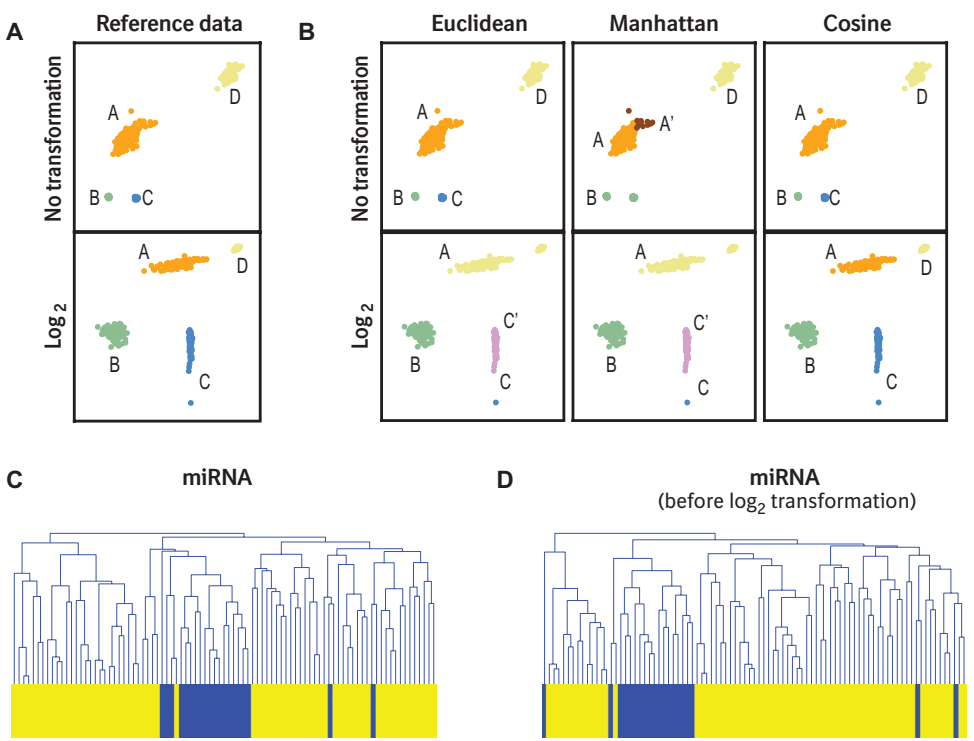

Fig. 3. Effect of transformations and distance metric on clustering results. (**A**) Demonstration of how transformations affect the relationship of data points in space. A toy data set (reference set, https://github.com/knaegle/clusteringReview) was clustered into four clusters with agglomerative clustering, average linkage, and Euclidean distance. The four reference clusters without transformation (upper panel) and after $\log_2$ transformation (lower panel). (**B**) Transformations and distance metrics change clustering results when compared to the reference clustering result. With no transformation (upper panels), Euclidean and cosine distance do not change cluster identity, but with Manhattan distance, a new cluster A′ is added, and cluster C is merged into cluster B. With the $\log_2$ transformation (lower panels), the Euclidean and Manhattan metrics caused cluster C′ to emerge and cluster D to be lost. (**C**) Dendrogram from the microRNA (miRNA) clustering experiment result from 89 cell lines and 217 microRNAs (*6*). Gastrointestinal-derived cell lines (blue bars) predominantly cluster together in the full-dimensional space. Note: The data were $\log_2$-transformed as part of the preclustering analysis. (**D**) Same microRNA data as in (C) but without $\log_2$ transformation.