Curriculum Vitae

Notarization. I have read the following and certify that this curriculum vitae is a current and accurate statement of my professional record.

_____   April 17, 2019
Signature   _____
                                   Date

I. Personal Information

I.A. UID, Last Name, First Name, Middle Name, Contact Information

Last Name: Corrada Bravo
First Name: Héctor
3114F Biomolecular Sciences Building #296
University of Maryland
College Park, MD 20745

hcorrada@umiacs.umd.edu
http://hcbravo.org
UID: 111544240

I.B. Academic Appointments at UMD

- Associate Professor with Tenure, Department of Computer Science
  University of Maryland, College Park, MD
  July 2016-Present

- Associate Professor with Tenure, Institute for Advanced Computer Studies
  University of Maryland, College Park, MD
  July 2016-Present

- Affiliate Associate Professor, Applied Math, Statistics and Scientific Computation
  University of Maryland, College Park, MD
  July 2016-Present

- Affiliate Associate Professor, Department of Mathematics
  University of Maryland, College Park, MD
  July 2018-Present

- Assistant Professor, Department of Computer Science
  University of Maryland, College Park, MD
  July 2010-June 2016

- Assistant Professor, Institute for Advanced Computer Studies
  University of Maryland, College Park, MD
  July 2010-June 2016

- Affiliate Assistant Professor, Applied Math, Statistics and Scientific Computation
  University of Maryland, College Park, MD
  July 2011-June 2016

I.D.   Other Employment

- Postdoctoral Fellow, Department of Biostatistics
  Johns Hopkins University School of Public Health, Baltimore, MD
  September 2008-June 2010

- Research and Teaching Assistant, Departments of Computer Science and Statistics
  University of Wisconsin, Madison, WI
  September 2003-August 2008

- Resarch Intern,
  IBM Research, Almaden, CA
  May 2005-August 2005

I.E.   Educational Background

- Ph.D., Computer Science, University of Wisconsin, Madison, WI
  September 2003 - August 2008
  Dissertation: Graph-based data analysis
  Advisor: Grace Wahba and Raghu Ramakrishnan

- D.M.A., Indiana University School of Music, Bloomington, IN (ABD)
  September 2000 - August 2003

- M.M., Peabody Institute of Music, Baltimore, MD
  September 1997 - May 1999

- B.M., Peabody Institute of Music, Baltimore, MD
  September 1993 - May 1997

II.   Research, Scholarly, Creative and Professional Activites

Annotation: corresponding authors*, lead authors^, advisee#

II.C.   Articles in Refereed Journals

1. N.D. Olson^#, N. Shah, J. Kancherla#, J.N. Paulson#, J. Wagner#, **H. Corrada Bravo*** (2019). metagenomeFeatures: An R package for working with 16S rRNA reference databases and marker-gene survey feature data. *Bioinformatics* btz136, doi:10.1093/bioinformatics/btz136.

2. J.S. Meisel^, D.J. Nasko, B. Brubach, V. Cepeda Espinoza, J. Chopyk, **H. Corrada-Bravo**, M. Fedarko, J. Ghurye, K. Javkar, N.D. Olson#, N. Shah, S.M. Allard, A.L. Bazinet, N.H. Bergman, A. Brown, J.G. Caporaso, S. Conlan, J. DiRuggiero, S.P. Forry, N.A. Hasan, J. Kralj, P.M. Leuthy, D.K. Milton, B.D. Ondov, S. Preheim, S. Ratnayake, S.M. Rogers, M.J. Rosovitz, E.G. Sakowski, N.O. Schliebs, D.D. Sommer, K.L. Ternus, G. Uritskiy, S.X. Zhang, M. Pop, T.J. Treangen* (2018). Current progress and future opportunities in applications of DNA sequencing for biodefense and pathogen detection: Report from the Winter Mid-Atlantic Microbiome Meet-up, College Park, MD January 10th, 2018. *Microbiome* 6:197, doi:10.1186/s40168-018-0582-5.

3. M.S. Kumar^#, E.V. Slud, K. Okrah^#, S.C. Hicks, S. Hannenhalli, **H. Corrada Bravo*** (2018). Analysis and correction of compositional bias in sparse sequencing count data. *BMC Genomics* 19:799, doi:10.1186/s12864-018-5160-5.

4. A.E. Nazario-Toole^, J. Robalino, K. Okrah^#, **H. Corrada Bravo**, S.M. Mount, L.P. Wu* (2018). The splicing factor RNA-binding Fox Protein 1 mediates the cellular immune response in Drosophila melanogaster. *The Journal of Immunology* , doi:10.4049/jimmunol.1800496.

5. S. Hicks^, K. Okrah^#, J. Paulson^#, J. Quackenbush, R. Irizarry, **H. Corrada Bravo*** (2018). Smooth quantile normalization. *Biostatistics* 19(2): 185-198, doi:10.1093/biostatistics/kxx028.

6. J. Wagner^#, F. Chelaru^#, J. Kancherla^, J.N. Paulson^#, A. Zhang^#, V. Felix, A. Mahurkar, N. Elmqvist, **H. Corrada Bravo*** (2018). Metaviz: interactive statistical and visual analysis of metagenomic data. *Nucleic Acids Research* gky136, doi:10.1093/nar/gky136.

7. J. Kancherla^#, A. Zhang^#, B. Gottfried^#, **H. Corrada Bravo*** (2018). Epiviz Web components: reusable and extensible component library to visualize genomic datasets. *F1000 Research* 7:1096, doi:10.12688/f1000research.15433.1.

8. S. Braid^, K. Okrah^#, A. Shetty, **H. Corrada Bravo*** (2017). Differences in DNA methylation patterns in the cord blood of neonates across gestational age are associated with differences in cell-type composition. *Nursing Research* 66(2): 115-122, doi:10.1097/NNR.0000000000000210.

9. S. Manimaran^, H.M. Selby, K. Okrah^#, C. Ruberman, J.T. Leek, J. Quackenbush, B. Haibe-Kains, **H. Corrada Bravo**, W.E. Johnson* (2016). BatchQC: interactive software for evaluating sample and batch effects in genomic data. *Bioinformatics* 32(24):3836-3838, doi:10.1093/bioinformatics/btw538.

10. M. Pop^, J.N. Paulson^#, S. Chakraborty, I. Astrovskaya, B.R. Lindsay, S. Li, **H. Corrada Bravo**, C. Harro, J. Parkhill, A.W. Walker, R.I. Walker, D.A. Sack, O.C. Stine* (2016). Individual-specific changes in the human gut microbiota after challenge with enterotoxigenic *Escherichia coli* and subsequent ciprofloxacin treatment. *BMC Genomics* 17(1):440, doi:10.1186/s12864-016-27777-0.

11. M. Sharmin^#, **H. Corrada bravo**, S. Hannenhalli* (2016). Heterogeneity of TF binding specificity models within and across cell lines. *Genome Research* 26(8):1110-23, doi:10.1101/gr.199166.115.

12. M.C. Fernandes^, L.A. Dillon^, A.T. Belew, **H. Corrada Bravo**, D.M. Mosser, N. El-Sayed* (2016). Dual transcriptome profiling of human macrophages infected with *Leishmania* parasites reveals transient reprogramming events followed by a generic phagocytic response. *mBio* 7(3). pii: e00027-16, doi:10.1128/mBio.00027-16.

13. Y. Li^, K.L. Caradonna^, K. Okrah^#, J. Choi, P. Padmanabhan, D.M. Ndegwa, S. Shah-Simpson, M.R. Temanni, **H. Corrada Bravo**, N. El-Sayed*, B.A. Burleigh* (2016). Transcriptome Remodeling in Trypanosoma cruzi and Human Cells during Intracellular Infection. *PLOS Pathogens* 12(4):e1005511, doi:10.1371/journal.ppat.1005511.

