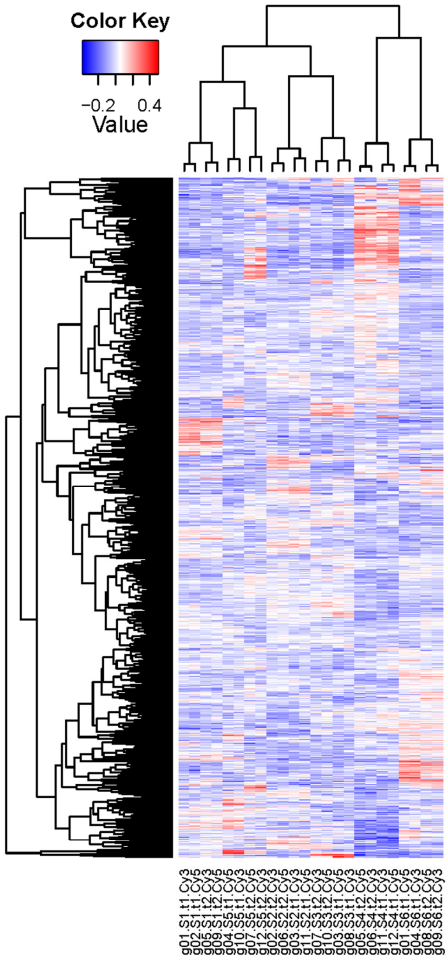
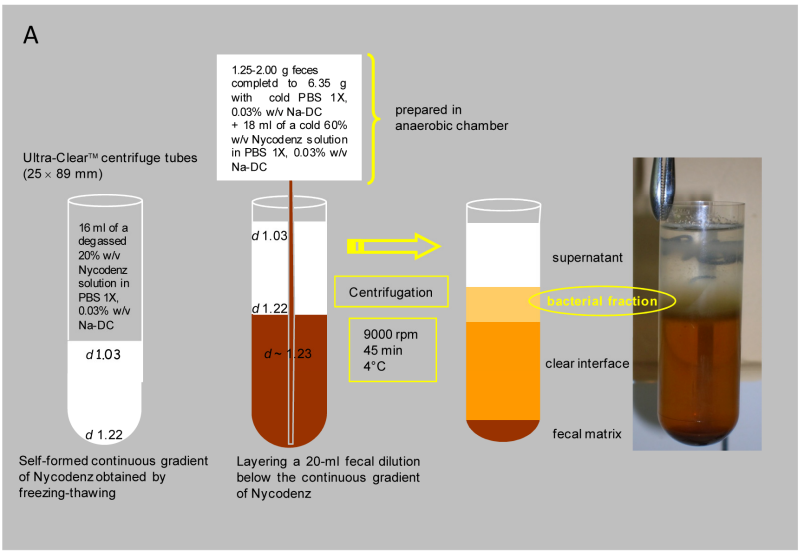
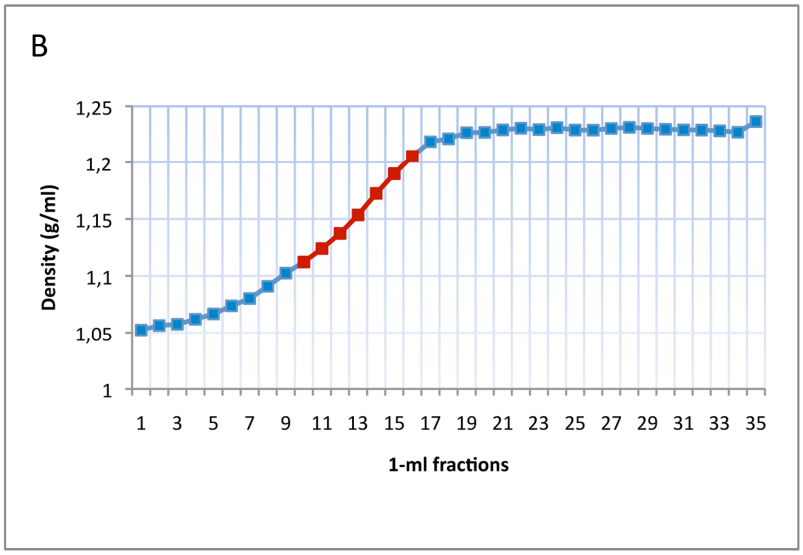
Supplementary figure 1



Supplementary figure 1 Cluster heat map constructed from the normalized volumes of spots (all spot variables centered at the mean) in six extra healthy subjects S1-S6 observed at a 1-month interval (t1 and t2); g01 to 12 denote gel numbers. Microbial fractions prepared in duplicate from the same stool specimen (Cy3 and Cy5) always clustered together, reflecting good technical reproducibility, and pairs of duplicates from a same subject observed at a 1-month interval always clustered together, indicating a clear stability in the proteomics data over time. The somewhat lower stability observed in subject S5 is due to freezing and thawing for transport at t2, a practice that was rejected for samples included in the study under evaluation.

