

SHANNON E. ELLIS

9500 Gilman Drive • CSB 243 • La Jolla, CA 92161
sellis@ucsd.edu • shanellis.com • (619) 736-0820

EDUCATION

- 2010—2016 **PhD** | Human Genetics
Johns Hopkins University School of Medicine • Baltimore, MD
Multi-omic Data Provide a More Complete Understanding of the Autistic Brain
- 2006—2010 **BS** | Biology & Spanish
King's College • Wilkes-Barre, PA
-

TEACHING

- 2022—Present **Associate Teaching Professor**
2018—2022 **Assistant Teaching Professor**
University of California San Diego • La Jolla, CA
Instructor for COGS 9 Introduction to Data Science, COGS 108 Data Science in Practice, COGS 18 Introduction to Python, COGS 137 Practical Data Science in R, and DSC 180A & DSC 180B Data Science Capstone Sequence - Genetics.
- Spring 2018 **Co-Instructor** | Health Data Analysis Practicum
Johns Hopkins University • Baltimore, MD
- Fall 2017 **Teaching Assistant** | Public Health Biostatistics
Johns Hopkins University • Baltimore, MD
- Summer 2017 **Instructor** | Genetics
Johns Hopkins Center for Talented Youth • Baltimore, MD
- Spring 2013 **Teaching Assistant** | Advanced Topics in Human Genetics
Johns Hopkins School of Medicine • Baltimore, MD

Mentoring

- 2021 Shubham Kulkarni | *Undergraduate Student*, UC San Diego
Arelly Vasquez | *Undergraduate Student*, UC San Diego
Bob Zhang | *Undergraduate Student*, UC San Diego
Jitarth Sheth | *Undergraduate Student*, UC San Diego
- 2020 Clara Ortez | *Undergraduate Student*, UC San Diego
Sirui Tao | *Undergraduate Student*, UC San Diego
Owen Zhang | *Undergraduate Student*, UC San Diego
Simran Bhatia | *Undergraduate Student*, UC San Diego
Brendan Wong | *Undergraduate Student*, UC San Diego
Jianan Liang | *Undergraduate Student*, UC San Diego
Qizoxuan (Josh) Wang | *Undergraduate Student*, UC San Diego
Ritwik Sinha | *Undergraduate Student*, UC San Diego
Sashwath Gollamudi | *Undergraduate Student*, UC San Diego
Weilun Yao | *Undergraduate Student*, UC San Diego

- 2019 Serene Issa | *Undergraduate Student*, UC San Diego
 Pratyush Khurana | *Undergraduate Student*, UC San Diego
 Anran Li | *Undergraduate Student*, UC San Diego
 Lauren Awaya | *Undergraduate Student*, UC San Diego
 Ayushi Gupta | *Undergraduate Student*, UC San Diego
 Zheng Tang | *Undergraduate Student*, UC San Diego
 Jaidev Mirchandani | *Undergraduate Student*, UC San Diego
 Qian Yin | *Undergraduate Student*, UC San Diego
 Weilun Yao | *Undergraduate Student*, UC San Diego
 Emily Le | *Undergraduate Student*, UC San Diego
- 2016 Augusto Ramirez | *Undergraduate Student*, Johns Hopkins
 Elizabeth Vincent | *Graduate Student*, Johns Hopkins
- 2015–2016 Rebecca Panitch | *Undergraduate Student*, Johns Hopkins
- 2014 Heather Wick | *Graduate Student*, Johns Hopkins
 Edward Pang | *Undergraduate Student*, Johns Hopkins
- 2013 James Miller | *Undergraduate Student*, Johns Hopkins

Thesis Committees

- 2019–2022 Sean Kross | *PhD Student*, Cognitive Science, UC San Diego
 2020 Rachel Goodridge | *MS Student*, Biology, UC San Diego
-