14. J. Wagner^#, J. Paulson^#, X.S. Wang, B. Bhattacharjee, **H. Corrada Bravo*** (2016). Privacy preserving human microbiome analysis using secure computation. *Bioinformatics* , doi:10.1093/bioinformatics/btw073.

15. M. Sharmin^, **H. Corrada Bravo***, S. Hannenhalli* (2016). Characterizing the regulatory landscape of large hypomethylated blocks in colon cancer. *BMC Cancer* 16:88, doi:10.1186/s12885-016-2128-1.

16. L.A. Dillon^, R. Suresh, K. Okrah^#, **H. Corrada Bravo**, D.M. Mosser, N.M. El-Sayed* (2015). Simultaneous transcriptional profiling of Leishmania major and its murine macrophage host cell reveals insights into host-pathogen interactions. *BMC Genomics* 16:1108, doi:10.1186/s12864-015-2237-2.

17. L. Dillon^, K. Okrah^#, V. Hughitt, R. Suresh, Y. Li, M. Fernandes, A. Belew, **H. Corrada Bravo**, D. Mosser, N. El-Sayed* (2015). Transcriptomic profiling of gene expression and RNA processing during *Leishmania major* differentiation. *Nucleic Acids Research* 43 (14): 6799-6813, doi:10.1093/nar/gkv656.

18. W. Dinalankara^#, **H. Corrada Bravo*** (2015). Gene expression signatures based on variability can robustly predict tumor progression and prognosis. *Cancer Informatics* 14:71-81.

19. W. Huber^, V. Carey, R. Gentleman, M. Carlson, B.S. Carvalho, **H. Corrada Bravo**, S. Davis, L. Gatto, T. Girke, R. Gottardo, F. Hahne, K. Hansen, R.A. Irizarry, M. Lawrence, M.I. Love, J. MacDonald, V. Obenchain, A.K. Olés, H. Pagés, P. Shannon, G. Smyth, D. Tenenbaum, L. Waldron, M. Morgan* (2015). Orchestrating high-throughput genomic analysis with Bioconductor. *Nature Methods* 12 115-121, doi:10.1038/nmeth.3252.

20. K. Okrah^#, **H. Corrada Bravo*** (2015). Shape analysis for high-throughput transcriptomics experiment data. *Biostatistics* 16(4):627-40, doi:10.1093/biostatistics/kxv018.

21. C. Ye^#, C. Hsiao^#, **H. Corrada Bravo*** (2014). BlindCall: ultra-fast base-calling of second-generation sequencing by blind deconvolution. *Bioinformatics* 30 (9):1214-9.

22. E. Alemu^, J.W. Carl, **H. Corrada Bravo***, S. Hannenhalli* (2014). Determinants of expression variability. *Nucleic Acids Research* 42 (6), 3503-14.

23. M. Pop^, A.W. Walker^, J.N. Paulson^#, B. Lindsey^, M. Antonio^, M.A. Hossain^, J. Oundo^, B. Tamboura, V. Mai, I. Astrovskaya, **H. Corrada Bravo**, R. Rance, M. Stares, M.M. Levine, S. Panchalingam, K. Kotloff,

U.N. Ikumapayi, C. Ebruke, D. Adeyemi, F. Ahmed, M.T. Alam, R. Amin, S. Siddiqui, J.B. Ochieng, E. Ouma, J. Juma, E. Mailu, R. Omore, J.G. Morris, R.F. Breiman, D. Saha, J. Parkhill, O.C. Stine*, J.P. Nataro* (2014). Diarrhea in young children from low-income countries leads to large-scale alterations in intestinal microbiota composition. *Genome Biology* 15, R76.

24. N. Akula^, J. Barb, X. Jiang, J. Wendland, K. Choi, S. Sen, L. Hou, D. Chen, G. Laje, K. Johnson, B. Lipska, J. Kleinman, **H. Corrada Bravo**, S. Detera-Wadleigh, P.J. Munson, F.J. McMahon* (2014). RNA-sequencing of brain transcriptome implicates dysregulation of neuroplasticity, circadian rhythms and GTPase binding in bipolar disorder. *Molecular Psychiatry* , doi:10.1038/mp.2013.170.

25. W. Timp^, **H. Corrada Bravo**^, O.G. McDonald, M. Goggins, C. Umbricht, M. Zeiger, A.P. Feinberg*, R.A. Irizarry* (2014). Large hypomethylated blocks related to large heterochromatin regions as a universal defining epigenetic alteration in human solid tumors. *Genome Medicine* 6:61, doi:10.1186/s13073-014-0061-y.

26. H.S. Parker^, **H. Corrada Bravo**, J.T. Leek* (2014). Removing batch effects for prediction problems with frozen surrogate variable analysis. *PeerJ* 2:e561, doi:10.7717/peerj.561.

27. M. Aryee^, A. Jaffe, **H. Corrada Bravo**, C. Ladd-Acosta, A. Feinberg, K. Hansen*, R.A. Irizarry* (2014). Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA Methylation microarrays. *Bioinformatics* 30 (10):1363-9.

28. F. Chelaru^#, L. Smith#, N. Goldstein#, **H. Corrada Bravo*** (2014). Epiviz: interactive visual analytics for epigenomics data. *Nature Methods* 11 938-940, doi:10.1038/nmeth.3038.

29. J.N. Paulson#, O.C. Stein, **H. Corrada Bravo***, M. Pop* (2013). Differential abundance analysis for microbial marker-gene surveys. *Nature Methods* 10 (12):1200-1202, doi:10.1038/nmeth.2658.

30. M.L. Nickerson^, K.M. Im, K.J. Misner, W. Tan, H. Lou, B. Gold, D.W. Wells, **H. Corrada Bravo**, K.M. Fredrikson, T.T. Harkins, P. Milos, B. Zbar, W.M. Linehan, M. Yeager, T. Andresson, M. Dean*, G.S. Bova* (2013). Somatic alterations contributing to metastasis of a castration-resistant prostate cancer. *Human Mutation* 34 (9): 1231-41, doi:10.1002/humu.22346.

31. X. He^, R. Chatterjee^, S. John, **H. Corrada Bravo**, B.K. Sathyanarayana, S.C. Biddle, P.C. Fitzgerald, J.A. Stamatoyannopoulos, G.L. Hager, C. Vinson* (2013). Contribution of nucleosome binding preferences and co-occurring DNA sequences to transcription factor binding. *BMC Genomics* 14 (428).

32. S. Boca^, **H. Corrada Bravo**, B. Caffo, J.T. Leek*, G. Parmigiani (2013). A decision-theory approach to interpretable set analysis for high-dimensional data. *Biometrics* 69 (3):614-23, doi:10.1111/biom.12060.

33. W. Shi^, G. Wahba, R.A. Irizarry, **H. Corrada Bravo**, S.J. Wright* (2012). The Partitioned LASSO-Patternsearch Algorithm with Application to Gene Expression Data. *BMC Bioinformatics* 13:98, doi:10.1186/1471-2105-13-98.

34. **H. Corrada Bravo**^, V. Pihur, M. McCall, R.A. Irizarry, J.T. Leek* (2012). Gene expression anti-profiles as a basis for cancer diagnostics. *BMC Bioinformatics* 13:272, doi:10.1186/1471-2105-13-272.

35. L. Shan^, H.C. Yang, S.A. Rabi, **H. Corrada Bravo**, J.D. Siliciano, R.A. Irizarry, H. Zhang, J. Margolick, R.F. Siliciano* (2011). Influence of host gene transcription level and orientation on HIV-1 latency in a primary cell model. *Journal of Virology* 85 (11):5384-93.

