

JMnorm: a novel Joint Multi-feature normalization method for integrative and comparative epigenomics



Guanjue Xiang, Yuchun Guo, David Bumcrot, and Alla Sigova

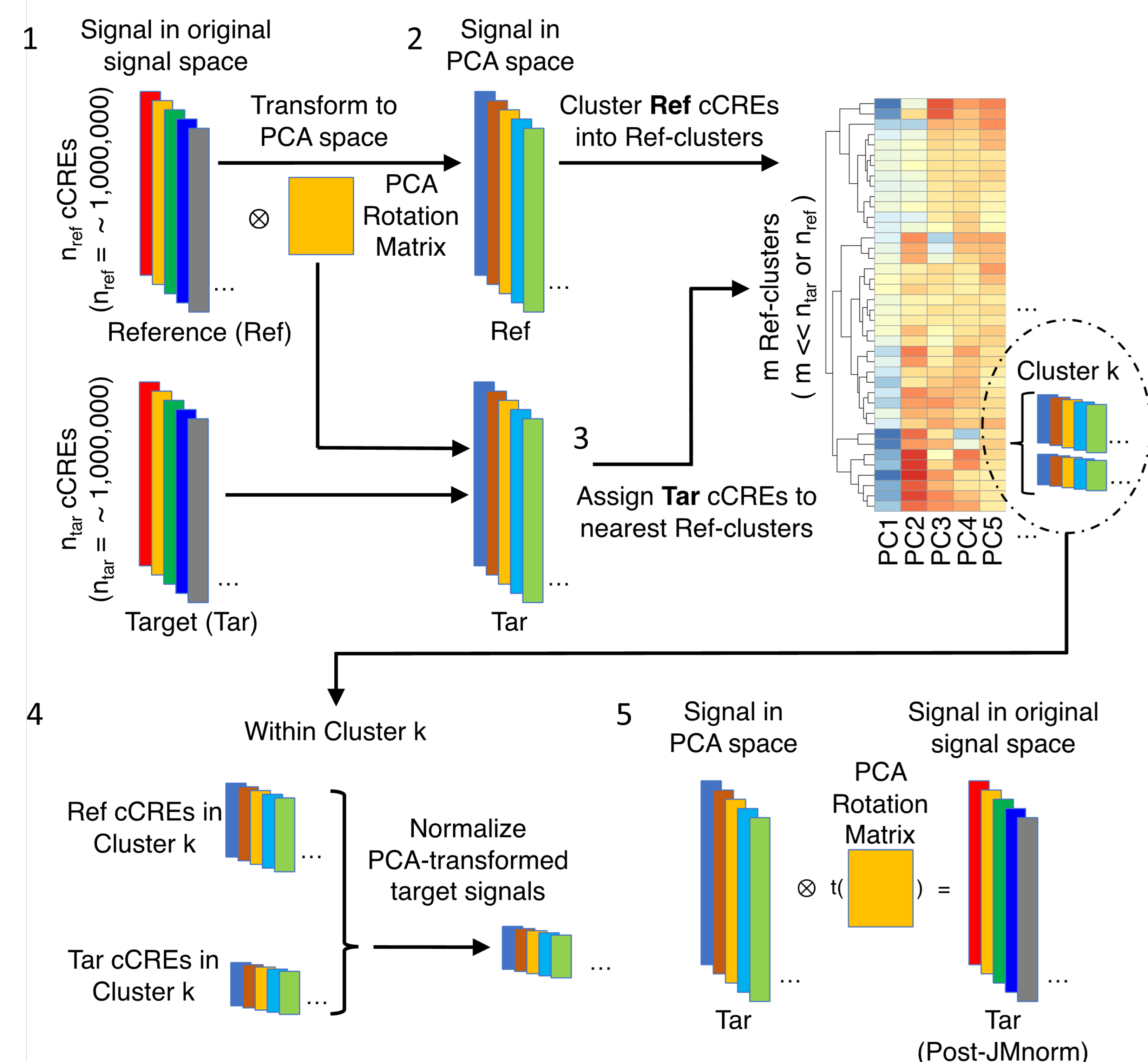
CAMP4 Therapeutics Corp., Cambridge, MA

Motivations

- Chromatin features provide important information in integrative and comparative epigenomic studies and exhibit correlated relationships.
- Robust normalization methods are required for comparison and integration of epigenomic data among different cell types and conditions.
- Existing normalization methods ignore cross-feature correlations and are inadequate for highly variable signal distributions that exist in different cell types.

