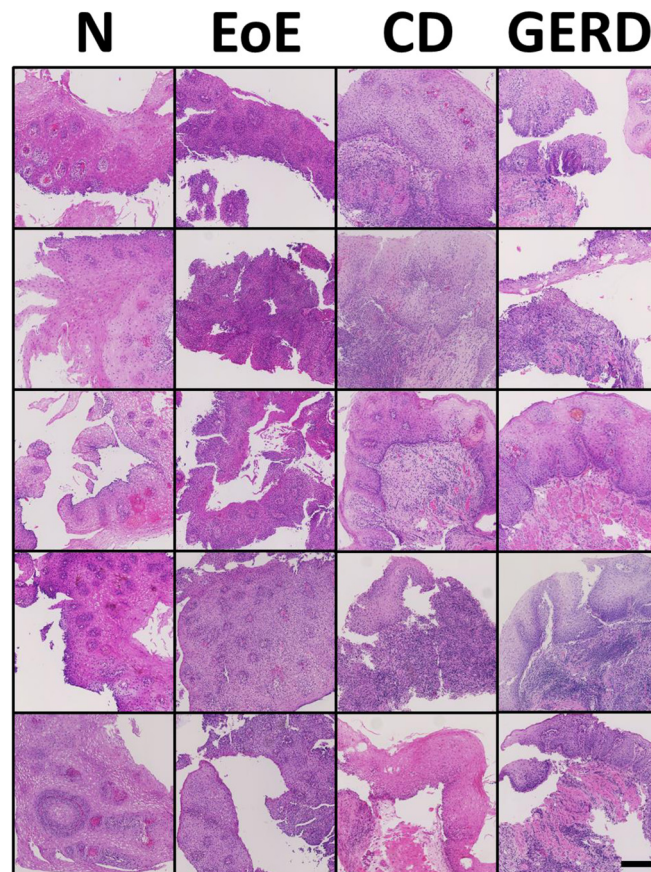


Expanded View Figures

**Figure EV1. Histology of reference samples.**

Representative micrographs of H&E stains of the reference samples for Normal (N), Eosinophilic esophagitis (EoE), Crohn's disease (CD), and gastro-esophageal reflux disease (GERD). Scale bar: 200 μ m.

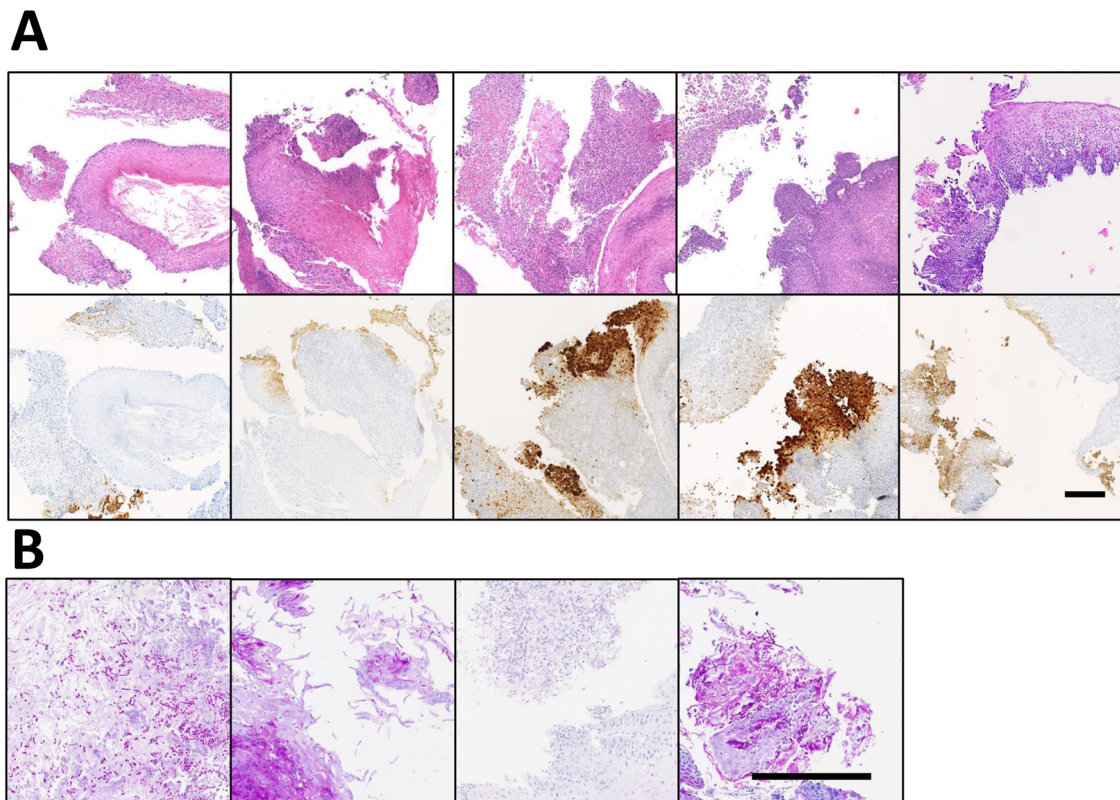


Figure EV2. Histology of infectious samples.

