Héctor Corrada Bravo

Center for Bioinformatics and Computational Biology
Assistant Professor, University of Maryland
Department of Computer Science, and
Institute for Advanced Computer Studies (UMIACS)

hcorrada@umiacs.umd.edu
301-405-2481
http://www.cbcb.umd.edu/~hcorrada

Education

- Ph.D., Computer Science, University of Wisconsin, Madison, 2008.
  Dissertation: *Graph-based data analysis*
  Advisors: Grace Wahba and Raghu Ramakrishnan
- D.M.A., Indiana University School of Music, August 2003 (ABD).
- M.M., Peabody Institute of Music, Johns Hopkins University, May 1999.
- B.M., Peabody Institute of Music, Johns Hopkins University, May 1997.

Professional Appointments

- Assistant Professor, Computer Science Department, UMCP.
  July 2010-Present
- Postdoctoral Fellow, Rafael A. Irizarry Group, Department of Biostatistics, Johns Hopkins University School of Public Health, Baltimore, MD.
  September 2008-June 2010

Publications

Selected Articles (out of 29) in Refereed Journals

1. **H. Corrada Bravo**, R.A. Irizarry (2009). Model-based quality assessment and base-calling for second-generation sequencing data. *Biometrics,* 66(3), 665-74. Published online before print November 13, 2009. doi:10.1111/j. 1541-0420.2009.01353.x

2. **H. Corrada Bravo**, K.E. Lee, B.E.K. Klein, R. Klein, S.K. Iyengar and 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, no. 20: 8128-8133.

3. K. H. Eng, **H. Corrada Bravo** and S. Keles (2009). A phylogenetic mixture model for the evolution of gene expression. *Molecular Biology and Evolution.* 26 (10):2363-2372.

4. H. Wu, RA. Irizarry, **H. Corrada Bravo** (2010). Intensity normalization improves color calling in SOLiD sequencing. *Nature Methods,* 7, 336-337.

5. 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.

6. 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.*

7. 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.*

8. **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. *Highly accessed.*

9. M.L. Nickerson, K.M. Im, K.J. Misner, A.L. Yates, D.W. Wells, **H. Corrada Bravo**, K. Fredrikson, W. Tan, M. Yeager, P. Milos, B. Zbar, M. Dean, G.S. Bova (2013). Exome sequencing of metastatic tumors from a patient with prostate cancer identifies novel nonsynonymous alterations and TET2 as a prostate cancer gene. *Human Mutation,* epub ahead of print. doi: 10.1002/humu.22346.

10. 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.

11. J. Paulson[§], O.C. Stein, **H. Corrada Bravo**[†], M. Pop. Robust statistical methods for differential abundance analysis of marker gene microbial survey data (2013). *Nature Methods,* epub ahead of print. *doi:10.1038/ nmeth.2658.*

12. 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. RNA-sequencing of brain transcriptome implicates dysregulation of neuroplasticity, circadian

rhythms and GTPase binding in bipolar disorder (2013). *Molecular Psychiatry,* (online before print) doi:10.1038/mp.2013.170.

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

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

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

16. Pop, M., Walker, A.W., Paulson, J.§, Lindsey, B., Antonio, M., Hossain, M.A., Oundo, J., Tamboura, B., Mai, V., Astrovskaya, I., **Corrada Bravo, H.**, Rance, R., Stares, M., Levine, M.M., Panchalingam, S., Kotloff, K., Ikumapayi, U.N., Ebruke, C., Adeyemi, D., Ahmed, F., Alam, M.T., Amin, R., Siddiqui, S., Ochieng, J.B., Ouma, E., Juma, J., Mailu, E., Omore, R., Morris, J.G., Breiman, R.F., Saha, D., Parkhill, J., Stine, O.C., Nataro, J.P., (2014). Diarrhea in young children from low-income countries leads to large-scale alterations in intestinal microbiota composition. *Genome Biology 15, R76.*

17. F. Chelaru§, L. Smith, N. Goldstein, **H. Corrada Bravo**† (2014)**.** Epiviz: interactive visual analytics for epigenomics data. *Nature Methods,* in press.

18. 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,* in press.

Book Reviews, Other Articles and Notes

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

Articles in Refereed Conferences and Workshops

1. **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. *Proc. of the 15th ILP Conf.* 69:86.

