

Siavash Mirarab

CONTACT INFORMATION	Department of Computer Science University of Texas at Austin 2317 Speedway, 2.302, Austin, TX 78712	mobile: 512 698 2967 e-mail: smirarab@cs.utexas.edu website: http://www.cs.utexas.edu/~smirarab google scholar: http://goo.gl/gEswQ5
RESEARCH INTERESTS	Computational biology and Bioinformatics, Large-scale phylogenetics, Phylogenomics, Multiple sequence alignment, Metagenomics, Parallel computing, Software engineering	
EDUCATION	University of Texas at Austin , Austin, Texas, United States Department of Computer Science <i>Ph.D</i> January 2011 – present <ul style="list-style-type: none">• GPA: 3.97• Advisors: Prof. Tandy Warnow, Prof. Keshav Pingali• Thesis: Novel scalable approaches for multiple sequence alignment and phylogenetic reconstruction University of Waterloo , Waterloo, Ontario, Canada Department of Electrical and Computer Engineering <i>Master of Science</i> September 2006 – September 2008 <ul style="list-style-type: none">• Advisors: Prof. Ladan Tahvildari• Thesis: A Bayesian Framework for Software Regression Testing University of Tehran , Tehran, Iran Department of Electrical and Computer Engineering <i>Bachelor of Science</i> September 2001 – December 2005 <ul style="list-style-type: none">• Computer engineering, specialized in software engineering.	
MAJOR AWARDS	Howard Hughes Medical Institute International Student Fellowship. 2012 - present National Science and Engineering Research Council of Canada (NSERC), Canada Graduate Scholarship (CGS-D) converted to PGS-D outside Canada. 2011- 2012 University of Texas at Austin, MCD fellowship (I declined to take CGS instead)	
OTHER AWARDS	University of Texas at Austin, College of Natural Sci., Dean's Excellence Award, 2011 University of Texas at Austin, Graduate Dean's Prestigious Fellowship Supplement Award, 2011, 2012, 2013, 2014 Travel awards from International Society for Computational Biology (ISCB) for ECCB, ISMB (2014, 2015), National Science Foundation (NSF) for RECOMB (2014), National Library of Medicine/National Institutes of Health for PSB (2011, 2102), Institute for Pure and Applied Mathematics (IPAM) for Mathematical and Computational Approaches in High-Throughput Genomic (2011), and University of Texas at Austin, Dept. of Computer Science, for conference presentations (2011, 2012, 2013, 2014)	

1. **Siavash Mirarab** and Tandy Warnow. ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. *Bioinformatics*, 31(12):i44–i52, 2015.
2. Nam Nguyen, **Siavash Mirarab**, Keerthana Kumar, and Tandy Warnow. Ultra-large alignments using phylogeny-aware profiles. *Genome Biology*, 16(1):124, 2015.
3. Md Shamsuzzoha Bayzid, **Siavash Mirarab**, Bastien Boussau, and Tandy Warnow. Weighted Statistical Binning: Enabling Statistically Consistent Genome-Scale Phylogenetic Analyses. *PLoS ONE*, 10(6):e0129183, 2015.
4. (Erich D Jarvis, **Siavash Mirarab**)*, 100 other authors, Edward L Braun, Tandy Warnow, Wang Jun, M Thomas P Gilbert, and Guojie Zhang. Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science*, 346(6215):1320–1331, 2014.
5. **Siavash Mirarab**, Md. Shamsuzzoha Bayzid, Bastien Boussau, and Tandy Warnow. Statistical binning enables an accurate coalescent-based estimation of the avian tree. *Science*, 346(6215), 2014.
6. (Norman J. Wickett, **Siavash Mirarab**)*, Nam Nguyen, Tandy Warnow, Eric Carpenter, Naim Matasci, Saravanaraj Ayyampalayam, Michael Barker, J. Gordon Burleigh, Matthew A. Gitzendanner, Brad R. Ruhfel, Eric Wafula, Joshua P. Der, Sean W. Graham, Sarah Mathews, Michael Melkonian, Douglas E. Soltish, Pamela S. Soltish, Nicholas W. Miles, Carl J. Rothfels, Lisa Pokorný, A. Jonathan Shaw, Lisa DeGironimo, Dennis W. Stevenson, Barbara Surek, Juan Carlos Villarreal, Béatrice Rourev, Hervé Philippe, Claude W. dePamphilis, Tao Chen, Michael K. Deyholos, Regina S. Baucom, Toni M. Kutchan, Megan M. Augustin, Jun Wang, Yong Zhang, Zhijian Tian, Zhixiang Yan, Xiaolei Wu, Xiao Sun, Gane Ka-Shu Wong, and James Leebens-Mack. Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proceedings of the National Academy of Sciences (PNAS)*, 111(45):E4859–4868, 2014.
7. **Siavash Mirarab**, Md. Shamsuzzoha Bayzid, and Tandy Warnow. Evaluating summary methods for multi-locus species tree estimation in the presence of incomplete lineage sorting. *Systematic Biology*, doi: 10.1093/sysbio/syu063, 2014.
8. **Siavash Mirarab**, Nam Nguyen, Sheng Guo, Li-San Wang, Junhyong Kim, and Tandy Warnow. PASTA: Ultra-Large Multiple Sequence Alignment for Nucleotide and Amino-Acid Sequences. *Journal of Computational Biology*, 22(05):377–386, 2015.