(A) H&E stains showing strong inflammation and heterogenous cytopathic changes (top row). IHC stains using an antibody against HSV1/2 highlight HSV-positive cells in brown (bottom row). (B) PAS reaction showing pseudo hyphae and spores of *Candida* spp. The epithelium displaying unspecific changes. Scale bars: 200 μ m.

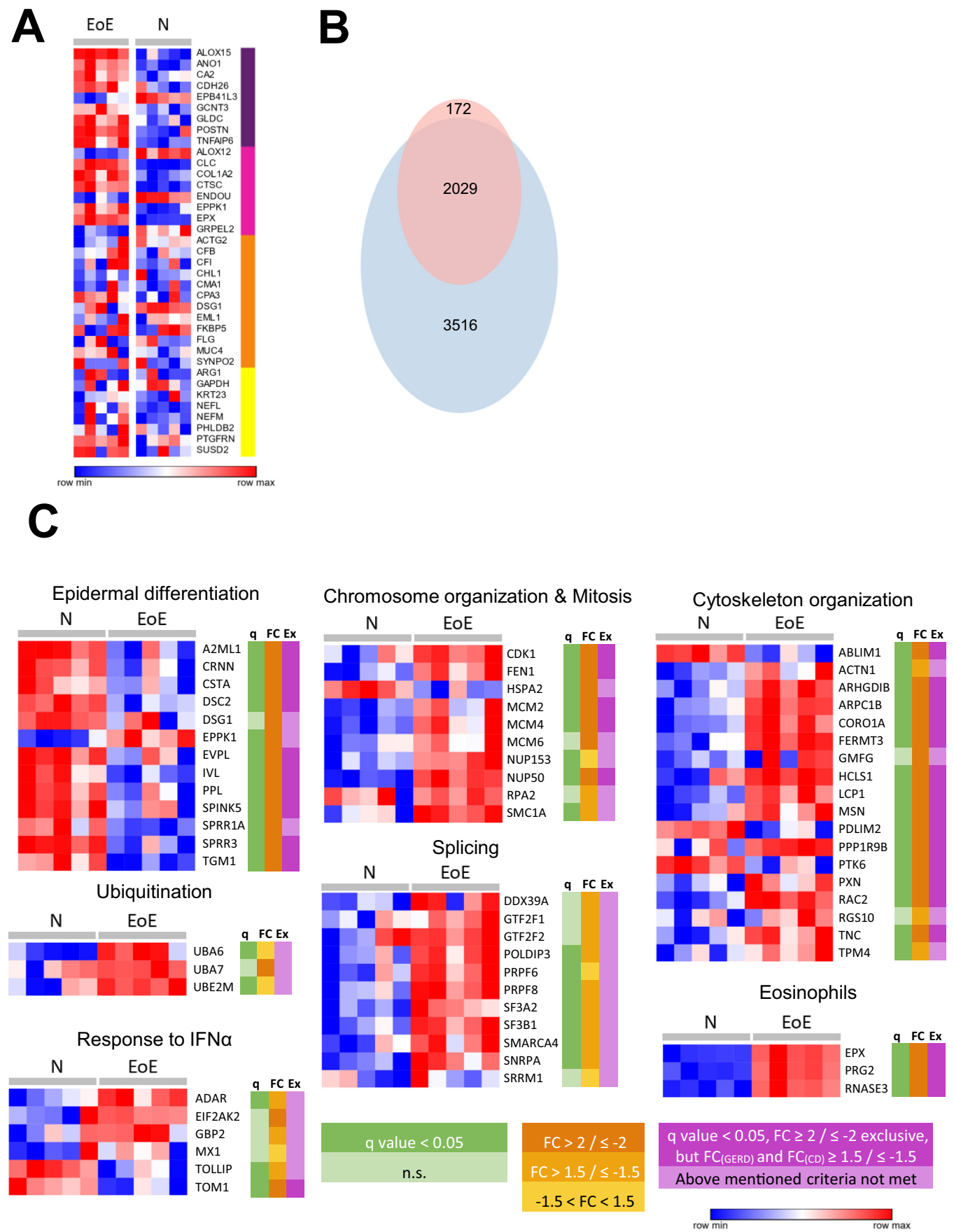


Figure EV3. Subset of proteins quantified in EoE versus N corresponding to previously published markers for EoE.

