

HOST RESPONSE

 OPEN ACCESS

The mouse gastric surface epithelial cell and its response to early *Helicobacter pylori* infection

Mattias Erhardsson ^a, Licinia Santos ^a, John Benktander ^a, Sinan Sharba ^a, Kaisa Thorell ^{b,c}, and Sara Lindén ^a

^aDepartment of Medical Biochemistry and Cell Biology, Institute of Biomedicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden; ^bDepartment of Chemistry & Molecular Biology, Faculty of Science and Technology, University of Gothenburg, Gothenburg, Sweden; ^cSciLifeLab, Department of Chemistry & Molecular Biology, Faculty of Science and Technology, University of Gothenburg, Gothenburg, Sweden

ABSTRACT

Helicobacter pylori infection is the main risk factor for gastric cancer. *H. pylori* easily develop antibiotic resistance and evade host defenses. In-depth knowledge of the first barrier that *H. pylori* encounter, the gastric surface mucus-producing epithelial cells (SMCs), may enable improved treatment and prevention. This study aimed to characterize SMC gene expression, mucus glycosylation, and identify how *H. pylori* colonization affects these parameters. The glycosylation of eight *H. pylori*-infected and eight sham control mice was characterized by mass spectrometry. SMCs from five infected and five sham control mice were extracted with laser microdissection (LCM) and sequenced with RNA sequencing (RNA-Seq). SMCs were characterized by high gene expression for proteins secreted into mucus (*Tff1*, *Gkn1*, *Gkn2*, *PscA*, and *Muc5ac*), mitochondria RNA, and cytoskeleton proteins. Mucin glycans were large, complex, heavily fucosylated, and dense with H-antigen motifs. Two main glycosylation pathways ending in H-antigen glycans were identified and corroborated with glycosyltransferase expression. Glycosylation was consistent between *H. pylori*-infected and sham control mice. RNA-Seq data was analysed for differential gene expression, gene set enrichment analysis, and network analysis of functionally-related genes. The analyses revealed that genes required for protein synthesis and oxidative phosphorylation were down-regulated in infected mice. Most up-regulated genes were either interferon-stimulated genes or able to induce interferon production themselves. Depletion of Nkx6-3 occurred in the infected mice, indicating initiation of a pre-cancerous cascade. LCM RNA-Seq of SMCs was thus feasible and enabled characterization of the SMC and definition of a gene set showing how *H. pylori* infection affects SMCs.

ARTICLE HISTORY

Received 27 June 2025
Revised 25 February 2026
Accepted 11 March 2026

KEYWORDS

Helicobacter pylori; gastric mucosa; surface mucus cells; foveolar cells; RNA-Seq; Laser capture microdissection; oxidative phosphorylation; interferons; glycosylation; gastric mucins

Introduction

Helicobacter pylori is a helical-shaped, gram-negative, carcinogenic bacterium that colonizes half of all human stomachs [1]. *H. pylori* infection is the primary risk factor for gastric cancer [2], which is the fourth leading cause of cancer mortality world-wide [3]. Chronic gastritis and gastro-duodenal ulcers can also be caused by *H. pylori* infection [2].

A wide assortment of virulence factors and strategies allow *H. pylori* to circumvent the host immune response in order to establish life-long colonization [4]. One of the challenges in treating *H. pylori* infection is its plastic genome and high propensity to develop antibiotic resistance [5]. The widespread use of antibiotics has thus contributed to the emergence of antibiotic-resistant strains of *H. pylori*, making widespread eradication of *H. pylori* through antibiotics alone unfeasible [6]. This highlights the need for alternative

treatment strategies and a better understanding of the mechanisms underlying *H. pylori* pathogenesis and colonization.

The mucus layer is the first layer of defense and protects the host through several mechanisms, among them by acting as a physical barrier between the bacteria and the gastric epithelium. The mucus layer consists mainly of water and mucins, and mucin glycosylation play crucial roles in defending against *H. pylori* infection [7–9]. A healthy stomach expresses 3 different mucin genes: MUC1, MUC5AC, and MUC6. MUC1 is a transmembrane mucin which can be shed when *H. pylori* bind to the glycans [9]. The secreted mucus layer is segmented so that MUC6 fills the glands, while MUC5AC makes up surface mucus, as well as a mixture of the two where mucins meet in the glandular isthmus [10]. MUC5AC is expressed by the gastric surface mucus-producing epithelial cells (SMCs), and

the protective mechanisms provided by the mucus produced by these cells must be overcome for *H. pylori* to be able to infect the host. These SMCs can also be called foveolar cells, but this term could also include the deeper, glandular MUC6-expressing cells. Another alias for the MUC5AC-expressing cell type is pit cell [11], as these cells are found on both the surface and in the pit, i.e. the upper region of the gastric glands/gastric units. SMCs was the nomenclature of choice for this study.

Mice infected with *H. pylori* strain SS1 is a widely-used animal model to study *H. pylori* pathogenesis, treatment [12], and immunization [13]. Other experiments with SS1-infected mice include the discovery of an immune regulatory mechanism where blocking programmed cell death ligand 1 (PD-L1), thus preventing stimulation of programmed death 1 receptors (PD-1) by gastric dendritic cells, worsened *H. pylori*-induced gastritis [14]. *H. pylori* induced inhibition of mucin production in SMCs, which likely contributes to long-term *H. pylori* colonization, has also been shown using the SS1-infection model [15].

Single-cell RNA sequencing (scRNA-Seq) has been instrumental in characterizing cell-type-specific responses in the gastric mucosa. In humans, scRNA-Seq of immune cells revealed expansion of NKp44⁺ ILC3s, CD11c⁺ myeloid cells, and activated CD4⁺ T and B cells in *H. pylori*-colonized compared to non-colonized individuals. In mice, analysis of gastric epithelial cells identified spasmolytic polypeptide – expressing metaplasia (SPEM) cells lacking *Gif* gene expression, and further demonstrated that *H. pylori* preferentially adheres to SMCs, likely due to chemotaxis toward urea produced by these cells [16–18].

Laser capture microdissection (LCM) offers another precise approach to studying specific gastric cell populations. LCM followed by microarray analysis revealed over 1000 differentially-expressed genes in parietal, chief, and mucus-producing cells from *H. pylori*-infected mice, including up-regulation of prostaglandin E₂ receptor 4 (EP4). Similar approaches identified *Calgranulin A* up-regulation as a marker of MALT lymphoma in gastric lymphoid aggregates and provided evidence that SPEM represents an early stage in gastric tumorigenesis [11,19–21]. Compared to microarrays, RNA sequencing (RNA-Seq) offers transcriptome-wide coverage, both relative and absolute expression levels instead of only relative, and a broader dynamic range [22]. However, the RNA-Seq-based transcriptional profile of normal SMCs during *H. pylori* remains unexplored. The present study addresses this gap, providing insight into early epithelial responses to *H. pylori* infection.

While gastric mucin glycosylation has been studied in mice [23–25] and humans [26], recent advances in glycobioinformatics enables more statistically sound and biologically relevant analysis of glycan data than previously possible, including motif-level differential glycomics [27].

The overall aims of this study were to characterize the normal SMC gene expression and mucus glycosylation as well as determine how *H. pylori* colonization affects these parameters, thereby advancing our understanding of *H. pylori* colonization and pathogenesis. In doing so, we also developed LCM-RNA-seq as a methodology for studying host responses to infection *in vivo*, identified key mucin glycan biosynthesis pathways by integrating glycomics with glycosyltransferase transcriptomics and validated some differentially-expressed genes using immunofluorescence.

Materials and methods

Overview

Reporting of this study followed the ARRIVE 2.0 reporting guidelines [28]. The reporting checklist is provided in supplemental file 10 which can be accessed at the following link: <https://github.com/mattias-erhardsson/lmpc-infection-rnaseq/tree/main/Virulence%20supplemental%20files>.

See Figure 1 for an overview of the methodology employed in this article. Ten mouse stomachs were harvested for the RNA-Seq cohort and 16 for the glycan cohort.

In the RNA-Seq cohort, corpus SMCs were collected using LCM, the RNA was extracted from the cells, and the isolated RNA sequenced. Nine of the samples were successfully sequenced and passed the MultiQC quality control. They were analysed with multiple bioinformatics tools for differential gene expression, gene set enrichment, and network analysis. A manual review of bioinformatics databases was also conducted for salient genes and gene sets. Histological differences between infected and sham control mice were clear in pilot experiments in our laboratory conducted at the same time points and with the same group sizes as the RNA-Seq cohort, which formed the basis for group sizes. Furthermore, group sizes of five are sufficient for RNA-Seq even if one to two samples would fail sequencing [29].

For both the RNA-Seq and glycan cohorts, all stomachs were assessed histologically for gastritis (histological index) and *H. pylori* density using fluorescent *in situ* hybridization (FISH).

Methodology overview

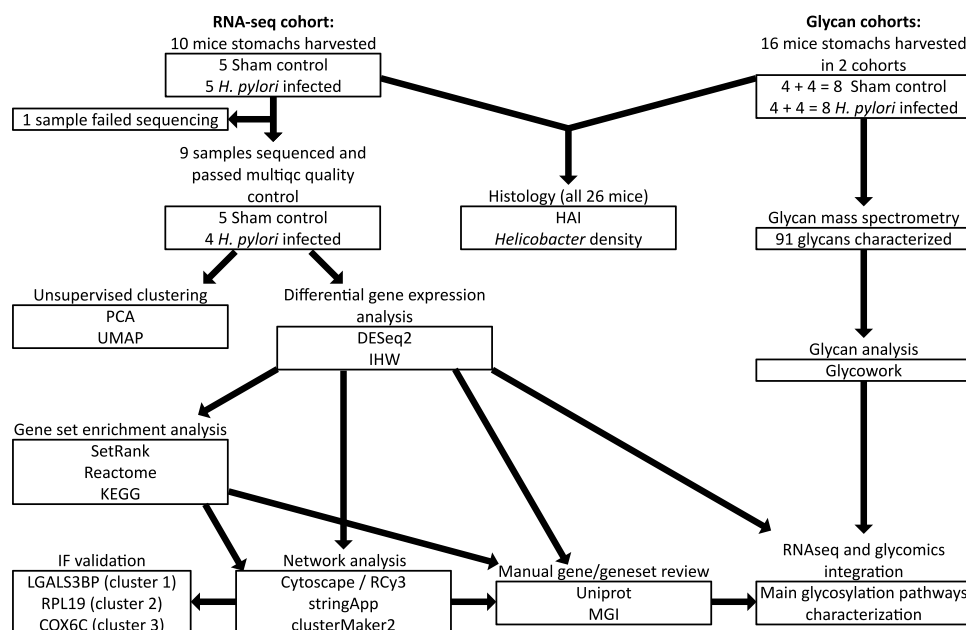


Figure 1. Flowchart of the path from harvested mouse stomachs to analysed gene expression data. PCA = Principal Component Analysis. UMAP = Uniform Manifold Approximation and Projection. DESeq2 = the R package DESeq2. IHW = independent Hypothesis Weighting. HAI = Histological Activity Index. SetRank = the R package SetRank. Reactome = the reactome pathway database was used in the SetRank analysis. KEGG = the Kyoto Encyclopedia of Genes and Genomes database was used in the SetRank analysis. Cytoscape/RCy3 = the R package RCy3 was used for programmatic network analysis in the software Cytoscape. stringApp = the app stringApp was used to annotate genes in the Cytoscape analysis with interactions from the string database. clusterMaker2 = the app clusterMaker2 was used to cluster gene interactions in the Cytoscape analysis. UniProt = the UniProt database accessed with the web interface. MGI = the Mouse Genome Informatics database accessed with the web interface. Glycowork = the Glycowork python package used for glycomics analyses. IF = Immunofluorescence.

Ethical approval

In Sweden, animal experiments require approval from regional ethics committees coordinated by the Swedish Board of Agriculture. Animal suffering is weighed against scientific value and the possibility of using non-animal alternatives. Ethical approval was granted by Göteborgs Djurförsöksetiska Nämnd (Ethic No. 52–2021).

Animals

RNA-Seq cohort

Ten male C57BL/6Ntac mice (Taconic Biosciences, Denmark), 6-weeks old and specific-pathogen-free, were kept at the Laboratory for Experimental Biomedicine, University of Gothenburg. The animals were kept in individually ventilated cages with 12-hour light/dark cycles and had free access to food, water, and nesting material. All five mice within each group were housed in the same cage to enable social interaction while preventing cross-contamination of *H. pylori*. The cages were kept in the same room. No

inclusion or exclusion criteria were set, apart from the ethical approval-related endpoints showing general symptoms (matte fur, sitting still, or weight loss of 10% or more – none of these occurred during the experiment). When sacrificed, the mice were sedated with isoflurane followed by cervical dislocation. The same mice were also used in another 2025 study [30].

Glycan cohorts

A further two cohorts with a total of 16 C57BL/6Ntac mice were kept and sacrificed according to the description above. There were eight mice in each cohort, of which four were *H. pylori* infected and four were sham control. The same mice were also used in another 2025 study [30].

Experimental *H. pylori* infection

A stock culture of *H. pylori* strain SS1 was kept at -80°C in Brain-Heart infusion (BHI) broth (CM1135, Oxoid) containing 20% glycerol. Bacteria from the stock culture was brought up and cultured on

Brucella agar (CM0169, Oxoid) supplemented with 10% horse blood in microaerobic conditions for four days. The bacteria were then transferred and cultured overnight in BHI with 10% fetal bovine serum (FBS) in microaerobic conditions. Bacteria were then suspended in phosphate-buffered solution (PBS) to a 600 nm optical density (OD_{600}) of 1.0. Quality control involved inspecting the suspension under a microscope to ensure the viability of bacteria and re-culturing the suspension on agar for four days and counting colony forming units (CFUs). The mice were infected through oral gavage with the *H. pylori* SS1 suspension, two doses of 200 μ L (3.5×10^7 CFU/mouse) given one day apart when mice were eight weeks old. The sham controls were given a corresponding sham treatment of oral gavages with 200 μ L PBS. The mice were harvested two weeks after the first *H. pylori* infection dose.

Snap-freezing tissue embedded in optimal cutting temperature compound (OCT)

Harvested stomachs were pinned to filter paper with the lumen facing up, using needles piercing the edges of the tissue at the forestomach and antrum to keep the tissue flat and spread out against the filter paper. A longitudinal strip, encompassing forestomach to antrum, was cut out slightly to the side of the middle/curvatura minor of the harvested stomach (Figure 2(A)).

Correct orientation of the tissue when embedded into OCT was crucial for the downstream LCM. To achieve proper orientation, the cut edge was placed

facing down in the embedding mold by balancing the tissue stuck to filter paper on its side, using two needles inserted into the tissue paper above to balance the tissue on the cut side (Figure 2(B)). OCT was carefully added to the mold with balanced tissue to avoid air bubbles. The filled mold was then snap-frozen by transferring the mold to a beaker with a mixture of dry ice and Cryospray Aerosol (Article number 00211, HistoLab, Gothenburg, Sweden). To minimize RNA degradation, the protocol was optimized so that it took at most ten minutes between the mouse being sacrificed to the gastric strip being frozen in OCT. Frozen molds were sealed together with desiccant bags in 50 mL centrifuge tubes and immediately put on dry ice until they could be stored at -80°C .

Formalin-fixed paraffin-embedded (FFPE) histological assays

Gastric corpus tissue for FFPE histological assays were collected in parallel to the OCT-embedded tissue from the same stomachs. The tissue was fixed in buffered formaldehyde 4% aqueous solution ($\text{pH } 7.0 \pm 0.2$, 9713.100, VWR) at harvest and embedded in paraffin.

Gastritis was assessed on Hematoxylin and Eosin (H&E) stained slides. Sectioned tissue (4 μm) was deparaffinized, hydrated, and stained first with hematoxylin, and then after a wash in water stained with eosin. The slides were washed again and then dehydrated with a series of baths with increasing concentrations of ethanol, then isopropanol, and finally xylene.

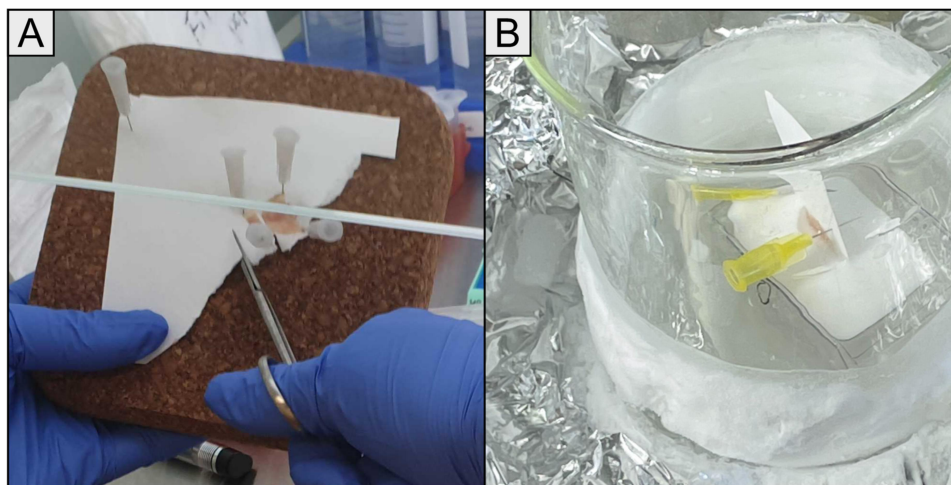


Figure 2. Orienting and embedding gastric tissue in OCT. (A): The stomach was opened at the curvature major and laid out on filter paper with the lumen facing up. Needles in the sides of the forestomach and antrum were used to keep the tissue flat on the tissue paper. A longitudinal strip for OCT embedding slightly to the side of the middle (curvature minor) was then cut from the antrum to the forestomach. (B): Needles in the filter paper were used to balance the tissue on its side, with the previously cut side facing down in the mold. OCT was carefully poured to avoid bubbles. The mold was transferred with forceps to a prepared beaker enclosed with aluminum foil containing a mixture of dry ice and Cryospray, which rapidly froze the OCT.

Mounted slides were scored blinded according to the Histological Activity Index (HAI) [31].

Immunofluorescence (IF) was used to validate RNA-Seq results on a protein level. Slides were sectioned (4 µm), deparaffinized, and hydrated. Antigen retrieval was performed by boiling slides in 10 mM citric acid buffer (pH 6) with 0.05% Tween 20 (v/v) for 20 minutes. Slides were allowed to cool for 30 minutes before drawing hydrophobic barriers around the tissues. Tissues were washed thrice in wash buffer (0.5% BSA (m/V) + 0.05% Tween-20 (V/V) in PBS) with a 1-minute incubation per wash, permeabilized (10-minute incubation with 0.5% BSA (m/V) + 0.5% Tween-20 (V/V) in PBS), washed as before, incubated with blocking buffer (1% BSA (m/V) + 5% Normal Goat Serum (V/V) + 0.01% Tween-20 (V/V) in PBS) for 1 h, incubated with primary antibody overnight, washed three times with wash buffer with 5 minute incubations, then incubated with secondary antibody for 1 h, washed again with three 5-minute wash buffer washes, and then an extra triple wash with PBS. 200 µL liquid was added to the tissues in all steps after drawing hydrophobic markers (washes, permeabilization, blocking, primary antibody incubation, secondary antibody incubation), and the liquids were removed by tilting the slides in an incubation chamber. All incubations were performed at room temperature (except primary antibody incubation which was at 4°C) and in the dark. The primary antibodies were RPL19 (Rabbit Polyclonal Antibody diluted 1:1000, article number PA5-101384, Fisher Scientific, Waltham, USA), LGALS3BP (Recombinant Rabbit Monoclonal Antibody diluted 1:100, article number MA5-50412, Fisher Scientific, Waltham, USA), and COX6C (Recombinant Rabbit Monoclonal Antibody diluted 1:50, article number MA5-63536, Fisher Scientific, Waltham, USA). The secondary antibody for all three primary antibodies was a Goat anti-Rabbit Alexa Fluor™ 594 diluted 1:200 (article number A-11037, Fisher Scientific, Waltham, USA).

Slides were blinded, and for each slide the SMCs in the gland with the most intense signal was measured with Fiji [32] to a depth of approximately 70 µm from the surface and the mean signal intensity of this gland used as the readout. For the antibodies against cytoplasmic proteins (LGALS3BP and RPL19) the nuclei were excluded from the region of interest (ROI), while the ROI included areas with DAPI signal for the mitochondrial antibody (COX6C) since most of the signal was expected to be in close proximity to the nuclei and thus coincide with much of the DAPI signal.

Visualization of bacteria was conducted on FISH slides with two probes. The probes were EUB 338 labeled with Cy3.5 (5'GCTGCCTCCCGTAGGAGT3') to detect all bacteria, while the other probe Alexa 488 (5' TCTCAGGCCGGATACCCGTCATAGCCT3') was

specific for *Helicobacter* spp. Slides were deparaffinized, hydrated, washed with PBS, and air-dried. Dried slides were hybridized with a solution containing 40% (v/v) formamide, 0.1% sodium dodecyl sulfate, 0.9 M NaCl, 20 mM Tris-HCl (pH 7.4) and 10 ng/µL probe. Slides were incubated for 18 h at 37 °C in a dark humidified chamber with 40% (v/v) formamide. After incubation, the slides were washed in a solution containing 0.9 M NaCl and 20 mM Tris-HCl (pH 7.4) for 20 minutes at 50°C. Slides were then dipped in ultrapure water and air-dried in darkness. Dried slides were mounted with Prolong Gold anti-fade reagent and DAPI (Thermo Fischer Scientific). Blinded counting of *Helicobacter* bacteria was conducted in the pits since this reduces the risk of falsely low counts due to the potential loss of surface mucus during processing. Counting was done with an Eclipse 90i fluorescence microscope (Nikon, Tokyo, Japan). The results show the average counts in five 40x-magnified fields.

Differences in histological assays between the two groups were tested with Mann-Whitney U, correlations between IF signal intensity and RNA-Seq gene expression levels were tested with Spearman Rank Correlation, and the *p*-values adjusted with the Benjamini-Hochberg method [33] in the R programming language [34].

Slide preparation for LCM

RNA degradation is slowed in fresh-frozen tissue stored at −80°C, yet not stopped. The median number of days it took between harvest and sectioning for the OCT-embedded tissue analysed in this article was eight, and within a range of six to nine days. All solutions and materials that encountered the tissue from this step and onwards were prepared to be RNase-free.

