List of publications

Siavash Mirarab

October 5, 2018

1. Maryam Rabiee and **Siavash Mirarab**. INSTRAL: Discordance-aware Phylogenetic Placement using Quartet Scores. *bioRxiv*, 432906, 2018.

2. Niema Moshiri, Manon Ragonnet-cronin, Joel O Wertheim, and **Siavash Mirarab**. FAVITES : simultaneous simulation of transmission networks , phylogenetic trees , and sequences. *BioRxiv*, 2018.

3. Shahab Sarmashghi, Kristine Bohmann, M. Thomas P. Gilbert, Vineet Bafna, and **Siavash Mirarab**. Assembly-free and alignment-free sample identification using genome skims. *bioRxiv (abstract appeared at RECOMB 2018)*, 230409, 2018.

4. Daniel McDonald and American Gut Consortium. American Gut: an Open Platform for Citizen Science Microbiome Research. *mSystems*, 3(3), may 2018.

5. Wolf L. Eiserhardt, Alexandre Antonelli, Dominic J. Bennett, Laura R. Botigué, J. Gordon Burleigh, Steven Dodsworth, Brian J. Enquist, Félix Forest, Jan T. Kim, Alexey M. Kozlov, Ilia J. Leitch, Brian S. Maitner, **Siavash Mirarab**, William H. Piel, Oscar A. Pérez-Escobar, Lisa Pokorny, Carsten Rahbek, Brody Sandel, Stephen A. Smith, Alexandros Stamatakis, Rutger A. Vos, Tandy Warnow, and William J. Baker. A roadmap for global synthesis of the plant tree of life. *American Journal of Botany*, 105(3):614–622, 2018.

6. Stefan Janssen, Daniel McDonald, Antonio Gonzalez, Jose A. Navas-Molina, Lingjing Jiang, Zhenjiang Zech Xu, Kevin Winker, Deborah M. Kado, Eric Orwoll, Mark Manary, **Siavash Mirarab**, and Rob Knight. Phylogenetic Placement of Exact Amplicon Sequences Improves Associations with Clinical Information. *mSystems*, 3(3):e00021–18, 2018.

7. Chao Zhang, Maryam Rabiee, Erfan Sayyari, and **Siavash Mirarab**. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics*, 19(S6):153, 2018.

8. Uyen Mai and **Siavash Mirarab**. TreeShrink: fast and accurate detection of outlier long branches in collections of phylogenetic trees. *BMC Genomics*, 19(S5):272, 2018.

9. Erfan Sayyari and **Siavash Mirarab**. Testing for Polytomies in Phylogenetic Species Trees Using Quartet Frequencies. *Genes*, 9(3):132, 2018.

10. Erfan Sayyari, James B. Whitfield, and **Siavash Mirarab**. DiscoVista: Interpretable visualizations of gene tree discordance. *Molecular Phylogenetics and Evolution*, 122:110–115, 2018.

11. Ali Akbari, Joseph J Vitti, Arya Iranmehr, Mehrdad Bakhtiari, Pardis C Sabeti, **Siavash Mirarab**, and Vineet Bafna. Identifying the favored mutation in a positive selective sweep. *Nature Methods*, 15(4):279–282, 2018.

12. Niema Moshiri and **Siavash Mirarab**. A Two-State Model of Tree Evolution and Its Applications to Alu Retrotransposition. *Systematic Biology*, 67(3):475–489, 2018.

13. Erfan Sayyari, James B Whitfield, and **Siavash Mirarab**. Fragmentary Gene Sequences Negatively Impact Gene Tree and Species Tree Reconstruction. *Molecular Biology and Evolution*, 34(12):3279–3291, 2017.

14. Luke R. Thompson, Jon G. Sanders, Daniel McDonald, Amnon Amir, Joshua Ladau, Kenneth J. Locey, Robert J. Prill, Anupriya Tripathi, Sean M. Gibbons, Gail Ackermann, Jose A. Navas-Molina, Stefan Janssen, Evguenia Kopylova, Yoshiki Vázquez-Baeza, Antonio González, James T. Morton, **Siavash Mirarab**, Zhenjiang Zech Xu, Lingjing Jiang, Mohamed F. Haroon, Jad Kanbar, Qiyun Zhu, Se Jin Song, Tomasz Kosciolek, Nicholas A. Bokulich, Joshua Lefler, Colin J. Brislawn, Gregory Humphrey, Sarah M. Owens, Jarrad Hampton-Marcell, Donna Berg-Lyons, Valerie McKenzie, Noah Fierer, Jed A. Fuhrman, Aaron Clauset, Rick L. Stevens, Ashley Shade, Katherine S. Pollard, Kelly D. Goodwin, Janet K. Jansson, Jack A. Gilbert, Rob Knight, and The Earth Microbiome Project Consortium. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature*, 551(7681), 2017.

