microPublication Biology and undergraduate research

Large CURE Project Results

BREWMOR 2022

Classroom CURE project data are microPublishable

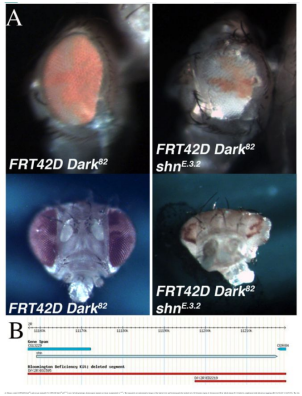
- Project is authentic hypothesis-driven research
- Outcome is unknown
- Whole classrooms of students address the research question
- Data are valuable to the research community

Good data should be discoverable like any other research data

All Submissions are Peer-reviewed

1. **Managing Editor** looks at article and sends it to a **Science Officer***
   **Science Officer is a known and respected research community member*
2. **Reviewer** invited
3. **Database Curator** invited to check community nomenclature use and data reporting standards
4. **Authors** submit revision addressing Reviewer and Curator comments
5. **Science Officer** makes the final decision

Two examples of Classroom CURE projects



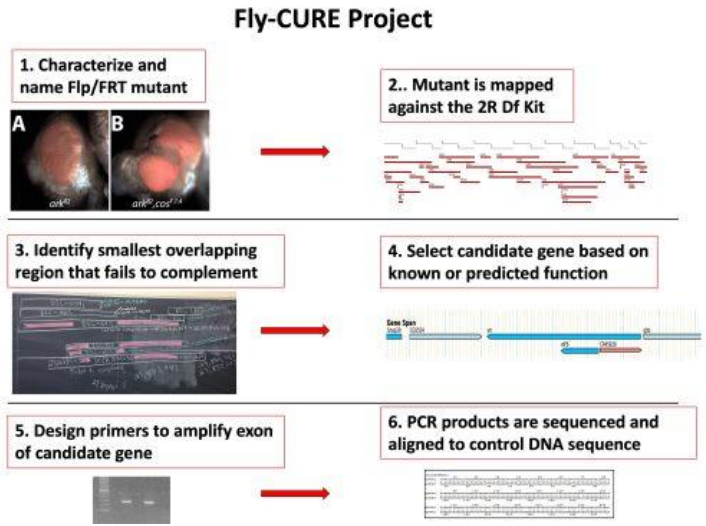
Fly-CURE mutant characterization articles

Genome Education Partnership (GEP)
Gene Model Annotation of Drosophila species

Fly-CURE

Started by Jacob Kagey (U Detroit, Mercy)

Project: Throughout a semester students work on a novel *Drosophila* mutant identified from a screen for conditional regulators of developmental signaling, cell growth control, and cell division (Kagey, Brown, & Moberg, 2012).



16 institutions (PUIs, MSIs, *CC*(1))

Over 500 undergraduate students participated

14 novel Drosophila mutants characterized

Jacob Kagey 3/22 https://thenode.biologists.com/fly-cure-a-nation-wide-cure-initiative/education/

Fly-CURE article workflow

1. Mapping Data generated in duplicate/triplicate by independent groups of UG

2. Article is drafted once the mutant is fully mapped to gene

    -Initially, articles drafted by PIs

    -Most recent 3-4 articles written by UGs who volunteered to draft the first version of the microPublication; these students share first authorship on the article

3. All eligible co-authors read, edit, and approve

4. Manuscript is submitted to microPublication Biology by a PI - goes through our normal Peer-review process

Frequency of submission is ~1 per semester

8 submissions, 1 rejection, 7 published, lots of authors

4/26/2019 - *Open Access*

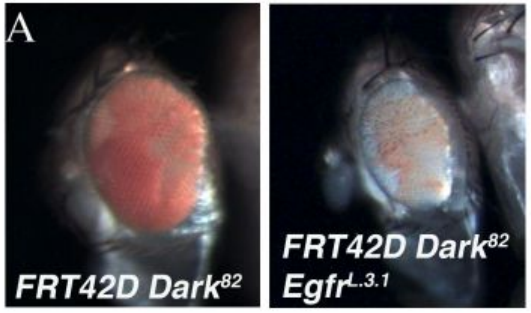
Genetic mapping of EgfrL.3.1 in Drosophila melanogaster

