BIOGRAPHICAL SKETCH

NAME	POSITION TITLE
Debraj GuhaThakurta	Senior Research Scientist

CONTACT INFORMATION
10930 Forbes Creek Drive, #S-307
Kirkland, WA 98033. USA.
Email: d.guha@comcast.net
Phone: (425)828-9761

EDUCATION

INSTITUTION AND LOCATION	DEGREE	YEAR(s)	FIELD OF STUDY
Presidency College, University of Calcutta, Calcutta, India	B.Sc.	1989-1992	Chemistry (Honors)
Indian Institute of Technology, Kanpur, India	M.Sc.	1992-1994	Chemistry (Organic)
The Johns Hopkins University, Baltimore, MD	Ph.D.	1994-1999	Chemistry (Biophysical)
Washington University School of Medicine, St. Louis, MO	Postdoctoral	1999-2001	Genetics (Computational Biology)

ADDITIONAL EDUCATION/TRAINING

1998: Short course in computational biology. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. Duration: 1 week.
2005: Summer course in statistical genetics. North Carolina State University, Raleigh, NC Duration: 3 weeks.

EMPLOYMENT

1994-1996:
Teaching Assistant, Department of Chemistry, The Johns Hopkins University, Baltimore, MD

1996-1999:
Research Assistant, Department of Chemistry, The Johns Hopkins University, Baltimore, MD
Advisor: David E. Draper
Research topic: Protein-RNA recognition.
Thesis title: "Ribosomal RNA recognition by ribosomal protein L11: The role of conserved residues and implications of protein-RNA sequence covariation."

1999-2001:
Research Associate, Department of Genetics, Washington University School of Medicine, St. Louis, MO
Advisor: Gary D. Stormo
Research topic: Computational methods for the identification of transcription regulatory elements in eukaryotic genomes.

2001-Present:
Senior Research Scientist, Department of Informatics and Department of Research Genetics, Rosetta Inpharmatics LLC (a wholly owned subsidiary of Merck & Co.), Seattle, WA
Research areas: Annotation of human genome using microarray technology, identification of causal genes for complex diseases through integrative genomics using mouse intercrosses and genetics of gene-expression.

GRANT REVIEW

2002: Grant Proposal Reviewer. The Biomedical Research Council, Government of Singapore.
2005: Grant Proposal Reviewer. Genome Canada, Competition III

JOURNAL REVIEW AND EDITORSHIP

2001: Associate Guest Editor. IEEE Intelligent Systems (Sep/Oct 2001), special issue on Intelligent Systems in Biology.
Reviewer. The Journal of Molecular Biology.

2002: Reviewer. Bioinformatics.
Associate Guest Editor. IEEE Intelligent Systems (March/April 2002), special issue on Intelligent Systems in Biology.
2003: Reviewer. Nucleic Acids Research.
2004: Reviewer. Nucleic Acids Research, Bioinformatics.
2005 Reviewer. Bioinformatics.

THESIS SUPERVISED

2004: Master's Thesis. Manish Anand, School of Informatics, Indiana University, Bloomington (worked as an extended intern at Rosetta Inpharmatics LLC, Seattle). Thesis title: "Effect of polymorphisms in transcriptional regulation in mice".

MEMBERSHIPS

1. International Society for Computational Biology (ISCB)
2. American Society of Human Genetics (AHSG)

TECHNICAL EXPERTISE

Bioinformatics and Statistical Genetics:

- Algorithms for machine learning and sequence analysis, pattern recognition, data clustering.
- Knowledge and use of methods for sequence analysis and structure prediction programs (BLAST, FASTA, Sim4, Grail, Genscan, HMMer, Mfold, and motif finding programs like Gibbs Sampler, MEME, Consensus etc.).
- Database design for genomic and genetic data using SQL-server.
- Experience of integration of annotation resources like LocusLink, KEGG, MGI (Mouse Gene Index), Transfac etc.
- Software for genetic QTL mapping in experimental populations (QTL Cartographer suite).
- Statistical programming with R. Data analysis with statistical software like JMP, Spotfire. Expression data analysis with Rosetta Resolver.

Platforms/Programming:

- UNIX/LINUX (including clusters), Windows NT/XP and Macintosh OS.
- Extensive experience with Perl and SQL-server.
- Programming experience in R, C++, Matlab, HTML, Oracle RDBMS.

Molecular Biology:

- Cloning, mutagenesis, PCR, RT-PCR, DNA gel sequencing, protein expression and purification, *in vitro* transcription and large scale RNA purification.

Biophysics:

- Experience in macro-molecular interaction, especially protein-nucleic acid recognition. Gel electrophoresis, gel retardation assays, determination of protein-nucleic acid binding affinities, NMR of small molecules, RNA UV-melting.
- Analysis of RNA structures and ligand-RNA interactions from melting experiments and circular dichroism (CD) spectroscopy.

