List of publications

Siavash Mirarab

December 5, 2023

1. Metin Balaban, Yueyu Jiang, Qiyun Zhu, Daniel McDonald, Rob Knight, and **Siavash Mirarab**. Generation of accurate, expandable phylogenomic trees with uDance. *Nature Biotechnology*, Online, July 2023.

2. Daniel McDonald, Yueyu Jiang, Metin Balaban, Kalen Cantrell, Qiyun Zhu, Antonio Gonzalez, James T. Morton, Giorgia Nicolaou, Donovan H. Parks, Søren M. Karst, Mads Albertsen, Philip Hugenholtz, Todd DeSantis, Se Jin Song, Andrew Bartko, Aki S. Havulinna, Pekka Jousilahti, Susan Cheng, Michael Inouye, Teemu Niiranen, Mohit Jain, Veikko Salomaa, Leo Lahti, **Siavash Mirarab**, and Rob Knight. Greengenes2 unifies microbial data in a single reference tree. *Nature Biotechnology*, Online, July 2023.

3. Chao Zhang, Rasmus Nielsen, and **Siavash Mirarab**. CASTER: Direct species tree inference from whole-genome alignments. *bioRxiv*, prepress, 2023.

4. Shayesteh Arasti, Puoya Tabaghi, Yasamin Tabatabaee, and Siavash Mirarab. Optimal Tree Metric Matching Enables Phylogenomic Branch Length Reconciliation. *bioRxiv*, 2023.11.13.566962, January 2023.

5. Ali Osman Berk Şapcı, Eleonora Rachtman, and **Siavash Mirarab**. CONSULT-II: Taxonomic Identification Using Locality Sensitive Hashing. In Katharina Jahn and Tomáš Vinař, editors, *Comparative Genomics*, 196–214, Cham, 2023. Springer Nature Switzerland.

6. Shayesteh Arasti and **Siavash Mirarab**. Optimal Subtree Prune and Regraft for Quartet Score in Sub-Quadratic Time. In Djamal Belazzougui and Aïda Ouangraoua, editors, *23rd International Workshop on Algorithms in Bioinformatics (WABI 2023)*, volume 273 of *Leibniz International Proceedings in Informatics (LIPIcs)*, 4:1–4:20,

Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. ISSN: 1868-8969.

7. Uyen Mai, Eduardo Charvel, and **Siavash Mirarab**. Expectation-Maximization enables Phylogenetic Dating under a Categorical Rate Model. In *RECOMB*, 2023.

8. Yasamin Tabatabaee, Chao Zhang, Tandy Warnow, and **Siavash Mirarab**. Phylogenomic branch length estimation using quartets. *Bioinformatics*, 39(Supplement_1):i185–i193, June 2023.

9. Afif Elghraoui, **Siavash Mirarab**, Krister M Swenson, and Faramarz Valafar. Evaluating impacts of syntenic block detection strategies on rearrangement phylogeny using Mycobacterium tuberculosi isolates. *Bioinformatics*, 39(1):btad024, January 2023.

10. Chao Zhang, Andrey V. Bzikadze, Yana Safonova, and **Siavash Mirarab**. A scalable model for simulating multi-round antibody evolution and benchmarking of clonal tree reconstruction methods. *Frontiers in Immunology*, 13:1014439, 2022.

11. Chao Zhang and **Siavash Mirarab**. Weighting by Gene Tree Uncertainty Improves Accuracy of Quartet-based Species Trees. *Molecular Biology and Evolution*, 39(12):msac215, October 2022.

12. Chao Zhang and **Siavash Mirarab**. ASTRAL-Pro 2: ultrafast species tree reconstruction from multi-copy gene family trees. *Bioinformatics*, 38(21):4949–4950, September 2022.

13. Metin Balaban, Nishat Anjum Bristy, Ahnaf Faisal, Md. Shamsuzzoha Bayzid, and **Siavash Mirarab**. Genome-wide alignment-free phylogenetic distance estimation under a no strand-bias model. *Bioinformatics Advances*, 2(1):2021.11.10.468111, January 2022.

14. Maryam Rabiee and **Siavash Mirarab**. QuCo: quartet-based co-estimation of species trees and gene trees. *Bioinformatics*, 38(Supplement_1):i413–i421, June 2022.

15. Yueyu Jiang, Puoya Tabaghi, and **Siavash Mirarab**. Learning Hyperbolic Embedding for Phylogenetic Tree Placement and Updates. *Biology*, 11(9):1256, 2022. ISBN: 2079-7737.

16. Navid Bin Hasan, Metin Balaban, Avijit Biswas, Md. Shamsuzzoha Bayzid, and **Siavash Mirarab**. Distance-Based Phylogenetic Placement with Statistical Support. *Biology*, 11(8):1212, August 2022. ISBN: 2079-7737.

17. Eleonora Rachtman, Shahab Sarmashghi, Vineet Bafna, and **Siavash Mirarab**. Quantifying the uncertainty of assembly-free genome-wide distance estimates and phylogenetic relationships using subsampling. *Cell Systems*, 13(10):817–829.e3, October 2022.

