MICHAEL I. LOVE
September 5, 2021

PERSONAL

University of North Carolina at Chapel Hill

Department of Biostatistics
4115E McGavran-Greenberg Hall
135 Dauer Dr, Chapel Hill, NC 27599

Department of Genetics
120 Mason Farm Rd, Chapel Hill, NC 27514

`milove [at] email.unc.edu`
http://mikelove.github.io

EDUCATION

Postdoctoral Research Fellow **2013 - 2016**
Supervisor: Rafael Irizarry
Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute
Department of Biostatistics, Harvard TH Chan School of Public Health

Dr. rer. nat., Computational Biology **2013**
Magna cum laude
Freie Universität, Berlin, Germany
Max Planck Institute for Molecular Genetics, Berlin, Germany
International Max Planck Research School for Computational Biology and Scientific Computing
Advisors: Prof. Dr. Martin Vingron, Prof. Dr. Knut Reinert, Dr. Stefan Haas
Dissertation title: *Statistical Analysis of High-Throughput Sequence Count Data*

M.S., Statistics **2010**
Stanford University, Stanford, CA

B.S., Mathematics **2005**
With distinction
Stanford University, Stanford, CA

PROFESSIONAL EXPERIENCE

Assistant Professor **2016 - present**
Department of Biostatistics,
Department of Genetics,
Member of Lineberger Comprehensive Cancer Center
University of North Carolina-Chapel Hill

HONORS

- MAQC 2021 Scientist Research Award (3rd place)
- UNC Gillings Departmental Teaching Excellence and Innovation Award for 2020-21.
- UNC Center for Environmental Health and Susceptibility (CEHS) Recruitment Award for 2019.
- UNC Junior Faculty Development Award for 2017.

BIBLIOGRAPHY

Books and chapters:

1. **Michael I. Love**, "Statistical modeling of high dimensional counts." *RNA Bioinformatics*, 2nd ed., edited by Ernesto Picardi, Springer US, 2021, pp. 97–134. doi: 10.1007/978-1-0716-1307-8
2. Rafael A. Irizarry and **Michael I. Love**, *Data Analysis for the Life Sciences with R*, Chapman and Hall/CRC, 2016, 376 pages. doi: 10.1201/9781315367002

Refereed papers/articles:

Key: **Primary method development**; **Interdisciplinary work**, with role noted; " * " for work first authored by an advisee.

1. Nil Aygün, Angela L. Elwell, Dan Liang, Michael J. Lafferty, Kerry E. Cheek, Kenan P. Courtney, Jessica Mory, Ellie Hadden-Ford, Oleh Krupa, Luis de la Torre-Ubieta, Daniel H. Geschwind, **Michael I. Love**, Jason L. Stein. Brain-trait-associated variants impact cell-type-specific gene regulation during neurogenesis.
   American Journal of Human Genetics, (2021). doi: 10.1016/j.ajhg.2021.07.011
2. Gieira S. Jones, Katherine A. Hoadley, Linnea T. Olsson, Alina M. Hamilton, Arjun Bhattacharya, Erin L. Kirk, Heather J. Tipaldos, Jodie M. Fleming, **Michael I. Love**, Hazel B. Nichols, Andrew F. Olshan, Melissa A. Troester Hepatocyte growth factor pathway expression in breast cancer by race and subtype.
   Breast Cancer Research, (2021).
   Role: Provided statistical and bioinformatic consultation. doi: 10.1186/s13058-021-01460-5
3. Sean D. McCabe*, Andrew B. Nobel, **Michael I. Love**. ACTOR: a latent Dirichlet model to compare expressed isoform proportions to a reference panel.
   Biostatistics, kxab013, (2021). doi: 10.1093/biostatistics/kxab013
4. Dan Liang, Angela L. Elwell, Nil Aygün, Michael J. Lafferty, Oleh Krupa, Kerry E. Cheek, Kenan P. Courtney, Marianna Yusupova, Melanie E. Garrett, Allison Ashley-Koch, Gregory E. Crawford, **Michael I. Love**, Luis de la Torre-Ubieta, Daniel H. Geschwind, Jason L. Stein. Cell-type specific effects of genetic variation on chromatin accessibility during human neuronal differentiation.
   Nature Neuroscience, (2021).
   Role: Provided statistical analysis for caQTL, allelic imbalance, and meta-analysis. doi: 10.1038/s41593-021-00858-w

5. Anqi Zhu* [1], Nana Matoba[1], Emmaleigh Wilson, Amanda L. Tapia, Yun Li, Joseph G. Ibrahim, Jason L. Stein, **Michael I. Love**. MRLocus: identifying causal genes mediating a trait through Bayesian estimation of allelic heterogeneity.
PLOS Genetics, 17(4):e1009455, (2021). doi: 10.1371/journal.pgen.1009455

6. Arjun Bhattacharya*, Yun Li, **Michael I. Love**. Multi-omic strategies for transcriptome-wide prediction and association studies.
PLOS Genetics, 17(3):e1009398, (2021). doi: 10.1371/journal.pgen.1009398

7. Arjun Bhattacharya*, Alina M. Hamilton, Melissa A. Troester, **Michael I. Love**. DeCompress: tissue compartment deconvolution of targeted mRNA expression panels using compressed sensing.
Nucleic Acids Research, 49(8):e48, (2021). doi: 10.1093/nar/gkab031

8. Scott Van Buren*, Hirak Sarkar, Avi Srivastava, Naim U. Rashid, Rob Patro, **Michael I. Love**. Compression of quantification uncertainty for scRNA-seq counts.
Bioinformatics, btab001, (2021). doi: 10.1093/bioinformatics/btab001