OCT-embedded tissue was sectioned to ten micrometer sections in a −20°C Cryostat and transferred to PEN-membrane slides (article number 415,190-9081-000, Carl Zeiss, Oberkochen, Germany). A pilot LCM with subsequent RNA-extraction of 12 sections yielded 15 ng purified RNA available for sequencing after quality control. Based on this pilot, we sectioned three slides with at least four sections on each slide for every sample. Slides were air-dried for two minutes in the cryostat after sectioning and then transferred to centrifuge tubes with desiccant bags, in which they were briefly stored at −80°C before staining.

Slides were stained with 1% Cresyl Violet (article number 10,045,203, Fisher Scientific, Waltham, USA) using ice-cold solutions with the following steps performed inside a hood. Slides were first transferred from the centrifuge tubes to a 70% ethanol bath for fixation

lasting two minutes. The slides were then put on a staining rack and stained with 1% Cresyl Violet with 200 μ L of the solution applied directly to each sample using a pipette. The staining solution was rinsed away with 70% ethanol after staining for 30 seconds. After rinsing, the slides were dipped in a series of baths consisting of 70% ethanol, 100% ethanol, and then another bath of 100%. The slides were then air-dried for one minute on a paper towel, taking care not to let the paper touch the tissue or the membrane. Ensuring a quick staining in cold liquids, followed by maximally dehydrating the samples was likely crucial to avoid obliteration of RNA. Finally, the slides were put in new centrifuge tubes with desiccant bags, and placed in dry ice before swiftly being stored at -80°C .

LCM

SMCs were collected using LCM with a PALM MicroBeam (Carl Zeiss, Oberkochen, Germany) at the Center for Cellular Imaging, Sahlgrenska Academy, Gothenburg, Sweden.

The slides were thawed to room temperature while sealed inside the centrifuge tubes. To ensure that we collected SMCs from the corpus we used the forestomach transitional zone as a reference point and stained the tissue with cresyl violet to avoid parietal cells (Figure 3(A)). The collection of SMCs was carried out across the entire corpus, with the exception of the first few crypts in the forestomach-corpus transitional zone as this area differs from the rest of the corpus [31]. We also avoided the corpus-antrum transitional zone to minimize the risk of collecting antral cells. LCM for each sample took four hours, but the staining process dehydrates the tissue so RNA damage caused by the extended exposure to room temperature was minimal. Cells were cut with the 20x objective using the settings cutting speed 70, cutting energy 44, focus 74, and cutting laser function JointCut. Cut areas were catapulted with the 5x objective and catapult energy 100. Catapulting was performed in cell-free areas wherever possible, and to achieve this, approximately one-third of the collected area consisted of cell-free PEN membrane on the luminal side of the epithelium. Catapulted areas were collected in AdhesiveCap 500 clear (article number 415,190-9211-000, Carl Zeiss, Oberkochen, Germany) (Figure 3(B)). Samples in adhesive caps were then lysed by adding 350 μ L of Buffer RLT from RNeasy Micro Kit (article number 74,004, Qiagen, Hilden, Germany) containing β -Mercaptoethanol to the tubes, vortexing and then incubated upside-down for 30 min. Lysates were then spun down for 10 minutes at 7300 rcf and stored at -80°C .

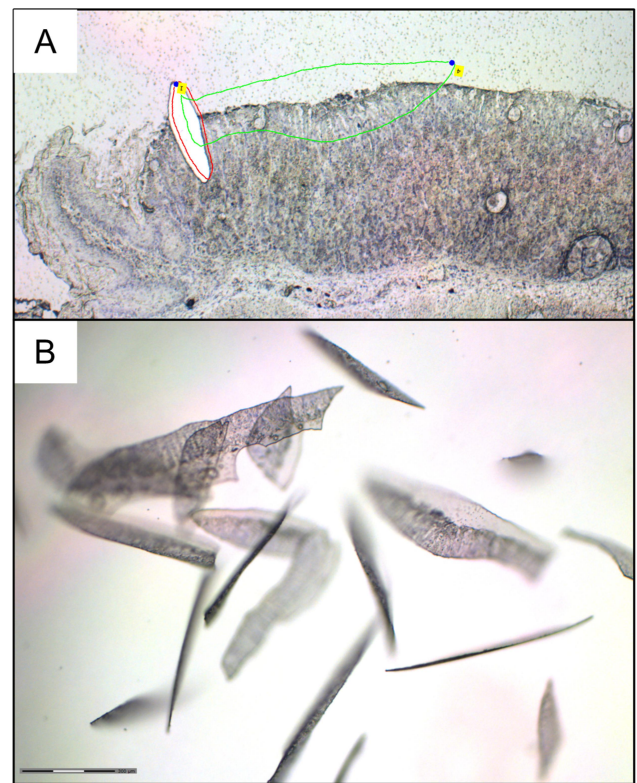


Figure 3. LCM collection of SMCs. A: Micrograph example of LCM cutting area. Areas for cutting were drawn starting from the forestomach transitional area as a navigational landmark. The transitional area itself (red outline) was removed but not collected. The green outline was one of the drawn areas of SMCs, which would later be cut and catapulted. The deeper layer containing parietal cells was avoided. Most cut areas contained cell-free parts of the membrane to minimize RNA damage from the laser. B: Micrograph of a collection cap after several areas had been catapulted.

RNA extraction

RNA was extracted using RNeasy Micro Kit (article number 74,004, Qiagen, Hilden, Germany). After thawing the lysed samples, the QIAGEN protocol “Total RNA Isolation from Microdissected Cryosections” was followed, starting from step five (RNeasy® Micro Handbook 04/2003, pp 20). Two μ L of purified RNA were aliquoted for quality control. RNA samples were put on dry ice immediately after aliquoting and then stored at -80°C .

RNA quality control and sequencing

Initial quality control of purified RNA Integrity Number (RIN-values) was performed with TapeStation on High Sensitivity RNA ScreenTape (Agilent, Santa Clara, USA) according to the manufacturer’s instructions.

Samples were shipped to National Genomics Infrastructure (NGI), Solna, Sweden for further quality control and subsequent sequencing. The second round of quality control was performed using Bioanalyzer (Agilent, Santa Clara, USA).

Libraries were prepared with SMARTer Total Stranded RNA-Seq, Pico input mammalian – V3 (Takara Bio, Kusatsu, Japan). Validation of constructed libraries were performed with Fragment Analyzer (Agilent, Santa Clara, USA).

Samples with successfully constructed libraries were sequenced with NovaSeq 6000 (Illumina, San Diego, USA). We aimed for a minimum sequencing depth of 30 million reads for each sample.

Bioinformatics pipeline of fastq-files to gene counts

The nf-core/rnaseq [35] analysis pipeline was run by NGI at Uppsala Multidisciplinary Center for Advanced Computational Science, Uppsala, Sweden (UPPMAX).

Analysis of gene expression data

Gene expression analysis consisted of three main approaches:

- (1) Network analysis of gene interactions between significantly differentially-expressed genes.
- (2) Gene set enrichment analysis of all genes ranked by adjusted p -values.
- (3) Manual review of significantly differentially-expressed genes with the most extreme fold changes and/or expression levels.

The programmatic parts of the analysis was performed using R 4.3.2 [34]. Figure 1 summarizes the analysis process.

First, counts from the bioinformatics pipeline were imported with tximport [36].

Differential gene expression was determined with DESeq2 and p -values adjusted with independent hypothesis weighting (IHW) [37,38]. Clustering of gene expression was investigated with Principal Component Analysis (PCA) [39] and Uniform Manifold Approximation and Projection (UMAP) [40].

Genes were annotated with gene sets from Reactome and Kyoto Encyclopedia of genes and Genomes (KEGG) [41,42]. To achieve as complete annotation as possible to Reactome, the union of annotation with both biomaRt [43,44] and AnnotationDbi [45] was used to annotate the genes with Reactome gene sets. With biomaRt, the Ensembl version 109 and Reactome release 85 were used. With AnnotationDbi, reactome.

db version 1.84.0 was used, which corresponds to Reactome release 84. KEGGREST was used to annotate the genes with KEGG release 107. Gene set enrichment analysis was performed with SetRank [46]. SetRank was run in ranked mode, with the gene list ranked by adjusted p -value from DESeq2/IHW. SetRank accounts for the overlap between gene sets, which was also used to adjust the p -values together with a multiple testing adjustment using the Holm method [46]. The background gene list consisted of all genes detected among the samples (genes with a count > 0).

Associations between significant genes and gene sets were explored programmatically in CytoScape using R package RCy3 with STRING and ClusterMaker2 apps installed [47–50]. STRING database version 12 was used for annotation. Data wrangling and visualization was facilitated with tidyverse packages [51].

Manual review of interesting genes and gene sets was performed in Uniprot and MGI, as well as by following the links provided in these database entries [52,53].

Marker genes were gathered from both PanglaoDB [54] and a scRNA-Seq dataset from mouse stomachs [17]. Marker genes from PanglaoDB were filtered based on the supplied variables canonical marker = 1, and specificity_mouse = 0, to maximize specificity. This gene list was supplemented with the marker genes found in the supplemental data of the aforementioned scRNA-Seq. All marker genes were selected for cell types which could be expected to be present in the collected tissue. Finally, marker duplicate genes which were assigned to multiple cell types in this selection were removed to further increase specificity. Before this removal of duplicates, the metaplastic cell marker genes were manually adjusted by removing *Muc6* [17], since this would otherwise overlap with the only marker gene for the important cell type mucus neck cells. The data-driven approach with strict filtering to maximize specificity, resulted in exclusion of some canonical marker genes that lacked sufficient specificity, e.g. *Cd11c* (also known as *Itgax*) a dendritic cell marker, which was also annotated as a canonical marker for macrophages [54].

Glycan analysis

Crude mucin extractions on corpus tissue (i.e. not LCM SMCs), followed by dot blot, β -elimination and LC-MS, were conducted as described previously [55].

Glycomics analyses were performed with glycowork [56] in Jupyter [57] running Python 3 [58], as well as hierarchical agglomerative clustering (HAC) in R [59] based on glycan relative abundances. The HAC was performed with default settings (Euclidean distance

metric and complete linkage) with the package pheatmap [59].

Glycans were visualized according to the Symbol Nomenclature for Glycans (SNFG) guidelines [60].

Randomization and blinding

Harvesting and LCM was carried out sequentially, with the sham control mice first and the infected mice last. All other steps were pseudorandomized through no order being applied, and consequently were carried out in different orders.

Histological assessments were blinded during scoring and signal quantification.

Results

H. pylori was only detected in infected mice, and infected samples had more inflammation than sham control mice

Infection status and histological activity index (HAI) for the mice were previously reported as part of a larger cohort [30]. In Figure 4, we summarize the data for animals included in this study. Infection status was

confirmed by FISH using a *Helicobacter*-specific probe (Figure 4(A) and (C)). *H. pylori*-infected mice obtained higher histological activity index (HAI) scores than their sham control counterparts (Figure 4(B, D and E)). Inflammation, characterized by the presence of leukocytes in the submucosa, was the parameter that contributed the most to the scores in infected mice. Histological changes were moderate (HAI ranged 2–9 in infected mice and 0–3 in sham control mice, theoretically possible maximum score in HAI was 24), and the surface epithelium was relatively intact with no sign of infiltration of other cell types.

RNA yield and quality from cells collected with LCM were sufficient for RNA-Seq, without amplification prior to library prep

The amount of purified RNA obtained from the 12 tissue sections processed from each mouse ranged from 16.71 to 128.58 ng from a collected area of 1.68 – 6.05 mm². There was a strong correlation between collected area and quantity of purified RNA (Supplemental Figure S1). Both the quality and quantity of RNA were sufficient for the kit SMARTer Total Stranded RNA-seq, Pico input mammalian – V3

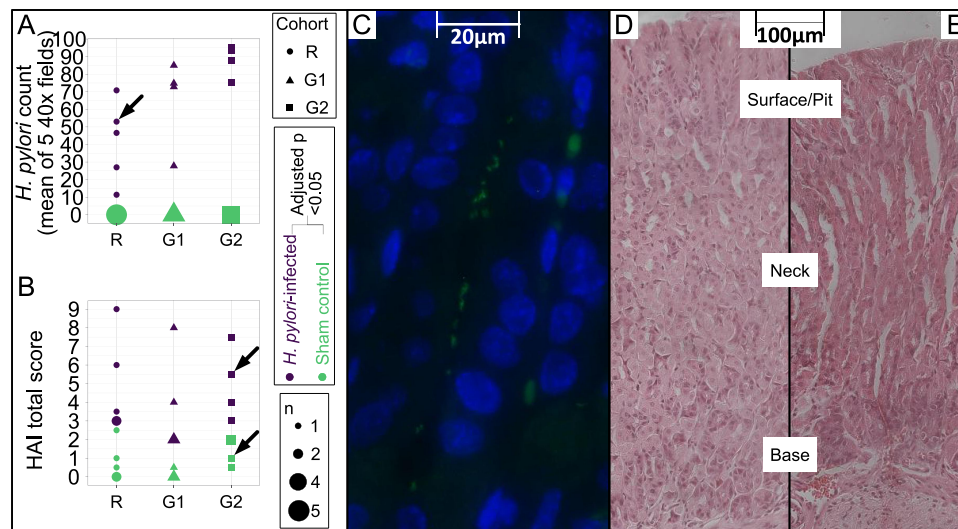


Figure 4. Histological assessments of *H. pylori* density and inflammation in *H. pylori*- and sham control mice. The RNA-Seq cohort was denoted as R, while the two glycan cohorts were denoted G1 and G2. A: Dotplot of *H. pylori* counts in gastric corpus. Slides were stained with FISH-probes for bacteria and *H. pylori* spp. The mean *Helicobacter* counts of 5 fields at 40x magnification was used. Counting was performed blinded. *n* = Number of datapoints on each coordinate of the graph to account for overlapping datapoints. The arrow points to the count obtained for the photomicrograph shown in panel C. B: Dotplot of histological activity index (HAI) total scores. Scoring was performed blinded on H&E-stained slides. Infected mice had significantly (adjusted *p*-value < 0.05, Mann-Whitney U) higher HAI total scores than sham control mice. *n* = Number of datapoints on each coordinate of the graph to account for overlapping datapoints. The arrows point to the scores obtained for the photomicrographs shown in panel D and E. C: Gastric section of *H. pylori*-infected mouse at 14 days after infection. *H. pylori* was visualized with a *Helicobacter* fluorescent probe (green) and nuclei of cells with DAPI (blue) at 40x magnification. D and E: Gastric section stained with H&E from a sham control (D) and an *H. pylori*-infected mouse (E), 20x magnification. The three major layers of epithelial cells surface/pit, neck, and base are marked.

(Supplemental Figure S1) which recommend a minimum DV200 > 50% and input amount > 10 ng.

RNA-Seq on purified RNA was successful for all samples except one from the infected group. This sample did not yield a successful library and thus could not be sequenced. We assessed that all sequenced samples passed quality control with MultiQC [61]. The mean sequencing depth was 60.9 million reads (standard deviation 6.87), and all sequenced samples were above our desired minimum sequencing depth of 30 million reads.

LCM captured SMCs with high precision

The histology-based cellular collection method ensured that a clearly visible transition to deeper tissue where non-SMC cell types are present could be avoided. Our hypothesis was that this should entail a high purity of SMCs in sequenced material. Therefore, we investigated the gene expression of marker genes for cell types which could either be present due to cutting too deep into the mucosa, or the presence of non-SMC cell types in this outer mucosal layer. See Supplemental Table S2 and Supplemental Table S3 for details of the resulting marker genes.

Visualization of the Transcripts per million (TPM) for the marker genes showed high expression of SMC marker genes *Tff1* [17,54] and *Muc5ac* [17,54], low expression of parietal cell marker genes *Atp4a* [17] and *Atp4b* [17,54] (sub-units of gastric hydrogen potassium ATPase) as well as endothelial marker gene *Tspan8* [54], and a negligible expression of other marker genes (Figure 5).

While the gene expression levels of other marker genes were small, there was a significant difference in gene expression (adjusted p -value < 0.05) between infected and sham control mice for four of the marker genes, namely *Zbp1* [54] (plasma cells), *Tet2* [54] (monocytes), *Xdh* [54] and *S100a13* [54] (both for endothelial cells). There were no differences between infected and sham control groups for the expression levels of the other 11 cell markers for plasma cells, the two other cell markers for monocytes, or the 42 other cell markers for endothelial cells. The search function in PanglaoDB web interface showed that the plasma cell marker *Zbp1* is also present in cell clusters annotated as epithelial cells [54]. Furthermore, the PanglaoDB web interface showed that monocyte marker *Tet2* is present in several different cell clusters of mucosal epithelia, and endothelial markers *Xdh* and *S100a13* are present in many cell types which include epithelial and mucus

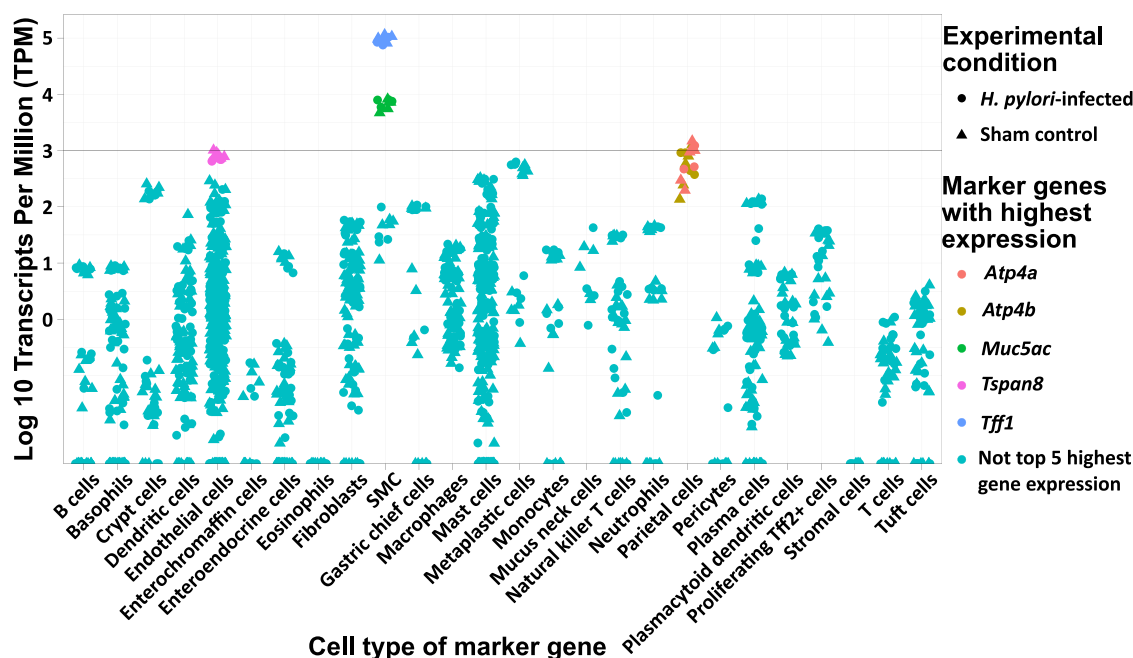


Figure 5. Horizontally jittered dotplot of absolute marker gene expression. Marker genes for cell types that could be present in collected tissue were chosen. The absolute gene expression was quantified as transcripts per million (TPM) and plotted on a 10-logarithmic scale since low expression genes would otherwise be drowned out by the most highly-expressed genes. The top 5 genes in terms of gene expression were highlighted through color annotation, where the gastric surface mucus-producing epithelial cell (SMC) marker gene *Tff1* dwarfed the gene expression of all other marker genes. These 5 genes were also the only genes with a TPM > 1000 in at least 1 mouse (horizontal line).

producing cells [54]. Finally, the log2 fold changes for the endothelial cell markers *Xdh* and *S100a13* were opposite of each other, where *Xdh* was up-regulated while *S100a13* was down-regulated with infection. Thus, we conclude that a vast majority of the cells captured with LCM are SMCs.

SMC gene expression was dominated by genes coding for secreted mucus proteins, mitochondrial function, and cellular structure, and mucin glycosylation was characterized by large and complex fucosylated O-glycans

We next investigated what signifies the typical gene expression of SMCs. The genes were ordered by the mean TPM in sham control mice. Genes were filtered based on the lowest expression of *Muc5ac* among sham control mice as a lower limit (Table 1). *Tff1* was the most highly-expressed gene in normal SMCs, and *Tff1* was also one of the genes coding for proteins secreted into the mucus. Products of the highly-expressed *Gkn2*, *Gkn1*, *Pscs*, *Muc5ac*, and *Lgals2* have also been shown to be secreted into mucus [64,72,73]. The genes *mt-Rnr1*, *mt-Rnr2*, *mt-Atp6*, and *mt-Co1* were also highly expressed, and the products of these genes are known to be integral for mitochondrial functions including oxidative phosphorylation (OXPHOS) [52]. Finally,

Krt8, *Krt19*, and *Actg1* were highly-expressed genes with known structural functions [52].

A primary function of SMCs is to secrete mucus containing MUC5AC. Mass spectrometry revealed that the mouse gastric mucin O-glycans were large and had a bimodal distribution where glycans of 5–6 or 10–11 monosaccharides in size had the greatest combined relative abundances (Figure 6). In total, 100 glycans were characterized (Supplemental File 4).

In line with the glycan size distribution, the two most abundant glycans were Fuc(α1–2)Gal(β1–3)[Fuc(α1–2)Gal(β1–4)GlcNAc(β1–6)]GalNAc and Fuc(α1–2)Gal(β1–3/4)GlcNAc(β1–3)[Fuc(α1–2)Gal(β1–4)GlcNAc(β1–6)]Gal(β1–3)[Fuc(α1–2)Gal(β1–3/4)GlcNAc(β1–6)]GalNAc, constituting 9 and 7% respectively of the glycan-relative abundance. These two glycans, composed of 6 and 11 monosaccharides respectively, were also representative of another pattern wherein the glycans were heavily fucosylated (Figure 7). The mean relative abundance of fucosylated glycans was 66%. A heatmap of all glycans clustered with hierarchical agglomerative clustering (HAC) is shown in supplemental figure S2. The 10 most abundant glycans made up 43.7% of the total glycan pool and the remaining 90 glycans had mean relative abundances no greater than 2.6%.