18. Uyen Mai and **Siavash Mirarab**. Completing gene trees without species trees in sub-quadratic time. *Bioinformatics*, 38(6):1532–1541, March 2022.

19. Metin Balaban, Yueyu Jiang, Daniel Roush, Qiyun Zhu, and **Siavash Mirarab**. Fast and accurate distance-based phylogenetic placement using divide and conquer. *Molecular Ecology Resources*, 22(3):1213–1227, April 2022.

20. Shahab Sarmashghi, Metin Balaban, Eleonora Rachtman, Behrouz Touri, **Siavash Mirarab**, and Vineet Bafna. Estimating repeat spectra and genome length from low-coverage genome skims with RESPECT. *PLOS Computational Biology*, 17(11):e1009449, November 2021.

21. **Siavash Mirarab**, Luay Nakhleh, and Tandy Warnow. Multispecies Coalescent: Theory and Applications in Phylogenetics. *Annual Review of Ecology, Evolution, and Systematics*, 52(1):247–268, November 2021.

22. Eleonora Rachtman, Vineet Bafna, and **Siavash Mirarab**. CONSULT: accurate contamination removal using locality-sensitive hashing. *NAR Genomics and Bioinformatics*, 3(3):10.1101/2021.03.18.436035, June 2021.

23. Chao Zhang, Yiming Zhao, Edward Louis Braun, and **Siavash Mirarab**. TAPER: Pinpointing errors in multiple sequence alignments despite varying rates of evolution. *Methods in Ecology and Evolution*, 12(11):2145–2158, November 2021.

24. Niema Moshiri, Davey M Smith, and **Siavash Mirarab**. HIV Care Prioritization Using Phylogenetic Branch Length. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 86(5):626–637, April 2021.

25. Maryam Rabiee and **Siavash Mirarab**. SODA: multi-locus species delimitation using quartet frequencies. *Bioinformatics*, 36(24):5623–5631, April 2021.

26. Shaohong Feng, Josefin Stiller, Yuan Deng, Joel Armstrong, Qi Fang, Andrew Hart Reeve, Duo Xie, Guangji Chen, Chunxue Guo, Brant C. Faircloth, Bent Petersen, Zongji Wang, Qi Zhou, Mark Diekhans, Wanjun Chen, Sergio Andreu-Sánchez, Ashot Margaryan, Jason Travis Howard, Carole Parent, George Pacheco, Mikkel-Holger S. Sinding, Lara Puetz, Emily Cavill, Ângela M. Ribeiro, Leopold Eckhart, Jon Fjeldså, Peter A. Hosner, Robb T. Brumfield, Les Christidis, Mads F. Bertelsen, Thomas Sicheritz-Ponten, Dieter Thomas Tietze, Bruce C. Robertson, Gang Song, Gerald Borgia, Santiago Claramunt, Irby J. Lovette, Saul J. Cowen, Peter Njoroge, John Philip Dumbacher, Oliver A. Ryder, Jérôme Fuchs, Michael Bunce, David W. Burt, Joel Cracraft, Guanliang Meng, Shannon J. Hackett, Peter G. Ryan, Knud Andreas Jønsson, Ian G. Jamieson, Rute R. da Fonseca, Edward L. Braun, Peter Houde, **Siavash Mirarab**, Alexander Suh, Bengt Hansson, Suvi Ponnikas, Hanna Sigeman, Martin Stervander, Paul B. Frandsen, Henriette van der Zwan, Rencia van der Sluis, Carina Visser, Christopher N. Balakrishnan, Andrew G. Clark, John W. Fitzpatrick, Reed Bowman, Nancy Chen, Alison Cloutier, Timothy B. Sackton, Scott V. Edwards, Dustin J. Foote, Subir B. Shakya, Frederick H. Sheldon, Alain Vignal, André E. R. Soares, Beth Shapiro, Jacob González-Solís, Joan Ferrer-Obiol, Julio Rozas, Marta Riutort, Anna Tigano, Vicki Friesen, Love Dalén, Araxi O. Urrutia, Tamás Székely, Yang Liu, Michael G. Campana, André Corvelo, Robert C. Fleischer, Kim M. Rutherford, Neil J. Gemmell, Nicolas Dussex, Henrik Mouritsen, Nadine Thiele, Kira Delmore, Miriam Liedvogel, Andre Franke, Marc P. Hoeppner, Oliver Krone, Adam M. Fudickar, Borja Milá, Ellen D. Ketterson, Andrew Eric Fidler, Guillermo Friis, Ángela M. Parody-Merino, Phil F. Battley, Murray P. Cox, Nicholas Costa Barroso Lima, Francisco Prosdocimi, Thomas Lee Parchman, Barney A. Schlinger, Bette A. Loiselle, John G. Blake, Haw Chuan Lim, Lainy B. Day, Matthew J. Fuxjager, Maude W. Baldwin, Michael J. Braun, Morgan Wirthlin, Rebecca B. Dikow, T. Brandt Ryder, Glauco Camenisch, Lukas F. Keller, Jeffrey M. DaCosta, Mark E. Hauber, Matthew I. M. Louder, Christopher C. Witt, Jimmy A. McGuire, Joann Mudge, Libby C. Megna, Matthew D. Carling, Biao Wang, Scott A. Taylor, Glaucia Del-Rio, Alexandre Aleixo, Ana Tereza Ribeiro Vasconcelos, Claudio V. Mello, Jason T. Weir, David Haussler, Qiye Li, Huanming Yang, Jian Wang, Fumin Lei, Carsten Rahbek, M. Thomas P. Gilbert, Gary R. Graves, Erich D. Jarvis, Benedict Paten, and Guojie Zhang. Dense sampling of bird diversity increases power of comparative genomics. *Nature*, 587(7833):252–257,