(A) Heatmap showing relative abundances of the indicated proteins in normal (N) and eosinophilic esophagitis (EoE) samples. The proteins correspond to markers (mRNAs) which were previously published (Wen et al, 2013) as part of an EoE diagnostic panel (EDP). Purple: exclusive for EoE v N; pink: significant, but not exclusive for EoE v N; orange: significant in CD v N or GERD v N, but not in EoE v N; yellow: not significantly altered, (B) Venn diagram of all quantified proteins in our dataset compared to a recent publication on tissue proteomics showing the differences between N and EoE samples (Rochman et al, 2023), (C) Heatmaps of the proteins mentioned in the enrichment analysis in Rochman et al, (2023) comparing N and EoE samples. Color coding showing different criteria for significance (q value, FC and exclusiveness in our dataset compared to GERD and CD samples) as stated on the bottom.

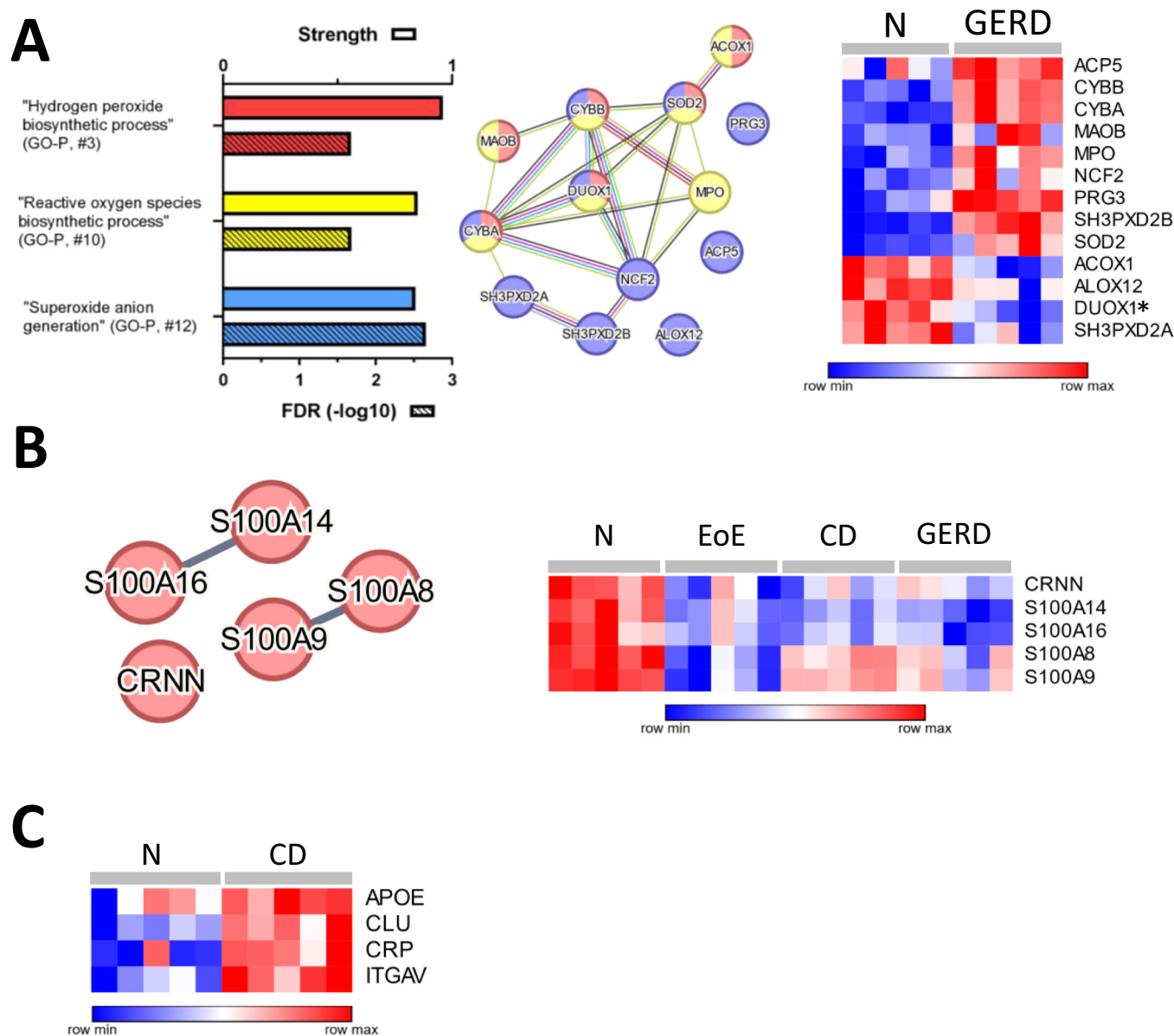


Figure EV4. Other relevant proteins of different comparisons.