Method overview

JMnorm simultaneously normalizes multiple epigenetic features across cell types/conditions by leveraging information from partially correlated features



Step 1 - Transform multi-dimension epigenomic signal matrices into orthogonal Principal Component Analysis (PCA) spaces.

Step 2 - Generate reference candidate cis-regulatory element (cCRE) clusters based on reference data in PCA space.

Step 3 - Assign target cCRE into reference clusters based on target data in PCA space

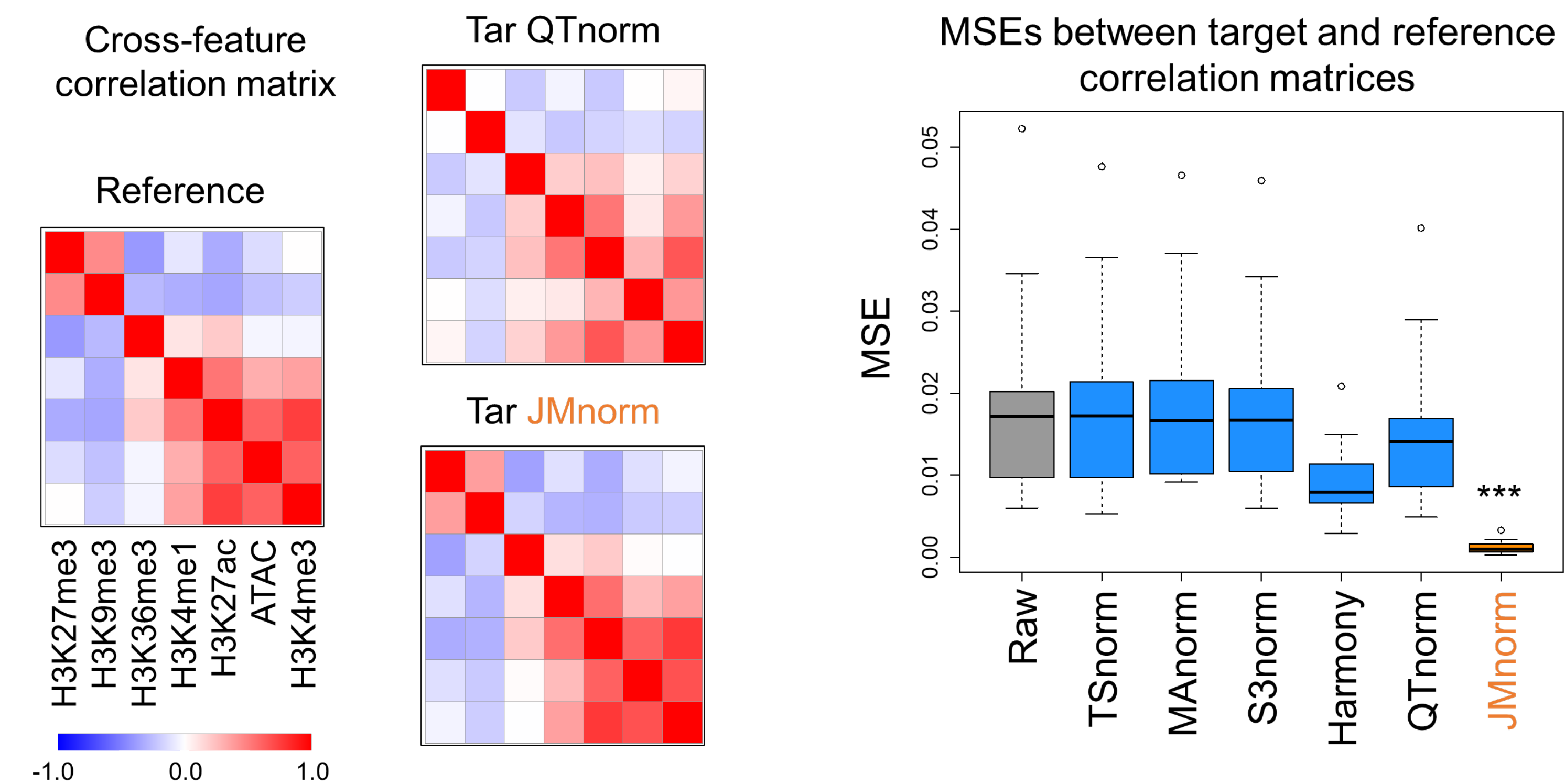
Step 4 - Within each cluster in PCA space, apply quantile normalization to normalize target data against reference data.