November 2020.

27. Uyen Mai and **Siavash Mirarab**. Log Transformation Improves Dating of Phylogenies. *Molecular Biology and Evolution*, 38(3):1151–1167, March 2021.

28. Chao Zhang, Celine Scornavacca, Erin K Molloy, and **Siavash Mirarab**. ASTRAL-Pro: Quartet-Based Species-Tree Inference despite Paralogy. *Molecular Biology and Evolution*, 37(11):3292–3307, November 2020.

29. Kristine Bohmann, **Siavash Mirarab**, Vineet Bafna, and M. Thomas P. Gilbert. Beyond DNA barcoding: The unrealized potential of genome skim data in sample identification. *Molecular Ecology*, mec.15507, 2020.

30. Metin Balaban and **Siavash Mirarab**. Phylogenetic double placement of mixed samples. *Bioinformatics*, 36(Supplement_1):i335–i343, July 2020.

31. Ekin Tilic, Erfan Sayyari, Josefin Stiller, **Siavash Mirarab**, and Greg W Rouse. More is needed—Thousands of loci are required to elucidate the relationships of the 'flowers of the sea' (Sabellida, Annelida). *Molecular Phylogenetics and Evolution*, 151:106892, October 2020.

32. Francesco Asnicar, Andrew Maltez Thomas, Francesco Beghini, Claudia Mengoni, Serena Manara, Paolo Manghi, Qiyun Zhu, Mattia Bolzan, Fabio Cumbo, Uyen May, Jon G Sanders, Moreno Zolfo, Evguenia Kopylova, Edoardo Pasolli, Rob Knight, **Siavash Mirarab**, Curtis Huttenhower, and Nicola Segata. Precise phylogenetic analysis of microbial isolates and genomes from metagenomes using PhyloPhlAn 3.0. *Nature Communications*, 11(1):2500, December 2020. Publisher: Springer US ISBN: 0348556977.

33. Eleonora Rachtman, Metin Balaban, Vineet Bafna, and **Siavash Mirarab**. The impact of contaminants on the accuracy of genome skimming and the effectiveness of exclusion read filters. *Molecular Ecology Resources*, 20(3):1755–0998.13135, May 2020.

34. Metin Balaban, Shahab Sarmashghi, and **Siavash Mirarab**. APPLES: Scalable Distance-Based Phylogenetic Placement with or without Alignments. *Systematic Biology*, 69(3):566–578, May 2020.

35. Maryam Rabiee and **Siavash Mirarab**. INSTRAL: Discordance-Aware Phylogenetic Placement Using Quartet Scores. *Systematic Biology*, 69(2):384–391, August 2020.

36. Maryam Rabiee and **Siavash Mirarab**. Forcing external constraints on tree inference using ASTRAL. *BMC Genomics*, 21(S2):218, April 2020. Publisher: BMC Genomics.

37. Qiyun Zhu, Uyen Mai, Wayne Pfeiffer, Stefan Janssen, Francesco Asnicar, Jon G. Sanders, Pedro Belda-Ferre, Gabriel A. Al-Ghalith, Evguenia Kopylova, Daniel McDonald, Tomasz Kosciolek, John B Yin, Shi Huang, Nimaichand Salam, Jian-yu Jiao, Zijun Wu, Zhenjiang Z. Xu, Kalen Cantrell, Yimeng Yang, Erfan Sayyari, Maryam Rabiee, James T Morton, Sheila Podell, Dan Knights, Wen-jun Li, Curtis Huttenhower, Nicola Segata, Larry Smarr, **Siavash Mirarab**, and Rob Knight. Phylogenomics of 10,575 genomes reveals evolutionary proximity between domains Bacteria and Archaea. *Nature Communications*, 10(1):5477, December 2019.

