

# T. Jeffrey Cole

github.com/tijeco

jffcole7@gmail.com

## Education

- ▣ **Ph.D. in Biological Sciences** May 2021  
East Carolina University, Greenville NC
  - Dissertation title: *Molecular evolution of venom proteins in Ctenidae spiders*
- ▣ **B.S. in Biology, *magna cum laude*** May 2016  
Samford University, Birmingham AL
  - **GPA:** 3.88
  - **Honors**
    - ◇ Beta Beta Beta National Biology Honors Society (2013-2015)
      - Vice President (2015)
    - ◇ Alpha Lambda Delta Honors Society (2013-2015)

## Skills

**Bioinformatics Pipelines:** Nextflow, nf-core tools, Snakemake

**Cloud Computing:** AWS (S3, EC2, Batch), Seqera Platform

**Machine Learning:** scikit-learn, TensorFlow, PyTorch

**Reproducible Analysis:** Conda, Docker

**Data Visualization:** tidyverse, ggplot2, plotly

**Programming Languages:** Python, R, BASH, Go

## Experience

- ▣ **Senior Bioinformatics Engineer** Feb 2022 - present  
ExosomeDx (Bio-Techne), Austin TX
  - **Multi-omic liquid biopsy bioinformatics platform**
    - ◇ Lead developer of bioinformatics pipelines for RNA fusions and splice variants in RNAseq datasets
    - ◇ Lead coordinator of AWS infrastructure and pipeline management through Seqera platform
    - ◇ Contributor to nextflow pipelines for RNAseq gene quantification, differential gene expression, and cfDNA methylation analysis
  - **Biomarker discovery machine learning platform**

- ◇ Project coordinator for interdivisional machine learning competition for biomarker discovery using TCGA data
- ◇ Lead developer of machine learning models and algorithms for biomarker discovery using multi-omic liquid biopsy data

#### □ **Postdoctoral Research Fellow**

Jan 2021 - Jan 2022

University of Texas at Austin, Austin TX

- **Discovery of antimicrobial peptides in bacterial genomic datasets**
  - ◇ Developer of bioinformatics pipeline to identify small novel antimicrobial peptides in bacterial genome assemblies
- **Design of antimicrobial peptides with protein language deep learning models**
  - ◇ Lead contributor of analysis to combine libraries of synthetic peptides with machine learning models to inform design of peptides with optimal antimicrobial properties
- **Integration of deep protein language models with 3D structural models**
  - ◇ Project coordinator for the development of machine learning models that aggregate information from sequence based models and structure based models to improve protein function prediction

#### □ **NSF Graduate Research Fellow**

Jun 2016 - Dec 2020

East Carolina University, Greenville NC

- **Molecular evolution of inhibitor cysteine knots in wandering spiders**
  - ◇ Independent research study involving specimen collection from the field, venom isolation and RNA extraction in the lab, and bioinformatics analysis of venom gland transcriptomics and proteomics
  - ◇ Analysis of molecular evolution of inhibitor cysteine knots using phylogenetic reconstruction, disulfide bond structural predictions, and tests for positive selection
- **Identification of families under positive selection in transcriptomic datasets**
  - ◇ Developer of bioinformatics pipeline to cluster transcripts into gene families and reconstruct phylogenetic trajectories to identify signatures of positive selection
- **Deep learning model for classifying animal toxins**
  - ◇ Developer of recurrent neural network model predict the toxicity of proteins expressed in animal venom

### Publications

Full list here: <https://scholar.google.com/citations?user=MbxmQhcAAAAAJ&hl=en>

Tutrone R, Lowentritt B, Neuman B, Donovan MJ, Hallmark E, **Cole TJ**, Yao Y, Biesecker C, Kumar S, Verma V, Sant GR. (2023). *Prostate Cancer and Prostatic Diseases*, 26(3), 596-601.

Randall JR, DuPai CD, **Cole TJ**, Davidson G, Groover KE, Slater SL, Mavridou DA, Wilke CO, Davies BW. (2023). Designing and identifying  $\beta$ -hairpin peptide macrocycles with antibiotic potential. *Science Advances*, 9 (2), eade0008.

Kulikova AV, Diaz DJ, Chen T, **Cole TJ**, Ellington AD, Wilke CO. (2023). Two sequence-and two structure-based ML models have learned different aspects of protein biochemistry. *Scientific Reports*, 13 (1), 13280.

- Cole TJ**, Parker JK , Feller AL, Wilke CO, Davies BW. (2022). Evidence for Widespread Class II Microcins in Enterobacterales Genomes. *Applied and Environmental Microbiology* 88 (23), e01486-22.
- Randall JR, Davidson G, Fleeman RM, Acosta SA, Riddington IM, **Cole TJ**, DuPai CD, Davies BW. (2022). Synthetic antibacterial discovery of symbah-1, a macrocyclic  $\beta$ -hairpin peptide antibiotic. *IScience*, 25 (1).
- Brewer MS, **Cole TJ**. (2023). Killer knots: Molecular evolution of inhibitor cystine knot toxins in Wandering spiders (Araneae: Ctenidae). *Toxins* 15 (2), 112.
- Cole TJ**, Brewer MS. (2019). toxify: a deep learning approach for classifying animal venom proteins. *PeerJ*, 7, p.e7200.
- Cole TJ**, Brewer MS. (2018) FUSTr: A tool for finding gene Families Under Selection in Transcriptome databases. *PeerJ*, 6, e4234.

## Awards & Grants

American Museum of Natural History Theodore Roosevelt Memorial Grant (\$1,000)	2017
East Carolina University biodiversity initiative internal grant for outreach (\$200)	2017
NSF Graduate Research Fellowship (\$132,000)	2016
Brakefield Textbook scholarship (\$100)	2016
NSF REU Summer research grant, University of Kansas (\$5,500)	2015
NSF REU Summer research grant, Samford University (\$5,500)	2014
Samford University Summer Propst Research scholar (\$2,000)	2013

## Teaching

- Principles of Biology Laboratory, BIOL 1101, graduate teaching assistant**
2018-2020
  - East Carolina University
  - Utilizing Argument Driven Inquiry to guide students in experiments related to molecular biology, bioenergetics, cellular structure, physiology, and the molecular basis of inheritance and control of gene expression
- Practical computing, BIOL 6220, guest lecturer**
2017-2019
  - East Carolina University
  - Best practices for reproducibility in bioinformatics
  - Phylogenetic tests for positive selection
- Population Biology, BIOL 350, guest lecturer**
2019
  - Samford University
  - Techniques and principles of population genomics