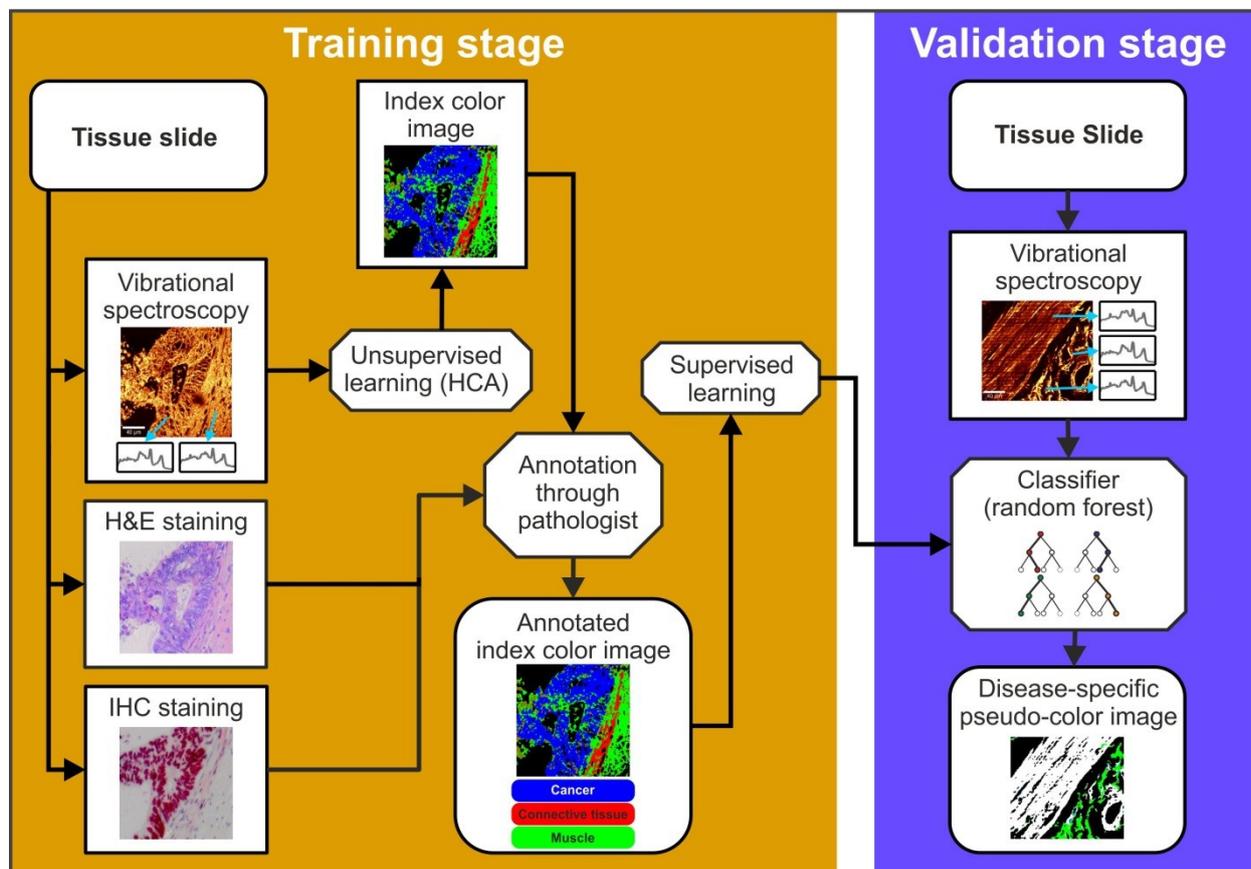
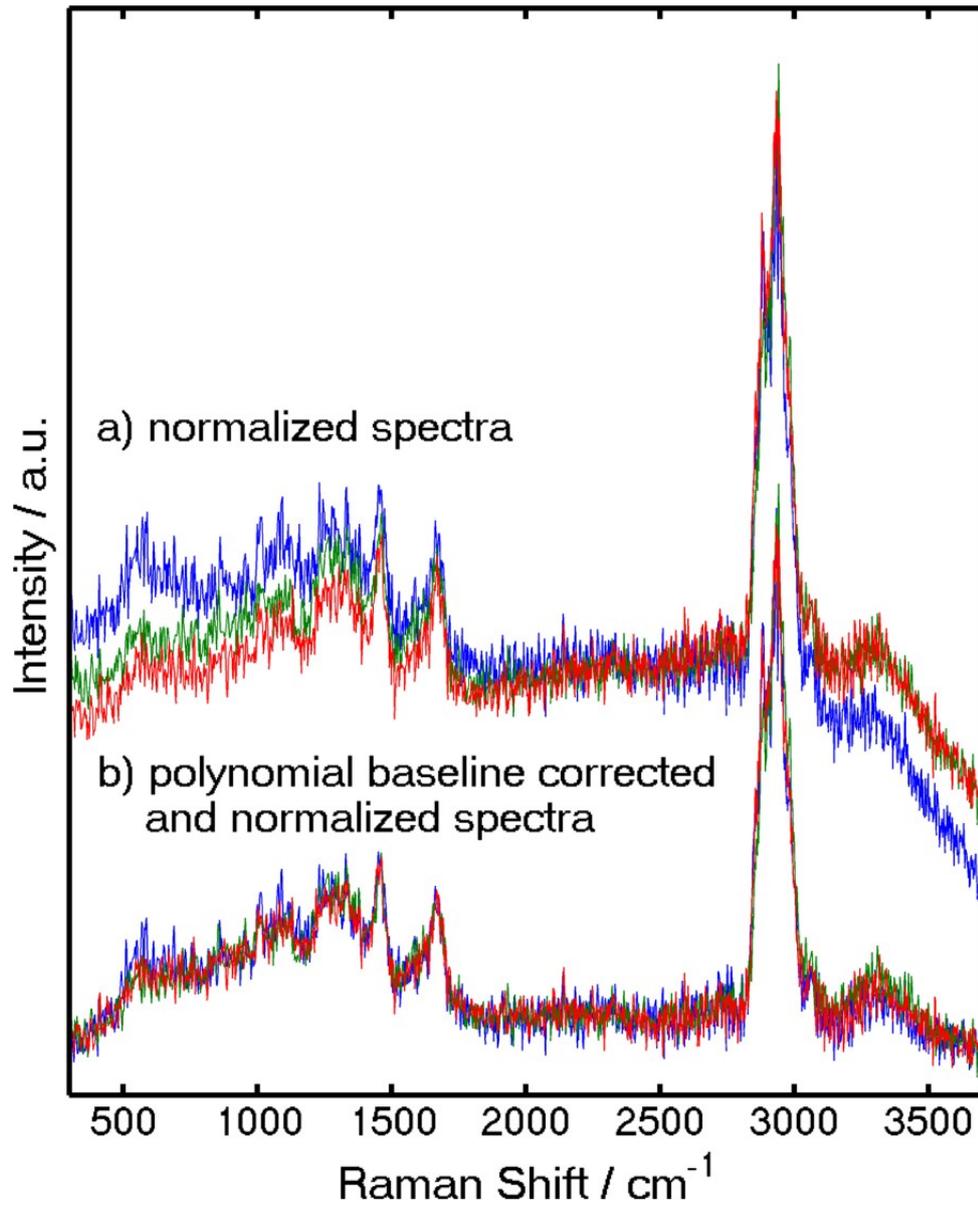


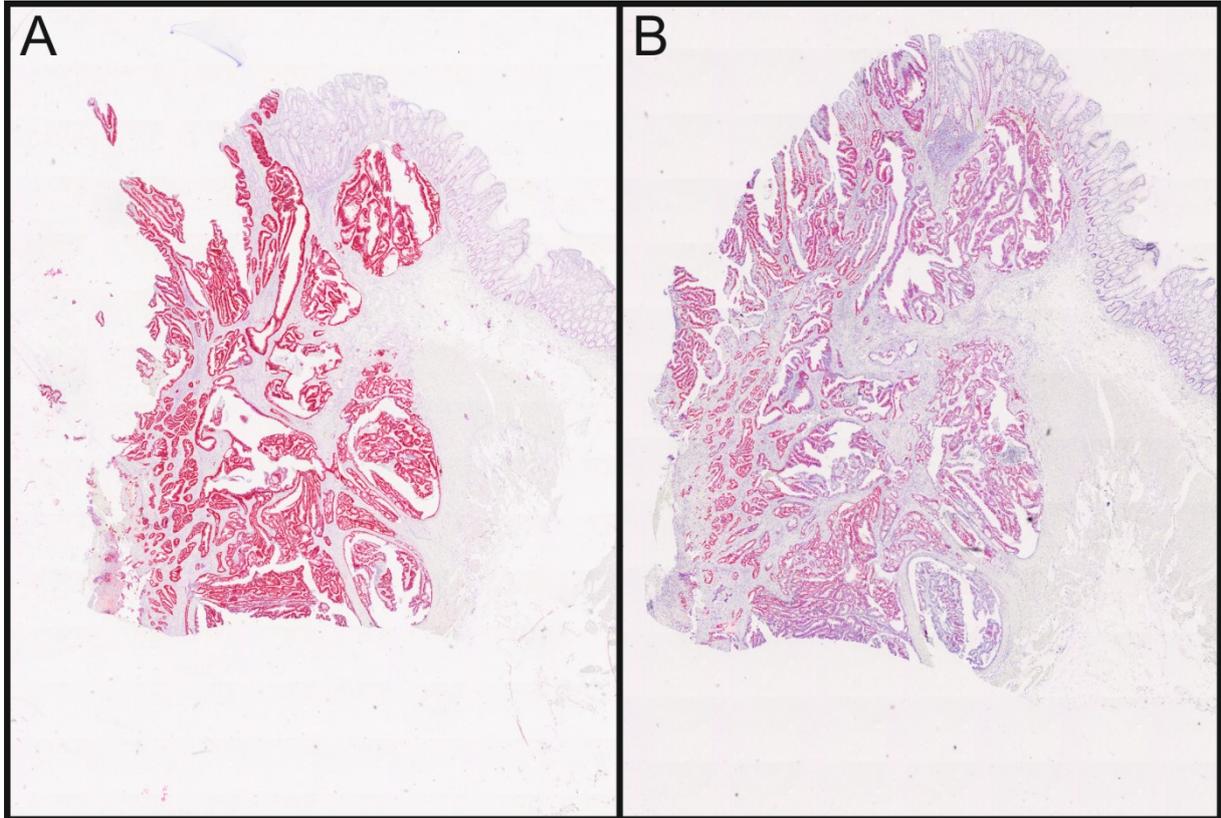
## Supplementary Information



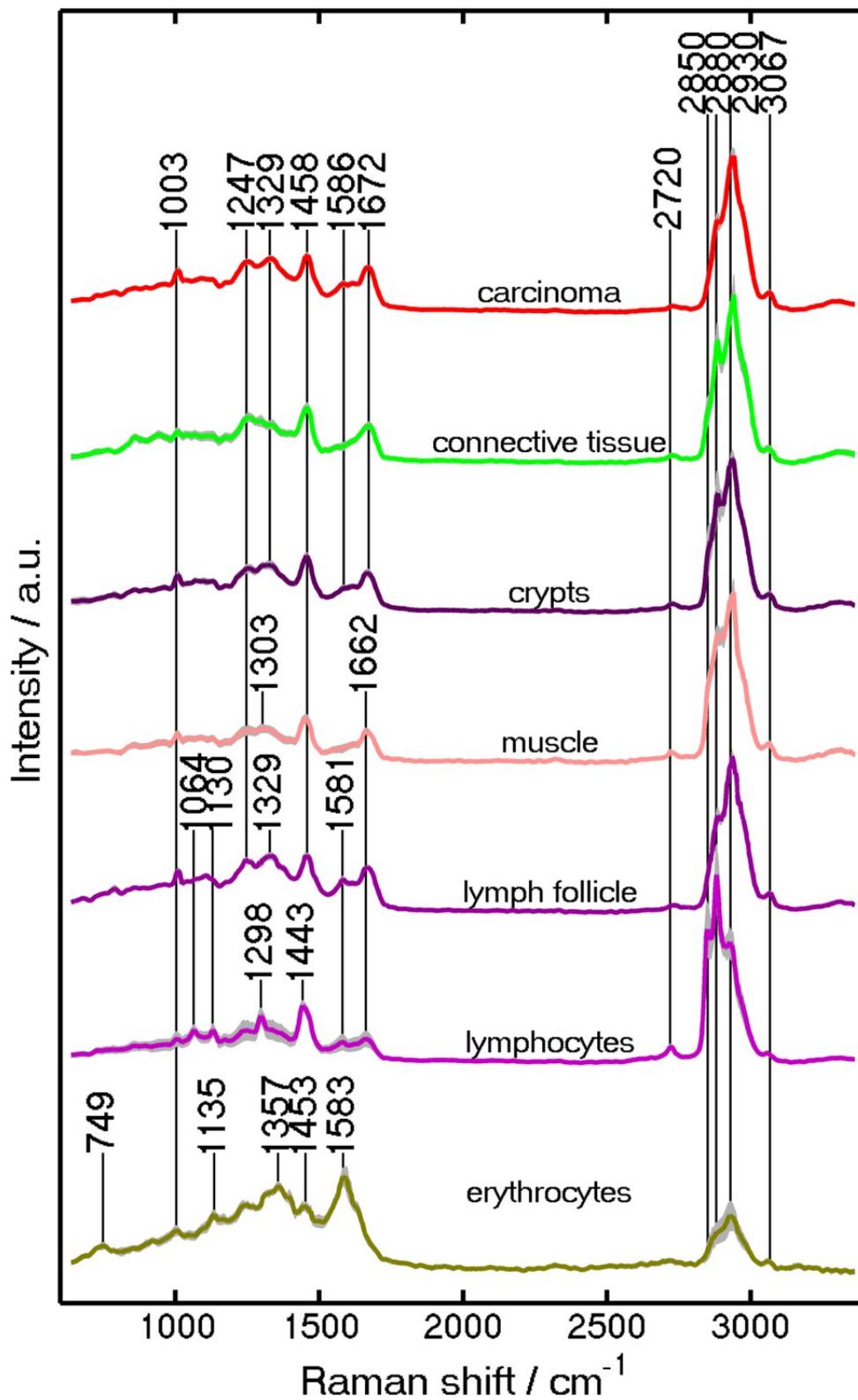
**Fig. S1.** Workflow describing the training and validation stages for Raman based SHP. The pseudo-colour images created by unsupervised learning algorithm are annotated by pathologist with help of H&E and IHC staining. These spectra build up a training data set for a supervised learning algorithm (RF), which can classify other tissue slides automatically.



