



**Fig. 2. Dimensionality reduction methods and effects.** Comparison of PCA and subspace clustering. **(A)** Three clusters are plotted in two dimensions. The dashed red line indicates the one-dimensional line upon which the original two-dimensional data is projected as determined by PCA. **(B)** The clusters are plotted in the new single dimension after reducing the dimensionality from two to one. **(C)** Three alternate one-dimensional projections (dashed red lines) onto which the data can be projected, each demonstrating better separability for some clusters than the projection identified using PCA. **(D to F)** Comparison of the original clustering results of 89 cell lines in ~14,000-dimensional mRNA data (D) to clustering results after PCA (E) and after subspace clustering (F). Blue bars, gastrointestinal cell lines; yellow bars, nongastrointestinal cell lines.