PUBLICATIONS

1. Maikap, G.C., **GuhaThakurta, D.**, and Iqbal, J., 1995. Cobalt catalyzed benzylic oxidation with molecular oxygen. *SYNLETT*, **2**, 189-191.
2. Xing, Y., **GuhaThakurta, D.**, and Draper, D.E., 1997. The RNA binding domain of ribosomal protein L11 is structurally similar to homeodomains. *Nature Structural Biology,* **4**, 24-27.
3. **GuhaThakurta, D.**, and Draper, D.E., 1999. Protein-RNA sequence co-variation in ribosomal RNA-L11 complex. *Biochemistry,* **38**, 3633-3640.
4. Draper, D.E., Conn, G.L., Gittis, A.G., **GuhaThakurta**, **D.**, Lattman, E.E., and Reynaldo, L.P., 2000. RNA tertiary structure and protein recognition in L11-RNA complex. In Garrett, R.A., *et. al.* (*eds.*) *The*

Ribosome: Structure Function, Antibiotic and Cellular Interactions. (American Society for Microbiology Press, Washington D.C.)
5. **GuhaThakurta, D.**, and Draper, D.E., 2000. Contributions of basic residues to ribosomal protein L11 recognition of RNA. *Journal of Molecular Biology,* **295**, 569-580.
6. **GuhaThakurta, D.**, and Stormo, G.D., 2001. Identifying target sites for cooperatively binding factors. *Currents in Computational Molecular Biology*, pp. 39-40, El-Mabrouk N., Lengauer, T., and Sankoff, D., (eds.) Les Publications CRM, Montreal, Canada. (Extended abstract. Full paper appears in *Bioinformatics*, **17**, 608-621).
7. **GuhaThakurta, D.**, and Stormo, G.D., 2001. Identifying target sites for cooperatively binding factors. *Bioinformatics*, **17**, 608-621.
8. **GuhaThakurta, D.**, Schriefer, L.A., Hresko, M.C., Waterston, R.H. and Stormo, G.D., 2002. Identifying muscle regulatory elements and genes in the nematode *Caenorhabditis elegans*. *Proceedings of the Pacific Symposium on Biocomputing,* **7**, 425-436.
9. **GuhaThakurta, D.**, Palomar, L., Stormo, G.D., Tedesco, P., Johnson, T.E., Walker, D., Lithgow, G., Kim Stuart., and Link, C.D., 2002. Identification of a n ovel *cis*-regulatory element involved the heat shock response in *C. elegans* using microarray gene expression and computational methods. *Genome Research,* **12**, 701-712.
10. Tata, P., Miles, S., **GuhaThakurta, D.**, Jemiolo, D., and Breeden, L. 2002. Conserved homeodomain proteins interact with MADS box protein Mcm1 to restrict ECB-dependent transcription to the M/G1 phase of the cell cycle. *Genes and Development,* **16**, 3034-3045.
11. Ying, L., Schadt, E., Svetnik, V., Holder, D., Edwards, S. and **GuhaThakurta D.** 2003. Identification of chromosomal regions containing transcribed sequences using microarrays and computational methods. *Proceedings of the Annual Meeting of the American Statistical Association,* pp. 4672-4677.
12. Schadt, E., Edwards, S., **GuhaThakurta, D.**, Holder D., Ying, L., et.al., 2004. A comprehensive transcript index of the human genome generated using microarrays and computational approaches. *Genome Biology, **5**(10),* R73.
13. **GuhaThakurta, D.**, Schriefer, L.A., Waterston, R.H., and Stormo, G.D. 2004. Identification of novel transcription regulatory elements in *Caenorhabditis elegans* muscle genes. *Genome Research,* **14**, 2457-2468.
14. Dunstan, M.S., **GuhaThakurta, D.**, Draper, D.E., and Conn, G.C., 2005. Co-evolution of protein and RNA structures within a highly conserved ribosome domain. *Chemistry and Biology*, **12**, 201-206.
15. Schadt, E.E., Lamb, J.R., Yang, X., Zhu, J., Edwards, S.E., **GuhaThakurta, D.**, Sieberts, S., Monks, S., et.al., 2005. An integrative genomics approach to infer causal associations between gene expression and disease. *Nature Genetics*, **37**, 710-717.

REFERENCES

1. Dr. Gary D. Stormo
Professor
Department of Genetics
Washington University
Campus Box 8510, Room 5410
4444 Forest Park Parkway
St. Louis, MO 63108
Email: stormo@ural.wustl.edu
Phone: (314) 747-5534

2. Dr. David E. Draper
Professor
Department of Chemistry
Remsen Hall, Room # 154
3400 N. Charles Street
Baltimore, MD 21218
Email: draper@jhu.edu
Phone: (410) 516-7448

3. Dr. Panayiotis (Takis) V. Benos
Assistant Professor
Department of Human Genetics
Center for Computational Biology & Bioinformatics
University of Pittsburgh
W1041 Biomedical Sciences Tower
200 Lothrop Street,
Pittsburgh, PA 15213
E-mail: benos@pitt.edu
Phone: (412) 648-3315

4. Dr. Christopher D. Link
Research Associate Professor
Institute for Behavioral Genetics
University of Colorado at Boulder
Campus Box 447
Boulder, CO 80309
Email: linkc@colorado.edu
Phone: (303) 735-5112

5. Dr. Chris Tsz-Kwong Man
Assistant Professor
Section of Hematology/ Oncology
Department of Pediatrics
Baylor College of Medicine
Head of Bioinformatics Laboratory
Cancer Genomics Group
Texas Children's Cancer Center
6621 Fannin St MC 3-3320
Houston, TX 77030-2399
E-mail: tman@bcm.tmc.edu
Phone: (832) 824-4682