan Crim, Stuart Ryan, ... > OPEN ACCESS 2019, bioRxiv - Preprint 2011 2013 2015 2017 Citations 1 Altmetric 2 Open Access Publications (total) RESEARCHERS Micah J Horwith Washington Department of Na While you can download files from Figshare, Laura also provided a url to GitHub: https://github.com/RobertsLab/Paper-oly-adult-oa RobertsLab / Paper-oly-adult-oa O Unwatch ▼ <> Code Wiki Wiki ! Issues 0 1 Pull requests 0 Actions Projects 0 Security ılı Insights Settings For files associated with the Olympia oyster adult OA exposure paper Edit Manage topics 6 commits 1 branch TO packages O releases 1 contributor Branch: master -New pull request Create new file Upload files Find file Clone or download ▼ [Displaying laura spencer and laura spencer revised plots, tested env. data on juvenile survival Latest commit 6c4969b on Sep 25, 2019 .Rproj.user revised plots, tested env. data on juvenile survival 5 months ago revised plots, tested env. data on juvenile survival Analyses 5 months ago revised plots, tested env. data on juvenile survival 5 months ago Data revised plots, tested env. data on juvenile survival 5 months ago Figures changed a few names, added .gitignore for DS_Store files 8 months ago Images .DS_Store revised plots, tested env. data on juvenile survival 5 months ago .RData revised plots, tested env. data on juvenile survival 5 months ago .Rhistory revised plots, tested env. data on juvenile survival 5 months ago gitignore.txt changed a few names, added .gitignore for DS_Store files 8 months ago Paper-oly-adult-oa.Rproj created new clean directory to publish along with manuscript 8 months ago README.md Create README.md 8 months ago Granted there are some unneeded files. But the repo has great structure and using "Project" Now anyone can just Clone into RStudio and this it what it looks like... 🛂 🕶 🌆 📹 🕶 🗐 Paper-oly-adult-oa 📗 👸 🕶 🚟 🕶 Addins 🕶 02_Temperatures.R History Connections Git (= 📦 | 🚈 | 🚃 🔲 Source on Save | 🔍 🏸 🗸 📗 ♣ | ♠ | 🕭 | 🌣 master - C Run Source -Diff Commit Staged Status Path .gitignore HOBO.data.temp <- read.csv("Data/temp-treatment-data.csv", header=TRUE, stringsAsFactors = FALSE)[,c(1:3)] HOBO.data.temp\$Time.Point <- as.POSIXct(HOBO.data.temp\$Time.Point,</pre> format = "%m/%d/%Y %H:%M") plot(x=HOBO.data.temp\$Time.Point, y=HOBO.data.temp\$Cold.temp) plot(x=HOBO.data.temp\$Time.Point, y=HOBO.data.temp\$Ambient.temp) temp.treat <- subset(HOBO.data.temp, Time.Point > "2016-12-21 13:00:00" & Time.Point < "2017-02-04 00:00:00") Files Plots Packages Help Viewer C 🛂 New Folder 👂 Delete 📑 Rename 🥻 More 🗸 plot(x=temp.treat\$Time.Point, y=temp.treat\$Cold.temp) #ambient temp over **®** ... ■ **↑** Home > Documents > GitHub > Paper-oly-adult-oa > Figures plot(x=temp.treat\$Time.Point, y=temp.treat\$Ambient.temp) #chilled temp over broodstock-height.pdf 7.7 KB Mar 4, 2020, 112 broodstock-height.png Mar 4, 2020, 11 155.5 KB cumulative larvae 1.pdf Mar 4, 2020, 112 5.2 KB Console ~/Documents/GitHub/Paper-oly-adult-oa/ 🖈 cumulative_larvae_2.pdf Mar 4, 2020, 11 5.7 KB Platform: x86_64-apple-darwin15.6.0 (64-bit) cumulative_larvae_3.pdf 5.6 KB Mar 4, 2020, 112 cumulative_larvae_4.pdf Mar 4, 2020, 112 6.2 KB R is free software and comes with ABSOLUTELY NO WARRANTY. deployment-survival-Cl.pdf Mar 4, 2020, 112 You are welcome to redistribute it under certain conditions. Type 'license()' or 'licence()' for distribution details. deployment-survival-envdata.pdf 10.5 KB Mar 4, 2020, 112 🤼 deployment–survival–FB.pdf Mar 4, 2020, 112 Natural language support but running in an English locale deployment-survival-legend.pdf Mar 4, 2020, 112 5.5 KB R is a collaborative project with many contributors. deployment-survival-PG.pdf Mar 4, 2020, 11 Type 'contributors()' for more information and deployment-survival-plot-title.pdf 11.5 KB Mar 4, 2020, 112 citation()' on how to cite R or R packages in publications. Case 2 Dimond JL and Roberts SB. (2020) Convergence of DNA Methylation Profiles of the Reef Coral Porites astreoides in a Novel Environment Frontiers in Marine Science, vol 6. doi:10.3389/fmars.2019.00792 Again we see a preprint, even with some nice discussion. bioRxiv bioRxiv @biorxivpreprint Convergence of DNA methylation profiles in a novel environment in the reef coral Porites astreoides https://t.co/b2cbVcm8jz #bioRxiv 28 Aug 2019 Privacy Policy 2 Comments BioRxiv Sort by Newest -Recommend Share Tweet Join the discussion... Mikhail V Matz · 5 months ago Nice one! One concern about Fig. 6: comparison of distance matrices must be based on Mantel test, not regular correlation, since data points are not independent. Should not draw trendline and especially the shaded credible interval. ∧ | ∨ · Reply · Share › James Dimond → Mikhail V Matz · 5 months ago Thank you! Yes, we had a similar comment about a similar figure from a previous paper. I've had conflicting advice on this and will look into it further. ∧ | ∨ · Reply · Share › There is a Data Availability Statement in the Frontiers article Conclusion This work shows that DNA methylation is an environmentally responsive epigenetic process that is reflective of the environment, and is consistent with its putative role in acclimatization. We were able to detect subtle changes in P. astreoides methylation associated with experimental transplantation, as well as evidence for heritability of methylation patterns. Loci responding to transplantation were associated with signaling, apoptosis, gene regulation and epigenetic crosstalk, yet much remains to be EDITED BY learned about the function of methylation changes in these differentially MAIN Andrew S. Mount methylated genes. This study helps set the stage for further work on both the Clemson University, United States functional genomics and molecular ecology of acclimatization processes in reef corals. REVIEWED BY Iliana B. Baums Pennsylvania State University **Data Availability Statement** (PSU), United States Kevin M. Johnson The datasets generated for this study can be found in the NCBI Sequence Louisiana State University, Read Archive, accession number SRP132538. Analysis archive at United States https://doi.org/10.5281/zenodo.3576526. Which sends us to Zenodo https://doi.org/10.5281/zenodo.3576526, another means to archive a GitHub repository, citable as Jay Dimond. (2019, December 15). jldimond/P.ast-transplant: Porites astreoides tra nsplant experiment (Version v2.0.0). Zenodo. http://doi.org/10.5281/zenodo.3576526 Q Upload Communities December 15, 2019 Software Open Access jldimond/P.ast-transplant: Porites views downloads astreoides transplant experiment See more details... Jay Dimond This is a repository detailing an experiment in which colonies of the coral Porites astreoides on the Belize Barrier Reef were transplanted to a common garden for one year and assessed for changes in DNA methylation. Available in Preview GitHub P.ast-transplant-v2.0.0.zip × ildimond-P.ast-transplant-149f269 gitignore 155 Bytes OpenAIRE **Publication date:** December 15, 2019 DOI 10.5281/zenodo.3576526 Related identifiers: Supplement to https://github.com/jldimond/P.asttransplant/tree/v2.0.0 License (for files): C Other (Open)