36. T.S. Niranjan^, A. Adamczyk^, **H. Corrada Bravo**^, M. Taub, S.J. Wheelan, R.A. Irizarry, T. Wang* (2011). Effective detection of rare variants in pooled DNA samples using Srfim and cross-pool tail-curve analysis. *Genome Biology* 12 (9):R93.

37. K. Hansen^, W. Timp^, **H. Corrada Bravo**^, S. Sabunciyan^, B. Langmead^, O.G. McDonald, B. Wen, H. Wu, D. Diep, E. Briem, K. Zhang, R.A. Irizarry*, A.P. Feinberg* (2011). Increased methylation variation in epigenetic domains across cancer types. *Nature Genetics* 43 (8):768-75.

38. A. Rivas^, K. Bohane, **H. Corrada Bravo**^, M. Tan, R. Tamargo, H.W. Francis (2011). A model for early prediction of facial nerve recovery after vestibular schwannoma surgery. *Otology & Neurotology* 32 (5):826-33.

39. M. Taub^, **H. Corrada Bravo**, R.A. Irizarry* (2010). Overcoming bias and systematic errors in next generation sequencing data. *Genome Medicine* 2 (12):87.

40. J.T. Leek^, R. Scharpf, **H. Corrada Bravo**, D. Simcha, B. Langmead, W.E. Johnson, D. Geman, K. Baggerly, R.A. Irizarry* (2010). Tackling the widespread and critical impact of batch effects in high-throughput data. *Nature Reviews Genetics* 11 (10):733-739.

41. H. Wu, R.A. Irizarry*, **H. Corrada Bravo**^(2010). Intensity normalization improves color calling in SOLiD sequencing. *Nature Methods* 7:336-337.

42. M. Acevedo^, T.M. Aide*, L. J. Villanueva-Rivera, **H. Corrada Bravo**^, C. J. Corrada-Bravo* (2009). Automated classification of bird and amphibian calls using machine learning: a comparison of methods. *Ecological Informatics* 4 (4):206-214.

43. **H. Corrada Bravo**^, R.A. Irizarry* (2009). Model-based quality assessment and base-calling for second-generation sequencing data. *Biometrics* 66(3):665-74, doi:10.1111/j.1541-0420.2009.01353.x.

44. **H. Corrada Bravo**^, K.E. Lee, B.E.K. Klein, R. Klein, S.K. Iyengar, G. Wahba* (2009). Examining the relative influence of familial, genetic and environmental covariate information in flexible risk models. *Proceedings of the National Academy of Science* 106 (20): 8128-8133.
45. K. H. Eng^, **H. Corrada Bravo**^, S. Keles* (2009). A phylogenetic mixture model for the evolution of gene expression. *Molecular Biology and Evolution* 26 (10):2363-2372.
46. A. T. Evan, R. Bennartz, V. Bennington, **H. Corrada Bravo**, A. K. Heidinger, N. M. Mahowald, C. S. Velde, G. Myhre, J. P. Kossin (2008). Ocean temperature forcing by aerosols across the Atlantic tropical cyclone development region. *Geochem, Geophys. Geosyst.* 9:Q05V04, doi:10.1029/2007GC001774.
47. C. Kuang, P. McMurry, A. McCormick, F. Eisele, S.H. Lee, L.H. Young, D.R. Benson, et al. (2007). A system for operational aerosol optical depth data assimilation over global oceans. *J. Geophys. Res* 113(D10):D10208.

II.D.   Published Conference Proceedings

II.D.1.   Refereed Conference Proceedings

1. M.K. Gunady^#, S. Cornwell#, S.M. Mount, **H. Corrada Bravo*** (2017). Yanagi: transcript segment library construction for RNA-seq quantification. *17th Workshop for Algorithms in Bioinformatics (WABI) 2017.*
2. F. Dorri^#, L. Mendelowitz, **H. Corrada Bravo*** (2015). methylFlow: cell-specific methylation pattern reconstruction from high-throughput bisulfite-converted DNA sequencing. *HitSEQ 2015; Bioinformatics 32(11):1618-24.*
3. F. Chelaru^*#, **H. Corrada Bravo*** (2015). Epiviz: a view inside the design of an integrated visual analysis software for genomics. *BioVis 2015; BMC Bioinformatics 16(Supl 11):S4.*
4. **H. Corrada Bravo**^, K. Eng, S. Keles, G. Wahba, S. Wright (2009). Estimating tree-structured covariance matrices via mixed integer programming. *Twelfth International Conference on Artificial Intelligence and Statistics (AISTATS '09); Journal of Machine Learning Research Workshop and Conference Proceedings, 5*33:40.
5. **H. Corrada Bravo**^, R. Ramakrishnan (2007). Optimizing MPF queries: decision support and probabilistic inference. *26th ACM SIGMOD Intl. Conf. on Management of Data*701:712.
6. **H. Corrada Bravo**^, D. Page, R. Ramakrishnan, J. Shavlik, V. Santos Costa (2005). A framework for set-oriented computation in inductive logic programming and its application in generalizing inverse entailment. *15th ILP Conf.*69:86.

II.E.   Conferences, Workshops and Talks

II.E.2.   Invited Talks

1. What's next? Beyond integration of visual and statistical analysis of genomic data. *Biological Data Science Meeting, Cold Spring Harbor Laboratory,* Long Island, NY. November 2018.
2. Is this Data Science? Adventures in metagenomic data analysis. *Computation and Informatics in Biology and Medicine Program Annual Retreat, University of Wisconsin,* Madison, WI. October 2017.
3. Interactive visual and computational analysis of epigenetic data. *VIZBI 2017,* Sindey, Australia. June 2017.
4. Interactive statistical and computational analysis of functional and metagenomic data. *University of Puerto Rico,* Rio Piedras, PR. April 2017.
5. Challenges and opportunities in metagenomic data analysis. *NE1602 Kickoff Meeting,* Washington, DC. March 2017.
6. Bias in Biomedical Data Science. *University of Maryland Diversity in Computing Summit,* College Park, MD. November 2016.
7. Statistics and visualization for metagenomics data. *Georgetown University ICBI Seminar Series,* Washington, DC. October 2016.
8. Statistics and visualization for metagenomics data. *Johns Hopkins Young Investigator Symposium,* Baltimore, MD. October 2016.
9. Visualization, Statistical Modeling and Discovery in Computational Epigenomics. *Howard University Precision Medicine Workshop,* Washington, DC. March 2016.
10. Visualization, statistical modeling and discovery in computational epigenomics. *Harvard University Department of Biostatistics and Computational Biology,* Boston, MA. February 2016.

