

## Michael N. Molla, Ph.D.

---

| Room 104A | Life Science and Engineering Building | Boston University | 24 Cummington St. | 02215 | (617)358-2506 |

---

### Education

**Ph.D.** in Computer Sciences with Genetics Minor, 2007  
The University of Wisconsin, Madison, Madison, WI

**M.Eng.** in Computer Science (AI concentration), 1999  
The University of Colorado at Boulder, Boulder, CO

**B.A.** in Computer Science, 1993  
Brandeis University, Waltham, MA

---

### Publications

- T. Albert, **M. Molla**, D. Muzny, L. Nazareth, D. Wheeler, X. Song, T. Richmond, C. Middle, M. Rodesch, C. Packard, G. Weinstock, and R. Gibbs (2007)  
Direct Selection of Human Genomic Loci by Microarray Hybridization *Nature Methods*, 4:11 pp. 907-909
- **M. Molla** (2007).  
Novel Uses for Machine Learning and Other Computational Methods for the Design and Interpretation of Genetic Microarrays. PhD thesis, Department of Computer Sciences, University of Wisconsin-Madison.  
(Also appears as UW Technical Report CS-TR-07-1612)
- E. Haag & **M. Molla** (2005).  
Compensatory Evolution of Interacting Gene Products Through Multifunctional Intermediates. *Evolution*, 59, pp. 1620-32
- T. Albert, D. Dailidienė, G. Dailidienė, J. Norton, A. Kalia, T. Richmond, **M. Molla**, J. Singh, R. Green & D. Berg (2005).  
Mutation Discovery in Bacterial Genomes: Metronidazole Resistance in Helicobacter Pylori. *Nature Methods*, 2:12 pp. 951-953
- **M. Molla**, J. Shavlik, T. Albert, T. Richmond & S. Smith (2004).  
A Self-Tuning Method for One-Chip SNP Identification. *Proceedings of the IEEE Conference on Computer Systems Bioinformatics* (CSB 2004), Stanford, CA
- **M. Molla**, M. Waddell, D. Page & J. Shavlik (2004).  
Using Machine Learning to Design and Interpret Gene-Expression Microarrays. *AI Magazine*, 25, pp. 23-44.
- **M. Molla**, P. Andrae & J. Shavlik (2004).  
Building Genome Expression Models using Microarray Expression Data and Text. Department of Computer Sciences, University of Wisconsin, Machine Learning Research Group Working Paper 04-1.

- **M. Molla**, P. Andrae, J. Glasner, F. Blattner & J. Shavlik (2002). Interpreting Microarray Expression Data Using Text Annotating the Genes. *Information Sciences*, 146, pp. 73-86. Also appears in: *Proceedings of the 4th Conference on Computational Biology and Genome Informatics*, Durham, NC
- J. Tobler, **M. Molla**, E. Nuwaysir, R. Green & J. Shavlik (2002). Evaluating Machine Learning Approaches for Aiding Probe Selection for Gene-Expression Arrays. *Bioinformatics*, 18, Special Issue Based on the Papers Presented at the *Tenth International Conference on Intelligent Systems for Molecular Biology (ISMB-02)*, Edmonton, Canada, pp. S164-S171.
- E. Nuwaysir, W. Huang, T. Albert, J. Singh, K. Nuwaysir, A. Pitas, T. Richmond, T. Gorski, J. Berg, J. Ballin, M. McCormick, J. Norton, T. Pollock, T. Sumwalt, L. Butcher, D. Porter, **M. Molla**, C. Hall, F. Blattner, M. Sussman, R. Wallace, F. Cerrina, and R. Green (2002). Gene Expression Analysis Using Oligonucleotide Arrays Produced by Maskless Photolithography. *Genome Research*, Vol. 12, Issue 11, 1749-1755.

## Employment Experience

**September 2007 – Present**

### **Center for BioDynamics Fellow**

#### **Boston University, Department of Biomedical Engineering in Boston, MA**

I am currently working with Professors Timothy Gardner and Simon Kasif to develop new tools and techniques for the analysis of microbes.

**May 2001 – August 2007**

### **Biotechnology Consultant**

#### **Nimblegen Systems (now Roche Nimblegen) in Madison, WI**

While in graduate school, I am worked part time with Nimblegen's Bioinformatics Department applying machine learning techniques to the design and interpretation of DNA expression arrays.

**September 1999 – August 2007**

### **Graduate Researcher**

#### **The University of Wisconsin's Computer Sciences Department in Madison, WI**

For my Ph.D. in Computer Science, I worked, in conjunction with my advisor, Jude Shavlik, and Nimblegen Systems, to apply machine learning techniques to design gene chips and to interpret microarray data.