15. Uyen Mai and **Siavash Mirarab**. TreeShrink: Efficient Detection of Outlier Tree Leaves. In Joao Meidanis and Luay Nakhleh, editors, *Lecture Notes in Computer Science*, volume 10562 LNBI, 116–140. 2017.

16. Chao Zhang, Erfan Sayyari, and **Siavash Mirarab**. ASTRAL-III: Increased Scalability and Impacts of Contracting Low Support Branches. In Joao Meidanis and Luay Nakhleh, editors, *Lecture Notes in Computer Science*, volume 10562 LNBI, 53–75. 2017.

17. Uyen Mai, Erfan Sayyari, and **Siavash Mirarab**. Minimum variance rooting of phylogenetic trees and implications for species tree reconstruction. *PLOS ONE*, 12(8):e0182238, 2017.

18. Shubhanshu Shekhar, Sebastien Roch, and **Siavash Mirarab**. Species tree estimation using ASTRAL: how many genes are enough? *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, PP(99):1–1, 2017.

19. **Siavash Mirarab**. Phylogenomics: Constrained gene tree inference. *Nature Ecology & Evolution*, 1:0056, 2017.

20. Erfan Sayyari and **Siavash Mirarab**. Anchoring quartet-based phylogenetic distances and applications to species tree reconstruction. *BMC Genomics*, 17(S10):101–113, 2016.

21. Erfan Sayyari and **Siavash Mirarab**. Fast Coalescent-Based Computation of Local Branch Support from Quartet Frequencies. *Molecular Biology and Evolution*, 33(7):1654–1668, 2016.

22. Nam-phuong Nguyen, Michael Nute, **Siavash Mirarab**, and Tandy Warnow. HIPPI: highly accurate protein family classification with ensembles of HMMs. *BMC Genomics*, 17(S10):765, 2016.

23. James E Tarver, Mario dos Reis, **Siavash Mirarab**, Raymond J Moran, Sean Parker, Joseph E. O'Reilly, Benjamin L King, Mary J. O'Connell, Robert J Asher, Tandy Warnow, Kevin J Peterson, Philip C.J. Donoghue, and Davide Pisani. The Interrelationships of Placental Mammals and the Limits of Phylogenetic Inference. *Genome Biology and Evolution*, 8(2):330–344, 2016.

24. **Siavash Mirarab**, Md. Shamsuzzoha Bayzid, Bastien Boussau, and Tandy Warnow. Response to Comment on "Statistical binning enables an accurate coalescent-based estimation of the avian tree". *Science*, 350(6257):171, 2015.

25. Joel Cracraft, Peter Houde, Simon Y W Ho, David P Mindell, Jon Fjeldså, Bent Lindow, Scott V Edwards, Carsten Rahbek, **Siavash Mirarab**, Tandy Warnow, M Thomas P Gilbert, Guojie Zhang, Edward L Braun, and Erich D Jarvis. Response to Comment on "Whole-genome analyses resolve early branches in the tree of life of modern birds". *Science*, 349(6255):1460, 2015.

26. Ruth Davidson, Pranjal Vachaspati, **Siavash Mirarab**, and Tandy Warnow. Phylogenomic species tree estimation in the presence of incomplete lineage sorting and horizontal gene transfer. *BMC Genomics*, 16(Suppl 10):S1, 2015.

27. Jed Chou, Ashu Gupta, Shashank Yaduvanshi, Ruth Davidson, Mike Nute, **Siavash Mirarab**, and Tandy Warnow. A comparative study of SVDquartets and other coalescent-based species tree estimation methods. *BMC Genomics*, 16(Suppl 10):S2, 2015.

28. **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.