Step 5 - Reconstruct JMnorm target signal matrix by transforming the normalized target signal in PCA space back to the original signal space.

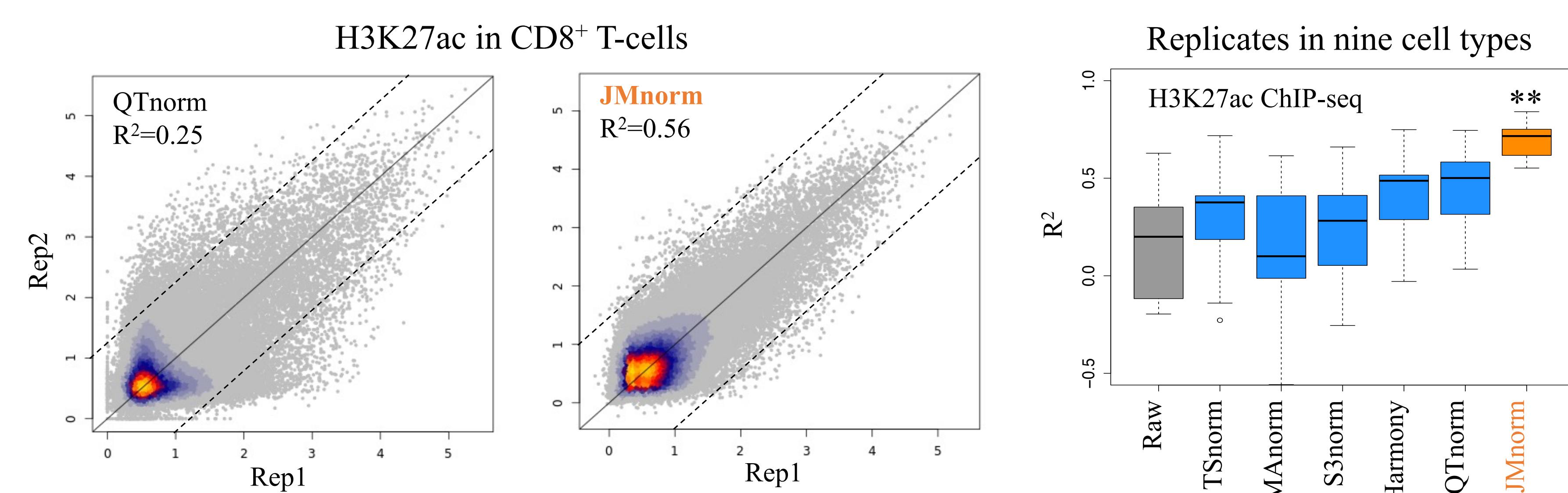
JMnorm outperforms existing normalization methods

- JMnorm better preserves cross-feature correlations.
- JMnorm improves signal consistency between biological replicates.
- JMnorm improves cross-feature combinatorial patterns and cross-cell-type RNA-seq predictions.
- JMnorm improves cross-cell-type differential DNase-seq peak calling.
- JMnorm improves differential glucocorticoid receptor (GR) peak calling in response to dexamethasone treatment.