38. James H Leebens-Mack, Michael S Barker, Eric J Carpenter, Michael K Deyholos, Matthew A Gitzendanner, Sean W Graham, Ivo Grosse, Zheng Li, Michael Melkonian, **Siavash Mirarab**, Martin Porsch, Marcel Quint, Stefan A Rensing, Douglas E Soltis, Pamela S Soltis, Dennis W Stevenson, Kristian K Ullrich, Norman J Wickett, Lisa DeGironimo, Patrick P Edger, Ingrid E Jordon-Thaden, Steve Joya, Tao Liu, Barbara Melkonian, Nicholas W Miles, Lisa Pokorny, Charlotte Quigley, Philip Thomas, Juan Carlos Villarreal, Megan M Augustin, Matthew D Barrett, Regina S Baucom, David J Beerling, Ruben Maximilian Benstein, Ed Biffin, Samuel F Brockington, Dylan O Burge, Jason N Burris, Kellie P Burris, Valérie Burtet-Sarramegna, Ana L Caicedo, Steven B Cannon, Zehra Çebi, Ying Chang, Caspar Chater, John M Cheeseman, Tao Chen, Neil D Clarke, Harmony Clayton, Sarah Covshoff, Barbara J Crandall-Stotler, Hugh Cross, Claude W DePamphilis, Joshua P Der, Ron Determann, Rowan C Dickson, Verónica S Di Stilio, Shona Ellis, Eva Fast, Nicole Feja, Katie J Field, Dmitry A Filatov, Patrick M Finnegan, Sandra K Floyd, Bruno Fogliani, Nicolás García, Gildas Gâteblé, Grant T Godden, Falicia (Qi Yun) Goh, Stephan Greiner, Alex Harkess, James Mike Heaney, Katherine E Helliwell, Karolina Heyduk, Julian M Hibberd, Richard G J Hodel, Peter M Hollingsworth, Marc T J Johnson, Ricarda Jost, Blake Joyce, Maxim V Kapralov, Elena Kazamia, Elizabeth A Kellogg, Marcus A Koch, Matt Von Konrat, Kálmán Könyves, Toni M Kutchan, Vivienne Lam, Anders Larsson, Andrew R Leitch, Roswitha Lentz, Fay-Wei Li, Andrew J Lowe, Martha Ludwig, Paul S Manos, Evgeny Mavrodiev, Melissa K McCormick, Michael McKain, Tracy McLellan, Joel R McNeal, Richard E Miller, Matthew N Nelson, Yanhui Peng, Paula

Ralph, Daniel Real, Chance W Riggins, Markus Ruhsam, Rowan F Sage, Ann K Sakai, Moira Scascitella, Edward E Schilling, Eva-Marie Schlösser, Heike Sederoff, Stein Servick, Emily B Sessa, A Jonathan Shaw, Shane W Shaw, Erin M Sigel, Cynthia Skema, Alison G Smith, Ann Smithson, C Neal Stewart, John R Stinchcombe, Peter Szövényi, Jennifer A Tate, Helga Tiebel, Dorset Trapnell, Matthieu Villegente, Chun-Neng Wang, Stephen G Weller, Michael Wenzel, Stina Weststrand, James H Westwood, Dennis F Whigham, Shuangxiu Wu, Adrien S Wulff, Yu Yang, Dan Zhu, Cuili Zhuang, Jennifer Zuidof, Mark W Chase, J Chris Pires, Carl J Rothfels, Jun Yu, Cui Chen, Li Chen, Shifeng Cheng, Juanjuan Li, Ran Li, Xia Li, Haorong Lu, Yanxiang Ou, Xiao Sun, Xuemei Tan, Jingbo Tang, Zhijian Tian, Feng Wang, Jun Wang, Xiaofeng Wei, Xun Xu, Zhixiang Yan, Fan Yang, Xiaoni Zhong, Feiyu Zhou, Ying Zhu, Yong Zhang, Saravanaraj Ayyampalayam, Todd J Barkman, Nam-phuong Nguyen, Naim Matasci, David R Nelson, Erfan Sayyari, Eric K Wafula, Ramona L Walls, Tandy Warnow, Hong An, Nils Arrigo, Anthony E Baniaga, Sally Galuska, Stacy A Jorgensen, Thomas I Kidder, Hanghui Kong, Patricia Lu-Irving, Hannah E Marx, Xinshuai Qi, Chris R Reardon, Brittany L Sutherland, George P Tiley, Shana R Welles, Rongpei Yu, Shing Zhan, Lydia Gramzow, Günter Theißen, Gane Ka-Shu Wong, and One Thousand Plant Transcriptomes Initiative. One thousand plant transcriptomes and the phylogenomics of green plants. *Nature*, 574(7780):679–685, October 2019.

39. Erfan Sayyari, Ban Kawas, and **Siavash Mirarab**. TADA: phylogenetic augmentation of microbiome samples enhances phenotype classification. *Bioinformatics*, 35(14):i31–i40, July 2019.