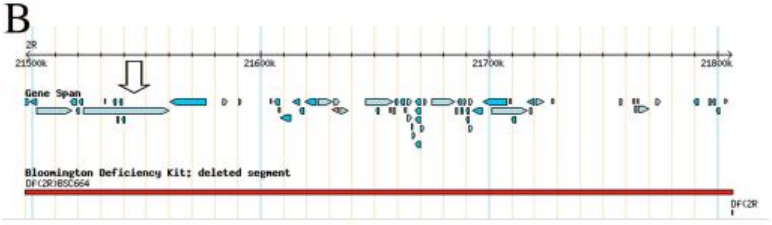
Joyce Stamm[1], Gnanda S Joshi[2], MA Anderson[2], Katie Bussing[1], Colton Houchin[1], Amber C Elinsky[2], Jacob T Flyte[2], Nadine Husseini[2], Dominika Jarosz[2], Chelsea L Johnson[2], Abby F Johnson[2], Christina E Jones[2], Taj P Kooner[2], Daniel Myhre[1], Thomas N Rafaill[2], Sarah Sayed[2], Kirby W Swan[2], Jonathan Toma[2] and Jacob D Kagey[2§]

[1]Department of Biology, University of Evansville

[2]Biology Department, University of Detroit Mercy

[§]To whom correspondence should be addressed: kageyja@udmercy.edu





A. Mosaic (*FRT42D Dark*^{82}), and *Dark*^{82} *Egfr*$^{L.3.1}$ (*FRT42D Dark*^{82}*Egfr*$^{L.3.1}$) eyes. In both eyes the homozygous mutant tissue is pigmented (w^{+mC}). B. Region of chromosome 2R that failed to complement L.3.1 by deficiency mapping (2R:21,497,290..21,806,350). Arrow denotes location of *Egfr* gene. B is adapted from flybase.org (Gramates et al., 2017).

Description

An EMS screen was conducted utilizing the Flp/FRT system to identify mutations that caused an array of phenotypic alterations in the size of the eye including the ratio of mutant to wild type tissue (red over white) or the developmental patterning of the mosaic eye. This screen was done in the genetic background of blocked apoptosis in the homozygous mutant cells to identify conditional regulators of cell growth and eye development (Kagey *et al.*, 2012). The block in apoptosis in the

