

Daniel Blankenberg

CURRICULUM VITAE

JANUARY 01, 2025

PERSONAL INFORMATION

Name; last, first middle
Blankenberg, Daniel J
Credentials; MD, PhD, etc.
PhD

Institution & Institute
Cleveland Clinic Research, Cleveland Clinic
Department
Center for Computational Life Science
Office Address & Mail Code
9500 Euclid Avenue / NE29, Cleveland, OH, 44195
Office Phone
216-445-4336
Office Email
blanked2@ccf.org

EDUCATION & TRAINING

Education

School & City, State / Country
Gettysburg College, Gettysburg, PA
Degree
BS, Biochemistry & Molecular Biology / Computer Science
Start-End Dates
08/2000-05/2004

School & City, State / Country
The Pennsylvania State University, University Park, PA
Degree
PhD, Biochemistry, Microbiology and Molecular Biology
Start-End Dates
08/2004-12/2009

PhD Thesis

Title
Galaxy, a web-based framework for the integration of genome analysis
Thesis Committee Members
Anton Nekrutenko, PhD, Webb C. Miller, PhD, Ross C. Hardison, PhD,
Stephan Schuster, PhD, Andrey Krasilnikov, PhD, Scott Selleck, MD, PhD

Post-Graduate Training

Institution & City, State / Country
The Pennsylvania State University, University Park, PA
Position
Postdoctoral Research Associate
Start-End Dates
01/2010-06/2016

PROFESSIONAL APPOINTMENTS

Position
Associate Staff
Institution & Institute
Cleveland Clinic Research, Cleveland Clinic
Department
Center for Computational Life Sciences
City, State / Country
Cleveland, OH
Start-End Dates
12/2024-present

Position
Assistant Staff
Institution & Institute
Cleveland Clinic Research, Cleveland Clinic
Department
Center for Computational Life Sciences
City, State / Country
Cleveland, OH
Start-End Dates
03/2022-11/2024

Position
Assistant Staff

Institution & Institute Cleveland Clinic, Lerner Research Institute
Department Genomic Medicine Institute
City, State / Country Cleveland, OH
Start-End Dates 08/2017-6/2023

Position Senior Research Associate
Institution & Institute The Pennsylvania State University, Eberly College of Science
Department Biochemistry and Molecular Biology
City, State / Country University Park, PA
Start-End Dates 07/2016-07/2017

ACADEMIC APPOINTMENTS

Rank / Department Assistant Professor of Molecular Medicine
Institution & City, State / Country Cleveland Clinic Lerner College of Medicine of Case Western Reserve
University, Cleveland, OH
Start-End Dates 05/2021-present

Rank / Department Adjunct Assistant Professor of Molecular Medicine
Institution & City, State / Country Cleveland Clinic Lerner College of Medicine of Case Western Reserve
University, Cleveland, OH
Start-End Dates 07/2020-05/2021

MEMBERSHIP IN PROFESSIONAL SOCIETIES

Name of Society International Society for Computational Biology
Role / Title Member
Start-End Dates 2011-present

Name of Society Open Bioinformatics Foundation
Role / Title Member
Start-End Dates 2011-present

Name of Society American Society for Human Genetics
Role / Title Member
Start-End Dates 2018-present

Name of Society Association for Computing Machinery
Role / Title Member
Start-End Dates 2019-present

PROFESSIONAL SERVICES

Study Sections/ Grant Review Committees

Organization NIH/NIAID
Section / Committee ZAI1 TC-X (S1) 1 - Early-stage Development of Data Science
Technologies for Infectious and Immune-mediated Diseases (U01);
Exploratory Data Science Methods and Algorithm Development in
Infectious and Immune-mediated Diseases (R21)
Start-End Dates 06/2022

Organization NIH
Section / Committee Biodata Management and Analysis (BDMA) study section; Early Career
Reviewer
Start-End Dates 02/2020

Organization
Section / Committee CCF Alumni Association Postdoc Travel Award review committee
Start-End Dates 2019-2020

Advisory Groups

Organization NIAMS, NIH
Title Member, External Advisory Board for NIH NIAMS P50 Psoriasis-focused Center of Research Translation (P50 AR070590)
Start-End Dates 2019-2022

Organization ACCESS (formerly XSEDE) Campus Champions
Title Campus Champion
Start-End Dates 2018-present

COMMITTEE SERVICE

International

Organization Open Bioinformatics Foundation
Committee Name / Role Member, scientific review committee
Start-End Dates 2015

Organization Galaxy Community Conference
Committee Name / Role Co-organizer of hackathon
Start-End Dates 2015

Organization Galaxy Project
Committee Name / Role Member, Galaxy Community Conference scientific review committee
Start-End Dates 2019-present

Organization Global Galaxy Project
Committee Name / Role Member, Galaxy Project Steering Committee
Start-End Dates 2019 (established)-present

Organization James P. Taylor Foundation for Open Science (JXTX Foundation)
Committee Name / Role Board Member
Start-End Dates 2022-present

Academic

Organization The Pennsylvania State University
Committee Name / Role Member, Administrative Evaluation Committee (AD14) for Biochemistry and Molecular Biology Department Head
Start-End Dates 2016

Organization Case Western Reserve University
Committee Name / Role Member, Yue Xu PhD thesis committee
Start-End Dates 2023-present

Cleveland Clinic

Committee Name / Role Member, Courtney Hershberger PhD thesis committee.
Start-End Dates 2017-2020

Committee Name / Role Member, Microbial Composition/Genomics and Microbial Culture/Engineering Core Directors Search Committee
Start-End Dates 2017-2018

Committee Name / Role Start-End Dates	Member, Brett Graham CCLCM thesis committee 2018-2020
Committee Name / Role Start-End Dates	Member, Center for Microbiome & Human Health Faculty Search Committee 2019-2022
Committee Name / Role Start-End Dates	Member, Abstract review committee for Bumpus Award in Basic Science 2019
Committee Name / Role Start-End Dates	Member, Raoul Wadhwa CCLCM thesis committee 2020-2021
Committee Name / Role Start-End Dates	Lead, Discovery Accelerator Quantum Work Stream 2021-present
Committee Name / Role Start-End Dates	Member, Discovery Accelerator CCF/IBM Joint Steering Committee 2022-present
Committee Name / Role Start-End Dates	Member, CCF Quantitative Data Sciences Think Team 2022
Committee Name / Role Start-End Dates	Member, LRI Computing Committee 2023-present
Committee Name / Role Start-End Dates	Member, FRIC Bioinformatics Faculty Search Committee 2023-2025
Committee Name / Role Start-End Dates	Member, LRI Computing Strategic Visioning and Planning: Financial Sustainability Taskforce 2024-2025
Committee Name / Role Start-End Dates	Member, Research Computing Faculty Search Committee 2024-2025
Committee Name / Role Start-End Dates	Member, Genomic Sciences and System Biology Populations omics Faculty Search Committee 2024-2025

TEACHING ACTIVITIES

Invited Lectures

International

1. Galaxy: Making NGS Analyses Accessible for All. NGx: Evolution of Next-Generation Sequencing. Providence, RI. 09/2010.
2. Galaxy: Making NGS Analyses Accessible for All. BioIT World Europe. Hannover, Germany. 10/2010.
3. Galaxy Tools and Services for Enabling Next-Generation Sequencing Analysis. UCHC Stem Cell and Translational Genomics Cores Joint Training Symposium. Farmington, CT. January 2011.
4. Galaxy for high-throughput NGS analysis. MipTec, Next Generation Sequencing for Research and Clinical Genomics User Group, Basel, Switzerland. September 2011.
5. Galaxy: open source NGS analyses. Keynote at Flemish Training Network Life Sciences Next Generation Sequencing workshop. KU Leuven, Belgium. September 2012.
6. Introduction to Variant Analysis with Galaxy. Royal College of Pathologists of Australasia Genomic Bioinformatics Workshop. Sydney, Australia. October, 2013.

