Supplementary material for

High-throughput Raman spectroscopy allows *ex vivo* characterization of murine small intestinal intra-epithelial lymphocytes (IEL)

Rustam R. Guliev¹, Tina Vogler^{2,5}, Natalie Arend¹, Simone Eiserloh^{1,3}, Alexander Wiede¹, Timo Kunert^{2,5}, Martin Dinkel^{2,5}, Jürgen Popp^{1,3,4}, Iwan W. Schie^{1,6}, Kai Hildner^{2,5#}, Ute Neugebauer^{1,3,4#*}

¹ Leibniz Institute of Photonic Technology, 07745 Jena, Germany

² University Hospital Erlangen, Medical Department 1, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91054 Erlangen, Germany

 ³ Center for Sepsis Control and Care, Jena University Hospital, 07747 Jena, Germany
⁴ Friedrich Schiller University Jena, Institute of Physical Chemistry and Abbe Center of Photonics, 07743 Jena, Germany

⁵ Deutsches Zentrum Immuntherapie (DZI), University Hospital Erlangen, Erlangen, Germany
⁶ University of Applied Sciences Jena, Department for Medical Engineering and
Biotechnology, 07745 Jena

shared senior authorship

* corresponding author: Ute Neugebauer Email: ute.neugebauer@uni-jena.de

Content:

Supplemental Figure S1: Scores of PCA models

Supplemental Figure S2: 2D scatter-plot of PC1 and PC2 split by mouse

Supplemental Figure S3: Comparison of PC1 loadings for the models shown in Figure S1

Supplemental Figure S4: Computed difference spectra of group 2 and group 1 mean spectra

Supplemental Figure S5: Ratio of Raman bands intensities at 788 cm⁻¹ and 2894 cm⁻¹

Supplemental Figure S6: Multivariate Curve Resolution-Alternating Least Squares (MCR-ALS) analysis of Raman data

Supplemental Figure S7: Dendrogram from HCA

Supplemental Figure S8: Details of cross-validation performance of the LDA model

Supplemental Figure S9: Accuracy of PCA-LDA model trained using different number PCs

Supplemental Table S1: Clarification of resamples in cross-validation



Supplemental Figure S1-A: PCA score plot for spectra originating from cells isolated from group 1 (with T-cell transfer) mice (M1-M3). The curves in the diagonal frames show density estimates. These are simply smoothed histogram plots allowing better visual comparison between different distributions. Non-diagonal frames show 2D scatter plots of the corresponding PC scores in X- (column) and Y- (row) axes. All plots are split by mouse using different colors to demonstrate that some patters are animal specific (M1: dark green, M2: turquoise, M3: green-yellow).



Supplemental Figure S1-B: PCA score plot for spectra originating from cells isolated from group 2 (control) mice (M4-M6). The curves in the diagonal frames show density estimates. These are simply smoothed histogram plots allowing better visual comparison between different distributions. Nondiagonal frames show 2D scatter plots of the corresponding PC scores in X- (column) and Y- (row) axes. All plots are split by mouse using different colors to demonstrate is some patters are animal specific (M4: orange, M5: red, M6: violet).



Supplemental Figure S1-C: PCA scores plot for all pre-processed spectra from both treatment groups (M1-M6). The curves in the diagonal frames show density estimates. These are simply smoothed histogram plots allowing better visual comparison between different distributions. Non-diagonal frames show 2D scatter plots of the corresponding PC scores in X- (column) and Y- (row) axes. All plots are split by mouse using different colors to demonstrate is some patters are animal specific. For visual convenience, mice from the same group are colored by close colors: green-like colors for group 1 (with T-cell transfer) and red-like colors for group 2 (control).



Supplemental Figure S2: 2D scatter-plot of the PC1 and PC2, when applied to all pre-processed spectra, split by mouse. The points are colored with respect to mouse number (with T cell transfer: M1: dark green, M2: turquoise, M3: green-yellow; control: M4: orange, M5: red, M6: violet). For visual convenience, mice from the same group are colored by close colors: green-like colors for group 1 (with T-cell transfer) and red-like colors for group 2 (control).