RESEARCH & PROJECTS

- 2021 **Co-Author** | Tidyverse Skills for Data Science in R
 Tidyverse Skills for Data Science in R (Wright, Ellis, Hicks & Peng, 2021, *Leanpub* and *Coursera*) helps readers develop insights from data with tidy tools. Through its chapters readers are introduced to the tidyverse before being taught how to import, wrangle, visualize, and model data in the tidyverse. This textbook includes case examples and case studies throughout to help teach and demonstrate how one can use the tidyverse suite of packages to carry out their data science tasks.
- 2017–2020 **Curriculum Lead** | Cloud-Based Data Science
 clouddatascience.org • Baltimore, MD
 Developed educational content for 13 free, online courses to teach the basics of data science to individuals with no prior computational experience. Supported development of technology built to support content development. Developed and implemented in-person tutoring and support program to help individuals who have earned their GED learn the basics of data science and obtain entry-level data science jobs.
- 2016–2018 **Postdoctoral Fellow** | *Biostatistics*
 Johns Hopkins University Bloomberg School of Public Health • Baltimore, MD
 Advisor: Jeffrey T. Leek, Ph.D.
 Used gene expression data from 70,000 human samples and machine learning to predict critical phenotype information. Developed content for Cloud-based Data Science.
- 2010–2016 **Graduate Student** | *Institute of Genetic Medicine*
 Johns Hopkins University School of Medicine • Baltimore, MD
 Advisor: Dan E Arking, Ph.D.

Analyzed RNA-Sequencing data and identified an upregulation of activated M2 microglia genes in autism brains. Developed a method to guide RNA-Sequencing analysis using eQTLs as a gold standard. Identified significant DNA hypermethylation at cytosines outside of the classically-studied CpG context in autism brains utilizing bisulfite sequencing.

Peer-reviewed Publications

- 2022 Donoghue T, Voytek B, and **Ellis SE**. Course Materials for Data Science in Practice. *The Journal of Open Source Education*. (2022)
- Lau S, Eldridge J, **Ellis SE**, Fraenkel A, Langlois M, Rampure S, Tiefenbruck J, and Guo P. The Challenges of Evolving Technical Courses at Scale: Four Case Studies of Updating Large Data Science Courses. (2022) *ACM Conference on Learning @ Scale (L@S)*.
- 2021 Hubbard J.A., Hoffman M.A., **Ellis S.E.**, Sobolesky P.M., Smith B.E., Suhandynata R.T., Sones E.G., Sanford S.K., Umlauf A., Huestis M.A., Grelotti D.J., Grant I., Marcotte R.D., and Fitzgerald R.L. (2021). Biomarkers of Recent Cannabis Use in Blood, Oral Fluid and Breath. *Journal of Analytical Toxicology*.
- Hoffman M.A., Hubbard J.A., Sobolesky P.M., Smith B.E., Suhandynata R.T., Sanford S., Sones E.G., **Ellis S.E.**, Umlauf A., Huestis M.A., Grelotti D.J., Grant I., Marcotte T.D., and Fitzgerald R.L. (2021). Blood and Oral Fluid Cannabinoid Profiles of Frequent and Occasional Cannabis Smokers. *Journal of Analytical Toxicology*.
- 2020 Donoghue T., Voytek B., and **Ellis S.E.** (2020). Teaching Creative and Practical Data Science at Scale. *Journal of Statistics and Data Science Education*.
- 2019 Madugundu A.K., Hyun Na C., Nirujogi R.S., Reunuse S., Kim KP, Burns KH, Langmead B., **Ellis S.E.**, Collado-Torres L., Halushka M.K., Kim M., and Pandey A. (2019). Integrated Transcriptomic and Proteomic Analysis of Primary Human Umbilical Vein Endothelial cells. *Proteomics*.
- 2018 **Ellis S.E.**, Collado-Torres L., Jaffe A., Leek J.T. (2018). Improving the value of public RNA-seq expression data by phenotype prediction. *Nucleic Acids Research*.
- Ellis S.E.** and Leek J.T. (2018). How to share data for collaboration. *The American Statistician*.
- 2017 Andrews S., **Ellis S.E.**, Bakulski K., Sheppard B., Croen L., Hertz-Piccioto I., schaffer C., Feinberg A., Arking D.E., Ladd-Acosta C., and Fallin M. (2017). Cross-tissue integration of genetic and epigenetic data offers insight into autism spectrum disorder. *Nature Communications*.
- Ellis S.E.**, Gupta S., Moes A., West A.B., Arking D.E. (2017). Exaggerated CpH Methylation in the Autism-Affected Brain. *Molecular Autism*.
- Collado-Torres L., Nellore A., Kammers K., **Ellis S.E.**, Taub M.A., Hansen K.D., Jaffe A.E., Langmead B., Leek J. (2017). Reproducible RNA-seq analysis using *recount2*. *Nature Biotechnology*.
- 2016 **Ellis S.E.**, Panitch R., West A.B., Arking D.E. (2016). Transcriptome Analysis of Cortical Tissue Reveals Shared Sets of Down-Regulated Genes in Autism and Schizophrenia. *Translational Psychiatry*.
- 2015 Huang C, Haritunians T, Okou DT, Cutler DJ, Zwick ME, Taylor KD, Datta LW, Maranville JC, Liu Z, **Ellis S**, Chopra P, Alexander JS, Baldassano RN, Cross RK, Dassopoulos T, Dhere TA, Duerr RH, Hanson JS, Hou JK, Hussain SZ,

