

SUPPLEMENTARY MATERIALS AND METHODS

Patient cohort and clinical information

The patient cohort subjected to 14-day oral gluten challenge has previously been described by Vikas et al. where full clinical information can be found.[1] (Trial registration number NCT02464150). For this study we have used material from formalin fixed paraffin embedded (FFPE) small intestinal biopsies collected at baseline for 19 of 20 enrolled patients before challenge and at day 14 of challenge. Histomorphometry (Marsh score and Vh:Cd ratio) and IEL counts was previously reported and were performed on sections from the same FFPE biopsy blocks from which we collected tissue for proteomics analysis. All patients were upon study initiation considered to be in complete clinical and mucosal remission at baseline before challenge as assessed by Marsh score and serum anti-transglutaminase IgA titers. One patient (P11, CD442) was initially evaluated to be in mucosal remission before onset of gluten challenge but was later revised to "Marsh 3" after a blinded re-evaluation of all biopsies.[1] Clinical biochemistry and cytokine measurements were performed at baseline for the 19 patients that completed gluten challenge. Measurement of CD4+ gluten specific T cells has previously been described and reported from blood and gut.[1,2]

Indirect Patient and Public Involvement

The oral gluten challenge clinical trial from which we have used archival biopsy material was performed at a translational research centre (K.G. Jebsen Centre for Coeliac Disease Research) where patient organisation representatives serve as advisory board members and aid in communication of research results to patients and the public. Study participant recruitment led by Dr. Knut E.A. Lundin and patient communication are integrated activities at the research centre.

Total tissue sample digestion and processing

For generation of total tissue digests, fifteen 5µm thick sections from each FFPE biopsy block were collected in a tube (Eppendorf, Hamburg, Germany) and dehydrated by 10 min incubations in 80%, 96% and 100% ethanol, respectively, followed by dewaxing by two incubations with xylene (2x3 min at 55°C) and two incubations in 100% ethanol (3min at room temperature). Dewaxed tissue was resuspended in 20µl 50mM ammonium bicarbonate with 0.2% ProteaseMAX surfactant (Trypsin enhancer; Promega, Madison, WI) followed by addition of 1µl 0.5 mol/L dithiotheriol (DTT) and 73.5µl 50mM ammonium bicarbonate. Formalin crosslinks were cleaved by heating the samples (98°C 90min) followed by sonication (60 min in a water bath). Protein amount was estimated by DirectDetect (Millipore, Merck, Darmstadt, Germany). Disulfide bonds were reduced with DTT (1 µl 0.5mol/L per sample, 20 min incubation with gentle agitation at 56°C) and alkylated with iodoacetamide (2.7 µl 0.55mol/L per sample, 15 min incubation with gentle agitation at room temperature in the dark). To digest proteins, 2 µl 1% ProteaseMAX and 1 µg Trypsin (Sequencing grade, ProMega) was added to each sample followed by incubation in a wet chamber over night at 37°C. Peptides were purified on C18 micro columns as previously described.[3] Purified samples were adjusted to a final volume of 11µl or 20µl depending on protein concentration.

Laser capture microdissection sample collection and processing

Eight micrometer tissue sections were adhered to PEN-covered slides (Zeiss) and dried at 37°C. Dry sections were dewaxed in xylene (3 min + 2 min) followed by 1 min in 100% ethanol, 95% ethanol and 70% ethanol, respectively followed by 2 x 1 min in water. Tissue was visualized by staining with Mayer's hematoxylin solution (Sigma) for ~30 s followed by rinsing in tap water. Stained sections were air-dried and stored dry until cutting. Samples were collected using a PALM MicroBeam laser capture microdissection system (Carl Zeiss MicroImaging, Munich, Germany), and isolated tissue collected in 0.5ml opaque adhesive cap tubes (Zeiss). For total epithelial cell layer analysis on average 250 000 μm^2 x 8 μm tissue was collected per sample, while for apical epithelial tissue on average 150 000 μm^2 x 8 μm tissue was collected per sample. For comparison of lamina propria, villus and crypt epithelium we collected two samples per region. Two cohorts of total epithelial cell layer samples were dissected and digested (LCM1; 17 samples from 13 biopsies; LCM2: 25 samples from 20 biopsies). For apical epithelial cell layer analysis, we dissected 24 samples from 12 biopsies collected before challenge. Dissected tissue was retrieved from adhesive caps with 10 μl 50mM NH_4HCO_3 with 0.2% ProteaseMax Surfactant followed by 10 μl 50mM ammonium bicarbonate and transferred to 0.5mL Low-Bind tubes (Eppendorf, Hamburg, Germany). All downstream processing of LCM samples was performed in the same tube to minimize sample loss from pipetting. Samples were heated to 98°C for 90 min followed by sonication in water bath for 60 min. Disulfide bridges were reduced by addition of 2 μl 0.1 mol/L DTT followed by incubation for 20 min at 56°C and alkylated by addition of 2 μl 0.55 mol/L iodoacetamide followed by incubation for 15 min in the dark at room temperature. Samples were digested by addition of trypsin (1.5 μl 0.01 $\mu\text{g}/\text{ml}$) and incubation in wet chamber over night at 37°C. Peptides were purified on C18 micro columns and eluted samples were adjusted to a final of 7 μl with 0.1% formic acid.