9. Arjun Bhattacharya* [1], Alina M. Hamilton[1], Helena Furberg, Eugene Pietzak, Mark P. Purdue, Melissa A. Troester, Katherine A. Hoadley[†], **Michael I. Love**[†]. An approach for normalization and quality control for NanoString RNA expression data.
Briefings in Bioinformatics, 22(3):bbaa163, (2021). doi: 10.1093/bib/bbaa163

10. Hyo Young Choi, Heejoon Jo, Xiaobei Zhao, Katherine A. Hoadley, Scott Newman, Jeremiah Holt, Michele C. Hayward, **Michael I. Love**, J. S. Marron, D. Neil Hayes. SCISSOR: a framework for identifying structural changes in RNA transcripts.
Nature Communications 12(286), (2021). doi: 10.1038/s41467-020-20593-3

11. Joshua P. Zitovsky, **Michael I. Love**. Fast effect size shrinkage software for beta-binomial models of allelic imbalance.
F1000Research, 8:2024, (2020). doi: 10.12688/f1000research.20916.2

12. Nana Matoba, **Michael I. Love**, Jason L. Stein. Evaluating brain structure traits as endophenotypes using polygenicity and discoverability.
Human Brain Mapping, Advance online publication, (2020). doi: 10.1002/hbm.25257
Role: Provided statistical analysis for quantifying trait polygenicity and discoverability.

13. Halei C. Benefield, Katherine E. Reeder-Hayes, Hazel B. Nichols, Benjamin C. Calhoun, **Michael I. Love**, Erin L. Kirk, Joseph Geradts, Katherine A. Hoadley, Stephen R. Cole, H. Shelton Earp, Andrew F. Olshan, Lisa A. Carey, Charles M. Perou, Melissa A. Troester. Outcomes of Hormone-Receptor Positive, HER2-Negative Breast Cancers by Race and Tumor Biological Features.
JNCI Cancer Spectrum, 5(1):pkaa072, (2020). doi: 10.1093/jncics/pkaa072
Role: Assisted with definition of biological subgroups among patients

14. Avi Srivastava, Laraib Malik, Hirak Sarkar, Mohsen Zakeri, Fatemeh Almodaresi, Charlotte Soneson, **Michael I. Love**, Carl Kingsford, Rob Patro. Alignment and mapping methodology influence transcript abundance estimation.
Genome Biology, 21(1):239, (2020). doi: 10.1186/s13059-020-02151-8

15. Hirak Sarkar, Avi Srivastava, Hector Corrada Bravo, **Michael I. Love**, Rob Patro. Terminus enables the discovery of data-driven, robust transcript groups from RNA-seq data.
Bioinformatics, (36)S1:i102–i110, (2020). doi: 10.1093/bioinformatics/btaa448

16. Sarah A. Reifeis*, Michael G. Hudgens, Mete Civelek, Karen L. Mohlke, **Michael I. Love**. Assessing exposure effects on gene expression.
Genetic Epidemiology, 44(6):601–610, (2020). doi: 10.1002/gepi.22324

17. Stuart Lee, Michael Lawrence, **Michael I. Love**. Fluent genomics with plyranges and tximeta.
F1000Research, 9:109, (2020). doi: 10.12688/f1000research.22259.1

18. Charlotte Soneson, Federico Marini, Florian Geier, **Michael I. Love**, Michael B. Stadler. ExploreModelMatrix: Interactive exploration for improved understanding of design matrices and linear models in R.
F1000Research, 9:512, (2020). doi: 10.12688/f1000research.24187.2

19. Arjun Bhattacharya*, Montserrat Garcia-Closas, Andrew F. Olshan, Charles M. Perou, Melissa A. Troester, **Michael I. Love**. A framework for transcriptome-wide association studies in breast cancer in diverse study populations.
Genome Biology, 21(1):42, (2020). doi: 10.1186/s13059-020-1942-6

20. **Michael I. Love**, Charlotte Soneson, Peter F. Hickey, Lisa K. Johnson, N. Tessa Pierce, Lori Shepherd, Martin Morgan, Rob Patro. Tximeta: Reference sequence checksums for provenance identification in RNA-seq.
PLOS Computational Biology, 16(2):e1007664, (2020). doi: 10.1371/journal.pcbi.1007664

21. Sean D. McCabe*, Dan-Yu Lin, **Michael I. Love**. Consistency and overfitting of multi-omics methods on experimental data.
Briefings in Bioinformatics, 21(4):1277–1284, (2020). doi: 10.1093/bib/bbz070

22. Lourdes Cruz-Garcia, Grainne O'Brien, Botond Sipos, Simon Mayes, **Michael I. Love**, Daniel J. Turner, Christophe Badie. Generation of a Transcriptional Radiation Exposure Signature in Human Blood Using Long-Read Nanopore Sequencing.
Radiation Research, 193(2):143–154, (2019). doi: 10.1667/RR15476.1
Role: Developed methods and pipeline for statistical inference of transcript counts.

