

Description of Additional Supplementary Files

File name: Supplementary Data 1

Description: Dataset 1

This dataset was used to generate the figure of the SNV-normalized spectra (Figure 1b).

File name: Supplementary Data 2

Description: Dataset 2

This dataset was used on fitting a PLS-DA model to classify lymphoma versus non-malignant tissue. A PLS-DA model is trained using five components, and predictions are made on the full dataset, training, and test sets, with a classification threshold of 0.450. The model's performance is evaluated through bootstrapping techniques, calculating key metrics like ROC-AUC and generating visualizations such as confusion matrices and ROC curves. Moreover, significant different wavenumbers between both groups are identified based on this dataset.

File name: Supplementary Data 3

Description: Dataset 3

This dataset was used on trying to differentiate distinct non-Hodgkin lymphoma subtypes from each other on the training set using PLS-DA models. Additionally, this dataset was used on trying to differentiate non-hodgkin lymphoma subtypes from Hodgkin lymphoma using PLS-DA models.