The main glycosylation pathway(s) were elucidated by identifying the top five most abundant glycans and

Table 1. Genes with the highest gene expression in the normal sham control mouse gastric corpus SMCs. The genes were ordered by the mean TPM in sham control mice. The filtering approach was to take all genes with a mean TPM at least as high as the lowest TPM of *Muc5ac* across all samples (TPM ≥ 4690). 18 genes remained after filtering. Genes marked with # (*Tff1* and *Muc5ac*) were annotated as both canonical marker genes in PanglaoDB, as well as marker genes in Bockerstett et al. 2020 [17]. Genes marked with * (*Gkn2*, *Gkn1*, *Pscs*, and *Lgals2*) were assigned as marker genes in Bockerstett et al. 2020 [17], but were annotated as non-canonical marker genes in PanglaoDB.

Gene	Mean TPM in sham control mice	Secreted to mucus	Mitochondrial function	Structural function		
<i>Tff1</i> [#]	99264	Secreted [62]	Mitochondrion 12S rRNA [63]	Type II cytoskeleton keratin [67]		
<i>mt-Rnr1</i>	96937					
<i>mt-Rnr2</i>	61903		Mitochondrion 16S rRNA [63]			
<i>Yam1</i>	56432	Secreted [64]	OXPHOS complex V subunit [66]	Potential organizer of membrane lipid rafts [69]		
<i>Gkn2</i> [*]	38260					
<i>Gkn1</i> [*]	37150	Secreted [64]				
<i>Gm26917</i>	22724	Secreted [65]				
<i>Pscs</i> [*]	17950					
<i>mt-Atp6</i>	13874					
<i>Dpcr1</i>	10123	Also known as Muc13, but not known to be secreted [52]	OXPHOS complex IV subunit [68]	Cytoskeleton γ-actin, which in epithelial cells is localized in the apical side [71]		
<i>Krt8</i>	6670	Secreted [10]				
<i>Muc5ac</i> [#]	6613					
<i>mt-Co1</i>	6243		Type I cytoskeleton keratin [67]			
<i>Mal</i>	5348	Secreted [70]				
<i>Lgals2</i> [*]	4948					
<i>Actg1</i>	4887					
<i>Krt19</i>	4833					
<i>Rn7sk</i>	4697					

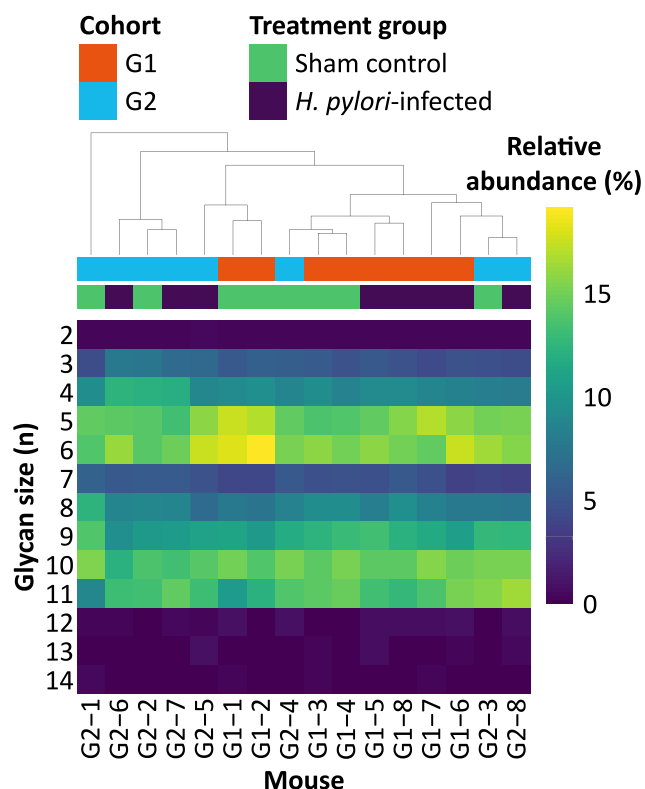


Figure 6. Relative abundance heatmap of glycan sizes. Columns are mice and rows are glycans of different sizes. Size was defined as the number of monosaccharides in the glycan. Each cell shows the pooled relative abundance of all glycans of the same size in that mouse. The dendrogram at the top of the heatmap shows the hierarchical agglomerative clustering (HAC) which was used to cluster the mice. Cohort and treatment denoted by color.

mapping all intermediate steps between them and the initial precursor glycan, commonly referred to as Tn [74]. Additional detected glycans were used to fill in steps linking the top five glycans and the Tn antigen, and any remaining gaps were bridged by interpolating presumed intermediary glycans. All these glycans were O-linked glycans. The RNA-Seq data were subsequently mined for glycosyltransferases by manually reviewing detected genes annotated with glycosyltransferase activity (GO term GO:0016757). These genes were ranked by mean expression (TPM) in sham control mice and annotated with supporting evidence from the literature until sufficient glycosyltransferases had been identified to account for all steps in the pathway (s). Ultimately, the 15 most abundant glycosyltransferases with known activity on mucin O-glycans could explain the complete biosynthetic chain of the top five most abundant glycans (Table 2).

The network resulting from mapping the glycosylation of the top five most abundant glycans formed two main pathways ending with either the most abundant

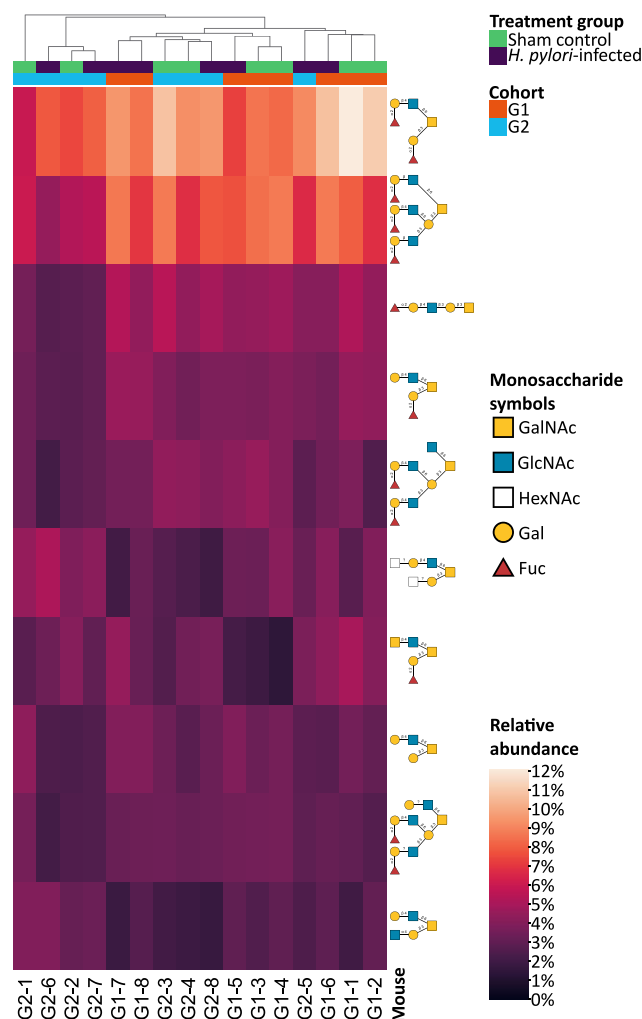


Figure 7. Heatmap of relative abundances of the top 10 most abundant glycans characterized with mass spectrometry. The dendrogram at the top of the heatmap shows the hierarchical agglomerative clustering (HAC) which was used to cluster the mice. Cohort and treatment denoted by color. Glycans were visualized using symbols according to the Symbol Nomenclature for Glycans (SNFG) guidelines. GalNAc = N-Acetylgalactosamine. GlcNAc = N-Acetylglucosamine. HexNAc = Either GalNAc or GlcNAc but it could not be determined which. Gal = Galactose. Fuc = Fucose.

glycan 031 (Fuc(α 1-2)Gal(β 1-3)[Fuc(α 1-2)Gal(β 1-4)GlcNAc(β 1-6)]GalNAcol), or the second most abundant glycan 085 (Fuc(α 1-2)Gal(β 1-3/4)GlcNAc(β 1-3)[Fuc(α 1-2)Gal(β 1-4)GlcNAc(β 1-6)]Gal(β 1-3)[Fuc(α 1-2)Gal(β 1-3/4)GlcNAc(β 1-6)]GalNAcol). Glycan 085 was capped with three H antigen (Fuc(α 1-2)Gal(β)GlcNAc) terminal motifs, while glycan 031 was capped with one H antigen terminal motif. See Figure 8 for a diagram of the main glycosylation pathways. Spectra annotation of diagnostic ions for glycan 085 is shown in Supplemental Figure S3.

The pathways were charted by identifying glycosylation steps between them and the initial “Tn” glycan

Table 2. The top 15 most abundant glycosyltransferases with known activity on mucin O-glycans. These glycosyltransferases could explain the biosynthesis of the top 5 most abundant glycans. 12 of these 15 glycosyltransferases could catalyze at least one of the steps required for the glycosylation of these top 5 most abundant glycans.

Gene	Mean TPM	Reaction(s)	Part of glycosylation of top 5 most abundant glycans	Evidence
Gcnt3	294	3 different transfers of GlcNAc to the following: 1: Core 1 GalNac =>Core 2 2: Core 3 GalNac =>Core 4 3: Linear polylactoseamine Gal =>I branch	Yes	Mouse [75]
C1galt1	190	Gal(β1-3) to Tn to create Core 1	Yes	Mouse [76]
B3gnt3	105	GlcNAc(β1-3) to Gal(β1-4)GlcNAc or Gal(β1-3)GlcNAc	Yes	Human [77,78]
Fut2	82	Fuc(α1-2) to terminal Gal	Yes	Mouse [25]
Galnt12	74	GalNac to serine or threonine ("Tn" Glycan)	Yes	Human [79]
Galnt7	60	Tn	Yes	Human [80]
Galnt4	60	Tn	Yes	Mouse [81]
Galnt2	60	Tn	Yes	Human [82]
B4galt3	54	GalNac(β1-4) to GlcNAc	No	Human [83]
Fut4	43	Fuc(α1-3) to GlcNAc	No	Mouse [84]
Galnt3	38	Tn	Yes	Mouse [85]
B4galt1	33	Gal(β1-4) to GlcNAc in Gal(β1-4)GlcNAc	Yes	Mouse [86]
Galnt6	26	Tn	Yes	Human [87]
B3gnt2	23	GlcNAc(β1-3) to Gal(β1-4)GlcNAc	No	Mouse [88]
B3galt5	20	Gal(β1-3) to either GlcNAc or GalNac	Yes	Human [89,90]

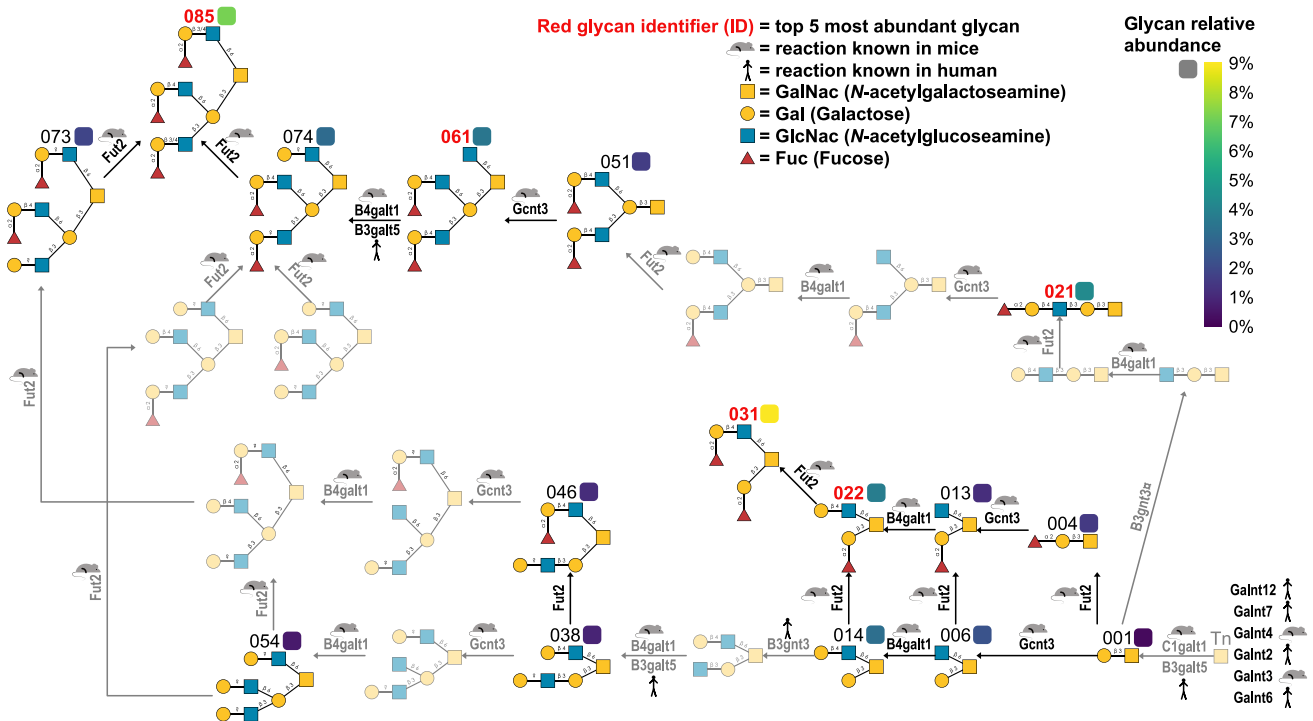


Figure 8. Main glycosylation pathways of surface mucins in the mouse gastric corpus. The top five most abundant glycans are denoted with glycan identifiers (id, see Supplemental File 4) colored red. The color of the boxes next to the id shows the mean relative abundance of that glycan.

(GalNAcol). Intermediate glycans which were identified by mass spectrometry have ID with black text. The transparent glycans are intermediate glycans which were not identified with mass spectrometry but rather are inferred, for example, the Tn glycan which is the precursor to all other O-linked glycans on mucins. The most abundant glycans, 031 and 085, form the end of the main glycosylation pathways.

Glycosyltransferases conducting these steps were identified by datamining the RNA-Seq data to identify the most abundant glycosyltransferases, and then which of these were known to catalyze the required reactions to connect the identified glycans. Glycosyltransferase activity evidence comes directly from experiments on mice (denoted with mouse symbol), or is inferred based on homology to human proteins where such a reaction

is known (denoted with stick figure symbol). Glycans were visualized using symbols according to the SNFG guidelines.

Mucin glycosylation did not change with *H. pylori* infection

Glycomics analysis was deepened in glycowork by analysing the glycan motif level, based on terminal motifs that were one to three monosaccharides in size. PCA did not cluster according to any discernible pattern, including treatment and cohort (Figure 9).

Likewise, HAC based on a heatmap of these terminal motifs did not demonstrate any clustering based on treatment or cohort (Figure 10). The heatmap also revealed that the relative abundances of fucosylated terminal motifs dwarfed the relative abundances of non-fucosylated motifs. A small cluster of four mice containing two infected and two sham control (the cluster on the right) contained a few percent less fucosylated terminal motifs. A large proportion of the terminal motifs were H antigens, primarily H type 2 (Fuc ($\alpha 1-2$)Gal($\beta 1-4$)GlcNAc). Glycans with at least one H-type antigen constituted 49.7% of the relative abundance. Among the 10 most abundant terminal epitopes was $\alpha 1,4$ GlcNAc, which has been shown to inhibit *H. pylori* growth [91]. Sialylated glycans were rare, with Neu5Ac and Neu5Gc moieties constituting 1.6 % and 0.63 % of the relative abundance respectively, while

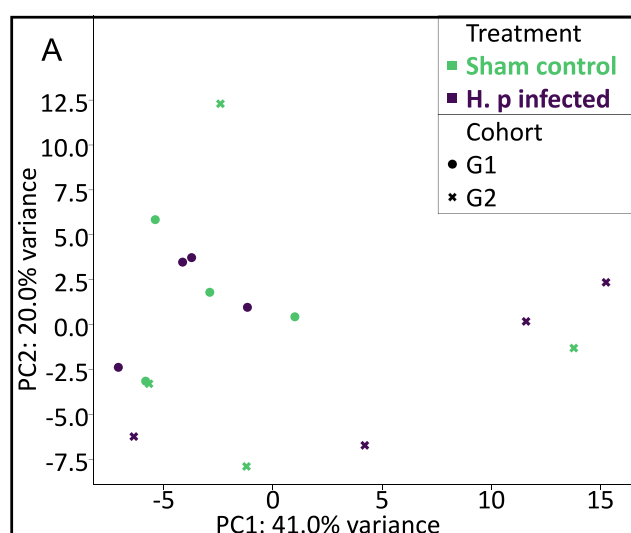


Figure 9. Principal component analysis (PCA) of terminal motif relative abundances using glycowork. Terminal structures of one to three monosaccharides in size were selected as features. 61% of the variance was captured principal components one (PC1) and two (PC2) together. Which cohort the mouse belonged to is denoted with shape (dot or cross), while the treatment that the mouse received is denoted by color.

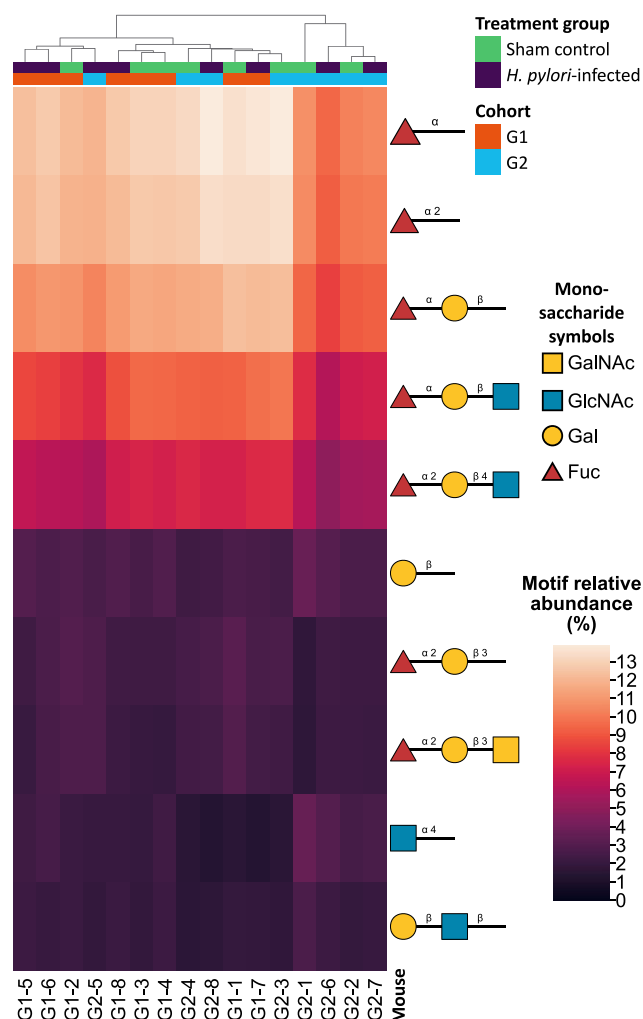


Figure 10. Heatmap of relative abundances of the top 10 most abundant terminal glycan motifs. The complete heatmap with all terminal motifs are shown in Supplemental Figure S4. Terminal motifs of one, two, or three monosaccharides in size were identified using the python package glycowork. The dendrogram at the top of the heatmap shows the hierarchical agglomerative clustering (HAC) which was used to cluster the mice. Cohort and treatment are denoted by color. Glycans are visualized using symbols according to the SNFG guidelines. GalNAc = N-Acetylgalactosamine. GlcNAc = N-Acetylglucosamine. Gal = Galactose. Fuc = Fucose.

sulfated glycans were even more rare at 0.098 % of the relative abundance.

Differential glycomics in glycowork yielded zero terminal motifs that had different relative abundances between sham control and *H. pylori*-infected mice (Supplemental File 5). There were significant differences in relative abundances of terminal motifs between the two cohorts, although they were small with no log2 fold changes (l2fc) above ± 2 and only one l2fc above ± 1 (Supplemental Figure S4 and Supplemental File 6).

Early *H. pylori* infection changes the gene expression of mouse SMCs

The differential gene expression was first assessed with the unsupervised machine learning methods PCA and UMAP, to determine if samples clustered together based on infection status.

PCA and subsequent UMAP dimensionality reduction of principal components 1 through 5 showed that infected and sham control mice clustered separately, but with some overlap (Figure 11(A and C)). Principal components for UMAP dimensionality reduction were chosen based on a plateauing of captured variance in subsequent principal components (Figure 11B). In the PCA, *H. pylori*-infected mice tended to have greater values of PC1 and lower values of PC2 as compared to the sham control mice (Figure 11A). Meanwhile, *H. pylori*-infected mice tended to have lower UMAP 2 values than the sham control mice, while no differences between the two groups was observed for UMAP 1 (Figure 11C).

Differentially-expressed genes between infected and sham control mice were identified using DeSeq2 and IHW. Of the 23,678 genes detected in the tissue (Supplemental File 7), 214 showed significant differential expression (adjusted p -value < 0.05 , Figure 11D, and Supplemental File 8). In *H. pylori*-infected mice, 33 of the significant genes were up-regulated with a log2 fold change greater than 2, while ten were down-regulated with a log2 fold change less than -2 . None of the very abundantly-expressed genes in normal SMCs (Table 1) were differentially expressed.

Gene set enrichment analysis indicated a broad perturbation of genes related to oxidative phosphorylation and infection

The ranked gene set enrichment analysis with SetRank yielded 119 significant gene sets (adjusted p -value < 0.05 , Supplemental File 9). To narrow down the significant gene sets to a more manageable number for analysis we applied a filter of the variable pSetRank (< 0.05) and analysed the resulting network of 8 gene sets in Cytoscape using RCy3 (Figure 12). A hierarchical layout was applied which showed that highly ranked genes in the two gene sets “Oxidative phosphorylation” (from KEGG) and “L13a-mediated translational silencing of Ceruloplasmin expression” (from Reactome) drives at least part of the significance [46] of the other 6 gene sets “non-alcoholic fatty liver disease,” “coronavirus disease – COVID-19,” “cellular responses to stress,” “cellular response to chemical stress,” “signalling by receptor tyrosine kinases,” and “intracellular signaling by second messengers.”