23. Chelsea K. Raulerson, Arthur Ko, John C. Kidd, Kevin W. Currin, Sarah M. Brotman, Maren E. Cannon, Ying Wu, Cassandra N. Spracklen, Anne U. Jackson, Heather M. Stringham, Ryan P. Welch, Christian Fuchsberger, Adam E. Locke, Narisu Narisu, Aldons J. Lusis, Mete Civelek, Terrence S. Furey, Johanna Kuusisto, Francis S. Collins, Michael Boehnke, Laura J. Scott, Dan-Yu Lin, **Michael I. Love**, Markku Laakso, Päivi Pajukanta, Karen L. Mohlke. Adipose Tissue Gene Expression Associations Reveal Hundreds of Candidate Genes for Cardiometabolic Traits.
Americal Journal of Human Genetics, 105(4):773–787, (2019). doi: 10.1016/j.ajhg.2019.09.001
Role: Developed methods for quantifying tissue heterogeneity of samples. Quality control assessment of gene expression data. Reviewed statistical and bioinformatic methods.

24. Anqi Zhu*, Avi Srivastava, Joseph G. Ibrahim, Rob Patro, **Michael I. Love**. Nonparametric expression analysis using inferential replicate counts.
Nucleic Acids Research, 47(18):e105, (2019). doi: 10.1093/nar/gkz622

25. Charlotte Soneson, **Michael I. Love**, Rob Patro, Shobbir Hussain, Dheeraj Malhotra, Mark D. Robinson. A junction coverage compatibility score to quantify the reliability of transcript abundance estimates and annotation catalogs.
Life Science Alliance, 2(1):e201800175, (2019). doi: 10.26508/lsa.201800175

26. Lindsay A. Williams, Katherine A. Hoadley, Hazel B. Nichols, Joseph Geradts, Charles M. Perou, **Michael I. Love**, Andrew F. Olshan, Melissa A. Troester. Differences in race, molecular and tumor

characteristics among women diagnosed with invasive ductal and lobular breast carcinomas.
Cancer Causes & Control, 30(1):31–39, (2019). doi: 10.1007/s10552-018-1121-1
Role: Reviewed statistical comparisons of molecular tumor characteristics.

27. Yoh Isogai, Zheng Wu, **Michael I.Love**, Michael Ho-Young Ahn, Dhananjay Bambah-Mukku, Vivian Hua, Karolina Farrell, Catherine Dulac. Multisensory Logic of Infant-Directed Aggression by Males.
Cell, 175(7):1827–1841, (2018). doi: 10.1016/j.cell.2018.11.032
Role: Performed statistical analysis of RNA sequencing data.

28. Anqi Zhu*, Joseph G. Ibrahim, **Michael I. Love**. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences.
Bioinformatics, 35(12):2084–2092, (2018). doi: 10.1093/bioinformatics/bty895

29. **Michael I. Love**, Charlotte Soneson, and Rob Patro. Swimming downstream: statistical analysis of differential transcript usage following Salmon quantification.
F1000Research, 7:952, (2018). doi: 10.12688/f1000research.15398.3

30. Yuchen Yang, Ruth Huh, Houston W. Culpepper, Yuan Lin, **Michael I. Love**, Yun Li. SAFE-clustering: Single-cell Aggregated (From Ensemble) Clustering for Single-cell RNA-seq Data.
Bioinformatics, 35(8):1269–1277, (2018). doi: 10.1093/bioinformatics/bty793

31. Alena van Bömmel, **Michael I. Love**, Ho-Ryun Chung, Martin Vingron. coTRaCTE predicts co-occurring transcription factors within cell-type specific enhancers.
PLOS Computational Biology, 14(8):e1006372, (2018). doi: 10.1371/journal.pcbi.1006372

32. Edward W. Pietryk, Kiristin Clement, Marwa Elnagheeb, Ryan Kuster, Kayla Kilpatrick, **Michael I. Love**, Folami Y. Ideraabdullah. Intergenerational response to the endocrine disruptor vinclozolin is influenced by maternal genotype and crossing scheme.
Reproducitive Toxicology, 78:9–19, (2018). doi: 10.1016/j.reprotox.2018.03.005
Role: Supervised statistical analysis of mixed effect model for mouse traits.

33. Koen Van den Berge[1], Fanny Perraudeau[1], Charlotte Soneson, **Michael I. Love**, Davide Risso, Jean-Philippe Vert, Mark D. Robinson, Sandrine Dudoit[†], Lieven Clement[†]. Observation weights unlock bulk RNA-seq tools for zero inflation and single-cell applications.
Genome Biology, 19(1):24, (2018). doi: 10.1186/s13059-018-1406-4

34. Lindsay A. Williams, Hazel B. Nichols, Katherine A. Hoadley, Chiu Kit Tse, Joseph Geradts, Mary Elizabeth Bell, Charles M. Perou, **Michael I. Love**, Andrew F. Olshan, Melissa A. Troester. Reproductive risk factor associations with lobular and ductal carcinoma in the Carolina Breast Cancer Study.
Cancer Causes & Control, 29(1):25–32, (2017). doi: 10.1007/s10552-017-0977-9
Role: Reviewed statistical comparisons of molecular tumor characteristics.

35. Doug H. Phanstiel, Kevin Van Bortle, Damek Spacek, Gaelen T. Hess, Muhammad S. Shamim, Ido Machol, **Michael I. Love**, Erez L. Aiden, Michael C. Bassik, Michael P. Snyder. Static and Dynamic DNA Loops form AP-1-Bound Activation Hubs during Macrophage Development.
Molecular Cell, 67(6):1037–1048, (2017). doi: 10.1016/j.molcel.2017.08.006
Role: Performed statistical analysis of differential chromatin conformation from DNA-DNA sequence counts.