scripts Updated analysis and R script Revert "New assembly & Update gitignore" .gitignore readme.md Update read III readme.md Branching Porites project repository This is a repository to accompany the manuscript "Genetic and epigenetic insight into morphospecies in a reef coral". The study involves analysis of ddRAD-seq and EpiRAD-seq data for 27 samples of branching Porites spp. corals collected in Belize. There is taxonomic uncertainty whether Caribbean branching Porites spp. comprise 3 different species or a single polymorphic species. The species are distinguished primarily by branch thickness.

any epigenetic structuring of these individuals using the EpiRAD data.

Jay relied more on Jupyter notebooks here as opposed to RStudio.

nbviewer A simple way to share Jupyter Notebooks Enter the location of a Jupyter Notebook to have it rendered here: URL | GitHub username | GitHub username/repo | Gist ID Go! JUPYTER P.ast-transplant / notebooks Extract read counts from .vcf file The following script uses vcftools to extract read count information from the .vcf file output from ipyrad. The read co ##INFO=<ID=NS, Number=1, Type=Integer, Description="Number of Samples With Data">

last week I uploaded a preprint on biorxiv where I tested the possibility of using EpiRADseq for genetic analysis as well, i.e. I compared estimates of genetic diversity and population structure with SNPs obtained from ddRAD and EpiRAD. I sequenced the fish I am working on, but I also used your data from the article "Genetic and epigenetic insight into morphospecies in a reef coral". I wanted to personally thank you for making your data available, as it was really good to have two lines of evidence to support the findings of my work. I also want to than you for making your code available, it was really useful, and I'll try to follow your standards in code-sharing for all my research. Please find below the link to the biorxiv preprint, in case you are

Case 3 General DNA methylation patterns and environmentally-induced differential methylation in the

T Nome. We declare no conflict of interest.

https://osf.io/yem8n/ OSF**HOME** ▼

 OSF Storage (United States) + annotation_resources + archive Recent Activity + bam_files