Left column (deleted):

1. Correlation coefficients are widely used to identify patterns in data that may be of particular interest.

2. In transcriptomics, genes with correlated expression often share functions or are part of disease-relevant biological processes.

3. Here we introduce the Clustermatch Correlation Coefficient (CCC), an efficient, easy-to-use and not-only-linear coefficient based on machine learning models.

4. CCC reveals biologically meaningful linear and nonlinear patterns missed by standard, linear-only correlation coefficients.

5. CCC captures general patterns in data by comparing clustering solutions while being much faster than state-of-the-art coefficients such as the Maximal Information Coefficient.

6. When applied to human gene expression data, CCC identifies robust linear relationships while detecting nonlinear patterns associated, for example, with sex differences that are not captured by linear-only coefficients.

7. Gene pairs highly ranked by CCC were enriched for interactions in integrated networks built from protein-protein interaction, transcription factor regulation, and chemical and genetic perturbations, suggesting that CCC could detect functional relationships that linear-only methods missed.

8. CCC is a highly-efficient, next-generation not-only-linear correlation coefficient that can readily be applied to genome-scale data and other domains across different data types.

Right column (added):

1. This paper presents the Clustermatch Correlation Coefficient (CCC), an efficient and not-only-linear correlation coefficient based on machine learning models, to identify linear and nonlinear patterns in transcriptomics data.

2. We aim to determine if CCC can detect meaningful linear and nonlinear relationships in gene expression data, including those missed by linear-only correlation coefficients, and if highly-ranked gene pairs by CCC are enriched for interactions in integrated networks.

3. When applied to human gene expression data, CCC identifies robust linear relationships and nonlinear patterns associated with sex differences.

4. Our results suggest that CCC can detect functional relationships not captured by linear-only methods.

5. CCC is a highly-efficient, next-generation not-only-linear correlation coefficient that can be applied to genome-scale data and other domains across different data types.