40. Lei Chen, Qiang Qiu, Yu Jiang, Kun Wang, Zeshan Lin, Zhipeng Li, Faysal Bibi, Yongzhi Yang, Jinhuan Wang, Wenhui Nie, Weiting Su, Guichun Liu, Qiye Li, Weiwei Fu, Xiangyu Pan, Chang Liu, Jie Yang, Chenzhou Zhang, Yuan Yin, Yu Wang, Yue Zhao, Chen Zhang, Zhongkai Wang, Yanli Qin, Wei Liu, Bao Wang, Yandong Ren, Ru Zhang, Yan Zeng, Rute R. da Fonseca, Bin Wei, Ran Li, Wenting Wan, Ruoping Zhao, Wenbo Zhu, Yutao Wang, Shengchang Duan, Yun Gao, Yong E. Zhang, Chunyan Chen, Christina Hvilsom, Clinton W. Epps, Leona G. Chemnick, Yang Dong, **Siavash Mirarab**, Hans Redlef Siegismund, Oliver A. Ryder, M. Thomas P. Gilbert, Harris A. Lewin, Guojie Zhang, Rasmus

Heller, and Wen Wang. Large-scale ruminant genome sequencing provides insights into their evolution and distinct traits. *Science*, 364(6446):eaav6202, June 2019.

41. Shahab Sarmashghi, Kristine Bohmann, M. Thomas P. Gilbert, Vineet Bafna, and **Siavash Mirarab**. Skmer: assembly-free and alignment-free sample identification using genome skims. *Genome Biology*, 20(1):34, December 2019.

42. John Yin, Chao Zhang, and **Siavash Mirarab**. ASTRAL-MP: scaling AS-TRAL to very large datasets using randomization and parallelization. *Bioinformatics*, 35(20):3961–3969, October 2019.

43. Metin Balaban, Niema Moshiri, Uyen Mai, Xingfan Jia, and **Siavash Mirarab**. TreeCluster: Clustering biological sequences using phylogenetic trees. *PLOS ONE*, 14(8):e0221068, 2019.

44. Niema Moshiri, Manon Ragonnet-Cronin, Joel O Wertheim, and **Siavash Mirarab**. FAVITES: simultaneous simulation of transmission networks, phylogenetic trees and sequences. *Bioinformatics*, 35(11):1852–1861, June 2019.

45. Peter Houde, Edward L Braun, Nitish Narula, Uriel Minjares, and **Siavash Mirarab**. Phylogenetic Signal of Indels and the Neoavian Radiation. *Diversity*, 11(7):108, July 2019.

46. Maryam Rabiee, Erfan Sayyari, and **Siavash Mirarab**. Multi-allele species reconstruction using ASTRAL. *Molecular Phylogenetics and Evolution*, 130:286–296, January 2019. Publisher: Elsevier.

47. Shyam Gopalakrishnan, Mikkel-Holger S. Sinding, Jazmín Ramos-Madrigal, Jonas Niemann, Jose A. Samaniego Castruita, Filipe G. Vieira, Christian Carøe, Marc de Manuel Montero, Lukas Kuderna, Aitor Serres, Víctor Manuel González-Basallote, Yan-Hu Liu, Guo-Dong Wang, Tomas Marques-Bonet, **Siavash Mirarab**, Carlos Fernandes, Philippe Gaubert, Klaus-Peter Koepfli, Jane Budd, Eli Knispel Rueness, Claudio Sillero, Mads Peter Heide-Jørgensen, Bent Petersen, Thomas Sicheritz-Ponten, Lutz Bachmann, Øystein Wiig, Anders J. Hansen, and M. Thomas P. Gilbert. Interspecific Gene Flow Shaped the Evolution of the Genus Canis. *Current Biology*, 28(21):3441–3449.e5, November 2018.