Mass spectrometry analysis

Three μl digest was injected per run for all samples. Total tissue digests were analysed with two technical replicates. LCM1 samples were analysed together (LCM cohort 1). LCM2 samples were run together with 13 samples from LCM1 resulting in total 38 samples from 21 biopsies in the LCM2 cohort 2 dataset (LCM cohort 2). All experiments were performed on an Easy nLC1000 nano-LC system connected to a quadrupole – Orbitrap (QExactive Plus) mass spectrometer (ThermoElectron, Bremen, Germany) equipped with a nanoelectrospray ion source (EasySpray/Thermo). For liquid chromatography separation we used an EasySpray column (C18, 2 μm beads, 100 \AA , 75 μm inner diameter) (Thermo) capillary of 50 cm bed length. The flow rate used was 0.3 $\mu\text{L}/\text{min}$. For total tissue samples, the solvent gradient was 2% to 5 % in 10 minutes, to 19% in 170 minutes and then to 35% B in 60 minutes followed by a wash with 90% B for 20min. For LCM isolated sample, the solvent gradient was 2 - 7% B in 10 minutes, then to 30 % B in 55 minutes and finally a wash with 90 % B in 20 minutes. Solvent A was aqueous 0.1 % formic acid, whereas solvent B was 100 % acetonitrile in 0.1 % formic acid. Column temperature was kept at 60°C.

The mass spectrometer was operated in the data-dependent mode to automatically switch between MS and MS/MS acquisition. Survey full scan MS spectra (total tissue: m/z 300 to 1,500; LCM samples: m/z 400 to 1,200) were acquired in the Orbitrap with resolution $R = 70,000$ at m/z 200 (after accumulation to a target of 3,000,000 ions in the quadrupole). The method used allowed sequential isolation of the most intense multiply-charged ions, up to ten, depending on signal intensity, for

fragmentation on the HCD cell using high-energy collision dissociation at a target value of 100,000 charges or maximum acquisition time of 100 ms. MS/MS scans were collected at 17,500 resolution at the Orbitrap cell. Target ions already selected for MS/MS were dynamically excluded for 30 seconds. General mass spectrometry conditions were: electrospray voltage, 2.1 kV; no sheath and auxiliary gas flow, heated capillary temperature of 250°C, normalized HCD collision energy 25%.

Protein identification, quantification and data analysis

MS raw files were processed in the MaxQuant environment [4] (version 1.6.1.0) with the integrated Andromeda search engine [5] for peptide and protein identification. The human UniProtKB FASTA database (September 2018) was used as forward database for protein identification. Match between runs was enabled and label-free protein quantification was performed using the MaxQuant's Label Free Quantification algorithm with a minimum ratio count of one. [6] Methionine oxidation and N-terminal acetylation was used as variable modification and carbamidomethyl cysteine as fixed modification. Separate MaxQuant searches for LC-MS/MS data from for total tissue, total epithelial LCM cohort 1, total epithelial LCM cohort 2 and apical epithelial LCM sample runs were done. Compartment samples (lamina propria, epithelial villi and crypts) were run and analysed together with LCM2 cohort 2. Statistical and bioinformatical analysis was performed in Perseus (version 1.6.2.2) [7] or the R framework (<https://www.r-project.org/>). For all datasets, proteins matched to the reverse decoy database, identified by site or identified as potential contaminant were removed. For the total tissue dataset, technical replicates were averaged and the dataset filtered to keep only proteins having valid LFQ-values in at least 70% of the samples in at least one of four sample-groups (responders or non-responders before or after challenge). Missing values were imputed based on normal distribution to simulate low abundant LFQ values. For LCM cohort 2, compartment specific samples were analysed separately and PC1 was subtracted from total epithelial samples prior to downstream analysis. After these steps, LCM cohort 1 and LCM cohort 2 datasets they were merged. LFQ values from LCM1 and LCM2 datasets were merged by matching protein groups from LCM1 (n = 2495) to protein groups from LCM2 (n = 3080). Only protein groups present in both datasets (n = 2404) were kept, and PC1 was subtracted from the combined LFQ values prior to downstream analysis. Unless otherwise stated, individual data points correspond to one sample or MS rawfile. Where indicated, median values per biopsy (i.e. per patient before or after challenge) is used. For apical epithelial LCM dataset we filtered for protein groups present in 50% of responder or non-responder samples followed by imputation of missing values as above. Apical epithelial data (apical LCM) were compared with total epithelial data z-scored across only before challenge samples.

Gene set annotation and 2D categorical enrichment to denote biological pathways

Small intestinal epithelial cell type gene sets were retrieved from Supplementary Table 4 from [8]. We converted mouse to human genes by matching mouse to human gene name orthologues (www.ensembl.org/biomart) and gene sets were generated for mature enterocytes, goblet cells and foetal enterocytes. Duplicates gene names and gene names that appeared in two groups were removed to give final gene sets for mature enterocytes (n=437), goblet cells (n=367) and foetal enterocytes (n=280) to which we matched out proteome datasets (**online supplementary table 3**). Biological pathway enrichment analysis was performed in Perseus (2D enrichment based on Student t-tests fold difference,

Benjamini-Hochberg adjustment for multiple testing, FDR = 0.02).[9] 2D enrichment data for Gene Ontology pathways (GOBP, GOCC, GOMF) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were exported and visualized in R. We filtered for pathways with less than 100 genes and student t-test fold difference (Responders after vs. before challenge and Before challenge responders vs. non-responders). Student t-test difference > 0.25; Shared UP (38 pathways of which 23 were GO Biological Pathways, 410 unique proteins.). Student t-test difference < 0.25; Shared DOWN (77 pathways of which 42 were GO Biological Pathways, 573 unique proteins). Expression of proteins that mapped to GOBP pathways in Shared UP or DOWN were compared between responder groups (ANOVA and Tukey's honest significance test, FDR = 0.05). For selected pathways, median z-score protein expression per responder groups was compared. Results from 2D enrichment analysis are uploaded with mass spectrometry raw data.