The Ceruloplasmin gene (*Cp*) itself, from the gene set bearing its name, did not show significant differential expression. This gene set corresponds to a Reactome pathway and was created on the basis of a study describing how interferon gamma (IFN- γ) silences *Cp* [92]; therefore, other components of this gene set are likely perturbed without directly suppressing *Cp* expression.

Up-regulated genes were often inducible by interferon type I, ii, or iii, and were mostly co-expressed in a functionally related cluster

By visualizing the interactions from the STRING database of all significant genes, 3 distinct clusters emerged (Figure 13). Most up-regulated genes in the network were present in cluster 1, while the other two clusters consisted mostly of slightly down-regulated genes. A large version of Figure 13 with full annotation of all differentially-expressed genes mapped to the STRING database is available in the supplementary (Supplemental Figure S5).

A large proportion of all significantly up-regulated genes was located in cluster 1 (41 of 90 genes; 46%), particularly among the most strongly up-regulated genes with a log2 fold change > 2 (27 of 33 genes; 82%). All up-regulated genes in cluster 1 were either interferon-stimulated genes (ISGs) inducible by at least one interferon type (I, II, or III) or capable of inducing interferon production themselves (Table 3). This aligned with the observation that one of the two significantly enriched gene sets from the SetRank analysis, “L13a-mediated translational silencing of Ceruloplasmin expression” (Figure 12), is related to IFN- γ (type II interferon) signaling [92].

Genes involved in oxidative phosphorylation are down-regulated in mouse SMCs infected with *H. pylori*

Cluster 2 consisted of 29 down-regulated genes and one up-regulated gene. When investigating these 30 genes, we discovered that they largely are related to OXPHOS (Table 4). OXPHOS was also one of the top two most important gene sets according to the SetRank analysis (Figure 12). Thus, we conclude that *H. pylori* infection leads to a down-regulation of OXPHOS in SMCs in vivo.

***H. pylori* infection down-regulates genes related to protein synthesis which are connected to OXPHOS and ISGs**

The 22 genes in cluster 3 were primarily associated with ribosomes and protein synthesis (Table 5). All genes encoding ribosomal subunits within this cluster also

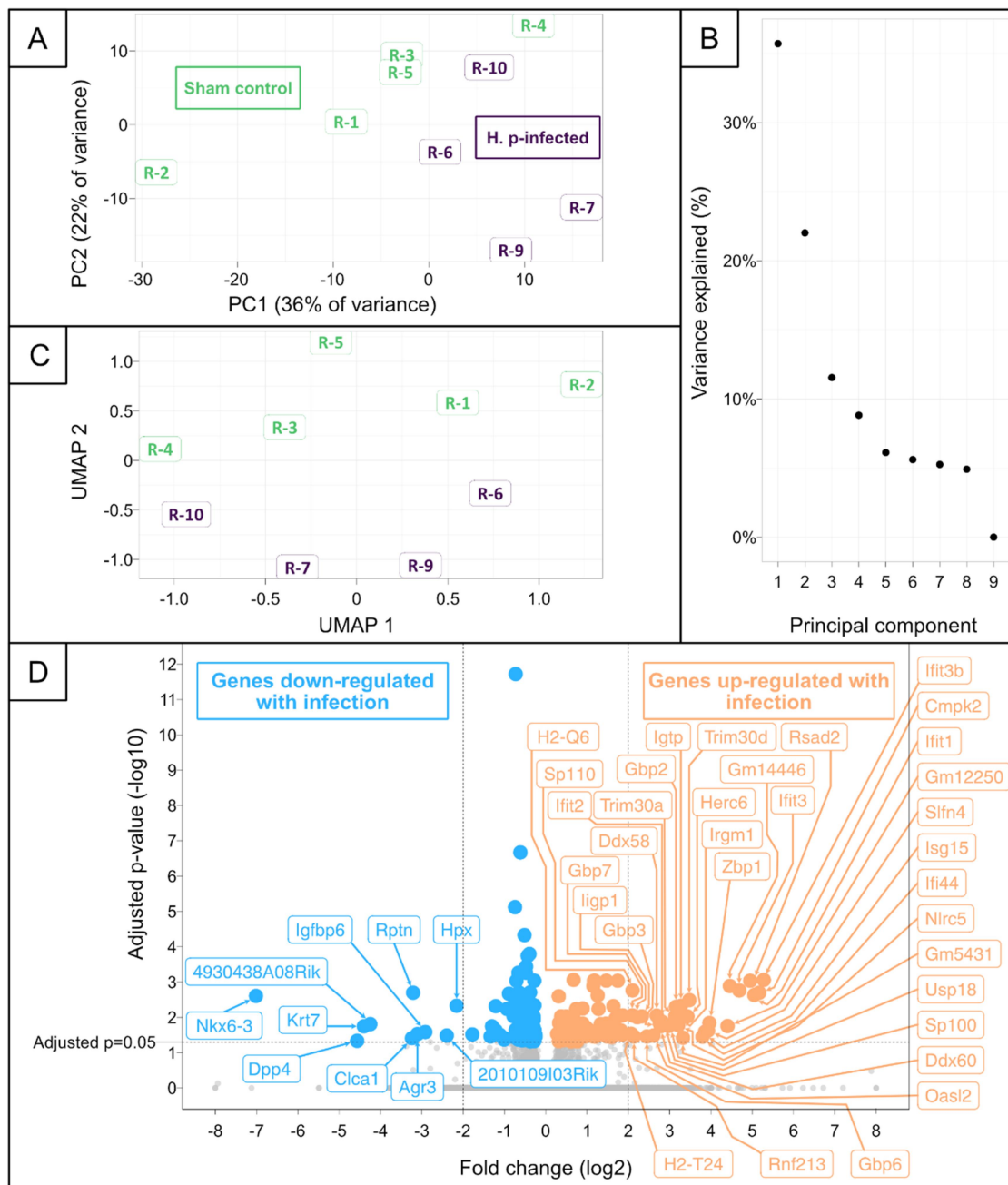


Figure 11. Differential gene expression between infected and sham control mice, as determined with DESeq2 and Independent Hypothesis Weighting (IHW). A: PCA of regularized log transformed gene expression data. Data from *H. pylori* infected mice was colored purple while data from sham control mice was colored green. PC1 and PC2 together captured 58% of the variance. B: Dotplot visualizing the proportion of variance captured by each principal component from the PCA in panel A. The captured variance plateaued after PC5. C: UMAP dimensionality reduction of principal components 1 through 5 from the PCA in panel A. D: Volcano plot of differential gene expression results from DESeq2 after adjusting with IHW. Significant genes (adjusted p -value < 0.05) are above the horizontal dotted line. Genes which are down-regulated with *H. pylori* infection are to the left and are colored blue, while genes which are up-regulated with *H. pylori* infection are to the right and are colored orange. Furthermore, genes with a $\log_2$ Fold change less than -2 or greater than 2 are highlighted with their respective gene symbols in text boxes.

belonged to one of the two most important gene sets identified by the SetRank analysis, namely “L13a-mediated translational silencing of Ceruloplasmin expression” (Figure 12). All but three genes in cluster 3 were down-regulated; among the up-regulated genes,

Zbp1 showed a notably high $\log_2$ fold change of 3.96. Like other up-regulated genes in cluster 1, *Zbp1* is linked to interferon stimulation [137].

By drawing a sub-network of the genes belonging to clusters 1–3 the interconnectivity between these 3

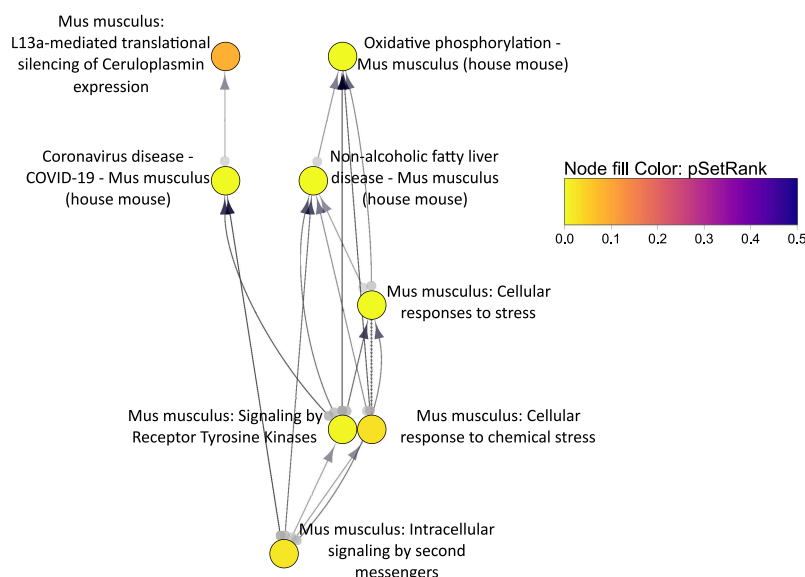


Figure 12. Cytoscape visualization with hierarchical layout of significant gene sets from SetRank network. The R package SetRank output a network of enriched gene sets. Gene sets with $p\text{SetRank} < 0.05$ were selected and analysed in Cytoscape through the R package RCy3. The network style was a modification of the SetRank style provided by the R package. The directed edges indicate that the gene sets overlap and the arrow points towards which of the two overlapping gene sets that contained a greater proportion of higher ranked genes. The hierarchical layout was applied with default settings, which arranged the nodes to avoid edge crossings and consider the directionality of the edges. The topmost nodes of the hierarchical layout were Oxidative phosphorylation and L13a-mediated translational silencing of Ceruloplasmin expression, which had incoming edges but no outgoing edges. The Ceruloplasmin gene itself (Cp) was not differentially-expressed.

clusters was visualized (Figure 14). Here, it was evident that cluster 3 had many connections with both the up-regulated interferon-inducible genes of cluster 1 and the down-regulated genes related to oxidative phosphorylation of cluster 2, while there were no direct interactions between clusters 1 and 2.

Inflammation-related non-coding miRNA *Gm23935* is the most abundant differentially-expressed gene

Highly-expressed genes might be important, despite a small $\log_2$ fold change. To discover such important genes, we investigated the mean TPM of all differentially-expressed genes in the two experimental groups (Figure 15). The non-coding miRNA *Gm23935* had the greatest overall gene expression. In contrast, the $\log_2$ fold change of *Gm23935* was modest at 0.9. *Gm23935* is known to be enriched in NLR family pyrin domain containing 3 (NLRP3) inflammasomes [139], however, *Nlrp3* was not differentially-expressed in our data set. *Gm23935* is also differentially-expressed in hypo-perfused mouse brain cells [140]. Thus, the most highly-expressed, differentially-expressed gene appears associated with inflammation.

***Nkx6-3* had the most extreme differential expression with a $\log_2$ Fold change of -7 , and depletion of this gene is a known contributor to gastric cancer development**

43 significantly differentially-expressed genes had an absolute $\log_2\text{FoldChange}$ ($l2fc$) of $>|2|$ (Table 6). Out of the 33 genes in Table 6 with a $l2fc > 2$ (up-regulated in *H. pylori*-infected mice), 32 genes were known to be inducible by interferons. Among the down-regulated genes with $l2fc < -2$, one gene was known to be essential for mucin expansion (*Cla1*) [141]. Finally, *Nkx6-3*, coding for a homeobox protein related to gastric cell differentiation and tumor suppression, had the most extreme $\log_2$ fold change (-7). Notably, the depletion of this protein has been shown to contribute to gastric cancer development [146,147].

***Lgals3bp* (cluster 1) and *Rpl19* (cluster 2), but not *Cox6c* (cluster 3) gene expression was validated at the protein level through immunofluorescence (IF)**

To validate the overall RNA-level results, we selected one protein-coding gene from each cluster 1–3. Protein choice was based on high gene expression and the

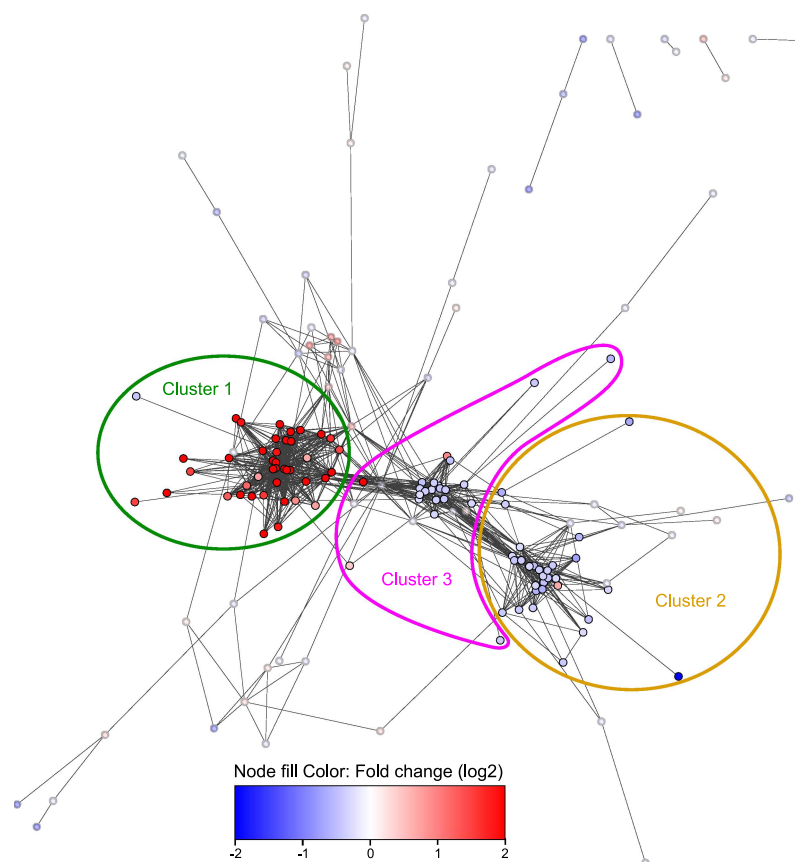


Figure 13. Cytoscape string network of significantly differentially-expressed genes. R packages were used to create a cytoscape network based on full string interactions with default confidence score of 400. Significant genes without a string interaction to another significant gene were not included in the network. Node fill was colored by log2 Fold changes on a continuous scale from blue (down-regulated) to red (up-regulated) with *H. pylori* infection. Minimum and maximum values in the color scale was set to -2 and 2 respectively to facilitate visualization, where the more extreme Fold-changes would otherwise distort the scale. There were 3 visually obvious clusters, and they were picked out with a systematic approach using MCL clustering with the default inflation parameter of 2.5 and selecting the top 3 clusters. To highlight these 3 clusters, the opacity of all nodes which did not belong to them was reduced. The colored loops indicate which cluster the full opacity nodes within belonged to. The supplemental file 1 cytoscape file contains further details, such as which genes the nodes consist of.

availability of adequately verified antibodies on mouse tissue. LGALS3BP (cluster 1), also known as Tumor-associated antigen 90K [52], is a type I and II interferon induced gene [114]. In humans, LGALS3BP has been implicated in the immune defense against cancer and pathogens [148]. RPL19 (cluster 2) is a component of the large ribosomal subunit [52]. The mitochondrial protein COX6C (cluster 3) is a cytochrome c oxidase (complex IV) subunit [52]. TNF α stimulation and *C. rodentium* infection decreases COX6C expression in the mouse colon, and this reduction is reversed with IL-4 and IL-6 treatment [149].

Mean signal intensity of LGALS3BP (Figure 16A, D, and G) and RPL19 (Figure 16B, E, and H) differed significantly between *H. pylori*-infected and sham control mice, while COX6C (Figure 16C, F, and I) did not.

Similarly, there was a significant positive correlation between LGALS3BP signal intensity and *Lgals3bp* gene expression, but not between COX6C signal intensity and *Cox6c* gene expression. The raw p value for the positive correlation (Spearman's $\rho = 0.72$) between RPL19 signal intensity and *Rpla19* gene expression was 0.037, but when adjusted for multiple comparisons it was bordering to significant (adjusted $p = 0.055$). IF Images and associated code are available in Supplemental File 11.

Discussion

Here, we demonstrated that the collection of the SMCs using LCM yielded sufficient RNA quantity and quality for RNA-Seq in 9/10 mice, without the need for

Table 3. Genes which constituted cluster 1 from Figure 13. 41 of the 42 genes in the cluster were up-regulated. All 41 up-regulated genes in cluster 1 were either inducible by at least one type of interferon (type I, type II, or type III), or induced interferon production itself.

Gene	log2 Fold Change	Type I interferon induced gene	Type II interferon induced gene	Type III interferon induced gene	Effect on type I interferon production
Bst2	1,57	Yes [93]	Yes [94]		
Cmpk2	5,17	Yes [95]	Yes [96]	Yes [97]	
Ddx58	2,69				Stimulates [98,99]
Ddx60	2,79	Yes [100]			
Eif2ak2	0,76	Yes [98]			
Gbp2	3,15	Yes [97]	Yes [101,102]	Yes [97]	
Gbp3	2,76	Yes [97]	Yes [101,102]	Yes [97]	
Gbp5	1,86		Yes [101,103]		
Gbp6	2,45	Yes [97]	Yes [101,102]	Yes [97]	
Gm14446	4,46	Yes [104]			
Helz2	0,89	Yes [105]			
Herc6	3,43	Yes [106]			
Ifi27	0,67	Yes [107]			
Ifi44	3,79	Yes [108]		Yes [97]	
Ifit1	5,05	Yes [99]		Yes [97]	
Ifit2	2,82	Yes [99]			
Ifit3	4,69	Yes [99]	Yes [109]	Yes [97]	
Ifit3b	5,28	Yes [104]			
Igtp	3,27		Yes [110]		
Iigp1	2,33	Yes [97]	Yes [111]	Yes [97]	
Irgm1	3,23	Yes [97]	Yes [112,113]	Yes [97]	
Isg15	3,86	Yes [98]		Yes [97]	
Lgals3bp	0,74	Yes [114]	Yes [114]		
Nlrc5	3,21		Yes [115]		Inhibits [115]
Oasl2	2,61	Yes [97]	Yes [116]	Yes [97]	
Parp14	1,86	Yes [97]	Yes [117]	Yes [97]	
Phf11d	1,82	Yes [97]		Yes [97]	
Psmb8	1,64	Yes [118]			
Psmb9	1,46	Yes [119]		Yes [97]	
Rnf213	2,14	Yes [120]	Yes [120]		
Rsad2	4,95	Yes [99]	Yes [108]	Yes [97]	
Slfn4	3,93	Yes [121]			
Sp100	2,82	Yes [99]	Yes [122]	Yes [97]	
Sp110	2,10	Yes [97]		Yes [97]	
Tlr3	1,20		Yes [123]		Stimulates [124]
Tnfsf10	1,50	Yes [97]	Yes [125]	Yes [97]	
Tpt1	-0,44				
Trim14	1,48	Yes [109]	Yes [109]		
Trim30a	3,14	Yes [97]		Yes [97]	
Trim30d	3,48	Yes [97]		Yes [97]	
Usp18	2,98	Yes [97]		Yes [97]	
Znfx1	1,28	Yes [126]			

amplification prior to library preparation. Genes coding for secreted mucus proteins, mitochondrial function, and cellular structure were hallmarks of the normal SMC. During early *H. pylori* infection, 214 genes were differentially-expressed, while 119 Reactome and KEGG gene-sets were perturbed. Network analyses of significant genes and gene-sets showed that *H. pylori* infection causes an up-regulation of interferon-stimulated genes (ISGs), a down-regulation of both oxidative phosphorylation (OXPHOS) and protein synthesis-related genes, and the initiation of an early pre-cancerous cascade with a depletion of *Nkx6-3*. Furthermore, glycan analysis using mass spectrometry expanded the known mouse gastric glycome and revealed two main glycosylation pathways characterized by large and heavily fucosylated glycans, and this was corroborated by glycosyltransferase expression. Finally, the mouse gastric glycome was remarkably unchanged by early *H. pylori* infection.

The normal mouse SMC of the gastric corpus contains abundant RNA coding for tumor suppressors, secreted proteins, components of cellular respiration, and structural proteins

The gene expression of normal (sham control) SMCs was dominated by a high expression of well-known marker genes *Muc5Ac*, *Tff1*, *Gkn1*, *Gkn2*, *PscA*, and *Lgals2* [17,54]. The proteins coded by these genes can all be secreted into mucus. Of these, *Tff1*, *Gkn1*, and *Gkn2* are also known as tumor suppressor genes [62,64,150]. Another highly-expressed, tumor-suppressing gene was *Mal* [69].

Although present in the data, a notable absence among the highly-expressed genes was *Il33*, which was a canonical marker gene in both PanglaoDB [54] and Bockerstett 2020 [17]. In Bockerstett 2020, which used Balb mice, there were two different clusters annotated as SMCs, cluster 0 and cluster 2, and healthy cells were part of both clusters. *Il33* defined cluster 0 together with *Tff1* and *Muc5ac*, but not

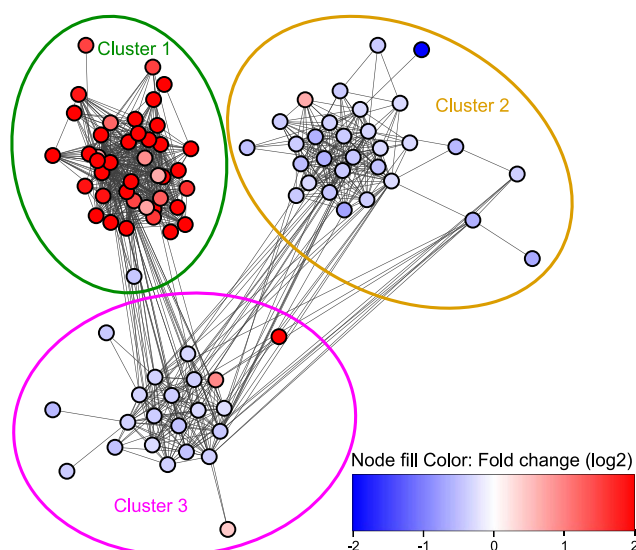


Figure 14. The main clusters 1–3 were highly interconnected. There was a high degree of string interaction interconnectivity between cluster 1 and 3, and between cluster 2 and 3. There were no direct interactions between cluster 1 and 2. The cytoscape file contains further details, such as which genes the nodes consist of (supplemental file 1).

