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

Aug 2011- May 2017	PhD. in Bioengineering | Advisor: Prof. Teresa Head-Gordon The UCB-UCSF Graduate Program in Bioengineering, University of California, Berkeley Dissertation: Biomolecular Dynamics and Function: A Study on Amino Acids and Enzymes
Aug 2007- May 2011	Bachelor of Technology (Hons.) in Chemical Engineering | Minor in Biosciences Indian Institute of Technology, Bombay, Mumbai, India Bachelor Thesis:Computational Prediction of Loop Structures in GPCR Proteins Advisor: Prof. Santosh Noronha

Work Experience

Mar 2020- Current	Postdoctoral Scholar, Kern-Ralph Colab, Institute for Ecology and Evolution, University of Oregon *Deep learning methods for population genetics*
July 2017- Jan 2020	Postdoctoral Scholar, Wall Lab, Institute for Human Genetics, University of California, San Francisco *Analysis of large scale human genomic data*

Publications & Presentations

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. (2018) **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 (2019) *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

Talks and Presentations

- 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 Reference Databases for Imputation in Humans. (November 2019) **Saurabh Belsare** (Bay Area Population Genomics Meeting 2019) [Talk]
- Evaluating Phasing and Imputation in Reference Databases. (November 2019) **Saurabh Belsare** (Center for Theoretical and Evolutionary Genomics Seminar, Berkeley) [Talk]
- Studying the dynamics and solvent interactions of amino acids and proteins through biomolecular simulations (January 2018) Department of Biosciences and Bioengineering, Indian Institute of Technology, Bombay [Talk]

[1]Equal Contribution

- Evaluating the phasing accuracy of the 1000 Genomes Project. (November 2017) **Saurabh Belsare** and Jeff Wall (Bay Area Population Genomics Meeting 2017) [Poster]
- Studying Solvation of Small Biomolecules via Molecular Dynamics using a Polarizable Force Field. (Feb 2017) **Saurabh Belsare**[1], Alexander Esser[1], Dominik Marx, and Teresa Head-Gordon (Proceedings of the Biophysical Society Meeting 2017) [Poster]
- Studying Solvation of Small Biomolecules via Molecular Dynamics using a Polarizable Force Field. (January 2017) Alexander Esser[1], **Saurabh Belsare**[1], Dominik Marx, and Teresa Head-Gordon. (Berkeley Statistical Mechanics Meeting) [Poster]
- Dynamic metrics for rationalizing directed evolution of enzymes. (August 2014) **Saurabh Belsare**, Teresa Head-Gordon (Proceedings of the American Chemical Society Meeting 2014) [Poster]
- Rationalizing Directed Evolution through Protein Dynamics. (Jan 2014) **Saurabh Belsare** (Proceedings of the Biophysical Society Meeting 2014) [Poster]
- Homology Modeling of Protein Structure using Fragment/Profile based search method in CASP9. M. Parthiban, S. Susdalzew, **S. Belsare**, et. al. 2010 Critical assessment of methods for protein structure prediction (CASP9) - Abstract Book

Research

Mar 2020-Current — Postdoctoral Research
Developing deep learning methods for genomic analysis
Advisor: Prof. Andrew Kern, Prof. Peter Ralph, Institute for Ecology and Evolution, University of Oregon
- Developed a neural network based method for ancestral haplotype prediction
- Improved genealogical reconstruction from ancestral haplotypes predicted using neural networks
- Contributed species and QC for community simulation software

July 2017-Jan 2020 — Postdoctoral Research
Analysis of haplotype phasing and genotype imputation methods applied to human genomic data
Advisor: Prof. Jeff Wall, Institute for Human Genetics, University of California, San Francisco
- 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

July 2012-June 2017 — Graduate Research
Studying dynamics of small biomolecules and enzymes, and their interactions with solvent
Advisor: Prof. Teresa Head-Gordon, Department of Bioengineering, University of California, Berkeley
- Devaloped computational methods to analyze dynamics of biomolecules
- Implemented codes in Fortran within existing simulation packages to enable calculations of dynamic observables
- Developed an understanding of the influence of specific residue modification on protein function
- Interpreted the influence of solvent on biomolecule behavior via solvent VDOS and THz observables
- Developed scripts in python and matlab for determining observables from simulation data

May 2016-June 2017 — Research Collaboration
Studying entropic contributions to enzyme solvation using a polarizable force field
Co-Advisor: Prof. Matthias Heyden, Max Planck Institut für Kohlenforschung, Mülheim
- Performed molecular dynamics simulations of Kemp Eliminase artificial enzymes and MMP enzyme using the AMOEBA polarizable force field
- Analyzed the solvation entropy trends over evolved variants in Kemp Eliminase to explain solvent contribution to improved catalytic activity
- Analyzed solvation around MMP active site with zinc cofactor, to understand influence of polarization and solvation in entropic contribution to MMP activity

May 2015-June 2017 — Research Collaboration
Studying small molecule modes via classical simulations
Co-Advisor: Prof. Dominik Marx, Lehrstuhl für Theoretische Chemie, Ruhr Universität, Bochum
- Modified existing classical simulation approaches to compare with *ab-initio* simulation approaches
- Developed a formulation for localizing charges to allow mode decomposition
- Obtained significant agreement of dynamic modes predicted from classical simulations to *ab-initio* simulations

Computer Skills

- Code: Python, R, awk, bash
- Bioinformatics: phasing, imputation, recombination rate estimation, liftover, genomic data analysis, molecular dynamics
- Computing: SLURM, git, snakemake, OpenMP
- Machine Learning: Tensorflow, Keras, Neural Networks (Dense, CNNs, RNNs), Logistic Regression

Peer Reviewing

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

Teaching Assistantship

BioE 131/231: Introduction to Computational Molecular Biology [Aug-Dec 2013]
Course Instructor: Prof. Ian Holmes, Department of Bioengineering, University of California, Berkeley

- Graduate student instructor (GSI) for a joint undergraduate-graduate course
- Conducted introductory programming classes, lab sessions, office hours, and graded homeworks and projects

Relevant Coursework

UC Berkeley: Computational Methods in Biology, Statistical Phylogenetics, Probabilistic Modeling in Computational Biology, Introduction to Protein Informatics, Thermodynamics and Statistical Mechanics, Software Engineering for Scientific Computing, Applications of Parallel Computers, Introduction to Statistical Computing

Indian Institute of Technology, Bombay: Introduction to Computational Biology, Chemistry of Biomolecules, Molecular Biology, Molecular Immunology, Genetic Engineering, Optimization, Applied Multivariate Statistics in Chemical Engineering