Mass spectrometry data have been deposited to the ProteomeXchange Consortium [10] via the PRIDE partner repository with the dataset identifier PXD(will be updated)

SUPPLEMENTARY FIGURES

Supplementary figure 1

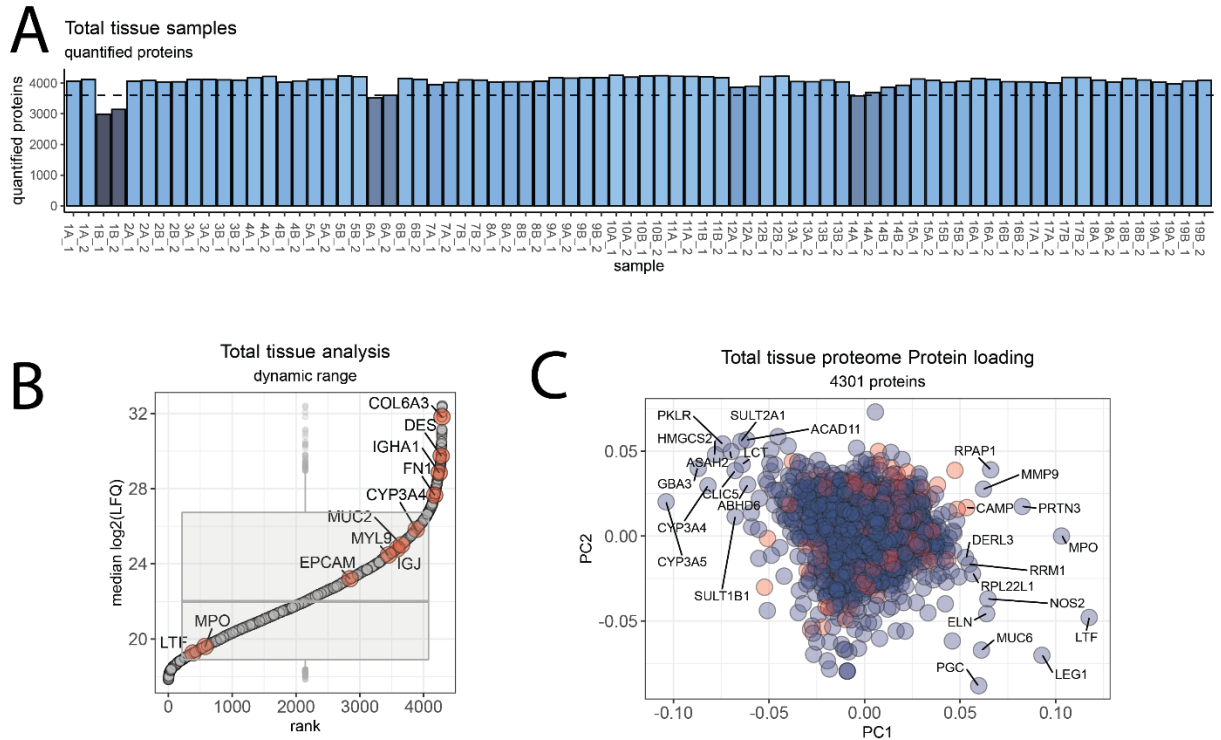


Figure S1. LC-MS/MS analysis of total tissue digest from biopsies collected before and after challenge

A. Samples analysed by LC-MS/MS from total tissue digests. Samples with <3600 valid values from LFQ protein quantification were removed from the dataset (1B, 6A, 14A) **B.** Dynamic range of LFQ quantified protein groups from total tissue digest samples. Boxplot denotes 5% and 95% quantile and error bars 1% and 99% quantile. **C.** Protein loadings that drive separation of samples in Figure 1A. Colour shows number of peptides used for protein quantification (blue = razor + unique >1, n = 3963; red; razor + unique = 1, n = 338)

Supplementary figure 2

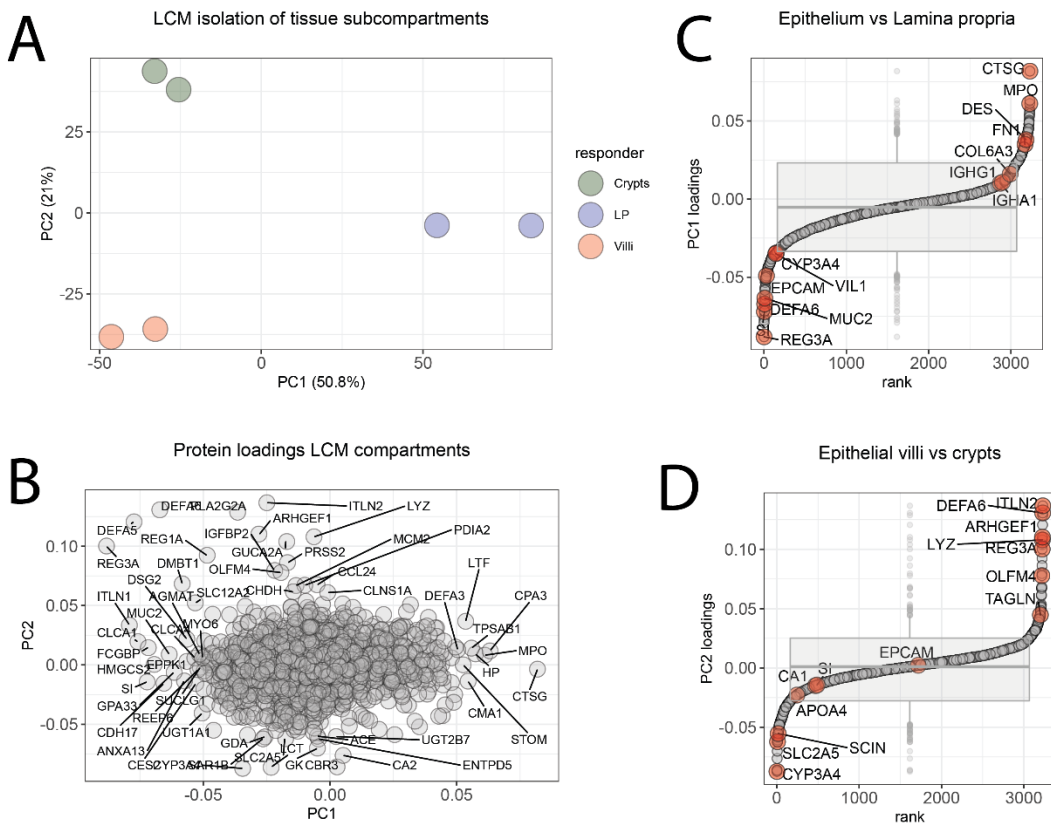


Figure S2. LCM isolation and proteomics analysis of distinct tissue subcompartments

