Subrata Paul – Résumé

Bioinformatics Analyst
Axle Informatics
860 Clermont St, APT 905
Denver, CO 80220

Mobile Phone	+1 (765) 631 5129
Email	subrata.paul@nih.gov
Linkedin	www.linkedin.com/in/paulsub/
Github	github.com/spaul-genetics

Experience:

Nov 2021

Ongoing

Bioinformatics Scientist, Axle Informatics
I am working on multiple projects at the National Institute of Allergy and Infectious Diseases (NIAID) and the National Institute of Arthritis and Musculoskeletal and Skin Diseases. The projects include analyzing whole genome sequencing, single-cell RNA-seq, and data mining. In one project we investigated the effect of aging on immune cells in HIV-infected individuals. In a clinical trial study, we are investigating the effect of stopping TNF inhibitors in rheumatoid arthritis remission. I am also analyzing WGS data of capillary syndrome and cryptococcal meningitis patients.

Aug 2021
to
Oct 2021

Bioinformatics Analyst III, Leidos Biomedical Research Inc.
My primary responsibility is to provide statistical genetics and bioinformatics support to the projects at the NIAID and the NIAIMS. The project I am currently working on is the rare variant association of scleroderma and Systemic Capillary Leak Syndrome (SCLS). In another project, I am studying the effect of arthritis in a patient's financial and social life.

Jan 2020
to
July 2021

Postdoctoral Associate, Institute for Behavioral Genetics, CU Boulder
In collaboration with multiple labs I am working on uncovering genetic and phenotypic correlations among different pain conditions using UK Biobank data. In another project, I am investigating performance of dosage data in estimating heritability compared to the hard-call genotype. I am co-supervising two graduate students in their projects on genetic genetic correlations of behavioral traits using genetic relationship matrix based approach and quantifying phenotypic assortment.

May 2015
to
Dec 2019

Research Assistant, Dept. of Mathematical and Statistical Science, CU Denver
I was involved in genetic risk prediction of generalized vitiligo. My responsibility was data pre-processing, genome-wide association analysis, model polygenic risk score, compare multiple risk score model, dissect risk score model using functional annotation. I was also involved in analysing secondary phenotypes and advanced method development to incorporate phenotypic heterogeneity into association framework.

May 2018
to
Jul 2018

Collaboration with Department of Electrical Engineering, CU Denver
Objective of this project was real time electricity price forecasting to optimize electricity cost. I applied artificial neural network on intuitively designed feature space for forecasting electricity price in real time.

June 2017
to
July 2017

Research Assistant, Dept. of Mathematical and Statistical Science, CU Denver
A functional data analysis approach was developed to model spatiotemporal data. In this project, I used sparse matrices, numerical linear algebraic methods, and parallel computing to reduce computational expenses.

Education

Aug 2014 to **Dec 2019**	University of Colorado Denver PhD in Applied Mathematics Concentration in Statistical Genetics. Through courses, I have learned advanced statistical models, Bayesian analysis, machine learning techniques, and human genetics etc.
Aug 2012 to **May 2014**	Ball State University, Indiana, USA MS in Mathematics MS thesis: *Numerical Multigrid algorithm for solving Integral Equation.* Focused on solving special types of integral equations with less computational complexity using multigrid method and to show its efficiency.
Mar 2003 to **Dec 2007**	University of Dhaka BS in Mathematics with minors in Computer Science, Statistics and Physics BS project: *On Binomial Asset Pricing Model with convergence to Black-Scholes model.* Discussed the power of binomial model on asset pricing and proven its strength by showing its convergence to Black-Scholes model.

Publications and Posters

2022	M. Ali Rai, Victoria Shi, Brooke D. Kennedy, Emily J. Whitehead, Jesse S. Justement, Subrata Paul, Jana Blazkova, and Tae-Wook Chun. Effect of aging on immune cells in male hiv-infected and -uninfected healthy individuals. *AIDS*, 9900 Danielle Niemann, Subrata Paul, and Humairat H Rahman. Avalanche preparedness and accident analysis among backcountry skier, sidecountry, and snowmobile fatalities in the united states: 2009 to 2019. *Wilderness & Environmental Medicine*, 33(2):197–203, 2022
2020	Subrata Paul and Stephanie A Santorico. Optimized genetic risk prediction for vitiligo and its use to define disease subtypes. In *GENETIC EPIDEMIOLOGY*, volume 44-5, page 515, 2020
2019	Genevieve H. L. Roberts, Subrata Paul, Daniel Yorgov, Stephanie A. Santorico, and Richard A. Spritz. Family clustering of autoimmune vitiligo results principally from polygenic inheritance of common risk alleles. *The American Journal of Human Genetics*, July 2019
2018	Subrata Paul and Stephanie A Santorico. Incorporation of heterogeneity through a mixture model to boost power of association tests. In *GENETIC EPIDEMIOLOGY*, volume 42-7, pages 722–722, 2018
2018	Md Habib Ullah, Subrata Paul, and Jae-Do Park. Real-time electricity price forecasting for energy management in grid-tied mtdc microgrids. In *2018 IEEE Energy Conversion Congress and Exposition (ECCE)*, pages 73–80. IEEE, 2018
2016	Stephanie A Santorico, Subrata Paul, Daniel Yorgov, Ying Jin, Tracey Ferrara, and Richard A Spritz. A comparison of genetic risk prediction and subtyping for generalized vitiligo. In *GENETIC EPIDEMIOLOGY*, volume 40-7, pages 657–658, 2016
2015	Stephanie A Santorico, Ying Jin, Daniel Yorgov, Subrata Paul, Tracey Ferrara, and Richard Spritz. Optimized genetic risk prediction for vitiligo and its use to define disease subtypes. In *GENETIC EPIDEMIOLOGY*, volume 39-7, pages 577–578, 2015
2009	Sheik Ahmed Ullah, Subrata Paul, and Md. Sharif Ullah Mozumder. On information carriage through sigma-algebra in binomial asset pricing model. *Journal of Statistical Studies*, 28:1–8, 2009

Skills

Core	Statistical Modeling Human Genetics scRNA-seq Survival Analysis Machine Learning Computational Math Pure Math	Programming Languages	R (Expert) Python (Advanced level) SQL, MATLAB, C++, FORTRAN (Intermediate)
Languages	English (Proficient) Bengali (Native) Hindi (Communication)	Others	Latex, Beamer Mathematica Data Visualization Inkscape

Referees

Stephanie A. Santorico, Ph.D.
Professor,
University of Colorado Denver
Department of Math and Statistical Sciences
P.O. Box 173364
Campus Box 170
Denver, CO 80217-3364
Email: stephanie.santorico@ucdenver.edu
Phone: (303) 315-1714

Matthew C. Keller, Ph.D.
Professor,
Institute for Behavioral Genetics,
University of Colorado Boulder
447 UCB
Boulder, CO 80309-0447
Email: matthew.c.keller@colorado.edu
Phone: (303) 492-5799