Samantha Klasfeld, PhD

Computational Biologist, Bioinformatics

✉ sjk314@gmail.com | 561-271-9934 | https://sklasfeld.netlify.app | Cambridge, MA

Summary

Bioinformatician with deep expertise in statistical genetics and large-scale NGS data analysis. Proven track record of integrating cutting-edge data analysis, statistical inference, and reproducible workflows. Known for a collaborative mindset, creative problem-solving, and translating genomic data into meaningful insights to drive scientific and clinical impact.

Technical Skills

- **Programming languages:** Python, R, command line interface (Bash), slurm, git, SQL, Apache spark, Matlab, Perl, Java
- **Workflow development**: published ChIP-seq methods, experience with Github, Docker, conda, Nextflow, Snakemake
- **NGS Analysis**: Burden testing, GWAS, ChIP-seq (proteomic profiling), bulk RNA-seq, MNase-seq, bulk ATAC-seq
- **Genomics software**: BWA, Bowtie, Bowtie2, STAR, kallisto**,** MACS2, UMAP, Blacklist, DiffBind, DESeq2, EdgeR, sleuth**,** DANPOS2, DANPOS3**,** FASTQC, bedtools, samtools, deeptools, MEME, HOMER, AHA, PBJelly, miRDeep2, miRDeep*, miREAP, miRdentify, Centroidfold, CAP-miRSeq
- **Databases**: UK Biobank, All of Us, FinnGen, gnomAD, GTex, ClinVar, dbNSFP, Clingen, Genebass, dbSNP
- **Genetics software**: PLINK, VCFtools, bcftools, VEP, LOFTEE, MutPred2, ESM1b, AlphaMissense

Professional Experience

Independent Contractor, Sep 2025 — Present
Genscience, Remote
- Conduct subject-level genetic analysis to investigate the role of genomic variation in disease risk and progression within the UK Biobank Research Analysis Platform (UKB-RAP) and All of Us Researcher Workbench
- Apply statistical and bioinformatics methods to these large-scale datasets to interpret findings and contribute to hypothesis-driven study design

Postdoctoral Fellow, Feb 2022 – Feb 2025 (3 years)
Pfizer, Internal Medicine Research Unit and Rare Disease Research Unit, Cambridge, MA
- Organized genetic and medical data leveraged from self-reports, electronic health records, and ~500K exomes in the UK Biobank and generated version-controlled code with git
- Calculated polygenic risk scores (PRS) of individuals in UK Biobank from genome-wide association studies (GWAS)
- Enhanced understanding of the genetic architecture of cardiomyopathy, a rare cardiac condition, by leveraging advanced rare variant effect prediction techniques and statistical modeling of disease penetrance, onset, and severity
- Delivered actionable genetic insights for drug target development and contributed to a first-author publication
- Utilized Olink proteogenomic data in UK Biobank to evaluate the utility of incorporating missense variant predictions into rare variant association analyses
- Articulated my work to multidisciplinary audiences including nonsubject matter experts (eg. biologists, leadership)

Graduate Student Researcher, Aug 2015 – Feb 2022 (6 years)
University of Pennsylvania, Philadelphia, PA
- Developed Greenscreen method to accurately identify and mask false-positive ChIP-seq peaks with efficacy in *Arabidopsis, Drosophila,* and *Oryza sativa* resulting in 1st author publication
- Designed and orchestrated Next Generation Sequencing (NGS) data analysis (eg. ChIP-Seq, RNA-seq, MNase-seq) to measure the activity and outcomes of mobile and DNA binding proteins functional in cellular reprogramming*,* resulting in three 2nd author publications

Postbaccalaureate Fellow, Aug 2014 – June 2015 (1 year)
National Institutes of Health, National Human Genome Research Institute, Bethesda, MD
- Collaborated with computational biologists to analyze the genomes of *Hydractinia echinata* and *Hydractinia symbiolongicarpus* by designing and implementing analysis of novel small RNAs during head regeneration for publication
- Employed a domain architecture search to identify TRAF immune genes in eukaryotic genomes to study cnidarian evolution

Summer Undergraduate Researcher, Jun 2013 – Aug 2013 (3 months)
J. Craig Venter Institute, Rockville, MD
- Investigated novel microRNAs in parasitic nematode *Rotylenchulus reniformis*
- Assessed AHA and PBJelly as gap closing and scaffold tools for *Medicago truncatula* PacBio sequencing data

Summer Undergraduate Researcher, Jun 2012 – Aug 2012 (3 months)
Boyce Thompson Institute, Ithaca, NY
- Improved methods to classify known and novel virus siRNA sequences from sweet potato samples more effectively

Undergraduate Researcher, Aug 2010 – May 2014 (4 years)
Cornell University, Plant Breeding Department, Ithaca, NY
- Compared Genotype by Sequencing data from different restriction enzymes to *Cucurbita pepo* transcripts with BLAST
- Developed and tested KASP markers for relevant genes in peppers to make available for breeding work
- Extracted DNA and used PCR/sequencing to link gene mutation to pepper phenotype.

