

Daniel E. Cook

✉ danielecook@gmail.com • 🌐 danielecook.com

Experience

- | | |
|---|--|
| Google
<i>AI Resident (Health)</i> | Remote
<i>2020-Present</i> |
| Swanton Lab, Crick Institute
<i>Principal Research Fellow (Senior Bioinformatician)</i> | London, UK
<i>2018-2020</i> |
| ◦ Development of bioinformatic pipelines as part of the TRACERx study. The project uses high-depth, whole-exome sequence data from multiple tumor regions to study many aspects of lung cancer. | |
| Vectorbase
<i>Scientific Developer</i> | Remote
<i>2018</i> |
| Tempus, Inc.
<i>Bioinformatics Intern</i> | Chicago, IL
<i>2017</i> |
| Andersen Lab, Northwestern University
<i>Graduate Student</i> | Evanston, IL
<i>2014-2018</i> |
| ◦ Assembled, organized, and processed the largest collection of <i>C. elegans</i> wild isolate sequence data in existence. | |
| ◦ Identified a gene and variant underlying natural variation in telomere length across the <i>C. elegans</i> species. | |
| ◦ Created a web-based platform for performing studies of natural diversity in <i>C. elegans</i> (elegansvariation.org). | |
| ◦ Co-discovered a new species: <i>C. oiwi</i> . | |
| Ober Lab, University of Chicago
<i>Research Assistant</i> | Chicago, IL
<i>2012-2013</i> |
| ◦ Assisted with development of databases for phenotypic and genotypic data. | |
| ◦ Helped disclose carrier status to individuals. | |
| Murray Lab, University of Iowa
<i>Research Assistant</i> | Iowa City, IA
<i>2009-2012</i> |
| ◦ Conducted geographic analysis of Iowa newborn screen data. | |
| ◦ Collected human placental tissue for genetic, epigenetic, and expression studies. | |

Computational Skills

Projects: GitHub.com/danielecook	Development: git, CI-testing, Docker
Programming: Python, nim, Go, R, Bash, Javascript, Stata, L ^A T _E X	Database: SQL, BigQuery, Google Datastore
Machine Learning: scikit-learn, scikit-image, pandas	Pipelines: Nextflow, Snakemake
Cloud: AWS, Google Cloud Platform	Misc: Web development, Django/Flask

Education

Northwestern University <i>Ph.D. Biological Sciences (Emphasis on Bioinformatics and Genomics)</i>	Evanston, IL <i>2013-2018</i>
University of Iowa <i>B.S. Biology with Honors</i>	Iowa City, IA <i>May 2010</i>

Honors and Awards

- **2014 - NSF-GRFP:** National Science Foundation Graduate Research Fellowship (Approx. \$130,000)

Software

Additional work can be seen on my GitHub profile: [GitHub.com/danielecook](https://github.com/danielecook)

Python Packages

- VCF-kit - Command line utility for analyzing genetic variation.
- CeNDR (elegansvariation.org) - Flask website that provides tools for working with wild isolates of the *C. elegans* species including a genome browser and a genome-wide association mapping pipeline.
- IEX-API-python - Python library for accessing data from the Investors Exchange (IEX) stock/financial data API.

R Packages

- rdatastore - provides an R-based interface for Google Datastore.

Open Source Contributions

- mosdepth - Added coverage summary for depth of coverage calculations of sequencing data.
- memoise - developed functionality for caching function results (memoisation) using cloud-based environments such as Amazon S3.

Side projects

- hntrending.com - A website that queries HackerNews stories and comments in unique ways, and sorts the most popular stories by week/month/year.
- upvote.pub [source] - A reddit-style clone specifically geared towards scientific publications. This website was experimental and is no longer active.

Stata Programs

- manhattan - Generation of Manhattan plots.
- ccmatch - Utility for case/control studies.

Professional Activities

- **2015, 2016** - NSF Data Science Workshop @ University of Washington

Teaching

- **Fall 2016** - *Biol. Sci. 390* - Advanced Molecular Biology
- **Spring 2015** - *Biol. Sci. 393* - Genetic Analysis
- **Summer 2014-2016** - IBiS Computational Biology Bootcamp

Mentorship

Joshua Roberts, 2015-2016 - *Computer Science (CS) Undergraduate* - Improved software engineering practices, data handling ability, and developed web-application skillset. Joshua is now a software engineer at Lyft.

Rohit Rastogi, 2016-2017 - *CS Undergraduate* - Fostered development of programming skills in R and Python. Rohit now works as a software engineer at Google.

