

## Ming-Ying Leung

Professor, Department of Mathematical Sciences and  
 Director, Bioinformatics and Computational Science Programs  
 The University of Texas at El Paso, 500 W. University Avenue  
 El Paso, TX 79968-0514; Phone: (915)747-6836; Fax: (915)747-6502  
 Email: [mleung@utep.edu](mailto:mleung@utep.edu); Web: [math.utep.edu/Faculty/mleung](http://math.utep.edu/Faculty/mleung)

### Education

1980:	University of Hong Kong	B.Sc.	Mathematics
1983:	University of Hong Kong	M.Phil.	Mathematics
1988:	Stanford University	M.S.	Computer Science
1989:	Stanford University	Ph.D.	Mathematics

### Positions

1980-1983 : Teaching Assistant, Department of Mathematics, University of Hong Kong  
 1983-1989 : Research Assistant and Teaching Fellow, Department of Mathematics, Stanford University  
 1989-2003 : Assistant and Associate Professor, Division of Mathematics and Statistics, The University of Texas at San Antonio  
 1993 : Visiting Research Fellow, Department of Statistics, University of California at Berkeley and Department of Pharmaceutical Chemistry, University of California at San Francisco  
 2001-2002 : Visiting Associate Professor, Department of Statistics, Rice University  
 2003- : Professor, Department of Mathematical Sciences, and Director, Bioinformatics Program, The University of Texas at El Paso  
 2013- : Director, Computational Science Program, The University of Texas at El Paso

### Honors and Services

1986-1987 : Andrew Mellon Foundation Research Award, Institute of Population and Resource Studies, Stanford University  
 1990-1991 : Consultant for the mathematical molecular biology groups at the University of Southern California and Stanford University  
 2002-2003 : Editorial Board Member, Advances and Applications in Statistics  
 2002-2004 : Co-chair, Organizing Committee for the International Workshop on Statistical Methods in Microarray Data Analysis, Institute of Mathematical Sciences, National University of Singapore  
 2004 : Professor Y.C. Wong Visiting Lectureship, University of Hong Kong  
 2005 : Chair, Joint Session in Bioinformatics, 2005 INFORMS Annual Meeting  
 2005-2007 : External Advisory Board Member, NSF and HHMI funded “Talent Expansion in Quantitative Biology” project at East Tennessee State University  
 2007-2008 : Outstanding Performance Award, Office of Research and Sponsored Programs, UTEP  
 2007-2013 : Associate Editor, INFORMS Journal on Computing  
 2008 : Chair, Invited paper session in Stochastic Models for Biological Processes, International Workshop on Applied Probability, July 2008, Compiègne, France  
 2008-2010 : Member, The University of Texas System Computational Biology Workgroup for the Cancer Prevention and Research Institute of Texas  
 2008- : Director, Bioinformatics Computing Core Facility, Border Biomedical Research Center, UTEP  
 2009- : Member, NIH-RCMI Translational Research Network Translational Informatics Subcommittee  
 2010- : External Review Panelist for West Virginia IDeA Network for Biomedical Research Excellence, Research Competitiveness Program, American Association for the Advancement of Science  
 2010 : Chair, Session on Stochastic Models for Biological Systems, 2010 INFORMS Annual Meeting

- 2011 : Chair for invited Session Computational Methods in Biomolecular and Phylogenetic Analyses, International Federation of Operational Research Societies (IFORS) Conference, July 2011, Melbourne, Australia.
- 2011 : Chair for invited Cluster on Computational Biology, Institute for Operations Research and Management Science (INFORMS) Annual Conference, November 2011, Charlotte.
- 2012- : Organizer, Joint UTEP/NMSU Workshop on Mathematics, Computer Science, and Computational Sciences
- 2013- : Mentor, National Alliance for Doctoral Studies in the Mathematical Sciences
- 2016 : Member, Scientific Program Committee, International Workshop on Applied Probability, Toronto
- 2017 : Student Choice Award for Outstanding Teaching, Department of Mathematical Sciences, UTEP