7. Galaxy: a Collaborative Web-based Workbench for the Analysis of Large-Scale Biomedical Data. Brain-CODE Workshop: Building Analytics for Integrated Neuroscience Data. Ontario Brain Institute. Toronto, ON. May 2014.
8. Workshop on Microbiome Data Analysis. Centro de Investigación en Biología Celular y Molecular, Universidad de Costa Rica. San José, San Pedro, Costa Rica. January 2018.
9. Scaling interoperability, reproducibility, and access of informatics resources. AI Medicine & Digital Hospital Joint Symposium. Yonsei University College of Medicine. January 2020.
10. Galaxy Community Update: The State of the Galaxy. Bioinformatics Community Conference. Toronto, Canada (Virtual). July 2020.
11. Galaxy bridge to Jupyter notebooks and RStudio. Freiburg, Germany (Virtual). March 2021.
12. Galaxy Resources for Tool Developers. Global Galaxy Webinar Series (Virtual). May 2021.
13. Galaxy Community Update: The State of the Galaxy. Galaxy Community Conference. Ghent, Belgium (Virtual). July 2021.
14. Perspectives on Challenges in Healthcare and Life Sciences. International Working Group on Quantum Computing in Healthcare Life Science. Cleveland, USA. April 13, 2023.
15. The Cleveland Clinic-IBM Discovery Accelerator: Leveraging quantum computing within an academic medical center. Opportunities of Quantum Computing in the Healthcare Industry and beyond. Garching, Germany (Virtual). September 12, 2023.
16. Quantum Computing for Healthcare and Life Sciences at the Cleveland Clinic. Arab-American Frontiers of Science Engineering and Medicine. Doha, Qatar. October 23, 2023.
17. Quantum Computing Applications for Healthcare and Life Science. Workshop on Quantum Applications - Software Stack for Scalable Heterogeneous NISQ Cluster (SSSHNC). Cleveland, USA. November 6, 2023.
18. Leveraging Quantum Computing for Protein Structure Prediction. IBM Quantum Partners Forum. Milan, Italy. May 15, 2024.

Other Presentations

International

1. Galaxy: A platform for interactive large-scale genome analysis. Evolution Meeting. Stony Brook, NY. June 2006. (platform presentation)
2. Genome analysis in 15 minutes: clicking recipes for making sense of complex data with Galaxy. Genome Informatics Meeting. Hinxton, UK. September 2006. (platform presentation)
3. Making the analysis of multiple-species whole-genome alignments accessible to everyone. The Society for Molecular Biology and Evolution Annual Meeting. Halifax, Nova Scotia, Canada. June 2007. (platform presentation)
4. Animal Genomics with Galaxy: Analyze, Publish, and Visualize. Plant and Animal Genome Conference (PAG XVIII). San Diego, CA. January 2010. (platform presentation)
5. Galaxy: Analyze, Visualize, Communicate. Plant and Animal Genome Conference (PAG XIX). San Diego, CA. January 2011. (platform presentation)
6. New Data Sources, New Tools. The 2011 Galaxy Community Conference. Lunteren, Netherlands. May 2011. (platform presentation)
7. NGS Best Practices through Galaxy: Cloud-based variant discovery with visual analytics. Intelligent Systems for Molecular Biology (ISMB). Vienna, Austria. July 2011. (platform presentation)
8. Best practices for variant discovery through Galaxy: Oscillating evolution within dual coding genes. Genome Informatics 2011, Cold Spring Harbor, New York, United States. November 2011. (platform presentation)
9. Staying on top of evolving best practices with Galaxy. Biology of Genomes, Cold Spring Harbor, New York, United States. May 2012. (poster presentation)
10. Integrating Tools & Data Sources. The 2012 Galaxy Community Conference. Chicago, USA. July 2012. (platform presentation)
11. Galaxy Toolshed for seamless software integration. Genome Informatics, Hinxton, UK. September 2012. (poster presentation)
12. Managing Galaxy's Built-in Data. The 2013 Galaxy Community Conference. Oslo, Norway. July 2013. (platform presentation)
13. Wrangling Galaxy's Reference Data. Genome Informatics, Cold Spring Harbor, New York, United States. November 2013. (poster presentation)

14. More Options, Less Time: Streamlining Access to Reference Datasets. Galaxy Community Conference. Baltimore, MD. July 2014. (platform presentation)
15. Enhancing the Galaxy Experience through Community Involvement. Bioinformatics Open Source Conference (BOSC). Boston, MA. July 2014. (platform presentation)
16. From the Ground to the Cloud in 25 minutes: Building a Customized Galaxy Analysis Server Using Only a Web Browser. Intelligent Systems for Molecular Biology (ISMB). Boston, MA. July 2014. (platform presentation)
17. Less Click, More Quick: Unattended Installation of Galaxy's Built-in Reference Data. 2015 Galaxy Community Conference. Norwich, UK. July 2015. (platform presentation)
18. Demystifying the Interoperability of Disparate Genomic Resources. Bioinformatics Open Source Conference (BOSC). Dublin, IE. July 2015. (platform presentation)
19. From the Ground to the Cloud in Just Minutes: Building a Customized Galaxy Analysis Server Using Only a Web Browser. Genome Informatics, Cold Spring Harbor, New York, United States. November 2015. (poster presentation)
20. Galaxy for the Analysis of Very Large Biomedical Data. Festival of Genomics. San Mateo, CA. November 2015. (platform presentation)
21. Sample Size Does Matter: Scaling up Analysis in Galaxy with Metagenomics. Galaxy Community Conference. Bloomington, IN. June 2016. (platform presentation)
22. Scaling up Analysis in Galaxy with Metagenomics. Galaxy Community Conference. Bioinformatics Open Source Conference (BOSC). Orlando, FL. July 2016 (platform presentation)
23. Automated Generation of Complex Toolkits for Galaxy. Galaxy Community Conference. Montpellier, France. June 2017. (platform presentation)
24. GCC 2017 Community Update. Galaxy Community Conference. Montpellier, France. June 2017. (platform presentation)
25. Choosing the best of all worlds — à la carte access to extant and emergent best-practice metagenomic pipelines. Genome Informatics, Cold Spring Harbor, New York, United States. November 2017. (poster presentation)
26. Scaling scalable infrastructure to analysis tools. Biological Data Science, Cold Spring Harbor, New York, United States. November 2018. (poster presentation)
27. RealTimeTools: Integrating, Customizing, and Accessing UI-based Tools in Galaxy. Galaxy Community Conference. Freiburg, Germany. July 2019. (platform presentation)
28. An integrative environment for microbiome analysis and visualization. Microbiomes, Cold Spring Harbor, New York, United States. July 2019. (poster presentation)
29. Workshop on Analyzing Large / Complex Experimental Designs with Galaxy. Galaxy Community Conference. Freiburg, Germany. 07/2019.
30. You wrote it, now get it used: Publishing your software with Galaxy and Bioconda. ACM Conference on Bioinformatics, Computational Biology, and Health Informatics (ACM BCB). Niagara Falls, NY. 09/2019. (platform presentation)
31. Galaxy as an integrative environment for microbiome analysis and visualization. American Society of Human Genetics (ASHG) Annual Meeting. Houston, TX. 10/2019. (poster presentation)
32. Interactive Workshop: Reproducible and Transparent Genomic Analysis with Galaxy. American Society of Human Genetics (ASHG) Annual Meeting. Houston, TX. 10/2019. (workshop presentation)
33. An integrative environment for microbiome analysis and visualization. Genome Informatics, Cold Spring Harbor, New York, United States. November 2019. (poster presentation)
34. Roundtable on Quantum Computing in Biomedical Science. Cleveland Clinic and the National Science Foundation workshop on artificial intelligence and biomedical research. June 14, 2021.
35. Interview at IBM Quantum Open House. April 12, 2022.
36. Panelist on Workforce Development at IBM Europe Quantum Partner Forum. Zurich, Switzerland. May 12, 2022.
37. Investigating the known and unknown microbial composition of metagenomic samples made easy with MetaSBT in Galaxy. Galaxy Community Conference 2023. Brisbane, Australia. July 10, 2023. (platform presentation)
38. Let the analysis be GiN: Galaxy in Notebooks: An update. Galaxy Community Conference 2023. Brisbane, Australia. July 12, 2023. (platform presentation)
39. Quantum computing comes to the Galaxy. Genome Informatics. December 7, 2023. (poster presentation).
40. Empowering computational chemistry outreach and training with the Galaxy Project. American Chemical Society Spring Meeting. March 20, 2024. (platform presentation)