48. Daniel McDonald, Embriette Hyde, Justine W. Debelius, James T. Morton, Antonio Gonzalez, Gail Ackermann, Alexander A. Aksenov, Bahar Behsaz, Caitriona Brennan, Yingfeng Chen, Lindsay DeRight Goldasich, Pieter C. Dorrestein, Robert R. Dunn, Ashkaan K. Fahimipour, James Gaffney, Jack A. Gilbert, Grant Gogul, Jessica L. Green, Philip Hugenholtz, Greg Humphrey, Curtis Huttenhower, Matthew A. Jackson, Stefan Janssen, Dilip V. Jeste, Lingjing Jiang, Scott T. Kelley, Dan Knights, Tomasz Kosciolek, Joshua Ladau, Jeff Leach, Clarisse Marotz, Dmitry Meleshko, Alexey V. Melnik, Jessica L. Metcalf, Hosein Mohimani, Emmanuel Montassier, Jose Navas-Molina, Tanya T. Nguyen, Shyamal Peddada, Pavel Pevzner, Katherine S. Pollard, Gholamali Rahnavard, Adam Robbins-Pianka, Naseer Sangwan, Joshua Shorenstein, Larry Smarr, Se Jin Song, Timothy Spector, Austin D. Swafford, Varykina G. Thackray, Luke R. Thompson, Anupriya Tripathi, Yoshiki Vázquez-Baeza, Alison Vrbanac, Paul Wischmeyer, Elaine Wolfe, Qiyun Zhu, Rob Knight, Allison E. Mann, Amnon Amir, Angel Frazier, Cameron Martino, Carlito Lebrilla, Catherine Lozupone, Cecil M. Lewis, Charles Raison, Chi Zhang, Christian L. Lauber, Christina Warinner, Christopher A. Lowry, Chris Callewaert, Cinnamon Bloss, Dana Willner, Daniela Domingos Galzerani, David J. Gonzalez, David A. Mills, Deepak Chopra, Dirk Gevers, Donna Berg-Lyons, Dorothy D. Sears, Doug Wendel, Elijah Lovelace, Emily Pierce, Emily TerAvest, Evan Bolyen, Frederic D. Bushman, Gary D. Wu, George M. Church, Gordon Saxe, Hanna D. Holscher, Ivo Ugrina, J. Bruce German, J. Gregory Caporaso, Jacob M. Wozniak, Jacqueline Kerr, Jacques Ravel, James D. Lewis, Jan S. Suchodolski, Janet K. Jansson, Jarrad T. Hampton-Marcell, Jason Bobe, Jeroen Raes, John H. Chase, Jonathan A. Eisen, Jonathan Monk, Jose C. Clemente, Joseph Petrosino, Julia Goodrich, Julia Gauglitz, Julian Jacobs, Karsten Zengler, Kelly S. Swanson, Kim Lewis, Kris Mayer, Kyle Bittinger, Lindsay Dillon, Livia S. Zaramela, Lynn M. Schriml, Maria G. Dominguez-Bello, Marta M. Jankowska, Martin Blaser, Meg Pirrung, Michael Minson, Mike Kurisu, Nadim Ajami, Neil R. Gottel, Nicholas Chia, Noah Fierer, Owen White, Patrice D. Cani, Pawel Gajer, Philip Strandwitz, Purna Kashyap, Rachel Dutton, Rachel S. Park, Ramnik J. Xavier, Robert H. Mills, Rosa Krajmalnik-Brown, Ruth Ley, Sarah M. Owens, Scott Klemmer, Sébastien Matamoros, **Siavash Mirarab**, Stephanie Moorman, Susan Holmes, Tara Schwartz, Tifani W. Eshoo-Anton, Tim Vigers, Vineet Pandey, Will Van Treuren, Xin Fang, Zhenjiang Zech Xu, Alan Jarmusch, Justin Geier, Nicolai Reeve, Ricardo Silva, Evguenia Kopylova, Dominic

Nguyen, Karenina Sanders, Rodolfo Antonio Salido Benitez, Arthur Cole Heale, Max Abramson, Jérôme Waldispühl, Alexander Butyaev, Chris Drogaris, Elena Nazarova, Madeleine Ball, and Beau Gunderson. American Gut: an Open Platform for Citizen Science Microbiome Research. *mSystems*, 3(3), May 2018.

49. 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, March 2018.

50. 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, April 2018.

51. 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, May 2018.

52. Uyen Mai and **Siavash Mirarab**. TreeShrink: fast and accurate detection of outlier long branches in collections of phylogenetic trees. *BMC Genomics*, 19(S5):272, May 2018. ISBN: 9783319679785.

53. Erfan Sayyari and **Siavash Mirarab**. Testing for Polytomies in Phylogenetic Species Trees Using Quartet Frequencies. *Genes*, 9(3):132, February 2018. Publisher: Multidisciplinary Digital Publishing Institute.

54. Erfan Sayyari, James B. Whitfield, and **Siavash Mirarab**. DiscoVista: Interpretable visualizations of gene tree discordance. *Molecular Phylogenetics and Evolution*, 122:110–115, May 2018. arXiv: 1709.09305.

55. 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, February 2018. Publisher: Nature Publishing Group.

56. 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*, 15(5):1738–1747, 2018.

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

58. 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, December 2017.

59. 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):457–463, 2017.

60. Uyen Mai and **Siavash Mirarab**. TreeShrink: Efficient Detection of Outlier Tree Leaves. In Joao Meidanis and Luay Nakhleh, editors, *Comparative Genomics. RECOMB-CG 2017. Lecture Notes in Computer Science, vol 10562*, 116–140. Springer International Publishing, Cham, 2017.

61. Dong-sheng Zhang, Ya-dong Zhou, Chun-sheng Wang, and Greg Rouse. A new species of Ophryotrocha (Annelida, Eunicida, Dorvilleidae) from hydrothermal vents on the Southwest Indian Ridge. *ZooKeys*, 687:1–9, August 2017.

62. Uyen Mai, Erfan Sayyari, and **Siavash Mirarab**. Minimum variance rooting of phylogenetic trees and implications for species tree reconstruction. *PLOS ONE*, 12(8):e0182238, August 2017. Publisher: Public Library of Science ISBN: 1111111111.

63. **Siavash Mirarab**. Phylogenomics: Constrained gene tree inference. *Nature Ecology & Evolution*, 1(2):0056, January 2017. Publisher: Nature Publishing Group.

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

65. 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.

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

67. 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, February 2016.

68. **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, October 2015.

69. 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, September 2015.

70. 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.

71. 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.

72. **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, June 2015.

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

74. 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, January 2015. Publisher: Public Library of Science.