### **Selected Publications and Software Released**

- Leung, M.Y., Blaisdell, B.E., Burge, C., and Karlin, S. (1991) An Efficient Algorithm for Identifying Matches with Errors in Multiple Long Molecular Sequences, *J. Mol. Biol.* 221, 1367-1378.
- Leung, M.Y., Schachtel, G., and Yu, H.S., (1994) Scan Statistics and DNA Sequence Analysis: The Search for an Origin of Replication in a Virus, *Nonlinear World* 1, 445-471.
- Leung, M.Y., Marsh, G.M., and Speed, T.P. (1996) Over- and Underrepresentation of Short Oligonucleotides in Herpes Virus Genomes. *J. Computational Biology*, 3(3), 345-360.
- Leung, M.Y. and Yamashita, T.E. (1999) Applications of the Scan Statistic in DNA Sequence Analysis, in *Scan Statistics and Applications*, pp.269-286. Ed. Balakrishnan, N. and Glaz, J., Birkhauser Publishers.
- Grosu, D., Chronopoulos, A.T., and Leung, M.Y. (2002) Load Balancing in Distributed Systems: An Approach Using Cooperative Games. *Proceedings of the 16<sup>th</sup> IEEE International Parallel and Distributed Processing Symposium, Fort Lauderdale, Florida, USA*, IEEE Computer Science Press, 52-61.
- Chew, D.S.H., Choi, K.P., Heidner, H., and Leung, M.Y. (2004) Palindromes in SARS and Other Coronaviruses, *INFORMS J. Comp.* 16(4), 331-340.
- Leung, M.Y., Choi, K.P., Xia, A. and Chen, L.H.Y. (2005) Nonrandom Clusters of Palindromes in Herpesvirus Genomes, *J. Computational Biology* 12(3), 331-354.
- Chew, D.S.H., Choi, K.P., and Leung, M.Y. (2005) Scoring Schemes of Palindrome Clusters for More Sensitive Prediction of Replication Origins in Herpesviruses, *Nucleic Acids Research* 33 (15), e134.
- McMullin, B.T., Leung, M.Y., Shanbag, A.S., McNulty, D., Mabrey, J.D., and Agrawal, C.M. (2006) Correlating Subjective and Objective Descriptors of Ultra High Molecular Weight Wear Particles from Total Joint Prostheses. *Biomaterials* 27, 752-757.
- Leung, M.Y. (2006) Scan Statistics, in *Encyclopedia of Measurement and Statistics*, pp.865-868, Salkind, N.J., Editor, SAGE Publications, Thousand Oaks.
- Chew, D.S.H., Leung, M.Y., and Choi, K.P. (2007) AT Excursion: a New Approach to Predict Replication Origins in Viral Genomes by Locating AT-rich Regions. *BMC Bioinformatics* 8, 163.
- Grosu, D., Chronopoulos, A.T., and Leung, M.Y. (2008) Cooperative Load Balancing in Distributed Systems. *Concurrency and Computation: Practice and Experience* 20:1953-1976.
- Taufer, M., Leung, M.Y., Solorio, T., Licon, A., Mireles, D., and Johnson, K.L. (2008) RNAVLab: A Virtual Laboratory for Studying RNA Secondary Structures Based on Grid Computing Technology, *Parallel Computing* 34: 661-680.
- Taufer, M., Licon, A., Araiza, R., Mireles, D., Gulyaev, A., Van Batenburg, F.H.D., and Leung, M.Y. (2009) PseudoBase++: An Extension of PseudoBase for Easy Searching, Formatting, and Visualization of Pseudoknots *Nucleic Acids Research* 37(Database Issue):D127-135.
- Cruz-Cano, R. and Leung, M.Y. (2009) Comparison of Feature Selection Techniques for Viral DNA Replication Origin Prediction. *Proceedings of the 2009 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology*, March 2009, Nashville, Tennessee, USA, 128-134.
- Greenberg, H.J., Holder, A.G., Leung, M.Y., and Schwartz, R. (2009) Computational Biology and Medical Applications: Opportunities for Operations Research, *OR/MS Today*, 36(3).