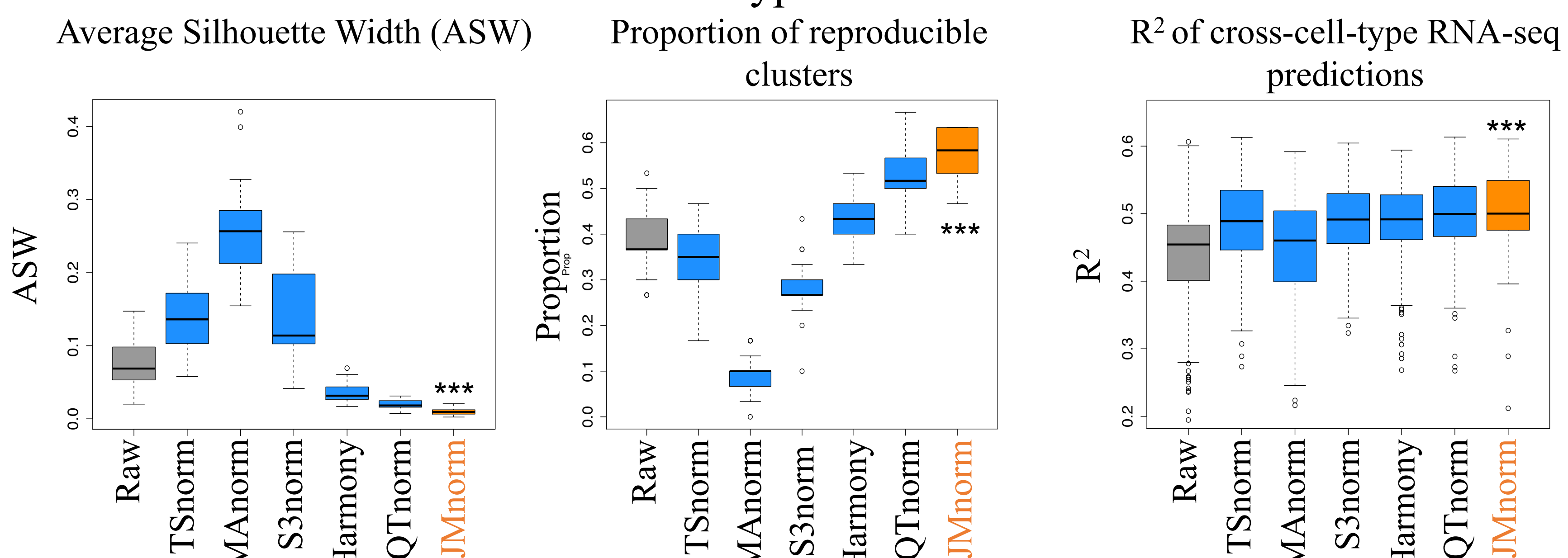
(A) JMnorm better preserves cross-feature correlations