36. Rob Patro, Geet Duggal, **Michael I. Love**, Rafael A. Irizarry, Carl Kingsford, Salmon provides fast and bias-aware quantification of transcript expression.
Nature Methods, 14(4):417–419, (2017). doi: 10.1038/nmeth.4197

37. **Michael I. Love**, Matthew Huska, Marcel Jurk, Robert Schopflin, Stephan Starick, Kevin Schwahn, Samantha Cooper, Keith Yamamoto, Morgane Thomas-Chollier, Martin Vingron, Sebastiaan Meijsing. Role of the chromatin landscape and sequence in determining cell type-specific genomic glucocorticoid receptor binding and gene regulation.
Nucleic Acids Research, 45(4):1805–1819, (2016). doi: 10.1093/nar/gkw1163

38. **Michael I. Love**, John B. Hogenesch, Rafael A. Irizarry, Modeling of RNA-seq fragment sequence bias reduces systematic errors in transcript abundance estimation.
Nature Biotechnology, 32(12):1287–1291, (2016). doi: 10.1038/nbt.3682

39. Leonardo Collado Torres, Abhinav Nellore, Alyssa C. Frazee, Christopher Wilks, **Michael I. Love**, Ben Langmead, Rafael A. Irizarry, Jeffrey Leek, Andrew E. Jaffe, Flexible expressed region analysis for RNA-seq with derfinder.
Nucleic Acids Research, 45(2):e9, (2016). doi: 10.1093/nar/gkw852

40. Mingxiang Teng, **Michael I. Love**, Carrie A. Davis, Sarah Djebali, Alexander Dobin, Brenton R. Graveley, Sheng Li, Christopher E. Mason, Sara Olson, Dmitri Pervouchine, Cricket A. Sloan, Xintao Wei, Lijun Zhan, Rafael A. Irizarry, A benchmark for RNA-seq quantification pipelines.
Genome Biology, 17(1):74, (2016). doi: 10.1186/s13059-016-0940-1

41. Charlotte Soneson, **Michael I. Love**, Mark D. Robinson, Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences.
F1000Research, 4:1521, (2015). doi: 10.12688/f1000research.7563.1

42. **Michael I. Love**, Simon Anders, Vladislav Kim, Wolfgang Huber, RNA-seq workflow: gene-level exploratory analysis and differential expression.
F1000Research, 4:1070, (2015). doi: 10.12688/f1000research.7035.1

43. Raman Kumar, Mark A. Corbett, Bregje WM van Bon, Alison Gardner, Joshua A. Woenig, Lachlan A. Jolly, Evelyn Douglas, Kathryn Friend, Chuan Tan, Hilde Van Esch, Maureen Holvoet, Martine Raynaud, Michael Field, Melanie Leffler, Bartłomiej Budny, Marzena Wisniewska, Magdalena Badura-Stronka, Anna Latos-Bieleńska, Jacqueline Batanian, Jill A. Rosenfeld, Lina Basel-Vanagaite, Corinna Jensen, Melanie Bienek, Guy Froyen, Reinhard Ullmann, Hao Hu, **Michael I. Love**, Stefan A. Haas, Pawel Stankiewicz, Sau Wai Cheung, Anne Baxendale, Jillian Nicholl, Elizabeth M Thompson, Eric Haan, Vera M Kalscheuer, Jozef Gecz, Increased STAG2 dosage defines a novel cohesinopathy with intellectual disability and behavioural problems.
Human Molecular Genetics, 24(25):7171–7181, (2015). doi: 10.1093/hmg/ddv414
Role: Designed and implemented algorithm for detection of de novo structural variants, applied to XLID patient cohort.

44. Stephan R. Starick, Jonas Ibn-Salem, Marcel Jurk, Céline Hernandez, **Michael I. Love**, Ho-Ryun Chung, Martin Vingron, Morgane Thomas-Chollier, Sebastiaan H. Meijsing, ChIP-exo signal associated with DNA-binding motifs provide insights into the genomic binding of the glucocorticoid receptor and cooperating transcription factors.
Genome Research, 25(6):825–835, (2015). doi: 10.1101/gr.185157.114
Role: Performed initial statistical analysis of ChIP-exo reads at transcription factor binding sites, produced Figures 1C-E.

45. Wolfgang Huber, Vincent J. Carey, Robert Gentleman, Simon Anders, Marc Carlson, Benilton S. Carvalho, Hector Corrada Bravo, Sean Davis, Laurent Gatto, Thomas Girke, Raphael Gottardo, Florian Hahne, Kasper D. Hansen, Rafael A. Irizarry, Michael Lawrence, **Michael I. Love**, James MacDonald, Valerie Obenchain, Andrzej K. Oleś, Hervé Pagès, Alejandro Reyes, Paul Shannon,

Gordon K. Smyth, Dan Tenenbaum, Levi Waldron, Martin Morgan, Orchestrating high-throughput genomic analysis with Bioconductor.
Nature Methods, 12(2):115–121, (2015). doi: 10.1038/nmeth.3252