- Cruz-Cano, R., Chew, D.S.H., Choi, K.P., and Leung, M.Y. (2010) Least-Squares Support Vector Machine Approach to Viral Replication Origin Prediction, *INFORMS J. Computing* 22(3), 457-470.
- Roskopf, J.J.; Upton, J.H.III, Rodarte, L., Romero, T.A., Leung, M.Y., Taufer, M. and Johnson, K.L. (2010) A 3' terminal stem-loop structure in Nodamura virus RNA2 forms an essential cis-acting signal for RNA replication. *Virus Research* 150(1-2):12-21.
- Schaubert, K.L., Price, D.A., Salkowitz, J.R., Sewell, A.K., Sidney, J, Asher, T.E., Blondelle, S.E., Adams, S., Marincola, F.M., Joseph, A., Sette, A., Douek, D.C., Ayyavoo, V, Storkus, W., Leung, M.Y., Ng, H.L., Yang, O.O., Goldstein, H., Wilson, D.B., Kan-Mitchell, J. (2010) Generation of robust CD8(+) T-cell responses against subdominant epitopes in conserved regions of HIV-1 by repertoire mining with mimotopes. *Eur. J. Immunol.* 40(7):1950-1962.
- Licon, A., Taufer, M., Leung, M.Y., Johnson, K.L. (2010) A Dynamic Programming Algorithm for Finding the Optimal Segmentation of an RNA Sequence in Secondary Structure Predictions. In: Proceedings of the 2nd International Conference on Bioinformatics and Computational Biology 2010 (BICoB-2010), pp.165-170. Honolulu, Hawaii.
- Zheng, J., Fuentes, O., Leung, M.Y., and Jackson, E. (2010) Mammogram Compression Using Super-Resolution. In J. Martí, A. Oliver, J. Freixenet, and R. Martí (Eds.), Vol. 6136, pp.46-53, Springer-Verlag, Berlin: Lecture Notes in Computer Science. Available at [link.springer.com/chapter/10.1007/978-3-642-13666-5\\_7](http://link.springer.com/chapter/10.1007/978-3-642-13666-5_7). Accessed 3/4/2021.
- Mohl, J., Licon, A., Viswakula, S., Kelley, P., Araiza, R., Kodimala, V., Vegesna, R., Saldivar, L., Yehadego, D., Cardenas, G., Vest, E., Taufer, M., Fuentes, O., Johnson, K. L., and Leung, M.Y. (2012) RNASSA 2.0: RNA Secondary Structure Analysis (Version 2.0.121208). Available at [navlab.utep.edu/downloads](http://navlab.utep.edu/downloads). Accessed 3/4/2021.
- Cruz-Cano, R., Lee, M.L.T., and Leung, M.Y. (2012) Logic Minimization and Rule Extraction for Identification of Functional Sites in Molecular Sequences. *BioData Mining*, 5(10), 1-21. Available at [biodatamining.org/content/5/1/10](http://biodatamining.org/content/5/1/10). Accessed 3/4/2021.
- Aguilar-Bonavides, C., Leung, M.Y., Nakayasu, E. S., Lopes, F. G., Almeida, I. C., (2012). Computational tool for large-scale GPIomic analysis. In: Proceedings of the ACM Conference on Bioinformatics, Computational Biology and Biomedicine, ACM, New York, NY, pp. 585-587. Available at [doi.org/10.1145/2382936.2383029](http://doi.org/10.1145/2382936.2383029). Accessed 3/4/2021.
- Yehdego, D.T., Zhang, B., Taufer, M., Kodimala, V.K.R., Vegesna, R., Viswakula, S., Johnson, K. L., and Leung, M.Y. (2012) Secondary Structure Predictions for Long RNA Sequences. In: Proceedings of the ACM Conference on Bioinformatics, Computational Biology and Biomedicine, ACM, New York, NY, pp. 545-547. Available at [doi.org/10.1145/2382936.2383016](http://doi.org/10.1145/2382936.2383016). Accessed 3/4/2021.
- Yehdego, D.T., Zhang, B., Kodimala, V.K.R., Johnson, K.L., Taufer, M., Leung, M.Y., (2013) Secondary Structure Predictions for Long RNA Sequences Based on Inversion Excursions and MapReduce. In: Proceedings of the 12th IEEE International Workshop on High Performance Computational Biology (HiCOMB 2013), Boston, MA, pp. 1-10. Available at [doi.org/10.1109/IPDPSW.2013.109](http://doi.org/10.1109/IPDPSW.2013.109). Accessed 3/4/2021.
- Viswakula, S., Kodimala, V., Yehadego, D., Vegesna, R., Taufer, M., Leung, M.Y., Johnson, K. L., (2012). Segmenta 2.0: A Bioinformatics Tool for RNA Segmentation (Version 2.0.121208). Available at [navlab.utep.edu/downloads](http://navlab.utep.edu/downloads). Accessed 3/4/2021.
- Aguilar-Bonavides, C., Cardenas, G. A., Nakayasu, E. S., Lopes, F. G., Almeida, I. C., Leung, M.Y., (2013) Automatic Annotation of GPI Structures Using Grid Computing. (pp. 1-6). Honolulu, Hawaii: Proceedings of the 5th International Conference on Bioinformatics and Computational Biology (BICoB 2013).
- Zhang, B., Yehdego, D. T., Johnson, K. L., Leung, M.Y., Taufer, M. (2013) Enhancement of accuracy and efficiency for RNA secondary structure prediction by sequence segmentation and MapReduce. *BMC Structural Biology*, 13 (Suppl. 1)(S3), 1-24. Available at [biomedcentral.com/1472-6807/13/S1/S3](http://biomedcentral.com/1472-6807/13/S1/S3). Accessed 3/4/2021.
- Leung, M.Y., Cardenas, G.A., Almeida, I.C., and Gerken, T.A. (2014) Isoform Specific O-Glycosylation Prediction (ISOGlyP), Version 1.2. Available at [isoglyp.utep.edu](http://isoglyp.utep.edu). Accessed 3/4/2021.