Regional

1. A database of mouse Ka/Ks sliding window profiles. Northeast Ecology and Evolution Conference. University Park, PA. March 2005. (platform presentation)
2. Collaborative Computational Biology. Cleveland Clinic - IBM Research Discovery Accelerator Workshop on Drug Discovery: Small Molecules and Biologics. Cleveland, OH. January 25, 2023

Trainees / Mentees

1. Vijay Krishna Nagampalli, PhD. Postdoctoral Fellow. 03/2018-2021.
2. Jayadev Joshi, PhD. Postdoctoral Fellow. 09/2018-08/2023. Research Associate. 09/2023-present.
3. Christopher Lowe. Undergraduate 2-month internship. 10-12/2018.
4. Bryan Raubenholt, PhD. Postdoctoral Fellow. 09/2021-present.
5. Fabio Cumbo, PhD. Postdoctoral Fellow. 03/2022-present.
6. Mussa Wisoba. Inaugural CCF Quantum High School Internship. Summer 2022.
7. Luis Gutierrez. CCF DA Quantum High School Internship. Summer 2023
8. Maeve Gaffney. CCF DA Quantum Undergraduate Internship. Summer 2024.
9. Alexander Holden. CCF DA Quantum Undergraduate Internship. Summer 2024. Undergraduate student researcher August 2024-present.
10. Hannah Linn. Visiting PhD student from Chalmers University of Technology, Applied Quantum Physics Laboratory. August 2024-November 2024.

Teaching Activities

1. Elementary Microbiology Laboratory (MICRB 107). Pennsylvania State University, Undergraduate students. 4-hours per week per semester, for two semesters. 2005.
2. Bioinformatics: Analyzing High Throughput Sequencing Data (BMMB 597D). 1-hour lecture. 11/2013.
3. Bioinformatics I (BMMB 598A). 1-hour lecture 11/2014, 11/2015
4. Foundations of data-driven life sciences (IBIOS/BMB554). 1-hour lecture. 11/2016
5. Mammalian Genetics, Genomics, and Bioinformatics (MMED 414). Sequence analysis and 2-hour lecture, with hands-on activities. Graduate students. 04/2018
6. 5-day, 2-hours per day workshop on bioinformatics and data analysis. Mammalian Genetics, Genomics, and Bioinformatics (MMED 414). 05/2018, 05/2019, 05/2020
7. Training on Clinical Cancer Genomics Fellowship Bioinformatics rotation for Takae Brewer, MD. 2-weeks. 03/2019
8. GTN Smörgåsbord: A Global Galaxy Course. Instructor. 02/15-02/19 2021.
9. GTN Smörgåsbord 2022: Tapas Edition. Instructor. 03/14-03/18 2022.
10. Lecture on bioinformatics and data analysis. Mammalian Genetics, Genomics, and Bioinformatics (MMED 414). 05/2022.
11. Guest Lecturer for Cleveland State CIS 492/593 Special Topic in Quantum Computing. October 21, 2022.
12. CCLCM Bioinformatics Thread Leader. 2022-present.

FUNDING / RESEARCH / GRANT SUPPORT

Current

Funding Agency and ID #	Wellcome Trust, 313498/Z/24/Z
Title of Grant	Automated Generation of Galaxy tools
PI	Blankenberg
Role	PI
Percent Effort	10%
Percent Salary Support	10%
Annual Direct Costs	\$173,913 Annual Direct Costs to Blankenberg Lab
Award time frame (funding period)	10/01/2024-09/31/2026
Funding Agency and ID #	NIH/NCI, U24 CA284167
Title of Grant	Developing a Cancer Galaxy Computational Workbench to Meet Emerging Cancer Data Analysis Needs
PI	Goecks

Role	Co-I
Percent Effort	5%
Percent Salary Support	5%
Annual Direct Costs	\$30,000 Annual Direct Costs to Blankenberg Lab
Award time frame (funding period)	06/01/2024-05/31/2029
Funding Agency and ID #	NIH/NHGRI, U24 HG006620
Title of Grant	Democratization of Data Analysis in Life Sciences Through Galaxy
PI	Nekrutenko; Schatz
Role	Co-I
Percent Effort	5%
Percent Salary Support	5%
Annual Direct Costs	\$65,000 Annual Direct Costs to Blankenberg Lab
Award time frame (funding period)	02/01/2021-01/30/2026

Pending

Funding Agency and ID #	National Science Foundation
Title of Grant	Collaborative Research: Framework: Integrating Avogadro and Galaxy: Enabling Complex Supramolecular Dynamics Simulations
Role	Co-PI with Hutchison
Percent Effort	9%
Percent Salary Support	9%
Annual Direct Costs	\$178,850 Annual Direct Costs to Blankenberg Lab
Award time frame (funding period)	07/01/2026-06/30/2029 - Pending

Funding Agency and ID #	NIH/NIGMS
Title of Grant	Transforming Biomedical Analysis via Machine Learning, FAIR Analytics, and Quantum Methods
Role	PI
Percent Effort	36%
Percent Salary Support	36%
Annual Direct Costs	\$275,000 Annual Direct Costs to Blankenberg Lab
Award time frame (funding period)	04/01/2026-03/31/2031 - Pending

Funding Agency and ID #	NIH/NIGMS
Title of Grant	Accelerating Biomedical Discovery through Machine Learning, FAIR Data Science, and Quantum Computing
Role	PI
Percent Effort	36%
Percent Salary Support	36%
Annual Direct Costs	\$275,000 Annual Direct Costs to Blankenberg Lab
Award time frame (funding period)	09/01/2026-08/31/2031 - Pending

Funding Agency and ID #	NIH/NHGRI, U24 HG006620
Title of Grant	Democratization of Data Analysis in Life Sciences Through Galaxy
PI	Nekrutenko; Schatz; Blankenberg; Goecks
Role	Co-PI
Percent Effort	5%
Percent Salary Support	5%
Annual Direct Costs	\$119,876 Annual Direct Costs to Blankenberg Lab
Award time frame (funding period)	07/01/2026-06/30/2031 - Pending