11. What is Biomedical Data Science?. *Symposium of Health Informatics in Latin America and the Caribbean,* San Juan, PR. November 2015.
12. Visualization, Statistical Modeling and Discovery in Computational Epigenomics. *Computer Science Colloquium Series,* College Park, MD. September 2015.
13. Interactive and exploratory visualization of epigenome-wide data. *BioIT World Conference,* Boston, MA. April 2015.
14. Interactive and exploratory visualization of epigenome-wide data. *UMD Campus Visualization Partnership Lecture Series,* College Park, MD. February 2015.
15. Interactive and exploratory visualization of epigenome-wide data. *Epigenomics in Disease, Molecular Medicine Tri-Con,* San Francisco, CA. February 2015.
16. Exploring tumor epigenetic heterogeneity by cell-specific methylation pattern reconstruction. *Department of Biostatistics and Computational Biology, Johns Hopkins Cancer Center,* Baltimore, MD. November 2014.
17. Exploring tumor epigenetic heterogeneity by cell-specific methylation pattern reconstruction. *CMU-Pitt Ph.D. Program in Computational Biology Seminar Series,* Pittsburgh, PA. April 2014.
18. Cell-specific methylation pattern reconstruction using minimum cost network flow algorithms. *Department of Mathematics, George Mason University,* Fairfax, VA. February 2014.
19. Gene expression anti-profiles as a basis for accurate universal cancer signatures. *Greenbaum Cancer Center, University of Maryland School of Medicine,* Baltimore, MD. September 2013.
20. Gene expression anti-profiles as a basis for accurate universal cancer signatures. *Institute for Genome Sciences, University of Maryland School of Medicine,* Baltimore, MD. March 2013.
21. Gene expression anti-profiles as a basis for accurate universal cancer signatures. *Department of Bioinformatics and Computational Biology, Genentech, Inc.,* South San Franciso, CA. February 2013.
22. Gene expression anti-profiles as a basis for accurate universal cancer signatures. *Innovation Center for Biomedical Informatics, Georgetown University,* Washington, DB. December 2012.
23. Statistical and computational methods for the analysis of pooled, targeted, second-generation re-sequencing data. *Biostatistics Department, University of Alabama,* Birmingham, AL. April, 2012.
24. Modeling gene expression variability for prediction in disease populations. *Department of Biostatistics, Columbia University School of Public Health,* New York, NY. December 2010.
25. Modeling gene expression variability for prediction in disease populations. *Johns Hopkins University School of Medicine,* Baltimore, MD. December 2010.
26. Gene expression variability in disease populations. *National Cancer Institute,* Bethesda, MD. October 2010.
27. Modeling uncertainty in second-generation sequencing data. *Dept. of Biostatistics, Harvard School of Public Health,* Boston, MA. November 2009.
28. Modeling and managing uncertainty in second-generation sequencing data. *Dept. of Computer Science, University of Maryland,* College Park, MD. October 2009.
29. Model-based quality assessment and base-calling for second-generation sequencing data. *University of Wisconsin-Milwaukee,* Milwaukee, WI. October 2009.
30. Model-based quality assessment and base-calling for second-generation sequencing data. *University of Manchester,* Manchester, England. October 2009.
31. Kernel methods for examining the relative influence of familial, genetic and environmental covariate information in risk models: results and (more importantly) extensions. *University of Wisconsin,* Madison, WI. May 2009.
32. Model-based quality assessment and base-calling for second-generation sequencing data. *University of Wisconsin,* Madison, WI. May 2009.
33. Model-based quality assessment and base-calling for second-generation sequencing data. *Case Western Reserve University,* Cleveland, OH. April 2009.
34. Data analysis at the computational/statistical sciences borderland: two examples from genomics. *Johns Hopkins Bloomberg School of Public Health,* Baltimore, MD. February 2009.
35. Estimating tree-structured covariance matrices via mixed-integer programming. *Johns Hopkins School of Public Health,* Baltimore, MD. January 2008.

II.E.3.   Refereed Presentations

1. Interactive Statistical and Visual Analysis of Longitudinal Metagenomic Data. *Joint Statistical Meetings,* Vancouver, Canada. August 2018.
2. Visualization and computation over hierarchically organized features for metagenomics and epigenomics.

Joint Statistical Meetings, Chicago, IL. August 2016.

3. Differential abundance analysis of metagenomic whole-genome sequencing. *Joint Statistical Meetings,* Seattle, WA. August 2015.
4. Addressing reproducibility in genomic signatures by characterizing variance and estimation stability. *Joint Statistical Meetings,* Seattle, WA. August 2015.
5. methylFlow: cell-specific methylation pattern reconstruction from high-throughput bisulfite-converted DNA sequencing. *HiTSeq,* Dublin, Ireland. July 2015.
6. Interactive and exploratory visual analytics of epigenome-wide data. *ISMB,* Dublin, Ireland. July 2015.
7. Epiviz: a view inside the design of an integrated visual analysis software for genomics. *BioVis,* Dublin, Ireland. July 2015.
8. Interactive and exploratory visualization of epigenome-wide data. *Joint Statistical Meetings,* Boston, MA. July 2014.
9. Interactive, Exploratory Visualization and Statistical Analysis of Genome-Scale Data. *International Biomteric Society ENAR Meeting,* Baltimore, MD. March 2014.
10. Gene expression anti-profiles as a basis for accurate universal cancer signatures. *ISMB '13,* Berlin, Germany. July 2013.
11. Srfim2: using basecalling model parameter estimates to understand sequencing bias. *2012 Joint Statistical Meetings,* San Diego, CA. August 2012.
12. Increased methylation variation in epigenetic domains across cancer types. *16th Annual International Conference on Research in Computational Molecular Biology (RECOMB),* Barcelona, Spain. April 2012.
13. Statistical and computational methods for the analysis of pooled, targeted, second-generation re-sequencing data. *2011 Joint Statistical Meetings,* Miami Beach, FL.. August 2011.
14. Model-based quality assessment and base-calling for second-generation sequencing data. *WNAR/IMS annual meeting,* Portland, OR. June 2009.
15. Tuning regularized kernel estimation parameters for prediction. *SIAM Conference on Optimization,* Boston, MA. May 2008.
16. Optimizing MPF queries: decision support and probabilistic inference. *26th ACM SIGMOD Intl. Conf. on Management of Data,* Beijing, China. June 2007.
17. A framework for set-oriented computation in inductive logic programming and its application in generalizing inverse entailment.. *15th ILP Conf.,* Bonn, Germany. August 2005.

II.E.7. Non-Refereed Presentations

1. Does Healthcare Insurance Matter? Establishing Insurance Status as a Risk Factor for Trauma Mortality. *INFORMS 2014,* San Francisco, CA. November 2014.
2. Epiviz(r): turning a genome browser into a display device. *Bioconductor conference,* Boston, MA. July 2014.
3. Statistical and computational methods for the analysis of pooled, targeted, second-generation re-sequencing data. *8th International Chinese Statistical Association (ICSA) International Conference,* Guangzhou, China. December 2010.
4. Model-based quality assessment and base-calling for second-generation sequencing data. *Conference on Next-Generation Sequencing,* Barcelona, Spain. October 2009.

II.E.11. Symposia

1. Metaviz: interactive statistical and visual analysis of Human Microbiome Project data. *Mid-Atlantic Metagenomics Meetup,* College Park, MD. January 2018.
2. Visualization and statistical modeling for metagenomics data. *Human Microbiome Project 2 Symposium,* Bethesda, MD. June, 2016.
3. Gene expression network anti-profiles. *UMD/NIST Network Science Mini-Symposium,* College Park, MD. January, 2014.
4. Increased methylation variation in epigenetic domains across cancer types. *Omics Day, University of Maryland,* Shady Grove, MD. May 2012.
5. Model-based quality assessment and base-calling for second-generation sequencing data. *Third Annual Young Investigators Symposium on Genomics and Bioinformatics,* Baltimore, MD. September 2009.

II.E.12. Workshops

1. Segment-based transcriptome analysis with Yanagi. *8th Workshop on Computational Advances for Next Generation Sequencing, IEEE International Conference on Computational Advances in Bio and Medical Sciences,* Las Vegas, NV. October 2018.
2. Interactive Collaborative Visualization for Genomics Design Session. *Bioconductor Conference,* Palo Alto, CA. July 2016.
3. How to publish in Biomedical Data Science. *Symposium of Health Informatics in Latin America and the Caribbean,* San Juan, PR. November 2015.
4. Statistical and visualization methods for metagenomics data analysis. *NIMBioS Workshop on Computational Methods for Metagenomics,* Knoxville, TN. July 2015.
5. Interactive visualization for genomic data. *NSF DALI workshop on dynamic languages for scalable data analytics,* Indianapolis, IN. November 2013.
6. Interactive visualization in Bioconductor using Epivizr. *BioC 2013,* Seattle, WA. July 2013.

