

Title: *A Bayesian method for time reannotation of transcriptomics data*

Background: *Transcriptomics experiments are often performed to capture changes in gene expression over time. However, time annotations may be missing, imprecise, or not reflect the same biological phase. Assigning accurate time points to these experiments by using a reference model is crucial for identifying differentially expressed genes, and understanding gene regulatory networks in order to elucidate the studied organism's physiology and life cycle.*

Method: *In this study, we propose a Bayesian approach based on Gaussian process regression modeling to address this challenge. We employ this method to perform time annotation in legacy *Clostridium botulinum* microarray experiments, which were initially annotated based on growth phases, utilizing recently collected RNA-Seq time series data comprising multiple replicates as reference. We also test the performance of the method on RNA-Seq data by using the experiments collected on even time points as the training set and the rest as the validation set. Furthermore, we assess the method's robustness to measurement errors by applying it to synthetically generated data with varying levels of noise.*

Results: *By reassigning the growth phases to the microarray experiments based on the new time annotations, incorporating heuristic knowledge about the expected time interval for each growth phase, we significantly enhance the description of the microarray data. Notably, the improved annotation allows for clear separation of experiments belonging to different growth phases, as demonstrated by principal component analysis (PCA). Consequently, we successfully identify experiments that appear likely to have been initially assigned to incorrect growth phases.*