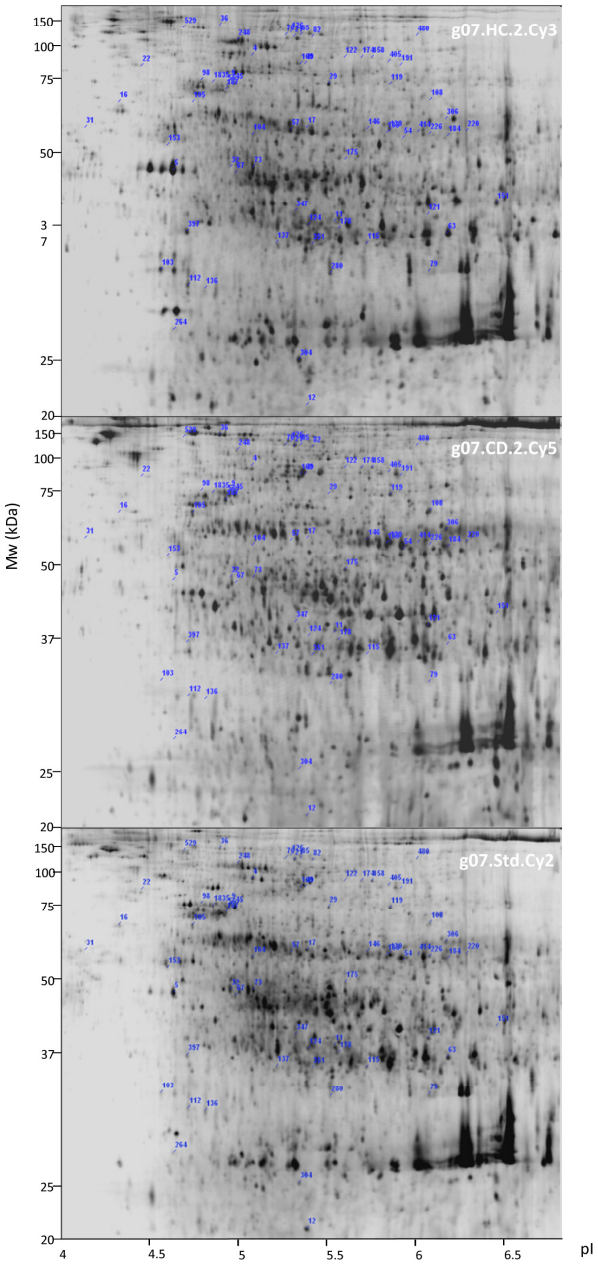
Supplementary figure 2





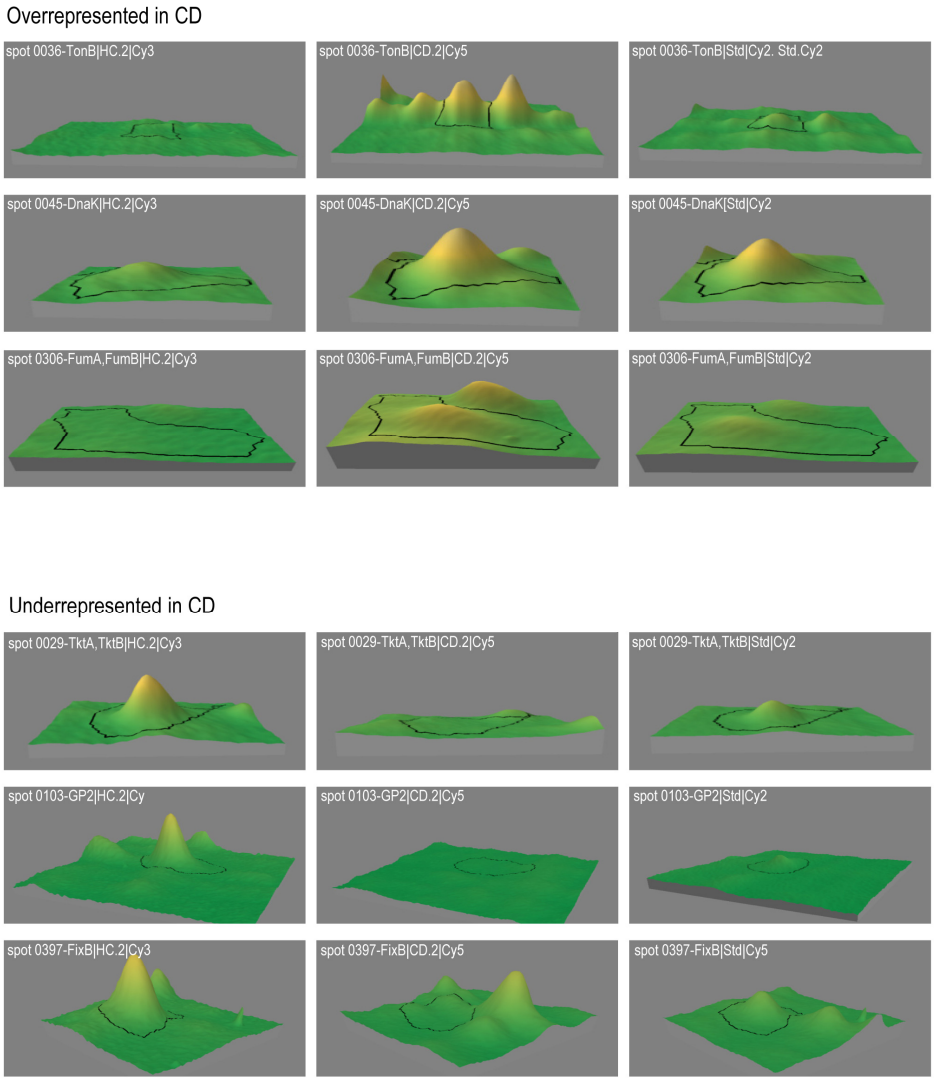
Supplementary figure 2 Schematic representation of the gradient procedure. (A) The overall strategy. (B) Flotation of bacterial cells in the middle of the gradient (seven 1-ml fractions, 1.112<*d*<1.206) after low-speed ultracentrifugation (9000 rpm for 45 min, 4°C) in a swing-out rotor (SW 28 Ti Rotor, Beckmann).

Supplementary figure 3



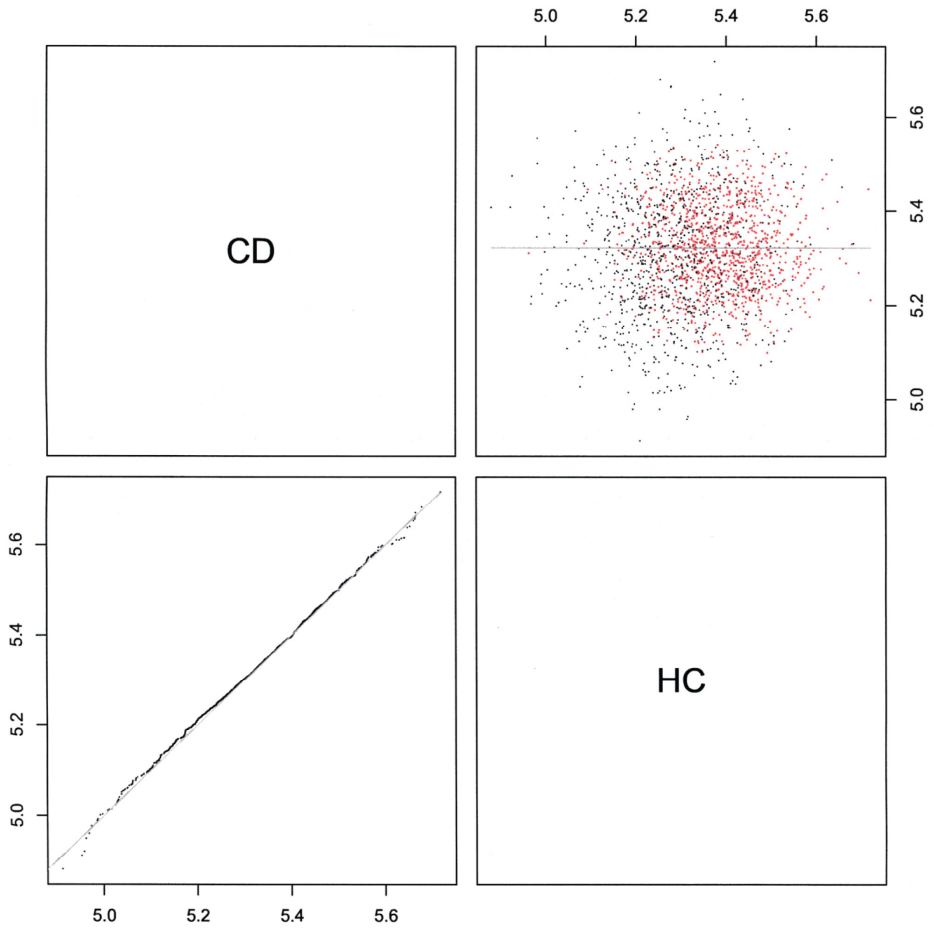
Supplementary figure 3 Cy3 (HC.2), Cy5 (CD.2) and Cy2 images of a representative multiplexed 2D gel. The internal standard was made of a pool of equal protein amounts from each of the 24 samples included in the study. Within the dynamic range of 2D-DIGE, the internal standard was not overwhelmed with additional spots as compared to the individual images Cy3 and Cy5, illustrating a great conservation of electrophoretic mobility for proteins serving main cellular functions across close members of the gut microbiome. Numbers correspond to the identified under- and overexpressed spots (reported in figure 3 and detailed in supplementary excel file). The strong saturated spots within the right bottom region were identified as pancreatic human elastase-3A with different post-translational modifications and were excluded from the analysis.