II.F. Professional Publications

II.F.2. Pre-print/Working Paper (Not Work in Progress)

1. M.K. Gunady^#, S. Mount, **H. Corrada Bravo*** (2018). Fast and interpretable alternative splicing and differential gene-level expression analysis using transcriptome segmentation with Yanagi. *biorxiv* https://doi.org/10.1101/364281.
2. N.D. Olson^#, M. Senthil Kumar#, S. Li, S. Hao, W. Timp, M.L. Salit, O.C. Stine, **H. Corrada Bravo*** (2018). Assessing 16S marker gene survey data analysis methods using mixtures of human stool sample DNA extracts. *biorxiv* https://doi.org/10.1101/400226.
3. M.S. Kumar^#, E.V. Slud, K. Okrah#, S.C. Hicks, S. Hannenhalli, **H. Corrada Bravo*** (2017). Analysis and correction of compositional bias in sparse sequencing count data. *biorxiv* https://doi.org/10.1101/142851.
4. J. Wanger^#, F. Chelaru^#, J. Kancherla^, J. Paulson^#, V. Felix, A. Mahurkar, **H. Corrada Bravo*** (2017). Metaviz: interactive statistical and visual analysis of metagenomic data. *biorxiv* https://doi.org/10.1101/105205.
5. J. Paulson^#, H. Talukder#, **H. Corrada Bravo*** (2017). Longitudinal differential abundance analysis of microbial marker-gene surveys using smoothing splines. *biorxiv* https://doi.org/10.1101/099457.
6. S. Hicks^, K. Okrah#, J. Paulson#, J. Quackenbush, R. Irizarry, **H. Corrada Bravo*** (2016). Smooth quantile normalization. *biorxiv* https://doi.org/10.1101/085175.
7. M. Sharmin^#, **H. Corrada Bravo***, S. Hannenhalli* (2015). Distinct genomic and epigenomic features demarcate hypomethylated blocks in colon cancer. *biorxiv* http://dx.doi.org/10.1101/028803.
8. M. Sharmin^#, **H. Corrada Bravo**, S. Hannenhalli* (2015). Heterogeneity of Transcription Factor binding specificity models within and across cell lines. *biorxiv* http://dx.doi.org/10.1101/028787.
9. J. Wagner^#, J.N. Paulson#, X.S. Wang, B. Bhattacharjee, **H. Corrada Bravo*** (2015). Privacy-Preserving Microbiome Analysis Using Secure Computation. *biorxiv* http://dx.doi.org/10.1101/025999.
10. W. Dinalankara^#, **H. Corrada Bravo** (2013). Anomaly classification with the anti-profile support vector machine. *arXiv preprint server* arXiv:1301.3514 [stat.ML].
11. H.S. Parker^, **H. Corrada Bravo**, J.T. Leek (2013). Removing batch effects for prediction problems with frozen surrogate variable analysis. *arXiv preprint server* arXiv:1031.3947 [stat.ME].
12. S. Boca^, **H. Corrada Bravo**, B. Caffo, J.T. Leek, G. Parmigiani (2010). A decision-theory approach to interpretable set analysis for high-dimensional data. *Johns Hopkins University, Dept. of Biostatistics Working Papers* Working paper 211. http://biostats.bepress.com/jhubiostat/paper211.
13. R.A. Irizarry, **H. Corrada Bravo**^(2009). Model-based quality assessment and base-calling for second-generation sequencing data. *Johns Hopkins University, Dept. of Biostatistics Working Papers* Working paper 184. http://biostats.bepress.com/jhubiostat/paper184.

II.H. Completed Creative Works

II.H.8. Software and Applications

1. Metavizr: Interactive visualization of metagenomics data in R/Bioconductor.
   Released December 2016. http://github.com/epiviz/metavizr.

2. Metaviz: Interactive visualization for metagenomics data.
   Released December 2016. http://metaviz.org.

3. qsmooth: smooth quantile normalization.
   Released October 2016. https://github.com/stephaniehicks/qsmooth.

4. methylFlow: cell-specific methylation pattern reconstruction.
   Released June 2015. http://github.com/hcorrada/methylFlow.

5. Epivizr: interactive visualization of genomics data in R/Bioconductor.
   Released June 2013. http://bioconductor.org/packages/release/bioc/html/epivizr.html.

6. Epiviz: Interactive visualization for genomics data.
   Released June 2013. http://github.com/epiviz.

7. antiProfiles: Gene expression anti-profiles as a basis for accurate universal cancer signatures.
   Released April 2013. http://bioconductor.org/packages/release/bioc/html/antiProfiles.html.

8. Healthvis: Interactive visualization in health.
   Released April 2013. http://healthvis.org.

9. metagenomeSeq: Statistical analysis for sparse high-throughput sequencing.
   Released February 2013. http://bioconductor.org/packages/release/bioc/html/metagenomeSeq.html.

10. Minfi: Analyze Illumina's 450k methylation arrays.
    Released Febraury 2013. http://bioconductor.org/packages/release/bioc/html/minfi.html.

11. bumphunter: Tools for finding bumps in genomic data.
    Released February 2013. http://bioconductor.org/packages/release/bioc/html/bumphunter.html.

12. Servic4e: Effective detection of rare variants in pooled DNA samples.
    Released November 2011. http://www.cbcb.umd.edu/~hcorrada/secgen.

13. Srfim: Model-based base-calling and quality assessment for second-generation sequencing.
    Released April 2010. http://www.cbcb.umd.edu/~hcorrada/secgen.

14. Rsolid: Intensity normalization for SOLiD sequencing.
    Released April 2010. http://www.cbcb.umd.edu/~hcorrada/secgen.

15. Rcsdp: An interface to the CSDP semidefinite programming library for R.
    Released Dec. 2008. http://cran.r-project.org/web/packages/Rcsdp/index.html.

16. Rcplex: An interface to the CPLEX optimization engine for R.
    Released Jan. 2008. http://cran.r-project.org/web/packages/Rcplex/index.html.

II.H.9. Websites

1. Metaviz: Interactive visualization for metagenomic data.
   Released September 2017. http://metaviz.org.

2. Epiviz: Interactive visualization for genomics data.
   Released May 2014. http://epiviz.org.

II.J.   Sponsored Research

II.J.1.   Grants

1. Title: R01: Overcoming bias and unwanted variability in next generation sequencing
   Funding Agency: NIH HG005220
   Recipient Institution: Dana Farber Cancer Institute, Rafael A. Irizarry (PI)
   Amount Awarded: $397,903 subcontracted to UMCP
   Dates: March 2015-February 2020
   Role: PI

2. Title: R01: Integrative visual and computational exploratory analysis of genomics data
   Funding Agency: NIH GM114267
   Recipient Institution: University of Maryland, College Park
   Amount Awarded: $1,800,421
   Dates: September 2015-August 2019
   Role: PI

3. Title: NRT-DESE; Network biology: from data to information to insights
   Funding Agency: NSF 1632976
   Recipient Institution: University of Maryland, College Park, Michelle Girvan (PI)
   Amount Awarded: $2,999,847
   Dates: September 2016-August 2021
   Role: co-PI