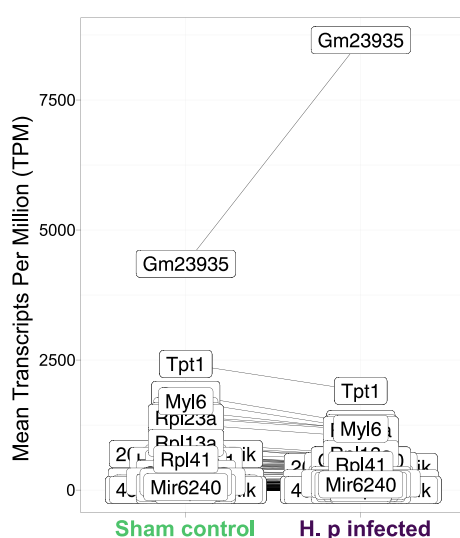


Figure 15. Significantly differentially-expressed genes mean TPM (Transcripts per Million) of infected vs. sham control mice. The most highly-expressed gene with a significant differential gene expression was the non-coding miRNA Gm23935. Gm23935 was up-regulated in *H. pylori* infected mice and had a modest log2 Fold change of 0.9.

cluster 2, which was only defined by *Tff1* and *Muc5ac*. In humans, according to the protein atlas, the expression of *Il33* is more concentrated in the isthmus, where the isthmus stem cells (IsthSCs) reside [151], then decreasing below detection limit toward the surface [152]. Thus, *Il33* might be associated with less differentiated SMCs, while fully

differentiated SMCs reduce or stop expressing *Il33*. Similarly, in C57BL/6 mice, *Il33* was expressed in the isthmus SMCs as well as the mucus neck cells, but not higher up in the pits or on the surface [153].

Il33 was not differentially expressed in the present study, however, in Balb mice infected with an 11 times higher dose of *H. pylori* strain 26,695, there was an increase of IL-33 and its receptor ST-2 in SMCs and other cells 2 weeks after infection [154]. The authors did not quantify the inflammation but described it as severe, and the H&E micrographs provided appeared more inflamed than in the present study. If *Il33* decreases with SMC differentiation, then one potential mechanism behind the differential expression could be reduced life-span of the SMCs. *In vitro* results from the same article showed that the IL-33 could be both unchanged, up-regulated, and down-regulated depending on infection time and dose with *H. pylori*, which might explain why *Il33* and *Il1rl1* (which codes for ST-2) was neither highly nor differentially expressed in the present study's data. Different *H. pylori* and mouse strains might also explain differences in results.

In humans, FCGBP is expressed in SMCs and forms a heteromer with TFF1 [62]. While *Fcgbp* was detected in the present mouse data, it was of very low abundance, indicating that another protein might fill the same role in mice. Another gene which differs between normal mouse and human gene expression was *Dpcr1* (Also known as *Muc13*), which was highly expressed in the present study's mouse SMCs, but not expressed in human gastric SMCs [155]. The function of *Dpcr1/Muc13* in SMCs is unknown, but is annotated to be present in the cytoplasm and serve as a transmembrane protein. The protein is likely glycosylated and could play a part in NF- κ B signaling and cell growth [52]. According to the STRING database, *Dpcr1/Muc13* interacts with *Gkn2*, which was another highly-expressed gene [48].

Other highly-expressed genes with unclear function were the long non-coding RNA (lincRNA) *Yam1/Gm26924* and *Gm26917* [63], as well as the small nuclear RNA (snRNA) *Rn7sk*. *Rn7sk* is known to be important for gene regulation of highly-expressed gene pairs which are co-expressed through a shared bi-directional promoter [156]. We postulate that *Gkn1/Gkn2* could be such a gene pair regulated by *Rn7sk*, since they are situated right next to each other in the genomes of both humans and mice, and they are also transcribed in opposite directions.

In the present study, mitochondrial genes were also highly expressed. *mt-Rnr1* and *mt-Rnr2* are mitochondriosome rRNA [63], and the only gene with a higher TPM than these two was *Tff1*. The OXPHOS genes *mt-Atp6* and *mt-Co1* were also highly expressed.

Table 4. Genes which formed cluster 2 (Figure 13). All genes except mt-Nd6 were down-regulated in this cluster. With few exceptions, these genes were either components of oxidative phosphorylation or directly affected this process in some other way.

Gene	log2FoldChange	Component of OXPHOS complex	Other direct effect on OXPHOS
2010107E04Rik	-0.73	Complex V	
2310036O22Rik	-0.40		
4930438A08Rik	-4.25	Complex V	
Atp5e	-0.37	Complex V	
Atp5h	-0.44	Complex V	
Atp5j	-0.48	Complex V	
Atp5j2	-0.30	Complex V	
Atp5o	-0.28	Complex V	
Atp6v1d	-0.30		
Chchd10	-0.42		Required for complex I assembly [127]
Chchd2	-0.41		Required for complex I assembly [127]
Cox4i1	-0.32	Complex IV	
Cox6a1	-0.40	Complex IV	
Cox6c	-0.61	Complex IV	
Cox7b	-0.40	Complex IV	
Gm3244	-0.61	Complex I	
Minos1	-0.29		Important for Cristae formation [128]
Mrpl43	-0.39		Component of mitochondrial ribosomes [129]
Mrps6	-0.61		Component of mitochondrial ribosomes [129]
mt-Nd6	0.66	Complex I	
Ndufa13	-0.51	Complex I	
Ndufa6	-0.45	Complex I	
Ndufb10	-0.29	Complex I	
Ndufs4	-0.31	Complex I	
Ndufs6	-0.41	Complex I	
Pet100	-0.49		Important for the assembly of oxidative phosphorylation complexes [130]
Ppif	-0.55		Modulates complex V [131]
Slc5a3	-0.66		
Tomm7	-0.42		Component of mitochondrial protein import TOM complex [132]
Uqcrcq	-0.36	Complex III	

Another category of highly-expressed genes were genes with a structural function, namely *Krt8*, *Krt19*, *Actg1*, and *Mal* [52]. Of these, the first three are intermediate filament genes. KRT8 and KRT19 likely forms a keratin pair in these SMCs, since KRT8 is a type II and KRT19 is a type I keratin, and they have similar expression levels. Co-expression of KRT8/KRT19 is previously known in other epithelial cells [157].

The mucus produced by SMCs is glycosylated with large and heavily fucosylated glycans, and this glycosylation does not change with early Helicobacter pylori infection in mice

While the gastric mucin glycosylation is characterized by extreme heterogeneity in humans [26], there was low inter-individual variation in the present mouse data, likely due to that the mice were an inbred strain. Furthermore, while the mucin glycans differed between chronically *H. pylori*-infected and non-infected humans, and early *H. pylori* infection has been shown to increase the expression of sialylated Lewis antigens in the rhesus monkey [158], the glycans did not change as a result of early infection in the mouse. This stable glycosylation was comprised of remarkably large and heavily fucosylated glycans, and such FUT2 dependent fucosylation binds to *H. pylori* [24]. A large proportion

of the terminal motifs consisted of H antigens, and in particular H type 2 antigens.

A limitation with the glycome analysis was that it was performed on corpus tissue rather than LCM-isolated SMCs. As a result, approximately 95% of detected glycans originated from MUC5AC-expressing SMCs, with a minor contribution from MUC6-expressing mucus neck cells [159]. This approach was chosen due to that in-depth glycome profiling requires ~100 µg of glycoprotein which is not feasible with LCM. O-glycomes obtained from tissues and isolated mucins are essentially identical, indicating that the glycomes reported here mainly reflects mucin glycosylation [55].

Another limitation with the present glycome study was that it was conducted on two cohorts different from the cohort which the RNA-Seq data was based on. Since 17 terminal motifs were differentially-expressed between the glycan cohorts, the glycosylation could have been different for the RNA-Seq cohort. However, the PCA and clustermap revealed that this difference was driven by a cluster of four mice in cohort G2, which had a slightly lower fucosylation. We could not find any explanation for this divergent cluster in information known to us about the mice (age, weight, experimental notes). Within this cluster of four divergent mice there were both *H. pylori*- and sham control mice. Of importance, the other four mice in

Table 5. Genes from cluster 3 (Figure 13). All genes in this cluster, except *Srp54a*, *Xrn2*, and *Zbp1*, were down-regulated. Most genes encoded ribosomal subunits or influenced protein synthesis through other mechanisms. All ribosomal genes in cluster 3 also belonged to the gene set “L13a-mediated translational silencing of Ceruloplasmin expression,” which was one of the two most important perturbed gene sets according to the SetRank analysis (Figure 12).

Gene symbol	log2FoldChange	Component of ribosomes	Part of gene set “L13a-mediated translational silencing of Ceruloplasmin expression”	Other relationships to protein synthesis
<i>Abtb1</i>	−0.56			Probable elongation factor protein [52]
<i>Btf3</i>	−0.32			Translational regulation, potentially protection against intracellular bacteria [133]
<i>Eef1b2</i>	−0.35			Probable elongation factor protein [52]
<i>Fkbp2</i>	−0.40			Probable PPIase which facilitates protein folding [52]
<i>Pdcd5</i>	−0.31			Induces apoptosis [134]
<i>Rpl13a</i>	−0.47	60S ribosomal subunit protein [52]	Yes	
<i>Rpl18</i>	−0.34	60S ribosomal subunit protein [52]	Yes	
<i>Rpl19</i>	−0.48	60S ribosomal subunit protein [52]	Yes	
<i>Rpl23a</i>	−0.35	60S ribosomal subunit protein [52]	Yes	
<i>Rpl29</i>	−0.37	60S ribosomal subunit protein [52]	Yes	
<i>Rpl34</i>	−0.39	60S ribosomal subunit protein [52]	Yes	
<i>Rplp0</i>	−0.40	60S ribosomal subunit protein [52]	Yes	
<i>Rplp2</i>	−0.38	60S ribosomal subunit protein [52]	Yes	
<i>Rps11</i>	−0.42	40S ribosomal subunit protein [52]	Yes	
<i>Rps26</i>	−0.28	40S ribosomal subunit protein [52]	Yes	
<i>Rps3a1</i>	−0.26	40S ribosomal subunit protein [52]	Yes	
<i>Rps9</i>	−0.38	40S ribosomal subunit protein [52]	Yes	
<i>Srp54a</i>	0.93			Signal recognition particle (SRP) complex subunit [52]
<i>Tubb4b</i>	−0.42			Might affect ribosomes through its association with RNA granules [135]
<i>Xrn2</i>	0.42			Exoribonuclease which may stop mRNA transcription [136]
<i>Zbp1</i>	3.96			Mediates type I interferon responses from DNA sensing, thus suppresses global protein translation [137]
<i>Zfp622</i>	−0.48			Helps assemble ribosomal 60S subunit [138]

that cohort had an indistinguishable glycosylation to that of cohort G1. Therefore, it is unlikely that the overall pattern, with unchanging glycosylation with *H. pylori* infection and the glycans being overall large and heavily fucosylated, would have been different in the RNA-Seq cohort. Finally, that the most abundant glycosyltransferases could neatly explain the main glycosylation pathways also lends credence to the notion that the overall glycosylation was the same between all three cohorts.

Holmén-Larsson et al. 2013 characterized gastric glycans in C57BL/6 mice using a similar method as the present study [23]. In the present study, more structures were discovered (100 vs 31), and they were also both larger and more fucosylated in comparison to Holmén-Larsson et al. 2013 [23]. In line with our observation that *Fut2* was the 4th most abundantly-expressed glycosyltransferase with activity on mucin O-glycans, Magalhães et al. 2009 described 40 glycan structures in the gastric mucosa glycans of C57BL/6 mice and determined that the glycans were overall large and heavily fucosylated, whereas there was almost no fucosylation in isogenic *Fut2*-null mice [24]. Furthermore, the two most abundant glycans in the

present study (Fuc(α1-2)Gal(β1-3)[Fuc(α1-2)Gal(β1-4)GlcNAc(β1-6)]GalNAcol and Fuc(α1-2)Gal(β1-3/4)GlcNAc(β1-3)[Fuc(α1-2)Gal(β1-4)GlcNAc(β1-6)]Gal(β1-3)[Fuc(α1-2)Gal(β1-3/4)GlcNAc(β1-6)]GalNAcol) correspond to abundant peaks at m/z 1331 and 2404 respectively in Magalhães et. al. 2009 Figure 4 [24]. The bimodal size distribution of the glycans in the present study was likely the result of B3GNT3 competing with FUT2, where core 1 and core 2 Gal(β1-3)GalNAcol were either extended with GlcNAc(β1-3) through B3GNT3 activity, or fucosylated by FUT2, thus locking the glycans into one of the two mutually exclusive main glycosylation pathways leading to either Fuc(α1-2)Gal(β1-3)[Fuc(α1-2)Gal(β1-4)GlcNAc(β1-6)]GalNAcol or Fuc(α1-2)Gal(β1-3/4)GlcNAc(β1-3)[Fuc(α1-2)Gal(β1-4)GlcNAc(β1-6)]Gal(β1-3)[Fuc(α1-2)Gal(β1-3/4)GlcNAc(β1-6)]GalNAcol.

The *Fut2*-dependent fucosylation was also shown to be important for BabA-mediated adhesion of *H. pylori* [24], and although we found no clear evidence in the published literature that glycan size or fucosylation status affect *H. pylori*-driven carcinogenesis, the BabA adhesin is connected to gastric cancer [160]. Sialylated glycans are also important for *H. pylori* binding to

Table 6. Top significantly differentially-expressed genes based on analysis with Deseq2 and IHW. The table shows which of the differentially-expressed genes (adjusted p -value < 0.05) that had a log2FoldChange > 2 or < -2. The up-regulated genes in the table were genes whose overall expression level could be associated with inflammation.

Gene symbol	log2FoldChange	Relevance to previous analyses
<i>Nkx6-3</i>	-7.02	Not in cluster 1–3.
<i>Dpp4</i>	-4.57	Not in cluster 1–3.
<i>Krt7</i>	-4.40	Not in cluster 1–3, but is a type II keratin [52] just like the highly-expressed gene <i>Krt8</i> (Table 1).
<i>4930438A08Rik</i>	-4.25	Codes for an OXPHOS complex V subunit cluster 2 (Table 4).
<i>Clca1</i>	-3.24	Not in cluster 1–3, but is essential for mucus expansion [141].
<i>Rptn</i>	-3.21	Not in cluster 1–3
<i>Agr3</i>	-3.11	Not in cluster 1–3
<i>Igfbp6</i>	-2.92	Not in cluster 1–3, but is a secreted protein [52].
<i>2010109I03Rik</i>	-2.40	Not in cluster 1–3. Also known as <i>Ly6m</i> [52].
<i>Hpx</i>	-2.16	Not in cluster 1–3
<i>H2-T24</i>	2.07	Not in cluster 1–3, but codes for a histocompatibility protein [52] which interacts with <i>Usp18</i> in cluster 1 (Figure 13). Inducible by interferon type I [142].
<i>Sp110</i>	2.10	In cluster 1, induced by interferon type 1/3 (Table 3).
<i>H2-Q6</i>	2.10	Not in cluster 1–3, but codes for a histocompatibility protein [52] which interacts indirectly with cluster 1 through <i>H2-T24</i> , <i>H2-Q4</i> , <i>Tap2</i> , and <i>B2m</i> (Figure 13). Inducible by interferon type I [143] and type II [144].
<i>Rnf123</i>	2.14	In cluster 1, induced by interferon type 1/2 (Table 3).
<i>Gbp7</i>	2.22	Not in cluster 1–3, but codes for an interferon-inducible GTPase that protects against bacterial infections [101].
<i>Ilgp1</i>	2.33	In cluster 1, induced by interferon type 1/2/3 (Table 3). Also known as <i>Irga6</i> [52].
<i>Gbp6</i>	2.45	In cluster 1, induced by interferon type 1/2/3 (Table 3).
<i>Oas2</i>	2.61	In cluster 1, induced by interferon type 1/2/3 (Table 3).
<i>Ddx58</i>	2.69	In cluster 1, stimulates type I interferon production (Table 3). Also known as <i>Rigi</i> [52].
<i>Gbp3</i>	2.76	In cluster 1, induced by interferon type 1/2/3 (Table 3).
<i>Ddx60</i>	2.79	In cluster 1, induced by interferon type 1 (Table 3).
<i>Ifit2</i>	2.82	In cluster 1, induced by interferon type 1 (Table 3).
<i>Sp100</i>	2.82	In cluster 1, induced by interferon type 1/2/3 (Table 3).
<i>Usp18</i>	2.98	In cluster 1, induced by interferon type 1/3 (Table 3).
<i>Trim30a</i>	3.14	In cluster 1, induced by interferon type 1/3 (Table 3).
<i>Gbp2</i>	3.15	In cluster 1, induced by interferon type 1/2/3 (Table 3).
<i>Nlr5</i>	3.21	In cluster 1, induced by interferon type 2, and inhibits production of interferon type I (Table 3).
<i>Irgm1</i>	3.23	In cluster 1, induced by interferon type 1/2/3 (Table 3).
<i>Igtp</i>	3.27	In cluster 1, induced by interferon type 2 (Table 3).
<i>Gm5431</i>	3.32	Not in cluster 1–3, but has predicted coding for a GTP-binding protein which is predicted to respond to interferon-beta stimulation [52].
<i>Herc6</i>	3.43	In cluster 1, induced by interferon type 1 (Table 3).
<i>Trim30d</i>	3.48	In cluster 1, induced by interferon type 1/3 (Table 3).
<i>Ifi44</i>	3.79	In cluster 1, induced by interferon type 1/3 (Table 3).
<i>Isg15</i>	3.86	In cluster 1, induced by interferon type 1/3 (Table 3).
<i>Sifn4</i>	3.93	In cluster 1, induced by interferon type 1 (Table 3).
<i>Zbp1</i>	3.96	In cluster 3, induced by IFN type I stimulation (Table 5).
<i>Gm12250</i>	4.40	Not in cluster 1–3, but is translated to <i>IRB10</i> which is induced by type 1 interferon [145].
<i>Gm14446</i>	4.46	In cluster 1, induced by interferon type 1 (Table 3).
<i>Ifit3</i>	4.69	In cluster 1, induced by interferon type 1/2/3 (Table 3).
<i>Rsad2</i>	4.95	In cluster 1, induced by interferon type 1/2/3 (Table 3).
<i>Ifit1</i>	5.05	In cluster 1, induced by interferon type 1/2 (Table 3).
<i>Cmpk2</i>	5.17	In cluster 1, induced by interferon type 1/2/3 (Table 3).
<i>Ifit3b</i>	5.28	In cluster 1, induced by interferon type 1 (Table 3).

mucins via the sialic acid binding adhesin (SabA) [161–163]. Additional adhesins have been identified in *H. pylori* [164], however, although some of these adhesins bind to salivary mucins, there is no observed distinct binding to human gastric mucins with these [161,163,165]. *H. pylori* strains expressing BabA and SabA bind to human mucins in a manner largely driven by mucin fucosylation, and this binding correlates with the relative abundance of α 1,2-fucosylated structures on mucins from both non-infected and *H. pylori*-infected individuals [166]. It is difficult to establish how the glycan size affects binding in the complex glycan environment of mucus. On the one hand there are specific requirements/minimal epitopes determining specificity for the BabA adhesin [167], and mucin glycan size has been correlated to pathogen binding in species with relatively small mucin glycans [168]. By contrast, both

human and mouse mucin glycans are large and complex and the high density of fucosylated glycans in a solution may confer both opportunities for binding that does not occur to a large extent when sparsely-located epitopes on a rigid surface are presented. Multivalency within the dense mucin-glycan environment likely also enables otherwise weak interactions to become biologically significant. In the present study with short-term infected mice, sialylated and sulfated glycans had very low abundance. In long-term infection in rhesus monkeys and humans there are increased levels of sialylated Lewis antigens as demonstrated with immunohistochemistry [158,169,170]. Meanwhile, sulfated glycans are common in gastric mucins from patients with intestinal metaplasia and gastric tumors, as shown by mass spectrometry [171,172]. There is no clear evidence in the literature on how long-term *H. pylori* infection

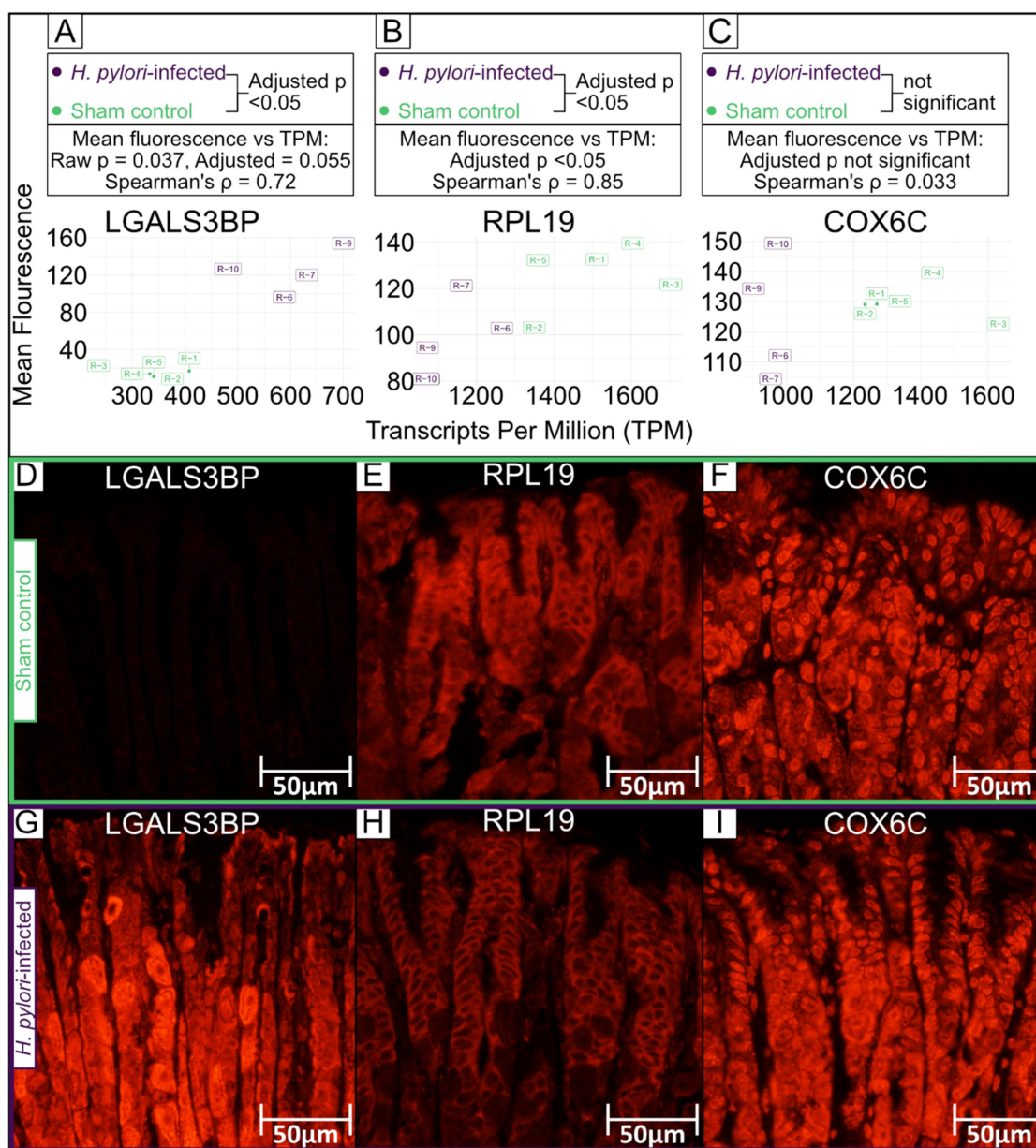


Figure 16. Immunofluorescence protein-level validation of protein coding genes belonging to RNA-Seq clusters 1–3. A–C: Mean fluorescence of SMCs correlated to the gene expression (TPM) of the gene coding for the protein. D–I: Representative IF images (40x magnification) from sham control (D–F) and *H. pylori*-infected mice (G–I). Cluster 1 (a, d, G) antibody recognizing RPL19. Cluster 2 (b, e, H) antibody recognizing LGALS3BP. Cluster 3 (C, F, I) antibody recognizing COX6C.

affects the large fucosylated glycans that dominate the glycome in the present study, but increases in sialylated and sulfated structures implies a corresponding reduction in the relative abundance of other structures.