29. Nam Nguyen, **Siavash Mirarab**, Keerthana Kumar, and Tandy Warnow. Ultra-large alignments using phylogeny-aware profiles. *Genome Biology*, 16(1):124, 2015.

30. 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.

31. Erich D Jarvis, **Siavash Mirarab**, Andre J Aberer, Bo Li, Peter Houde, Cai Li, Simon Y W Ho, Brant C Faircloth, Benoit Nabholz, and Jason T Howard. Phylogenomic analyses data of the avian phylogenomics project. *GigaScience*, 4(1):4, 2015.

32. Erich D Jarvis*, **Siavash Mirarab***, 100 other authors, Edward L Braun, Tandy Warnow, Wang Jun, M Thomas P Gilbert, and Guojie Zhang. Whole-genome analy-

---

*co-first author

ses resolve early branches in the tree of life of modern birds. *Science*, 346(6215):1320–1331, 2014.

33. **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):1250463, 2014.

34. Naim Matasci, Ling-Hong Hung, Zhixiang Yan, Eric J Carpenter, Norman J Wickett, **Siavash Mirarab**, Nam Nguyen, Tandy Warnow, Saravanaraj Ayyampalayam, and Michael Barker. Data access for the 1,000 Plants (1KP) project. *GigaScience*, 3(1):17, 2014.

35. 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 Pokorny, 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.

36. **Siavash Mirarab**, Md. Shamsuzzoha Bayzid, and Tandy Warnow. Evaluating Summary Methods for Multilocus Species Tree Estimation in the Presence of Incomplete Lineage Sorting. *Systematic Biology*, 65(3):366–380, 2014.

37. **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.

38. **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)*, 177–191, 2014.

[†]co-first author

39. **Siavash Mirarab**, Rezwana Reaz, Md. Shamsuzzoha Bayzid, Théo Zimmermann, M Shel Swenson, and Tandy Warnow. ASTRAL: Genome-Scale Coalescent-Based Species Tree. *Bioinformatics (ECCB)*, 30(17):i541–i548, 2014.

40. Nam Nguyen, **Siavash Mirarab**, Bo Liu, Mihai Pop, and Tandy Warnow. TIPP: Taxonomic Identification and Phylogenetic Profiling. *Bioinformatics*, 30(24):3548–3555, 2014.

41. Théo Zimmermann, **Siavash Mirarab**, and Tandy Warnow. BBCA: improving the scalability of *BEAST using random binning. *BMC Genomics*, 15(Suppl 6):S11, 2014.

42. 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.

43. 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)*, 250–61, 2013.

44. Nam Nguyen, **Siavash Mirarab**, and Tandy Warnow. MRL and SuperFine+ MRL: new supertree methods. *Algorithms for Molecular Biology*, 7(1):3, 2012.

45. **Siavash Mirarab**, Nam Nguyen, and Tandy Warnow. SEPP: SATé-Enabled Phylogenetic Placement. In *Proceedings of the Pacific Symposium on Biocomputing (PSB)*, 247–58, 2012.

46. **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.

47. **Siavash Mirarab** and Tandy Warnow. FastSP: linear time calculation of alignment accuracy. *Bioinformatics*, 27:3250–8, 2011.

48. 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.

49. **Siavash Mirarab** and Ladan Tahvildari. An empirical study on bayesian network-based approach for test case prioritization. In *Proceedings of ICST*, 278–287. IEEE, 2008.

50. **Siavash Mirarab**, Afshar Ganjali, Ladan Tahvildari, Shimin Li, Weining Liu, and Mike Morrissey. A requirement-based software testing framework: An industrial practice. In *Proceedings ICSM*, 452–455. IEEE, 2008.

51. 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 FSE*, 71–82. ACM, 2008.

52. Mehdi Amoui, Mazeiar Salehie, **Siavash Mirarab**, and Ladan Tahvildari. Adaptive action selection in autonomic software using reinforcement learning. In *Proceedings of ICAS*, 175–181. IEEE, 2008.

53. **Siavash Mirarab** and Ladan Tahvildari. A prioritization approach for software test cases based on bayesian networks. In *Proceedings of FASE*, 276–290, 2007.

54. **Siavash Mirarab**, Alaa Hassouna, and Ladan Tahvildari. Using bayesian belief networks to predict change propagation in software systems. In *Proceedings of ICPC*, 177–188. IEEE, 2007.

55. 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.