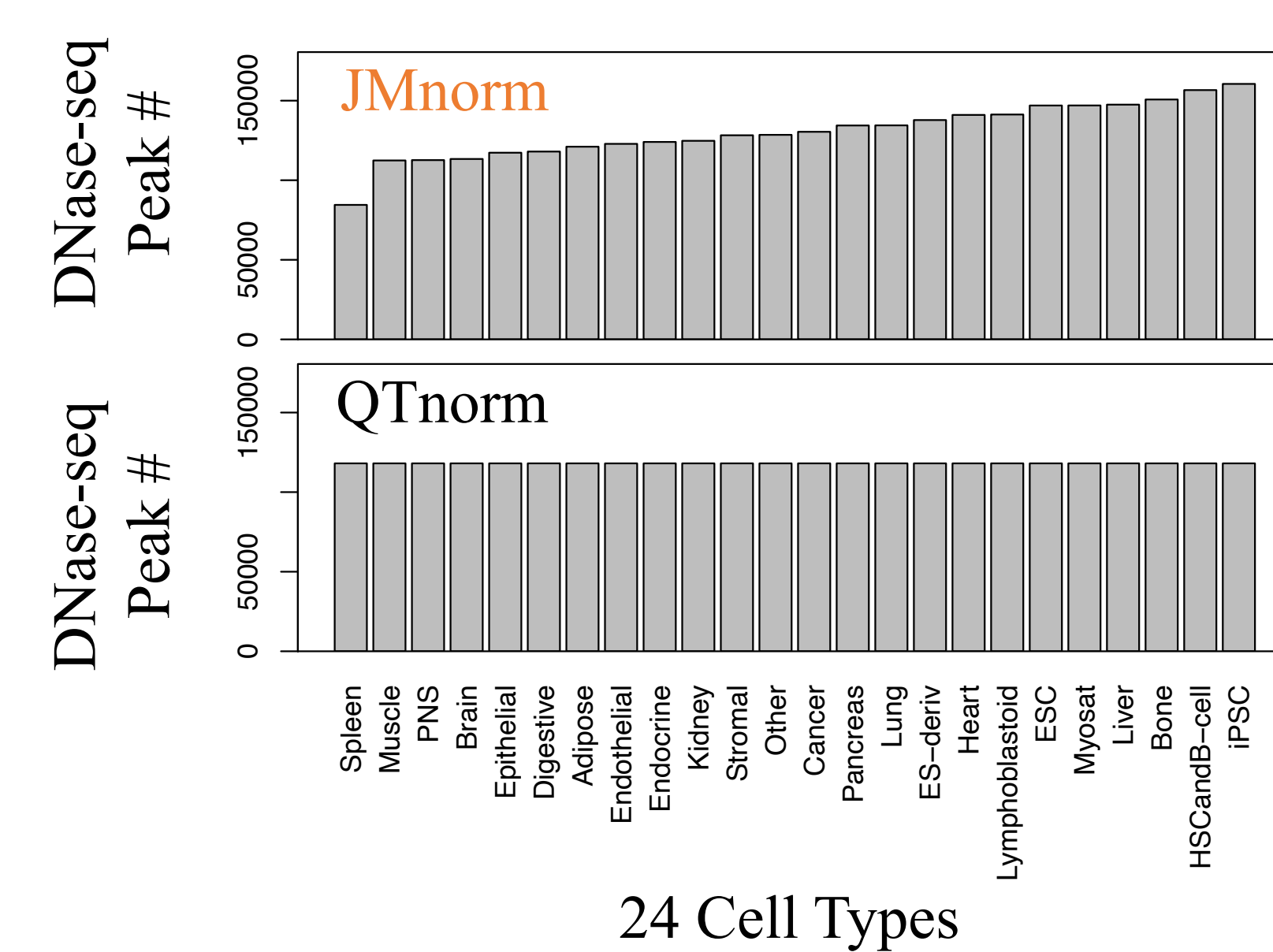
(B) JMnorm improves signal consistency between biological replicates