A. PCA analysis of six samples collected from three distinct tissue compartments (Epithelial villi = red; Epithelial crypts = green; Lamina Propria = blue). **B.** Protein loadings that drive the separation in A. **C.** Protein distribution along PC1 that separates epithelium from lamina propria. Selected proteins representative for intestinal epithelium and lamina propria are indicated in red **D.** Protein distribution along PC2 that separates total epithelium from crypts. Selected proteins representative for enterocyte absorptive function (villi) and epithelial crypt proliferation and Paneth cells (crypts) are indicated in red.

Supplementary figure 3

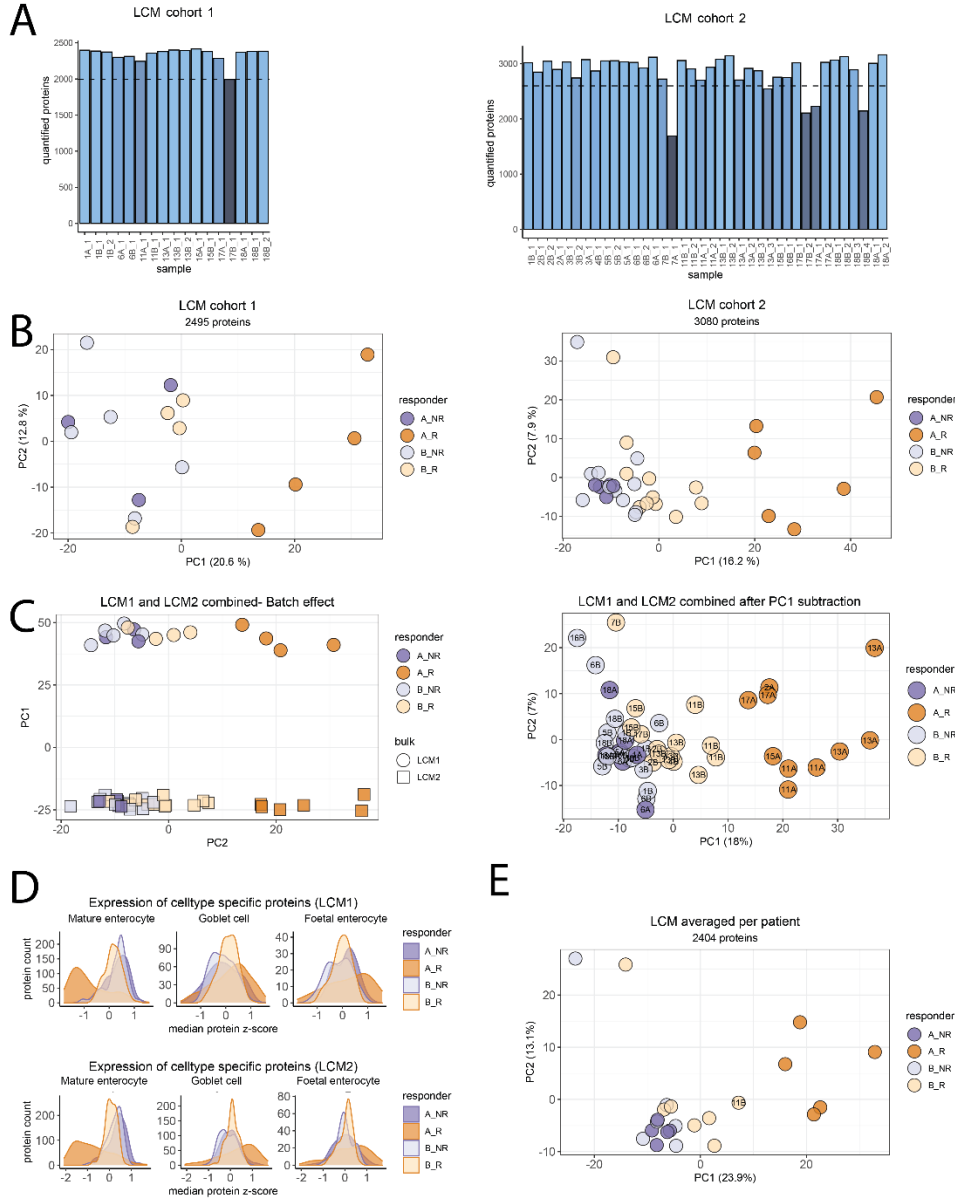


Figure S3. LCM isolation of epithelial cell layer samples from biopsies collected before and after gluten challenge

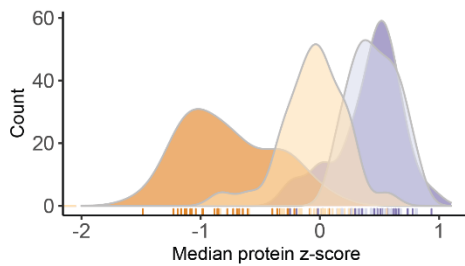
A. Number of proteins with LFQ values (quantified proteins) for each sample in the two cohorts of LCM isolated epithelium. Protein identification and label-free quantification was performed separately for the two cohorts to avoid excessive reliance on data extrapolation from match-between-runs. Samples were filtered based on LFQ valid values (LCM1 <2000, one sample removed; LCM2 <2600, 5 samples removed) **B.** PCA plot show distribution of samples from the independently analyzed LCM cohorts. **C** Data from LCM1 and LCM2 was combined (left) and PC1 subtracted. The resulting PCA plot from 2404 proteins (right) shows similar distribution of individual samples as observed for total tissue digest. **D.**

The distribution of cell type specific proteins is similar for both LCM cohorts with skewed distribution in responder after challenge samples. E. PCA plot from median protein expression per biopsy (compared to distribution of all samples as shown in C).