- Guerrero, F., Kellogg, A., Ogrey, A. N., Heekin, A. M., Barrero, R., Bellgard, M. I., Dowd, S. E., Leung, M.Y. (2016) Prediction of G protein-coupled receptor encoding sequences from the synganglion transcriptome of the cattle tick, *Rhipicephalus microplus*. *Ticks and Tick-borne Diseases*. 7(5), 670-677. Available at [doi.org/10.1016/j.ttbdis.2016.02.014](https://doi.org/10.1016/j.ttbdis.2016.02.014). Accessed 3/4/2021.
- Leung, M.Y., Knapka, J.A., Wagler, A.E., Rodriguez, G., Kirken, R.A. (2016) OncoMiner: A pipeline for bioinformatics analysis of exonic sequence variants in cancer. In: *Big Data Analytics in Genomics*, Wong, K.C. (Ed.), pp. 373-396, Springer, New York. Available at [link.springer.com/chapter/10.1007/978-3-319-41279-5\\_12](https://link.springer.com/chapter/10.1007/978-3-319-41279-5_12). Accessed 3/4/2021.
- Munoz, S., Guerrero, F., Kellogg, A., Heekin, A. M., Leung, M.Y. (2017) Bioinformatic prediction of G protein-coupled receptor encoding sequences from the transcriptome of the foreleg, including the Haller's organ, of the cattle tick, *Rhipicephalus australis*. *PLOS ONE* 12(2):e0172326. Available at [doi.org/10.1371/journal.pone.0172326](https://doi.org/10.1371/journal.pone.0172326). Accessed 3/4/2021.
- Leung, M.Y. (2017) Scan Statistics Applications in Genomics. In: *Handbook of Scan Statistics*, Glaz, J. and Koutras, M.V. (Eds.); Springer, New York. Available at [doi.org/10.1007/978-1-4614-8414-1\\_42-1](https://doi.org/10.1007/978-1-4614-8414-1_42-1). Accessed 3/4/2021.
- Rivas, J.A., Mohl, J.E., Van Pelt, R.S., Leung, M., Wallace, R.L., Gill, T. E. and Walsh, E. J. (2018) Evidence for regional aeolian transport of freshwater micrometazoans in arid regions. *Limnology and Oceanography Letters*. Available at [doi.org/10.1002/lo2.10072](https://doi.org/10.1002/lo2.10072). Accessed 3/4/2021.
- Vasquez, M., Mohl, J.E., Leung, M.Y. (2018) Parsing Next Generation Sequencing Data in Parallel Environments for Downstream Genetic Variation Analysis. *Journal of Computational Science Education* 9(2):37-45. Available at [doi.org/10.22369/issn.2153-4136/9/2/5](https://doi.org/10.22369/issn.2153-4136/9/2/5). Accessed 3/4/2021.
- Patil, A. R., Ayivor, F., Mohl, J. E., Leung, M.Y., Kim, S. (2019) Analyzing epigenetics data using feature ranking with various classifiers. *Computational and Mathematical Biophysics*, 7, 98-120. [doi.org/10.1515/cmb-2019-0008](https://doi.org/10.1515/cmb-2019-0008). Accessed 3/4/2021.
- Begum, K., Mohl, J. E., Ayivor, F.\*, Perez, E. E.\*, Leung, M.Y., (2020). GPCR-PEnDB: A database of protein sequences and derived features to facilitate prediction and classification of G protein-coupled receptors. *Database*. [academic.oup.com/database/article/doi/10.1093/database/baaa087/5995841](https://academic.oup.com/database/article/doi/10.1093/database/baaa087/5995841). Accessed 3/4/2021.
- Mohl, J. E., Gerken, T., Leung, M.Y., (2020). ISOglyP: de novo prediction of isoform specific mucin-type O-glycosylation. *Glycobiology*. [doi.org/10.1093/glycob/cwaa067](https://doi.org/10.1093/glycob/cwaa067). Accessed 3/4/2021.
- Wang, B., Mohl, J. E., Leung, M.Y., (2020). Computational Prediction of Functional Effects for Cancer Related Genetic Sequence Variants. (pp. 2999-3001). Seoul: IEEE International Conference on Bioinformatics and Biomedicine (BIBM). [doi.ieeecomputersociety.org/10.1109/BIBM49941.2020.9313326](https://doi.ieeecomputersociety.org/10.1109/BIBM49941.2020.9313326). Accessed 3/4/2021.
- Mohl, J. E., Gerken, T., Leung, M.Y., (2020). Predicting mucin-type O-Glycosylation using enhancement value products from derived protein features. *Journal of Theoretical and Computational Chemistry*, 19(3), 2040003. [doi.org/10.1142/S0219633620400039](https://doi.org/10.1142/S0219633620400039). Accessed 3/4/2021.