Past

Funding Agency and ID #	DOD/DARPA, HR001121S0026
Title of Grant	QPerf: A benchmarking framework for quantum computers
PI	Shehab
Role	Site PI
Percent Effort	10%
Percent Salary Support	10%
Annual Direct Costs	\$301,644 Average Annual Direct Costs to CCF
Award time frame (funding period)	5/18/2022 – 8/31/2023
Funding Agency and ID #	NIH/NCI, U24 CA231877
Title of Grant	A Federated Galaxy for user-friendly large-scale cancer genomics research
PI	Goecks; Schatz
Role	Co-I
Percent Effort	5%
Percent Salary Support	5%
Annual Direct Costs	\$41,500/yr Annual Direct Costs to Blankenberg Lab
Award time frame (funding period)	09/16/2018-08/31/2023
Funding Agency and ID #	Ohio Aerospace Institute / US Army Research Office
Title of Grant	Assessment of Quantum Computing Impacts on Air Force Digital Transformation
PI	Erdemir
Role	Co-I
Percent Effort	25%
Percent Salary Support	25%
Annual Direct Costs	\$299,960 total costs
Award time frame (funding period)	10/2023-07/2024
Funding Agency and ID #	Wellcome Leap
Title of Grant	Protein Conformation Prediction with Quantum Computing
Role	PI
Percent Effort	20%
Percent Salary Support	20%
Annual Direct Costs	\$1,034,971 Annual Direct Costs
Award time frame (funding period)	09/01/2023-09/30/2024
Funding Agency and ID #	Wellcome Leap
Title of Grant	Quantum Computing for Photon-Drug Interactions in Cancer Prevention and Cancer Treatment
PI	Maniscalco & Krishna
Role	Co-I
Percent Effort	10%
Percent Salary Support	10%
Annual Direct Costs	\$311,229 Average Annual Direct Costs to CCF
Award time frame (funding period)	09/01/2023-02/28/2026
Funding Agency and ID #	NIH/NCI, U24 CA248457
Title of Grant	GenePattern and GenePattern Notebook: Integrative ‘Omic Analysis for Cancer Research
PI	Mesirov; Tamayo
Role	Co-I

Percent Effort	5%
Percent Salary Support	5%
Annual Direct Costs	\$39,588/yr (1, 2), \$20,348/yr (3,4,5) Annual Direct Costs to Blankenberg Lab
Award time frame (funding period)	06/01/2020-05/31/2025

BIBLIOGRAPHY

Underlined when senior/corresponding author
*Joint first or senior authorship denoted with **
Trainees denoted with #

Peer-reviewed articles of original research

1. Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P, Zhang Y, Blankenberg D, Albert I, Taylor J, Miller W, Kent WJ, Nekrutenko A. Galaxy: a platform for interactive large-scale genome analysis. *Genome Res.* 2005 Oct;15(10):1451-5. doi: 10.1101/gr.4086505. Epub 2005 Sep 16. PubMed PMID: 16169926; PubMed Central PMCID: PMC1240089.
2. Blankenberg D, Taylor J, Schenck I, He J, Zhang Y, Ghent M, Veeraraghavan N, Albert I, Miller W, Makova KD, Hardison RC, Nekrutenko A. A framework for collaborative analysis of ENCODE data: making large-scale analyses biologist-friendly. *Genome Res.* 2007 Jun;17(6):960-4. doi: 10.1101/gr.5578007. PubMed PMID: 17568012; PubMed Central PMCID: PMC1891355.
3. Sangar V, Blankenberg DJ, Altman N, Lesk AM. Quantitative sequence-function relationships in proteins based on gene ontology. *BMC Bioinformatics.* 2007 Aug 8;8:294. doi: 10.1186/1471-2105-8-294. PubMed PMID: 17686158; PubMed Central PMCID: PMC1976327.
4. Miller W, Rosenbloom K, Hardison RC, Hou M, Taylor J, Raney B, Burhans R, King DC, Baertsch R, Blankenberg D, Kosakovsky Pond SL, Nekrutenko A, Giardine B, Harris RS, Tyekucheva S, Diekhans M, Pringle TH, Murphy WJ, Lesk A, Weinstock GM, Lindblad-Toh K, Gibbs RA, Lander ES, Siepel A, Haussler D, Kent WJ. 28-way vertebrate alignment and conservation track in the UCSC Genome Browser. *Genome Res.* 2007 Dec;17(12):1797-808. doi: 10.1101/gr.6761107. Epub 2007 Nov 5. PubMed PMID: 17984227; PubMed Central PMCID: PMC2099589.
5. Blankenberg D, Gordon A, Von Kuster G, Coraor N, Taylor J, Nekrutenko A. Manipulation of FASTQ data with Galaxy. *Bioinformatics.* 2010 Jul 15;26(14):1783-5. doi: 10.1093/bioinformatics/btq281. Epub 2010 Jun 18. PubMed PMID: 20562416; PubMed Central PMCID: PMC2894519.
6. Blankenberg D, Coraor N, Von Kuster G, Taylor J, Nekrutenko A. Integrating diverse databases into a unified analysis framework: a Galaxy approach. *Database (Oxford).* 2011;2011:bar011. doi: 10.1093/database/bar011. Print 2011. PubMed PMID: 21531983; PubMed Central PMCID: PMC3092608.
7. Blankenberg D, Taylor J, Nekrutenko A. Making whole genome multiple alignments usable for biologists. *Bioinformatics.* 2011 Sep 1;27(17):2426-8. doi: 10.1093/bioinformatics/btr398. Epub 2011 Jul 19. PubMed PMID: 21775304; PubMed Central PMCID: PMC3157923.
8. Stamatoyanopoulos JA, Snyder M, Hardison R, Ren B, Gingeras T, Gilbert DM, Groudine M, Bender M, Kaul R, Canfield T, Giste E, Johnson A, Zhang M, Balasundaram G, Byron R, Roach V, Sabo PJ, Sandstrom R, Stehling AS, Thurman RE, Weissman SM, Cayting P, Hariharan M, Lian J, Cheng Y, Landt SG, Ma Z, Wold BJ, Dekker J, Crawford GE, Keller CA, Wu W, Morrissey C, Kumar SA, Mishra T, Jain D, Byrka-Bishop M, Blankenberg D, Lajoie BR, Jain G, Sanyal A, Chen KB, Denas O, Taylor J, Blobel GA, Weiss MJ, Pimkin M, Deng W, Marinov GK, Williams BA, Fisher-Aylor KI, Desalvo G, Kiralusha A, Trout D, Amrhein H, Mortazavi A, Edsall L, McCleary D, Kuan S, Shen Y, Yue F, Ye Z, Davis CA, Zaleski C, Jha S, Xue C, Dobin A, Lin W, Fastuca M, Wang H, Guigo R, Djebali S, Lagarde J, Ryba T, Sasaki T, Malladi VS, Cline MS, Kirkup VM, Learned K, Rosenbloom KR, Kent WJ, Feingold EA, Good PJ, Pazin M, Lowdon RF, Adams LB. An encyclopedia of mouse DNA elements (Mouse ENCODE). *Genome Biol.* 2012 Aug 13;13(8):418. doi: 10.1186/gb-2012-13-8-418. PubMed PMID: 22889292; PubMed Central PMCID: PMC3491367.
9. Minevich G, Park DS, Blankenberg D, Poole RJ, Hobert O. CloudMap: a cloud-based pipeline for analysis of mutant genome sequences. *Genetics.* 2012 Dec;192(4):1249-69. doi: 10.1534/genetics.112.144204. Epub 2012 Oct 10. PubMed PMID: 23051646; PubMed Central PMCID: PMC3512137.
10. Blankenberg D, Von Kuster G, Bouvier E, Baker D, Afgan E, Stoler N, Taylor J, Nekrutenko A. Dissemination of scientific software with Galaxy ToolShed. *Genome Biol.* 2014 Feb 20;15(2):403. doi: 10.1186/gb4161. PubMed PMID: 25001293; PubMed Central PMCID: PMC4038738.