46. Hao Hu, Stefan A Haas, Jamel Chelly, Hilde Van Esch, Martine Raynaud, Arjan PM de Brouwer, Stefanie Weinert, Guy Froyen, Suzanna GM Frints, Frédéric Laumonnier, Tomasz Zemojtel, **Michael I. Love**, Hughes Richard, Anne-Katrin Emde, Melanie Bienek, Corinna Jensen, Melanie Hambrock, Ute Fischer, Claudia Langnick, Mirjam Feldkamp, Willemijn Wissink-Lindhout, Nicolas Lebrun, Laetitia Castelnau, Julien Rucci, Rodrick Montjean, Olivier Dorseuil, Pierre Billuart, Till Stuhlmann, Marie Shaw, Mark A Corbett, Alison Gardner, Saffron Willis-Owen, Chuan Tan, Kathryn L Friend, Stefanie Belet, Kees EP van Roozendaal, Mélanie Jimenez-Pocquet, Marie-Pierre Moizard, Nathalie Ronce, Ruping Sun, Sean O'Keeffe, Ramu Chenna, Alena van Bömmel, Jonathan Göke, Anna Hackett, Michael Field, Louise Christie, Jackie Boyle, Eric Haan, John Nelson, Gillian Turner, Gareth Baynam, Gabriele Gillessen-Kaesbach, Ulrich Müller, Daniela Steinberger, Bartlomiej Budny, Magdalena Badura-Stronka, Anna Latos-Bieleńska, Lilian B Ousager, Peter Wieacker, Germán Rodríguez Criado, Marie-Louise Bondeson, Göran Annerén, Andreas Dufke, Monika Cohen, Lionel Van Maldergem, Catherine Vincent-Delorme, Bernard Echenne, Brigitte Simon-Bouy, Tjitske Kleefstra, Michèl Willemsen, Jean-Pierre Fryns, Koenraad Devriendt, Reinhard Ullmann, Martin Vingron, Klaus Wrogemann, Thomas F Wienker, Andreas Tzschach, Hans van Bokhoven, Jozef Gécz, Thomas J Jentsch, Wei Chen, Hans-Hilger Ropers, Vera M Kalscheuer, X-exome sequencing of 405 unresolved families identifies seven novel intellectual disability genes.
Molecular Psychology, 21(1):133, (2015). doi: 10.1038/mp.2014.193
Role: Designed and implemented algorithm for detection of de novo structural variants, applied to XLID patient cohort.

47. **Michael I. Love**, Wolfgang Huber, Simon Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.
Genome Biology, 15(12):550+, (2014). doi: 10.1186/s13059-014-0550-8

48. Wei Li, Han Xu, Tengfei Xiao, Le Cong, **Michael I. Love**, Feng Zhang, Rafael A. Irizarry, Jun S. Liu, Myles Brown, Xiaole X. Liu, MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens.
Genome Biology, 15(12):554+, (2014). doi: 10.1186/s13059-014-0554-4

49. Jonas Ibn-Salem[1], Sebastian Köhler[1], **Michael I. Love**, Ho-Ryun Chung, Ni Huang, Matthew E. Hurles, Melissa Haendel, Nicole L. Washington, Damian Smedley, Christopher J. Mungall, Suzanna E. Lewis, Claus-Eric Ott, Sebastian Bauer, Paul N. Schofield, Stefan Mundlos, Malte Spielmann, Peter N. Robinson, Deletions of chromosomal regulatory boundaries are associated with congenital disease.
Genome Biology, 15(9):423+, (2014). doi: 10.1186/s13059-014-0423-1
Role: Created cell-type-specific map of chromatin accessibility across Roadmap Epigenomic samples, leveraging sparse data storage and statistical computation.

50. Owen D. Solberg[1], Edwin J. Ostrin[1], **Michael I. Love**, Jeffrey C. Peng, Nirav R. Bhakta, Lydia Hou, Christine Nguyen, Margaret Solon, Cindy Nguyen, Andrea J. Barczak, Lorna T. Zlock, Denitza P. Blagev, Walter E. Finkbeiner, K. Mark Ansel, Joseph R. Arron, David J. Erle, Prescott G. Woodruff, Airway epithelial miRNA expression is altered in asthma.
American Journal of Respiratory and Critical Care Medicine, 186(10):965–974, (2012).
doi: 10.1164/rccm.201201-0027oc
Role: Performed statistical analysis of miRNA microarray data.

51. Ruping Sun, **Michael I. Love**, Tomasz Zemojtel, Anne-Katrin Emde, Ho-Ryun Chung, Martin Vingron, Stefan A. Haas, Breakpointer: Using local mapping artifacts to support sequence breakpoint discovery from single-end reads.
Bioinformatics, 28(7):1024–1025, (2012). doi: 10.1093/bioinformatics/bts064

52. **Michael I. Love**, Alena Mysickova, Ruping Sun, Vera Kalscheuer, Martin Vingron, Stefan A. Haas, Modeling read counts for cnv detection in exome sequencing data.
Statistical Applications in Genetics and Molecular Biology, 10(1):52, (2011). doi: 10.2202/1544-6115.1732

Digital and other novel forms of scholarship:

Open source software

The following list comprises open source software packages maintained by the Love Lab. The number of monthly unique IP downloads are listed for software packages maintained by the Love Lab that are widely used in genomics research pipelines in academia and industry.

Key:
(Bioc) - R package within the Bioconductor Project;
(GitHub) - software hosted as a GitHub repository.

