Sean Davis, MD, PhD

2754 Thornbrook Road, Ellicott City, MD 21042 | seandavi@gmail.com | seandavi.github.io

Google Scholar | i10-index: 98 | h-index: 66 | 47455 citations (20240 since 2021; h-index 41 since 2021)

EDUCATION AND PROFESSIONAL EXPERIENCE

2021–Present **Professor and Rifkin Chair of Cancer Informatics**
Departments of Medicine and Biomedical Informatics, University of Colorado Anschutz School of Medicine

2017–2020 **Senior Associate Scientist**
Center for Cancer Research, National Cancer Institute

2009–2016 **Staff Scientist**
Center for Cancer Research, National Cancer Institute

2007–2008 **Research Fellow**
National Cancer Institute

2005–2007 **Research Fellow**
National Human Genome Research Institute

2002–2005 **Clinical Fellow**
Combined Johns Hopkins and National Cancer Institute Pediatric Hematology/Oncology Fellowship

1999–2002 **Pediatric Resident**
Children's Hospital and Regional Medical Center, University of Washington

1993–1999 **MD**
University of Pittsburgh School of Medicine

1995–1997 **PhD**
University of Pittsburgh Graduate School of Public Health, Department of Human Genetics

1989–1993 **B.S.E., With Honors**
Princeton University, Mechanical and Aerospace Engineering

LEADERSHIP AND SERVICE

Institutional Strategy

April 2026 Author, *Implementing AI in Academic Medicine* — 196-page open AI governance reference distributed to CU Anschutz Chancellor, CIO, and Dean (amc-ai-governance)

2026-Present Member, Education and Research AI Working Groups, Anschutz Medical Campus

2022-Present Member, UCHealth Clinical Intelligence Steering Committee

2021–Present Founding Member and Faculty, Department of Biomedical Informatics, University of Colorado School of Medicine (chaired by Casey Greene)

2021–Present Associate Director for Informatics and Data Science, University of Colorado Comprehensive Cancer Center

2022–Present Member, University of Colorado Health System Clinical Intelligence Committee

2021–2023 Acting Lead, Center for Health AI software engineering team, University of Colorado School of Medicine

2021 Interviewer for Medical Oncology Fellowship program, University of Colorado Department of Medicine

2021–2022 Chair, open rank faculty search committee, Center for Health AI, University of Colorado School of Medicine

2021–2024 Deputy Director, Center for Health Artificial Intelligence, University of Colorado Anschutz School of Medicine

2021–2023 Led development of Alpine, the Anschutz Medical Campus high-performance computing environment (700+ researcher users in three years), in partnership with CU Anschutz OIT, School of Medicine IT, and CU Boulder Research Computing

2021–2024 Member, Compass Clinical Data Warehouse Steering Committee, University of Colorado School of Medicine

2021–Present Member, Governance Committee, University of Colorado Cancer Center

2021–Present Member, Executive Committee, University of Colorado Cancer Center

2015–Present CCR Representative to CBIIT Strategic Planning Committee

2010–Present Steering Committee, NCI Center of Excellence in Integrative Cancer Biology and Genomics

2010–2012 Scientific Liaison, Center for Cancer Research Bioinformatics Core

2009–2010 Chair, Center for Cancer Research Bioinformatics Planning and Implementation Committee

2009–2016 Sequencing Facility Steering and Review Committee, Center for Cancer Research, NCI

National & International Policy

February 2026–Present Member, NCCN Digital Oncology Forum (nominated by CU Cancer Center Associate Director for Clinical Research)

April 2026 Project Lead, NCI-DOE Genesis Mission (MOSSAIC): Advanced Computing and AI Foundations for Childhood and Young Adult Cancer Research

June 2021 NIH/CSR special review of DSI-Africa U54 Center Grants

2020–Present Member of NHGRI AnVIL project Executive Committee

2016–2020 Founding Member, NIH Data Science Special Interest Group (>600 members)

2012–2019 Founding Member, NIH High Performance and Scientific Computing Working Group

2009–2020 Contributed to development of Biowulf, NIH's institutional high-performance computing environment (top-100 supercomputer worldwide)

2017 NIH Intramural Representative, NIH Data Commons working group

December 2017 Co-organizer, NIH Hour of Code, Data Science Special Interest Group, NIH, Bethesda, MD

November 2017 NIH Representative to US Department of Agriculture, Blueprint for USDA Efforts in Agricultural Animal Genomics, Beltsville, MD

August 2017 NIH Intramural Representative, NIH Data Commons Review Committee

February 2017 Organizer, NIH/NIST Medical Devices Cybersecurity Workshop, Bethesda, MD

2017–2020 Cancer Moonshot Blue Ribbon Panel Implementation Working Group, National Cancer Data Ecosystem

January 2017 Organizer, Globus Data Platform Hackathon and Workshop, NIH, Bethesda, MD

January 2017 NCI Representative, NHLBI TopMed Data Commons Planning Workshop

2016–2017 NCI Representative, NIH Data Commons Reference Dataset Working Group

2016 Presidential Subcommittee on AI and Machine Learning, Cancer Moonshot Initiative

December 2015 NCI representative and panel member, FDA Informatics and Precision Medicine Workshop

2015–2017 NCI Cancer Cloud Pilot, Lead for Intramural Research Program evaluation and implementation

2014–2018 NIH and NCI Genomic Data Sharing Policy Implementation working groups

2014 NCI Center for Cancer Genomics Genomic Data Commons (GDC) Review Committee

2012–2016 High Throughput Molecular Data Working Group, National Cancer Institute

Community Infrastructure & Governance

2025–2026 Organizing Committee, Bioconductor Annual Conference (BioC2026)

April 2026 Co-Organizer, Bioconductor Spatial Data and Image Analysis Hackathon, Venice, Italy

January 2026 Co-Organizer, NVIDIA × Carnegie Mellon University Biomedical AI Hackathon, Pittsburgh, PA

2008–2021 Organizer/Co-Organizer, Bioconductor Annual Meeting and Developer Conference (annual; Fred Hutchinson, Boston, Stanford, Seattle, Toronto, NYU, virtual)

2008–2020 Bioconductor Technical Advisory Board — shared responsibility for ongoing technical leadership of the Bioconductor Project

May 2019 Co-organizer, Artificial Intelligence for Data Reuse, an NSF-funded symposium, Carnegie Mellon University, Pittsburgh, PA

October 2018 Organizer, Informatics Training and Education for Cancer Centers, New Orleans, LA

May 2018 Co-organizer, Renal Cell Cancer Hackathon, in collaboration with sv.ai and Google, San Francisco, CA

July 2016 Co-organizer, Frontiers of Predictive Oncology and Computing Symposium, Washington, DC

January 2016 Co-organizer, NCBI Genomics and Bioinformatics Hackathon

January 2014 Organizer and Instructor, Bioinformatics Summer Course, Ribeirão Preto Medical School, University of São Paulo, Brazil

April 2008 Organizer, European Bioconductor Developer Conference, Lausanne, Switzerland

TEACHING AND TRAINING

Recurring Courses

2024–Present **Frontiers of AI in Medicine** · *Course Director*
University of Colorado School of Medicine (medical student course)

2013–Present **Statistical Methods for Functional Genomics** · *Senior Instructor and Course Organizer*
Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
Annual two-week intensive; topics include exome sequencing, methylation arrays, CGH, public data access, and data integration

2005–2012 **AACR Molecular Biology in Clinical Oncology** · *Instructor*
Aspen & Snowmass, CO (annual, except 2008)

2015 **Harvard School of Engineering and Applied Science: CS290 Extreme Computing** · *Course Organizer and Faculty*
Harvard University, Cambridge, MA

2014 **Bioinformatics Summer Course** · *Organizer and Instructor*
Ribeirão Preto Medical School, University of São Paulo, Brazil

Spring 2010 **Microarray Data Analysis Using R and Bioconductor** · *Instructor*
Department of Biostatistics, Bioinformatics, and Biomathematics, Georgetown University

Mentorship and Advising

2021–Present **Faculty Mentor, Medical Student Research Track**
University of Colorado School of Medicine

2021–Present **Faculty Member, Computational Biosciences Graduate Training Program**
University of Colorado School of Medicine

2015–2018 **PhD Thesis Committee** · *Vincent Laufer*
University of Alabama (graduated 2018)

Summer 2016 **Undergraduate Research Mentor** · *Rosa Choe, computer science*
UC Berkeley (graduated 2019)

Summer 2015 **Undergraduate Research Mentor** · *Olivia Zhang, computer science*
Princeton University (graduated 2020)

Summer 2013 **Undergraduate Research Mentor** · *Peter Hansen, biology*
Cornell University (graduated 2018)

Selected Invited Presentations

August 2025 **Bioconductor: Cancer Genomics — Integrative and Scalable Solutions**
Bioconductor Annual Conference (BioC2025) / NCI ITCR session

November 2023 **Artificial Intelligence & Machine Learning**
Monroe Hospital, Monroe, Colorado

September 2022 **Artificial Intelligence & Machine Learning**
Department of Medicine Grand Rounds, University of Colorado Anschutz Medical Campus