11. Dickins B, Rebolledo-Jaramillo B, Su MS, Paul IM, Blankenberg D, Stoler N, Makova KD, Nekrutenko A. Controlling for contamination in re-sequencing studies with a reproducible web-based phylogenetic approach. *Biotechniques*. 2014;56(3):134-141. doi: 10.2144/000114146. eCollection 2014. PubMed PMID: 24641477; PubMed Central PMCID: PMC4377138.
12. Blankenberg D, Johnson JE, Taylor J, Nekrutenko A. Wrangling Galaxy's reference data. *Bioinformatics*. 2014 Jul 1;30(13):1917-9. doi: 10.1093/bioinformatics/btu119. Epub 2014 Feb 28. PubMed PMID: 24585771; PubMed Central PMCID: PMC4071198.
13. Rebolledo-Jaramillo B, Su MS, Stoler N, McElhoe JA, Dickins B, Blankenberg D, Korneliusen TS, Chiaromonte F, Nielsen R, Holland MM, Paul IM, Nekrutenko A, Makova KD. Maternal age effect and severe germ-line bottleneck in the inheritance of human mitochondrial DNA. *Proc Natl Acad Sci U S A*. 2014 Oct 28;111(43):15474-9. doi: 10.1073/pnas.1409328111. Epub 2014 Oct 13. PubMed PMID: 25313049; PubMed Central PMCID: PMC4217420.
14. Qu K, Garamszegi S, Wu F, Thorvaldsdottir H, Liefeld T, Ocana M, Borges-Rivera D, Pochet N, Robinson JT, Demchak B, Hull T, Ben-Artzi G, Blankenberg D, Barber GP, Lee BT, Kuhn RM, Nekrutenko A, Segal E, Ideker T, Reich M, Regev A, Chang HY, Mesirov JP. Integrative genomic analysis by interoperation of bioinformatics tools in GenomeSpace. *Nat Methods*. 2016 Mar;13(3):245-247. doi: 10.1038/nmeth.3732. Epub 2016 Jan 18. PubMed PMID: 26780094; PubMed Central PMCID: PMC4767623.
15. Afgan E, Baker D, van den Beek M, Blankenberg D, Bouvier D, Čech M, Chilton J, Clements D, Coraor N, Eberhard C, Grüning B, Guerler A, Hillman-Jackson J, Von Kuster G, Rasche E, Soranzo N, Turaga N, Taylor J, Nekrutenko A, Goecks J. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res*. 2016 Jul 8;44(W1):W3-W10. doi: 10.1093/nar/gkw343. Epub 2016 May 2. PubMed PMID: 27137889; PubMed Central PMCID: PMC4987906.
16. Grüning B, Dale R, Sjödin A, Chapman BA, Rowe J, Tomkins-Tinch CH, Valieris R, Köster J, Bioconda Team. Bioconda: sustainable and comprehensive software distribution for the life sciences. *Nat Methods*. 2018 Jul;15(7):475-476. doi: 10.1038/s41592-018-0046-7. PMID: 29967506.
17. Grüning B, Sallou O, Moreno P, da Veiga Leprevost F, Ménager H, Søndergaard D, Röst H, Sachsenberg T, O'Connor B, Madeira F, Dominguez Del Angel V, Crusoe MR, Varma S, Blankenberg D, Jimenez RC, Perez-Riverol Y. Recommendations for the packaging and containerizing of bioinformatics software. *F1000Res*. 2018;7. doi: 10.12688/f1000research.15140.2. eCollection 2018. PubMed PMID: 31543945; PubMed Central PMCID: PMC6738188.
18. Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Cech M, Chilton J, Clements D, Coraor N, Grüning BA, Guerler A, Hillman-Jackson J, Hiltemann S, Jalili V, Rasche H, Soranzo N, Goecks J, Taylor J, Nekrutenko A, Blankenberg D. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res*. 2018 Jul 2;46(W1):W537-W544. doi: 10.1093/nar/gky379. PubMed PMID: 29790989; PubMed Central PMCID: PMC6030816.
19. Craig SJC, Blankenberg D, Parodi ACL, Paul IM, Birch LL, Savage JS, Marini ME, Stokes JL, Nekrutenko A, Reimherr M, Chiaromonte F, Makova KD. Child Weight Gain Trajectories Linked To Oral Microbiota Composition. *Sci Rep*. 2018 Sep 19;8(1):14030. doi: 10.1038/s41598-018-31866-9. PubMed PMID: 30232389; PubMed Central PMCID: PMC6145887.
20. Parthibane V, Acharya D, Srideshikan SM, Lin J, Myerscough DG, Abimannan T, Vijaykrishna N[#], Blankenberg D, Bondada L, Klarmann KD, Fox SD, Andresson T, Tassarollo L, Acharya U, Keller JR, Acharya JK. Sptlc1 is essential for myeloid differentiation and hematopoietic homeostasis. *Blood Adv*. 2019 Nov 26;3(22):3635-3649. doi: 10.1182/bloodadvances.2019000729. PubMed PMID: 31751474; PubMed Central PMCID: PMC6880889.
21. Blankenberg D, Coraor N, Chilton J. Galaxy External Display Applications: Closing a dataflow interoperability loop. *Nat Methods*. 2020 Feb;17(2):123-124. doi: 10.1038/s41592-019-0727-x. PubMed PMID: 31959996.
22. Chapman LM, Spies N, Pai P, Lim CS, Carroll A, Narzisi G, Watson CM, Proukakis C, Clarke WE, Nariai N, Dawson E, Jones G, Blankenberg D, Brueffer C, Xiao C, Kolara SRR, Alexander N, Wolujewicz P, Ahmed AE, Smith G, Shehreen S, Wenger AM, Salit M, Zook JM. A crowdsourced set of curated structural variants for the human genome. *PLoS Comput Biol*. 2020 Jun 19;16(6):e1007933. doi:10.1371/journal.pcbi.1007933. PubMed PMID: 32559231; PubMed Central PMCID: PMC7329145.
23. Jalili V, Afgan E, Gu Q, Clements D, Blankenberg D, Goecks J, Taylor J, Nekrutenko A. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2020 update. *Nucleic Acids Res*. 2020 Jun 25;gkaa434. doi:10.1093/nar/gkaa434. PMID: 32479607.
24. Baker D, van den Beek M, Blankenberg D, Bouvier D, Chilton J, Coraor N, Coppens F, Eguinoa I, Gladman S, Grüning B, Keener N, Larivière D, Lonie A, Pond S, Maier W, Nekrutenko A, Taylor J, Weaver S. No more

