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: Hilton, Jason

eRA COMMONS USER NAME (credential, e.g., agency login): 14452550

POSITION TITLE: Director, Lattice

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
University of Miami, Coral Gables, FL	BS	05/2006	Marine Science, Biology
University of California, Santa Cruz, CA	PHD	06/2014	Ocean Science

A. Personal Statement

Through my graduate studies, I learned how to critically analyze research, both my own and that of others, preparing me well to apply scrutiny on data submissions at the level that is expected from data presented on community resources. My graduate research also provided me with the experience of end-to-end genomics research, from sample acquisition to nucleic acid purification to computational analysis of genomic features. This intimate knowledge of the complete data flow prepared me well to translate data generation procedures into a structured format within a database. My time as a data wrangler for the ENCODE project expanded my scientific knowledge in the fields of human genomics, as well as the assays deployed in the community. I proved to be a valuable project leader first while coordinating a large-scale effort to import and re-process data from the NIH Roadmap Epigenomics project in order to enhance the findability and accessibility of a widely-used data corpus. I also spent time as the Scrum Master of the ENCODE DCC, coordinating releases of the codebase and corresponding portal, providing guidance on individual tasks while balancing short-term needs & long-term goals, and offering process-oriented solutions. My leadership has also shown through mediation of scientific working groups, leading weekly group meetings for the ENCODE DCC, and mentoring data wranglers. Activities like organizing data utilization workshops and representing the ENCODE project within the International Human Epigenome Consortium (IHEC) have prepared me well for my current role in the single-cell genomics community. Since starting the Lattice data coordination team, I have applied my leadership skills well beyond our lab and toward the whole single-cell community. By forming relationships with funders, data generators, computational biologists, data resource managers, and ontologists, I have been able to catalyze the design of the first standards for single-cell data and invigorate collaborations across single-cell data resources.

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. Hitz BC, Rowe LD, Podduturi NR, Glick DI, Baymuradov UK, Malladi VS, Chan ET, Davidson JM, Gabdank I, Narayana AK, Onate KC, Hilton J, Ho MC, Lee BT, Miyasato SR, Dreszer TR, Sloan CA, Strattan JS, Tanaka FY, Hong EL, Cherry JM. SnoVault and encodeD: A novel object-based storage system and applications to ENCODE metadata. PLoS One. 2017;12(4):e0175310. PubMed Central PMCID: PMC5389787.

3. Stunnenberg HG, Hirst M. The International Human Epigenome Consortium: A Blueprint for Scientific Collaboration and Discovery. Cell. 2016 Nov 17;167(5):1145-1149. PubMed PMID: 27863232.
4. Hong EL, Sloan CA, Chan ET, Davidson JM, Malladi VS, Strattan JS, Hitz BC, Gabdank I, Narayanan AK, Ho M, Lee BT, Rowe LD, Dreszer TR, Roe GR, Podduturi NR, Tanaka F, Hilton JA, Cherry JM. Principles of metadata organization at the ENCODE data coordination center. Database (Oxford). 2016;2016 PubMed Central PMCID: PMC4792520.

B. Positions, Scientific Appointments and Honors

Positions and Scientific Appointments

2019 -	Director, Lattice, Department of Genetics/Department of Biomedical Data Science, Stanford University, Stanford, CA
2015 - 2019	Senior Data Wrangler, ENCODE Data Coordination Center, Department of Genetics, Stanford University, Stanford, CA
2015 - 2015	Genomics Data Consultant, Independent
2014 - 2015	Science Department Head and Teacher, Fusion Academy, Los Gatos, CA
2006 - 2014	Graduate Researcher, Department of Ocean Sciences, University of California, Santa Cruz, CA
2005 - 2006	Undergraduate Researcher, Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, FL