December 2021 **Orchestra: a Cloud Platform and Ecosystem for Scalable Computational Training**
VIB/ELIXIR Framework workshop on cloud migration, Ghent, Belgium

November 2021 **Bioinformatics, HPC and AI: Convergence, Perspectives and the Future for Biomedical Applications**
Expert panel, Supercomputer 2021 (SC21), St. Louis, MO

September 2021 **Translation of AI and ML to a Learning Healthcare System**
Data Management Track, Bio-IT World, Boston, MA

April 2021 **Leveraging Public Genomic Data Resources to Improve Science**
Virtual hands-on workshop for 100+ researchers, Human Heredity and Health in Africa (H3Africa) and H3Bionet

September 2020 **sars2pack: An open software package providing fit-for-use COVID-19 data resources**
Cornell Weill School of Medicine, New York, NY

July 2018 **Cloud computing approaches to genomic data science**
American Statistical Association, Joint Statistical Meeting, Vancouver, Canada

January 2018 **A Data Ecosystem for Biomedical Big Data**
Grand Rounds, Wake Forest School of Medicine, Winston-Salem, NC

January 2017 **A cloud-based data ecosystem for cancer research**
Dana Farber Cancer Institute, Boston, MA

October 2016 **Democratizing access to Big Cancer Data**
Midatlantic Bioinformatics Conference, University of Pennsylvania, Philadelphia, PA

November 2011 **High-resolution Views of the Cancer Genome Using Next-Generation Sequencing Approaches**
Lombardi Cancer Center, Georgetown University

March 2012 **Featured Speaker, Bioinformatics for Medical Genetics Symposium**
American College of Medical Genetics, Charlotte, NC

February 2012 **Advanced R and Bioconductor Workshop on High-Throughput Genetic Analysis**
Fred Hutchinson Cancer Research Center, Seattle, WA

February 2009 **Genomics for the Pediatrician: An Overview of Genomics Technologies**
Pediatric Grand Rounds, Oklahoma University Health Sciences Center, Oklahoma City, OK

PROGRAM OF RESEARCH — SELECTED PUBLICATIONS

Full publication list: Google Scholar (h-index 66, 47455 citations; h-index 41 and 20240 citations since 2021). Software tools downloaded by 50,000+ users/year. Complete bibliography at end of document.

Data Infrastructure & Bioconductor

Orchestrating high-throughput genomic analysis with Bioconductor
Nature Methods, 2015 · 4191 citations — *Landmark reference describing the Bioconductor project infrastructure*

Bioconductor's Computational Ecosystem for Genomic Data Science in Cancer
Methods in Molecular Biology, 2025

MultiAssayExperiment: Software for the Integration of Multi-Omics Experiments in Bioconductor
Cancer Research, 2017 · 139 citations

GenomicDataCommons: a Bioconductor Interface to the NCI Genomic Data Commons
Bioinformatics, 2017 · 11 citations

GenomicSuperSignature facilitates interpretation of RNA-seq experiments through robust, efficient comparison to public databases
Nature Communications, 2022 · 13 citations

BugSigDB captures patterns of differential abundance across a broad range of host-associated microbial signatures
Nature Biotechnology, 2024 · 22 citations

GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor
Bioinformatics, 2007 · 3394 citations — *Canonical R/Bioconductor interface to NCBI GEO; widely-used community resource*

RCircos: an R package for Circos 2D track plots
BMC Bioinformatics, 2013 · 966 citations

Translational Research: Cancer Genomics & Oncology

Pan-cancer genome and transcriptome analyses of 1,699 paediatric leukaemias and solid tumours
Nature, 2018 · 988 citations — *International pediatric cancer consortium*

Super-enhancers delineate disease-associated regulatory nodes in T cells
Nature, 2015 · 420 citations

Meta-analysis of 22,710 human microbiome metagenomes defines an oral-to-gut microbial enrichment score and associations with host health and disease
Nature Communications, 2025 · 6 citations

Integration of 168,000 samples reveals global patterns of the human gut microbiome
Cell, 2025 · 108 citations
A pivotal role for Wnt antagonists in constraining Wnt activity to promote digit joint specification
bioRxiv (preprint), 2025
ASXL3 Is a Novel Pluripotency Factor in Human Respiratory Epithelial Cells and a Potential Therapeutic Target in Small Cell Lung Cancer
Cancer Research, 2017 · 33 citations
Lineage of origin in rhabdomyosarcoma informs pharmacological response
Genes & Development, 2014 · 125 citations

Policy, Governance & Data Ethics

The Common Fund Data Ecosystem (CFDE)
bioRxiv (preprint), 2026 — *Senior and corresponding author; descriptor paper for the NIH Common Fund Data Ecosystem*
Lifecycle-aware evaluation of biomedical data ecosystems
Zenodo (preprint; submitted to Cell Trends Open), 2026 — *Multidimensional evaluation framework for NIH CFDE; preprint at doi.org/10.5281/zenodo.20314128*
From Patient Engagement to Precision Oncology: Leveraging Informatics to Advance Cancer Care
Yearbook of Medical Informatics, 2020 · 20 citations
Resources for Interpreting Variants in Precision Genomic Oncology Applications
Frontiers in Oncology, 2017 · 23 citations
Public data and open source tools for multi-assay genomic investigation of disease
Briefings in Bioinformatics, 2016 · 66 citations
Ten simple rules for large-scale data processing
PLOS Computational Biology, 2022 · 5 citations

Community Building & Training

Learning and teaching biological data science in the Bioconductor community
PLOS Computational Biology, 2025 · 5 citations
Orchestrating a community-developed computational workshop and accompanying training materials
F1000Research, 2018
Exploring Cancer in Colorado using a novel data platform: the ECCO experience
medRxiv, 2026 — *Community data platform for cancer disparities research*

SELECTED SOFTWARE

Software tools downloaded by **50,000+ users per year**.

scriptorium
Agentic operating system for scholarly writing — citation audit, reviewer simulation, structural analysis.
quartobot
Quarto-based scholarly publishing toolchain; manubot successor, on PyPI.
fda-approval-app
FDA approval lookup for clinical-trials operations (openFDA + LLM arbiter).
pubmed-grader
Bibliometric dashboard for cancer center portfolio analysis (iCite + Unpaywall + Altmetric).
nextflow_telemetry
Cloud-native observability stack for Nextflow workflow runs.

lifeos-template
Public template for AI-augmented research operating systems.
awesome-single-cell
Community-curated reference list of single-cell software and data resources (3,700+ GitHub stars).
awesome-cancer-variant-resources
Community-maintained catalog of cancer-variant clinical knowledgebases and databases.
OmicIDX
Cloud-native metadata index for 80M+ SRA runs and 8M GEO samples.
GEOquery
Canonical R/Bioconductor interface to NCBI Gene Expression Omnibus.
Genomic Data Commons
R/Bioconductor client for the NCI Genomic Data Commons.
BiocPkgTools
Tools for analyzing and reporting on the Bioconductor package ecosystem.
GenomicSuperSignatures
Interpret RNA-seq experiments via robust comparison to public databases.
SRAdb
SQLite-backed query interface to the NCBI Sequence Read Archive.
GEOmetadb
SQLite-backed query interface to NCBI GEO sample metadata.
biomaRt
Programmatic R interface to BioMart databases (e.g., Ensembl).
RCircos
Circular genome-data visualization in R (CRAN).
methylumi
R/Bioconductor infrastructure for Illumina methylation array analysis.
ACME
Algorithms for calculating microarray enrichment (R/Bioconductor).
ngCGH
Next-gen sequencing copy-number analysis (early HTS-era tool).
sars2pack
Fit-for-use COVID-19 data resources for R (2020 pandemic response).
state-cancer-profiles
Programmatic access to NCI State Cancer Profiles data.

AWARDS AND HONORS

2018	NIH Director's Award
2017	NCI Technology Transfer Award
2016	NCI Technology Transfer Award
2016	U.S. Department of Health and Human Services Director's Award
2015	NIH Director's Award
2012	Staff Scientist/Staff Clinician Travel Award, Center for Cancer Research, NCI
2002–2007	NIH General Loan Repayment Program
2002	Family-Centered Care Award, University of Washington, Children's Hospital of Seattle

1995 W.M. Keck Fellowship for Advanced Scientific Computing

1989 National Merit Scholar

1988 Pennsylvania Governor's School for Science

FUNDING

Current

Comprehensive genetic ancestry annotation of existing human molecular data
R01 Supplement | NIH | 12/01/2025–11/30/2030 | Total: $214K
PI: Alex Krasnitz Role: *Co-Investigator*
Site: Cold Spring Harbor Laboratory
Annotates public human molecular data with ancestry predictions to improve understanding of diversity across human genomics datasets.

Big Data Training for Cancer Research
R25 CA233429 | NIH | 09/01/2024–08/31/2029 | Total: $1615K
PI: Min Zhang Role: *Co-Investigator*
Site: University of Colorado Denver
Supports a two-week intensive, hands-on cancer informatics training workshop for biologists learning large-scale cancer data analysis.