4. Title: CGDNet: Cancer Gene Drug Network: Using patient-specific drug-gene networks for recommending targeted cancer therapies
   Funding Agency: NIH R21CA220398
   Recipient Institution: Georgetown University, Simina Boca (PI)
   Amount Awarded: $149,051 subcontracted to UMCP
   Dates: September 2017-August 2019
   Role: PI

5. Title: Illuminating neurodevelopment through integrate analysis and visualization of multi-omic data
   Funding Agency: NIH R24MH114815
   Recipient Institution: University of Maryland, Baltimore, Ronna Hertzano (PI)
   Amount Awarded: $415,443 subcontracted to UMCP
   Dates: September 2018-May 2021
   Role: PI

6. Title: Epigenomic regulation of hippocampal transcriptome after experimental traumatic brain injury
   Funding Agency: UMD Seed Grant NA
   Recipient Institution: University of Maryland, College Park
   Amount Awarded: $50,000
   Dates: August 2017-July 2019
   Role: co-PI

7. Title: U54: Longitudinal multiomics microbial profiling in healthy and disease individuals
   Funding Agency: NIH DK102556
   Recipient Institution: Stanford University, Michael P. Snyder (PI)

Amount Awarded: $100,702 subcontracted to UMCP
Dates: September 2015-December 2016
Role: PI

8. Title: R01: Analysis tools and software for second generation sequencing
   Funding Agency: NIH HG005220
   Recipient Institution: Johns Hopkins University, Rafael A. Irizarry (PI)
   Amount Awarded: $380,400 subcontracted to UMCP ($1,230,000 total awarded by NIH)
   Dates: August 2010-May 2013
   Role: PI

9. Title: R01: Alignment software for second generation sequencing
   Funding Agency: NIH HG006102
   Recipient Institution: Johns Hopkins University, Steven Salzberg (PI)
   Amount Awarded: $432,259.00 subcontracted to UMCP ($3,585,436 total awarded by NIH)
   Dates: May 2011-April 2014
   Role: PI

II.K. Fellowships, Gifts and Other Funded Research

II.K.1. Fellowships

1. Ford Fellowship, National Academies of Science.
2. Advanced Opportunity Fellowship, University of Wisconsin-Madison.

II.L. Submissions and Works in Progress

II.L.1. Current Grant Applications

1. Title: FACT: Application of big data in food safety-integrating whole genome sequencing information into a risk assessment framework
   Funding Agency: USDA-National Institute of Food and Agriculture
   Recipient Institution: University of Maryland, College Park, Abani K. Pradhan (PI)
   Amount Requested: $499,998
   Dates:
   Role: co-PI

2. Title: HDR DSC: Project experiences to ready the Data Scientist
   Funding Agency: National Science Foundation
   Recipient Institution: University of Maryland, College Park
   Amount Requested: $1,047,608
   Dates:
   Role: PI

II.L.2. Manuscripts in Preparation

II.L.3. Manuscripts under Review

1. N.D. Olson^#, M. Senthil Kumar#, S. Li, S. Hao, W. Timp, M.L. Salit, O.C. Stine, **H. Corrada Bravo***. Assessing 16S marker gene survey data analysis methods using mixtures of human stool sample DNA extracts. *Microbiome, revision submitted.*

2. M.K. Gunady^#, S. Mount, **H. Corrada Bravo***. Fast and interpretable alternative splicing and differential gene-level expression analysis using transcriptome segmentation with Yanagi. *BMC Bioinformatics, revision submitted.*

3. Z. Cui^#, J. Kancherla#, N. Elmqvist, **H. Corrada Bravo***. Proactive Visual and Statistical Analysis of Genomic Data in Epiviz. *Bioinformatics, submitted.*

4. A.A. Ismail^#, T. Wood#, **H. Corrada Bravo***. Improving long-horizon forecasts with expectation-biased LSTM networks. *Scientific Reports, submitted.*

5. Z. Cui^#, J. Kancherla#, **H. Corrada Bravo**, N. Elmqvist*. Sherpa: Leveraging User Attention for Computational Steering in Visual Analytics. *IEEE VIS 2019, under review.*

III.   Teaching, Advising and Mentoring

III.A.   Courses Taught

Term	Course	Enrollment	Description
Spring 2019	CMSC320	230	Introduction to Data Science
Spring 2019	CMSC644	16	Algorithms for Data Science
Spring 2018	CMSC 320	183	Introduction to Data Science
Fall 2018	CSMC798E	9	Graduate Seminar in Computer Science: Research Colloqium
Fall 2018	CMSC643	18	Machine Learning & Data Mining (Data Mining Professional Master's Certificate
Fall 2018	CMSC828O	27	Analysis of Biological Networks across Scales
Fall 2018	CMSC798	1	Graduate Seminar in Computer Science (Individual instruction)
Spring 2017	CMSC 423	52	Bioinformatic Databases, Tools and Algorithms
Fall 2017	CSMC828O	31	Analysis of Biological Networks across Scales
Fall 2017	CMSC643	11	Machine Learning & Data Mining (Data Mining Professional Master's Certificate
Fall 2017	CSMC798E	23	Graduate Seminar in Computer Science: Research Colloqium
Fall 2017	CMSC798	1	Graduate Seminar in Computer Science (Individual instruction)
Fall 2017	CMSC898	1	Pre-candidacy Research (Individual instruction)
Fall 2017	CMSC899	2	Doctoral Dissertation Research (Individual instruction)
Fall 2017	CMSC499A	1	Independent Undergraduate Research (Individual instruction)
Spring 2016	CMSC 320	50	Introduction to Data Science
Spring 2016	CMSC 898	1	Pre-candidacy Research (Individual instruction)
Spring 2016	CMSC 899	1	Doctoral Dissertation Research (Individual instruction)
Fall 2016	CMSC 320	83	Introduction to Data Science
Fall 2016	CMSC 899	2	Doctoral Dissertation Research (Individual instruction)
Spring 2015	CMSC 498T	32	Introduction to Data Science II
Spring 2015	CMSC 898	1	Pre-candidacy research (Individual instruction)
Spring 2015	CMSC 899	2	Doctoral dissertation research (Individual instruction)
Spring 2015	AMSC 899	2	Doctoral dissertation research (Individual instruction)
Fall 2015	CMSC 423	49	Bioinformatics Databases, Tools and Algorithms
Fall 2015	CMSC 898	3	Pre-candidacy research (Individual instruction)
Fall 2015	CMSC 499A	1	Independent Undergraduate Research (Individual instruction)
Fall 2014	CMSC 423	49	Bioinformatics Databases, Tools and Algorithms
Fall 2014	CMSC 423	35	Bioinformatics Databases, Tools and Algorithms