1. **DESeq2**: Differential expression analysis for RNA-seq. (Bioc)
   - **Greater than 19,000 monthly unique IP downloads**
   - 5th most accessed article published by *Genome Biology*
   - Awarded and funded as a CZI *Essential Open Source Software* (EOSS) project, out of 72 projects
   - Press: "Love wins CZI funding for open source software to analyze genomics data"
     UNC Gillings School of Public Health News
2. **apeglm**: Approximate posterior estimation for GLM coefficients. (Bioc)
   - **Greater than 3,000 monthly unique IP downloads**
3. **tximport**: Import and summarization of transcript-level estimates for gene-level analysis. (Bioc)
   - **Greater than 3,000 monthly unique IP downloads**
4. **tximeta**: Transcript data import with automatic metadata for computational reproducibility. (Bioc)
   - **Greater than 1,200 monthly unique IP downloads**
5. **fishpond**: Nonparametric differential transcript expression with inferential replicates. (Bioc)
6. **airpart**: Differential cell-type-specific allelic imbalance using single-cell allelic count data. (Bioc)
7. **MRLocus**: Mendelian Randomization analysis at the genomic locus, or *cis*-MR. (GitHub)
8. **MOSTWAS**: Multi-Omic Strategies for TWAS. (GitHub)
9. **DeCompress**: Tissue compartment deconvolution for targeted RNA panels. (GitHub)
10. **ACTOR**: LDA model to compare isoform expression to a reference panel. (GitHub)

11. **alpine**: Correction of fragment sequence bias for RNA-seq transcript abundance estimation. (Bioc)
12. **exomeCopy**: Detection of copy number variants for Exome-seq. (Bioc)

Data science workflows

Reproducible genomic data science workflows maintained or co-authored by the Love Lab. The workflows are "literate programming" documents, which weave prose, figures, and code together to demonstrate completion of a data science task. These are submitted and hosted both on Bioconductor servers, where they are tested against current R and package ecosystem, and on *F1000Research* where they are subjected to peer-review.

1. **RNA-seq workflow: gene-level exploratory analysis and differential expression**
   - Bioconductor rendered workflow (the top downloaded Bioconductor workflow, out of 29)
   - F1000Research peer-reviewed publication (more than 61,000 views)
   - Press: Gibbs, W. Wayt, "A test drive of a DNA-analysis toolkit in the cloud". *Nature*, 552:137–138, (2017). doi: 10.1038/d41586-017-07833-1
     "*...Also available are peer-reviewed tutorials, known as workflows, which are updated as the platform evolves. One, co-authored by Love, walks readers through a differential-expression analysis of RNA-sequencing data. I used his workflow to guide my exploration.*"
2. **RNA-seq workflow for differential transcript usage following Salmon quantification**
   - Bioconductor rendered workflow
   - F1000Research peer-reviewed publication (more than 14,000 views)
3. **Fluent genomics with plyranges and tximeta**
   (first author: Stuart Lee, of Monash University)
   - Bioconductor rendered workflow
   - F1000Research peer-reviewed publication (more than 900 views)

Invited presentations

1. September 2021 - Dept. of Biostatistics, Virginia Commonwealth University
2. September 2021 - ICSA Applied Statistics Symposium - Organized Invited Sesssion "Integrative multi-omics inference"
3. August 2021 - JSM, "Genomics in Neuroscience"
4. June 2021 - Genetics, Genomics and Informatics Virtual Seminar, The University of Tennessee Health Science Center
5. December 2020 - Dept. of Statistics, Penn State University
6. December 2020 - Institute for Personalized Medicine Seminar Series, Icahn School of Medicine, Mount Sinai Hospital

7. June 2020, Banff International Research Station (BIRS), "Mathematical Frameworks for Integrative Analysis of Emerging Biological Data Types"
8. March 2020 - ENAR, "Recent advanced and opportunities in large scale and multi-omic single cell data analysis"
9. March 2020 - Human Genetics Seminar Series, Dept. of Genetics, University of Utah
10. December 2019 - Statistical Genetics Seminar, Dept. of Biostatistics, Johns Hopkins University
11. November 2019 - APHA, Spiegelman Awardee Invited Session
12. July 2019 - JSM, "Making an Impact in Statistics Education through Innovation and Outreach"
13. May 2019 - ASA Symposium on Data Science & Statistics, "Democratizing Data Science through Workflows"
14. March 2019 - ENAR, "Teaching Data Science through Case Studies"
15. March 2019 - Dept. of Biostatistics and Bioinformatics, Emory University
16. February 2019 - Epigenetics and Stem Cell Biology, NIEHS
17. November 2018 - Epigenetics and Epigenomics Program, Duke University
18. October 2018 - Computational Biology & Bioinformatics, Duke University
19. September 2018 - School of Public Health Statistics Seminar, Brown University
20. May 2018 - Bioinformatics Research Center, NCSU
21. March 2018 - ENAR, "Teaching Data Science at all Levels"
22. February 2018 - Center for Public Health Genomics, UVA
23. December 2016 - Computational Biology, Bioinformatics & Genomics, UMD
24. October 2016 - Triangle Statistical Genetics Meeting, Cary, NC

Non-refereed papers/articles:

The following are non-refereed papers/articles and all titles are indicated with an asterisk.

1. Kim-Anh Lê Cao, Al J. Abadi, Emily F. Davis-Marcisak, Lauren Hsu, Arshi Arora, Alexis Coullomb, Atul Deshpande, Yuzhou Feng, Pratheepa Jeganathan, Melanie Loth, Chen Meng, **Wancen Mu**, Vera Pancaldi, Kris Sankaran, Amrit Singh, Joshua S. Sodicoff, Genevieve L. Stein-O'Brien, Ayshwarya Subramanian, Joshua D. Welch, Yue You, Ricard Argelaguet, Vincent J. Carey, Ruben Dries, Casey S. Greene, Susan Holmes, **Michael I. Love**, Matthew E. Ritchie, Guo-Cheng Yuan, Aedin C. Culhane, Elana Fertig. Community-wide hackathons to identify central themes in single-cell multi-omics.
Genome Biology, (2021). doi: 10.1186/s13059-021-02433-9
2. Anushka Rajesh, Yutong Chang, Malak S. Abedalthagafi, Annie Wong-Beringer, **Michael I. Love**, Serghei Mangul. Improving the completeness of public metadata accompanying omics studies.
Genome Biology, 22(1):106, (2021). doi: 10.1186/s13059-021-02332-z