Exploiting Public Metagenomic Data to Uncover Cancer-Microbiome Relationships
R01 CA230551 | NCI | 09/01/2024–08/31/2029 | Total: $2219K
PI: Levi Waldron Role: *Co-Investigator*
Site: University of Colorado Denver
Investigates the role of the human microbiome in cancer using published metagenomic data and returns improved, more reusable resources to the research community.

Integration, Dissemination, and Evaluation (BRIDGE) Center for the NIH Bridge to Artificial Intelligence (BRIDGE2AI) Program
U54 HG012513 | NIH | 07/01/2022–04/30/2026 | Total: $10518K
PI: Monica Cecilia Munoz-Torres Role: *Co-Investigator*
Site: University of Colorado Denver
Coordinates the BRIDGE2AI program to ensure data are ethically sourced, machine-understandable, and interoperable for AI and machine learning applications.

Cancer Genomics: Integrative and Scalable Solutions in R/Bioconductor
U24 CA289073 | NIH | 07/01/2024–06/30/2029 | Total: $5225K
PI: Levi Waldron, Sean Davis Role: *MPI*
Site: University of Colorado Denver
Develops, expands, and sustains Bioconductor software and data resources for cancer genomics research, analysis, and training.

Oculomics as a Biomarker for Comprehensive and Non-Invasive Patient Health Assessment
AWD-AAI2401 | Anschutz Acceleration Initiative, Anschutz Foundation | 07/01/2024–02/28/2029 | Total: $6979K
PI: Sean Davis Role: *Co-Investigator*
Site: University of Colorado Anschutz School of Medicine
Uses ophthalmic imaging and AI to identify non-invasive biomarkers for systemic disease risk and earlier patient assessment.

CONNECT: Collaborative Network for Nurturing Ecosystems of Common Fund Team Science
U54 OD036472 | NIH | 04/01/2024–03/31/2029 | Total: $2937K
PI: Jake Yue Chen, Sean Davis, Casey S. Greene, Peipei Ping, Wei Wang Role: *MPI*
Site: University of Colorado Denver
Builds collaborative infrastructure and coordination capacity for the NIH Common Fund Data Ecosystem so investigators can better reuse and integrate Common Fund resources.

Leveraging whole-exome sequence data from diverse biobanks and cohorts to study rare coding variation in prostate cancer
R01 CA276826 | NCI | 09/01/2023–08/31/2028 | Total: $450K
PI: Christopher Alan Haiman Role: *Co-Investigator*
Site: University of Colorado Denver
Studies rare and common genetic variation associated with prostate cancer susceptibility, progression, and staging using globally diverse biobank cohorts.

University of Colorado Cancer Center Support Grant
P30 CA046934 | NCI | 02/01/2023–01/31/2027 | Total: $4137K
PI: Richard D. Schulick Role: *Co-Investigator*
Site: University of Colorado Denver
Supports the University of Colorado Cancer Center's cancer discovery, translation, patient-centered care, training, and catchment-area impact.

Pending

Genetic Ancestry knowledgebase for human molecular data in SRA
NIH pending | NIH | 12/01/2026–11/30/2030 | Total: $144K
PI: Alex Krasnitz Role: *Co-Investigator*
Site: Cold Spring Harbor Laboratory
Proposes a repository of ancestry annotations for human Sequence Read Archive data to enable ancestry-aware analyses at population scale.

A comprehensive biomedical metadata knowledgebase to democratize data discovery
R24 HG014772 | NHGRI | 12/01/2025–11/30/2029 | Total: $2142K
PI: Arjun Krishnan Role: *Co-Investigator*
Site: University of Colorado Denver
Proposes a metadata knowledgebase to support metadata enrichment, deep search, and machine learning over public biomedical datasets.

A Knowledgebase Streamlining Access to Public Omics Data through Curated Metadata
R24 (pending) | NIH | 12/01/2025–11/30/2029 | Total: $70K
PI: Arjun Krishnan Role: *Co-Investigator*
Site: University of Colorado Denver
Proposes a public knowledgebase that uses controlled vocabularies and ontologies to improve access to and reuse of public omics data.

Completed

Bioconductor: high quality training and support for a worldwide community
Chan Zuckerberg Initiative, Essential Open Source Software for Science, Cycle 4 | 01/01/2022–12/31/2023
PI: Aedin Culhane Role: *Co-Investigator*
Uses technology to extend Bioconductor data science and bioinformatics training efforts to better support diversity, inclusivity, and global reach.

BigRNA: Accelerating Research By Increasing Access and Use of Public Transcriptomic Data
Amazon Research Grant | 06/01/2019–06/01/2020
PI: Sean Davis Role: *PI*
The BigRNA project reprocesses and distributes publicly available transcriptomic data to facilitate data reuse, reanalysis, and machine learning applications.

Symposium on Personal Control of Genomic Data for Research
Cancer Moonshot Initiative | September 2019
PI: Sean Davis, on behalf of the Network for Direct Patient Engagement Role: *PI*
Brought together patients, advocates, lawyers, ethicists, medical care providers, and data scientists to share expertise on personal control of genomic data.

Increasing the value of public transcriptomic data through harmonized processing and metadata
XSEDE (PSC) MCB200027 | 01/30/2020–01/29/2021
PI: Sean Davis Role: *PI*
Allocation to prototype using XSEDE resources to perform large-scale, harmonized omics data processing, including transcriptomics and metagenomics.

Increasing the value of public transcriptomic data through harmonized processing and metadata
Administrative Supplement to U01 CA230551 | 10/01/2021–09/30/2022
PI: Levi Waldron Role: *Co-Investigator*
Adds support for >100,000 homogeneously processed microbiome samples to the public microbiome data collection.

EDITORIAL RESPONSIBILITIES

2015–Present	Editor, f1000research Bioconductor Channel
2010–Present	Associate Editor, BMC Bioinformatics
2009	Book reviewer, CRC Press

Peer Reviewer — selected journals:

- Bioinformatics
- Cancer Research
- Genome Research
- Molecular Cancer Research
- Nucleic Acids Research
- PLOS One
- Gigascience
- BMC Bioinformatics
- Clinical Cancer Research
- Genomics
- Nature Methods
- PLOS Computational Biology
- Genetic Epidemiology
- Briefings in Bioinformatics

PUBLICATIONS

Complete list. Google Scholar link provides citation metrics and open-access links.

Davis, S. (2026). *Implementing AI in Academic Medicine: A guide to governance and principled practice*. https://seandavi.github.io/amc-ai-governance/

Lowery, J. T., Alquaddoomi, F., Rubinetti, V., Burus, T., Jardine, C., Warren, A. C., Walsh, J., Borrayo, E., & Davis, S. (2026, February 4). *Exploring Cancer in Colorado using a novel data platform: the ECCO experience.* https://doi.org/10.64898/2026.02.03.26345489

Abdill, R. J., Graham, S. P., Rubinetti, V., Ahmadian, M., Hicks, P., Chetty, A., McDonald, D., Ferretti, P., Gibbons, E., Rossi, M., Krishnan, A., Albert, F. W., Greene, C. S., Davis, S., & Blekhman, R. (2025). Integration of 168,000 samples reveals global patterns of the human gut microbiome. *Cell*, *188*(4), 1100–1118.e17. https://doi.org/10.1016/j.cell.2024.12.017

Drnevich, J., Tan, F. J., Almeida-Silva, F., Castelo, R., Culhane, A. C., Davis, S., Doyle, M. A., Geistlinger, L., Ghazi, A. R., Holmes, S., Lahti, L., Mahmoud, A., Nishida, K., Ramos, M., Rue-Albrecht, K., Shih, D. J. H., Gatto, L., & Soneson, C. (2025). Learning and teaching biological data science in the Bioconductor community. *PLoS Computational Biology*, *21*(4), e1012925. https://doi.org/10.1371/journal.pcbi.1012925

Huang, Y. P., Harmon, L., Deering-Gardner, E., Ma, X., Harsh, J., Xue, Z., Wen, H., Ramos, M., Davis, S., & Triche, T. J. J. (2025). bamSliceR: a Bioconductor package for rapid, cross-cohort variant and allelic bias analysis. *Bioinformatics Advances*, *5*(1), vbaf98. https://doi.org/10.1093/bioadv/vbaf098

Manghi, P., Antonello, G., Schiffer, L., Golzato, D., Wokaty, A., Beghini, F., Mirzayi, C., Long, K., Gravel-Pucillo, K., Piccinno, G., Gamboa-Tuz, S. D., Bonetti, A., D'Amato, G., Azhar, R., Eckenrode, K., Zohra, F., Giunchiglia, V., Keller, M., Pedrotti, A., … Waldron, L. (2025). Meta-analysis of 22,710 human microbiome metagenomes defines an oral-to- gut microbial enrichment score and associations with host health and disease. *Nature Communications*, *17*(1), 196. https://doi.org/10.1038/s41467-025-66888-1

Ramos, M., Shepherd, L., Sheffield, N. C., Mahmoud, A., Pages, H., Wokaty, A., Righelli, D., Risso, D., Davis, S., Oh, S., Waldron, L., Morgan, M., & Carey, V. (2025). Bioconductor's Computational Ecosystem for Genomic Data Science in Cancer. *Methods in Molecular Biology (Clifton, N.J.)*, *2932*, 1–46. https://doi.org/10.1007/978-1-0716-4566-6_1