75. Erich D Jarvis, **Siavash Mirarab**, Andre J Aberer, Bo Li, Peter Houde, Cai Li, Simon Y W Ho, Brant C Faircloth, Benoit Nabholz, Jason T Howard, Alexander Suh, Claudia C Weber, Rute R da Fonseca, Alonzo Alfaro-Núñez, Nitish Narula, Liang Liu, Dave Burt, Hans Ellegren, Scott V Edwards, Alexandros Stamatakis, David P Mindell, Joel Cracraft, Edward L Braun, Tandy Warnow, Wang Jun, M Thomas Pius Gilbert, and Guojie Zhang. Phylogenomic analyses data of the avian phylogenomics project. *GigaScience*, 4(1):1–9, December 2015. Publisher: BioMed Central.

76. **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(5):377–86, May 2015.

77. E. D. Jarvis, **Siavash Mirarab**, A. J. Aberer, B. Li, Peter Houde, C. Li, S. Y. W. Ho, B. C. Faircloth, Benoit Nabholz, J. T. Howard, Alexander Suh, C. C. Weber, R. R. da Fonseca, J. Li, F. Zhang, H. Li, L. Zhou, Nitish Narula, L. Liu, Ganeshkumar Ganapathy, Bastien Boussau, M. S. Bayzid, Volodymyr Zavidovych, Sankar Subramanian, T. Gabaldon, S. Capella-Gutierrez, Jaime Huerta-Cepas, Bhanu Rekepalli, Kasper Munch, M. Schierup, Bent Lindow, W. C. Warren, David Ray,

R. E. Green, M. W. Bruford, Xiangjiang Zhan, Andrew Dixon, S. Li, N. Li, Yinhua Huang, E. P. Derryberry, M. F. Bertelsen, F. H. Sheldon, R. T. Brumfield, C. V. Mello, P. V. Lovell, Morgan Wirthlin, M. P. C. Schneider, Francisco Prosdocimi, J. A. Samaniego, A. M. V. Velazquez, A. Alfaro-Nunez, P. F. Campos, Bent Petersen, Thomas Sicheritz-Ponten, An Pas, Tom Bailey, Paul Scofield, Michael Bunce, D. M. Lambert, Qi Zhou, Polina Perelman, A. C. Driskell, Beth Shapiro, Zijun Xiong, Yongli Zeng, S. Liu, Zhenyu Li, Binghang Liu, Kui Wu, Jin Xiao, Xiong Yinqi, Qiuemei Zheng, Y. Zhang, Huanming Yang, Jian Wang, Linnea Smeds, F. E. Rheindt, M. Braun, Jon Fjeldsa, Ludovic Orlando, F.K. Keith Barker, K. A. Jonsson, Warren Johnson, K.-P. Koepfli, S. O'Brien, David Haussler, O. A. Ryder, Carsten Rahbek, Eske Willerslev, Gary R Graves, T. C. Glenn, J. McCormack, D. Burt, Hans Ellegren, P. Alstrom, S. V. Edwards, Alexandros Stamatakis, D. P. Mindell, Joel Cracraft, E. L. Braun, Tandy Warnow, Wang Jun, M.T.P. 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, December 2014.

78. **Siavash Mirarab**, M. S. Bayzid, Bastien Boussau, and Tandy Warnow. Statistical binning enables an accurate coalescent-based estimation of the avian tree. *Science*, 346(6215):1250463–1250463, December 2014.

79. Naim Matasci, Ling-Hong L.-H. Hung, Zhixiang Yan, E.J. Eric J Carpenter, N.J. Norman J Wickett, **Siavash Mirarab**, Nam Nguyen, Tandy Warnow, Saravanaraj Ayyampalayam, Michael S Barker, J.G. Burleigh, M.A. Gitzendanner, E. Wafula, J.P. Der, C.W. dePamphilis, B. Roure, H. Philippe, B.R. Ruhfel, N.W. Miles, S.W. Graham, S. Mathews, B. Surek, M. Melkonian, D.E. Soltis, P.S. Soltis, C. Rothfels, L. Pokorny, J.A. Shaw, L. DeGironimo, D.W. Stevenson, J.C. Villarreal, T. Chen, T.M. Kutchan, M. Rolf, R.S. Baucom, M.K. Deyholos, R. Samudrala, Z. Tian, X. Wu, X. Sun, Y. Zhang, J. Wang, J. Leebens-Mack, and G.K.S. Wong. Data access for the 1,000 Plants (1KP) project. *GigaScience*, 3(1):17, 2014. Publisher: BioMed Central.