**June 1997 - June 1999**

### **Software Engineer**

#### **Whitehead/MIT Center for Genome Research in Cambridge, MA**

A member of the center's sequencing informatics team, I designed and created custom software still in use primarily by scientists sequencing and finishing the human genome. In some cases my software makes tasks easier for humans. In others, it takes them out of the process altogether. I also designed and performed numerous virtual assembly experiments including the 1998/1999 Intermediate Product/Draft Sequence viability experiments for the NIH and helped to write the center's SNP identification software. This work involved C, C++, Java and Perl in a UNIX/WinNT environment.

**December 1992 - May 1997**

### **Programmer**

#### **IBM in Tucson, AZ and Boulder, CO**

*May 1996 - May 1997*: I was a member of IBM's network printer driver department, responsible for IBM's PCL printer driver development team for Windows3.1 and Windows95. I won a \$1000 award for this effort.

*May 1995 - May 1996*: I helped to architect and implement the AFCCU2, the AFCCU's object oriented counterpart,

which is currently in use on IBM's midrange printers. I filed an invention disclosure regarding one of my modules in this system.

*October 1994 - May 1995:* I was the technical team leader for the adaptation of the AFCCU's user interface to our midrange printers. I won a \$2500 award for this effort.

*January 1993 - October 1994:* I helped to design and implement the user IBM's Advanced Function Common Control Unit(AFCCU) which has since been used to control printers from 30 to 300 ppm.

All of the high-end and midrange development was done in C and C++ on a RISC6000 using AIX(IBM's version of UNIX) in an XWindows environment. The network printer driver development was in C and involved Windows(3.1 & 95) and OS/2.

#### **Summer 1992**

##### **Internship**

##### **National Library of Medicine in Bethesda, MD**

During this time, I helped to update the MEDLINE user interface and implemented data compression/decompression and search algorithms. I used an IBM PC and a Power Mac to accomplish this.

#### **Summer 1991**

##### **Internship**

##### **Roswell Park Cancer Institute in Buffalo, NY**

My main responsibility was computer-aided molecular modeling. The object was to guide researchers in the production of genetic probes in order to someday cure human autoimmune disorders. This was done on an IBM PC, a Macintosh and a remote mainframe. I was also the laboratory's computer troubleshooter.

#### **Summer 1990**

##### **Internship**

##### **University of Arizona's Mathematics Department in Tucson, AZ**

I wrote programs to aid professors in their original research projects. One program solved differential equations to model the effect of incident laser pulses on the human eye as part of a project for ANSI. The other modeled chaotic functions in order to solve a structural engineering problem. They were written in Pascal on an IBM mainframe and in FORTRAN on a SUN workstation, respectively.

### **Honors**

- NLM Computation and Informatics in Biology and Medicine Fellowship
- NIH Biotechnology Training Program Fellowship
- University of Wisconsin Advanced Opportunity Fellowship
- Two IBM informal awards for technical achievement
- Two varsity letters for NCAA intercollegiate track&field
- National Achievement Scholarship
- Scholarships Won in the National Academic Decathlon

### **Special Skills**

**Languages:** C, C++, C\*, FORTRAN, HTML, Java, LISP, Pascal, perl, PCL, PostScript, Prolog, Scheme, SQL  
**Operating Systems:** DOS, OS/2, UNIX, Windows (NT, XP, etc.), XWindows  
**Tools/Packages:** CGI, Lex, YACC, sed, awk, make, Mosaic, MSNSDK, MSNDDK, IBM collection class and GUI libraries  
**Library Systems:** SCCS, CMVC, RCS

**Outside Interests:** I ran intercollegiate track and cross-country and was a college radio DJ. I have been playing the drums in various local (to me) musical groups since junior high school.

## References

Jude W. Shavlik  
Professor  
Departments of Computer Sciences and Biostatistics & Medical Informatics  
University of Wisconsin-Madison  
1210 West Dayton Street  
Madison, WI 53706  
(608)262-7784

C. David Page Jr.  
Associate Professor  
Department of Biostatistics and Medical Informatics and Department of Computer Sciences  
University of Wisconsin Medical School  
Medical Science Center, Rm 6743  
1300 University Avenue  
Madison, WI 53706  
(608)265-6168

Roland Green, Ph.D.  
Vice President of Research and Development and Chief Technology Officer  
Nimblegen Systems, Inc.  
1 Science Court  
Madison, WI 53711  
(608)218-7600

Frederick Blattner  
Oliver Smithies Professor of Genetics  
University of Wisconsin-Madison  
5430 Genetics/Biotech  
425-G Henry Mall  
Madison, WI 53706  
(608)890-0191

Additional references furnished upon request.