* co-first authorship

9. **Siavash Mirarab**, Nam Nguyen, and Tandy Warnow. PASTA: ultra-large multiple sequence alignment. In *Proceedings of the 18th International Conference on Research in Computational Molecular Biology (RECOMB)*, pages 177–191, 2014.
10. **Siavash Mirarab**, Rezwana Reaz, Md. Shamsuzzoha Bayzid, Théo Zimmermann, M Shel Swenson, and Tandy Warnow. ASTRAL: Genome-Scale Coalescent-Based Species Tree. *Bioinformatics - ECCB Special Issue*, 30(17):i541–i548, 2014.
11. Nam Nguyen, **Siavash Mirarab**, Bo Liu, Mihai Pop, and Tandy Warnow. TIPP: Taxonomic identification and phylogenetic profiling. *Bioinformatics*, doi: 10.1093/bioinformatics/btu721, 2014.
12. Théo Zimmermann, **Siavash Mirarab**, and Tandy Warnow. BBICA: improving the scalability of *BEAST using random binning. *BMC Genomics*, 15(Suppl 6):S11, 2014.
13. Arlin Stoltzfus, Hilmar Lapp, Naim Matasci, Helena Deus, Brian Sidlauskas, Christian M Zmasek, Gaurav Vaidya, Enrico Pontelli, Karen Cranston, Rutger Vos, Campbell O Webb, Luke J Harmon, Megan Pirrung, Brian O’Meara, Matthew W Pennell, **Siavash Mirarab**, Michael S Rosenberg, James P Balhoff, Holly M Bik, Tracy A Heath, Peter E Midford, Joseph W Brown, Emily Jane McTavish, Jeet Sukumaran, Mark Westneat, Michael E Alfaro, Aaron Steele, and Greg Jordan. Phylotastic! Making tree-of-life knowledge accessible, reusable and convenient. *BMC bioinformatics*, 14(1):158, 2013.
14. Md. Shamsuzzoha Bayzid, **Siavash Mirarab**, and Tandy Warnow. Inferring optimal species trees under gene duplication and loss. In *Proceedings of the Pacific Symposium on Biocomputing (PSB)*, pages 250–61, 2013.
15. Nam Nguyen, **Siavash Mirarab**, and Tandy Warnow. MRL and SuperFine+ MRL: new supertree methods. *Algorithms for Molecular Biology*, 7(1):3, 2012.
16. **Siavash Mirarab**, Nam Nguyen, and Tandy Warnow. SEPP: SATé-Enabled Phylogenetic Placement. In *Proceedings of the Pacific Symposium on Biocomputing (PSB)*, pages 247–58, 2012.
17. **Siavash Mirarab**, Soroush Akhlaghi, and Ladan Tahvildari. Size-constrained regression test case selection using multicriteria optimization. *IEEE Transactions on Software Engineering*, 38(4):936–956, 2012.
18. **Siavash Mirarab** and Tandy Warnow. FastSP: linear time calculation of alignment accuracy. *Bioinformatics*, 27:3250–8, 2011.
19. Hyunsook Do, **Siavash Mirarab**, Ladan Tahvildari, and Gregg Rothermel. The effects of time constraints on test case prioritization: A series of controlled experiments. *IEEE Transactions on Software Engineering*, 36(5):593–617, 2010.