Publications

- 2021 [Preprint] Balancing selection maintains hyper-divergent haplotypes in *C. elegans***
Lee, Daehan, Stefan Zdraljevic, Lewis Stevens, Ye Wang, Robyn E. Tanny, Timothy A. Crombie, **Cook, Daniel E.**, Amy K. Webster, Rojin Chirakar, L. Ryan Baugh, Mark G. Sterken, Christian Braendle, Marie-Anne Félix, Matthew V. Rockman, and Erik C. Andersen
2021 *bioRxiv*. DOI: 10.1101/2020.07.23.218420.
- 2020 [Preprint] The distribution of mutational effects on fitness in *Caenorhabditis elegans* inferred from standing genetic variation**
Gilbert, Kimberly J., Stefan Zdraljevic, Daniel E. Cook, Asher D. Cutter, Erik C. Andersen, and Charles F. Baer
2020 *bioRxiv*. DOI: 10.1101/2020.10.26.355446.
- 2019 Long-read sequencing reveals intra-species tolerance of substantial structural variations and new subtelomere formation in *C. elegans***
Kim, Chuna, Jun Kim, Sunghyun Kim, **Cook, Daniel E.**, Kathryn S Evans, Erik C Andersen, and Junho Lee
2019 *Genome research* 29.6. PMID: 31123081 DOI: 10.1101/gr.246082.118.
- A Novel Gene Underlies Bleomycin-Response Variation in *Caenorhabditis elegans***
Brady, Shannon C, Stefan Zdraljevic, Karol W Bisaga, Robyn E Tanny, **Cook, Daniel E.**, Daehan Lee, Ye Wang, and Erik C Andersen
2019 *Genetics*. PMID: 31171655 DOI: 10.1534/genetics.119.302286.
- Selection and gene flow shape niche-associated variation in pheromone response**
Lee, Daehan, Stefan Zdraljevic, **Cook, Daniel E.**, Lise Frézal, Jung-Chen Hsu, Mark G Sterken, Joost AG Riksen, John Wang, Jan E Kammenga, Christian Braendle, et al.
2019 *Nature ecology & evolution* 3.10. PMID: 31548647 DOI: 10.1038/s41559-019-0982-3.
- Population Selection and Sequencing of *Caenorhabditis elegans* Wild Isolates Identifies a Region on Chromosome III Affecting Starvation Resistance**
Webster, Amy K, Anthony Hung, Brad T Moore, Ryan Guzman, James M Jordan, Rebecca EW Kaplan, Jonathan D Hibshman, Robyn E Tanny, **Cook, Daniel E.**, Erik Andersen, and Ryan Baugh
2019 *G3: Genes, Genomes, Genetics* 9.10. PMID: 31444297 DOI: 10.1534/g3.119.400617.
- Deep sampling of Hawaiian *Caenorhabditis elegans* reveals high genetic diversity and admixture with global populations**
Crombie, Tim A, Stefan Zdraljevic, **Cook, Daniel E.**, Robyn E Tanny, Shannon C Brady, Ye Wang, Kathryn S Evans, Steffen Hahnel, Daehan Lee, Briana C Rodriguez, et al.
2019 *eLife* 8. PMID: 31793880 DOI: 10.7554/eLife.50465.
- 2018 Discovery of genomic intervals that underlie nematode responses to benzimidazoles**
Zamanian, Mostafa, **Cook, Daniel E.**, Stefan Zdraljevic, Shannon C Brady, Daehan Lee, Junho Lee, and Erik C Andersen
2018 *PLoS neglected tropical diseases* 12.3. PMID: 29601575 DOI: 10.1371/journal.pntd.0006368.