Term	Course	Enrollment	Description
Spring 2014	CMSC 702	30	Computational Systems Biology and Functional Genomics
Spring 2014	AMSC 899	3	Doctoral dissertation research (Individual instruction)
Spring 2014	CMSC 899	1	Doctoral dissertation research (Individual instruction)
Spring 2014	CMSC 898	3	Pre-candidacy research (Individual instruction)
Fall 2014	CMSC 899	2	Doctoral dissertation research (Individual instruction)
Fall 2014	CMSC 898	1	Pre-candidacy research (Individual instruction)
Fall 2014	AMSC 899	2	Doctoral dissertation research (Individual instruction)
Fall 2014	CMSC 798	1	Graduate seminar (Individual instruction)
Fall 2014	AMSC 760	1	Applied Statistics Practicum (Individual instruction)
Spring 2013	CMSC 702	35	Computational Systems Biology and Functional Genomics
Fall 2013	CMSC 423	48	Bioinformatics Databases, Tools and Algorithms
Fall 2013	AMSC 689	1	Research Interactions: Regularized Regression Methods
Fall 2013	AMSC 898	1	Pre-candidacy research (Individual instruction)
Fall 2013	CMSC 898	4	Pre-candidacy research (Individual instruction)
Fall 2013	AMSC 899	3	Doctoral dissertation research (Individual instruction)
Spring 2013	AMSC 899	1	Pre-candidacy research (Individual instruction)
Spring 2013	CMSC 898	2	Pre-candidacy research (Individual instruction)
Spring 2013	CMSC 798	1	Graduate seminar (Individual instruction)
Spring 2013	AMSC 898	3	Pre-candidacy research (Individual instruction)
Fall 2012	CMSC 726	48	Machine Learning
Fall 2012	CMSC 898	1	Pre-candidacy research (Individual instruction)
Fall 2012	AMSC 899	1	Doctoral dissertation research (Individual instruction)
Fall 2012	AMSC 898	3	Pre-candidacy research (Individual instruction)
Spring 2012	CMSC 858B	26	Computational Systems Biology and Functional Genomics
Spring 2012	CMSC 351	90	Introduction to Algorithms
Spring 2012	CMSC 898	1	Pre-candidacy research (Individual instruction)
Spring 2012	AMSC 898	4	Pre-candidacy research (Individual instruction)
Spring 2011	CMSC 858P	17	Computational Methods for High-Throughput Analysis of Biological Systems \
Spring 2011	AMSC 898	1	Pre-candidacy research (Individual instruction)
Fall 2011	CBMG 688P	12	Team-taught graduate 'Programming for Biologists' course
Fall 2011	AMSC 898	3	Pre-candidacy research (Individual instruction)
3rd Term 2010	140.644	16	Practical Machine Learning (Johns Hopkins University School of Public Health, Department of Biostatistics) \
Fall 2010	CBMG 688P	12	Team-taught graduate 'Programming for Biologists' course

III.B. Teaching Innovations

III.B.5. Course or Curriculum Development

CSMC828O - Analysis of Biological networks across scales

Developed new course in support of our NSF NRT training grant. This course introduces tools and methods for analysis of network data focusing on the study of molecular, neurological and ecological networks. Students are introduced to methods and tools from network science, statistics, Machine Learning and visualization.

CMSC643 - Machine Learning and Data Mining

Developed a new course as part of the Data Science Professional Master's Certificate program in the Computer Science Department. This course introduced students to tools and methods for data mining and Machine Learning.

It covered mathematical and conceptual aspects of these tools, while introducing practical use of these tools through newly-design projects.

CMSC 320 - Introduction to Data Science

Developed a new undergraduate course covering the practice of data science, and a required course on a new Data Science specialization within the undergraduate program of the Department of Computer Science. This course focuses on data management, exploratory and statistical data analysis, data and information visualization. It is heavily assignment-based and draws extensively from applications.

CMSC 498 - Introduction to Data Science II, Exploring, Modeling and Communicating with Data

This is the second of two new courses covering the practice of data science. This course focuses on exploratory and statistical data analysis, data and information visualization, and the presentation and communication of analysis results. It is heavily assignment-based and draws extensively from applications.

CMSC 702 - Computational Systems Biology and Functional Genomics

Developed this new advanced graduate-level course concentrating on the use of Statistical Learning methods and algorithms in the analysis of large high-throughput biomedical assays. Core course in the Computational Biology concentration area.

CMSC 858P - Computational Methods for High-Throughput Analysis of Biological Systems

Developed this new advanced graduate-level course concentrating on the use of Statistical Learning methods and algorithms in the analysis of large high-throughput biomedical assays. First offering attracted 21 students, including students from ECE, AMSC and BISI. This course was merged with other content to become CMSC 702 (Computational Systems Biology) to be offered as a regular course in the CS department starting in Spring 2012.

140.644 - Practical Machine Learning

Developed this new advanced graduate-level course as postdoc at the Johns Hopkins University School of Public Health. The course provided students from a variety of backgrounds with practical knowledge of Machine Learning models and applications.

III.C. Advising

III.C.1. Undergraduate

Name	Department	Stage	Role
Yifan Yang	CS	Current	advisor
Brian Gottfried	CS	Completed Institute of Genome Science, University of Maryland, Baltimore	advisor
Alexander Zhang	CS	Completed Facebook	advisor
Tim Wood	CS	Completed Broad Institute	advisor
Daniel Konecki	CS	Completed Baylor College of Medicine	advisor
Mihai Sirbu	CS	Completed Capital One	advisor
Lan Tran	CS	Current	advisor
Morgan Walter	Bioengineering	Current	advisor

III.C.2. Master's

Name	Department	Stage	Role
Graham Antoszewski	Math	Current	advisor
Nicholas Thieme	CS	Completed	advisor
Charles Dawkins	CBMG	Completed	committee member

III.C.3.   Doctoral

Name	Department	Stage	Role
Aya Ismail	CS	Pre-candidacy	advisor
Mohammed Gunady	CS	Candidacy	advisor
Domenick Braccia	CBMG	Pre-candidacy	advisor
Zhe Cui	EE	Candidacy	co-advisor
Faezeh Dorri	CS	Candidacy	advisor
Senthil Muthiah	CBMG	Candidacy	advisor
Nathanael Olson	CBMG	Completed Research Scientist, NIST	advisor
Justin Wagner	CS	Completed Resarch Scientist, NIST	advisor
Mahfuza Sharmin	CS	Completed Postdoc, Stanford University	co-advisor
Florin Chelaru	CS	Completed Postdoc, MIT Broad Institute	advisor
Chiao-Wen Hsiao	AMSC	Completed Postdoc, University of Chicago	advisor
Kwame Okrah	AMSC	Completed Biostatistician, Genentech Inc.	advisor
Wikum Dinalankara	CS	Completed Postdoc, Johns Hopkins University Medicine	advisor
Hisham Talukder	AMSC	Completed Data Scientist, Dow Jones	advisor
Joseph Paulson	AMSC	Completed Postdoc, Harvard/Dana-Farber Cancer Institute	co-advisor
Chengxi Ye	CS	Pre-candidacy	committee member
Viet-An Nguyen	CS	Completed Data Scientist, Facebook	committee member
Justin Malin	CBMG	Completed Postdoc, National Institutes of Health	committee member
Ted Gibbons	CBMG	Completed	committee member
Steven Smith	CBMG	Completed	committee member
Cody Buntain	CS	Completed	committee member
David Kelley	CS	Completed Postdoc, Broad Institute	committee member
Guillaume Marcais	CS	Completed Research Scientist, University of Maryland	committee member
Bhargav Kanagal	CS	Completed	committee member
Yuan Li	CBMG	Completed	committee member

Name	Department	Stage	Role
Mohammadreza Ghodsi	CS	Completed Software Engineer, Facebook	
Ginger Houston-Ludlam	CBMG	Completed	committee member
Louis Licamele	CS	Completed	committee member
Benjamin Langmead	CS	Completed Assistant Professor, Johns Hopkins University	committee member
Rob Partro	CS	Completed Postdoc, Carnegie Mellon University	committee member
Bo Liu	CS	Completed Software Engineer, Google	committee member
Daehwan Kim	CS	Completed Postdoc, Johns Hopkins University	committee member
Laura Dillon	CBMG	Completed	committee member
Shane Squires	Physics	Completed	committee member
Srutii Sarda	CBMG	Completed	committee member
Kun Wang	CBMG	Candidacy	committee member
Vahid Liaghat	CS	Completed	committee member
Adam Bazinet	CS	Completed Research Scientist, University of Maryland	committee member
Sameh Khamis	CS	Completed	committee member
Snighda Chaturvedi	CS	Completed	committee member
Alex Malozemoff	CS	Completed	committee member
Chris Hill	CS	Completed Postdoc, University of Washington	committee member
Jose Carrillo	ANSC	Completed	committee member
Stephen Xi Chen	CS	Completed	committee member
Yang Shen	Chemistry	Completed	committee member
Jay Pujara	CS	Completed	committee member
Jayanta Mondal	CS	Completed	committee member
Mossaab Bagdouri	CS	Completed	committee member
Avinash Sahu	CS	Completed	committee member
Mohamed Rastegari	CS	Completed	committee member
Sana Malek	CS	Completed	committee member