Functional network analysis indicates that interferon stimulation might cause a reduction of oxidative phosphorylation in SMCs

The overwhelming majority of the most up-regulated genes were inducible by either type I, II, or III-interferon. These ISGs formed a functionally related cluster according

to STRING database annotation and cluster analysis. This cluster was highly interconnected with another cluster of down-regulated genes related to protein synthesis, which in turn was functionally connected to a third cluster of down-regulated genes related to OXPHOS. This pattern suggests a chain of events where interferon stimulation cause a reduction of protein translation, resulting in reduced OXPHOS activity. Furthermore, the absence of any clear gene expression pattern related to cell death indicates that this OXPHOS inhibition is not driven by apoptosis or another form of cell death. Further experiments are necessary to investigate this process.

Similar findings to this proposed chain of events have been shown experimentally in another model system, a gastrointestinal mouse infection model with *Citrobacter rodentium*. Both infection with *C. rodentium* and IFN- γ stimulation caused a reduction of OXPHOS activity in the large intestine [149]. Meanwhile, the *H. pylori* infection itself leads to OXPHOS inhibition in systems independent of exogenous interferon stimulation [173]. In line with this, the *H. pylori* toxin VacA has *in vitro* been shown cause a reduction of OXPHOS activity, which was not due to cell death [174]. Essentially all *H. pylori* strains express VacA, however VacA is considered weak in the SS1 strain [175].

The differential gene expression profile provides a snapshot of both the early *H. pylori* infection and early gastric precancerous cascade in mice

H. pylori infection causes gastritis, and when chronic, this inflammation can continue to a pre-cancerous cascade which ends in gastric adenocarcinoma [176]. Unsurprisingly, much of the differential gene expression profile from the experimental early *H. pylori* infection described in this article was related to inflammation. Apart from the up-regulated genes being dominated by ISGs as discussed above, this was also evident in the SetRank gene set enrichment results where one of the top two gene sets was related to IFN- γ stimulation, the other being related to OXPHOS. When assessing gene set enrichment analysis results, it is important to keep in mind that the only gene sets that can be highlighted by the analysis are the gene sets which the analysis was built upon [46]. One of the few curated *H. pylori* infection gene sets in existence is the KEGG pathway “hsa05120 Epithelial cell signalling in *Helicobacter pylori* infection - *Homo sapiens*,” but there is no mouse variant [42]. While this human KEGG gene-set involves the mitochondria, it is not related to OXPHOS down-regulation itself, but rather how the virulence factor VacA targets mitochondria and induces apoptosis. The gene-set further describes how VacA induces vacuolation with the help of V-ATPase, of which one of the sub-units (ATP6V1D) was differentially expressed in our data. Much of the hsa05120 gene-set pertains to the effects of CagA on inflammation, cell morphology, proliferation, motility, and apoptosis. However, the SS1 *H. pylori* strain does not have a functional type IV secretion system (T4SS) and thus cannot inject cells with CagA [177]. The dataset we have formed can form the genesis for new gene sets, thus improving future gene set enrichment analysis.

A notable absence of differentially-expressed genes in the early *H. pylori* infection were the most highly-expressed genes, which can be seen as function-defining SMC genes. Taken together with the functional network analysis results, this suggests the possibility that the SMCs performs the same functions during early *H. pylori* infection, but at a reduced rate due to OXPHOS reduction.

When examining individual genes, the inflammation-related gene *Gm23935* [139,140] was the most abundant differentially-expressed gene. Also noteworthy was *Nkx6-3*, which showed a stark depletion in *H. pylori*-infected mice and had the most extreme log2 fold change (less than -7). *Nkx6-3* is associated with the precancerous cascade, as a loss of the encoded protein, Homeobox protein Nkx-6.3, contributes to the development of gastric adenocarcinoma [146,147]. NKX6-3 suppresses *NF-kB* and *Dnmt1* [147], which in turn seems to induce *Hace1* expression and Reactive Oxygen Species (ROS) production [147]. However, in our data, there was no significant difference in gene expression for *Hace1*, *Nf-kB*, or *Dnmt1*. This may be explained by the short infection duration in the present study or suggest an alternative mechanism operating in mouse SMCs *in vivo*.

Even though mouse models are common for both gastric cancer and *H. pylori* infection, they do not perfectly recapitulate human pathogenesis. The major shortcoming of the various mouse gastric cancer models is the poor representation of the later stages of cancer development, but they do perform better for studying the earlier stages of the pre-cancerous cascade [178]. *H. pylori* has a host tropism toward humans and do neither readily infect mice, nor create the same level of gastritis. Mouse-adapted strains such as SS1 used in the experiment described here, are often used to better and successfully recapitulate disease phenotypes such as loss of parietal cells [178]. CagA is an important virulence factor for human *H. pylori* infection pathogenesis; however, the previously described SS1 strain does not have a functional T4SS and therefore cannot inject CagA. Furthermore, *H. pylori* strains tend to lose T4SS function when chronically infecting mice [178]. The tendency of *H. pylori* to drop CagA function when adapting to mouse hosts could indicate that CagA is not important for *H. pylori* when infecting mice. Another notable difference between the human and mouse *H. pylori* pathogenesis is the differing microbiota which can at least partially be explained by the presence of a squamous epithelia forestomach and coprophagic behavior in mice, as well as a higher pH of 3–4 rather than pH around 1 in humans [179].

RNA-Seq of cells collected through LCM is a viable alternative to both scRNA-Seq and in situ spatial transcriptomics to study specific gastrointestinal epithelial cells in vivo

The results from the present work shared some similarity to a microarray-based LCM study of *H. pylori*-infected mice [11]. Similarly, both IFN- α and IFN- γ inducible genes and other inflammation-related genes were up-regulated during *H. pylori* infection. Relating these findings to humans, a mechanism that could contribute to IFN- α stimulation is TLR8 signaling in monocytes in response to *H. pylori* infection [180]. Another similarity in the present study to the microarray-based LCM study was the high purity of cellular-type. Dissimilarly, OXPHOS-related genes did not seem to be greatly perturbed in SMCs when subjected to *H. pylori* infection, although the specific genes included in the microarray are unavailable since the dataset where this was listed is no longer accessible, thus preventing direct comparisons. A strength of RNA-Seq as compared to microarrays is a greater coverage of the transcriptome and the ability to cover all important genes such as *Muc5Ac*, which otherwise could be absent in microarrays [11].

A limitation of LCM-derived material to analyse gene expression is the low quantity of RNA that can be purified. This can be mitigated by harvesting from more tissue sections, but this is labor-intensive and can be limited by both the total amount of tissue available and time due to the degradation of RNA. RNA amplification prior to library prep is another option, but risks introducing bias in the transcripts [11]. We opted not to amplify the RNA prior to library prep, but rather use the low-input library kit *SMARTer Total Stranded RNA-Seq, Pico input mammalian – V3* (Takara Bio, Kusatsu, Japan) to minimize bias [181]. Furthermore, this kit prepares a library independent of poly(A)tails, thereby improving the detection of differentially-expressed, non-coding RNAs.

The specificity of the LCM cellular collection method is high, but not perfect. The overwhelming majority of sequenced cells were SMCs, while a small minority were parietal and endothelial cells. There was a miniscule yet non-zero presence of marker genes for all other cell-types, and it is uncertain if this indicates a real presence of these cell types or not. Considering the histological proximity between parietal and endothelial cells to the SMCs it is plausible that they indeed were present in a small yet negligible quantity in the sequenced material.

The marker gene expression for the SMCs or parietal cells between infected and sham control samples did not differ. In contrast, among the other marker genes

there was a significant difference in marker gene expression for plasma cells (*Zbp1*), monocytes (*Tet2*), and endothelial (*Xdh* and *S100a13*) cell types. Notably for the 43 endothelial cell markers, the only high-abundance gene (*Tspan8*) was not differentially expressed. However, none of the other marker genes for these cell types had a significant difference in gene expression levels between infected and sham control samples. Furthermore, these differentially-expressed marker genes are also present in other cell clusters which can be related to surface mucus cells, according to the PanglaoDB web browser. Thus, it is unclear if these marker genes indicate an actual difference in non-surface mucus cell populations that were potentially present in the sequenced tissue.

A potentially more specific collection method would be cell-sorting, and this method has been used to provide gastric mouse cells for downstream scRNA-Seq to study the pre-cancerous cascade in a chemically, rather than *H. pylori*, induced model [17,182,183]. scRNA-Seq allows the study of all viable cells that can be sorted quickly after the tissue has been removed, in contrast to just the cells that can reliably be separated with LCM which requires that the desired cell-types can be readily identified with the staining used on the slides in the microdissection apparatus. A downside with scRNA-Seq is the methodological and logistical challenge of the cell-sorting which is more difficult and time-consuming than simply snap-freezing the tissue for later processing and might thus limit experimental design for collecting the tissue. For example, the time-constraint derived from the need to avoid harvest-induced changes in gene expression limits the throughput and possibilities to combine scRNA-Seq with other experiments in the same harvest. Another downside is that scRNA-Seq remains a more expensive method than LCM and RNA-Seq combined.

Another alternative method for spatial transcriptomics is *in situ* RNA-Seq, wherein the transcriptome of a microscopy slide can currently be seen on a resolution as detailed as 0.5 μm , depending on the method [184]. This not only allows the analysis of gene expression with a cellular (or even sub-cellular) resolution, but also how the gene expression depends on where the cells are in relation to other cells and in the tissue. While this method is expensive, it might enable the exploration of research questions otherwise not possible to answer with other methodology. A substantial drawback of spatial transcriptomics is the difficulty in handling and analysing the data, which has yet to reach a state of maturity with standardized data formats, processing pipelines, and quality controls [184].

The RNA-Seq results were partially validated on the protein level with IF. Two out of three tested proteins (LGALS3BP and RPL19) differed significantly between *H. pylori*-infected and sham control mice, and there was also a significant positive correlation between protein and mRNA levels for LGALS3BP/*Lgals3bp*. Although the correlation between protein and mRNA levels for RPL19/*Rpl19* was not significant (adjusted $p = 0.055$), considering that the raw p -value was 0.037 and Spearman's ρ was 0.72 the correlation between protein and gene expression would likely have been significant with a larger number of samples. In contrast there was no significant difference between the two groups in COX6C signal intensity, nor a correlation between COX6C signal and *Cox6c* gene expression. Possible explanations for the results include post-transcriptional regulation maintaining COX6C protein levels despite declining mRNA levels, that unlike the more diffuse cytoplasmic proteins LGALS3BP and RPL19, the mitochondrial protein COX6C would be unevenly distributed close to the nuclei, and there could be substantial differences in what percentage of the visible cells in each sectioned gland that consists of the mitochondrial compartment, thus introducing a degree of noise into the readout. Even though the COX6C IF did not validate the differential expression of the OXPHOS-cluster, downregulation of OXPHOS activity has been shown experimentally in *C. rodentium* infection, in vitro IFN- γ stimulation and in vitro *H. pylori* infection [149,174]. Thus, it is likely that other validation approaches could have confirmed that the downregulation of the 29 OXPHOS genes observed in this study had biological consequences. The LCM RNA-Seq workflow described in the present study contributes to a greater usefulness of *H. Pylori* mouse models, and potentially other mucosal infection models as well, by providing another feasible tool for spatial transcriptomics.

Formation of new gene sets based upon this dataset

Gene set enrichment analysis (GSEA) and related bioinformatics analysis such as SetRank are predicated on previously established gene sets. Current gene sets related to *H. pylori* infection and/or SMCs are limited, which begs for an expansion of such gene sets. We plan to take all differentially-expressed genes presented in the current study and create a child term of “response to Gram-negative bacterium” (GO:0140460), which could be named “Mouse gastric corpus surface mucus cell response to *Helicobacter pylori* infection.”

Conclusion

This article shows that RNA-Seq of cells collected with LCM is a viable method to investigate host-pathogen interactions *in vivo* and provides a detailed protocol for it. The hallmarks of mouse SMCs both with and without early *H. pylori* infection have been revealed in greater detail than previous studies [11,17,23–25,183]. The gene expression is dominated by genes coding for secreted proteins such as TFF1, GKN1, GKN2, and MUC5AC, mitochondria RNA, and cytoskeleton proteins. Another gene which characterized the SMCs was *Rn7sk*, which we postulate regulates the gene pair *Gkn1/Gkn2*. Furthermore, the known murine gastric mucin glycome was expanded and shown to be dominated by large, heavily fucosylated glycans, whose biosynthesis could be accounted for by the 15 most abundant glycosyltransferases.

We found that *H. pylori* infection caused OXPHOS-related genes to be down-regulated. These genes were functionally associated with down-regulated, protein-synthesis-related genes, which then were functionally associated with up-regulated, IFN-induced genes. While the sequential steps of this hypothesis remain to be validated experimentally, the results suggest that early *H. pylori* infection elicits IFN stimulation of the SMCs *in vivo*, subsequently diminishing protein synthesis, particularly of proteins required for OXPHOS. Along this pathway there may be potential drug targets. We also found depletion of *Nkx6-3* gene expression to be one of the earliest events of the *H. pylori*-induced pre-cancerous cascade. This article contributes a piece to the greater puzzle of *H. pylori* pathogenesis, which in turn can lead to improved treatment and prevention of *H. pylori*-related death and disease in humans. Finally, the dataset can form the genesis for new gene sets, thus improving future gene set enrichment analysis.

Acknowledgements

We acknowledge the Centre for Cellular Imaging at the University of Gothenburg and the National Microscopy Infrastructure, NMI (VR-RFI 2016–00968) for providing assistance in microscopy.

The authors acknowledge support from the National Genomics Infrastructure in Stockholm funded by Science for Life Laboratory, the Knut and Alice Wallenberg Foundation and the Swedish Research Council, and SNIC/Uppsala Multidisciplinary Center for Advanced Computational Science for assistance with massively parallel sequencing and access to the UPPMAX computational infrastructure.

The computations were enabled by resources in project NAISS 2023/22–759 provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS) at

UPPMAX, funded by the Swedish Research Council through grant agreement no. 2022–06725.

We would also like to thank Kristina Guix Erhardsson for her help with manuscript editing during review.

Author contributions

CRediT: **Mattias Erhardsson**: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Visualization, Writing – original draft, Writing – review & editing; **Licinia Santos**: Investigation, Methodology, Writing – review & editing; **John Benktander**: Data curation, Formal analysis, Investigation, Methodology, Writing – review & editing; **Sinan Sharba**: Investigation, Methodology, Writing – review & editing; **Kaisa Thorell**: Funding acquisition, Methodology, Resources, Supervision, Validation, Writing – review & editing; **Sara Lindén**: Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – review & editing.

Funding

This work was financially supported by Cancerfonden [19 0301 Pj] and Vetenskapsrådet [2019–01152] to SL, and Stiftelsen Assar Gabrielssons Fond [FB21-89] and Magnus Bergvalls Stiftelse to KT. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Data availability statement

All data is freely available under permissive open licenses permitting reuse (CC-BY, CC0, or MIT).

The RNA-Seq FastQ files are available at the European Nucleotide Archive (ENA) with the project accession number PRJEB70775: <https://www.ebi.ac.uk/ena/browser/view/PRJEB70775> [185]. The NGI data delivery that support the findings of this study are openly available in Svensk Nationell Datatjänst (SND) at <https://doi.org/10.5878/ejy9-2895>, reference number 2023–166/1 [186].

The R-scripts, Jupyter notebook, and other files used for the data analysis is available on GitHub: <https://github.com/mattias-erhardsson/lmpc-infection-rnaseq> [187]. The GitHub repository has been archived at Zenodo: <https://doi.org/10.5281/zenodo.18151929> [188].

Supplemental files, including the ARRIVE checklist, are available in the same GitHub repository as above: <https://github.com/mattias-erhardsson/lmpc-infection-rnaseq/tree/main/Virulence%20supplemental%20files> [185].

The glycomics data are available in GlycoPost (<https://glycopost.glycosmos.org/>) at <http://doi.org/10.50821/GLYCOPOST-GPST000464> [189].

Open scholarship



This article has earned the Center for Open Science badges for Open Data and Open Materials through Open Practices Disclosure. The data and materials are openly accessible at <https://doi.org/10.5878/ejy9-2895>

ORCID

Mattias Erhardsson  <http://orcid.org/0000-0002-5631-2033>

Licinia Santos  <http://orcid.org/0000-0003-0056-4490>

John Benktander  <http://orcid.org/0009-0009-8464-1211>

Sinan Sharba  <http://orcid.org/0000-0003-0531-2137>

Kaisa Thorell  <http://orcid.org/0000-0002-4447-8968>

Sara Lindén  <http://orcid.org/0000-0002-7751-9586>

References

- [1] Hooi JKY, Lai WY, Ng WK, et al. Global prevalence of *Helicobacter pylori* infection: systematic review and meta-analysis. *Gastroenterology*. 2017;153(2):420–429. doi: [10.1053/j.gastro.2017.04.022](https://doi.org/10.1053/j.gastro.2017.04.022)
- [2] Kotilea K, Bontems P, Epidemiology TE. Diagnosis and risk factors of *Helicobacter pylori* infection. In: Kamiya S S Backert, editors. *Helicobacter pylori* in human diseases: advances in microbiology, infectious diseases and public health volume 11. Cham: Springer International Publishing; 2019. p 17–33.
- [3] Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: gLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209–249. doi: [10.3322/caac.21660](https://doi.org/10.3322/caac.21660)
- [4] Sharndama HC, Mba IE. *Helicobacter pylori*: an up-to-date overview on the virulence and pathogenesis mechanisms. *Braz J Microbiol*. 2022;53(1):33–50. doi: [10.1007/s42770-021-00675-0](https://doi.org/10.1007/s42770-021-00675-0)
- [5] Pohl D, Keller PM, Bordier V, et al. Review of current diagnostic methods and advances in *Helicobacter pylori* diagnostics in the era of next generation sequencing. *World J Gastroenterol*. 2019;25(32):4629–4660. doi: [10.3748/wjg.v25.i32.4629](https://doi.org/10.3748/wjg.v25.i32.4629)
- [6] Yuan C, Yu C, Sun Q, et al. Research on antibiotic resistance in *Helicobacter pylori*: a bibliometric analysis of the past decade. *Front Microbiol*. 2023;14:1208157. doi: [10.3389/fmicb.2023.1208157](https://doi.org/10.3389/fmicb.2023.1208157)
- [7] Muthupalani S, Ge Z, Joy J, et al. Muc5ac null mice are predisposed to spontaneous gastric antro-pyloric hyperplasia and adenomas coupled with attenuated *H. pylori*-induced corpus mucous metaplasia. *Lab Invest*. 2019;99(12):1887–1905. doi: [10.1038/s41374-019-0293-y](https://doi.org/10.1038/s41374-019-0293-y)
- [8] McGuckin MA, Every AL, Skene CD, et al. Muc1 mucin limits both *Helicobacter pylori* colonization of the murine gastric mucosa and associated gastritis. *Gastroenterology*. 2007;133(4):1210–1218. doi: [10.1053/j.gastro.2007.07.003](https://doi.org/10.1053/j.gastro.2007.07.003)
- [9] Linden SK, Sheng YH, Every AL, et al. Muc1 limits *Helicobacter pylori* infection both by steric hindrance