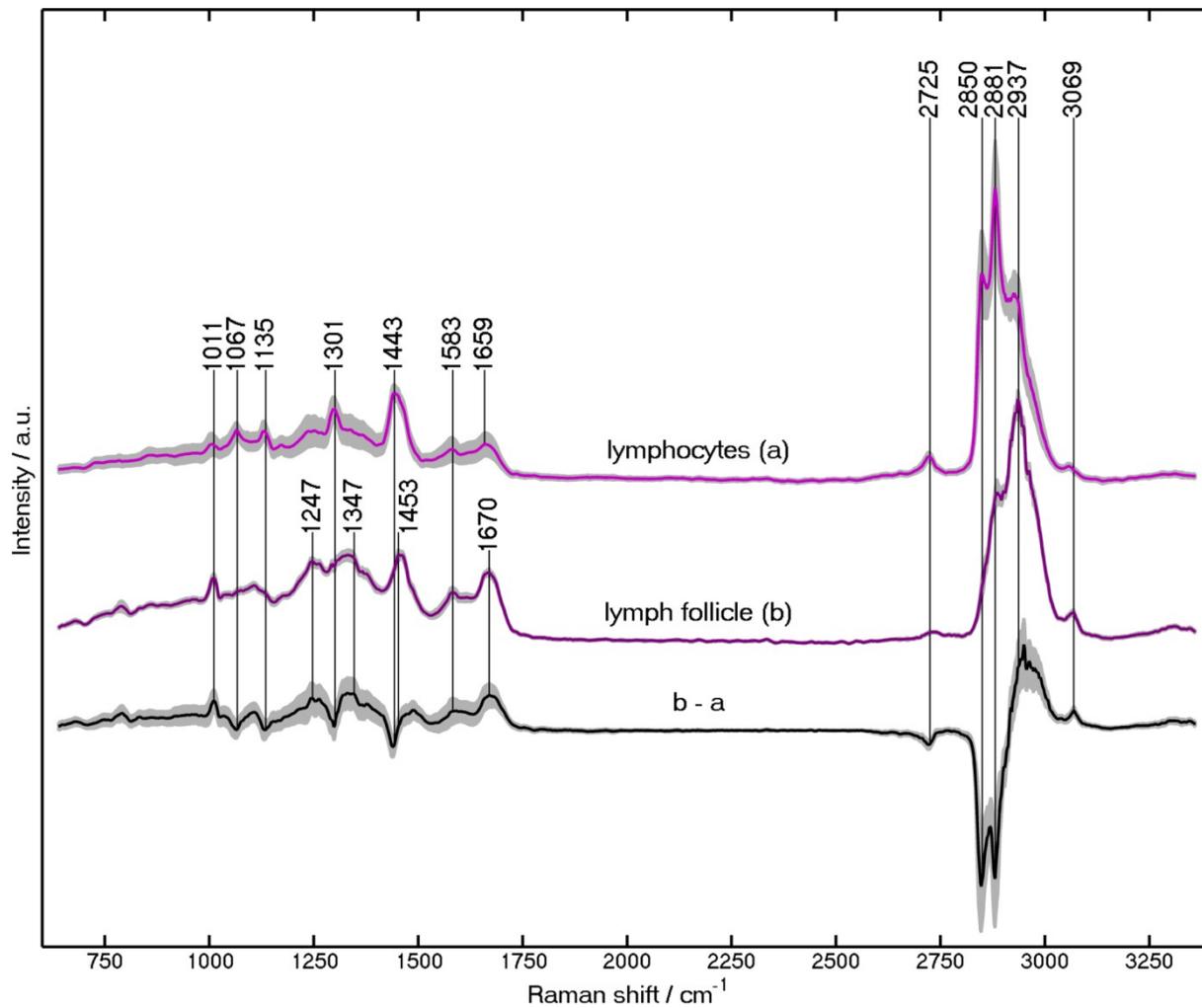
**Fig. S2.** Effect of baseline correction on Raman spectra. Raman spectra of tissue at different pixels without (a) and with baseline correction (b).