Author Contributions

Alysia D Vrailas-Mortimer: Data curation, Formal analysis, Investigation, Funding acquisition, Supervision, Validation, Writing - original draft, Writing - review and editing
Nitish Aggarwal: Formal analysis, Data curation, Writing - review and editing, Investigation
Neha N Ahmed: Data curation, Formal analysis, Investigation, Writing - review and editing
Ian M Alberts: Data curation, Formal analysis, Investigation, Writing - review and editing
Masa Alhawasli: Data curation, Formal analysis, Investigation, Writing - review and editing
Ibrahim A Aljerdi: Data curation, Formal analysis, Investigation, Writing - review and editing
Brooke M Allen: Data curation, Formal analysis, Investigation, Writing - review and editing
Ahmad M Alnajar: Data curation, Formal analysis, Investigation, Writing - review and editing
Michael A Anderson: Data curation, Formal analysis, Investigation, Writing - review and editing
Ryan Armstong: Data curation, Formal analysis, Investigation, Writing - review and editing
Chace C Avery: Data curation, Formal analysis, Investigation, Writing - review and editing
Eric J Avila: Data curation, Formal analysis, Investigation, Writing - review and editing
Tyra N Baker: Data curation, Formal analysis, Investigation, Writing - review and editing
Soolmaz Basardeh: Data curation, Formal analysis, Investigation, Writing - review and editing
Natasha A Bates: Data curation, Formal analysis, Investigation, Writing - review and editing
Farrah N Beidas: Data curation, Formal analysis, Investigation, Writing - review and editing
Ashton C Bosler: Data curation, Formal analysis, Investigation, Writing - review and editing
Deja M Brewer: Data curation, Investigation, Formal analysis, Writing - review and editing
Ryan S Buenaventura: Formal analysis, Investigation, Data curation, Writing - review and editing
Natalie JL Burrell: Data curation, Formal analysis, Investigation, Writing - review and editing
Alyn P Cabrera-Lopez: Data curation, Formal analysis, Investigation, Writing - review and editing
Amairani B Cervantes-Gonzalez: Data curation, Formal analysis, Investigation, Writing - review and editing
Raymond P Cezar: Data curation, Formal analysis, Investigation, Writing - review and editing
Joselyn Coronel: Data curation, Formal analysis, Investigation, Writing - review and editing
Corinne Croslyn: Data curation, Formal analysis, Investigation, Writing - review and editing
Konnor R Damery: Data curation, Formal analysis, Investigation, Writing - review and editing
Lucero Diaz-Alavez: Data curation, Formal analysis, Investigation, Writing - review and editing
Nupur P Dixit: Data curation, Formal analysis, Investigation, Writing - review and editing
Dilcia L Duarte: Data curation, Formal analysis, Investigation, Writing - review and editing
Amanda R Emke: Data curation, Formal analysis, Investigation, Writing - review and editing
Katherine English: Data curation, Formal analysis, Investigation, Writing - review and editing
Audrey A Eshun: Data curation, Formal analysis, Investigation, Writing - review and editing
Samuel R Esterly: Data curation, Formal analysis, Investigation, Writing - review and editing
Arnold J Estrada: Data curation, Formal analysis, Investigation, Writing - review and editing
Mark Feng: Data curation, Formal analysis, Investigation, Writing - review and editing
Meghan M Freund: Data curation, Formal analysis, Investigation, Writing - review and editing
Niko Garcia: Data curation, Formal analysis, Investigation, Writing - review and editing
Charandeep S Ghotra: Data curation, Formal analysis, Investigation, Writing - review and editing
Haddia Ghyasi: Data curation, Formal analysis, Investigation, Writing - review and editing
Christa SA Hale: Data curation, Formal analysis, Investigation, Writing - review and editing
Luci Hulsman: Data curation, Formal analysis, Investigation, Writing - review and editing
Lindsey Jamerson: Data curation, Formal analysis, Investigation, Writing - review and editing
Austin K Jones: Data curation, Formal analysis, Investigation, Writing - review and editing
Madeline Kuczynski: Data curation, Formal analysis, Investigation, Writing - review and editing
Takailah N Lacey-Kennedy: Data curation, Formal analysis, Investigation, Writing - review and editing
Mindy J Lee: Data curation, Formal analysis, Investigation, Writing - review and editing
Tala Mahjoub: Data curation, Formal analysis, Investigation, Writing - review and editing
Molly C Mersinger: Data curation, Formal analysis, Investigation, Writing - review and editing
Alyssa D Muckerheide: Data curation, Formal analysis, Investigation, Writing - review and editing
David W Myers: Data curation, Formal analysis, Investigation, Writing - review and editing
Karen Nielsen: Data curation, Formal analysis, Investigation, Writing - review and editing
Paul J Nosowicz: Data curation, Formal analysis, Investigation, Writing - review and editing
Jacinda A Nunez: Data curation, Formal analysis, Investigation, Writing - review and editing
Amy C Ortiz: Data curation, Formal analysis, Investigation, Writing - review and editing
Tulsi T Patel: Data curation, Formal analysis, Investigation, Writing - review and editing
Natalie N Perry: Data curation, Formal analysis, Investigation, Writing - review and editing
W Storm A Poser: Data curation, Formal analysis, Investigation, Writing - review and editing
Deisy M Puga: Data curation, Formal analysis, Investigation, Writing - review and editing
Cathrine Quam: Data curation, Formal analysis, Investigation, Writing - review and editing
Paulina Quintana-Lopez: Data curation, Formal analysis, Investigation, Writing - review and editing
Piper Rennerfeldt: Data curation, Formal analysis, Investigation, Writing - review and editing
Nicholas M Reyes: Data curation, Formal analysis, Investigation, Writing - review and editing
Ian G Rines: Data curation, Formal analysis, Investigation, Writing - review and editing
Cally Roberts: Data curation, Formal analysis, Investigation, Writing - review and editing
Daniel B Robinson: Data curation, Formal analysis, Investigation, Writing - review and editing
Kalina M Rossa: Data curation, Formal analysis, Investigation, Writing - review and editing
Gloria J Ruhlmann: Data curation, Formal analysis, Investigation, Writing - review and editing
Jeremy Schmidt: Data curation, Formal analysis, Investigation, Writing - review and editing
John R Sherwood: Data curation, Formal analysis, Investigation, Writing - review and editing
Dimyana H Shonoda: Data curation, Formal analysis, Investigation, Writing - review and editing
Hannah Soellner: Data curation, Formal analysis, Investigation, Writing - review and editing
Juan C Torrez: Data curation, Formal analysis, Investigation, Writing - review and editing
Mayukah Velide: Data curation, Formal analysis, Investigation, Writing - review and editing
Zachary Weinzapfel: Data curation, Formal analysis, Investigation, Writing - review and editing
Alisa C Ward: Data curation, Formal analysis, Investigation, Writing - review and editing
Kayla L Bieser: Data curation, Formal analysis, Funding acquisition, Investigation, Project administration, Me
Julie A Merkle: Data curation, Formal analysis, Investigation, Project administration, Supervision, Validation,
Joyce C Stamm: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Projec
Richard L Tillett: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, S
Jacob D Kagey: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Metho
review and editing

