

Saurabh Sinha

The Rockefeller University
Center for Studies in Physics and Biology, Box 25
1230 York Ave, New York, NY 10021
Phone : (212) 327-8835

<http://www.cs.washington.edu/homes/saurabh>
saurabh@lonnrot.rockefeller.edu
Cell Phone : (609) 947-8376
Home Phone : (609) 430-2226

RESEARCH INTERESTS

Computational Biology : Algorithms to solve problems in molecular biology.
Gene Regulation and its Evolution. Comparative Genomics.

EDUCATION

Ph.D. (Computer Science and Engineering), University of Washington, Seattle. (1997-2002.)
Advisor : Prof. Martin Tompa. Topic : *Finding Regulatory Elements in Genomic Sequences*.
B.Tech. (Computer Science and Engineering)
Indian Institute of Technology, Kanpur. (1993-1997.)

WORK EXPERIENCE

Postdoctoral Fellow, Rockefeller University, New York. (2002-present.)
– Advisor : Prof. Eric Siggia.
– Research in Computational Biology; study of regulatory modules and their evolution.
Research Assistant, University of Washington, Seattle. (2000-2002.)
– Advisor : Prof. Martin Tompa.
– Research in Computational Biology; discovery of transcription factor binding sites.

HONORS AND ACHIEVEMENTS

Best Paper Award (2nd prize). Intelligent Systems for Molecular Biology (ISMB) 2003, Brisbane, Australia. “A Probabilistic Method to Detect Regulatory Modules” – S. Sinha, E. van Nimwegen, E. D. Siggia.
W.M. Keck Foundation Presidential Fellowship. (2002-2004.) Awarded in support of mathematical and computational biology.
Microsoft Research Fellowship. (2001-2003.) Awarded to ~ 12 graduate students from Computer Science, Electrical Engineering or Mathematics departments in the U.S. or Canada.
News report in WIRED magazine (“MS May Have File-Trading Answer”, May 1, 2001) on work done at Microsoft Research – “A Graph-Theoretic Approach to Software Watermarking”.
Placed 25th (nation-wide), I.I.T. Entrance Examination (India), 1993. (~ 100,000 candidates.)
Certificate of Merit for Academic Performance, I.I.T. Kanpur. (1993-1994.)
N.C.E.R.T. Scholarship, awarded by the Govt. of India. (1993-1997.)

INDUSTRIAL EXPERIENCE

Research Internship, Microsoft Research, Redmond. (1999, 2000.)

- Advisor : Dr. R. Venkatesan.
- Designed and developed a software system deployed in Microsoft products for protecting digital goods using randomized and automatic code obfuscation.
- Designed and developed a tool that patches program binaries by finding local similarities in their control flow graphs.
- Designed algorithms for software anti-piracy.

TEACHING EXPERIENCE

Teaching Assistant, University of Washington, Seattle. (1997-2000.)

- *Machine Organization & Assembly Language* (CSE 378), *Computer Programming I* (CSE 142), *Computer Programming II* (CSE 143).
Taught quiz sections.
- *Design and Analysis of Algorithms I* (CSE 521, Graduate course), *Computability and Complexity* (CSE 531, Graduate course), *Compiler Construction* (CSE 401), *Applied Algorithms* (CSE 589PM, Professional Masters Degree course), *Introduction to Theory of Computation* (CSE 431).

PROFESSIONAL SERVICES

Journal reviewing Bioinformatics, BMC Bioinformatics, BMC Genetics.

Conference reviewing Research in Computational Molecular Biology (RECOMB 2004), Pacific Symposium on Biocomputing (PSB 2004, PSB 2005).

Program Committee Member, RECOMB Satellite Workshop on Regulatory Genomics, 2004.

PUBLIC SOFTWARE AND WEB TOOLS FOR RESEARCH

YMF Software and Web Server : The YMF and FindExplainers programs for motif-finding have been downloaded by over 500 users all over the world. The programs are also regularly accessed via our web server at <http://abstract.cs.washington.edu/~saurabh/YMFWeb/YMFInput.pl>

Stubb Software : The Stubb software for module detection has over 50 users world-wide, to date. It is available for download at <http://edsc.rockefeller.edu/cgi-bin/stubb/download.pl>

PhyME Software : The PhyME software for motif-finding in orthologous sequences is available for download at <http://edsc.rockefeller.edu/cgi-bin/phyme/download.pl>

THESES

Ph.D. thesis : “Finding Regulatory Elements in Genomic Sequences”. 2002.

Masters thesis : “Finding Motifs in Promoter Regions of Co-expressed Genes”. 2000.