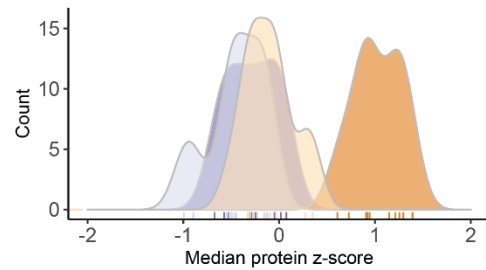
Supplementary figure 4

Pathway protein expression without samples from P11 before challenge (Marsh 3)

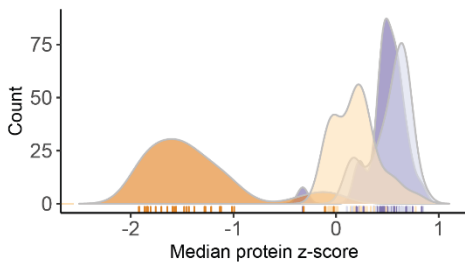
GOBP Mitochondrial electron transport NADH to ubuiqinone



GOBP Type I interferon-mediated signaling pathway



GOBP fatty acid beta-oxidation



GOBP Protein activation cascade

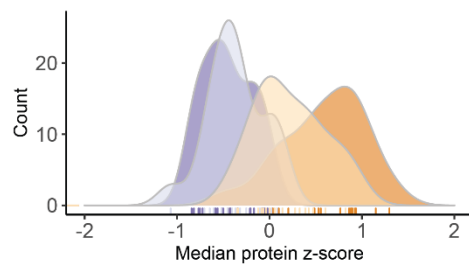


Figure S4 Baseline difference in responders does not depend on samples from patient P11

Samples from patient 11B baseline biopsies were removed from the dataset prior to calculation of responder group protein z-score and visualization of proteins in biological pathways from Figure 3D and E.

Supplementary figure 5

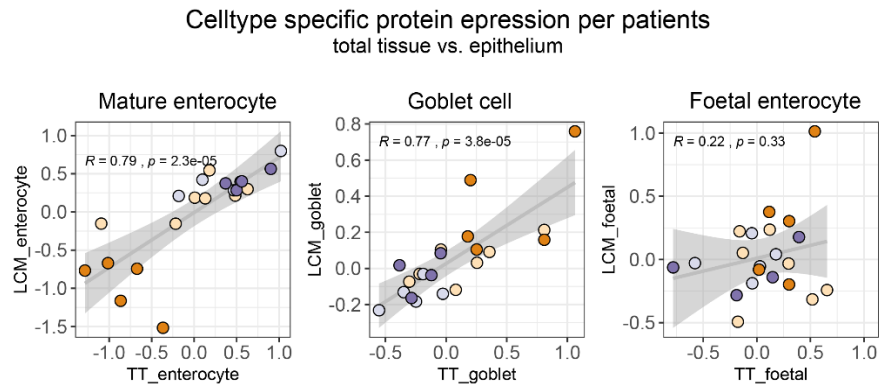


Figure S5 Comparison of cell type specific protein expression per patient

Correlation between median cell type specific protein expression per patient from total tissue (TT) and epithelium (LCM) proteome datasets.

Supplementary figure 6

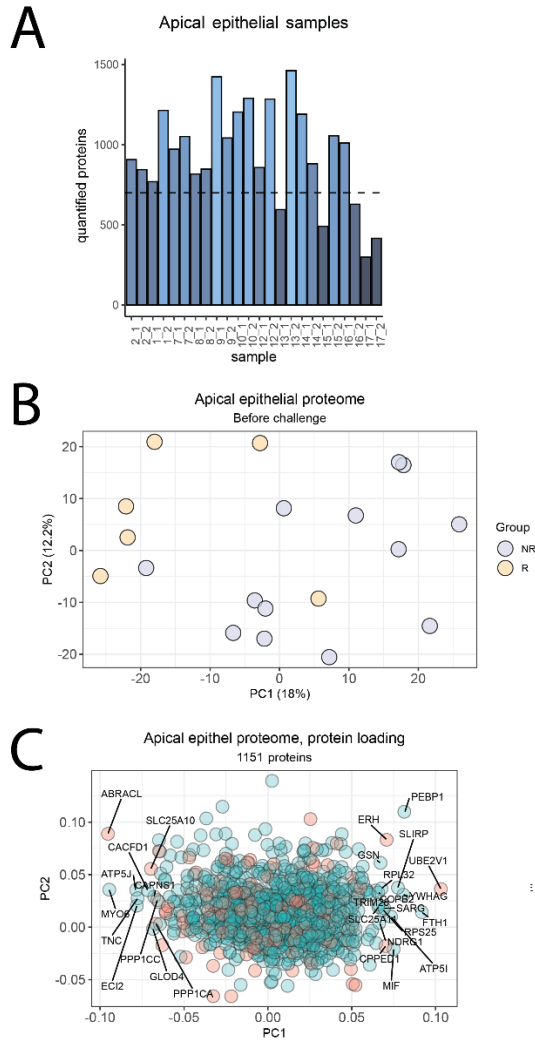


Figure S6 Quality control of apical proteome dataset

A. Number of proteins with LFQ values (quantified proteins) for the apical epithelial samples ($n = 24$, from 12 biopsies). Samples with valid values < 700 were removed (4 samples) from the dataset. **B.** PCA plot show distribution of all samples **C.** Protein loadings driving the separation in B (blue, razor + unique peptides > 1 , red = razor + unique = 1).

