



# Data Science Coast To Coast

## SEMINAR SERIES

**The DS C2C seminar series, hosted jointly by seven academic data science institutes, provides a unique opportunity to foster a broad-reaching data science community.**

In the first half of 2021, we will host five seminars, each featuring one faculty member and one postdoctoral fellow from two universities. Each speaker will give a 20-minute talk about ongoing projects and motivating issues, followed by 20 minutes of discussion with the audience. These seminars will be the launching point for follow-on research discussion meetings which will hopefully lead to fruitful collaborative research.

We strongly encourage you to [sign up ahead of time](#), and indicate your area of research and your interest in follow-up discussions.

**ALL TALKS WILL BE AT 3 PM EST/NOON PST.**





Dr. Rosemary Gillespie



Dr. Shelly Trigg

## ~ 19 May 2021 ~ Biodiversity

**Rosemary Gillespie** (Professor & Schlinger Chair in Systematic Entomology, University of California, Berkeley), *Data Integration Across Space and Time to Infer Biodiversity Dynamics*

**Shelly Trigg** (Data Science Postdoctoral Fellow, University of Washington), *Diversity in Animal Response to Environmental Change*

**ZOOM LINK**

## Biodiversity

### Dr. Rosemary Gillespie and Dr. Shelly Trigg

#### DATA INTEGRATION ACROSS SPACE AND TIME TO INFER BIODIVERSITY DYNAMICS

The world's ecosystems are under serious threat due to ongoing stressors of the Anthropocene, notably habitat destruction, climate change, loss of biodiversity, disease, and the spread of invasive species. Biodiversity in particular is suffering catastrophic decline and tracking and understanding the factors affecting change is a major challenge that we are currently not meeting. Unless we develop new approaches, it will take centuries to document biodiversity and identify attributes that

render ecological communities robust and resilient to change, and by then it will likely be too late. Here, we examine insights we can gain into biodiversity dynamics by looking at ways that we can first assess spatial patterns of diversity, abundance, and foodwebs, and determine the response of the organisms within these communities, to the changing environments that surround them. We have piloted an environmental DNA approach to generate estimates of abundance and interactions of macroorganisms in terrestrial systems across different spatial scales. By applying various theoretical and modeling approaches to the vast amounts of genetic data, we can encapsulate the “status” of a biological community in terms of its integrity and potential resilience to change. Moreover, by analyzing these data through slices in time (months, years, decades, or longer), we can assess how the community might accommodate, adapt, or collapse in response to change. These changes include habitat transformation, climate modification, fire, or disease. The critical data challenge is to integrate data that characterize the biological community, genomic data that reveal the response of any given taxon to that change, with past, present, and modeled climate change data. We highlight the role of historic collections from museums, and the physical record they provide of past environments.

## DIVERSITY IN ANIMAL RESPONSE TO ENVIRONMENTAL CHANGE

How will ecosystems tolerate the climate and ocean change occurring now and predicted for the future? To begin addressing this question, we can subject different animals to different anthropogenic pressures and evaluate their responses. We can more sensitively and comprehensively assess responses by performing molecular surveys using omics technologies (e.g. genomics, proteomics, metabolomics, etc.), which allow us to more clearly see the cellular processes that underlie environmental tolerance and intolerance. This data can also help us compare between species since all species have these general molecules (DNA, proteins, metabolites) in common. I'm going to present data from different studies on marine invertebrates exposed to different environmental conditions, and describe how I used multiple data science approaches to distill large omics datasets into dominant biological pathways associated with environmental tolerance and intolerance. After summarizing responses across species and conditions, I will propose future directions and data science applications for the wealth of environmental omics data being generated.



Dr. Miguel Jimenez-Urrea