Wavenumber / cm⁻¹

Supplemental Figure S3: Loadings plots. The lines illustrate the loadings of PC1 when PCA was applied to Raman spectra of cells originating only from treatment group 1 (after T cell transfer, green line), only to Raman spectra of cells from treatment group 2 (control, red line), and to Raman spectra of both cells from both groups (all data combined, blue line). Respective scores plots are presented in Supplemental Figure S1 and for combined data also in Figure 3C, main manuscript.



Supplemental Figure S4: Difference spectra of mean group 2 (control) spectrum subtracted from mean group 1 (with T cell transfer) for A) high PC1 (\geq 5) region and B) low PC1 (\leq -15) region.



Supplemental Figure S5: Distribution of ratios of two Raman bands intensities: at 788 cm⁻¹ and 2894 cm⁻¹. Top row corresponds to ratio of intensity at 788 cm⁻¹ divided by intensity 2894 cm⁻¹, and the bottom row to the inverse. The ratio is calculated using pre-processed spectra intensities. The distributions are provided separately for three subsets of data: all pre-processed spectra combined (first column), only subpopulation with higher PC1 scores (PC1 \geq 5; second column), and only subpopulation with lower PC1 scores (PC1 \leq -15; third column). The distributions are colored with respect to a mouse using the same colors as in Supplemental Figures S2-C and S3.



Supplemental Figure S6-1: Scatter plot of contributions of pure components calculated by MCR-ALS. The corresponding pure components spectra are provided in Supplemental Figure S9-2. The curves in the diagonal frames show density estimates. Non-diagonal frames show 2D scatter plots of the corresponding concentrations in X- (column) and Y- (row) axes. The frames below the diagonal to colored with respect to mice using the same color as in other figures. The frames above the diagonal are colored with respect to the PC1 cluster: high PC1 cluster – red, low PC1 cluster – blue.



Supplemental Figure S6-2: Spectra of pure components estimated by MCR-ALS. The components are assigned to, from top to bottom, lipids, nuclei, and water/cytoplasm. The corresponding contributions are provided in Supplemental Figure S9-1.



Supplemental Figure S7: A dendrogram of the hierarchical clustering. The leaves of the dendrogram are colored with respect to a mouse using the same colors as in Supplemental Figures S2-C and S3. The distribution of the spectra of the two clusters with respect to PCA space is given in Figure 5, main manuscript.



Supplemental Figure S8: The prediction of the PCA-LDA model, using 2 principal components, for test sets of each cross-validation resample. The values are grouped by mouse and colored by the cross-validation resample. It can be found, that each resample had two mice, one control and one colitis mice, in its test set. The predictions are shown in a form of a box plot, i.e. a rectangle represents the range from 25 % (Q1) to 75 % (Q3) percentile, the black line inside the rectangle indicates the median (Q2), Whiskers indicate the range from Q1-1.5*IR to Q3+1.5*IR, where IR=Q3-Q1 (interquartile range), and the black dots show the outlier values outside the whiskers range.



Supplemental Figure S9: Classification accuracy of PCA-LDA model trained on A) low PC1 cluster spectra and B) high PC1 cluster spectra using different number of principal components. The values are calculated based on cross-validation training (black) and testing (blue) predictions. The classification accuracy measured by area under ROC curve.

Supplemental Table S1: Resamples used during cross-validation. Each row corresponds to one cross-validation iteration, i.e. resample and each column correspond to a mouse. Text in a cell and the background color indicate in which dataset the corresponding spectra were included: testing or training. Total number of spectra 3377.

CV Resample	M1	M2	M3	M4	M5	M6
	434 spectra	655 spectra	466 spectra	613 spectra	574 spectra	635 spectra
Resample 1	Test	Train	Train	Train	Train	Test
Resample 2	Test	Train	Train	Test	Train	Train
Resample 3	Test	Train	Train	Train	Test	Train
Resample 4	Train	Test	Train	Train	Train	Test
Resample 5	Train	Test	Train	Test	Train	Train
Resample 6	Train	Test	Train	Train	Test	Train
Resample 7	Train	Train	Test	Train	Train	Test
Resample 8	Train	Train	Test	Test	Train	Train
Resample 9	Train	Train	Test	Train	Test	Train