Supplementary figure 7

Haemoglobin proteins

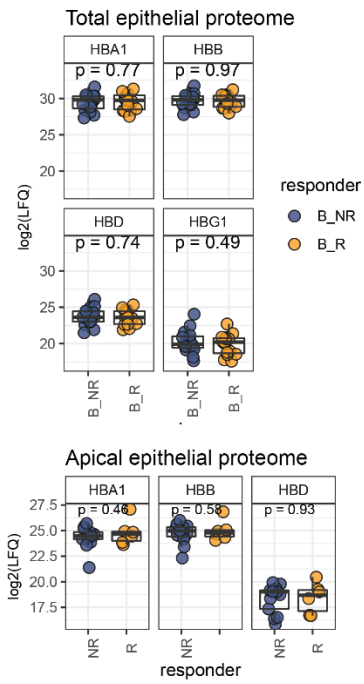


Figure S7 Haemoglobin proteins expression

Expression of haemoglobin proteins in the total epithelial and apical epithelial datasets.

Supplementary figure 8

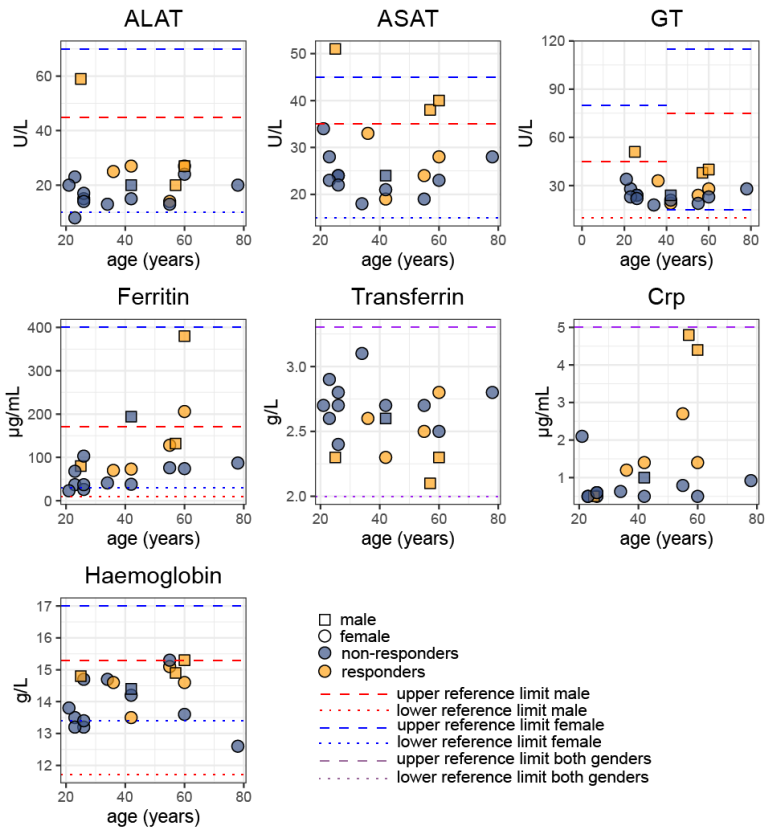


Figure S8 Biochemistry values as effect of age and gender

Reference values are from the local guidelines of Oslo University Hospital, of which most can be found here [11]

Supplementary figure 9

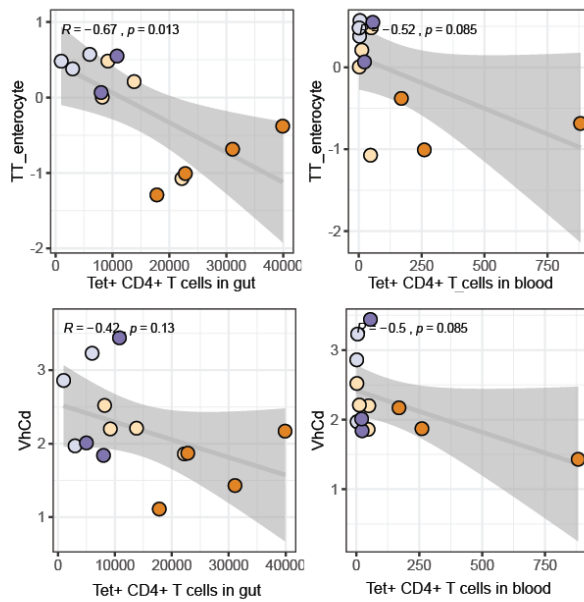


Figure S9 Correlation of mature enterocyte protein expression and Vh:Cd ratio with CD4+ tetramer positive T-cell frequency

Comparison of before and after challenge in gut (baseline and day 14) and blood (before and day 6) (Pearson's correlation)

SUPPLEMENTARY TABLES

Table S1 Patient variables.

Responder = responders as defined based on tissue proteomics. B = baseline before challenge. Clinical variables are from [1] except gluten specific T-cell frequencies in gut which are from [2].