C. Contribution to Science

1. My early research explored diatom-associated cyanobacteria in the open ocean. These unique symbiotic relationships are vital to both nitrogen cycling, through their ability to fix atmospheric nitrogen, and carbon cycling, through a highly efficient carbon export system into the deep ocean. Using flow cytometry and genomic sequencing technologies, this research revealed the evolutionary pressure a symbiont undergoes as a result of a long-term association with unicellular diatoms. Most striking was unprecedented loss of nitrogen metabolism pathway genes in the symbiont genome. This research was extended to environmental transcriptomics and revealed the symbiont population had significant reduction in photosystem II gene expression and regulation on the cyclic electron transport, surprising discoveries in photosynthetic organisms living near the surface ocean. These analyses revealed much about the ecology and evolution of diatom symbionts, and carry implications into global oceanic nutrient cycling, as well as plant-microbe associations, in general. I designed these studies and conducted the majority of wet-lab and analysis work.
   a. Hilton JA, Satinsky BM, Doherty M, Zielinski B, Zehr JP. Metatranscriptomics of N2-fixing cyanobacteria in the Amazon River plume. ISME J. 2015 Jul;9(7):1557-69. PubMed Central PMCID: PMC4478696.
   b. Hilton JA, Foster RA, Tripp HJ, Carter BJ, Zehr JP, Villareal TA. Genomic deletions disrupt nitrogen metabolism pathways of a cyanobacterial diatom symbiont. Nat Commun. 2013;4:1767. PubMed Central PMCID: PMC3667715.
2. My recent scientific career has focused on providing user-friendly community resources in order to maximize the value of data generated by other researchers. Guided by the FAIR principles, data resources can extend the lifecycle of data, enabling them to be used to test far more scientific hypotheses than just what the data generator had in mind. I first worked primarily on the development of the ENCODE data resource and the standardization of bulk epigenomic data. These publications represent my contribution to the design of metadata standards, utilization of ontologies to enhance data accessibility, and collaborations across consortia.
   a. Gabdank I, Chan ET, Davidson JM, Hilton JA, Davis CA, Baymuradov UK, Narayanan A, Onate KC, Graham K, Miyasato SR, Dreszer TR, Strattan JS, Jolanki O, Tanaka FY, Hitz BC, Sloan CA, Cherry JM. Prevention of data duplication for high throughput sequencing repositories. Database

(Oxford). 2018 Jan 1;2018 PubMed Central PMCID: PMC5829560.

   b. Hitz BC, Rowe LD, Podduturi NR, Glick DI, Baymuradov UK, Malladi VS, Chan ET, Davidson JM, Gabdank I, Narayana AK, Onate KC, Hilton J, Ho MC, Lee BT, Miyasato SR, Dreszer TR, Sloan CA, Strattan JS, Tanaka FY, Hong EL, Cherry JM. SnoVault and encodeD: A novel object-based storage system and applications to ENCODE metadata. PLoS One. 2017;12(4):e0175310. PubMed Central PMCID: PMC5389787.
   c. Stunnenberg HG, Hirst M. The International Human Epigenome Consortium: A Blueprint for Scientific Collaboration and Discovery. Cell. 2016 Nov 17;167(5):1145-1149. PubMed PMID: 27863232.
   d. Hong EL, Sloan CA, Chan ET, Davidson JM, Malladi VS, Strattan JS, Hitz BC, Gabdank I, Narayanan AK, Ho M, Lee BT, Rowe LD, Dreszer TR, Roe GR, Podduturi NR, Tanaka F, Hilton JA, Cherry JM. Principles of metadata organization at the ENCODE data coordination center. Database (Oxford). 2016;2016 PubMed Central PMCID: PMC4792520.

3. I have continued my career as a large-scale data coordinator by leading a data coordination team in the field of single-cell genomics. The nascent nature of this field, with constant development of new data formats and analysis tools, has made it vital for open data resources to establish and disseminate standards, as my team has done while leading curation for CZ CELLxGENE Discover. We have also served as data stewards for researchers funded through specific Chan Zuckerberg Initiative Cell Science grant programs, partnering with them to stay tightly connected to data generation and reuse cases while also ensuring that the value of their data is maximized through open data sharing.
   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.
   b. Sarfraz I, Wang Y, Shastry A, Teh WK, Sokolov A, Herb BR, Creasy HH, Virshup I, Dries R, Degatano K, Mahurkar A, Schnell DJ, Madrigal P, Hilton J, Gehlenborg N, Tickle T, Campbell JD. MAMS: matrix and analysis metadata standards to facilitate harmonization and reproducibility of single-cell data. Genome Biol. 2024 Aug 1;25(1):205. PubMed Central PMCID: PMC11292877.