Supplementary figure 4



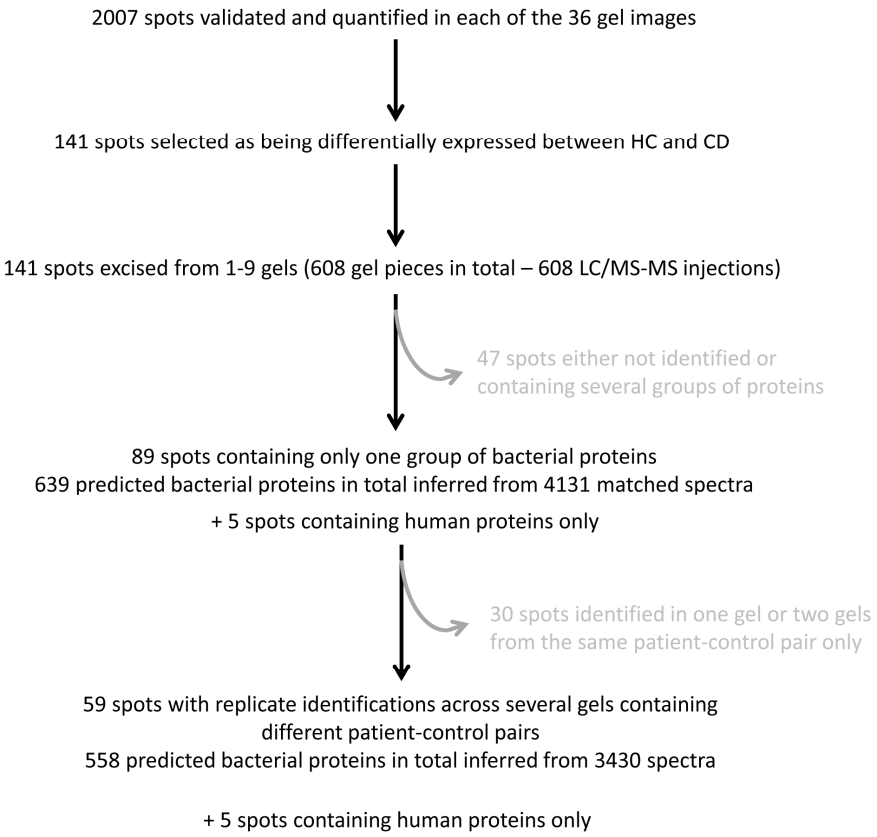
Supplementary figure 4 Three dimensional images of some relevant over- and under-represented spots.

Supplementary figure 5



Supplementary figure 5 Comparative scatter plots. The top-right panel holds a traditional M(A) plot, displaying the difference in protein abundance ('M', on the y-axis) *versus* the average protein abundance ('A', on the x-axis). A positive difference indicates higher abundance in CD *versus* HC samples. Each dot corresponds to one protein spot. The red dots indicate the non-varying subset of proteins identified and used for normalization purposes. The bottom-left panel provides a quantile-quantile plot, where the ordered protein abundances in CD samples (x-axis) are compared to the ordered protein abundances in HC samples (y-axis). The diagonal line indicates identical signal distributions. One can see from both plots that most proteins were unchanged between patients and controls.

Supplementary figure 6



Supplementary figure 6 Sequential spot selection process. Refer to supplementary excel file for lists of predicted proteins and corresponding peptides