80. Norman J. Wickett, **Siavash Mirarab**, Nam Nguyen, Tandy Warnow, Eric J Carpenter, Naim Matasci, Saravanaraj Ayyampalayam, Michael S. Barker, J. Gordon Burleigh, Matthew A. Gitzendanner, Brad R. Ruhfel, Eric Wafula, Joshua P. Der, Sean W. Graham, Sarah Mathews, Michael Melkonian, Douglas E. Soltis, Pamela S. Soltis, Nicholas W. Miles, Carl J. Rothfels, Lisa Pokorny, A. Jonathan Shaw, Lisa

DeGironimo, Dennis W. Stevenson, Barbara Surek, Juan Carlos Villarreal, Béatrice Roure, Hervé Philippe, Claude W. DePamphilis, Tao Chen, Michael K. Deyholos, Regina S. Baucom, Toni M. Kutchan, Megan M. Augustin, Jian Jun Wang, Yong Zhang, Zhijian Tian, Zhixiang Yan, Xiaolei Wu, Xiao Sun, Gane Ka-Shu Wong, and James Jim Leebens-Mack. Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proceedings of the National Academy of Sciences*, 111(45):E4859–4868, October 2014. arXiv: 1301.5277v2 ISBN: 0027-8424.

81. **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, May 2016. Publisher: Oxford University Press ISBN: 1063-5157.

82. **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.

83. **Siavash Mirarab**, Rezwana Reaz, Md. Shamsuzzoha Bayzid, Théo Zimmermann, M. S. Swenson, and Tandy Warnow. ASTRAL: genome-scale coalescent-based species tree estimation. *Bioinformatics*, 30(17):i541–i548, September 2014.

84. Nam-phuong Nguyen, **Siavash Mirarab**, Bo Liu, Mihai Pop, and Tandy Warnow. TIPP: taxonomic identification and phylogenetic profiling. *Bioinformatics*, 30(24):3548–3555, December 2014.

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

86. Arlin Stoltzfus, Hilmar Lapp, Naim Matasci, Helena Deus, Brian Sidlauskas, Christian M C.M. M Zmasek, Gaurav Vaidya, Enrico Pontelli, Karen a. Cranston, Rutger Vos, Campbell O C.O. O Webb, Luke J L.J. Harmon, Megan Pirrung, Brian O'Meara, Matthew W M.W. W Pennell, **Siavash Mirarab**, Michael S M.S. S Rosenberg, James P J.P. P Balhoff, Holly M H.M. M Bik, Tracy A Heath, P.E. Peter E Midford, J.W. Joseph W Brown, Emily Jane E.J. McTavish, Jeet Sukumaran, Mark Westneat, M.E. Michael E Alfaro, Aaron Steele, and Greg Jordan. Phylotastic! Making tree-of-life knowledge accessible, reusable and convenient. *BMC*

bioinformatics, 14(1):158, 2013. Publisher: BioMed Central Ltd ISBN: 1471-2105 (Electronic)\r1471-2105 (Linking).

87. 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.

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

89. **Siavash Mirarab**, Nam Nguyen, and Tandy Warnow. SEPP: SATé-Enabled Phylogenetic Placement. In *Pacific Symposium On Biocomputing*, 247–58. WORLD SCIENTIFIC, December 2012.

90. **Siavash Mirarab**, Soroush Akhlaghi, and Ladan Tahvildari. Size-constrained regression test case selection using multicriteria optimization. *Software Engineering, IEEE Transactions on*, 38(4):936–956, 2012. Publisher: IEEE.

91. **Siavash Mirarab** and Tandy Warnow. FastSP: linear time calculation of alignment accuracy. *Bioinformatics*, 27(23):3250–8, 2011. ISBN: 1367-4811; 1367-4803.

92. Hyunsook Do, **Siavash Mirarab**, Ladan Tahvildari, and Gregg Rothermel. The effects of time constraints on test case prioritization: A series of controlled experiments. *Software Engineering, IEEE Transactions on*, 36(5):593–617, 2010. Publisher: IEEE.

93. **Siavash Mirarab** and Ladan Tahvildari. An empirical study on bayesian network-based approach for test case prioritization. In *Software Testing, Verification, and Validation, 2008 1st International Conference on*, 278–287. IEEE, 2008.

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

95. 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*, 71–82. ACM, 2008.

96. Mehdi Amoui, Mazeiar Salehie, **Siavash Mirarab**, and Ladan Tahvildari. Adaptive action selection in autonomic software using reinforcement learning. In *Autonomic and Autonomous Systems, 2008. ICAS 2008. Fourth International Conference on*, 175–181. IEEE, 2008.

97. **Siavash Mirarab** and Ladan Tahvildari. A prioritization approach for software test cases based on bayesian networks. In *Fundamental Approaches to Software Engineering*, volume 4422 LNCS, 276–290. Springer Berlin Heidelberg, 2007. ISSN: 03029743.

98. **Siavash Mirarab**, Alaa Hassouna, and Ladan Tahvildari. Using bayesian belief networks to predict change propagation in software systems. In *Program Comprehension, 2007. ICPC'07. 15th IEEE International Conference on*, 177–188. IEEE, 2007.

99. Mehdi Amoui, **Siavash Mirarab**, Sepand Ansari, and Caro Lucas. A genetic algorithm approach to design evolution using design pattern transformation. *International Journal of Information Technology and Intelligent Computing*, 1(2):235–244, 2006.