Name	Department	Stage	Role
Fan Yang	CS	Completed	committee member
Bahadir Ozdemir	CS	Completed	committee member
Prachi Kulkarni	School of Public Health	Completed	committee member
Deok Gun Park	CS	Candidacy	committee member
Fan Du	CS	Completed	committee member
Zhenpeng Zhao	CS	Candidacy	committee member
Sriram Karthik Badam	CS	Candidacy	committee member
Hua He	CS	Candidacy	committee member
Yue-Hei Ng	CS	Completed	committee member
Xiyang Dai	CS	Completed	committee member
Hui Miao	CS	Completed	committee member
Ahmed Kosba	CS	Completed	committee member
Jay Ghurye	CS	Completed	committee member

III.F. Professional and Extension Education

III.F.3. Workshops

1. Cloud-based analysis of Human Microbiome Project data using Metaviz. *University of Maryland, Baltimore Institute of Genomic Science,* Baltimore, MD. June 2017.
2. Quantitative methods for humanities and social science with R. *Monterrey Technological College,* Mexico City, Mexico. June 2017.
3. Big Data in Social Sciences: Introduction to Data Science with R. *Monterrey Technological College,* Mexico City, Mexico. June 2015.
4. Computational and statistical challenges in second-generation sequencing data analysis. *Emerging Biotechnologies, Georgetown University Masters in Biotechnology program,* Washington, DC. February and March 2012.
5. Topics in modern biology: omics: new tools and new areas. *Guest lecturer, Department of Biology, University of Puerto Rico,* San Juan, PR. March 2011.
6. Computational and statistical challenges in second-generation sequencing data analysis. *Emerging Biotechnologies, Georgetown University Masters in Biotechnology program,* Washington, DC. March 2011.
7. Model-based base-calling and quality assessment for second-generation sequencing data. *Integrative Statistical Analysis of Genome Scale Data. Cold Spring Harbor Laboratory,* Cold Spring Harbor, NY. June 17, 2010.
8. Introduction to R and Bioconductor for Microarray Analysis. *PASI,* Guanajuato, México. May 4-5, 2010.
9. Clustering and Classification of Gene Expression Data. *Workshop. Johns Hopkins University Center for Computational Genomics,* Baltimore, MD. May 14, 2010.
10. Interactive Visualization of Genomic Data with epiviz(r). *Bioconductor Conference,* Palo Alto, CA. July 2016.
11. Machine Learning for Biomedical Data Analysis. *Centro Internacional de Entrenamiento e Investigaciónes Medicas,* Cali, Colombia. October 2018.

IV. Service and Outreach

IV.A. Editorships, Editorial Boards, and Reviewing Activities

IV.A.1. Editorships

1. Special Issue Editor, *Proceedings IEEE.* 2013-2016

IV.A.3. Reviewing Activities for Journals and Presses

1. *Bioinformatics*
2. *Annals of Applied Statistics*
3. *Journal of Machine Learning Research*
4. *Biostatistics*
5. *Journal of Artificial Intelligence Research*
6. *Genome Biology*
7. *Journal of the Royal Statistical Society (Series C)*
8. *IEEE Transactions on Computational Biology and Bioinformatics*
9. *BMC Genomics*
10. *Biometrics*
11. *BMC Bioinformatics*
12. *Genome Research*
13. *Nucleic Acids Research*
14. *Nature Communications*
15. *Nature Methods*
16. *New England Journal of Medicine*
17. *Nature Scientific Reports*
18. *Manning Publications Press*
19. *Chapman & Hall/CRC*

IV.A.4. Reviewing Activities for Agencies and Foundations

1. *Standing member NIH GNOM-G review panel 2017-*
2. *NSF review panel 2011, 2016 (2x)*
3. *NIH ad-hoc review panels 2014 (NIAID), 2016 (NCII, BD2K Training, NIAID), 2017 (BDMA)*
4. *Austrian Science Fund, 2015*

IV.A.5. Reviewing Activities for Conferences

1. *ISMB/ECCB 2013-2016*
2. *ACM-BCB 2014-2017*

IV.B.   Committees, Professional and Campus Service

IV.B.1.   Campus Service - Department

1. Outstanding Graduate Student Selection Committee. 2017
2. Capital One Chairs Search Committee. 2018-
3. Department Faculty Search Committee. 2017-2018
4. Department Council. 2016-2017
5. Computer Science Recruitment Committee. 2014-
6. Computational Biology Field Committee. 2010-
7. Center for Women in Computing Advisory Board. 2014-
8. UMIACS APT Committee. 2013-2015, 2016-
9. CBBG Graduate Admissions Committee. 2012-2014
10. Strategic Planning Subcommittee on Diversity and Outreach. 2014-2015

IV.B.2.   Campus Service - College

1. Computer Science Chair Search Committee. December 2011-March 2012, November 2016-March 2017

IV.B.3.   Campus Service - University

1. Steering Committee, Year of Data Science. 2017-2018

IV.B.6.   Leadership Roles in Meetings and Conferences

1. Organizing Committee, Biovis Workshop. 2015-2016
2. Session Organizer, Joint Statistical Meetings. 2016

IV.C.   External Service and Consulting

IV.C.1.   Community Engagements, Local, State, National, International

1. City of Hyattsville Council Educational Advisory Committee. September 2014-December 2016
2. Outreach: computer science education after-school club, Hyattsville Elementary School. March-April 2015

IV.C.5.   Consultancies

1. Training in Machine Learning. *Fannie Mae*, June-September 2017
2. Visualization in Bioconductor. *Genentech*, September 2013-December 2014

V.   Awards, Honors and Recognition

V.D.   Recognition in Media

- Researchers develop tool to better visualize, analyze human genomic data, *Genetics Times*, 11 Aug 2014.
- New Tool To Better Visualize, Analyze Human Genomic Data, *Counsel and Heal*, 04 Aug 2014.
- UMD Researchers Develop Tool to Better Visualize, Analyze Human Genomic Data, *Science Newsline*, 04 Aug 2014.
- UMD Researchers Develop Tool to Better Visualize, Analyze Human Genomic Data, *Technobahn*, 04 Aug 2014.

- New web-based tool to better analyse human genomic data, *Vancouver Verdesi*, 04 Aug 2014.
- New web-based tool to better analyse human genomic data, *Yahoo! News India*, 04 Aug 2014.
- New web-based tool to better analyse human genomic data, *Business Standard*, 04 Aug 2014.
- New web-based tool enables researchers to quickly visualize genomic information, *The Medical News*, 04 Aug 2014.
- UMD researchers develop tool to better visualize, analyze human genomic data, *Innovations Report*, 04 Aug 2014.
- Researchers Develop Tool to Analyze Human Genomic Data, *Medindia*, 04 Aug 2014.
- Tool to better visualize, analyze human genomic data developed, *Science Daily*, 03 Aug 2014.
- Team develops tool to better visualize, analyze human genomic data, *MedicalXpress*, 03 Aug 2014.
- UMD researchers develop tool to better visualize, analyze human genomic data, *EurekAlert!*, 03 Aug 2014.