- and by acting as a releasable decoy. *PLoS Pathog.* 2009;5(10):e1000617. doi: [10.1371/journal.ppat.1000617](https://doi.org/10.1371/journal.ppat.1000617)
- [10] Ho SB, Takamura K, Anway R, et al. The adherent gastric mucous layer is composed of alternating layers of MUC5AC and MUC6 mucin proteins. *Dig Dis Sci.* 2004;49(10):1598–1606. doi: [10.1023/B:DDAS.0000043371.12671.98](https://doi.org/10.1023/B:DDAS.0000043371.12671.98)
 - [11] Mueller A, Merrell DS, Grimm J, et al. Profiling of microdissected gastric epithelial cells reveals a cell type-specific response to *Helicobacter pylori* infection. *Gastroenterology.* 2004;127(5):1446–1462. doi: [10.1053/j.gastro.2004.08.054](https://doi.org/10.1053/j.gastro.2004.08.054)
 - [12] Ansari S, Yamaoka Y. Animal models and *Helicobacter pylori* infection. *J Clin Med.* 2022;11(11):3141. doi: [10.3390/jcm11113141](https://doi.org/10.3390/jcm11113141)
 - [13] MohabatiMobarez A, Salmanian AH, Hosseini AZ, et al. Clearance of *Helicobacter pylori* with formulation rCagA and LPS in a mouse model. *Gene Rep.* 2020;19:100588. doi: [10.1016/j.genrep.2020.100588](https://doi.org/10.1016/j.genrep.2020.100588)
 - [14] Go DM, Lee SH, Lee SH, et al. Programmed death ligand 1-expressing classical dendritic cells mitigate *Helicobacter*-induced gastritis. *Cell Mol Gastroenterol Hepatol.* 2021;12(2):715–739. doi: [10.1016/j.jcmgh.2021.04.007](https://doi.org/10.1016/j.jcmgh.2021.04.007)
 - [15] Navabi N, Johansson ME, Raghavan S, et al. *Helicobacter pylori* infection impairs the mucin production rate and turnover in the murine gastric mucosa. *Infect Immun.* 2013;81(3):829–837. doi: [10.1128/IAI.01000-12](https://doi.org/10.1128/IAI.01000-12)
 - [16] Sorini C, Tripathi KP, Wu S, et al. Metagenomic and single-cell RNA-seq survey of the *Helicobacter pylori*-infected stomach in asymptomatic individuals. *JCI Insight.* 2023;8(4). doi: [10.1172/jci.insight.161042](https://doi.org/10.1172/jci.insight.161042)
 - [17] Bockerstett KA, Lewis SA, Wolf KJ, et al. Single-cell transcriptional analyses of spasmolytic polypeptide-expressing metaplasia arising from acute drug injury and chronic inflammation in the stomach. *Gut.* 2020;69(6):1027–1038. doi: [10.1136/gutjnl-2019-318930](https://doi.org/10.1136/gutjnl-2019-318930)
 - [18] Aguilar C, Pauzuolis M, Pompaiah M, et al. *Helicobacter pylori* shows tropism to gastric differentiated pit cells dependent on urea chemotaxis. *Nat Commun.* 2022;13(1):5878. doi: [10.1038/s41467-022-33165-4](https://doi.org/10.1038/s41467-022-33165-4)
 - [19] Podgorny OV, Lazarev VN. Laser microdissection: a promising tool for exploring microorganisms and their interactions with hosts. *J Microbiol Methods.* 2017;138:82–92. doi: [10.1016/j.mimet.2016.01.001](https://doi.org/10.1016/j.mimet.2016.01.001)
 - [20] Mueller A, O'Rourke J, Grimm J, et al. Distinct gene expression profiles characterize the histopathological stages of disease in *Helicobacter*-induced mucosa-associated lymphoid tissue lymphoma. *Proc Natl Acad Sci USA.* 2003;100(3):1292–1297. doi: [10.1073/pnas.242741699](https://doi.org/10.1073/pnas.242741699)
 - [21] Nomura S, Baxter T, Yamaguchi H, et al. Spasmolytic polypeptide expressing metaplasia to preneoplasia in *H. felis*-infected mice. *Gastroenterology.* 2004;127(2):582–594. doi: [10.1053/j.gastro.2004.05.029](https://doi.org/10.1053/j.gastro.2004.05.029)
 - [22] Wang Z, Gerstein M, Snyder M. RNA-seq: a revolutionary tool for transcriptomics. *Nat Rev Genet.* 2009;10(1):57–63. doi: [10.1038/nrg2484](https://doi.org/10.1038/nrg2484)
 - [23] Holmen Larsson JM, Thomsson KA, Rodriguez-Pineiro AM, et al. Studies of mucus in mouse stomach, small intestine, and colon. III. Gastrointestinal Muc5ac and Muc2 mucin O-glycan patterns reveal a regiospecific distribution. *Am J Physiol Gastrointest Liver Physiol.* 2013;305(5):G357–63. doi: [10.1152/ajpgi.00048.2013](https://doi.org/10.1152/ajpgi.00048.2013)
 - [24] Magalhaes A, Gomes J, Ismail MN, et al. Fut2-null mice display an altered glycosylation profile and impaired BabA-mediated *Helicobacter pylori* adhesion to gastric mucosa. *Glycobiology.* 2009;19(12):1525–1536. doi: [10.1093/glycob/cwp131](https://doi.org/10.1093/glycob/cwp131)
 - [25] Magalhaes A, Rossez Y, Robbe-Masselot C, et al. Muc5ac gastric mucin glycosylation is shaped by FUT2 activity and functionally impacts *Helicobacter pylori* binding. *Sci Rep.* 2016;6(1):25575. doi: [10.1038/srep25575](https://doi.org/10.1038/srep25575)
 - [26] Chahal G, Padra M, Erhardsson M, et al. A complex connection between the diversity of human gastric mucin O-glycans, *Helicobacter pylori* binding, *Helicobacter* infection and fucosylation. *Mol Cell Proteomics.* 2022;21(11):100421. doi: [10.1016/j.mcpro.2022.100421](https://doi.org/10.1016/j.mcpro.2022.100421)
 - [27] Lundstrom J, Urban J, Bojar D. Decoding glycomics with a suite of methods for differential expression analysis. *Cell Rep Meth.* 2023;3(12):100652. doi: [10.1016/j.crmeth.2023.100652](https://doi.org/10.1016/j.crmeth.2023.100652)
 - [28] Percie du Sert N, Hurst A V, Alam A, et al. The arrive guidelines 2.0: updated guidelines for reporting animal research. *PLOS Biol.* 2020;18(7):e3000410.
 - [29] Li D, Zand MS, Dye TD, et al. An evaluation of RNA-seq differential analysis methods. *PLOS One.* 2022;17(9):e0264246. doi: [10.1371/journal.pone.0264246](https://doi.org/10.1371/journal.pone.0264246)
 - [30] Santos L, Sharba S, Benktander J, et al. Treatment with (R)-alpha-methylhistamine or IL4 stimulates mucin production and decreases *Helicobacter pylori* density in the murine stomach. *Virulence.* 2025;16(1):2530173. doi: [10.1080/21505594.2025.2530173](https://doi.org/10.1080/21505594.2025.2530173)
 - [31] Rogers AB. Histologic scoring of gastritis and gastric cancer in mouse models. *Methods Mol Biol.* 2012;921:189–203. doi: [10.1007/978-1-62703-005-2_22](https://doi.org/10.1007/978-1-62703-005-2_22)
 - [32] Schindelin J, Arganda-Carreras I, Frise E, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods.* 2012;9(7):676–682. doi: [10.1038/nmeth.2019](https://doi.org/10.1038/nmeth.2019)
 - [33] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B stat Methodol.* 1995;57(1):289–300. doi: [10.1111/j.2517-6161.1995.tb02031.x](https://doi.org/10.1111/j.2517-6161.1995.tb02031.x)
 - [34] R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2020.
 - [35] Ewels PA, Peltzer A, Fillinger S, et al. The nf-core framework for community-curated bioinformatics pipelines. *Nat Biotechnol.* 2020;38(3):276–278. doi: [10.1038/s41587-020-0439-x](https://doi.org/10.1038/s41587-020-0439-x)
 - [36] Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res.* 2015;4:1521. doi: [10.12688/f1000research.7563.1](https://doi.org/10.12688/f1000research.7563.1)
 - [37] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550. doi: [10.1186/s13059-014-0550-8](https://doi.org/10.1186/s13059-014-0550-8)
 - [38] Ignatiadis N, Klaus B, Zaugg JB, et al. Data-driven hypothesis weighting increases detection power in genome-scale multiple testing. *Nat Methods.* 2016;13(7):577–580. doi: [10.1038/nmeth.3885](https://doi.org/10.1038/nmeth.3885)

- [39] Ma S, Dai Y. Principal component analysis based methods in bioinformatics studies. *Brief Bioinform.* **2011**;12(6):714–722. doi: [10.1093/bib/bbq090](https://doi.org/10.1093/bib/bbq090)
- [40] Healy J, McInnes L. Uniform manifold approximation and projection. *Nat Rev Methods Primers.* **2024**;4(1):82. doi: [10.1038/s43586-024-00363-x](https://doi.org/10.1038/s43586-024-00363-x)
- [41] Fabregat A, Sidiropoulos K, Garapati P, et al. The Reactome pathway knowledgebase. *Nucleic Acids Res.* **2016**;44(D1):D481–7. doi: [10.1093/nar/gkv1351](https://doi.org/10.1093/nar/gkv1351)
- [42] Kanehisa M, Goto S. Kegg: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **2000**;28(1):27–30.
- [43] Durinck S, Spellman PT, Birney E, et al. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc.* **2009**;4(8):1184–1191. doi: [10.1038/nprot.2009.97](https://doi.org/10.1038/nprot.2009.97)
- [44] Durinck S, Moreau Y, Kasprzyk A, et al. Biomart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics.* **2005**;21(16):3439–3440. doi: [10.1093/bioinformatics/bti525](https://doi.org/10.1093/bioinformatics/bti525)
- [45] Pagès HCM, Falcon S, Li N. AnnotationDbi: manipulation of SQLite-based annotations in Bioconductor. R package version 1.62.2 ed. **2023**.
- [46] Simillion C, Liechti R, Lischer HE, et al. Avoiding the pitfalls of gene set enrichment analysis with SetRank. *BMC Bioinf.* **2017**;18(1):151. doi: [10.1186/s12859-017-1571-6](https://doi.org/10.1186/s12859-017-1571-6)
- [47] Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**;13(11):2498–2504. doi: [10.1101/gr.1239303](https://doi.org/10.1101/gr.1239303)
- [48] Szklarczyk D, Gable AL, Lyon D, et al. String v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **2019**;47(D1):D607–D13.
- [49] Utriainen M, Morris JH. ClusterMaker2: a major update to ClusterMaker, a multi-algorithm clustering app for Cytoscape. *BMC Bioinf.* **2023**;24(1):134. doi: [10.1186/s12859-023-05225-z](https://doi.org/10.1186/s12859-023-05225-z)
- [50] Gustavsen JA, Pai S, Isserlin R, et al. Rcy3: network biology using Cytoscape from within R. *F1000Res.* **2019**;8(1774):1774. doi: [10.12688/f1000research.20887.2](https://doi.org/10.12688/f1000research.20887.2)
- [51] Wickham H, Averick M, Bryan J, et al. Welcome to the tidyverse. *J Open Source Softw.* **2019**;4(43):1686. doi: [10.21105/joss.01686](https://doi.org/10.21105/joss.01686)
- [52] Pundir S, Martin MJ, O'Donovan C. UniProt protein knowledgebase. In: Wu CH, CN Arighi KE Ross, editors. *Protein bioinformatics: from protein modifications and networks to proteomics*. New York (NY): Springer New York; **2017**. p 41–55.
- [53] Blake JA, Baldarelli R, Kadin JA, et al. Mouse genome database (MGD): knowledgebase for mouse-human comparative biology. *Nucleic Acids Res.* **2021**;49(D1):D981–D7.
- [54] Franzen O, Gan LM, Bjorkegren JLM. PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. *Database (Oxford).* **2019**;2019. doi: [10.1093/database/baz046](https://doi.org/10.1093/database/baz046)
- [55] Benktander J, Venkatakrishnan V, Padra JT, et al. Effects of size and geographical origin on Atlantic salmon, *Salmo salar*, mucin O-glycan repertoire. *Mol Cell Proteomics.* **2019**;18(6):1183–1196. doi: [10.1074/mcp.RA119.001319](https://doi.org/10.1074/mcp.RA119.001319)
- [56] Thomes L, Burkholz R, Bojar D. Glycowork: a Python package for glycan data science and machine learning. *Glycobiology.* **2021**;31(10):1240–1244. doi: [10.1093/glycob/cwab067](https://doi.org/10.1093/glycob/cwab067)
- [57] Kluyver T, Ragan-Kelley B, Pérez F, et al. Jupyter notebooks – a publishing format for reproducible computational workflows. In: Loizides F B Schmidt, editors. *20th International Conference on Electronic Publishing*; 01/01/16). IOS Press; **2016**. p. 87–90.
- [58] Van Rossum G, Fred L. Python 3 reference manual. Scotts Valley (CA): CreateSpace; **2009**.
- [59] Kolde R. pheatmap. *Pretty Heatmaps*; **2019**.
- [60] Neelamegham S, Aoki-Kinoshita K, Bolton E, et al. Updates to the symbol nomenclature for glycans guidelines. *Glycobiology.* **2019**;29(9):620–624. doi: [10.1093/glycob/cwz045](https://doi.org/10.1093/glycob/cwz045)
- [61] Ewels P, Magnusson M, Lundin S, et al. Multiqc: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics.* **2016**;32(19):3047–3048. doi: [10.1093/bioinformatics/btw354](https://doi.org/10.1093/bioinformatics/btw354)
- [62] Heuer J, Heuer F, Sturmer R, et al. The tumor suppressor TFF1 occurs in different forms and interacts with multiple partners in the human gastric mucus barrier: indications for diverse protective functions. *Int J Mol Sci.* **2020**;21(7):2508. doi: [10.3390/ijms21072508](https://doi.org/10.3390/ijms21072508)
- [63] Baldarelli RM, Smith CL, Ringwald M, et al. Mouse genome informatics g. Mouse genome informatics: an integrated knowledgebase system for the laboratory mouse. *Genetics.* **2024**;227(1). doi: [10.1093/genetics/iyae031](https://doi.org/10.1093/genetics/iyae031)
- [64] Menheniott TR, Kurklu B, Giraud AS. Gastrokines: stomach-specific proteins with putative homeostatic and tumor suppressor roles. *Am J Physiol Gastrointest Liver Physiol.* **2013**;304(2):G109–21. doi: [10.1152/ajpgi.00374.2012](https://doi.org/10.1152/ajpgi.00374.2012)
- [65] Fernandez-Blanco JA, Fakih D, Arike L, et al. Attached stratified mucus separates bacteria from the epithelial cells in COPD lungs. *JCI Insight.* **2018**;3(17). doi: [10.1172/jci.insight.120994](https://doi.org/10.1172/jci.insight.120994)
- [66] Moreno-Loshuertos R, Movilla N, Marco-Brualla J, et al. A mutation in mouse MT-ATP6 gene induces respiration defects and opposed effects on the cell tumorigenic phenotype. *Int J Mol Sci.* **2023**;24(2):1300. doi: [10.3390/ijms24021300](https://doi.org/10.3390/ijms24021300)
- [67] Jacob JT, Coulombe PA, Kwan R, et al. Types I and II keratin intermediate filaments. *Cold Spring Harb Perspect Biol.* **2018**;10(4):a018275. doi: [10.1101/cshperspect.a018275](https://doi.org/10.1101/cshperspect.a018275)
- [68] Machado AM, Figueiredo C, Touati E, et al. *Helicobacter pylori* infection induces genetic instability of nuclear and mitochondrial DNA in gastric cells. *Clin Cancer Res.* **2009**;15(9):2995–3002.
- [69] Rubio-Ramos A, Labat-de-Hoz L, Correias I, et al. The MAL protein, an integral component of specialized membranes, in normal cells and cancer. *Cells.* **2021**;10(5):1065. doi: [10.3390/cells10051065](https://doi.org/10.3390/cells10051065)
- [70] Sasaki T, Saito R, Oyama M, et al. Galectin-2 has bactericidal effects against *Helicobacter pylori* in a β -galactoside-dependent manner. *Int J Mol Sci.* **2020**;21(8):2697. doi: [10.3390/ijms21082697](https://doi.org/10.3390/ijms21082697)

- [71] Dugina V, Zwaenepoel I, Gabbiani G, et al. Beta and gamma-cytoplasmic actins display distinct distribution and functional diversity. *J Cell Sci.* **2009**;122(Pt 16):2980–2988.
- [72] Rodriguez-Pineiro AM, Bergstrom JH, Ermund A, et al. Studies of mucus in mouse stomach, small intestine, and colon. ii. Gastrointestinal mucus proteome reveals MUC2 and MUC5AC accompanied by a set of core proteins. *Am J Physiol Gastrointest Liver Physiol.* **2013**;305(5):G348–56.
- [73] Tamura M, Tanaka T, Fujii N, et al. Potential interaction between galectin-2 and MUC5AC in mouse gastric mucus. *Biol Pharm Bull.* **2020**;43(2):356–360. doi: [10.1248/bpb.b19-00705](https://doi.org/10.1248/bpb.b19-00705)
- [74] Varki A, Cummings RD, Esko JD, et al. Essentials of glycobiology. 4th ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; **2022**.
- [75] Stone EL, Ismail MN, Lee SH, et al. Glycosyltransferase function in core 2-type protein O glycosylation. *Mol Cell Biol.* **2009**;29(13):3770–3782. doi: [10.1128/MCB.00204-09](https://doi.org/10.1128/MCB.00204-09)
- [76] Fu J, Wei B, Wen T, et al. Loss of intestinal core 1-derived O-glycans causes spontaneous colitis in mice. *J Clin Invest.* **2011**;121(4):1657–1666.
- [77] Shiraishi N, Natsume A, Togayachi A, et al. Identification and characterization of three novel beta 1,3-N-acetylglucosaminyltransferases structurally related to the beta 1,3-galactosyltransferase family. *J Biol Chem.* **2001**;276(5):3498–3507.
- [78] Yeh JC, Hiraoka N, Petryniak B, et al. Novel sulfated lymphocyte homing receptors and their control by a core1 extension beta 1,3-N-acetylglucosaminyltransferase. *Cell.* **2001**;105(7):957–969. doi: [10.1016/S0092-8674\(01\)00394-4](https://doi.org/10.1016/S0092-8674(01)00394-4)
- [79] Guo JM, Zhang Y, Cheng L, et al. Molecular cloning and characterization of a novel member of the UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase family, pp-GalNAc-T12. *FEBS Lett.* **2002**;524(1–3):211–218.
- [80] Bennett EP, Hassan H, Hollingsworth MA, et al. A novel human UDP-N-acetyl-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase, GalNAc-T7, with specificity for partial GalNAc-glycosylated acceptor substrates. *FEBS Lett.* **1999**;460(2):226–230. doi: [10.1016/S0014-5793\(99\)01268-5](https://doi.org/10.1016/S0014-5793(99)01268-5)
- [81] Hagen FK, Ten Hagen KG, Beres TM, et al. cDNA cloning and expression of a novel UDP-N-acetyl-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase. *J Biol Chem.* **1997**;272(21):13843–13848.
- [82] Wandall HH, Hassan H, Mirgorodskaya E, et al. Substrate specificities of three members of the human UDP-N-acetyl-alpha-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase family, GalNAc-T1, -T2, and -T3. *J Biol Chem.* **1997**;272(38):23503–23514.
- [83] Ikehara Y, Sato T, Niwa T, et al. Apical Golgi localization of N,N'-diacetyllactosamine synthase, beta4GalNAc-T3, is responsible for LacdiNAc expression on gastric mucosa. *Glycobiology.* **2006**;16(9):777–785. doi: [10.1093/glycob/cwl005](https://doi.org/10.1093/glycob/cwl005)
- [84] Homeister JW, Thall AD, Petryniak B, et al. The alpha(1,3) fucosyltransferases FucT-IV and FucT-VII exert collaborative control over selectin-dependent leukocyte recruitment and lymphocyte homing. *Immunity.* **2001**;15(1):115–126. doi: [10.1016/S1074-7613\(01\)00166-2](https://doi.org/10.1016/S1074-7613(01)00166-2)
- [85] Zara J, Hagen FK, Ten Hagen KG, et al. Cloning and expression of mouse UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase-T3. *Biochem Biophys Res Commun.* **1996**;228(1):38–44.
- [86] Arike L, Holmen-Larsson J, Hansson GC. Intestinal Muc2 mucin O-glycosylation is affected by microbiota and regulated by differential expression of glycosyltransferases. *Glycobiology.* **2017**;27(4):318–328. doi: [10.1093/glycob/cww134](https://doi.org/10.1093/glycob/cww134)
- [87] Bennett EP, Hassan H, Mandel U, et al. Cloning and characterization of a close homologue of human UDP-N-acetyl-alpha-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase-T3, designated GalNAc-T6. Evidence for genetic but not functional redundancy. *J Biol Chem.* **1999**;274(36):25362–25370.
- [88] Zhou D, Dinter A, Gutierrez Gallego R, et al. A beta-1,3-N-acetylglucosaminyltransferase with poly-N-acetylglucosamine synthase activity is structurally related to beta-1,3-galactosyltransferases. *Proc Natl Acad Sci U S A.* **1999**;96(2):406–411.
- [89] Zhou D, Berger EG, Hennen T. Molecular cloning of a human UDP-galactose: glcNAc beta1,3GalNAc beta1, 3 galactosyltransferase gene encoding an O-linked core3-elongation enzyme. *Eur J Biochem.* **1999**;263(2):571–576.
- [90] Zhou D, Henion TR, Jungalwala FB, et al. The beta 1,3-galactosyltransferase beta 3galt-v is a stage-specific embryonic antigen-3 (ssea-3) synthase. *J Biol Chem.* **2000**;275(30):22631–22634.
- [91] Lee H, Wang P, Hoshino H, et al. Alpha1,4GlcNAc-capped mucin-type O-glycan inhibits cholesterol alpha-glucosyltransferase from *Helicobacter pylori* and suppresses *H. pylori* growth. *Glycobiology.* **2008**;18(7):549–558. doi: [10.1093/glycob/cwn037](https://doi.org/10.1093/glycob/cwn037)
- [92] Mazumder B, Sampath P, Seshadri V, et al. Regulated release of L13a from the 60S ribosomal subunit as a mechanism of transcript-specific translational control. *Cell.* **2003**;115(2):187–198. doi: [10.1016/S0092-8674\(03\)00773-6](https://doi.org/10.1016/S0092-8674(03)00773-6)
- [93] Holmgren AM, Miller KD, Cavanaugh SE, et al. Bst2/tetherin is induced in neurons by type I interferon and viral infection but is dispensable for protection against neurotropic viral challenge. *J Virol.* **2015**;89(21):11011–11018. doi: [10.1128/JVI.01745-15](https://doi.org/10.1128/JVI.01745-15)
- [94] Yoo H, Park SH, Ye SK, et al. IFN-gamma-induced BST2 mediates monocyte adhesion to human endothelial cells. *Cell Immunol.* **2011**;267(1):23–29.
- [95] Zhu M, Lv J, Wang W, et al. CMPK2 is a host restriction factor that inhibits infection of multiple coronaviruses in a cell-intrinsic manner. *PLoS Biol.* **2023**;21(3):e3002039. doi: [10.1371/journal.pbio.3002039](https://doi.org/10.1371/journal.pbio.3002039)
- [96] Wang H, Peng W, Wang J, et al. Human cytomegalovirus UL23 antagonizes the antiviral effect of interferon-gamma by restraining the expression of specific IFN-stimulated genes. *Viruses.* **2023**;15(4):1014. doi: [10.3390/v15041014](https://doi.org/10.3390/v15041014)
- [97] Selvakumar TA, Bhushal S, Kalinke U, et al. Identification of a predominantly interferon-lambda-induced transcriptional profile in murine intestinal epithelial cells. *Front Immunol.* **2017**;8:1302. doi: [10.3389/fimmu.2017.01302](https://doi.org/10.3389/fimmu.2017.01302)