Isaacs KL, Kachelries KE, Kader H, Kappelman MD, Katz J, Kellermayer R, Kirschner BS, Kuemmerle JF, Kumar A, Kwon JH, Lazarev M, Mannon P, Moulton DE, Osuntokun BO, Patel A, Rioux JD, Rotter JI, Saeed S, Scherl EJ, Silverberg MS, Silverman A, Targan SR, Valentine J, Wang MH, Simpson CL, Bridges SL, Kimberly RP, Rich SS, Cho JH, Di Rienzo A, Kao LW, McGovern DP, Brant SR, and Kugathasan S. (2015). Characterization of Genetic Loci That Affect Susceptibility to Inflammatory Bowel Diseases in African Americans. *Gastroenterology*.

2014 Gupta, S., **Ellis, S.E.**, Ashar, F.N., Moes, A., Bader, J.S., West, A.B., and Arking, D.E. (2014). Transcriptome Analysis Reveals Deregulation of Innate Immune Response Genes and Neuronal Activity-Dependent Genes in Autism. *Nature Communications*.

2013 **Ellis, S.E.**, Gupta, S., Ashar, F.N., Bader, J.S., West, A.B., and Arking, D.E. (2013). RNA-Seq optimization with eQTL gold standards. *BMC Genomics* 14, 892.

Invited Talks & Workshops

2022 **Ellis, S.E.** (Feb 23, 2022) Structuring & Managing Group Projects in Large-Enrollment Undergraduate Data Science Course. *Toronto Reproducibility Workshop* - Toronto, Canada.

2021 **Ellis, S.E.** (June 18, 2021) Capstone and Project-Based Data Science Courses at UC San Diego. *National Workshop on Data Science Education - Berkeley, CA*.

2020 **Ellis, S.E.** (Sept 9, 2020) The {Art} of Effective Data Visualization. *Data Science Alliance*.

Ellis, S.E. (Feb 10, 2020). Improving the value of public data with *recount2* and phenotype prediction. *CSSA*.

Ellis, S.E. (Jan 29, 2020). RStudio Cloud in the Classroom. *ASA K-12 Virtual Workshop*.

2019 **Ellis, S.E.** (Jul 29, 2019). Overcoming the barriers of entry into data science for nontraditional learners with cloud computing. *Joint Statistical Meeting*.

Ellis, S.E. (Jun 1, 2019). Teaching Data Science: A Workshop for High School Teachers. *Symposium on Data Science and Statistics*.

Ellis, S.E. (Apr 2, 2019). Making a Case for Why Programming Will Save You Time. *Annual Congress in Clinical Mass Spectrometry*.

2018 **Ellis, S.E.** (Nov 28, 2018). Designing MOOCs to Democratize Data Science Education. *Design@Large*.

Ellis, S.E. (Feb 1, 2018). Improving the value of public data with *recount2* and phenotype prediction. *Joint Statistical Meeting*.

