Supplementary Materials for:

Investigating neoplastic progression of ulcerative colitis with label-free comparative proteomics

Summary of Materials:

spectralcount_comparison.doc: This Word document demonstrates the correlation between protein ratios calculated via our peptide ion intensity-based approached and ratios calculated by simple spectral counting.

supplementary_tables.xlsx: This Excel spreadsheet contains four worksheets, distributed in a single file for convenience. These worksheets correspond to Supplementary Tables 1-4 referred to in the manuscript. These tables are:

Table 1

UC patient clinical characteristics: the vital information about the samples used in this work.

Table 2

This table summarizes the protein groups discovered in our study that fit the criteria used to determine candidate status. A protein group is an entity based on peptide evidence that may be associated with one or more gene symbols and/or IPI identifiers and descriptions. Each protein group has at least two unique peptide observations, a quantitative ratio (in the PHGD:NP comparison, the PNEG:NP comparison, or both) ≤ 0.5 or ≥ 2.0, and either $q \leq 0.1$ or no q-value calculated. Proteins are sorted by the ratio from the PNEG:NP comparison

Each row represents one set of peptide evidence that may be explained by a single protein, or by multiple proteins that cannot be distinguished based on the evidence observed. Some fields may have multiple values; these are separated by semicolons. The columns of this table are as follows:

Column	Explanation
Gene(s)	Gene symbol(s)
Protein Identifier(s)	Unique identifier(s) in the IPI database, version 3.65
Ratio	Quantitative ratio, as described in the manuscript
q-value	q-value, as described in the manuscript
Peptides	Number of unique peptide sequences identified
Description	Protein description(s)

Table 3

Results of the top 10 clusters from DAVID Functional Annotation Clustering analysis for the P-NEG vs. NP comparison. This spreadsheet is divided into sections based on the enrichment score, with highest enrichment scores first. The Functional Annotation Clustering function uses a novel algorithm to measure relationships among the annotation terms based on the degrees of their co-association genes to group the similar, redundant, and heterogeneous annotation contents from the same or different resources into annotation groups. This reduces the burden of associating similar redundant terms and makes the biological interpretation more focused in a group level. Please refer references 24 and 25 in the manuscript for a detailed description of these tools.

Table 4

Same as Table 3, but for the P-HGD vs. NP comparison.

Table 5
Ingenuity Canonical Pathways analysis of P-NEG. The pathways are sorted by their p values.

Table 6
Ingenuity Canonical Pathways analysis of P-HGD. The pathways are sorted by their p values.

Table 7
List of HIS scores for CPS1.