(A) STRING networks of reactive oxygen species (ROS)-associated proteins and corresponding heatmaps (relative abundances of indicated proteins in each N and GERD sample) as well as related enrichment terms highlighted (for complete list see Dataset EV3B) by bar diagrams with the respective strength and false discovery rate (FDR, $-\log_{10}$, hatched bar) as well as the rank in the respective database. (B) STRING networks of S100A- and associated proteins and corresponding heatmaps (relative abundances of indicated proteins in each N, EoE, CD, and GERD sample). All of those proteins are significantly altered between N and each of the inflammatory groups. Except for S100A8 and S100A9 (each in N vs CD) all other proteins exceed $\log_2FC$ 1/−1 (Dataset EV1C). (C) Heatmap (relative abundances of indicated proteins in each N and CD sample) of proteins previously described as predictive markers for anti-TNF- α therapy (Gazouli et al, 2013; Kalla et al, 2021; Kumar et al, 2024). The lowest value in each row in the heatmaps is displayed in dark blue, the highest value in dark red. The scale traverses white. The heatmaps were created using Morpheus.

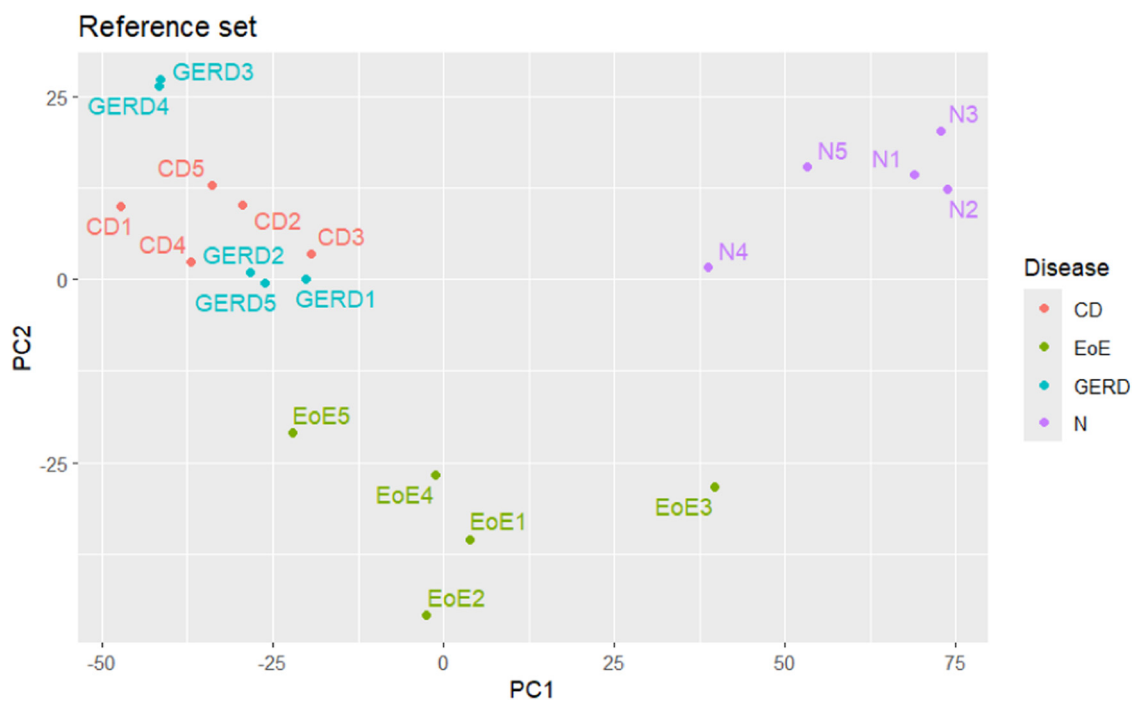


Figure EV5. PCA plot of reference dataset.

Principal component analysis (PCA) performed on the reference dataset illustrates separation of N, EoE, and the mixed cluster of GERD and CD samples.