### **Online Education and Curriculum Modules**

- Bioinformatics Information Modules for Education and Research (BIMER): First Release in 2000; Updated Version Available at [math.utep.edu/Faculty/mleung/BIMER](https://math.utep.edu/Faculty/mleung/BIMER). Accessed 3/4/2021.
- Probability Research Information Statistics Education Modules (PRISEM): First Release in 2001; Updated Version Available at [math.utep.edu/Faculty/mleung/PRISEM](https://math.utep.edu/Faculty/mleung/PRISEM). Accessed 3/4/2021.
- Undergraduate Participation in Bioinformatics Training (UPBiT) Curriculum Modules: First Release in 2010; Updated Version Available at [bioinformatics.utep.edu/UPBiT/curriculum](https://bioinformatics.utep.edu/UPBiT/curriculum). Accessed 3/4/2021.
- Bioinformatics Education for Agricultural Science (BEAS) Curriculum Modules: First Release in 2012; Updated Version Available at [bioinformatics.utep.edu/agriculture/curriculum](https://bioinformatics.utep.edu/agriculture/curriculum). Accessed 3/4/2021.

### **Selected Invited Presentations**

- Leung, M.Y., NMSU Department of Mathematical Sciences Colloquium, "Statistical models for predicting functional effects of genetic sequence variants in cancer," New Mexico State University at Las Cruces. (November 6, 2020).

