Saurabh Belsare, PhD

Email: saurabhmbelsare@gmail.com
GitHub: https://github.com/saurabhbelsare
Website: http://saurabhbelsare.com

Summary

I'm a researcher in Computational Biology, with significant experience in computational genomic analysis, developing machine learning methods for genomic data and biophysical simulations. My research experience spans both, large scale data analysis for gaining biological insights from empirical data, as well as methods development for new applications.

Education

University of California, Berkeley	PhD, Bioengineering (Joint degree with University of California, San Francisco)	Aug 2011-May 2017
Indian Institute of Technology, Bombay	Bachelor of Technology (Hons.), Chemical Engineering; Minor in Biosciences	Aug 2007-May 2011

Work Experience

Postdoctoral Scholar, Kern-Ralph Colab, Institute for Ecology and Evolution, University of Oregon — Mar 2020 - Current
Deep learning methods for population genetics

- Developed a neural network based method for ancestral haplotype prediction
- Explored visualization and analysis techniques for best practices in machine learning for population genetics
- Contributed species and QC for community simulation software stdpopsim

Postdoctoral Scholar, Wall Lab, Institute for Human Genetics, University of California, San Francisco — Jul 2017 - Jan 2020
Analysis of large scale human genomic data

- Assessed the accuracy of public databases for phasing and imputation
- Estimated recombination rates in humans and primates
- Evaluated the accuracy of a population specific reference panel for imputation
- Performed liftover for comparison between species

Graduate Student Researcher, Head-Gordon Lab, University of California, Berkeley — Jul 2012 - Jun 2017
Studying dynamics of small biomolecules and enzymes, and their interactions with solvent

- Developed computational methods to analyze dynamics of biomolecules
- Implemented codes in Fortran within existing simulation packages to enable calculations of dynamic observables
- Explained solvent contribution to improved catalytic activity Kemp Eliminase via analysis of solvent entropy via molecular dynamics using AMOEBA polarizable force field
- Generated improved small molecule parameterization and formulation for localizing charges to allow mode decomposition to classical simulation approaches to with *ab-initio* simulation approaches

Journal Publications

[Google Scholar]

- Expanding the stdpopsim species catalog, and lessons learned for realistic genome simulations (2022) M. Elise Lauterbur, ..., **Saurabh Belsare**, ..., Ilan Gronau *bioRxiv* DOI: 10.1101/2022.10.29.514266
- South Asian Patient Population Genetics Reveal Strong Founder Effects and High Rates of Homozygosity – New Resources for Precision Medicine (2021) Jeffrey D. Wall, ... , **Saurabh Belsare**, ..., Andrew S. Peterson *bioRxiv* DOI: 10.1101/2020.10.02.323238
- Genetic profiles of 103,106 individuals in the Taiwan Biobank provide insights into the health and history of Han Chinese (2021) Chun-Yu Wei, ... , **Saurabh Belsare**, ..., Pui-Yan Kwok *npj Genomic Medicine 6, 1* DOI: 10.1038/s41525-021-00178-9
- Evaluating the quality of the 1000 Genomes Project data (2019) **Saurabh Belsare**, Michal Sakin-Levy, Yulia Mostovoy, Steffen Durinck, Subhra Chaudhry, Ming Xiao, Andrew S. Peterson, Pui-Yan Kwok, Somasekar Seshagiri and Jeffrey D. Wall *BMC Genomics 20, 1* DOI: 10.1186/s12864-019-5957-x
- Analysis of 100 high-coverage genomes from a pedigreed captive baboon colony (2019) Jacqueline A. Robinson, **Saurabh Belsare**, Shifra Birnbaum, Deborah E. Newman, Jeannie Chan, Jeremy P. Glenn, Betsy Ferguson, Laura A. Cox and Jeffrey D. Wall *Genome Res. 29: 848-856* DOI:10.1101/gr.247122.118
- Solvent Entropy Contributions to Catalytic Activity in Designed and Optimized Kemp Eliminases (2017) **Saurabh Belsare**[1], Viren Pattni[1], Matthias Heyden, Teresa Head-Gordon, *J. Phys. Chem. B 122, 21, 5300-5307* DOI: 10.1021/acs.jpcb.7b07526
- Mode Specific THz Spectra of Solvated Amino Acids using the AMOEBA Polarizable Force Field (2017) Alexander Esser[1], **Saurabh Belsare**[1], Dominik Marx, and Teresa Head-Gordon, *Phys. Chem. Chem. Phys. 19, 5579-5590* DOI: 10.1039/C6CP07388C

Selected Recent Posters and Talks

- Neural Network based Ancestral Haplotype Prediction (April 2021) Saurabh Belsare, Andrew Kern (Probablistic Modeling in Genetics Meeting 2021) [Poster]
- Imputation in South Asian individuals using a high-coverage South Asian reference panel. (October 2019) Saurabh Belsare et al (American Society of Human Genetics Meeting 2019) [Poster]
- Evaluating Phasing and Imputation in Reference Databases (November 2019) (Center for Theoretical and Evolutionary Genetics Seminar, University of California, Berkeley) [Talk]
- Evaluating Reference Databases for Imputation in Humans (November 2019) (Bay Area Population Genomics Meeting) [Talk]

Computer Skills

- Code: R, python, awk, bash
- Bioinformatics: phasing, imputation, recombination rate estimation liftover, genomic data anlysis, molecular dynamics
- Computing: SLURM, git, snakemake, OpenMP
- Machine Learning: Tensorflow, Keras, scikit-learn Neural networks (Dense, CNNs, RNNs), logistic regression

Peer reviewer for journals:

- Human Genetics and Genomics Advances
- Genetics
- G3: Genes | Genomes | Genetics
- Nucleic Acids Research
- Molecular Ecology Resources
- Integrative and Comparative Biology

[1]Equal Contribution