3. Achal Patel, Montserrat García-Closas, Andrew F. Olshan, Charles M. Perou, Melissa A. Troester, **Michael I. Love**, Arjun Bhattacharya. Transcriptome-wide association study of risk of recurrence in Black and White breast cancer patients*.
medRxiv, March 2021. doi: 10.1101/2021.03.19.21253983

4. Miheer Dewaskar, John Palowitch, Mark He, **Michael I. Love**, Andrew Nobel. Finding stable groups of cross-correlated features in multi-view data*.
arXiv, September 2020. arXiv: 2009.05079

5. Koen Van Den Berge, Katharina M. Hembach, Charlotte Soneson, Simone Tiberi, Lieven Clement, **Michael I. Love**, Rob Patro, and Mark D. Robinson. RNA sequencing data: hitchhiker's guide to expression analysis*.
Annual Review of Biomedical Data Science, October 2018. doi: 10.1146/annurev-biodatasci-072018-021255

TEACHING RECORD

Classroom teaching:

- BIOS 784 / BCB 784: *Computational Biology*, elective, 3 credits.
Instructor, Spring 2021.
21 students enrolled.
Notes and syllabus: https://biodatascience.github.io/compbio

- BIOS 663: *Intermediate Linear Models*, required for Masters, 4 credits.
Instructor, Spring 2020.
46 students enrolled.

- BIOS 735: *Statistical Computing*, required for PhD, 4 credits.
Co-Instructor with Dr. Naim Rashid, Spring 2019.
23 students enrolled.
Notes and syllabus: https://biodatascience.github.io/statcomp

- BIOS 784: *Computational Biology*, elective, 3 credits.
Instructor, Fall 2017.
28 students enrolled.
Notes and syllabus: https://biodatascience.github.io/compbio

Graduate students supervised:

- Euphy Wu, PhD in Biostatistics.
- Wancen Mu, PhD in Biostatistics.
- Ji-Eun Park, PhD in Biostatistics. (Co-supervised with Di Wu)
- Alex Lockhart, M.S. in Biostatistics, Graduated Fall 2020.
- Sarah Reifeis, PhD in Biostatistics, Graduated Fall 2020. (Co-supervised with Dr. Hudgens)
- Arjun Bhattacharya, PhD in Biostatistics, Graduated Summer 2020.

- Scott Van Buren, PhD in Biostatistics, Graduated Summer 2020. (Co-supervised with Dr. Rashid)
- Sean McCabe, PhD in Biostatistics, Graduated Spring 2020. (Co-supervised with Dr. Lin)
- Anqi Zhu, PhD in Biostatistics, Graduated Fall 2019. (Co-supervised with Dr. Ibrahim)

GRANTS

All award amounts are direct costs for the current project year.

Ongoing:

- **Co-Investigator** 1-R01-CA253450-01A1 (Troester) 04/01/2021 - 03/31/2026
  NIH (NCI) $391,413
  P53, DNA Repair, and Immune Response in Breast Cancer Mortality Disparities
  10% effort
- **Co-Investigator** 1-R01-MH125236-01 (Crawford) 04/01/2021 - 01/31/2026
  NIH (NIMH) $1,225,816
  Beyond GWAS: High Throughput Functional Genomics and Epigenome Editing to Elucidate the Effects of Genetic Associations for Schizophrenia
  8% effort
- **Principal Investigator** Essential Open Source Software (Love) 01/01/2021 - 12/31/2021
  Chan Zuckerberg Initiative $74,000
  10% effort
- **Principal Investigator** 1-R01-MH118349-01 (Stein) 12/10/2018 - 11/30/2023
  NIH (NIMH) $303,438
  pathQTL: Integrative Multi-Omics Causal Inference of Molecular Mechanisms Leading to Neuropsychiatric Illness
  25% effort
- **Principal Investigator** 1-R01-HG009937-01A1 (Patro) 09/18/2018 - 06/30/2023
  NIH (NHGRI) $256,486
  A Modular Framework for Accurate, Efficient, and Reproducible Analysis of RNA-Seq Data
  35% effort
- **Co-Investigator** 5-P50-CA058223-26 (Perou) 09/01/2018 - 08/31/2023
  NIH (NCI) $1,531,225
  SPORE in Breast Cancer - Project 1: The Carolina Breast Cancer Study (CBCS)
  5% effort
- **Co-Investigator** 5-R01-DK093757-09 (Mohlke) 08/01/2017 - 05/31/2022
  NIH (NIDDK) $477,414
  Genetic Epidemiology of Rare and Regulatory Variants for Metabolic Traits
  5% effort

Completed:

- **Co-Investigator** 2-P30-ES010126-20 (Troester) 04/01/2021 - 02/28/2026
  NIH (NIEHS) $302,838
  UNC-CH Center for Environmental Health and Susceptibility - Biostatistics and Bioinformatics Facility Core
  15% effort

- **Co-Investigator** 5-P01-CA142538-10 (Lin) 04/01/2019 - 03/31/2020
  NIH (NCI) $416,534
  Statistical Methods for Cancer Clinical Trials - Project 3: Statistical/Computational Methods for Pharmacogenomics and Individualized Therapy
  10% effort