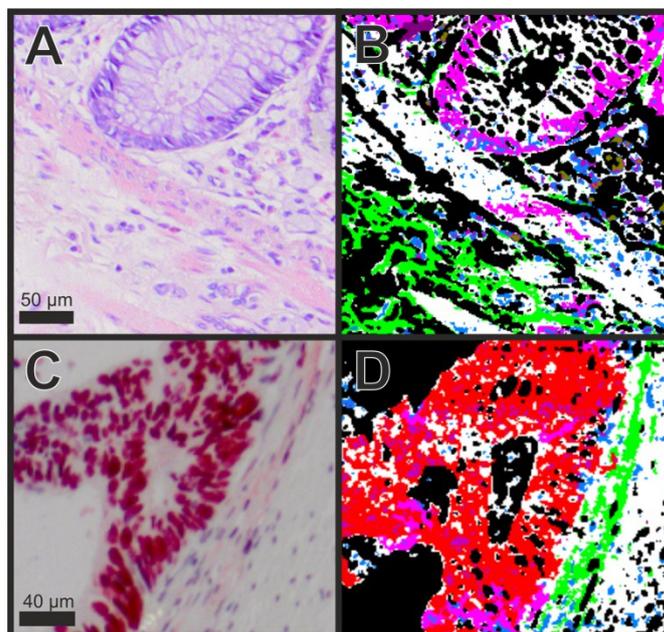
**Fig. S3.** Immunohistochemical staining of colorectal carcinoma tissue slide presented in Fig. 1. (A) Immunohistochemical staining of a colon tissue section by p53 antibody. Accumulation of p53 is shown in red. (B) Immunohistochemical staining of a colon tissue section by MiB-1 (Ki-67) antibody. Proliferating cells are shown in red.



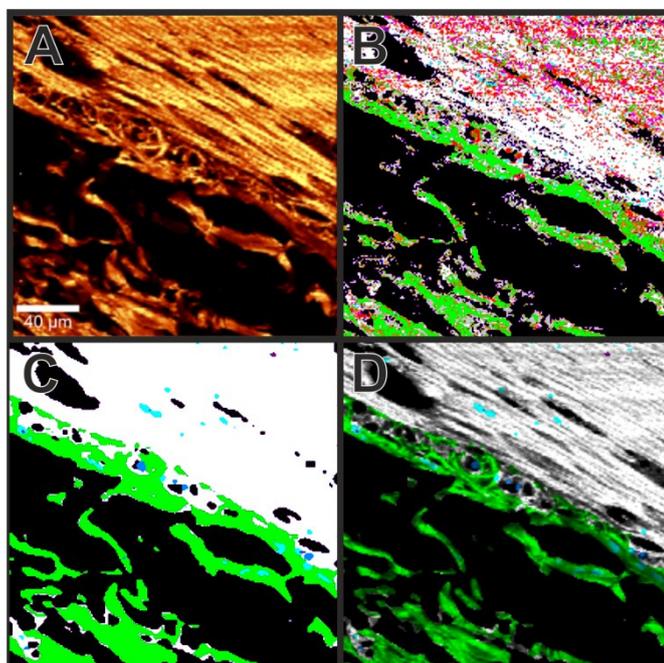
**Fig. S4.** Raman mean spectra of tissue components used in Raman RF classifier. The spectra are wavelet-dennoised and the standard-deviation is marked in grey. The spectra are shown from 700-3500  $\text{cm}^{-1}$ .