Huang, B.-L., Davis, S., Koyama, E., Pacifici, M., & Mackem, S. (2025, July 21). *A pivotal role for Wnt antagonists in constraining Wnt activity to promote digit joint specification.* https://doi.org/10.1101/2025.07.17.665381

Long, K., Gravel-Pucillo, K., Waldron, L., Davis, S., & Oh, S. (2025, December 1). *Large-scale Manual Curation and Harmonization of Metadata from Metagenomic and Cancer Genomic Repositories: Challenges and Solutions.* https://doi.org/10.1101/2025.11.26.689816

Geistlinger, L., Mirzayi, C., Zohra, F., Azhar, R., Elsafoury, S., Grieve, C., Wokaty, J., Gamboa-Tuz, S. D., Sengupta, P., Hecht, I., Ravikrishnan, A., Goncalves, R. S., Franzosa, E., Raman, K., Carey, V., Dowd, J. B., Jones, H. E., Davis, S., Segata, N., … Waldron, L. (2024). BugSigDB captures patterns of differential abundance across a broad range of host-associated microbial signatures. *Nature Biotechnology*, *42*(5), 790–802. https://doi.org/10.1038/s41587-023-01872-y

Oh, S., Gravel-Pucillo, K., Ramos, M., Schatz, M. C., Davis, S., Carey, V., Morgan, M., & Waldron, L. (2024). AnVILWorkflow: A runnable workflow package for Cloud-implemented bioinformatics analysis pipelines. *F1000research*, *13*, 1257. https://doi.org/10.12688/f1000research.155449.1

Fungtammasan, A., Lee, A., Taroni, J., Wheeler, K., Chin, C.-S., Davis, S., & Greene, C. (2022). Ten simple rules for large-scale data processing. *PLoS Computational Biology*, *18*(2), e1009757. https://doi.org/10.1371/journal.pcbi.1009757

Oh, S., Geistlinger, L., Ramos, M., Blankenberg, D., Beek, M. van den, Taroni, J. N., Carey, V. J., Greene, C. S., Waldron, L., & Davis, S. (2022). GenomicSuperSignature facilitates interpretation of RNA-seq experiments through robust, efficient comparison to public databases. *Nature Communications*, *13*(1), 3695. https://doi.org/10.1038/s41467-022-31411-3

Geistlinger, L., Csaba, G., Santarelli, M., Ramos, M., Schiffer, L., Turaga, N., Law, C., Davis, S., Carey, V., Morgan, M., Zimmer, R., & Waldron, L. (2021). Toward a gold standard for benchmarking gene set enrichment analysis. *Briefings in Bioinformatics*, *22*(1), 545–556. https://doi.org/10.1093/bib/bbz158

Griffin, A. C., Topaloglu, U., Davis, S., & Chung, A. E. (2020). From Patient Engagement to Precision Oncology: Leveraging Informatics to Advance Cancer Care. *Yearbook of Medical Informatics*, *29*(1), 235–242. https://doi.org/10.1055/s-0040-1701983

Marie, K. L., Sassano, A., Yang, H. H., Michalowski, A. M., Michael, H. T., Guo, T., Tsai, Y. C., Weissman, A. M., Lee, M. P., Jenkins, L. M., Zaidi, M. R., Pérez-Guijarro, E., Day, C.-P., Ylaya, K., Hewitt, S. M., Patel, N. L., Arnheiter, H., Davis, S., Meltzer, P. S., … Mishra, P. J. (2020). Melanoblast Transcriptome Analysis Reveals Pathways Promoting Melanoma Metastasis. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-019-14085-2

Oh, S., Abdelnabi, J., Al-Dulaimi, R., Aggarwal, A., Ramos, M., Davis, S., Riester, M., & Waldron, L. (2020). HGNChelper: identification and correction of invalid gene symbols for human and mouse. *F1000research*, *9*, 1493. https://doi.org/10.12688/f1000research.28033.2

Collado-Torres, L., Burke, E. E., Peterson, A., Shin, J., Straub, R. E., Rajpurohit, A., Semick, S. A., Ulrich, W. S., Price, A. J., Valencia, C., Tao, R., Deep-Soboslay, A., Hyde, T. M., Kleinman, J. E., Weinberger, D. R., & Jaffe, A. E. (2019). Regional Heterogeneity in Gene Expression, Regulation, and Coherence in the Frontal Cortex and Hippocampus across Development and Schizophrenia. *Neuron*, *103*(2), 203–216.e8. https://doi.org/10.1016/j.neuron.2019.05.013

Gopaulakrishnan, S., Pollack, S., Stubbs, B. J., Pages, H., Readey, J., Davis, S., Waldron, L., Morgan, M., & Carey, V. (2019). restfulSE: A semantically rich interface for cloud-scale genomics with Bioconductor. *F1000research*, *8*, 21. https://doi.org/10.12688/f1000research.17518.1

Razmara, A., Ellis, S. E., Sokolowski, D. J., Davis, S., Wilson, M. D., Leek, J. T., Jaffe, A. E., & Collado-Torres, L. (2019). *recount-brain : a curated repository of human brain RNA-seq datasets metadata*.

Su, S., Carey, V. J., Shepherd, L., Ritchie, M., Morgan, M. T., & Davis, S. (2019). BiocPkgTools: Toolkit for mining the Bioconductor package ecosystem. *F1000research*, *8*, 752. https://doi.org/10.12688/f1000research.19410.1

Davis, S., Ramos, M., Shepherd, L., Turaga, N., Geistlinger, L., Morgan, M. T., Haibe-Kains, B., & Waldron, L. (2018). Orchestrating a community-developed computational workshop and accompanying training materials. *F1000research*, *7*, 1656. https://doi.org/10.12688/f1000research.16516.1

Ma, X., Liu, Y., Liu, Y., Alexandrov, L. B., Edmonson, M. N., Gawad, C., Zhou, X., Li, Y., Rusch, M. C., Easton, J., Huether, R., Gonzalez-Pena, V., Wilkinson, M. R., Hermida, L. C., Davis, S., Sioson, E., Pounds, S., Cao, X., Ries, R. E., … Zhang, J. (2018). Pan-cancer genome and transcriptome analyses of 1,699 paediatric leukaemias and solid tumours. *Nature*, *555*(7696), 371–376. https://doi.org/10.1038/nature25795

Liu-Chittenden, Y., Jain, M., Gaskins, K., Wang, S., Merino, M. J., Kotian, S., Kumar Gara, S., Davis, S., Zhang, L., & Kebebew, E. (2017). RARRES2 functions as a tumor suppressor by promoting beta-catenin phosphorylation/degradation and inhibiting p38 phosphorylation in adrenocortical carcinoma. *Oncogene*, *36*(25), 3541–3552. https://doi.org/10.1038/onc.2016.497

Morgan, M. T., & Davis, S. R. (2017). *GenomicDataCommons: a Bioconductor Interface to the NCI Genomic Data Commons*.

Ramos, M., Schiffer, L., Re, A., Azhar, R., Basunia, A., Rodriguez, C., Chan, T., Chapman, P., Davis, S. R., Gomez-Cabrero, D., Culhane, A. C., Haibe-Kains, B., Hansen, K. D., Kodali, H., Louis, M. S., Mer, A. S., Riester, M., Morgan, M., Carey, V., & Waldron, L. (2017). Software for the Integration of Multiomics Experiments in Bioconductor. *Cancer Research*, *77*(21), e39–e42. https://doi.org/10.1158/0008-5472.CAN-17-0344

Shukla, V., Rao, M., Zhang, H., Beers, J., Wangsa, D., Wangsa, D., Buishand, F. O., Wang, Y., Yu, Z., Stevenson, H. S., Reardon, E. S., McLoughlin, K. C., Kaufman, A. S., Payabyab, E. C., Hong, J. A., Zhang, M., Davis, S., Edelman, D., Chen, G., … Schrump, D. S. (2017). ASXL3 Is a Novel Pluripotency Factor in Human Respiratory Epithelial Cells and a Potential Therapeutic Target in Small Cell Lung Cancer. *Cancer Research*, *77*(22), 6267–6281. https://doi.org/10.1158/0008-5472.CAN-17-0570

Tsang, H., Addepalli, K., & Davis, S. R. (2017). Resources for Interpreting Variants in Precision Genomic Oncology Applications. *Frontiers in Oncology*, *7*, 214. https://doi.org/10.3389/fonc.2017.00214

Zhu, F., Willette-Brown, J., Song, N.-Y., Lomada, D., Song, Y., Xue, L., Gray, Z., Zhao, Z., Davis, S. R., Sun, Z., Zhang, P., Wu, X., Zhan, Q., Richie, E. R., & Hu, Y. (2017). Autoreactive T Cells and Chronic Fungal Infection Drive Esophageal Carcinogenesis. *Cell Host & Microbe*, *21*(4), 478–493.e7. https://doi.org/10.1016/j.chom.2017.03.006

Comas-Garcia, M., Davis, S. R., & Rein, A. (2016). On the Selective Packaging of Genomic RNA by HIV-1. *Viruses*, *8*(9), E246. https://doi.org/10.3390/v8090246