Patient ID	CD number	Responder	Marsh score		Vh: Cd ratio		IEL (/100 EC)		Gluten specific CD4+ T cells			
			base line	day 14	base line	day 14	base line	day 14	blood* base line	blood day 6	gut# baseline	Gut day 14
P1	1284	NR	1	1	2.78	2.7	27.2	26.8	5.2	-	-	-
P2	1295	R	0	3	2.2	1.11	22.4	45.4	49.5	n.d.	9200	17800
P3	1294	NR	0	0	2.47	3.07	22.7	23.6	0.1	-	-	-
P4	1296	R	0	3	2.77	1.63	21.9	48.6	1.4	-	-	-
P5	1299	NR	0	1	3.23	3.44	23.9	26.5	5	55.9	6000	10800
P6	1298	NR	0	0	3.27	2.75	17.9	19.1	1.4	-	-	-
P7	1300	R	0	3	2.21	1.43	15.8	52.9	12.1	881.4	13800	31100
P8	1178	NR	0	1	2.88	2.52	24.4	33.1	3.1	-	-	-
P9	1302	NR	0	1	2.46	2.06	21.2	37.5	3	-	-	-
P10	1303	NR	1	1	2.76	2.59	33.6	28.8	5.6	-	-	-
P11	442	R	3	3	1.86	1.87	26.8	56.9	46.8	260.6	22200	22800
P12	1339	NR	1	1	1.97	1.84	29.7	48.4	3.3	22.9	3000	8000
P13	1340	R	0	1	2.52	2.17	21.4	59.5	2.2	169.4	8200	39900
P14	1342	NR	0	1	2.86	2.01	16.2	72.4	1.6	21.0	1000	5000
P15	1343	R	0	3	2.78	1.52	12.4	50.6	6.9	-	-	-
P16	1351	NR	0	1	3	2.4	31.8	50.8	2.4	-	-	-
P17	1353	R	0	1	2.73	2.16	22.3	38.8	10.7	-	-	-
P18	1366	NR	1	0	2.91	3.19	31.5	21.7	0	-	-	-
P19	1379	NR	0	1	3.69	2.53	23.4	36	6.5	-	-	-

* Number of HLA-DQ:gluten tetramer-binding effector-memory gut-homing CD4+ T cells per million CD4+ T cells in blood # Number of HLA-DQ:gluten tetramer-binding CD4+ T cells per million CD4+ T cells in gut biopsy. n.d. = not done

Table S2 LC-MS/MS samples

Tissue sample names as shown in figures. Number of samples analysed per patient biopsy per cohort. Total tissue: injection of two technical replicates from the same digest is reported as one sample. LCM: samples collection and analysis is described in material and methods. QC = poor quality sample removed prior to data analysis. - = no samples collected.

Patient ID	Responder	Biopsy Baseline (Before)	Biopsy Day 14 (After)	Total Tissue Before	Total Tissue After	Total Ep. LCM1 Before	Total Ep. LCM1 After	Total Ep. LCM2 Before	Total Ep. LCM2 After	Apical ep. LCM Before
P1	NR	1B	1A	QC	1	2	1	1	-	2
P2	R	2B	2A	1	1	-	-	2	1	2
P3	NR	3B	3A	1	1	-	-	2	1	-
P4	R	4B	4A	1	1	-	-	1	-	-
P5	NR	5B	5A	1	1	-	-	2	1	-
P6	NR	6B	6A	1	QC	1	1	2	1	-
P7	R	7B	7A	1	1	-	-	1	QC	2
P8	NR	8B	8A	1	1	-	-	-	-	2
P9	NR	9B	9A	1	1	-	-	-	-	2
P10	NR	10B	10A	1	1	-	-	-	-	2
P11	R	11B	11A	1	1	1	1	2	2	-
P12	NR	12B	12A	1	1	-	-	-	-	2
P13	R	13B	13A	1	1	2	1	3	2	1
P14	NR	14B	14A	1	QC	-	-	-	-	2
P15	R	15B	15A	1	1	1	1	1	-	1
P16	NR	16B	16A	1	1	-	-	1	-	1
P17	R	17B	17A	1	1	QC	1	QC	2	-
P18	NR	18B	18A	1	1	2	1	3	2	-
P19	NR	19B	19A	1	1	-	-	-	-	-

Table S3 Epithelial cell type gene sets

Mature enterocytes (n=437), goblet cells (n=367) and foetal enterocytes (n= 280) genes are provided as an xlsx-file

Table S4 Cell type protein expression per patient from total tissue and epithelial cell layer proteome.

Cell type specific protein expression per patient from biopsies collected before (B) or after (A) challenge. For each patient, median protein expression (z-score) was calculated from cell type specific proteins in the dataset. n = number of matched cell type specific proteins, NA = no value.