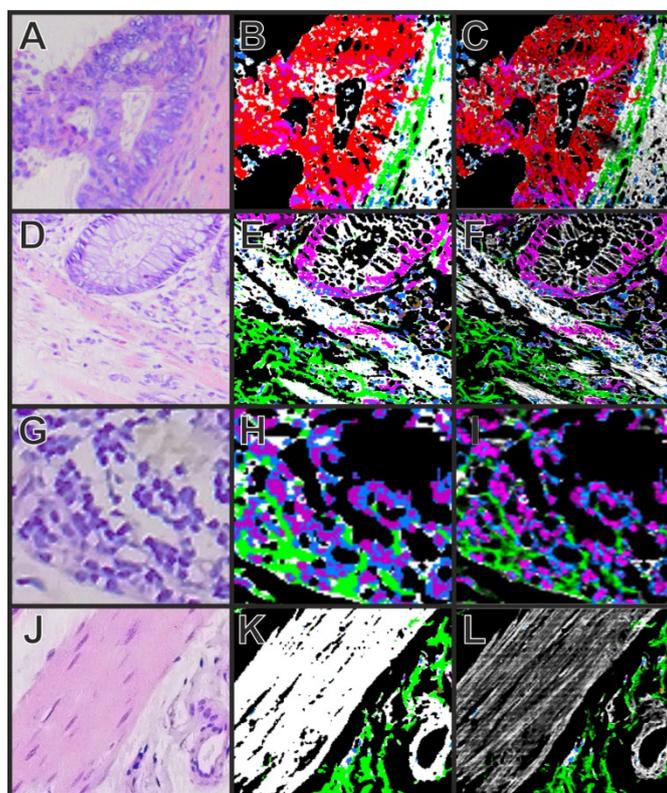
**Fig. S5.** Raman mean spectra of lymphocytes and lymph follicle used in Raman RF. The spectra are wavelet-denoised and the standard-deviation is marked in grey. The spectra are shown from 700-3500 cm<sup>-1</sup>. The difference spectrum is showed in black.



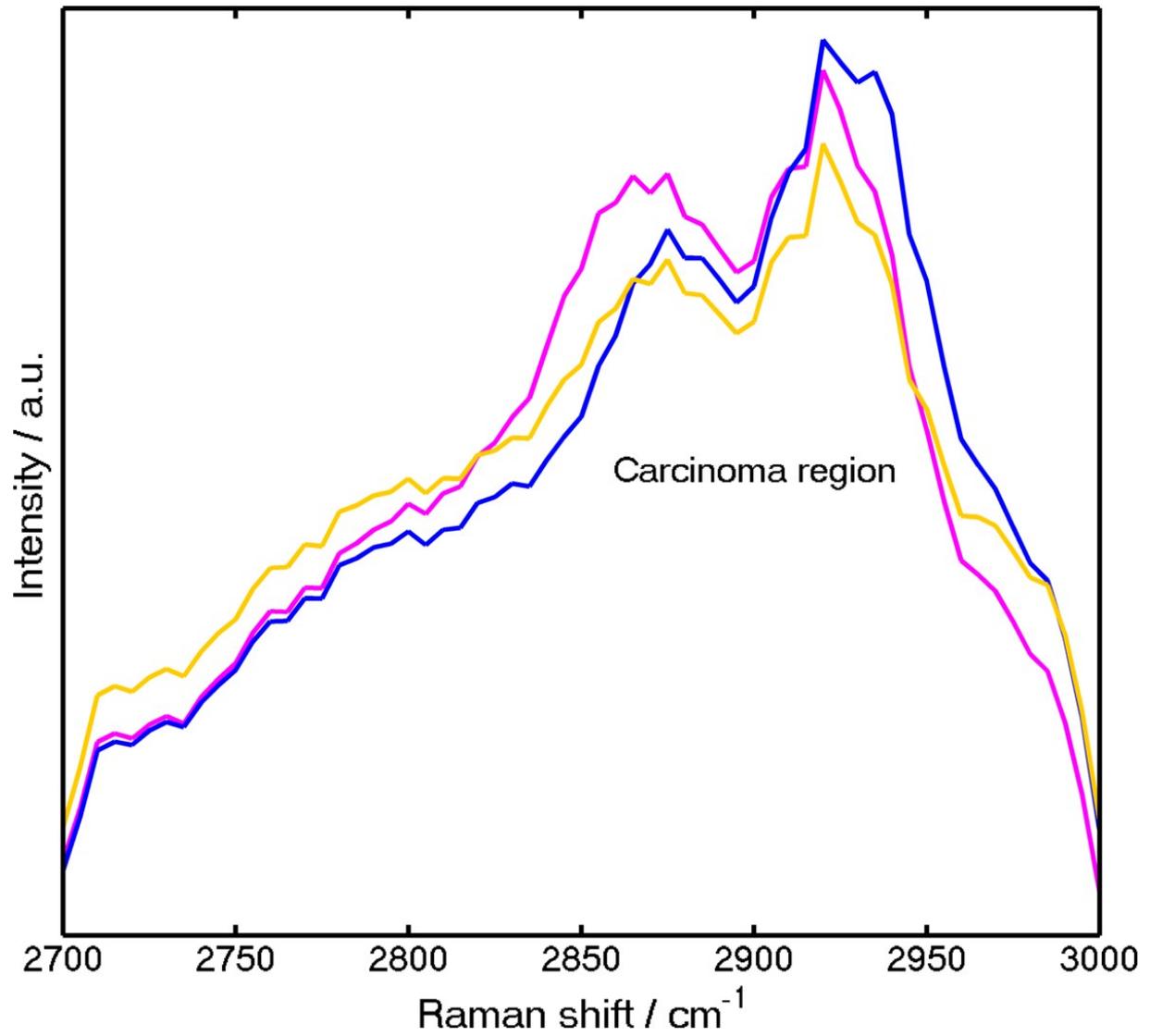
**Fig. S6.** Alternative representation for Fig. 3. Direct comparison of H&E and immunohistochemical staining with Raman virtual staining of selected regions from colorectal carcinoma tissue slide presented in Fig. 2.