Reviewed By

Cale Whitworth

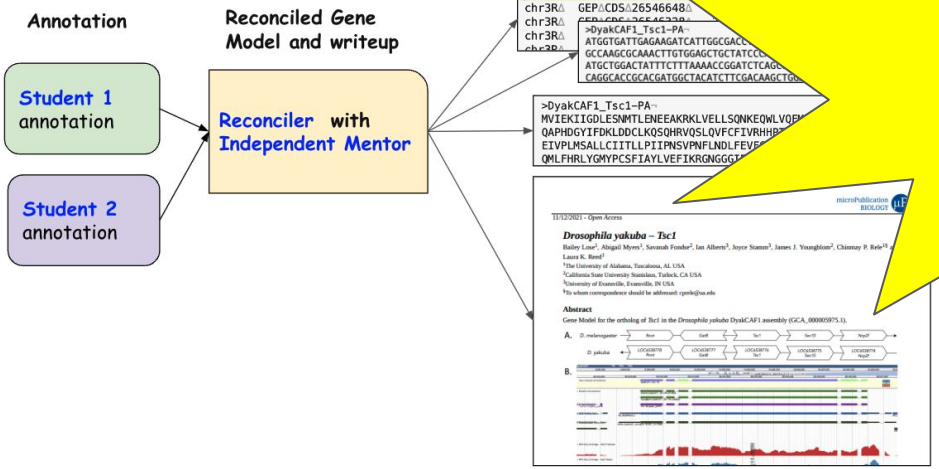
History

Received: 12/19/2020
Revision Received: 1/9/2021

GEP Gene Model Annotation of Drosophila species

GEP Started by Sarah Elgin at Wash U, now headed by Laura Reed at U Alabama

Project: Students use evidence from multiple sources to generate a gene model. Sources include gene model predic… RNA …lata … Encode data, homology calls, etc.). Currently focusing on Pat… …g)



1728

…logs across 28 …osophila and …

2… …ubmit… …in the nea… future
2… submitte… near the end of every summer

Rele CP, Sandlin KM, Leung W, Reed LK. 2020. Manual Annotation of Genes within Drosophila Species: the Genomics Education Partnership protocol. bioRxiv 2020.12.10.420521 doi:10.1101/2020.12.10.420521