Hakim, F. T., Memon, S., Jin, P., Imanguli, M. M., Wang, H., Rehman, N., Yan, X.-Y., Rose, J., Mays, J. W., Dhamala, S., Kapoor, V., Telford, W., Dickinson, J., Davis, S., Halverson, D., Naik, H. B., Baird, K., Fowler, D., Stroncek, D., … Gress, R. E. (2016). Upregulation of IFN-Inducible and Damage-Response Pathways in Chronic Graft-versus-Host Disease. *Journal of Immunology (Baltimore, Md. : 1950)*, *197*(9), 3490–3503. https://doi.org/10.4049/jimmunol.1601054

Kannan, L., Ramos, M., Re, A., El-Hachem, N., Safikhani, Z., Gendoo, D. M. A., Davis, S., Gomez-Cabrero, D., Castelo, R., Hansen, K. D., Carey, V. J., Morgan, M., Culhane, A. C., Haibe-Kains, B., & Waldron, L. (2016). Public data and open source tools for multi-assay genomic investigation of disease. *Briefings in Bioinformatics*, *17*(4), 603–615. https://doi.org/10.1093/bib/bbv080

Mathe, E., & Davis, S. D. (Eds.). (2016). *Statistical Genomics: Methods and Protocols* (1st ed.). Springer New York. https://doi.org/10.1007/978-1-4939-3578-9

Sorber, R., Teper, Y., Abisoye-Ogunniyan, A., Waterfall, J. J., Davis, S., Killian, J. K., Pineda, M., Ray, S., McCord, M. R., Pflicke, H., Burkett, S. S., Meltzer, P. S., & Rudloff, U. (2016). Whole Genome Sequencing of Newly Established Pancreatic Cancer Lines Identifies Novel Somatic Mutation (c.2587G>A) in Axon Guidance Receptor Plexin A1 as Enhancer of Proliferation and Invasion. *PloS One*, *11*(3), e149833. https://doi.org/10.1371/journal.pone.0149833

Zhang, H., Meltzer, P. S., & Davis, S. R. (2016). caOmicsV: an R package for visualizing multidimensional cancer genomic data. *BMC Bioinformatics*, *17*, 141. https://doi.org/10.1186/s12859-016-0989-6

Boufraqech, M., Nilubol, N., Zhang, L., Gara, S. K., Sadowski, S. M., Mehta, A., He, M., Davis, S., Dreiling, J., Copland, J. A., Smallridge, R. C., Quezado, M. M., & Kebebew, E. (2015). miR30a inhibits LOX expression and anaplastic thyroid cancer progression. *Cancer Research*, *75*(2), 367–377. https://doi.org/10.1158/0008-5472.CAN-14-2304

Figueroa, J. D., Yang, H., Garcia-Closas, M., Davis, S., Meltzer, P., Lissowska, J., Horne, H. N., Sherman, M. E., & Lee, M. (2015). *Integrated analysis of DNA methylation, immunohistochemistry and mRNA expression, data identifies a methylation expression index (MEI) robustly associated with survival of ER-positive breast cancer patients*. *150*(2), 457–466.

Huber, W., Carey, V. J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B. S., Bravo, H. C., Davis, S., Gatto, L., Girke, T., Gottardo, R., Hahne, F., Hansen, K. D., Irizarry, R. A., Lawrence, M., Love, M. I., MacDonald, J., Obenchain, V., Oles, A. K., … Morgan, M. (2015). Orchestrating high-throughput genomic analysis with Bioconductor. *Nature Methods*, *12*(2), 115–121. https://doi.org/10.1038/nmeth.3252

Huber, W., Carey, V., Davis, S., Hansen, K. D., & Morgan, M. (2015). The Bioconductor channel in F1000Research. *F1000research*, *4*, 217. https://doi.org/10.12688/f1000research.6758.2

Mirabello, L., Koster, R., Moriarity, B. S., Spector, L. G., Meltzer, P. S., Gary, J., Machiela, M. J., Pankratz, N., Panagiotou, O. A., Largaespada, D., Wang, Z., Gastier-Foster, J. M., Gorlick, R., Khanna, C., Toledo, S. R. C. de, Petrilli, A. S., Patino-Garcia, A., Sierrasesumaga, L., Lecanda, F., … Savage, S. A. (2015). A Genome-Wide Scan Identifies Variants in NFIB Associated with Metastasis in Patients with Osteosarcoma. *Cancer Discovery*, *5*(9), 920–931. https://doi.org/10.1158/2159-8290.CD-15-0125

Shakhova, O., Cheng, P., Mishra, P. J., Zingg, D., Schaefer, S. M., Debbache, J., Hausel, J., Matter, C., Guo, T., Davis, S., Meltzer, P., Mihic-Probst, D., Moch, H., Wegner, M., Merlino, G., Levesque, M. P., Dummer, R., Santoro, R., Cinelli, P., & Sommer, L. (2015). Antagonistic cross-regulation between Sox9 and Sox10 controls an anti- tumorigenic program in melanoma. *PLoS Genetics*, *11*(1), e1004877. https://doi.org/10.1371/journal.pgen.1004877

Vahedi, G., Kanno, Y., Furumoto, Y., Jiang, K., Parker, S. C. J., Erdos, M. R., Davis, S. R., Roychoudhuri, R., Restifo, N. P., Gadina, M., Tang, Z., Ruan, Y., Collins, F. S., Sartorelli, V., & O'Shea, J. J. (2015). Super-enhancers delineate disease-associated regulatory nodes in T cells. *Nature*, *520*(7548), 558–562. https://doi.org/10.1038/nature14154

Abraham, J., Nunez-Alvarez, Y., Hettmer, S., Carrio, E., Chen, H.-I. H., Nishijo, K., Huang, E. T., Prajapati, S. I., Walker, R. L., Davis, S., Rebeles, J., Wiebush, H., McCleish, A. T., Hampton, S. T., Bjornson, C. R. R., Brack, A. S., Wagers, A. J., Rando, T. A., Capecchi, M. R., … Keller, C. (2014). Lineage of origin in rhabdomyosarcoma informs pharmacological response. *Genes & Development*, *28*(14), 1578–1591. https://doi.org/10.1101/gad.238733.114

Ellis, R. J., Wang, Y., Stevenson, H. S., Boufraqech, M., Patel, D., Nilubol, N., Davis, S., Edelman, D. C., Merino, M. J., He, M., Zhang, L., Meltzer, P. S., & Kebebew, E. (2014). Genome-wide methylation patterns in papillary thyroid cancer are distinct based on histological subtype and tumor genotype. *The Journal of Clinical Endocrinology and Metabolism*, *99*(2), E329–337. https://doi.org/10.1210/jc.2013-2749

Kang, Z., Yu, Y., Zhu, Y. J., Davis, S., Walker, R., Meltzer, P. S., Helman, L. J., & Cao, L. (2014). Downregulation of IGFBP2 is associated with resistance to IGF1R therapy in rhabdomyosarcoma. *Oncogene*, *33*(50), 5697–5705. https://doi.org/10.1038/onc.2013.509

Reinhold, W. C., Varma, S., Sousa, F., Sunshine, M., Abaan, O. D., Davis, S. R., Reinhold, S. W., Kohn, K. W., Morris, J., Meltzer, P. S., Doroshow, J. H., & Pommier, Y. (2014). NCI-60 whole exome sequencing and pharmacological CellMiner analyses. *PloS One*, *9*(7), e101670. https://doi.org/10.1371/journal.pone.0101670

Waterfall, J. J., Arons, E., Walker, R. L., Pineda, M., Roth, L., Killian, J. K., Abaan, O. D., Davis, S. R., Kreitman, R. J., & Meltzer, P. S. (2014). High prevalence of MAP2K1 mutations in variant and IGHV4-34-expressing hairy-cell leukemias. *Nature Genetics*, *46*(1), 8–10. https://doi.org/10.1038/ng.2828

Abaan, O. D., Polley, E. C., Davis, S. R., Zhu, Y. J., Bilke, S., Walker, R. L., Pineda, M., Gindin, Y., Jiang, Y., Reinhold, W. C., Holbeck, S. L., Simon, R. M., Doroshow, J. H., Pommier, Y., & Meltzer, P. S. (2013). The exomes of the NCI-60 panel: a genomic resource for cancer biology and systems pharmacology. *Cancer Research*, *73*(14), 4372–4382. https://doi.org/10.1158/0008-5472.CAN-12-3342

Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., Marshall, K. A., Phillippy, K. H., Sherman, P. M., Holko, M., Yefanov, A., Lee, H., Zhang, N., Robertson, C. L., Serova, N., Davis, S., & Soboleva, A. (2013). NCBI GEO: archive for functional genomics data sets–update. *Nucleic Acids Research*, *41*(Database issue), D991–995. https://doi.org/10.1093/nar/gks1193

Bilke, S., Schwentner, R., Yang, F., Kauer, M., Jug, G., Walker, R. L., Davis, S., Zhu, Y. J., Pineda, M., Meltzer, P. S., & Kovar, H. (2013). Oncogenic ETS fusions deregulate E2F3 target genes in Ewing sarcoma and prostate cancer. *Genome Research*, *23*(11), 1797–1809. https://doi.org/10.1101/gr.151340.112