Ellis, S.E. (Feb 1, 2018). Improving the value of public data with *recount2* and phenotype prediction. *University of Washington Biostatistics Seminar*.

2017 **Ellis, S.E.** (Oct 11, 2017). Improving the value of public data with *recount2* and phenotype prediction. *Genomics and Bioinformatics Symposium*.

Ellis, S.E. (June 5-7, 2017). *In silico* phenotyping to improve the usefulness of public data. *Graybill Conference in Statistical Genomics and Genetics*.

Ellis, S.E. (March 26-31, 2017). *In silico* phenotyping to improve the usefulness of public data. *BIRS Statistical and Computational Challenges in Large Scale Molecular Biology*.

Ellis, S.E. (Feb 16-17, 2017). Increasing the value of public data with *in-silico* phenotyping. *Statistical and Algorithmic Challenges in Microbiome Data Analysis*.

PROFESSIONAL DEVELOPMENT & SERVICE

Pedagogy Training

Spring 2017 **Participant** | Johns Hopkins Teaching Institute

Completed an intensive workshop in pedagogy focused on enhancing instruction in higher education. Topics covered included: Teaching as Scholarship, Inclusive Classrooms, Active Learning, Planning a Course, Assessment, and Evaluation.

Scientific Meetings Attended

2019 Symposium on Data Science and Statistics
2018—2019 Joint Statistical Meeting
2017 Graybill Conference in Statistical Genomics and Genetics
2017 rOpenSci Unconference
2017 BIRS Statistical & Computational Challenges in Large Scale Molecular Biology
2017 Statistical and Algorithmic Challenges in Microbiome Data Analysis
2010—2015 American Society for Human Genetics
2013—2014 Society for Neuroscience
2009—2010 American Society for Microbiology

Volunteering

2020 **Essay Judge** | Stemanities
2019, 2020 **Career Panelist** | UCSD EPIC Postdoc Bootcamp
2015—Present **Volunteer** | Icing Smiles
2015—2020 **DNA Day Essay Judge** | The American Society of Human Genetics
2016—2018 **Volunteer** | Science Outside the Lines
2017 **Abstract Reviewer** | The American Society of Human Genetics

Service & Leadership

2021—Present **Academic Integrity Review Board** | UC San Diego
2021—Present **LPSOE Promotion Standards Committee** | Cognitive Science Department
2020—Present **DEI Committee** | Halicioğlu Data Science Institute
2018—Present **Undergraduate Curriculum Committee** | Cognitive Science Department
2019—2020 **LPSOE in Data Science Search Committee** | Halicioğlu Data Science Institute
2019—2020 **LPSOE in Machine Learning Search Committee** | Cognitive Science Department
2013—2016 **Graduate Student Representative** | Institute of Genetic Medicine
2013—2016 **Student Leader** | Barton Childs Lecture Planning Committee
2011—2015 **Committee Leader** | Graduate Program Student Recruitment
Spring 2014 **Student Leader** | McKusick Lecture Planning Committee

Journal Editor

Journal of Open Source Education | <https://jose.theoj.org/>

Journal Reviewer

European Journal of Human Genetics | www.nature.com/ejhg/

F1000 | <https://f1000.com/>

Harvard Data Science Review | <https://hdsr.mitpress.mit.edu/>

Journal of Statistics and DS Education | <https://www.tandfonline.com/toc/ujse21/>

Nature Biotechnology | www.nature.com/nbt/

Peer J | peerj.com/

PLoS Computational Biology | journals.plos.org/ploscompbiol/

PLoS One | journals.plos.org/plosone

Stat | <https://onlinelibrary.wiley.com/journal/20491573>

Professional Societies

2018–2020 **Member** | American Statistical Association

2010–2017 **Member** | The American Society of Human Genetics

—

HONORS AND AWARDS

2017 **Teaching Award** | Johns Hopkins Center for Talented Youth

2008 **Undergraduate Research Fellowship** | National Science Foundation

2006–2010 **Presidential Scholarship** (a full academic scholarship) | King's College