2. **H. Corrada Bravo**, R. Ramakrishnan (2007). Optimizing MPF queries: decision support and probabilistic inference. *Proc. of the 26th ACM SIGMOD Intl. Conf. on Management of Data* 701:712.

3. **H. Corrada Bravo**, K. Eng, S. Keles, G. Wahba and S. Wright (2009). Estimating tree-structured covariance matrices via mixed integer programming. *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics (AISTATS '09); Journal of Machine Learning Research Workshop and Conference Proceedings, 5,* 33:40.

Invited Talks (last 5 years)

1. Gene expression variability in disease populations. *National Cancer Institute,* Bethesda, MD. October 2010.

2. Modeling gene expression variability for prediction in disease populations. *Computational Genomics Seminar, Johns Hopkins University School of Medicine.* December 2010.

3. Modeling gene expression variability for prediction in disease populations. *Invited talk in departmental seminar, Department of Biostatistics, Columbia University School of Public Health.* December 2010.

4. 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.

5. Statistical and computational methods for the analysis of pooled, targeted, second-generation re-sequencing data. *2011 Joint Statistical Meetings, Miami Beach, FL.*, August 2011.

6. Statistical and computational methods for the analysis of pooled, targeted, second-generation re-sequencing data. *Biostatistics Department Seminar Series, University of Alabama-Birmingham,* April, 2012.

7. Increased methylation variation in epigenetic domains across cancer types. *16th Annual International Conference on Research in Computational Molecular Biology (RECOMB),* April 2012.

8. Increased methylation variation in epigenetic domains across cancer types. *Omics Day, University of Maryland,* May 2012.

9. Srfim2: using basecalling model parameter estimates to understand sequencing bias. *2012 Joint Statistical Meetings, San Diego, CA*, August 2012.

10. Gene expression anti-profiles as a basis for accurate universal cancer signatures. *Innovation Center for Biomedical Informatics, Georgetown University*, December 2012.

11. Gene expression anti-profiles as a basis for accurate universal cancer signatures. *Department of Bioinformatics and Computational Biology, Genentech, Inc.*, February 2013.

12. Gene expression anti-profiles as a basis for accurate universal cancer signatures. *Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore*, March 2013.

13. Gene expression anti-profiles as a basis for accurate universal cancer signatures. *ISMB '13*, July 2013.

14. Interactive and exploratory visualization of epigenome-wide data. BioIT World Conference, Boston, MA. April 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.

Research Software:

1. Epiviz: Interactive visualization for epigenomics. Released April 2013. http://epiviz.cbcb.umd.edu/help/

2. Healthvis: Interactive visualization in health. Released April 2013. http://healthvis.org
3. 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
4. metagenomeSeq: Statistical analysis for sparse high-throughput sequencing. http://bioconductor.org/packages/release/bioc/html/metagenomeSeq.html
5. Servic4e: Effective detection of rare variants in pooled DNA samples. Released November 2011. http://www.cbcb.umd.edu/~hcorrada/secgen
6. Srfim: Model-based base-calling and quality assessment for second-generation sequencing. Released April 2010. http://www.cbcb.umd.edu/~hcorrada/secgen
7. Rsolid: Intensity normalization for SOLiD sequencing. Released April 2010. http://www.cbcb.umd.edu/~hcorrada/secgen
8. Rcplex: An interface to the CPLEX optimization engine for R. Released Jan. 2008. http://cran.r-project.org/web/packages/Rcplex/index.html
9. Rcsdp: An interface to the CSDP semidefinite programming library for R. Released Dec. 2008. http://cran.r-project.org/web/packages/Rcsdp/index.html

Fellowships, Prizes and Awards

- Ford Fellowship, National Academies of Science. 2003-2007

- Advanced Opportunity Fellowship, University of Wisconsin-Madison, Dept. of Computer Sciences. 2004-2005

Reviewing Activities:

Journals: *Bioinformatics, Annals of Applied Statistics, Journal of Machine Learning Research, Biostatistics, Journal of Artificial Intelligence Research, Genome Biology, Journal of the Royal Statistical Society (Series C), IEEE Transactions on Computational Biology and Bioinformatics, BMC Genomics, Biometrics, BMC Bioinformatics, Genome Research, Nucleic Acids Research*

Conference Program Committees: *ISMB/ECCB 2013, ISMB 2014, ISMB 2015.*