- business as usual: Agile and effective responses to emerging pathogen threats require open data and open analytics. *PLoS Pathog.* 2020 Aug 13;16(8): e1008643. doi: 10.1371/journal.ppat.1008643.
25. Tekman M, Batut B, Ostrovsky A, Antoniewski C, Clements D, Ramirez F, Etherington GJ, Hotz HR, Scholtalbers J, Manning JR, Bellenger L, Doyle MA, Heydariyan M, Huang N, Soranzo N, Moreno P, Mautner S, Papatheodorou I, Nekrutenko A, Taylor J, Blankenberg D, Backofen R, Grüning B. A single-cell RNA-sequencing training and analysis suite using the Galaxy framework. *Gigascience.* 2020 Oct 20;9(10):giaa102. doi: 10.1093/gigascience/giaa102. PMID: 33079170; PMCID: PMC7574357.
 26. Eren AM, Kiefl E, Shaiber A, Veseli I, Miller S, Schechter M, Fink I, Pan J, Yousef M, Fogarty E, Trigodet F, Watson A, Esen O, Moore R, Clayssen Q, Lee M, Kivenson V, Graham E, Merrill B, Karkman A, Blankenberg D, Eppley J, Sjödin A, Scott J, Vázquez-Campos X, McKay L, McDaniel E, Stevens S, Anderson R, Fuessel J, Fernandez-Guerra A, Maignien L, Delmont T, Willis A. Community-led, integrated, reproducible multi-omics with anvio. *Nat Microbiol.* 6, 3–6 (2021). doi: 10.1038/s41564-020-00834-3. PMID: 33349678.
 27. Gramm M, Pérez-Palma E, Schumacher-Bass S, Dalton J, Leu C, Blankenberg D*, Lal D*. SimText: A text mining framework for interactive analysis and visualization of similarities among biomedical entities. *Bioinformatics.* 2021 May 25:btab365. doi: 10.1093/bioinformatics/btab365. PMID: 34037702
 28. Dubey R, Kumar Patra A, Joshi J[#], Blankenberg D, Sekhara Rao Kolluru S, Madhu B, Raval S. Evaluation of low-cost particulate matter sensors OPC N2 and PM Nova for aerosol monitoring. *Atmospheric Pollution Research.* Volume 13, Issue 3, 2022, 101335, ISSN 1309-1042. doi: /doi.org/10.1016/j.apr.2022.101335.
 29. Sharma V, Joshi J[#], Yeh I-J, Doughman Y, Blankenberg D, Wald D, Montano M. Re-expression of ERalpha and AR in Receptor Negative Endocrine Cancers via GSK3 inhibition. *Front Oncol.* 2022 Mar 24;12:824594. doi: 10.3389/fonc.2022.824594. PMID: 35402240; PMCID: PMC8988137.
 30. Dubey R, Patra AK, Joshi J[#], Blankenberg D, Allaudeen N. Evaluation of vertical and horizontal distribution of particulate matter near an urban roadway using an unmanned aerial vehicle. *Science of The Total Environment.* 25 August 2022. Volume 836, 2022, 155600, ISSN 0048-9697. doi: 10.1016/j.scitotenv.2022.155600.
 31. Galaxy Community (including Raubenolt B[#], Joshi J[#], Vijaykrishna N[#], Blankenberg D*). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. *Nucleic Acids Res.* 2022 Apr 21: gkac247. doi: 10.1093/nar/gkac247. PMID: 35446428.
 32. Vijaykrishna N[#], Joshi J[#], Coraor N, Hillman-Jackson J, Bouvier D, van den Beek M, Eguinoa I, Coppens F, Davis J, Stolarczyk M, C. Sheffield N, Gladman S, Cuccuru G, Grüning B, Soranzo N, Rasche H, Langhorst BW, Bernt M, Fornika D, Anderson de Lima Morais D, Barrette M, van Heusden P, Petrillo M, Puertas-Gallardo A, Patak A, Hotz H, Blankenberg D. Expanding the Galaxy's reference data. *Bioinform Adv.* 2022 Apr 29;2(1):vbac030. doi: 10.1093/bioadv/vbac030. PMID: 35669346; PMCID: PMC9155181.
 33. Joshi J[#], Blankenberg D. PDAUG - a Galaxy based toolset for peptide library analysis, visualization, and machine learning modeling. *BMC Bioinformatics.* 2022 May 28;23(1):197. doi: 10.1186/s12859-022-04727-6. PMID: 35643441; PMCID: PMC9148462.
 34. Oh S, Geistlinger L, Ramos M, Blankenberg D, van den Beek, M, Taroni J, Carey V, Greene C, Waldron L, Davis S. GenomicSuperSignature: interpretation of RNA-seq experiments through robust, efficient comparison to public databases. *Nat Commun.* 2022 Jun 27;13(1):3695. doi: 10.1038/s41467-022-31411-3. PMID: 35760813.
 35. Kunduri G, Le SH, Baena V, Vijaykrishna N[#], Harned A, Nagashima K, Blankenberg D, Yoshihiro I, Narayan K, Bamba T, Acharya U, Acharya JK. Delivery of ceramide phosphoethanolamine lipids to the cleavage furrow through the endocytic pathway is essential for male meiotic cytokinesis. *PLoS Biol.* 2022 Sep 28;20(9):e3001599. doi: 10.1371/journal.pbio.3001599. PMID: 36170207; PMCID: PMC9550178.
 36. Vasu K, Khan D, Ramachandiran I, Blankenberg D*, Fox PL*. Analysis of nested alternate open reading frames and their encoded proteins. *NAR Genom Bioinform.* 2022 Oct 19;4(4):lqac076. doi: 10.1093/nargab/lqac076. PMID: 36267124; PMCID: PMC9580016.
 37. Apollonio N, Blankenberg D, Cumbo F[#], Franciosa PG, Santoni D. Evaluating homophily in networks via HONTO (HOMophily Network TOol): a case study of chromosomal interactions in human PPI networks. *Bioinformatics.* 2023 Jan 1;39(1):btac763. doi: 10.1093/bioinformatics/btac763. PMID: 36440918; PMCID: PMC9805585.
 38. Hiltmann S, Rasche H, Gladman S, Hotz HR, Larivière D, Blankenberg D, Jagtap PD, Wollmann T, Bretaudeau A, Goué N, Griffin TJ, Royaux C, Le Bras Y, Mehta S, Syme A, Coppens F, Driesbeke B, Soranzo N, Bacon W, Psomopoulos F, Gallardo-Alba C, Davis J, Föll MC, Fahrner M, Doyle MA, Serrano-Solano B, Fouilloux AC, van Heusden P, Maier W, Clements D, Heyl F; Galaxy Training Network; Grüning B, Batut B. Galaxy Training: A powerful framework for teaching! *PLoS Comput Biol.* 2023 Jan 9;19(1):e1010752. doi: 10.1371/journal.pcbi.1010752. PMID: 36622853; PMCID: PMC9829167.

