Feature extraction from 2D TOF-SIMS data of biological samples by using multivariate analysis

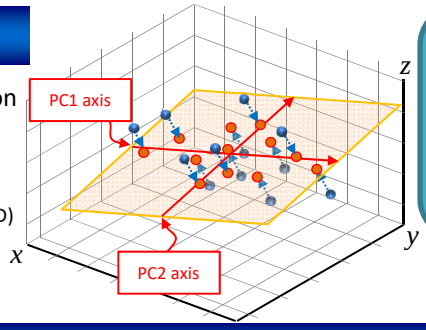
TOF-SIMS data obtained from biological samples, whose composition is complicated, are generally huge and difficult to interpret. By using multivariate analysis (e.g. Principal Component Analysis: PCA), the features of original TOF-SIMS data can be extracted as easy-to-interpret data.

The principle of PCA

(Ex.) Dimensionality reduction to 2D from 3D

● : Original data (3 dimensions)

● : Score (projection to 2D from 3D)

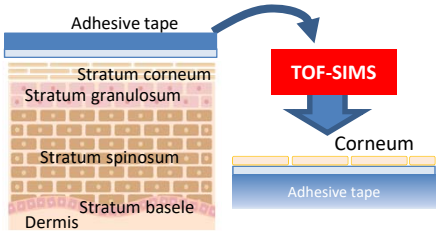


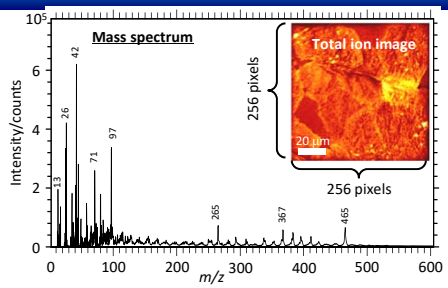
- The direction that accounts for the largest amount of variance in the original data is set as the 1st axis.
- The direction, orthogonal to the 1st axis, that accounts for the largest amount of remaining variance is set as the 2nd axis.
- If it were a higher-dimensional data set, PCA would also set a 3rd axis, orthogonal to both previous axes, and a 4th, a 5th, and so on – as many axes as the number of dimensions in the data set.

⇒ Eigenvalue problem of covarianace matrix

The application of PCA to TOF-SIMS image data of Stratum corneum

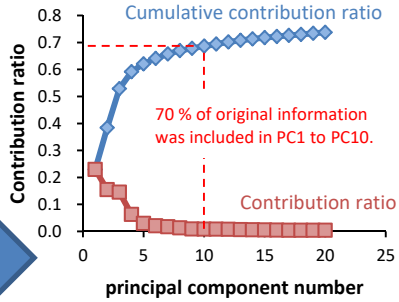
◆ TOF-SIMS imaging of corneum sampled by tape-stripping





○Size of the data analyzed by PCA:
468 peaks × 256 pixels × 256 pixels

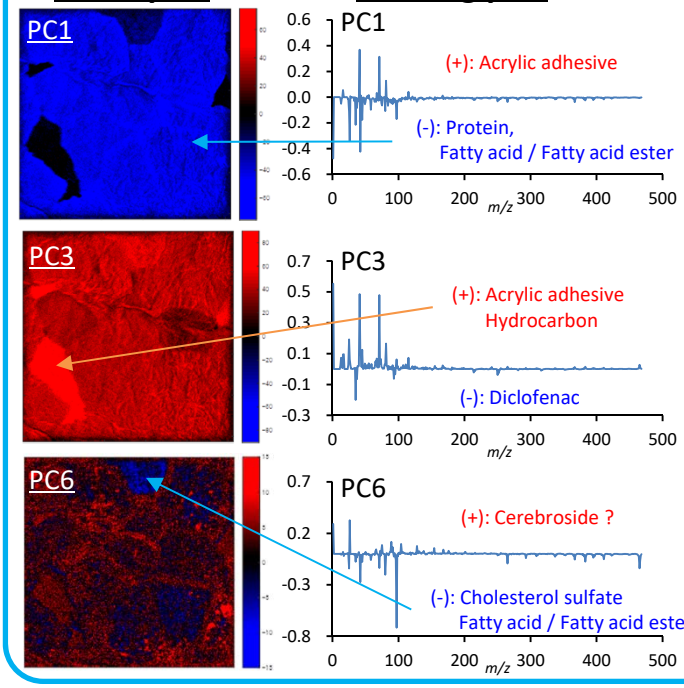
PCA



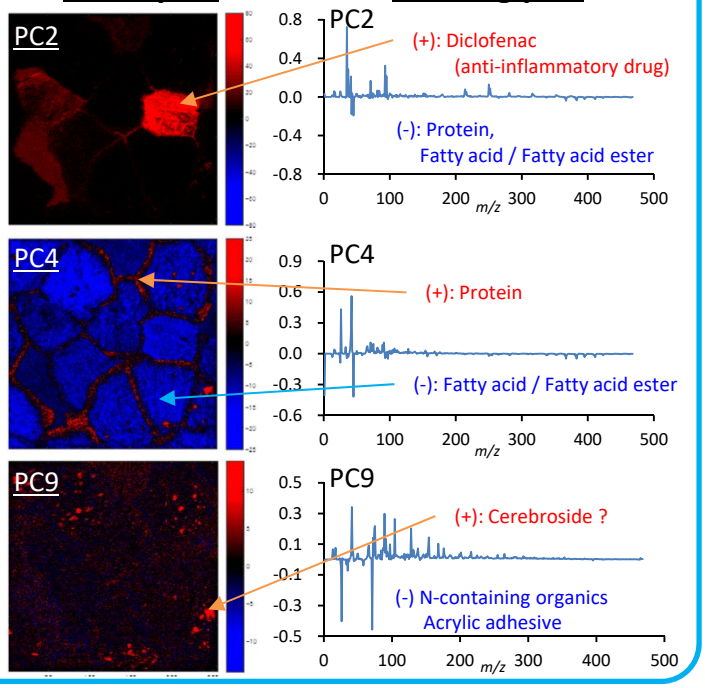
10 principal components (PCs), which is easy to interpret, were extracted from 468 images.

Dimensionality reduction
468 ⇒ 10 dimensions

Score plot / Loading plot



Score plot / Loading plot



The characteristic distribution of drug (diclofenac), intercellular lipids (fatty acid, cholesterol sulfate, etc.), and exposed acrylic adhesive were extracted by conducting PCA.

Toray Research Center, Inc.