Supporting Information

Potential use of MCR-ALS for the Identification of coeliac-related biochemical changes in Hyperspectral Raman Maps from Pediatric Intestinal Biopsies

Stefano Fornasaro^{1,*}, Annalisa Vicario¹, Luigina De Leo², Alois Bonifacio¹, Tarcisio Not³ and Valter Sergo¹

*Corresponding Author: E-mail: sformasaro@units.it

¹Department of Engineering and Architecture, University of Trieste, Trieste, Italy

²Institute for Maternal and Child Health - IRCCS "Burlo Garofolo"

³Institute for Maternal and Child Health - IRCCS "Burlo Garofolo", University of Trieste, Trieste, Italy



Figure S1 Baseline correction using asymmetric least squares fitting. A) 20 random raw spectra from CD samples group with fitted baseline (dashed line); B) same 20 spectra with baseline subtracted; C) Average Raman spectrum of raw CD samples group (n=21315); D) Average Raman spectrum of CD samples group with baseline subtracted. Intensity standard deviations are reported as shaded areas.



Figure S2 Average normalized Raman spectra of CrD samples group (red, n = 17600), compared with HC samples group (black, n = 20172), together with their difference spectra (CrD-HC). Intensity standard deviations are reported as shaded areas. Identified Raman bands are indicated by the corresponding Raman shift.



Figure S3 Pure component (1–6) distribution maps (A) and the corresponding spectral profiles (B), of Crohn Disease (CrD_{1-3}) multiset, following MCR-ALS. Distributions maps use a gradual colour scale where the dark blue colour refers to small concentration values and the red colour to large values.



Figure S4 Comparison between the reference spectra of pure compounds with those obtained by MCR-ALS for the CrD dataset. Sat, nonedecanoic acid; dna, ; alb, human serum albumine; col, collagen; oct, optimal cutting temperature compound; uns, palmitoleic acid; igg, Immunoglobulin G.



Figure S5 Detection of total anti-transglutaminase immunoglobulin deposits in a(A) healthy control and (B) celiac patient.



Figure S6 Average cluster spectra from the of HC and CD multisets following k-means clustering (KMC) and Hierarchical Cluster Analysis (HCA). Centroids were obtained after subtraction of the 2th percentile spectrum from all spectra.



Figure S7 Segmentation maps of CrD multiset following (A) k-means clustering and (B) Hierarchical Cluster Analysis. Cluster centroids were obtained after subtraction of the 2th percentile spectrum from all spectra.