39. Macnee M, Pérez-Palma E, Brünger T, Klöckner C, Platzer K, Stefanski A, Montanucci L, Bayat A, Radtke M, Collins RL, Talkowski M, Blankenberg D, Møller RS, Lemke JR, Nothnagel M, May P, Lal D. CNV-ClinViewer: enhancing the clinical interpretation of large copy-number variants online. *Bioinformatics*. 2023 May 4;39(5):btad290. doi: 10.1093/bioinformatics/btad290. PMID: 37104749; PMCID: PMC10174702.
40. Virmani R, Pradhan P, Joshi J[#], Wang AL, Joshi HC, Sajid A, Singh A, Sharma V, Kundu B, Blankenberg D, Molle V, Singh Y, Arora G. Phosphorylation-mediated regulation of the Bacillus anthracis phosphoglycerate mutase by the Ser/Thr protein kinase PrkC. *Biochem Biophys Res Commun*. 2023 Apr 25;665:88-97. doi: 10.1016/j.bbrc.2023.04.039. Epub ahead of print. PMID: 37149987.
41. Sharma D, Singh M, Joshi J[#], Garg M, Chaudhary V, Blankenberg D, Chandna S, Kumar V, Rani R. Design and synthesis of thiazole scaffold-based small molecules as anticancer agents targeting human lactate dehydrogenase A enzyme. *ACS Omega*. 2023 8 (20), 17552-17562 doi: 10.1021/acsomega.2c07569
42. Kaur H, Minchella P, Alvarez-Carbonell D, Purandare N, Nagampalli VK[#], Blankenberg D, Hulgán T, Gerschenson M, Karn J, Aras S, Kallianpur AR. Contemporary Antiretroviral Therapy Dysregulates Iron Transport and Augments Mitochondrial Dysfunction in HIV-Infected Human Microglia and Neural-Lineage Cells. *Int J Mol Sci*. 2023 Jul 31;24(15):12242. doi: 10.3390/ijms241512242. PMID: 37569616; PMCID: PMC10419149.
43. Abimannan T, Parthibane V, Le SH, Vijaykrishna N[#], Fox SD, Karim B, Kunduri G, Blankenberg D, Andresson T, Bamba T, Acharya U, Acharya JK. Sphingolipid biosynthesis is essential for metabolic rewiring during TH17 cell differentiation. *Sci Adv*. 2024 Apr 26;10(17):eadk1045. doi: 10.1126/sciadv.adk1045. Epub 2024 Apr 24. PMID: 38657065; PMCID: PMC11042737.
44. Galaxy Community (including Raubenolt B[#], Joshi J[#], Cumbo F[#], Blankenberg D*). The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update. *Nucleic Acids Res*. 2024 May 20:gkae410. doi: 10.1093/nar/gkae410. Epub ahead of print. PMID: 38769056.
45. Khan D, Ramachandiran I, Vasu K, China A, Khan K, Cumbo F[#], Halawani D, Terenzi F, Zin I, Long B, Costain G, Blaser S, Carnevale A, Gogonea V, Dutta R, Blankenberg D, Yoon G, Fox PL. Homozygous EPRS1 missense variant causing hypomyelinating leukodystrophy-15 alters variant-distal mRNA m6A site accessibility. *Nat Commun*. 2024 May 20;15(1):4284. doi: 10.1038/s41467-024-48549-x. PMID: 38769304; PMCID: PMC11106242.
46. Raubenolt B[#], Blankenberg D. Generalized open-source workflows for atomistic molecular dynamics simulations of viral helicases. *Gigascience*. 2024 Jan 2;13. doi: 10.1093/gigascience/giae026. PubMed PMID: 38869150; PubMed Central PMCID: PMC11170216.
47. Huang C, Guo Y, Zhu Z, S Mi, Blankenberg D, Liu R. Quantum Exploration-based Reinforcement Learning for Efficient Robot Path Planning in Sparse-Reward Environment. *2024 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN)*, Pasadena, CA, USA, 2024, pp. 516-521, doi: 10.1109/RO-MAN60168.2024.10731199.
48. Dubey R, Patra AK, Joshi J[#], Blankenberg D, Nazneen, Penchala A, Lee X. Vertical distribution of particulate matter near a national highway and influence of roadside tree canopy: A drone/UAV based study. *Atmospheric Environment*. doi: doi.org/10.1016/j.atmosenv.2024.120761. 2024 Nov 1.
49. Ozmen BB, Raubenolt BA[#], Blankenberg D, Schwarz GS. The next frontier in plastic surgery innovation: Quantum computing. *J Plast Reconstr Aesthet Surg*. 2025 Jan;100:282-284. doi: 10.1016/j.bjps.2024.12.001. Epub 2024 Dec 8. Review. PubMed PMID: 39675246.
50. Cumbo F[#], Truglia S, Weitschek E, Blankenberg D. Feature selection with vector-symbolic architectures: a case study on microbial profiles of shotgun metagenomic samples of colorectal cancer. *Brief Bioinform*. 2025 Mar 4;26(2):bbaf177. doi: 10.1093/bib/bbaf177. PMID: 40269516; PMCID: PMC12018301.
51. Reich M, Tabor T, Liefeld T, Joshi J[#], Kim F, Huang E, Thorvaldsdottir H, Blankenberg D, Mesirov JP. Genomics to Notebook (g2nb): Extending the Electronic Notebook to Address the Challenges of Bioinformatics Analysis. *J Bioinform Syst Biol*. 2025;8(1):1-7. doi: 10.26502/jbsb.5107096. Epub 2025 Jan 27. PMID: 40191160; PMCID: PMC11970915.
52. Joshi J[#], Bhutada S, Martin DR, Guzowski J, Blankenberg D, Apte SS. DICED (Database of Identified Cleavage Sites Endemic to Diseases States): A Searchable Web Interface for Terminomics/Degradomics. *Proteomics*. 2025 Jul;25(13):e202500007. doi: 10.1002/pmic.202500007. Epub 2025 May 12. PMID: 40351053; PMCID: PMC12246766.
53. Li X, Kulkarni VR, Nana J, Pu S, Xie N, Guan Q, Li R[#], Zhang S, Xu S, Blankenberg D, Chaudhary V. Quantum Circuit Optimization for Protein Structure Prediction. 2025. *55th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W)*, Naples, Italy, 2025, pp. 220-223, doi: 10.1109/DSN-W65791.2025.00065.

54. Raubenolt B[#], [Blankenberg D](#). GalaxyQ: A Platform for Reproducible and Provenance-Driven Hybrid Quantum-Classical Scientific Workflows. *2025 IEEE International Conference on Quantum Computing and Engineering (QCE)*. 2025 August; 2(2):264-267. doi: 10.1109/QCE65121.2025.10331.

Pending peer-reviewed articles of original research (submitted and in-review only)

1. Nasr E, Pechlivanis N, Strepis N, Amato P, Bernt M, Bhardwaj A, Blankenberg D, Brites D, Cumbo F[#], Do KT, Ferrari E, Griffin TJ, Grüning B, Hiltemann S, Hyde CJ, Jagtap P, Mehta S, Métris KL, Momin S, Nelson TM, Oba A, Pavloudi C, Péguilhan R, Price GR, Psomopoulos F, Rosic N, Schatz MC, Schiml VC, Siguret C, Soranzo N, Stubbs A, van Heusden P, Vohra M; microGalaxy Community; Zierrep P, Batut B. Microbiology Galaxy Lab: The first community-driven gateway for reproducible and FAIR analysis of microbial data. *bioRxiv [Preprint]*. 2025 Sep 16:2024.12.23.629682. doi: 10.1101/2024.12.23.629682. PMID: 39764050; PMCID: PMC11703195. (*submitted*)
2. Cumbo F[#], [Blankenberg D](#). Characterization of microbial dark matter at scale with MetaSBT and taxonomy-aware Sequence Bloom Trees. *bioRxiv [Preprint]*. 2025 Aug 30:2025.08.25.672238. doi: 10.1101/2025.08.25.672238. PMID: 40909705; PMCID: PMC12407952. (*under revision at Nature Biotech*)
3. Li R[#], Doga H^{*}, Raubenolt B[#], Mostame S, DiSanto N, Cumbo F[#], Joshi J[#], Linn H[#], Gaffney M[#], Holden A[#], Kulkarni V[#], Chaudhary V, Merz Jr KM, Saki AA, Radivoyevitch T, DiFilippo F, Qin J, Shehab O, [Blankenberg D](#). Quantum Algorithm for Protein Structure Prediction Using the Face-Centered Cubic Lattice. [preprint]. 2025 July. Available from: <https://arxiv.org/abs/2507.08955>. (*under review at npj quantum information*)
4. Cumbo F[#], Dhillon K[#], Joshi J[#], Chicco D, Aygun S, [Blankenberg D](#). A novel Vector-Symbolic Architecture for graph encoding and its application to viral pangenome-based species classification. [preprint]. 2025 September. Available from: <https://www.biorxiv.org/content/10.1101/2025.09.08.674958v1>. (*under review at BioData Mining*)
5. Cumbo F[#], Chicco D, Aygun S, [Blankenberg D](#). Designing Vector-Symbolic Architectures for Biomedical Applications: Ten Tips and Common Pitfalls. [preprint]. 2025 October. Available from: <https://www.preprints.org/manuscript/202510.0117>. (*accepted PeerJ*)
6. Cumbo F, Li R, Raubenolt B, Joshi J, Mohammad Masum A, Aygun S, Blankenberg D. Quantum Hyperdimensional Computing: a foundational paradigm for quantum neuromorphic architectures. [preprint]. 2025 November. Available from: <https://arxiv.org/abs/2511.12664>. (*under revision at npj unconventional computing*)