Biopsy	TT Mature enterocyte median z-score n = 214	TT Goblet cell median z-score n = 149	TT Foetal enterocyte median z-score n = 75	LCM Mature enterocyte median z-score, n = 174	LCM Goblet cell median z-score, n = 113	LCM Foetal enterocyte median z-score, n = 40
1B	NA	NA	NA	0.282	-0.222	-0.227
1A	0.500	-0.285	0.394	0.284	-0.164	0.177
2B	0.483	-0.219	-0.131	0.214	-0.030	0.052
2A	-1.285	0.201	0.020	-0.767	0.489	-0.080
3B	0.468	-0.349	-0.047	0.283	-0.131	0.207
3A	0.370	-0.384	0.148	0.375	0.017	-0.141
4B	-0.217	0.080	0.296	-0.154	-0.118	-0.032
4A	-0.667	-0.067	0.411	NA	NA	NA
5B	0.545	-0.027	-0.043	0.395	-0.141	-0.189
5A	0.561	-0.121	-0.189	0.401	-0.037	-0.283
6A	NA	NA	NA	0.371	-0.117	0.090
6B	-0.181	-0.189	0.178	0.210	-0.031	0.041
7B	0.180	-0.043	0.120	0.545	0.104	0.235
7A	-0.661	1.051	1.048	NA	NA	NA
8B	0.828	-0.280	-0.384	NA	NA	NA
8A	0.364	0.004	-0.243	NA	NA	NA
9B	0.294	-0.154	0.160	NA	NA	NA
9A	0.046	0.330	0.213	NA	NA	NA
10B	0.028	0.171	0.276	NA	NA	NA
10A	0.188	0.215	0.307	NA	NA	NA
11B	-1.093	0.809	0.519	-0.153	0.213	-0.316
11A	-1.013	0.810	0.302	-0.670	0.159	0.303
12B	0.387	0.185	0.016	NA	NA	NA
12A	0.071	-0.021	-0.695	NA	NA	NA
13B	0.007	0.357	-0.161	0.187	0.091	0.222
13A	-0.367	1.063	0.541	-1.517	0.759	1.013
14B	0.473	0.293	0.446	NA	NA	NA
14A	NA	NA	NA	NA	NA	NA
15B	0.628	0.255	0.655	0.299	0.031	-0.242
15A	-0.864	0.252	0.116	-1.163	0.105	0.376
16B	1.020	-0.248	-0.578	0.799	-0.184	-0.029
16A	0.200	-0.123	-0.370	NA	NA	NA
17B	0.131	-0.304	-0.177	0.178	-0.073	-0.493
17A	-0.672	0.179	0.301	-0.743	0.177	-0.198
18B	0.098	-0.553	0.029	0.422	-0.232	-0.055
18A	0.902	-0.045	-0.781	0.564	0.084	-0.063
19B	0.311	-0.415	0.030	NA	NA	NA
19A	-0.102	-0.127	-0.208	NA	NA	NA

Table S5 Clinical biochemistry and cytokine levels

Serum biochemistry measured at baseline before gluten challenge. TNF- α was measured in plasma as part of the Bio-Plex Pro Human Cytokine 27-plex Assay as reported in [1].

*Crp: LOD = 0.6. Values <0.6 are shown as 0.5 as this value was used for calculations in Figure 6.

Patient ID	ALAT (U/L)	ASAT (U/L)	ALP (U/L)	GT (U/L)	Transferrin (g/L)	Hemoglobin (g/dL)	Ferritin (μ g/L)	Crp* (mg/L)	TNF- α
P1	13	18	64	64	3.1	14.7	41	0.63	72.4
P2	27	19	62	62	2.3	13.5	73	1.4	117.25
P3	23	28	42	42	2.9	13.5	37	0.5	61.78
P4	14	24	71	71	2.5	15.1	128	2.7	124.03
P5	20	34	50	50	2.7	13.8	23	2.1	86.38
P6	15	21	48	48	2.7	14.2	38	0.5	89.84
P7	27	28	70	70	2.8	14.6	206	1.4	75.91
P8	13	19	51	51	2.7	15.3	76	0.79	93.3
P9	15	24	43	43	2.7	13.2	26	0.5	86.38
P10	20	28	72	72	2.8	12.6	87	0.92	68.87
P11	20	38	82	82	2.1	14.9	132	4.8	90.13
P12	24	23	42	42	2.5	13.6	74	0.5	65.57
P13	27	40	68	68	2.3	15.3	380	4.4	103.66
P14	17	24	32	32	2.8	14.7	37	0.6	98.62
P15	59	51	66	66	2.3	14.8	80	0.5	101.99
P16	8	23	38	38	2.6	13.2	68	0.5	61.95
P17	25	33	33	33	2.6	14.6	70	1.2	65.57
P18	20	24	53	53	2.6	14.4	194	1	86.69
P19	14	22	70	70	2.4	13.4	103	0.6	n.d.

REFERENCES

1. Sarna VK, Skodje GI, Reims HM, et al. HLA-DQ:gluten tetramer test in blood gives better detection of coeliac patients than biopsy after 14-day gluten challenge. *Gut* 2018;**67**:1606-13. doi: 10.1136/gutjnl-2017-314461
2. Risnes LF, Christophersen A, Dahal-Koirala S, et al. Disease-driving CD4+ T cell clonotypes persist for decades in celiac disease. *J Clin Invest* 2018;**128**:2642-50. doi: 10.1172/JCI98819
3. Tutturen AEV, Dorum S, Clancy T, et al. Characterization of the small intestinal lesion in celiac disease by label-free quantitative mass spectrometry. *The American journal of pathology* 2018;**188**:1563-79. doi: 10.1016/j.ajpath.2018.03.017
4. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 2008;**26**:1367-72. doi: 10.1038/nbt.1511
5. Cox J, Neuhauser N, Michalski A, et al. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res* 2011;**10**:1794-805. doi: 10.1021/pr101065j
6. Cox J, Hein MY, Luber CA, et al. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* 2014;**13**:2513-26. doi: 10.1074/mcp.M113.031591
7. Tyanova S, Temu T, Sinitcyn P, et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* 2016;**13**:731-40. doi: 10.1038/nmeth.3901
8. Nusse YM, Savage AK, Marangoni P, et al. Parasitic helminths induce fetal-like reversion in the intestinal stem cell niche. *Nature* 2018;**559**:109-13. doi: 10.1038/s41586-018-0257-1
9. Cox J, Mann M. 1D and 2D annotation enrichment: a statistical method integrating quantitative proteomics with complementary high-throughput data. *BMC Bioinformatics* 2012;**13 Suppl 16**:S12. doi: 10.1186/1471-2105-13-S16-S12
10. Vizcaino JA, Csordas A, del-Toro N, et al. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res* 2016;**44**:D447-56. doi: 10.1093/nar/gkv1145
11. Rustad P, Felding P, Franzson L, et al. The Nordic Reference Interval Project 2000: recommended reference intervals for 25 common biochemical properties. *Scand J Clin Lab Invest* 2004;**64**:271-84. doi: 10.1080/00365510410006324