

Sage M. Wright

Senior Bioinformatics Engineer • she/her

sagemwright@gmail.com • sage-wright.com
github.com/sage-wright • www.linkedin.com/in/sagewright
ORCID 0000-0002-0735-5504 • Google Scholar

EXPERIENCE

Theiagen Genomics — Fully Remote 2022–present

Senior Bioinformatics Engineer (Engineer III) 2026–present

- Lead and mentor developers, resolving technical blockers and delivering new features for public health bioinformatics pipelines on schedule.
- Lead development of Theiagen’s bioinformatics solutions, overseeing code review, testing, and documentation to improve scalability, reliability, and maintainability.

Senior Bioinformatics Developer (Scientist II) 2024–2026

- Mentored developers through formal and informal training, dedicating 20% of work time to technical guidance and support.
- Developed and maintained Google Cloud Workflows and Batch jobs to orchestrate data processing across BigQuery, Looker, and other Google Cloud services.
- Provided technical support to Theiagen partners, resolving at least 85% of reported bioinformatics issues within partner-specific service-level timelines.
- Led development of novel bioinformatics software ([tbp-parser](#)), evaluated emerging technologies for adoption, and presented findings in quarterly internal or external knowledge-sharing sessions.

Bioinformatics Scientist 2022–2024

- Developed and maintained bioinformatics pipelines for analyzing viral, bacterial, and eukaryotic pathogens (see the [Theiagen GitHub repositories](#)).
- Supported public health scientists worldwide with bioinformatics analyses, troubleshooting, and rapid-response bug fixes.
- Trained public health scientists in Georgia, Ghana, Mozambique, Oman, Ukraine, and the United States to perform bioinformatics analyses on locally collected pathogen data.

Pennsylvania Department of Health, Bureau of Laboratories — Exton, PA 2021–2022

Bioinformatician

- Established cloud-based bioinformatics infrastructure for rapid analysis of pathogens of public health concern, enabling outbreak response and epidemiological surveillance.
- Analyzed SARS-CoV-2 genomes for quality control and lineage assignment, and generated phylogenetic trees to track viral evolution.
- Contributed SARS-CoV-2 genomes and associated metadata to national and international databases, such as NCBI and GISAID.

SKILLS

Languages & workflows: Python, Bash, R, WDL, Nextflow

Additional languages: C++, C#, Java

Cloud & data: Google Cloud (Artifact Registry, Batch, BigQuery, Cloud Run, Looker, Workflows), Terra.bio, Docker, Git

Sequencing data: Whole-genome, RNA-seq, ChIP-seq, Hi-C, Oxford Nanopore long-read, Bionano optical mapping

Public health genomics: Pathogen surveillance, lineage assignment, phylogenetics, quality control, NCBI and GISAID submission

RESEARCH

Doctoral Researcher — Penn State College of Medicine, Hershey, PA 2017–2021

Dr. Sinisa Dovat Lab 2020–2021

- Analyzed heterochromatin histone-modification occupancy following CK2 inhibition using RNA-seq and ChIP-seq.

Dr. James Broach Lab 2019–2020

- Identified genetic variants associated with disease phenotypes in a newborn infant using whole-genome sequencing data from the family.
- Analyzed data produced from optical mapping technology (Bionano).
- Identified genomic variants in HPV-integrated cancers.

Dr. Feng Yue Lab 2017–2019

- Identified structural variants using data from Hi-C, Bionano optical mapping, and Oxford Nanopore long-read sequencing.
- Maintained and updated the databases underlying the [3D Genome Browser](#).
- Developed a machine-learning algorithm to identify topologically associating domains (TADs) in 3D genomic data.

PROFESSIONAL SERVICE

Public Health Alliance for Genomic Epidemiology (PHA4GE) — Bioinformatics Pipelines and Visualization Working Group 2021–present

Contributed to the development of the following PHA4GE guidance documents:

- [Ten Best Practices for Public Health Bioinformatics Pipelines](#) (2024)
- [Identifying SARS-CoV-2 Recombinants](#) (2023)
- [Bioinformatics Solutions for Mpox Genomic Analysis](#) (2022)

EDUCATION

M.Sc. Bioinformatics and Genomics — Penn State University, Hershey, PA 2017–2021

Advanced to candidacy and passed comprehensive examinations; completed with an M.Sc. after my dissertation advisor left the university.