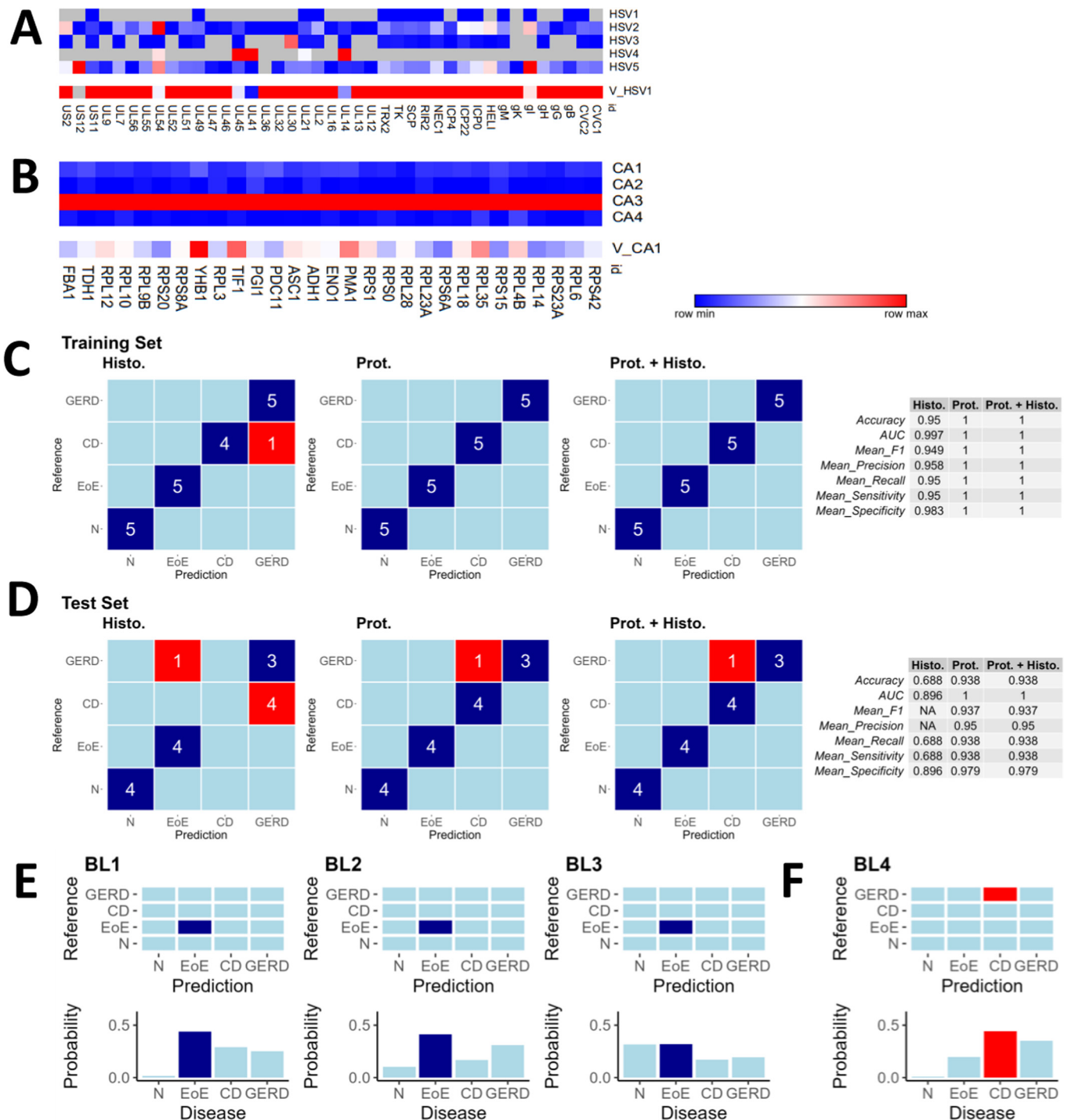


Figure EV6. Validation of proteomic approach in infectious conditions, illustration of the random forest (RF) parameters in training and test samples, as well as results of borderline cases.

(A) HSV-specific proteins are detected in the HSV esophagitis test sample (V_HSV1). (B) Candida-specific proteins are detected in the Candida esophagitis test sample (V_CA1). The lowest value in each column of the heatmaps is displayed in dark blue, the highest value in dark red. The scale traverses white. Heatmaps were created using Morpheus. (C) Results of the RF approach in the training dataset for histological features (Histo.) alone, proteome (Prot.) alone and combined (Prot. + Histo.). (D) Results of the RF with the aforementioned conditions in the test cohort. (E) Results of testing the approach in different borderline conditions. (F) Borderline case of a patient with histology and presentation favor GERD. Squares in the confusion matrices and bars colored in blue indicate the correct classification for each case by the RF, red markings indicating an incorrect classification.