- Leung, M.Y., AMS Central Fall Virtual Sectional Meeting, "Scan statistics for finding inversion clusters in genomic sequences," American Mathematical Society. (September 12, 2020).
- Mohl, J. E., Gerken, T., Leung, M.Y., AMS Central Fall Virtual Sectional Meeting, "Using derived protein feature enhancement values to calculate the propensity for mucin-type O-glycosylation with the ISOGLyP program," American Mathematical Society. (September 12, 2020).
- Leung, M.Y., Mohl, J. E., GlyGen Monthly Talks, "ISOGLyP: de novo prediction of isoform specific mucin-type O-glycosylation," www.glygen.org, Online. (April 7, 2020).
- Leung, M.Y., Mean and Variance of Inversion Counts in Nucleotide Sequences. The 8th International Workshop on Applied Probability (IWAP2016), Toronto, Canada. (June 22, 2016).
- Leung, M.Y., Identifying Inversion Clusters in Viral Genomes by Scan Statistics. The 8th International Workshop on Applied Probability (IWAP2016), Toronto, Canada. (June 20, 2016).

### **Ongoing Research Support**

2019-2024: NIMHD (RCMI) – “Border Biomedical Research Center: Research Infrastructure Core”;  
\$1,461,063, Co-Director

### **Past Research Support**

- 2014-2019: NIH (U01)/Case Western University – “Initiation and Regulation of Mucin Type O-glycosylation”;  
\$190,720, Subcontract PI
- 2014-2019: NIH – “Border Biomedical Research Center: Bioinformatics Core Facility”;  
\$505,376, Bioinformatics Core Director
- 2014-2017: NIH – “Mechanisms in Viral RNA Replication Complex Assembly: Novel Targets For Antivira”;  
\$453,000, Co-I
- 2012-2016: USDA – “Bioinformatics Education for Agricultural Science”;  
\$295,000, PI
- 2010-2016: NSF – “Recruiting and Keeping Undergraduate Students in the Sciences”;  
\$552,000, Co-PI
- 2009-2016: NSF Interdisciplinary Training for Undergraduates in Biological and Mathematical Sciences (UBM) – “UBM Institutional: Undergraduate Training in Bioinformatics”;  
\$870,000, PI
- 2009-2014: NIH/NCRR – “Border Biomedical Research Center: Bioinformatics Computing Core Facility”;  
\$279,605, Bioinformatics Core Director
- 2008-2013: NIH (MARC) – “Enhancement of Quantitative Science in Biology Curricula”;  
\$1,187,256, Co-PI
- 2008-2013: NSF Mathematical Sciences – “Mathematical Models for RNA”;  
\$415,632, PI
- 2008-2012: NIH (R01) – “Mapping Novel Subdominal B\*5701 Epitopes in Conserved Regions of the HIV Proteome”;  
\$2,353,543; Co-I
- 2007-2011: NIH Support for Continuous Research Excellence (SCORE) Grant – “Computational Prediction of RNA Viral Genome Structures”;  
\$606,952, Subproject PI
- 2006-2010: THECB Norman Hackerman Advanced Research Program, \$99,982 and \$117,780, PI
- 2005-2007: NIH/Pittsburgh Supercomputer Center; \$79,968, Subcontract PI
- 2005-2007: IBM Shared University Research (SUR) Awards Program; \$600,000, Subproject Co-PI
- 2003-2007: NIH Support for Continuous Research Excellence (SCORE) Grant; \$475,946, Subproject PI
- 2004-2006: NSF Major Research Instrumentation; \$207,152, Co-I
- 2004-2005: UTEP Advance GRA Grant; \$4,400, Co-PI
- 2004-2005: NIH Minority Access to Research Careers (MARC) Program; \$54,000, Subproject Co-I
- 2001-2003: NIH Support for Continuous Research Excellence (SCORE) Grant; \$147,798, Subproject PI
- 2000-2003: NSF Course, Curriculum, and Laboratory Improvement Grant; \$122,351, PI
- 1996-1999: NSF Research Improvement in Minority Institutions Grant (RIMI); \$134,863, Co-PI
- 1992-1994: Texas Higher Education Coordinating Board Advanced Research Program; \$74,852, PI
- 1991-1995: San Antonio Area Foundation Research Grants, \$9,500 (1992); \$13,775 (1994), PI
- 1991-1993: NSF Research Grant, Probability and Statistics Program; \$70,352, PI
- 1990-1992: Texas Higher Education Coordinating Board Advanced Research Program; \$90,000, PI