Gartner, J. J., Parker, S. C. J., Prickett, T. D., Dutton-Regester, K., Stitzel, M. L., Lin, J. C., Davis, S., Simhadri, V. L., Jha, S., Katagiri, N., Gotea, V., Teer, J. K., Wei, X., Morken, M. A., Bhanot, U. K., Chen, G., Elnitski, L. L., Davies, M. A., Gershenwald, J. E., … Samuels, Y. (2013). Whole-genome sequencing identifies a recurrent functional synonymous mutation in melanoma. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(33), 13481–13486. https://doi.org/10.1073/pnas.1304227110

Hirsch, D., Kemmerling, R., Davis, S., Camps, J., Meltzer, P. S., Ried, T., & Gaiser, T. (2013). Chromothripsis and focal copy number alterations determine poor outcome in malignant melanoma. *Cancer Research*, *73*(5), 1454–1460. https://doi.org/10.1158/0008-5472.CAN-12-0928

Kikuchi, K., Taniguchi, E., Chen, H.-I. H., Svalina, M. N., Abraham, J., Huang, E. T., Nishijo, K., Davis, S., Louden, C., Zarzabal, L. A., Recht, O., Bajwa, A., Berlow, N., Suelves, M., Perkins, S. L., Meltzer, P. S., Mansoor, A., Michalek, J. E., Chen, Y., … Keller, C. (2013). Rb1 loss modifies but does not initiate alveolar rhabdomyosarcoma. *Skeletal Muscle*, *3*(1), 27. https://doi.org/10.1186/2044-5040-3-27

Petrini, I., Rajan, A., Pham, T., Voeller, D., Davis, S., Gao, J., Wang, Y., & Giaccone, G. (2013). Whole genome and transcriptome sequencing of a B3 thymoma. *PloS One*, *8*(4), e60572. https://doi.org/10.1371/journal.pone.0060572

Praetorius, C., Grill, C., Stacey, S. N., Metcalf, A. M., Gorkin, D. U., Robinson, K. C., Van Otterloo, E., Kim, R. S. Q., Bergsteinsdottir, K., Ogmundsdottir, M. H., Magnusdottir, E., Mishra, P. J., Davis, S. R., Guo, T., Zaidi, M. R., Helgason, A. S., Sigurdsson, M. I., Meltzer, P. S., Merlino, G., … Steingrimsson, E. (2013). A polymorphism in IRF4 affects human pigmentation through a tyrosinase- dependent MITF/TFAP2A pathway. *Cell*, *155*(5), 1022–1033. https://doi.org/10.1016/j.cell.2013.10.022

Webster, L. R., Provan, P. J., Graham, D. J., Byth, K., Walker, R. L., Davis, S., Salisbury, E. L., Morey, A. L., Ward, R. L., Hawkins, N. J., Clarke, C. L., Meltzer, P. S., & Balleine, R. L. (2013). Prohibitin expression is associated with high grade breast cancer but is not a driver of amplification at 17q21.33. *Pathology*, *45*(7), 629–636. https://doi.org/10.1097/PAT.0000000000000004

Yauk, C. L., Lucas Argueso, J., Auerbach, S. S., Awadalla, P., Davis, S. R., DeMarini, D. M., Douglas, G. R., Dubrova, Y. E., Elespuru, R. K., Glover, T. W., Hales, B. F., Hurles, M. E., Klein, C. B., Lupski, J. R., Manchester, D. K., Marchetti, F., Montpetit, A., Mulvihill, J. J., Robaire, B., … Bishop, J. B. (2013).

Harnessing genomics to identify environmental determinants of heritable disease. *Mutation Research*, *752*(1), 6–9. https://doi.org/10.1016/j.mrrev.2012.08.002

Zhang, H., Meltzer, P., & Davis, S. (2013). RCircos: an R package for Circos 2D track plots. *BMC Bioinformatics*, *14*, 244. https://doi.org/10.1186/1471-2105-14-244

Zhu, Y., Stephens, R. M., Meltzer, P. S., & Davis, S. R. (2013). SRAdb: query and use public next-generation sequencing data from within R. *BMC Bioinformatics*, *14*, 19. https://doi.org/10.1186/1471-2105-14-19

Debbache, J., Zaidi, M. R., Davis, S., Guo, T., Bismuth, K., Wang, X., Skuntz, S., Maric, D., Pickel, J., Meltzer, P., Merlino, G., & Arnheiter, H. (2012). In vivo role of alternative splicing and serine phosphorylation of the microphthalmia-associated transcription factor. *Genetics*, *191*(1), 133–144. https://doi.org/10.1534/genetics.111.135996

Gartner, J. J., Davis, S., Wei, X., Lin, J. C., Trivedi, N. S., Teer, J. K., Meltzer, P. S., Rosenberg, S. A., & Samuels, Y. (2012). *Comparative exome sequencing of metastatic lesions provides insights into the mutational progression of melanoma*. *13*(1), –.

Killian, J. K., Walker, R. L., Bilke, S., Chen, Y., Davis, S., Cornelison, R., Smith, W. I., & Meltzer, P. S. (2012). Genome-wide methylation profiling in archival formalin-fixed paraffin- embedded tissue samples. *Methods in Molecular Biology (Clifton, N.J.)*, *823*, 107–118. https://doi.org/10.1007/978-1-60327-216-2_8

Kouprina, N., Lee, N. C. O., Pavlicek, A., Samoshkin, A., Kim, J.-H., Lee, H.-S., Varma, S., Reinhold, W. C., Otstot, J., Solomon, G., Davis, S., Meltzer, P. S., Schleutker, J., & Larionov, V. (2012). Exclusion of the 750-kb genetically unstable region at Xq27 as a candidate locus for prostate malignancy in HPCX1-linked families. *Genes, Chromosomes & Cancer*, *51*(10), 933–948. https://doi.org/10.1002/gcc.21977

O'Shea, P. J., Kim, D. W., Logan, J. G., Davis, S., Walker, R. L., Meltzer, P. S., Cheng, S.-Y., & Williams, G. R. (2012). Advanced bone formation in mice with a dominant-negative mutation in the thyroid hormone receptor beta gene due to activation of Wnt/beta-catenin protein signaling. *The Journal of Biological Chemistry*, *287*(21), 17812–17822. https://doi.org/10.1074/jbc.M111.311464

Zoppoli, G., Solier, S., Reinhold, W. C., Liu, H., Connelly, J. W. J., Monks, A., Shoemaker, R. H., Abaan, O. D., Davis, S. R., Meltzer, P. S., Doroshow, J. H., & Pommier, Y. (2012). CHEK2 genomic and proteomic analyses reveal genetic inactivation or endogenous activation across the 60 cell lines of the US National Cancer Institute. *Oncogene*, *31*(4), 403–418. https://doi.org/10.1038/onc.2011.283

Grohar, P. J., Woldemichael, G. M., Griffin, L. B., Mendoza, A., Chen, Q.-R., Yeung, C., Currier, D. G., Davis, S., Khanna, C., Khan, J., McMahon, J. B., & Helman, L. J. (2011). Identification of an inhibitor of the EWS-FLI1 oncogenic transcription factor by high-throughput screening. *Journal of the National Cancer Institute*, *103*(12), 962–978. https://doi.org/10.1093/jnci/djr156

Killian, J. K., Bilke, S., Davis, S., Walker, R. L., Jaeger, E., Killian, M. S., Waterfall, J. J., Bibikova, M., Fan, J.-B., Smith, W. I. J., & Meltzer, P. S. (2011). A methyl-deviator epigenotype of estrogen receptor-positive breast carcinoma is associated with malignant biology. *The American Journal of Pathology*, *179*(1), 55–65. https://doi.org/10.1016/j.ajpath.2011.03.022

Martin, M. M., Ryan, M., Kim, R., Zakas, A. L., Fu, H., Lin, C. M., Reinhold, W. C., Davis, S. R., Bilke, S., Liu, H., Doroshow, J. H., Reimers, M. A., Valenzuela, M. S., Pommier, Y., Meltzer, P. S., & Aladjem, M. I. (2011). Genome-wide depletion of replication initiation events in highly transcribed regions. *Genome Research*, *21*(11), 1822–1832. https://doi.org/10.1101/gr.124644.111

Rubin, B. P., Nishijo, K., Chen, H.-I. H., Yi, X., Schuetze, D. P., Pal, R., Prajapati, S. I., Abraham, J., Arenkiel, B. R., Chen, Q.-R., Davis, S., McCleish, A. T., Capecchi, M. R., Michalek, J. E., Zarzabal, L. A., Khan, J., Yu, Z., Parham, D. M., Barr, F. G., … Keller, C. (2011). Evidence for an unanticipated relationship between undifferentiated pleomorphic sarcoma and embryonal rhabdomyosarcoma. *Cancer Cell*, *19*(2), 177–191. https://doi.org/10.1016/j.ccr.2010.12.023

Valenzuela, M. S., Chen, Y., Davis, S., Yang, F., Walker, R. L., Bilke, S., Lueders, J., Martin, M. M., Aladjem, M. I., Massion, P. P., & Meltzer, P. S. (2011). *Preferential Localization of Human Origins of DNA Replication at the 5′-Ends of Expressed Genes and at Evolutionarily Conserved DNA Sequences*. *6*(5), e17308–.