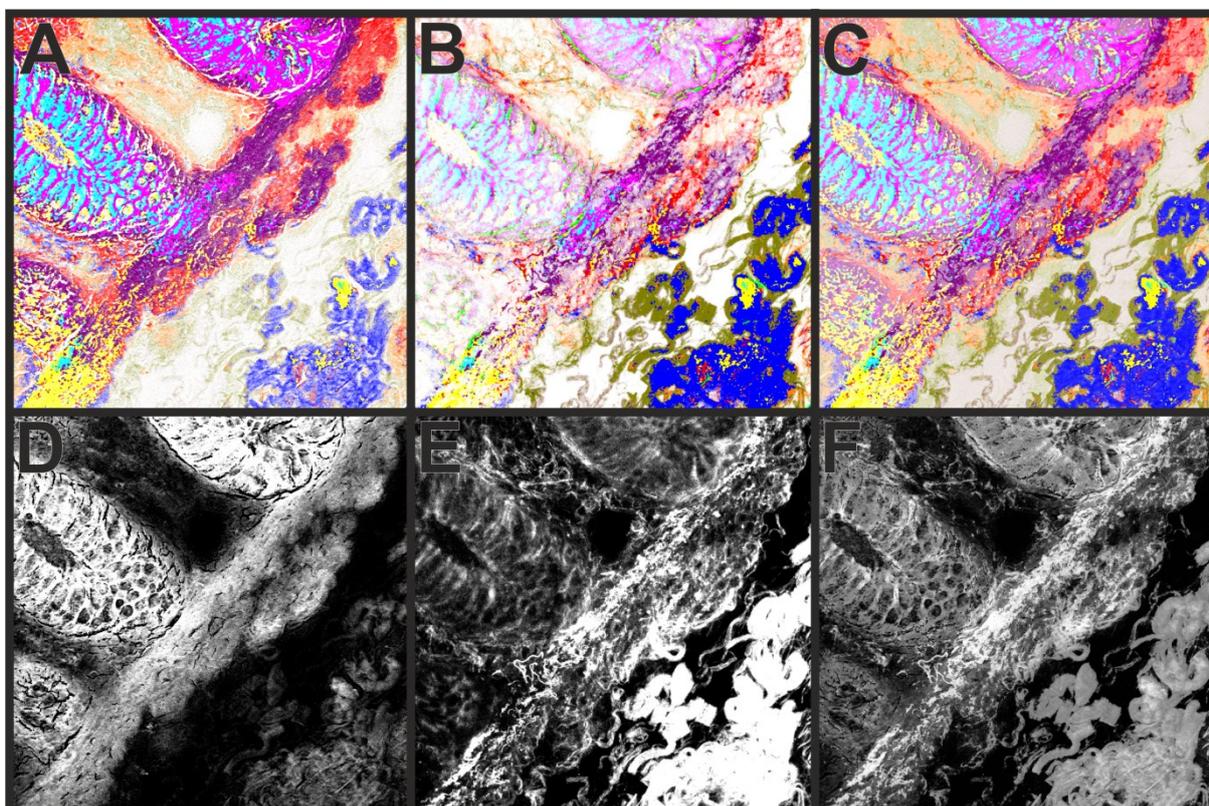
**Fig. S7.** Alternative representation for Fig. 5. Improvement of representation of Raman based SHP and establishment of Raman virtual staining.



**Fig. S8.** Alternative representation for Fig. 6. (A,D,G,J) H&E staining of selected regions of interest shown in Fig. 2. (B,E,H,K) Raman SHP of the same regions shown in (A,D,G,J). (C,F,I,L) Raman virtual staining, constructed from Raman SHP shown in (B,E,H,K) and their corresponding integrated Raman intensities in the 2800-3050  $\text{cm}^{-1}$  region.

