

Sourav Chatterji

1980 Barrymore Common
Apt O
Fremont, CA 94538
USA

Phone: +1 (510) 8960083
souravc@gmail.com

Education

Ph.D. in Computer Science, August 2006
(with *Designated Emphasis in Computational and Genomic Biology*)
University of California, Berkeley, CA, USA
Advisor : Professor Lior Pachter

B.Tech in Computer Science and Engineering, May 2001
Indian Institute of Technology, Kanpur, UP, India

Experience

Postdoctoral Researcher, August 2006-November 2008
Genome Center
University of California, Davis, CA, USA

Member of Technical Staff, December 2008-present
Chai Labs, Inc.
Suite A20, 148 Castro Street
Mountain View, CA 94041

Publications

Sourav Chatterji, Ichitaro Yamazaki, Zhaojun Bai and Jonathan Eisen, CompostBin: A DNA composition-based algorithm for binning environmental shotgun reads, Proceedings of the Twelfth Annual International Conference on Computational Molecular Biology, (2008) pp 17-28.

Drosophila Comparative Genome Sequencing and Analysis Consortium, Evolution of genes and genomes in the context of the Drosophila phylogeny, *Nature*, 450 (2007), p 203–218.

Sourav Chatterji and Lior Pachter, Patterns of gene duplication and intron loss in the ENCODE regions suggest a confounding factor, *Genomics*, 90 (2007), p44-48.

Sourav Chatterji and Lior Pachter, GeneMapper: Reference based gene annotation, *Genome Biology*, (2006), 7(4):R29.

Sourav Chatterji and Lior Pachter, Large multiple organism genefinding by collapsed Gibbs sampling, *Journal of Computational Biology*, 12 (2005), p 599–608.

Chicken Genome Sequencing Consortium, Sequence and comparative analysis of chicken genome provide unique perspectives into vertebrate evolution, *Nature* vol 433 (2005), p

695-716.

Sourav Chatterji and Lior Pachter. Multiple Organism Gene Finding by collapsed Gibbs sampling , Proceedings of the Eighth Annual International Conference on Computational Molecular Biology, (2004) p 187-193.

Rat Genome Sequencing Consortium, Genome sequence of the Brown Norway rat yields insights into mammalian evolution, Nature vol. 428 (2004), p 493-521.

Adam M. Breier, Sourav Chatterji, and Nicholas R. Cozzarelli. Prediction of *Saccharomyces cerevisiae* replication origins , Genome Biology, (2004) 5(4):R22.

Sourav Chatterji, Manikandan Narayanan, Jason Duell and Leonid Oliker. Performance Evaluation of Two Emerging Media Processors: VIRAM and Imagine, 17th International Parallel and Distributed Processing Symposium, (2003) p 229.

Sourav Chatterji, Sai Surya Kiran Evani, Sumit Ganguly and Mahesh Datt Yemmanuru. On the Complexity (Non-Approximability) of Query Optimization, Proceedings of the Twenty-first ACM Symposium on Principles of Database Systems (2002) p 282-292.

Awards and Honors

- DOE Travel Fellowship to RECOMB 2008.
- UC Regent's Fellowship for graduate studies in UC Berkeley, 2001.
- President's Gold Medal for the best academic performance in the graduating class in all disciplines of the undergraduate program at IIT Kanpur, 2001.
- General Proficiency Silver Medal award for the best academic performance in Computer Science and Engineering at IIT Kanpur, 2001.
- Proficiency award for the best final year thesis in Computer Science and Engineering at IIT Kanpur, 2001.
- Dr. V. Rajaraman Scholarship for excellent academic performance at IIT Kanpur (for the academic year 1999-2000).
- Ranked 2nd in I.I.T Joint Entrance Examination held nationwide in India, 1997.
- Ranked 1st in West Bengal Joint Entrance Examination for selection to engineering colleges in the state of West Bengal, 1997.
- Indian National Mathematics Olympiad (INMO) awardee, 1995. INMO is a nationwide examination held by National Board of Higher Mathematics in India.
- National Talent Search Scholarship awarded by the Government of India, 1995-2001.

Software Systems Developed

GeneMapper : A reference based gene prediction system.

CompostBin : A DNA-composition based metagenomic binning program.

Software for Quality Control in Single-Cell Genomics.

ZORRO : Probabilistic Masking/Trimming program to remove poorly aligned regions during phylogenetic inference.

Computer Skills

Programming Languages : C/C++, Java, (Bio)Perl, Python.

Scientific Computing Tools : Matlab, R.

Operating Systems : Windows (2000, XP, Vista), Mac OS X, Linux.

Professional Service

Invited Talk, Cadence Berkeley Labs, Berkeley, February 2007.

Invited Talk, Genome Center Colloquium, UC Davis, June 2007.

Invited Talk, CSB 2007, San Diego, August 2007.

Invited Talk, Single Cell Genomics Workshop, Bigelow Labs, September 2007.

Program Committee, RECOMB Comparative Genomics, 2007.

Reviewer, RECOMB 2007, PNAS, PLOS Biology, PLOS Computational Biology, PLOS One, Journal of Computational Biology, Nucleic Acids Research, Genome Biology, IEEE/ACM Transactions on Computational Biology and Bioinformatics(TCBB).