Drosophila grimshawi – Rheb

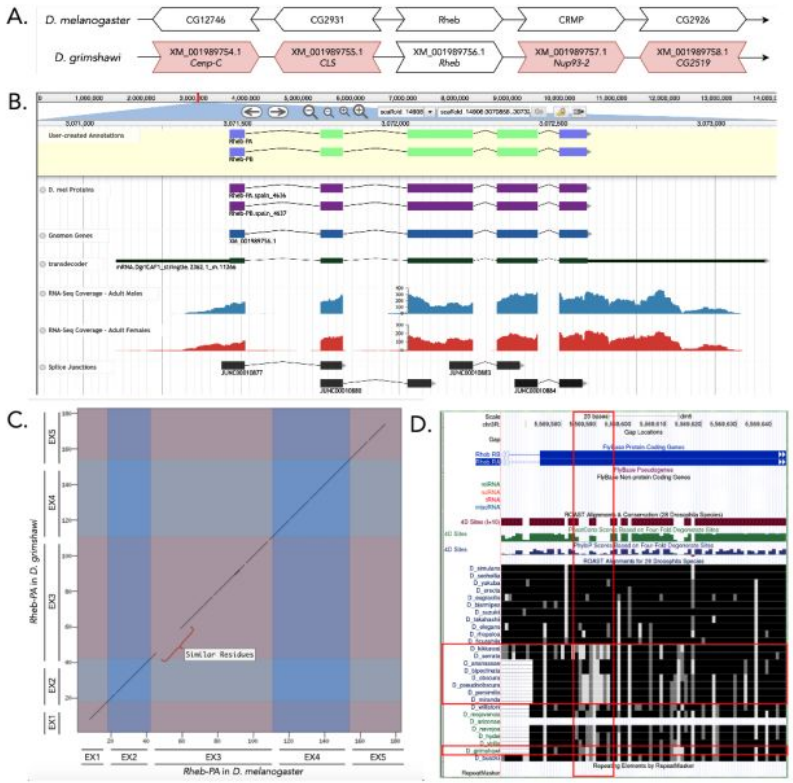
Chinmay P. Rele[1§], Jared Williams[2], Laura K Reed[1], James J Youngblom[2] and Wilson Leung[4]

[1]The University of Alabama, Tuscaloosa, AL 35487

[2]California State University Stanislaus, Turlock, CA 95382

[4]Washington University in St. Louis, St. Louis, MO 63130

[§]To whom correspondence should be addressed: cprele@ua.edu



Description

Introduction

The Ras homolog enriched in brain (*Rheb*) encodes a Ras homolog (Karassek *et al.* 2010). Ras is family of genes that make proteins involved in cell signaling pathways that control cell growth and cell death (Banerjee and Resat 2016). The gene model reported here (dgri_Rheb) was developed for the May 2011 assembly of *D. grimshawi* Agencourt dgri_caf1/DgriCAF1 (GCA_000005155.1) to describe the ortholog to *D. melanogaster Rheb*(FBgn0041191) at the locus previously annotated as XM_001989756.1. The general protocol and datasets for the genome browser tracks used to establish this reported gene model are reported in Rele *et al.* 2020. Additional tools and resources used in generating and confirming this model include HISAT (Kim *et al.* 2015), BEDTools (Quinlan *et al.* 2010), and the Sequence Read Archive (trace.ncbi.nlm.nih.gov/Traces/sra/?study=SRP073087).

Synteny

The *Rheb* gene on chromosome 3R in *D. melanogaster* is surrounded by the genes *CG12746*, *CG2931*, *CRMP*, and *CG2926*. Upon a *tblastn* search, the *Rheb* ortholog gene in *D. grimshawi*, on scaffold 14906, is surrounded by the genes XM_001989754.1, XM_001989755.1, XM_001989575.1, and XM_001989576.1 (orthologous to *Cenp-C* , *CLS*, *Nup93-2*, and *CG2519)* in *D. melanogaster*, Figure 1A). Though none of these flanking genes appear to be orthologous between the two species, we determined this region to contain the ortholog for *Rheb* in *D. grimshawi* because this location had a substantially better *blastp* hit to the *Rheb* protein sequence than to the second-best hit.

Gene Model

This gene model contains two isoforms of the *Rheb* protein in *D. grimshawi*, Rheb-PA and Rheb-PB (Figure 1B). Each of these isoforms contains five identical coding exons. The model in *D. melanogaster* and *D. grimshawi* have the same length and number of exons, and are similar in peptide sequence. However, there is a dissimilarity at the start of the third exon; but the peptides replacing the ones in *D. melanogaster* have similar properties. The coordinates of the corrected gene model can be found in NCBI at GenBank/BankIt using the accession BK014396 and archived in CaltechData here.