- **Co-Investigator** 5-UL1-TR002489-02 (Buse) 03/01/2019 - 02/29/2020
  NIH (NCATS) $6,281,669
  NC TraCS Institute
  10% effort

- **Co-Investigator** 5-R01-HG009125-03 (Nobel) 09/07/2016 - 06/30/2020
  NIH (NHGRI) $299,936
  Multi-Tissue and Network Models for Next-Generation EQTL Studies
  10% effort

SERVICE

Professional service:

- **Member** of Technical Advisory Board for Bioconductor Project (June 2020 - present)
- **Member** of organizing committee for BioC conference 2021.
- **Reviewer** for BRAIN Initiative, NIH NIMH Review Panel, June 11, 2021.
- **Temporary Member** of Biodata Management and Analysis (BDMA) NIH Study Section, October 22-23, 2020.
- **Developer and maintainer** for widely-used open-source software packages. Maintenance of software packages entails addressing feature requests as well as software support requests. DESeq2 typically receives more than 10 support requests per week on https://support.bioconductor.org.
- **Referee** for the following journals:

  Bioinformatics, Nucleic Acids Research, Genome Biology, Genome Research, Journal of the American Statistical Association, Biometrics, American Journal for Human Genetics, PLOS Computational Biology, PLOS Genetics, Statistical Applications in Genetics and Molecular Biology, Nature Biotechnology, Nature Methods, Nature Communications, Nature Reviews Molecular Cell Biology, Cell Reports, BMC Bioinformatics, Briefings in Bioinformatics, F1000Research, R Journal.

Departmental service:

- **Member** - Biostatistics Masters Exam Committee (January 2020 - present)
- **Member** - Biostatistics Computing Committee (August 2016 - present)
- **Co-Chair** - Biostatistics Seminar Committee (August 2016 - July 2020)
- **Member** - PhD / DrPH dissertation committees for 28 graduated trainees, and 17 currently enrolled trainees, predominantly in Biostatistics (18 trainees) and the BCB program (16 trainees). Chair for 5 PhD committees.
- **Member** - Joint Biostatistics/Genetics Faculty Search Committee (2017)
- **Member** - Genetics Statistical / Computational Faculty Search Committee (2017)
- **Judge** for UNC MPH Core Poster Session (2020, 2021).
- **Reviewer** for University Cancer Research Fund (UCRF) School of Public Health Student Award (2018).

Community engagement:

- **UNC PREP Mentor** to Kwame Forbes (PREP year 2020-2021).

  Diversity, equity, and inclusion: The UNC Postbaccalaureate Research Education Program (PREP) prepares post-baccalaureates from groups traditionally under-represented (UR) in the sciences for entry and success in top Biomedical PhD programs. UNC PREP has been running for 12 years,

and I served as a mentor to Mr. Forbes for the previous program year. Mr. Forbes applied to Bioinformatics PhD programs, and will enroll in UNC Bioinformatics and Computational Biology as a PhD student in Fall 2021.

I have also written two proposal for seeking additional funding to support the program itself, one of which is currently under review by the Chan Zuckerberg Initiative (RAMSeS proposals: 20-3116 and 21-4888, the latter with direct cost of $338,238 over two years).

- **Instructor/organizer** of "Statistical Methods for Functional Genomics" (two weeks NIH-sponsored course, 24 international students), Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, for the years 2017, 2018, 2019.

  Diversity, equity, and inclusion: The Statistical Methods for Functional Genomics course has a diversity mandate from its funders (NIH NIGMS funded via R25-GM100958) to support a diverse cohort of 24 students, including representation of women, individuals from under-represented groups and geographical regions. Each year, the organizers would take special care to select individual from up to 200 applications that would support a diverse cohort.

- **Instructor/organizer** of "Statistics and Computing in Genome Data Science" (5 day course, Bioconductor-sponsored event), Bressanone-Brixen, Italy, for the years: 2015, 2016, 2018.

- **Workshop lead** for RNA-seq and genomic data science topic at numerous events including: BioC conference 2014, 2015, 2016, 2018, 2019, 2020, R-Ladies Tunis 2020 (Tunisian local chapter of R-Ladies Global), ENAR 2018, UC Davis ANGUS 2017, Harvard Medical School 2015, Harvard School of Public Health 2015. I have also given 1 hour lectures in genomic data science to the UNC PREP cohort (2020) and the UNC Educational Pathways to Diversity in Genomics (EDGE) cohort (2021).

  Diversity, equity, and inclusion: the R-Ladies Tunis workshop was designed to provide introductory bioinformatics training to the R-Ladies chapter in Tunis, Tunisia. The mission of R-Ladies is "to achieve proportionate representation by encouraging, inspiring, and empowering people of genders currently underrepresented in the R community." Likewise, I have and will continue to find opportunities to present introductory genomic data science lectures and workshops to local programs encouraging diversity in Biostatistics, Computational Biology, and Genetics.

  Examples of publicly available workshop material:

  - BioC2020 - Importing alevin scRNA-seq counts into R/Bioconductor
  - BioC2019 - Overlapping differential expression and differential chromatin accessibility
  - BioC2018 - RNA-seq data analysis with DESeq2, MultiQC, Glimma, and ZINB-WaVE
  - BioC2016 - Low-level exploratory data analysis and methods development for RNA-seq
  - BioC2015 - Differential expression, manipulation, and visualization of RNA-seq reads

PROFESSIONAL SOCIETY MEMBERSHIPS

- Eastern North American Region (ENAR), 2016 - present.
- American Statistical Association (ASA), 2015 - present.
- American Society of Human Genetics (ASHG), 2021.