Wei, X., Walia, V., Lin, J. C., Teer, J. K., Prickett, T. D., Gartner, J., Davis, S., Stemke-Hale, K., Davies, M. A., Gershenwald, J. E., Robinson, W., Robinson, S., Rosenberg, S. A., & Samuels, Y. (2011). Exome sequencing identifies GRIN2A as frequently mutated in melanoma. *Nature Genetics*, *43*(5), 442–446. https://doi.org/10.1038/ng.810

Zaidi, M. R., Davis, S., Noonan, F. P., Graff-Cherry, C., Hawley, T. S., Walker, R. L., Feigenbaum, L., Fuchs, E., Lyakh, L., Young, H. A., Hornyak, T. J., Arnheiter, H., Trinchieri, G., Meltzer, P. S., De Fabo, E. C., & Merlino, G. (2011). Interferon-gamma links ultraviolet radiation to melanomagenesis in mice. *Nature*, *469*(7331), 548–553. https://doi.org/10.1038/nature09666

Bolton, K. L., Garcia-Closas, M., Pfeiffer, R. M., Duggan, M. A., Howat, W. J., Hewitt, S. M., Yang, X. R., Cornelison, R., Anzick, S. L., Meltzer, P., Davis, S., Lenz, P., Figueroa, J. D., Pharoah, P. D. P., & Sherman, M. E. (2010). Assessment of automated image analysis of breast cancer tissue microarrays for epidemiologic studies. *Cancer Epidemiology, Biomarkers & Prevention : a Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology*, *19*(4), 992–999. https://doi.org/10.1158/1055-9965.EPI-09-1023

Killian, J. K., Walker, R. L., Suuriniemi, M., Jones, L., Scurci, S., Singh, P., Cornelison, R., Harmon, S., Boisvert, N., Zhu, J., Wang, Y., Bilke, S., Davis, S., Giaccone, G., Smith, W. I. J., & Meltzer, P. S. (2010). Archival fine-needle aspiration cytopathology (FNAC) samples: untapped resource for clinical molecular profiling. *The Journal of Molecular Diagnostics : JMD*, *12*(6), 739–745. https://doi.org/10.2353/jmoldx.2010.090238

Liu, F., Killian, J. K., Yang, M., Walker, R. L., Hong, J. A., Zhang, M., Davis, S., Zhang, Y., Hussain, M., Xi, S., Rao, M., Meltzer, P. A., & Schrump, D. S. (2010). Epigenomic alterations and gene expression profiles in respiratory epithelia exposed to cigarette smoke condensate. *Oncogene*, *29*(25), 3650–3664. https://doi.org/10.1038/onc.2010.129

Morisot, S., Wayne, A. S., Bohana-Kashtan, O., Kaplan, I. M., Gocke, C. D., Hildreth, R., Stetler-Stevenson, M., Walker, R. L., Davis, S., Meltzer, P. S., Wheelan, S. J., Brown, P., Jones, R. J., Shultz, L. D., & Civin, C. I. (2010). High frequencies of leukemia stem cells in poor-outcome childhood precursor-B acute lymphoblastic leukemias. *Leukemia*, *24*(11), 1859–1866. https://doi.org/10.1038/leu.2010.184

John, S., Johnson, T. A., Sung, M.-H., Biddie, S. C., Trump, S., Koch-Paiz, C. A., Davis, S. R., Walker, R., Meltzer, P. S., & Hager, G. L. (2009). Kinetic complexity of the global response to glucocorticoid receptor action. *Endocrinology*, *150*(4), 1766–1774. https://doi.org/10.1210/en.2008-0863

Kauer, M., Ban, J., Kofler, R., Walker, B., Davis, S., Meltzer, P., & Kovar, H. (2009). A molecular function map of Ewing's sarcoma. *PloS One*, *4*(4), e5415. https://doi.org/10.1371/journal.pone.0005415

Killian, J. K., Bilke, S., Davis, S., Walker, R. L., Killian, M. S., Jaeger, E. B., Chen, Y., Hipp, J., Pittaluga, S., Raffeld, M., Cornelison, R., Smith, W. I. J., Bibikova, M., Fan, J.-B., Emmert-Buck, M. R., Jaffe, E. S., & Meltzer, P. S. (2009). Large-scale profiling of archival lymph nodes reveals pervasive remodeling of the

follicular lymphoma methylome. *Cancer Research*, *69*(3), 758–764. https://doi.org/10.1158/0008-5472.CAN-08-2984

Palavalli, L. H., Prickett, T. D., Wunderlich, J. R., Wei, X., Burrell, A. S., Porter-Gill, P., Davis, S., Wang, C., Cronin, J. C., Agrawal, N. S., Lin, J. C., Westbroek, W., Hoogstraten-Miller, S., Molinolo, A. A., Fetsch, P., Filie, A. C., O'Connell, M. P., Banister, C. E., Howard, J. D., … Samuels, Y. (2009). Analysis of the matrix metalloproteinase family reveals that MMP8 is often mutated in melanoma. *Nature Genetics*, *41*(5), 518–520. https://doi.org/10.1038/ng.340

Palmieri, D., Fitzgerald, D., Shreeve, S. M., Hua, E., Bronder, J. L., Weil, R. J., Davis, S., Stark, A. M., Merino, M. J., Kurek, R., Mehdorn, H. M., Davis, G., Steinberg, S. M., Meltzer, P. S., Aldape, K., & Steeg, P. S. (2009). *Analyses of Resected Human Brain Metastases of Breast Cancer Reveal the Association between Up-Regulation of Hexokinase 2 and Poor Prognosis*. *7*(9), 1438–1445.

Palmieri, D., Lockman, P. R., Thomas, F. C., Hua, E., Herring, J., Hargrave, E., Johnson, M., Flores, N., Qian, Y., Vega-Valle, E., Taskar, K. S., Rudraraju, V., Mittapalli, R. K., Gaasch, J. A., Bohn, K. A., Thorsheim, H. R., Liewehr, D. J., Davis, S., Reilly, J. F., … Steeg, P. S. (2009). Vorinostat inhibits brain metastatic colonization in a model of triple- negative breast cancer and induces DNA double-strand breaks. *Clinical Cancer Research : an Official Journal of the American Association for Cancer Research*, *15*(19), 6148–6157. https://doi.org/10.1158/1078-0432.CCR-09-1039

Paoloni, M., Davis, S., Lana, S., Withrow, S., Sangiorgi, L., Picci, P., Hewitt, S., Triche, T., Meltzer, P., & Khanna, C. (2009). Canine tumor cross-species genomics uncovers targets linked to osteosarcoma progression. *BMC Genomics*, *10*, 625. https://doi.org/10.1186/1471-2164-10-625

Rahman, M., Davis, S. R., Pumphrey, J. G., Bao, J., Nau, M. M., Meltzer, P. S., & Lipkowitz, S. (2009). TRAIL induces apoptosis in triple-negative breast cancer cells with a mesenchymal phenotype. *Breast Cancer Research and Treatment*, *113*(2), 217–230. https://doi.org/10.1007/s10549-008-9924-5

Balleine, R. L., Webster, L. R., Davis, S., Salisbury, E. L., Palazzo, J. P., Schwartz, G. F., Cornfield, D. B., Walker, R. L., Byth, K., Clarke, C. L., & Meltzer, P. S. (2008). Molecular grading of ductal carcinoma in situ of the breast. *Clinical Cancer Research : an Official Journal of the American Association for Cancer Research*, *14*(24), 8244–8252. https://doi.org/10.1158/1078-0432.CCR-08-0939

Bhoumik, A., Fichtman, B., Derossi, C., Breitwieser, W., Kluger, H. M., Davis, S., Subtil, A., Meltzer, P., Krajewski, S., Jones, N., & Ronai, Z. (2008). Suppressor role of activating transcription factor 2 (ATF2) in skin cancer. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(5), 1674–1679. https://doi.org/10.1073/pnas.0706057105

Boyle, A. P., Davis, S., Shulha, H. P., Meltzer, P., Margulies, E. H., Weng, Z., Furey, T. S., & Crawford, G. E. (2008). High-resolution mapping and characterization of open chromatin across the genome. *Cell*, *132*(2), 311–322. https://doi.org/10.1016/j.cell.2007.12.014

John, S., Sabo, P. J., Johnson, T. A., Sung, M.-H., Biddie, S. C., Lightman, S. L., Voss, T. C., Davis, S. R., Meltzer, P. S., Stamatoyannopoulos, J. A., & Hager, G. L. (2008). Interaction of the glucocorticoid receptor with the chromatin landscape. *Molecular Cell*, *29*(5), 611–624. https://doi.org/10.1016/j.molcel.2008.02.010

Walsh, T., McClellan, J. M., McCarthy, S. E., Addington, A. M., Pierce, S. B., Cooper, G. M., Nord, A. S., Kusenda, M., Malhotra, D., Bhandari, A., Stray, S. M., Rippey, C. F., Roccanova, P., Makarov, V., Lakshmi, B., Findling, R. L., Sikich, L., Stromberg, T., Merriman, B., … Sebat, J. (2008). Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science (New York, N.Y.)*, *320*(5875), 539–543. https://doi.org/10.1126/science.1155174