20. **Siavash Mirarab** and Ladan Tahvildari. An empirical study on bayesian network-based approach for test case prioritization. In *Proceedings of the 1st International Conference on Software Testing, Verification, and Validation (ICST)*, pages 278–287. IEEE, 2008.
21. **Siavash Mirarab**, Afshar Ganjali, Ladan Tahvildari, Shimin Li, Weining Liu, and Mike Morrissey. A requirement-based software testing framework: An industrial practice. In *Proceedings of the 24th International Conference on Software Maintenance (ICSM)*, pages 452–455. IEEE, 2008.
22. Hyunsook Do, **Siavash Mirarab**, Ladan Tahvildari, and Gregg Rothermel. An empirical study of the effect of time constraints on the cost-benefits of regression testing. In *Proceedings of the 16th ACM SIGSOFT International Symposium on Foundations of Software Engineering (FSE)*, pages 71–82. ACM, 2008.
23. Mehdi Amoui, Mazeiar Salehie, **Siavash Mirarab**, and Ladan Tahvildari. Adaptive action selection in autonomic software using reinforcement learning. In *Proceedings of the Fourth International Conference on Autonomic and Autonomous Systems (ICAS)*, pages 175–181. IEEE, 2008.
24. **Siavash Mirarab** and Ladan Tahvildari. A prioritization approach for software test cases based on bayesian networks. In *Fundamental Approaches to Software Engineering (FASE)*, pages 276–290. Springer, 2007.
25. **Siavash Mirarab**, Alaa Hassouna, and Ladan Tahvildari. Using bayesian belief networks to predict change propagation in software systems. In *Proceedings of the 15th International Conference on Program Comprehension (ICPC)*, pages 177–188. IEEE, 2007.
26. Mehdi Amoui, **Siavash Mirarab**, Sepand Ansari, and Caro Lucas. A genetic algorithm approach to design evolution using design pattern transformation. *Journal of Information Technology and Intelligent Computing*, 1(2):235–244, 2006.

SOFTWARE
(PUBLICLY
AVAILABLE)

ASTRAL (<https://github.com/smirarab/ASTRAL>)

Coalescent-based species tree estimation from gene trees

PASTA (<https://github.com/smirarab/pasta>)

Multiple sequence alignment of ultra-large datasets (up to a million sequences)

SEPP (<https://github.com/smirarab/sepp>)

Phylogenetic placement of metagenomic reads using families of HMMs

TIPP (<https://github.com/smirarab/sepp/blob/master/README.TIPP.md>)

Taxonomic profiling of metagenomic samples using phylogenetic placement

UPP (<https://github.com/smirarab/sepp/blob/master/README.UPP.md>)

Multiple sequences alignment using families of HMMs for ultra-large datasets

FastSP (<https://github.com/smirarab/FastSP>)
 Fast comparison of very large multiple sequence alignments

See <https://github.com/smirarab/?tab=repositories> for more.

PROFESSIONAL EXPERIENCE	International Business Machines (IBM) , Vancouver, British Columbia, Canada	
	<i>IT Specialist</i>	September 2008 – December 2010
	Worked on a software platform for wireless pace-maker devices	
	Research In Motion (RIM) , Waterloo, Ontario, Canada	
	<i>Research Associate (part-time)</i>	September 2007 – September 2008
	Cisco Systems, ISBU , San Jose, California, US	
	<i>College Intern as Software Engineer</i>	May 2007 – September 2007
TEACHING EXPERIENCE	Teaching Assistance:	
	University of Waterloo, Software Testing and Quality Assurance	Winter 2007
	University of Tehran, Scientific & Technical Communication	Fall 2005
	Workshops, Tutorials, and Short Courses:	
	Ann Arbor, Michigan, Phylogenomics Symposium and Software School	2015
	UT Austin, Summer school on Phylogenomics and Metagenomics	Summer 2014
	Evolution meetings (NC), Software School on ASTRAL and PASTA	Summer 2014
	UT Austin, Workshop on New Methods for Phylogeny and Alignment Estimation, Tutorials on SEPP and SATe	Winter 2013
	Guest Lecturer:	
	University of Texas at Austin (Biology) Computation Biology	2014
	University of Texas at Austin (CCBB) Introduction to Biological Computing	2014
	University of Texas at Austin (Computer Science) Algorithm of Comp. Biology	2013
HIGHLIGHTS TALKS	RECOMB, Warsaw, Poland, “Statistical binning enables an accurate ...”	2015
	IPAM, UCLA, “Ultra-large multiple sequence alignments”.	2015
PROFESSIONAL SERVICES	Program Committee: ISMB/ECCB	
	2015	
	Paper Review: Trans. on Comp. Bio. and Bioinform. (TCBB) , Bioinformatics, ISMB, RECOMB-CG, Theoretical Population Genetics, Trans. on Software Engi- neering (TSE), Intern. Conf. on Program Comprehension	
	2007-2015	
	Mentorship: Undergrad Research Assistants at University of Texas at Austin	2013
	Mentorship: Undergrad Research Assistants at University of Waterloo	2008
NATIONALITY	Citizen of Canada and Iran	