BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. DO NOT EXCEED FIVE PAGES.

NAME: Rutherford, Erica

eRA COMMONS USER NAME (credential, e.g., agency login):

POSITION TITLE: Data Wrangler

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)*

INSTITUTION AND LOCATION	DEGREE (if applicable)	END DATE MM/YYYY	FIELD OF STUDY
Longwood University	BS	05/2010	Biology
San Francisco State University	MS	08/2016	Ecology and Evolutionary Biology

A. Personal Statement

My career spanning nine years as a data curator has given me a lot of experience and perspective into the workings of scientific data and metadata, and its organization. During my time before graduate school, I attained experience on a variety of fieldwork projects in ecology (2008-2013). During these temporary seasonal assignments, the importance of precision and care in data collection was impressed in me. When I went to graduate school at San Francisco State University (2013-2016), I gained experience in all parts of a molecular biology experiment, from fieldwork to labwork to data analysis. After graduation, I worked at a microbiome focused startup company, Second Genome, as their data curator (2016-2021). While there, I was responsible for curation of metadata for both internal studies for R&D and for clients, and for external studies being brought in for our internal Knowledgebase. While there, I developed an appreciate for ontologies, and developed a custom Second Genome Ontology to handle our metadata. I moved on to the Lattice group, located at Stanford University, where I continued to expand my skills in data curation (2021-present). I have gained experience in handling single cell datasets and their associated metadata, and curating them to meet precise standards. I strive to work collaboratively with data contributors in order to ensure FAIR data standards.

1. Abdulla S, Aevermann B, Assis P, Badajoz S, Bell SM, Bezzi E, Cakir B, Chaffer J, Chambers S, Cherry JM, Chi T, Chien J, Dorman L, Garcia-Nieto P, Gloria N, Hastie M, Hegeman D, Hilton J, Huang T, Infeld A, Istrate AM, Jelic I, Katsuya K, Kim YJ, Liang K, Lin M, Lombardo M, Marshall B, Martin B, McDade F, Megill C, Patel N, Predeus A, Raymor B, Robatmili B, Rogers D, Rutherford E, Sadgat D, Shin A, Small C, Smith T, Sridharan P, Tarashansky A, Tavares N, Thomas H, Tolopko A, Urisko M, Yan J, Yeretssian G, Zamanian J, Mani A, Cool J, Carr A. CZ CELLxGENE Discover: a single-cell data platform for scalable exploration, analysis and modeling of aggregated data. Nucleic Acids Res. 2025 Jan 6;53(D1):D886-D900. PubMed Central PMCID: PMC11701654.
2. Ravichandar JD, Rutherford E, Chow CT, Han A, Yamamoto ML, Narayan N, Kaplan GG, Beck PL, Claesson MJ, Dabbagh K, Iwai S, DeSantis TZ. Strain level and comprehensive microbiome analysis in inflammatory bowel disease via multi-technology meta-analysis identifies key bacterial influencers of disease. Front Microbiol. 2022;13:961020. PubMed Central PMCID: PMC9614153.
3. Yin X, Altman T, Rutherford E, West KA, Wu Y, Choi J, Beck PL, Kaplan GG, Dabbagh K, DeSantis TZ, Iwai S. A Comparative Evaluation of Tools to Predict Metabolite Profiles From Microbiome Sequencing Data. Front Microbiol. 2020;11:595910. PubMed Central PMCID: PMC7746778.
4. Rutherford EM, Ontano A, Kantor C, Routman EJ. Genetic variation across trophic levels: A test of the correlation between population size and genetic diversity in sympatric desert lizards. PLoS One. 2019;14(12):e0224040. PubMed Central PMCID: PMC6894812.

B. Positions, Scientific Appointments and Honors

Positions and Scientific Appointments

2021 -	Data Wrangler, Stanford University
2016 - 2021	Data Curator, Second Genome Inc.

Honors

2016	Outstanding Graduate Assistant award, San Francisco State University
2014	ARCS (Achievement Rewards for College Scientists) Scholar, ARCS Foundation

C. Contribution to Science

1. During my master's research, I focused on population genetics of nonmodel organisms in natural populations. My intention was to test the neutral theory of genetic variation, by testing the relationship between population size and genetic diversity in four desert lizard species. I used trophic level as a proxy for long term population size, as well as field observations of abundance. I found the inverse of what I expected, with no greater genetic diversity in larger sized populations. As I did not find evidence of selection on the loci sequenced for this study, my results did not support neutral theory predictions.
   a. Rutherford EM, Ontano A, Kantor C, Routman EJ. Genetic variation across trophic levels: A test of the correlation between population size and genetic diversity in sympatric desert lizards. PLoS One. 2019;14(12):e0224040. PubMed Central PMCID: PMC6894812.
2. My career after graduate school has focused on the curation of scientific data and metadata to FAIR principles (emphasizing Findability, Accessibility, Interoperability, and Reuse). I worked at a startup company, Second Genome Inc, focused on the development of drugs and other products based on microbiome data. I was responsible for curating metadata for inclusion in their proprietary Knowledgebase, used for data analysis and drug discovery. I developed a custom ontology to fit the needs of the company. The biggest focus of the company was on human health, although I also curated studies with animal data and worked on an agricultural product partnership. In this way, I was able to support multiple projects in different branches of science and industry. I worked with scientists to help analyze their data and develop better models.
   a. Ravichandar JD, Rutherford E, Chow CT, Han A, Yamamoto ML, Narayan N, Kaplan GG, Beck PL, Claesson MJ, Dabbagh K, Iwai S, DeSantis TZ. Strain level and comprehensive microbiome analysis in inflammatory bowel disease via multi-technology meta-analysis identifies key bacterial influencers of disease. Front Microbiol. 2022;13:961020. PubMed Central PMCID: PMC9614153.
   b. Yin X, Altman T, Rutherford E, West KA, Wu Y, Choi J, Beck PL, Kaplan GG, Dabbagh K, DeSantis TZ, Iwai S. A Comparative Evaluation of Tools to Predict Metabolite Profiles From Microbiome Sequencing Data. Front Microbiol. 2020;11:595910. PubMed Central PMCID: PMC7746778.
3. I have continued my work in data curation by working on the Lattice team, which focuses on the curation of single-cell genomics datasets. My focus has been on data generated for the Chan Zuckerberg Initiative's Cell Science program, which aims to make the funded data publicly available through CZ CELLxGENE and through the Human Cell Atlas Data Portal. I contribute by working with data generating groups to get their data, and then transform it into a standardized data/metadata format that can be accepted into these public repositories for later reuse.
   a. Abdulla S, Aevermann B, Assis P, Badajoz S, Bell SM, Bezzi E, Cakir B, Chaffer J, Chambers S, Cherry JM, Chi T, Chien J, Dorman L, Garcia-Nieto P, Gloria N, Hastie M, Hegeman D, Hilton J, Huang T, Infeld A, Istrate AM, Jelic I, Katsuya K, Kim YJ, Liang K, Lin M, Lombardo M, Marshall B, Martin B, McDade F, Megill C, Patel N, Predeus A, Raymor B, Robatmili B, Rogers D, Rutherford E, Sadgat D, Shin A, Small C, Smith T, Sridharan P, Tarashansky A, Tavares N, Thomas H, Tolopko A, Urisko M, Yan J, Yeretssian G, Zamanian J, Mani A, Cool J, Carr A. CZ CELLxGENE Discover: a single-cell data platform for scalable exploration, analysis and modeling of aggregated data. Nucleic Acids Res. 2025 Jan 6;53(D1):D886-D900. PubMed Central PMCID: PMC11701654.