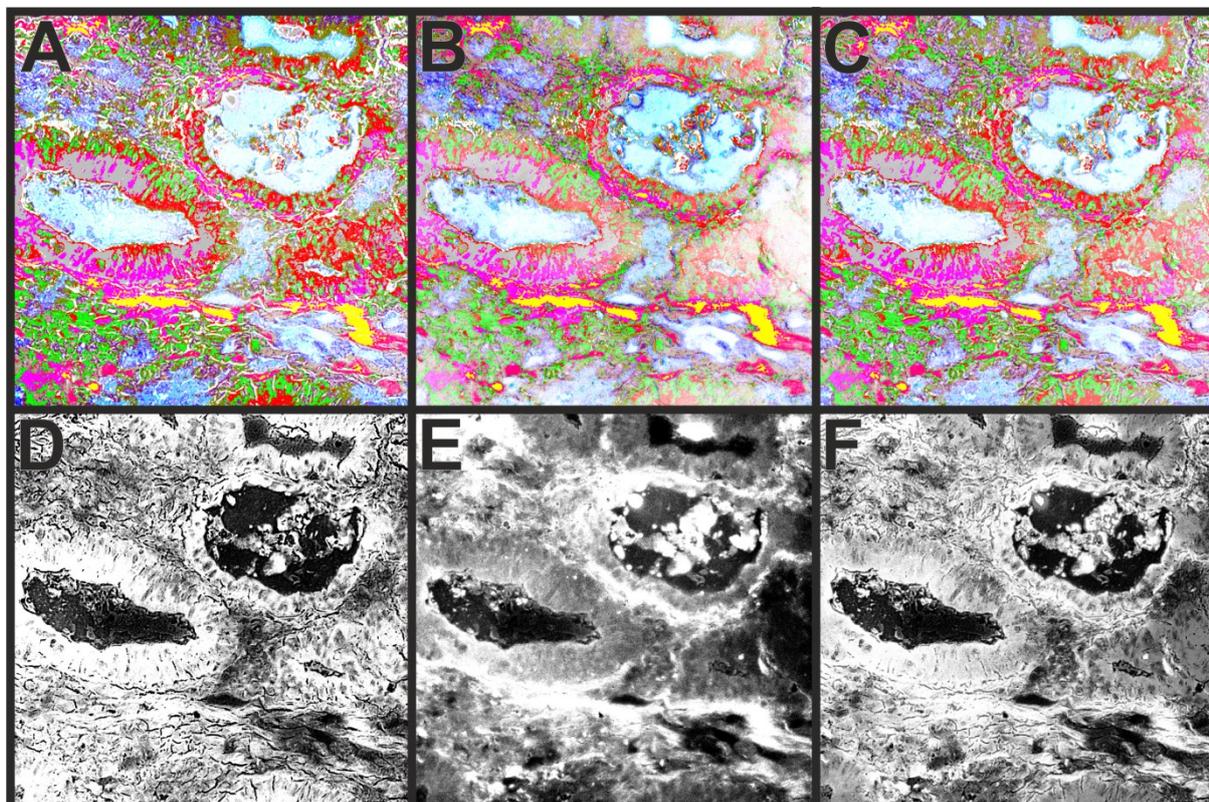


**Fig. S9.** CARS mean spectra of carcinoma (pink), connective tissue (blue), and tumor micro-environment (yellow). These spectra are produced using *k*-means clustering of CARS spectral dataset. For illustration the spectra are median filtered in the frequency domain.



**Fig. S10.** Constructed images from CARS *k*-means clustering analysis weighted with intensities of CARS (A), SHG (B) and both together (C). The pure intensity images of CARS at  $2850\text{ cm}^{-1}$  (D), SHG (E) and combined CARS and SHG (F) are also displayed.

Similar to the Raman virtual staining, we created intensity weighted pseudo-color images from *k*-means cluster analysis of CARS datasets. The pseudo-color image in Fig. 7B was weighted here with the CARS intensity at  $2850\text{ cm}^{-1}$  (A,D), with the SHG intensity (B,E) and with a combination of both intensities (C,F). It is important to note that the intensities are non-linear in comparison to the linear Raman signal.



**Fig. S11.** Constructed images from CARS *k*-means clustering analysis weighted with intensities of CARS (A), SHG (B) and both together (C). The pure intensity images of CARS at 2850  $\text{cm}^{-1}$  (D), SHG (E) and combined CARS and SHG (F) are also displayed.