PUBLICATIONS

Peer-Reviewed Journals

- [1] S. Sinha, M. Blanchette, M. Tompa. “PhyME: A Probabilistic Algorithm for Finding Motifs in Sets of Orthologous Sequences”. *BMC Bioinformatics* 2004, 5(170). (**Research Highlight.**)
- [2] S. Sinha, M. Schroeder, U. Unnerstall, U. Gaul, E. D. Siggia. “Cross-species comparison significantly improves genome-wide prediction of cis-regulatory modules in *Drosophila*”. *BMC Bioinformatics* 2004, 5(129).
- [3] S. Sinha (within a group, M. Tompa *et al.*). “An Assessment of Computational Tools for the Discovery of Transcription Factor Binding Sites”. *Nature Biotechnology* 2004. (To appear.)
- [4] S. Sinha. “Discriminative Motifs”. *Journal of Computational Biology* 2003, 10(3), 599-615. A preliminary version appeared in RECOMB, 2002.
- [5] S. Sinha, E. van Nimwegen, and E.D. Siggia. “A Probabilistic Method to Detect Regulatory Modules”. *Bioinformatics*, 2003, 19(S1). (Special issue on *ISMB* 2003.)
- [6] S. Sinha and M. Tompa. “YMF, a Program for Discovery of Novel Transcription Factor Binding Sites by Statistical Overrepresentation”. *Nucleic Acids Research*, 2003, 31(13), 3586-8.
- [7] S. Sinha and M. Tompa. “Discovery of Novel Transcription Factor Binding Sites by Statistical Overrepresentation”. *Nucleic Acids Research*, 2002, 30(24), 5549-60.
- [8] M. Blanchette and S. Sinha (equal authors). “Separating Real Motifs from their Artifacts”. *Bioinformatics*, 2001, 17(S1). (Special issue on *ISMB* 2001.)

Peer-Reviewed Conferences

- [9] S. Sinha and M. Tompa. “A Statistical Method for Finding Transcription Factor Binding Sites”. *Eighth International Conf. on Intelligent Systems for Molecular Biology (ISMB)*, 2000.
- [10] A. Prakash, M. Blanchette, S. Sinha, M. Tompa. “Motif Discovery in Heterogeneous Sequence Data”. *Ninth Pacific Symposium on Biocomputing (PSB)*, 2004.
- [11] S. Sinha and M. Tompa. “Performance Comparison of Algorithms for Finding Transcription Factor Binding Sites”. *Third IEEE Symposium on Bioinformatics and Bioengineering (BIBE)*, 2003.
- [12] Y. Chen, R. Venkatesan, M. Cary, R. Pang, S. Sinha, M. Jakubowski. “Oblivious hashing: Silent Verification of Code Execution”. *Fifth International Workshop on Information Hiding (IHW)*, 2002.
- [13] R. Venkatesan, V. Vazirani, S. Sinha. “A Graph-Theoretic Approach to Software Watermarking”. *Fourth International Workshop on Information Hiding (IHW)*, 2001.

Patents

- [14] S. Sinha, M.H. Jakubowski, R. Venkatesan, Y. Chen, M. Cary, R. Pang. “Integrity ordainment and ascertainment of computer-executable instructions”. Patent Application (20030191942) filed April, 2002.
- [15] S. Sinha, M.H. Jakubowski, R. Venkatesan, Y. Chen, M. Cary, R. Pang. “Integrity ordainment and ascertainment of computer-executable instructions with consideration for execution context”. Patent Application (20030191940) filed April, 2002.

Papers Submitted to Peer-Reviewed Journals

[16] S. Sinha and E. D. Siggia. “Sequence Turnover and Tandem Repeats in cis-Regulatory Modules in *Drosophila*”. Submitted to *Molecular Biology & Evolution*. 2004.

Manuscripts in Preparation

[17] S. Sinha, M. Pearce, E. D. Siggia, U. Gaul. “Evolution in the regulation of A-P patterning between *D. melanogaster* and *D. pseudoobscura*”. We examined cis-regulatory modules using computational methods for cross-species comparison, and found several modules with a substantial change in regulatory content (and inferred function). The changes fell into distinct categories that point toward general principles of evolution of transcriptional control. Experimental verification of the predicted differences has been completed.

[18] S. Sinha and R. Venkatesan. “GraphDiff: Matching and Patching Program Binaries”.

INVITED LECTURES

- Rutgers University, BioMaPS Summer School - course on transcription. July 2004.
- Cornell University, Tri-Institutional computational biology program. August 2004.
- University of Pennsylvania, Penn. Bioinformatics Forum. November 2003.
- Cold Spring Harbor Laboratory, Genome Informatics meeting. May 2003.

REFERENCES

Prof. Martin Tompa, Ph.D. advisor.
Department of Computer Science and Engineering
University of Washington
Box 352350
Seattle, WA 98195-2350.
tompa@cs.washington.edu. (206) 543-9263

Prof. Eric Siggia, Post-doctoral advisor.
Center for Studies in Physics and Biology
Box 25
Rockefeller University
1230 York Avenue
New York, N.Y. 10021.
siggiae@mail.rockefeller.edu. (212) 327 8546

Prof. Ulrike Gaul, Research Collaborator.
Laboratory of Developmental Neurogenetics
Rockefeller University
1230 York Avenue
New York, N.Y. 10021.
gaul@mail.rockefeller.edu. (212) 327 7621

Dr. Ramarathnam Venkatesan, Internship Supervisor.
Cryptographer
Microsoft Research
Cryptography and Anti-Piracy Research Group
One Microsoft Way, Redmond, WA 98052.
venkie@microsoft.com. (425) 703-9537