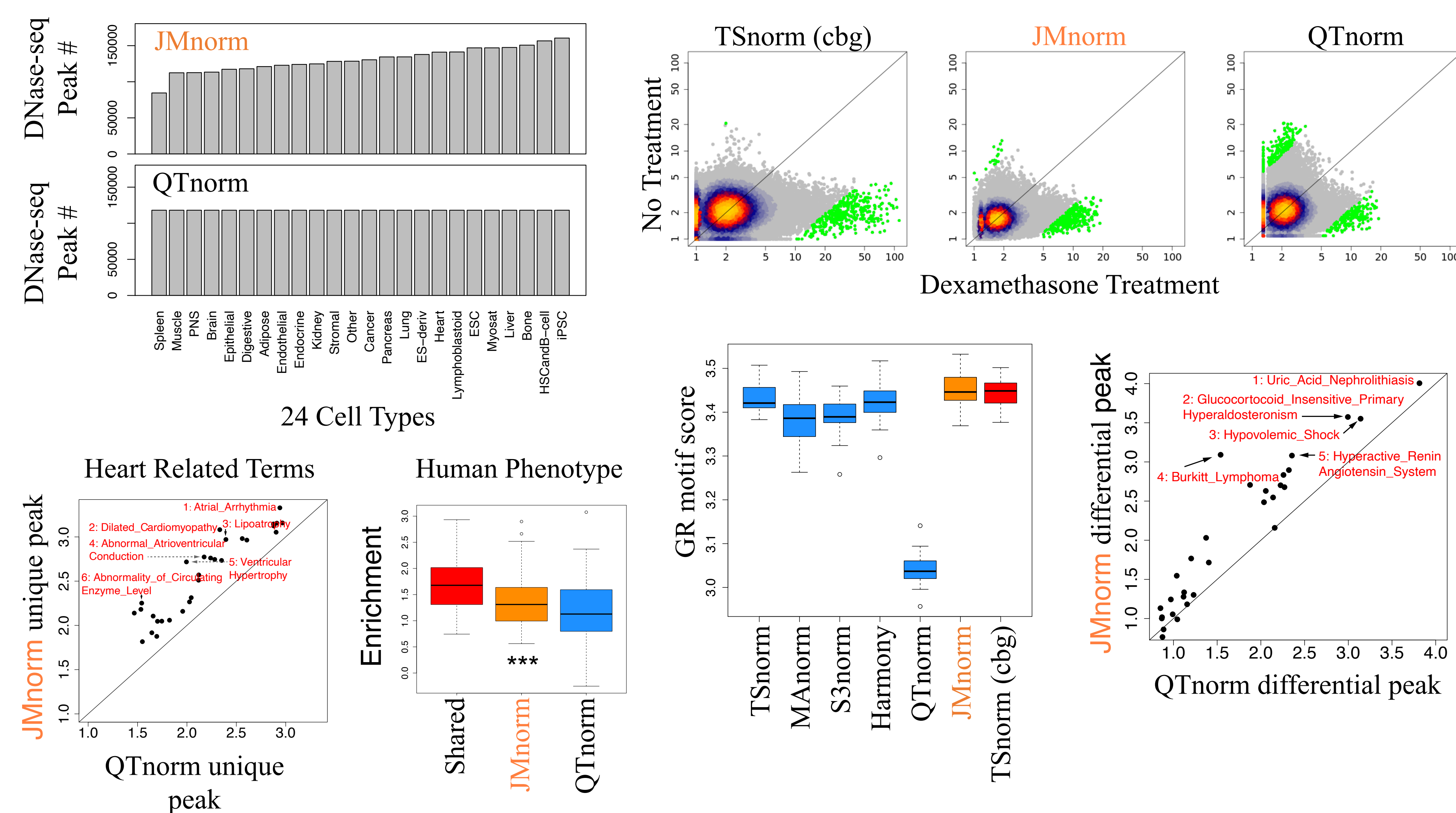
(C) JMnorm improves combinatorial pattern transferring and RNA-seq predictions across cell types



(D) JMnorm allows more accurate open chromatin peak calling



(E) JMnorm improves differential GR peak calling in response to dexamethasone treatment



JMnorm is available at GitHub (<https://github.com/camp4tx/JMnorm>) under GNU General Public License, version 2.0 or later.