Other peer-reviewed publications (reviews, method papers, commentaries, editorials)

1. [Blankenberg D](#), Taylor J, Nekrutenko A. Online resources for genomic analysis using high-throughput sequencing. *Cold Spring Harb Protoc*. 2015 Feb 5;2015(4):324-35. doi: 10.1101/pdb.top083667. PubMed PMID: 25655493.
2. Nekrutenko A, Team G, Goecks J, Taylor J, Blankenberg D. Biology Needs Evolutionary Software Tools: Let's Build Them Right. *Mol Biol Evol*. 2018 Jun 1;35(6):1372-1375. doi: 10.1093/molbev/msy084. PubMed PMID: 29688462; PubMed Central PMCID: PMC5967460.
3. Grüning B, Lampa S, Vaudel M, [Blankenberg D](#). Software engineering for scientific big data analysis. *Gigascience*. 2019 May 1;8(5). doi: 10.1093/gigascience/giz054. PubMed PMID: 31121028; PubMed Central PMCID: PMC6532757.
4. Cumbo F[#], Weitschek E, [Blankenberg D](#). hdlib: A Python library for designing Vector-Symbolic Architectures. *Journal of Open Source Software*. 2023 Sept 11; 8(89), 5704, <https://doi.org/10.21105/joss.05704>.
5. Doga H, Raubenolt B[#], Cumbo F[#], Joshi J[#], DiFilippo FP, Qin J, Blankenberg D, Shehab O. A Perspective on Protein Structure Prediction Using Quantum Computers. *J Chem Theory Comput*. 2024 May 14;20(9):3359-3378. doi: 10.1021/acs.jctc.4c00067. Epub 2024 May 4. PMID: 38703105; PMCID: PMC11099973.
6. Doga H, Bose A, Sahin ME, Bettencourt-Silva J, Pham A, Kim E, Andress A, Saxena S, Parida L, Robertus JL, Kawaguchi H, Soliman R, [Blankenberg D](#). How can quantum computing be applied in clinical trial design and optimization? *Trends Pharmacol Sci*. 2024 Oct;45(10):880-891. doi: 10.1016/j.tips.2024.08.005. Epub 2024 Sep 23. PMID: 39317621.
7. Flöther FF, Blankenberg D, Demidik M, Jansen K, Krishnakumar R, Krishnakumar R, Laanait N, Parida L, Saab CY, Utro F. How quantum computing can enhance biomarker discovery. *Patterns (N Y)*. 2025 Apr 29;6(6):101236. doi: 10.1016/j.patter.2025.101236. PMID: 40575130; PMCID: PMC12191739.

Non-peer reviewed publications and submissions to pre-print servers

1. Joshi J[#], Cumbo F[#], Blankenberg D. Large-scale classification of metagenomic samples: a comparative analysis of classical machine learning techniques vs a novel brain-inspired hyperdimensional computing approach. *bioRxiv* [Preprint]. 2025 Dec 2:2025.07.06.663394. doi: 10.1101/2025.07.06.663394. PMID: 40672168; PMCID: PMC12265723.
2. Cumbo F[#], Dhillon K[#], Joshi J[#], Raubenolt B[#], Chicco D, Aygun S, Blankenberg D. Predicting the toxicity of chemical compounds via Hyperdimensional Computing. [preprint]. 2025 September. Available from: <https://www.biorxiv.org/content/10.1101/2025.09.12.675894v1>.
3. Linn H, Li R[#], Holden A[#], Saki AA, DiFilippo F, Radivoyevitch T, Blankenberg D, García-Álvarez L, Johansson G. Efficient Quantum Protein Structure Prediction with Problem-Agnostic Ansatzes. [preprint]. 2025 September. Available from: <https://arxiv.org/abs/2509.18263>.

Books, book chapters, journal volumes as author or editor/guest editor

1. Taylor J, Schenck I, Blankenberg D, Nekrutenko A. Using galaxy to perform large-scale interactive data analyses. *Curr Protoc Bioinformatics*. 2007 Sep;Chapter 10:Unit 10.5. doi: 10.1002/0471250953.bi1005s19. PubMed PMID: 18428782; PubMed Central PMCID: PMC3418382.
2. Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, Mangan M, Nekrutenko A, Taylor J. Galaxy: a web-based genome analysis tool for experimentalists. *Curr Protoc Mol Biol*. 2010 Jan;Chapter 19:Unit 19.10.1-21. doi: 10.1002/0471142727.mb1910s89. PubMed PMID: 20069535; PubMed Central PMCID: PMC4264107.
3. Hillman-Jackson J, Clements D, Blankenberg D, Taylor J, Nekrutenko A. Using Galaxy to perform large-scale interactive data analyses. *Curr Protoc Bioinformatics*. 2012 Jun;Chapter 10:Unit10.5. doi: 10.1002/0471250953.bi1005s38. PubMed PMID: 22700312; PubMed Central PMCID: PMC4282168.
4. Blankenberg D, Hillman-Jackson J. Analysis of next-generation sequencing data using Galaxy. *Methods Mol Biol*. 2014;1150:21-43. doi: 10.1007/978-1-4939-0512-6_2. PubMed PMID: 24743989.
5. Ostrovsky A, Hillman-Jackson, J, Bouvier D, Clements D, Afgan E, Blankenberg D, Schatz M, Nekrutenko A, Taylor J, Galaxy Team, Lariviere D. Using Galaxy to Perform Large-Scale Interactive Data Analyses - An Update. *Curr Protoc*. 2021 Feb;1(2):e31. doi: 10.1002/cpz1.31. PMID: 33583104.

Software or application development (include URL/website if no journal citation)

1. <https://github.com/blankenberg>
2. <https://github.com/BlankenbergLab>

Patents and invention disclosures

1. APPLICATION # 63/520,261. 08/17/2023. MUTATION INDUCED CONFORMATIONAL CHANGES IN mRNA THAT PREVENT OR ALTER m6A METHYLATION AT DISTAL SITES. Paul Fox, Debjit Khan, Daniel Blankenberg and Fabio Cumbo. *Filed*.