I hope this email finds you well.

I am in the process of preparing a manuscript for submission to PNAS. The paper presents three novel types of dissimilarity measure that leverage moment and sparsity differences. This measure can be applied across various omics datasets, offering a powerful tool for uncovering the pivotal factors that distinguish different groups.

Enclosed is the abstract for your perusal:

Matrix dissimilarities based on differences in moments and sparsity

Generating a dissimilarity matrix is typically the first step in big data analysis. Although numerous methods exist, such as Euclidean distance, Minkowski distance, Manhattan distance, Bray–Curtis dissimilarity, Jaccard similarity and Dice dissimilarity, it remains unclear which factors drive dissimilarity between groups. In this paper, we introduce an approach based on differences in moments and sparsity. We show that this method can delineate the key factors underlying group differences. For example, in biology, mean dissimilarity indicates differences driven by up/down-regulated gene expressions, standard deviation dissimilarity reflects the heterogeneity of response to treatment, and sparsity dissimilarity corresponds to differences prompted by the activation/silence of genes. Through extensive reanalysis of transcriptome, proteome, metabolome, immune profiling, microbiome, and social science datasets, we demonstrate insights not captured in previous studies. For instance, without considering metadata such as age, BMI, sex, or biomarkers, it is feasible to predict COVID-19 patient mortality based solely on matrix dissimilarities observed during the first week with high accuracy.

Should you choose to review and contribute to that paper, I would be honored to include you as an author.

Your time and consideration are deeply appreciated. I am eager to receive your insights and perhaps engage in a fruitful collaboration that will enhance the impact of our shared research endeavors.

Sincerely,