Thesis: [The Effect of CK2 Inhibition on Heterochromatin in AML](#)

Thesis advisor: Dr. Sinisa Dovat

B.S. Bioinformatics — Brigham Young University, Provo, UT 2017

SELECTED PUBLICATIONS

- Elder J, Ryder R, Panditrao M, Grosgebauer K, Katz R, Tello L, Carroll E, Borthwick D, Kaur C, Smith R, Shiau V, Wheeler W, Reilly E, Myers J, Nelson L, Lim E, Arunleung P, Baylis E, Gilliam S, Hennesy-Burt T, Bregman B, Silver E, Kapsak C, **Wright SM**, Leon T, Bell J, Morales C, Wadford DA. 2026. The SARS-CoV-2 Integrated Genomic Epidemiology Database (IGED): Linking viral genomes with patient-level metadata to advance statewide genomic surveillance in California. *medRxiv*. doi:10.64898/2026.05.14.26353263
- Wadford DA, Baumrind N, Baylis EF, Bell JM, Bouchard EL, Crumpler M, Foote EM, Gilliam S, Glaser CA, Hacker JK, Ledin K, Messenger SL, Morales C, Smith EA, Sevinsky JR, Corbett-Detig RB, DeRisi J, Jacobson K, **the COVIDNet Consortium**. 2023. Implementation of California COVIDNet – a multi-sector collaboration for statewide SARS-CoV-2 genomic surveillance. *Front Public Health* 11:1249614. doi:10.3389/fpubh.2023.1249614
- Ambrosio FJ, Scribner MR, **Wright SM**, Otieno JR, Doughty EL, Gorzalski A, Siao DD, Killian S, Hua C, Schneider E, Tran M, Varghese V, Libuit KG, Pandori M, Sevinsky JR, Hess D. 2023. TheiaEuk: A Species-Agnostic Bioinformatics Workflow for Fungal Genomic Characterization. *Front Public Health* 11:1198213. doi:10.3389/fpubh.2023.1198213
- Libuit KG, Doughty EL, Otieno JR, Ambrosio F, Kapsak CJ, Smith EA, **Wright SM**, Scribner MR, Petit RA, Mendes CI, Huergo M, Legacki G, Loreth C, Park DJ, Sevinsky JR. 2023. Accelerating bioinformatics implementation in public health. *Microb Genom* 9:001051. doi:10.1099/mgen.0.001051
- Smith EA, Libuit KG, Kapsak CJ, Scribner MR, **Wright SM**, Bell J, Morales C, Crumpler M, Messenger S, Hacker JK, Ledin K, Glaser C, Jacobson K, Sevinsky JR, Wadford DA. 2023. Pathogen genomics in public health laboratories: successes, challenges, and lessons learned from California's SARS-CoV-2 Whole-Genome Sequencing Initiative, California COVIDNet. *Microb Genom* 9:001027. doi:10.1099/mgen.0.001027
- Salameh TJ, Wang X, Song F, Zhang B, **Wright SM**, Khunsriraksakul C, Ruan Y, Yue F. 2020. A supervised learning framework for chromatin loop detection in genome-wide contact maps. *Nat Commun* 11:3428. doi:10.1038/s41467-020-17239-9
- Wright SM**, Jensen SL, Cockriel KL, Davis B, Tschanz JT, Munger RG, Corcoran CD, Kauwe JSK. 2019. Association study of rs3846662 with Alzheimer's disease in a population-based cohort: the Cache County Study. *Neurobiol Aging* 84:242.e1–242.e6. doi:10.1016/j.neurobiolaging.2019.03.004
- Wright SM**, Carroll C, Walters A, Newell PD, Chaston JM. 2017. Genome sequence of *Leuconostoc citreum* DmW_111, isolated from wild *Drosophila*. *Genome Announc* 5:e00507-17. doi:10.1128/genomeA.00507-17

COMMUNITY

Violinist —First violin section	2017–2026
Attended weekly rehearsals and performed concerts with each orchestra.	
Narragansett Bay Symphony Community Orchestra, East Providence, RI	2022–2026
West Shore Symphony Orchestra, Mechanicsburg, PA	2017–2021