FLORIDA ATLANTIC
U N I V E R S I T Y

Announces the Ph.D. Dissertation Defense of

Sandipan Paul Arnab

for the degree of Doctor of Philosophy (Ph.D.)

From Specialized to Generalized Frameworks: Broadening Machine Learning Approaches to Detect Natural Selection

October 24, 2026, Friday 2:00 PM – 3:30 PM
Engineering East, Room #405
777 Glades Road
Boca Raton, FL

DEPARTMENT:

Electrical Engineering and Computer Science

ADVISOR:

Michael DeGiorgio, Ph.D.

PH.D. SUPERVISORY COMMITTEE:

Michael DeGiorgio, Ph.D., Chair
Hari Kalva, Ph.D.
Raquel Assis, Ph.D.
Behnaz Ghoraani, Ph.D.

ABSTRACT OF DISSERTATION:

Natural selection leaves characteristic footprints in genomic variation, and detecting these patterns is fundamental to understanding evolutionary history and adaptation in humans. In this thesis, I present a progression of machine learning frameworks, beginning with highly specialized feature engineering and complex modeling, advancing to approaches that require minimal training, and culminating in methods that remain robust to model misspecification. First, I developed *SISSSCO*, a spectral feature extraction framework that applies wavelet transforms, multitaper spectral analysis, and S-transform to genomic summary statistics, converting one-dimensional signals into two-dimensional spectral images analyzed by convolutional neural networks. *SISSSCO* achieved high accuracy across varied evolutionary scenarios, remained resilient to missing data, and uncovered both established and novel sweep candidates in European genomes. Building on this foundation, I introduced *TrIdent*, a transfer learning method that leverages pre-trained deep CNNs to efficiently extract features from multilocus genomic images. *TrIdent* reduced simulation requirements while improving detection of adaptive regions, provided interpretability through class activation maps, and revealed novel disease-associated candidate genes in European and African populations. Finally, I developed *PULSe*, a positive-unlabeled learning framework that bypasses the need for explicit negative training data, enabling robust sweep detection under domain shift and demographic misspecification. Applied to European and Bengali genomes, *PULSe* recovered well-supported sweep candidates and demonstrated strong generalizability across complex genomic landscapes. Together, these projects trace a trajectory from specialized yet powerful frameworks toward more flexible, generalized methodologies. This body of work advances machine learning strategies for detecting natural selection, extending their applicability from well-characterized to understudied human populations and enhancing our capacity to uncover the genetic basis of adaptation.

College of Engineering and Computer Science
Office of the Dean
777 Glades Road, EE96, Room 308
Boca Raton, FL 33431
561.297.3400

BIOGRAPHICAL SKETCH

Born in Mymensingh, Bangladesh.
B.S., BRAC University, Dhaka, Bangladesh, 2018
M.S., Florida Atlantic University, Boca Raton, FL, USA, 2022

CONCERNING PERIOD OF PREPARATION
& QUALIFYING EXAMINATION

Time in Preparation:

2021-2025

Qualifying Examination Passed:

Fall 2021

Published Papers:

1. R Kanjilal, AL Campelo dos Santos, **SP Arnab**, M DeGiorgio, R Assis. Genomic Anomaly Detection with Functional Data Analysis. *Genes* 16 (6), 710. 2025.
2. **SP Arnab**, AL Campelo dos Santos, M Fumagalli, M DeGiorgio. Efficient detection and characterization of targets of natural selection using transfer learning. *Molecular Biology and Evolution* 42 (5), msaf094. 2025.
3. MR Amin, M Hasan, **SP Arnab**, M DeGiorgio. Tensor Decomposition-based Feature Extraction and Classification to Detect Natural Selection from Genomic Data. *Molecular Biology and Evolution* 40 (10), msad216. 2023.
4. **SP Arnab**, MR Amin, M DeGiorgio. Uncovering footprints of natural selection through spectral analysis of genomic summary statistics. *Molecular Biology and Evolution* 40 (7), msad157. 2023.