Leadership Experience

2024 – Present	Women in Bioinformatics in Boston Web Developer (https://boston-wib.org) and Event Organizer
2024 – Present	Volunteer for Cambridge Tutoring Plus (weekly during the school year)
2019 – 2021	University of Pennsylvania Data Science Group kaggle team VP (online **machine learning** competitions with a focus on text data). For more information: https://www.kaggle.com/sklasfeld
2020 – 2021	Weekly lessons to teach lab mates ChIP-seq and RNA-seq analysis (22+ lectures)
2020 – 2021	University of Pennsylvania Data Science Group Journal Club VP
2016 – 2020	Choreographed python programming course for graduate students (6-8 lectures/year)
2018	Genomics and Computational Biology Graduate Program Representative
2016	Content developer: discoveringthegenome.org (2 modules)
2015	Philadelphia Science Education Academy teacher:1st grade science education (4 classes)

Education

University of Pennsylvania | Perelman School of Medicine | Philadelphia, PA
Doctor of Philosophy in Genomics and Computational Biology, 2015-2021

Cornell University | College of Agriculture and Life Sciences | Ithaca, NY
Bachelor of Science in Biology, Concentration in Computational Biology, 2010-2014

Publications

- **Klasfeld, S.**, Knutson, K.A., Miller, M. R., Fauman, E., Berghout, J., Moccia, R., & Kim, H. I. (2025). Common genetic modifiers influence cardiomyopathy susceptibility among the carriers of rare pathogenic variants. *Human Genetics and Genomics Advances*.
- **Klasfeld, S.**, Roulé, T., & Wagner, D. (2022). Greenscreen: A simple method to remove artifactual signals and enrich for true peaks in genomic datasets including ChIP-seq data. The Plant Cell, 34(12), 4795-4815.
- Zhu, Y., **Klasfeld, S.**, & Wagner, D. (2021). Molecular regulation of plant developmental transitions and plant architecture via PEPB family proteins–an update on mechanism of action. *Journal of Experimental Botany*.
- Jin, R., **Klasfeld, S.**, ... & Wagner, D., (2021). LEAFY is a pioneer transcription factor and licenses cell reprogramming to floral fate. *Nature communications*, 12(1), 1-14.
- Zhu, Y., **Klasfeld, S.**, …, & Wagner, D. (2020). TERMINAL FLOWER 1-FD complex target genes and competition with FLOWERING LOCUS T. *Nature communications*, 11(1), 1-12.
- Xiao, J., …, **Klasfeld, S.**, ..., & Wagner, D., (2017). Cis and trans determinants of epigenetic silencing by Polycomb repressive complex 2 in Arabidopsis. *Nature genetics*, 49(10), p.1546.
- Schnitzler, C. E., …, **Klasfeld, S.**, ..., & Baxevanis, A. D. (2023). The genome of the colonial hydroid *Hydractinia* reveals their stem cells utilize a toolkit of evolutionarily shared genes with all animals. *BioRxiv*.

Featured Presentations

- **Klasfeld, S.**, Fauman, E.B., Miller, M.R., Moccia, R., & Kim, H.I. (2024, November 5). Coding cis pQTLs from proteogenomic data allow evaluation of the performance of missense variant effect predictions and the utility of their application to rare variant association analyses. [Poster]. American Society of Human Genetics, Denver, CO.
- **Klasfeld, S.**, Moccia, R., Knutson, A., Berghout, J., & Kim, H.I. (2023, November 2). Genetic modifiers illuminate clinical phenotype penetrance in carriers of rare cardiomyopathy-causal variants [PowerPoint slides]. Pfizer Worldwide Research, Development and Medical Postdoc Symposium, Cambridge, MA.
- **Klasfeld, S.**, Moccia, R., & Berghout, J. (2022, November 2). Genetic modifiers illuminate frequency and penetrance of clinical and subclinical phenotypes in carriers of rare cardiomyopathy-causal variants [Poster]. Pfizer Worldwide Research, Development and Medical Postdoc Symposium, Cambridge, MA.
- **Klasfeld, S.**, & Wagner, D. (2019, June 12). Improving resolution of protein binding sites by filtering conserved ultra-high signal in Arabidopsis [Poster]. University of Pennsylvania Genomics and Computational Biology Graduate Group Retreat, The College of Physicians of Philadelphia, Philadelphia, PA.
- **Klasfeld, S.**, & Wagner, D. (2018, November 15). Resolving the mechanism of PRC2 during plant stress [Poster]. University of Pennsylvania Epigenetics Retreat, Independence Seaport Museum, Philadelphia, PA.
- **Klasfeld, S.**, & Wagner, D. (2017, October 13). The characterization and discovery of novel Polycomb and Trithorax Recruitment Elements in plants [Poster]. University of Pennsylvania Epigenetics Retreat, Citizens Bank Park, Philadelphia, PA.
- **Klasfeld, S.,** Baxevanis, A., & Schnitzler, C. (2015, October 22). Small RNA discovery and expression during regeneration in the colonial hydroid, Hydractinia [PowerPoint slides]. Bioinformatics Scientific Interest Group (BSIG), NIH.

Selected Awards

- Outstanding Poster Award at NIH Post Baccalaureate Poster Day 2015
- Bioinformatics-Scientific Interest Group 3rd Annual Poster Day Winner (First Place)
- Honorable Mention of the 2016 National Science Foundation (NSF) Graduate Research Fellowship