- Shared genomic regions underlie natural variation in diverse toxin responses**
 Evans, Kathryn S, Shannon C Brady, Joshua S Bloom, Robyn E Tanny, **Cook, Daniel E**, Sarah E Giuliani, Stephen W Hippleheuser, Mostafa Zamanian, and Erik C Andersen
2018 *Genetics* 210.4. PMID: 30341085 DOI: 10.1534/genetics.118.301311.
- 2017 Natural variation in the distribution and abundance of transposable elements across the *Caenorhabditis elegans* species**
 Laricchia, KM, S Zdraljevic, **Cook, DE**, and EC Andersen
May 2017 *Mol. Biol. Evol.* PMID: 28486636 DOI: 10.1093/molbev/msx155.
- Natural variation in a single amino acid substitution underlies physiological responses to topoisomerase II poisons**
 Zdraljevic, Stefan, Christine Strand, Hannah S Seidel, **Cook, Daniel E**, John G Doench, and Erik C Andersen
July 2017 *PLoS Genet.* 13.7. PMID: 28700616 DOI: 10.1371/journal.pgen.1006891.
- VCF-kit: Assorted utilities for the variant call format.**
Cook, DE and EC Andersen
Jan. 2017 *Bioinformatics (Oxford, England)*. PMID: 28093408 DOI: 10.1093/bioinformatics/btx011.
- 2016 CeNDR, the *Caenorhabditis elegans* natural diversity resource**
Cook, Daniel E, Stefan Zdraljevic, Joshua P Roberts, and Erik C Andersen
Oct. 2016 *Nucleic Acids Research*. PMID: 27701074 DOI: 10.1093/nar/gkw893.
- The Genetic Basis of Natural Variation in *Caenorhabditis elegans* Telomere Length**
Cook, Daniel E, Stefan Zdraljevic, Robyn E Tanny, Beomseok Seo, David D Riccardi, Luke M Noble, Matthew V Rockman, Mark J Alkema, Christian Braendle, Jan E Kammenga, John Wang, Leonid Kruglyak, Marie-Anne Félix, Junho Lee, and Erik C Andersen
Sept. 2016 *Genetics* (204(1):371-83). PMID: 27449056 DOI: 10.1534/genetics.116.191148.
- 2015 DYRK1A controls the transition from proliferation to quiescence during lymphoid development by destabilizing Cyclin D3**
 Thompson, Benjamin J, Rahul Bhansali, Lauren Diebold, **Daniel E Cook**, Lindsay Stolzenburg, Anne-Sophie Casagrande, Thierry Besson, Bertrand Leblond, Laurent Désiré, Sébastien Malinge, and John D Crispino
2015 *Journal of Experimental Medicine* 212.6. DOI: 10.1084/jem.20150002.
- 2013 The influence of maternal disease on metabolites measured as part of newborn screening**
 Ryckman, Kelli K, Oleg A Shchelochkov, **Daniel E Cook**, Stanton L Berberich, Sara Copeland, John M Dagle, and Jeffrey C Murray
Sept. 2013 *The J. of Maternal-Fetal & Neonatal Medicine* (26(14):1380-3). PMID: 23550828 DOI: 10.3109/14767058.2013.791267.
- The heritability of metabolic profiles in newborn twins.**
 Alul, Farah Y, **Cook, Daniel E**, Oleg A Shchelochkov, Lauren G Fleener, Stanton L Berberich, Jeffrey C Murray, and Kelli K Ryckman
Mar. 2013 *Heredity* (110(3):253-8). PMID: 23149456 DOI: 10.1038/hdy.2012.75.

Clinical and environmental influences on metabolic biomarkers collected for newborn screening

Ryckman, Kelli K, Stanton L Berberich, Oleg A Shchelochkov, **Daniel E Cook**, and Jeffrey C Murray

Jan. 2013 *Clinical Biochemistry* (46(1-2):133-8). PMID: 23010448 DOI: 10.1016/j.clinbiochem.2012.09.013.

Generating Manhattan plots in Stata

Cook, Daniel E, Kelli R Ryckman, and Jeffrey C Murray

2013 *Stata Journal* (13(2):323-328).

2012 Replication of clinical associations with 17-hydroxyprogesterone in preterm newborns

Ryckman, Kelli K, **Daniel E Cook**, Stanton L Berberich, Oleg A Shchelochkov, Susan K Berends, Tamara Busch, John M Dagle, and Jeffrey C Murray

Jan. 2012 *J. of Ped. Endo. and Met.* (25(3-4):301-5). PMID: 22768660 DOI: 10.1515/jpem-2011-0456.

Presentations

2016 The genetic basis of natural variation in *C. elegans* telomere length

Selected Talk – International Conference on Quantitative Genetics, Madison WI.

The *C. elegans* natural diversity resource

Selected Talk – Midwest *C. elegans* meeting

Van Andel Research Institute, Grand Rapids, MI.

2015 Genome-Wide and Species-Wide Variation in *C. elegans* Reveals Association of Telomere Length With Population Differences in *pot-2*

Selected Talk – International Worm Meeting 2015, UCLA.

2012 Mapping Newborn Screening Data: A Geographic Exploration

Poster – Pediatric Academic Societies Meeting, Boston, MA.

Volunteer Activities

Americorps

Boston, MA

Volunteer

2010-2011

- Worked for ten months on a team with 15 fellow volunteers at the Maurice J. Tobin K-8 school.
- Tutored and mentored a 6th-grade classroom.
- Helped run an after-school program in partnership with community organizations.

Table-To-Table (2007-2010) - Salvaged food every week from grocery stores and distributors in central Iowa for distribution at area food pantries.

Appalachia Service Project - Went on eight annual one-week trips during high-school and college to make homes safer, warmer, and drier for disadvantaged families in the Appalachian mountains of Kentucky and Tennessee.

Alpha Phi Omega - Member, 3.5 years; Engaged in a large variety of volunteer work, including performing environmental cleanups and teaching first-aid to boy scouts.