Zhu, Y., Davis, S., Stephens, R., Meltzer, P. S., & Chen, Y. (2008). GEOmetadb: powerful alternative search engine for the Gene Expression Omnibus. *Bioinformatics (Oxford, England)*, *24*(23), 2798–2800. https://doi.org/10.1093/bioinformatics/btn520

Birney, E., Stamatoyannopoulos, J. A., Dutta, A., Guigo, R., Gingeras, T. R., Margulies, E. H., Weng, Z., Snyder, M., Dermitzakis, E. T., Thurman, R. E., Kuehn, M. S., Taylor, C. M., Neph, S., Koch, C. M., Asthana, S., Malhotra, A., Adzhubei, I., Greenbaum, J. A., Andrews, R. M., … Jong, P. J. de. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*, *447*(7146), 799–816. https://doi.org/10.1038/nature05874

Davis, S. R., & Meltzer, P. S. (2007). Modeling synovial sarcoma: timing is everything. *Cancer Cell*, *11*(4), 305–307. https://doi.org/10.1016/j.ccr.2007.03.016

Davis, S., & Meltzer, P. S. (2007). GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics (Oxford, England)*, *23*(14), 1846–1847. https://doi.org/10.1093/bioinformatics/btm254

Johnston, J. J., Walker, R. L., Davis, S., Facio, F., Turner, J. T., Bick, D. P., Daentl, D. L., Ellison, J. W., Meltzer, P. S., & Biesecker, L. G. (2007). Zoom-in comparative genomic hybridisation arrays for the characterisation of variable breakpoint contiguous gene syndromes. *Journal of Medical Genetics*, *44*(1), e59. https://doi.org/10.1136/jmg.2006.042473

Kamradt, J., Jung, V., Wahrheit, K., Tolosi, L., Rahnenfuehrer, J., Schilling, M., Walker, R., Davis, S., Stoeckle, M., Meltzer, P., & Wullich, B. (2007). Detection of novel amplicons in prostate cancer by comprehensive genomic profiling of prostate cancer cell lines using oligonucleotide-based arrayCGH. *PloS One*, *2*(8), e769. https://doi.org/10.1371/journal.pone.0000769

Kim, K.-T., Baird, K., Davis, S., Piloto, O., Levis, M., Li, L., Chen, P., Meltzer, P., & Small, D. (2007). Constitutive Fms-like tyrosine kinase 3 activation results in specific changes in gene expression in myeloid leukaemic cells. *British Journal of Haematology*, *138*(5), 603–615. https://doi.org/10.1111/j.1365-2141.2007.06696.x

Sutter, N. B., Bustamante, C. D., Chase, K., Gray, M. M., Zhao, K., Zhu, L., Padhukasahasram, B., Karlins, E., Davis, S., Jones, P. G., Quignon, P., Johnson, G. S., Parker, H. G., Fretwell, N., Mosher, D. S., Lawler, D. F., Satyaraj, E., Nordborg, M., Lark, K. G., … Ostrander, E. A. (2007). *A Single IGF1 Allele Is a Major Determinant of Small Size in Dogs*. *316*(5821), 112–115.

Crawford, G. E., Davis, S., Scacheri, P. C., Renaud, G., Halawi, M. J., Erdos, M. R., Green, R., Meltzer, P. S., Wolfsberg, T. G., & Collins, F. S. (2006). DNase-chip: a high-resolution method to identify DNase I hypersensitive sites using tiled microarrays. *Nature Methods*, *3*(7), 503–509. https://doi.org/10.1038/nmeth888

Crawford, G. E., Holt, I. E., Whittle, J., Webb, B. D., Tai, D., Davis, S., Margulies, E. H., Chen, Y., Bernat, J. A., Ginsburg, D., Zhou, D., Luo, S., Vasicek, T. J., Daly, M. J., Wolfsberg, T. G., & Collins, F. S. (2006). Genome-wide mapping of DNase hypersensitive sites using massively parallel signature sequencing (MPSS). *Genome Research*, *16*(1), 123–131. https://doi.org/10.1101/gr.4074106

Davis, S., & Meltzer, P. S. (2006). Ewing's sarcoma: general insights from a rare model. *Cancer Cell*, *9*(5), 331–332. https://doi.org/10.1016/j.ccr.2006.05.003

Scacheri, P. C., Crawford, G. E., & Davis, S. (2006). [14] Statistics for ChIP-chip and DNase Hypersensitivity Experiments on NimbleGen Arrays. In *Methods in Enzymology* (pp. 270–282). Elsevier. https://doi.org/10.1016/s0076-6879(06)11014-9

Scacheri, P. C., Davis, S., Odom, D. T., Crawford, G. E., Perkins, S., Halawi, M. J., Agarwal, S. K., Marx, S. J., Spiegel, A. M., Meltzer, P. S., & Collins, F. S. (2006). Genome-wide analysis of menin binding provides insights into MEN1 tumorigenesis. *PLoS Genetics*, *2*(4), e51. https://doi.org/10.1371/journal.pgen.0020051

Baird, K., Davis, S., Antonescu, C. R., Harper, U. L., Walker, R. L., Chen, Y., Glatfelter, A. A., Duray, P. H., & Meltzer, P. S. (2005). *Gene Expression Profiling of Human Sarcomas: Insights into Sarcoma Biology*. *65*(20), 9226–9235.

Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A., & Huber, W. (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics (Oxford, England)*, *21*(16), 3439–3440. https://doi.org/10.1093/bioinformatics/bti525

Son, C. G., Bilke, S., Davis, S., Greer, B. T., Wei, J. S., Whiteford, C. C., Chen, Q.-R., Cenacchi, N., & Khan, J. (2005). Database of mRNA gene expression profiles of multiple human organs. *Genome Research*, *15*(3), 443–450. https://doi.org/10.1101/gr.3124505

Meltzer, P., Davis, S., Baird, K., & Chen, Y. (2004). *Are we there yet? Genomic profiling and mechanism in cancer research*.

Davis, S., & Nimgaonkar, V. L. (2001). Impact of overlapping recruitment on linkage analysis of complex disorders: simulation studies. *American Journal of Medical Genetics*, *105*(2), 141–144. https://doi.org/10.1002/1096-8628(2001)9999:9999<::aid-ajmg1162>3.0.co;2-5

Duerr, R. H., Barmada, M. M., Zhang, L., Davis, S., Preston, R. A., Chensny, L. J., Brown, J. L., Ehrlich, G. D., Weeks, D. E., & Aston, C. E. (1998). Linkage and association between inflammatory bowel disease and a locus on chromosome 12. *American Journal of Human Genetics*, *63*(1), 95–100. https://doi.org/10.1086/301929

Davis, S., & Weeks, D. E. (1997). Comparison of nonparametric statistics for detection of linkage in nuclear families: single-marker evaluation. *American Journal of Human Genetics*, *61*(6), 1431–1444. https://doi.org/10.1086/301635

Davis, S., Sobel, E., Marinov, M., & Weeks, D. E. (1997). Analysis of bipolar disorder using affected relatives. *Genetic Epidemiology*, *14*(6), 605–610. https://doi.org/10.1002/(SICI)1098-2272(1997)14:6<605::AID-GEPI9>3.0.CO;2-Y

Julier, C., Delepine, M., Keavney, B., Terwilliger, J., Davis, S., Weeks, D. E., Bui, T., Jeunemaitre, X., Velho, G., Froguel, P., Ratcliffe, P., Corvol, P., Soubrier, F., & Lathrop, G. M. (1997). Genetic susceptibility for human familial essential hypertension in a region of homology with blood pressure linkage on rat chromosome 10. *Human Molecular Genetics*, *6*(12), 2077–2085. https://doi.org/10.1093/hmg/6.12.2077

O'Connell, J. R., Davis, S., & Weeks, D. E. (1997). Analysis of a complex oligogenic disease. *Genetic Epidemiology*, *14*(6), 861–866. https://doi.org/10.1002/(SICI)1098-2272(1997)14:6<861::AID-GEPI50>3.0.CO;2-K

Weeks, D. E., Davis, S., Schroeder, M., & Goldin, L. R. (1997). Nonparametric simulation based linkage statistics for general pedigrees. *The Journal of Rheumatology*, *24*(1), 206–207.

Davis, S., Schroeder, M., Goldin, L. R., & Weeks, D. E. (1996). Nonparametric simulation-based statistics for detecting linkage in general pedigrees. *American Journal of Human Genetics*, *58*(4), 867–880.

Taylor, T. D., Litt, M., Kramer, P., Pandolfo, M., Angelini, L., Nardocci, N., Davis, S., Pineda, M., Hattori, H., Flett, P. J., Cilio, M. R., Bertini, E., & Hayflick, S. J. (1996). Homozygosity mapping of Hallervorden-Spatz syndrome to chromosome 20p12.3-p13. *Nature Genetics*, *14*(4), 479–481. https://doi.org/10.1038/ng1296-479