Special Characters of Gene Model

Improper alignment at start of coding exon three: As shown in Figure 1D, there is a stretch of amino acids (Figure 1D) that are not identical (black color), but similar (gray/white), between *D. melanogaster* and *D. grimshawi*. This pattern is also evident in a few other species as compared to *D. melanogaster* (Figure 1D). Though this is more obvious in *D. kikkawai*, *D. serrata*, *D. obscura*, *D. pseudoobscura*, *D. persimilis*, and *D. miranda* (Figure 1D) as compared to *D. grimshawi* (Figure 1D), it nonetheless exists.

Changes in the splice donor site for coding exon four: Rheb-PA in *D. grimshawi* has a canonical GT splice donor site at the end of coding exon four. In contrast, the end of coding exon four of Rheb-PA in *D. melanogaster*, and in most of the other *Drosophila* species across the genus *Sophophora*, use a non-canonical GC splice donor site.

Extended data:

Fasta (.faa and .fna) and GFF files are archived and available on CaltechData:
D_grimshawi_Rheb_geneModel_gff_faa_fna.
This file contains

- Peptide Sequence of *Rheb* in *D. grimshawi*
- Nucleotide Sequence of *Rheb* in *D. grimshawi*
- GFF of *Rheb* in *D. grimshawi*

Acknowledgments: We would like to acknowledge Katie M. Sandlin, Gary Williams, and Terence Murphy for their constructive criticism of this article.

References

Aspuria P. J., and F. Tamanoi, 2004. The Rheb family of GTP-binding proteins. Cell. Signal. Volume 16, Issue 10: 1105-1112. DOI: https://doi.org/10.1016/j.cellsig.2004.03.019

Large CURE Project Manuscript Issues

- Articles follow a course template
    - Template needed to be adapted for publishing, Science Officer and Curators defined a gold standard for data interoperability
- Reviewers for so many small articles - need to maintain high standards for peer-review. Reviewers need to come from the larger science community, not from within the project.
- Reviews and revisions might need to be done after the course is over, can be complicated by large number of authors

GEP Review Panel (for batch of 29 articles)

- ★ **Panel Core**
  - **Managing Editor** Karen Yook (microPublication)
  - **Science Officer** Brain Oliver (NIH/NIDDK)
  - **Curators/Data Experts** -
    - nomenclature/community standards check - Lynn Crosby (FlyBase)
    - NCBI RNAseq - Terrence Murphy (NCBI)
    - 12 Drosophila Genomes - Stephen Wade Schaeffer (UPenn)
- ★ **Independent Subject Experts** based on homologs under review
  - **Liz Rideout** UBC, CN
  - **Irene Miguel-Aliaga** ICL, UK
  - **Alex Gould** FCI, UK
- ★ **Project representatives**
  - Reconciler - Chinmay Rele (Author)
  - PI -Laura Reed

Advantages of Review Panel

Review Panel
- Panel Core (Staff, Science Officer, Curators)
- Independent Subject Experts
- Project representatives

- Consistency in review
  - Reviewers found issues across all articles that could be discussed at once
  - Allows GEP to adjust future articles based on generalized comments
- Efficiency in review process
  - Decreases reviewer fatigue
  - Eliminates need to make and track 1700 individual reviewer invitations
- Provides area for open dialog between data producers and data users
  - Speeds revision process

Why microPublication works for Large CURE data

- Undergraduates have better chance of getting credit quickly
  - Only require data that can be generated during a course, so manuscript is small
  - Short articles easier to review and revise, so potential for quick turnaround
- microPublications are peer-reviewed and can be found in PubMed
  - faculty, especially PUI faculty, can get credit for promotion and tenure
- APCs are sustainable
- microPublication team is willing to publish undergraduate work and help make sure specific workflow needs are addressed

Thank you Jacob Kagey, Chinmay Rele, and Laura Reed for last minute information

“…producing the micropublication with these students ***encouraged one to finished undergrad and go to graduate school*** *for research,* ***three others transferred to 4 years <sic> schools to continue research*** *as well. And, we were able to* ***use the publication as a proof of concept and got two more grants*** *to continue offering summer program collaborations between UNC and Durham Technical Community College! “*

- Eric Hastie

Lina Dahlberg WWU

Using microPublications for teaching and for biology education research