- [98] McNab F, Mayer-Barber K, Sher A, et al. Type I interferons in infectious disease. *Nat Rev Immunol*. 2015;15(2):87–103. doi: [10.1038/nri3787](https://doi.org/10.1038/nri3787)
- [99] Schoggins JW. Interferon-stimulated genes: roles in viral pathogenesis. *Curr Opin Virol*. 2014;6:40–46. doi: [10.1016/j.coviro.2014.03.006](https://doi.org/10.1016/j.coviro.2014.03.006)
- [100] Schoggins JW, Wilson SJ, Panis M, et al. A diverse range of gene products are effectors of the type I interferon antiviral response. *Nature*. 2011;472(7344):481–485. doi: [10.1038/nature09907](https://doi.org/10.1038/nature09907)
- [101] Kim BH, Shenoy AR, Kumar P, et al. A family of IFN-gamma-inducible 65-kD GTPases protects against bacterial infection. *Science*. 2011;332(6030):717–721. doi: [10.1126/science.1201711](https://doi.org/10.1126/science.1201711)
- [102] Goers L, Kim K, Stedman TC, et al. Shigella IpaH9.8 limits GBP1-dependent LPS release from intracytosolic bacteria to suppress caspase-4 activation. *Proc Natl Acad Sci USA*. 2023;120(15):e2218469120. doi: [10.1073/pnas.2218469120](https://doi.org/10.1073/pnas.2218469120)
- [103] Li Z, Qu X, Liu X, et al. GBP5 is an interferon-induced inhibitor of respiratory syncytial virus. *J Virol*. 2020;94(21). doi: [10.1128/JVI.01407-20](https://doi.org/10.1128/JVI.01407-20)
- [104] Mears HV, Sweeney TR. Mouse Ifit1b is a cap1-RNA-binding protein that inhibits mouse coronavirus translation and is regulated by complexing with Ifit1c. *J Biol Chem*. 2020;295(51):17781–17801.
- [105] Huntzinger E, Sinteff J, Morlet B, et al. Helz2: a new, interferon-regulated, human 3'-5' exoribonuclease of the RNB family is expressed from a non-canonical initiation codon. *Nucleic Acids Res*. 2023;51(17):9279–9293.
- [106] Jacquet S, Pontier D, Etienne L. Rapid evolution of HERC6 and duplication of a chimeric HERC5/6 gene in rodents and bats suggest an overlooked role of HERCs in mammalian immunity. *Front Immunol*. 2020;11:605270. doi: [10.3389/fimmu.2020.605270](https://doi.org/10.3389/fimmu.2020.605270)
- [107] Cheriya V, Leaman DW, Borden EC. Emerging roles of FAM14 family members (G1P3/ISG 6–16 and ISG12/IFI27) in innate immunity and cancer. *J Interferon Cytokine Res*. 2011;31(1):173–181.
- [108] Kurokawa C, Iankov ID, Galanis E. A key anti-viral protein, RSAD2/VIPERIN, restricts the release of measles virus from infected cells. *Virus Res*. 2019;263:145–150. doi: [10.1016/j.virusres.2019.01.014](https://doi.org/10.1016/j.virusres.2019.01.014)
- [109] Metz P, Dazert E, Ruggieri A, et al. Identification of type I and type II interferon-induced effectors controlling hepatitis C virus replication. *Hepatology*. 2012;56(6):2082–2093. doi: [10.1002/hep.25908](https://doi.org/10.1002/hep.25908)
- [110] Tachibana M, Hashino M, Watanabe K, et al. Interferon gamma-induced GTPase promotes invasion of *Listeria monocytogenes* into trophoblast giant cells. *Sci Rep*. 2015;5(1):8195. doi: [10.1038/srep08195](https://doi.org/10.1038/srep08195)
- [111] Uthaiyah RC, Praefcke GJ, Howard JC, et al. IIGP1, an interferon-gamma-inducible 47-kDa GTPase of the mouse, showing cooperative enzymatic activity and GTP-dependent multimerization. *J Biol Chem*. 2003;278(31):29336–29343.
- [112] Feng CG, Weksberg DC, Taylor GA, et al. The p47 GTPase Irg-47 (Irgm1) links host defense and hematopoietic stem cell proliferation. *Cell STEM Cell*. 2008;2(1):83–89. doi: [10.1016/j.stem.2007.10.007](https://doi.org/10.1016/j.stem.2007.10.007)
- [113] Feng CG, Zheng L, Lenardo MJ, et al. Interferon-inducible immunity-related GTPase Irgm1 regulates IFN gamma-dependent host defense, lymphocyte survival and autophagy. *Autophagy*. 2009;5(2):232–234. doi: [10.4161/auto.5.2.7445](https://doi.org/10.4161/auto.5.2.7445)
- [114] Lodermeier V, Suhr K, Schrott N, et al. 90k, an interferon-stimulated gene product, reduces the infectivity of HIV-1. *Retrovirology*. 2013;10(1):111. doi: [10.1186/1742-4690-10-111](https://doi.org/10.1186/1742-4690-10-111)
- [115] Chou WC, Jha S, Linhoff MW, et al. The NLR gene family: from discovery to present day. *Nat Rev Immunol*. 2023;23(10):635–654. doi: [10.1038/s41577-023-00849-x](https://doi.org/10.1038/s41577-023-00849-x)
- [116] Eskildsen S, Justesen J, Schierup MH, et al. Characterization of the 2'-5'-oligoadenylate synthetase ubiquitin-like family. *Nucleic Acids Res*. 2003;31(12):3166–3173.
- [117] Caprara G, Prosperini E, Piccolo V, et al. Parp14 controls the nuclear accumulation of a subset of type I IFN-inducible proteins. *J Immunol*. 2018;200(7):2439–2454.
- [118] Jiang H, Li Y, Shen M, et al. Interferon-alpha promotes MHC I antigen presentation of islet beta cells through STAT1-IRF7 pathway in type 1 diabetes. *Immunology*. 2022;166(2):210–221. doi: [10.1111/imm.13468](https://doi.org/10.1111/imm.13468)
- [119] Nakamura K, Jinnin M, Kudo H, et al. The role of PSMB9 upregulated by interferon signature in the pathophysiology of cutaneous lesions of dermatomyositis and systemic lupus erythematosus. *Br J Dermatol*. 2016;174(5):1030–1041. doi: [10.1111/bjd.14385](https://doi.org/10.1111/bjd.14385)
- [120] Thery F, Martina L, Asselman C, et al. Ring finger protein 213 assembles into a sensor for ISGylated proteins with antimicrobial activity. *Nat Commun*. 2021;12(1):5772. doi: [10.1038/s41467-021-26061-w](https://doi.org/10.1038/s41467-021-26061-w)
- [121] Mavrommatis E, Arslan AD, Sassano A, et al. Expression and regulatory effects of murine schlafen (slfn) genes in malignant melanoma and renal cell carcinoma*. *J Biol Chem*. 2013;288(46):33006–33015. doi: [10.1074/jbc.M113.460741](https://doi.org/10.1074/jbc.M113.460741)
- [122] Chelbi-Alix MK, Pelicano L, Quignon F, et al. Induction of the PML protein by interferons in normal and APL cells. *Leukemia*. 1995;9(12):2027–2033.
- [123] Kajita AI, Morizane S, Takiguchi T, et al. Interferon-gamma enhances TLR3 expression and anti-viral activity in keratinocytes. *J Invest Dermatol*. 2015;135(8):2005–2011. doi: [10.1038/jid.2015.125](https://doi.org/10.1038/jid.2015.125)
- [124] Chen Y, Lin J, Zhao Y, et al. Toll-like receptor 3 (TLR3) regulation mechanisms and roles in antiviral innate immune responses. *J Zhejiang Univ Sci B*. 2021;22(8):609–632.
- [125] Sanmarco LM, Wheeler MA, Gutierrez-Vazquez C, et al. Gut-licensed IFNgamma(+) NK cells drive LAMP1(+)TRAIL(+) anti-inflammatory astrocytes. *Nature*. 2021;590(7846):473–479. doi: [10.1038/s41586-020-03116-4](https://doi.org/10.1038/s41586-020-03116-4)
- [126] Wang Y, Yuan S, Jia X, et al. Mitochondria-localised ZNFX1 functions as a dsRNA sensor to initiate antiviral responses through MAVS. *Nat Cell Biol*. 2019;21(11):1346–1356. doi: [10.1038/s41556-019-0416-0](https://doi.org/10.1038/s41556-019-0416-0)
- [127] Straub IR, Janer A, Weraarpachai W, et al. Loss of CHCHD10-CHCHD2 complexes required for respiration underlies the pathogenicity of a CHCHD10 mutation in ALS. *Hum Mol Genet*. 2018;27(1):178–189.
- [128] Anand R, Reichert AS, Kondadi AK. Emerging roles of the MICOS complex in cristae dynamics and

- biogenesis. *Biol (Basel)*. 2021;10(7):600. doi: [10.3390/biology10070600](https://doi.org/10.3390/biology10070600)
- [129] Huang G, Li H, Zhang H. Abnormal expression of mitochondrial ribosomal proteins and their encoding genes with cell apoptosis and diseases. *Int J Mol Sci*. 2020;21(22):8879. doi: [10.3390/ijms21228879](https://doi.org/10.3390/ijms21228879)
- [130] Signes A, Fernandez-Vizarra E. Assembly of mammalian oxidative phosphorylation complexes I-V and supercomplexes. *Essays Biochem*. 2018;62(3):255–270.
- [131] Giorgio B V, Soriano E, Dabbeni-Sala ME, et al. Cyclophilin D modulates mitochondrial F0F1-ATP synthase by interacting with the lateral stalk of the complex*. *J Biol Chem*. 2009;284(49):33982–33988.
- [132] Young C, Batkovskytė D, Kitamura M, et al. A hypomorphic variant in the translocase of the outer mitochondrial membrane complex subunit TOMM7 causes short stature and developmental delay. *HGG Adv*. 2023;4(1):100148. doi: [10.1016/j.xhgg.2022.100148](https://doi.org/10.1016/j.xhgg.2022.100148)
- [133] Rawat K, Das S, Vivek Vinod BS, et al. Targeted depletion of BTF3a in macrophages activates autophagic pathway to eliminate *Mycobacterium tuberculosis*. *Life Sci*. 2019;220:21–31. doi: [10.1016/j.lfs.2019.01.035](https://doi.org/10.1016/j.lfs.2019.01.035)
- [134] Li P, Fei H, Wang L, et al. Pdcd5 regulates cell proliferation, cell cycle progression and apoptosis. *Oncol Lett*. 2018;15(1):1177–1183. doi: [10.3892/ol.2017.7401](https://doi.org/10.3892/ol.2017.7401)
- [135] Kipper K, Mansour A, Pulk A. Neuronal RNA granules are ribosome complexes stalled at the pre-translocation state. *J Mol Biol*. 2022;434(20):167801. doi: [10.1016/j.jmb.2022.167801](https://doi.org/10.1016/j.jmb.2022.167801)
- [136] Eaton JD, Davidson L, Bauer DLV, et al. Xrn2 accelerates termination by RNA polymerase II, which is underpinned by CPSF73 activity. *Genes Dev*. 2018;32(2):127–139. doi: [10.1101/gad.308528.117](https://doi.org/10.1101/gad.308528.117)
- [137] Takaoka A, Wang Z, Choi MK, et al. Dai (DLM-1/ZBP1) is a cytosolic DNA sensor and an activator of innate immune response. *Nature*. 2007;448(7152):501–505. doi: [10.1038/nature06013](https://doi.org/10.1038/nature06013)
- [138] Lv K, Gong C, Antony C, et al. HectD1 controls hematopoietic stem cell regeneration by coordinating ribosome assembly and protein synthesis. *Cell STEM Cell*. 2021;28(7):1275–90 e9. doi: [10.1016/j.stem.2021.02.008](https://doi.org/10.1016/j.stem.2021.02.008)
- [139] Zhang P, Cao L, Zhou R, et al. The lncRNA Neat1 promotes activation of inflammasomes in macrophages. *Nat Commun*. 2019;10(1):1495. doi: [10.1038/s41467-019-09482-6](https://doi.org/10.1038/s41467-019-09482-6)
- [140] Takase H, Hamanaka G, Ohtomo R, et al. Transcriptome profiling of mouse corpus callosum after cerebral hypoperfusion. *Front Cell Dev Biol*. 2021;9:685261. doi: [10.3389/fcell.2021.685261](https://doi.org/10.3389/fcell.2021.685261)
- [141] Nystrom EEL, Birchenough GMH, van der Post S, et al. Calcium-activated chloride channel regulator 1 (CLCA1) controls mucus expansion in colon by proteolytic activity. *EBioMedicine*. 2018;33:134–143. doi: [10.1016/j.ebiom.2018.05.031](https://doi.org/10.1016/j.ebiom.2018.05.031)
- [142] Xu HC, Grusdat M, Pandya AA, et al. Type I interferon protects antiviral CD8+ T cells from NK cell cytotoxicity. *Immunity*. 2014;40(6):949–960. doi: [10.1016/j.immuni.2014.05.004](https://doi.org/10.1016/j.immuni.2014.05.004)
- [143] West PK, McCorkindale AN, Guennewig B, et al. The cytokines interleukin-6 and interferon-alpha induce distinct microglia phenotypes. *J Neuroinflammation*. 2022;19(1):96. doi: [10.1186/s12974-022-02441-x](https://doi.org/10.1186/s12974-022-02441-x)
- [144] Hobson BD, Stanley AT, De Los Santos MB, et al. Conserved and cell type-specific transcriptional responses to IFN-gamma in the ventral midbrain. *Brain Behav Immun*. 2023;111:277–291. doi: [10.1016/j.bbi.2023.04.008](https://doi.org/10.1016/j.bbi.2023.04.008)
- [145] Christgen S, Place DE, Zheng M, et al. The IFN-inducible GTPase IRGB10 regulates viral replication and inflammasome activation during influenza A virus infection in mice. *Eur J Immunol*. 2022;52(2):285–296. doi: [10.1002/eji.202149305](https://doi.org/10.1002/eji.202149305)
- [146] Yoon JH, Kim O, Eun JW, et al. Multiple genetic mutations caused by NKX6.3 depletion contribute to gastric tumorigenesis. *Sci Rep*. 2018;8(1):17609. doi: [10.1038/s41598-018-35733-5](https://doi.org/10.1038/s41598-018-35733-5)
- [147] Yoon JH, Kim O, Nam SW, et al. NKX6.3 regulates reactive oxygen species production by suppressing NF-κB and DNMT1 activities in gastric epithelial cells. *Sci Rep*. 2017;7(1):2807. doi: [10.1038/s41598-017-02901-y](https://doi.org/10.1038/s41598-017-02901-y)
- [148] Ullrich A, Sures I, D'Egidio M, et al. The secreted tumor-associated antigen 90k is a potent immune stimulator. *J Biol Chem*. 1994;269(28):18401–18407. doi: [10.1016/S0021-9258\(17\)32322-0](https://doi.org/10.1016/S0021-9258(17)32322-0)
- [149] Maiti AK, Sharba S, Navabi N, et al. IL-4 protects the mitochondria against TNFalpha and IFNgamma induced insult during clearance of infection with *Citrobacter rodentium* and *Escherichia coli*. *Sci Rep*. 2015;5(1):15434. doi: [10.1038/srep15434](https://doi.org/10.1038/srep15434)
- [150] Alarcon-Millan J, Martinez-Carrillo DN, Peralta-Zaragoza O, et al. Regulation of GKN1 expression in gastric carcinogenesis: a problem to resolve (review). *Int J Oncol*. 2019;55(3):555–569.
- [151] Han S, Fink J, Jorg DJ, et al. Defining the identity and dynamics of adult gastric isthmus stem cells. *Cell STEM Cell*. 2019;25(3):342–56 e7. doi: [10.1016/j.stem.2019.07.008](https://doi.org/10.1016/j.stem.2019.07.008)
- [152] The Human Protein Atlas. IL33 [internet]. Stockholm, Sweden: The Human Protein Atlas; 2024 [cited 2024 Aug 15]. Available from: <https://www.proteinatlas.org/ENSG00000137033-IL33/tissue/stomach>
- [153] Buzzelli JN, Chalinor HV, Pavlic DI, et al. IL33 is a stomach alarmin that initiates a skewed Th2 response to injury and infection. *Cell Mol Gastroenterol Hepatol*. 2015;1(2):203–21 e3. doi: [10.1016/j.jcmgh.2014.12.003](https://doi.org/10.1016/j.jcmgh.2014.12.003)
- [154] Kuo CJ, Chen CY, Lo HR, et al. *Helicobacter pylori* induces IL-33 production and recruits ST-2 to lipid rafts to exacerbate inflammation. *Cells*. 2019;8(10):1290. doi: [10.3390/cells8101290](https://doi.org/10.3390/cells8101290)
- [155] Lee HJ, Nam KT, Park HS, et al. Gene expression profiling of metaplastic lineages identifies CDH17 as a prognostic marker in early stage gastric cancer. *Gastroenterology*. 2010;139(1):213–25 e3. doi: [10.1053/j.gastro.2010.04.008](https://doi.org/10.1053/j.gastro.2010.04.008)
- [156] Bandiera R, Wagner RE, Britto-Borges T, et al. Rn7sk small nuclear RNA controls bidirectional transcription of highly expressed gene pairs in skin. *Nat Commun*. 2021;12(1):5864. doi: [10.1038/s41467-021-26083-4](https://doi.org/10.1038/s41467-021-26083-4)
- [157] Schreurs O, Karatsaidis A, Balta MG, et al. Expression of keratins 8, 18, and 19 in epithelia of atrophic oral lichen planus. *Eur J Oral Sci*. 2020;128(1):7–17. doi: [10.1111/eos.12666](https://doi.org/10.1111/eos.12666)
- [158] Linden S, Mahdavi J, Semino-Mora C, et al. Role of ABO secretor status in mucosal innate immunity and

- H. pylori infection. PLOS Pathog. 2008;4(1):e2. doi: [10.1371/journal.ppat.0040002](https://doi.org/10.1371/journal.ppat.0040002)
- [159] Squire JM, Guerreiro MJ, Sidebotham RL, et al. Quantitative MUC5AC and MUC6 mucin estimations in gastric mucus by a least-squares minimization method. Anal Biochem. 2013;439(2):204–211. doi: [10.1016/j.ab.2013.04.013](https://doi.org/10.1016/j.ab.2013.04.013)
- [160] Gerhard M, Lehn N, Neumayer N, et al. Clinical relevance of the *Helicobacter pylori* gene for blood-group antigen-binding adhesin. Proc Natl Acad Sci. 1999;96(22):12778–12783. doi: [10.1073/pnas.96.22.12778](https://doi.org/10.1073/pnas.96.22.12778)
- [161] Linden SK, Wickstrom C, Lindell G, et al. Four modes of adhesion are used during *Helicobacter pylori* binding to human mucins in the oral and gastric niches. Helicobacter. 2008;13(2):81–93. doi: [10.1111/j.1523-5378.2008.00587.x](https://doi.org/10.1111/j.1523-5378.2008.00587.x)
- [162] Aspholm-Hurtig M, Dailide G, Lahmann M, et al. Functional adaptation of Baba, the H. pylori ABO blood group antigen binding adhesin. Science. 2004;305(5683):519–522. doi: [10.1126/science.1098801](https://doi.org/10.1126/science.1098801)
- [163] Linden S, Nordman H, Hedenbro J, et al. Strain- and blood group-dependent binding of *Helicobacter pylori* to human gastric MUC5AC glycoforms. Gastroenterology. 2002;123(6):1923–1930. doi: [10.1053/gast.2002.37076](https://doi.org/10.1053/gast.2002.37076)
- [164] Padra M, Benktander J, Robinson K, et al. Carbohydrate-dependent and antimicrobial peptide defence mechanisms against *Helicobacter pylori* infections. In: Backert S, editor. Molecular mechanisms of inflammation: induction, resolution and escape by *Helicobacter pylori* current topics in microbiology and immunology. Springer; 2019. p.179–207. doi:[10.1007/978-3-030-15138-6_8](https://doi.org/10.1007/978-3-030-15138-6_8).
- [165] Quintana-Hayashi MP, Rocha R, Padra M, et al. Baba-mediated adherence of pediatric ulcerogenic H. pylori strains to gastric mucins at neutral and acidic pH. Virulence. 2018;9(1):1699–1717. doi: [10.1080/21505594.2018.1532243](https://doi.org/10.1080/21505594.2018.1532243)
- [166] Chahall G, Padra M, Erhardsson M, et al. A complex connection between the diversity of human gastric mucin O-glycans *Helicobacter pylori* binding, *Helicobacter* infection and fucosylation. Submitted.
- [167] Moonens K, Van den Broeck I, Okello E, et al. Structural insight in the inhibition of adherence of F4 fimbriae producing enterotoxigenic *Escherichia coli* by llama single domain antibodies. Vet Res. 2015;46:14. doi: [10.1186/s13567-015-0151-x](https://doi.org/10.1186/s13567-015-0151-x)
- [168] Padra JT, Sundh H, Jin C, et al. *Aeromonas salmonicida* binds differentially to mucins isolated from skin and intestinal regions of Atlantic salmon in an N-acetylneuraminic acid-dependent manner. Infect Immun. 2014;82(12):5235–5245. doi: [10.1128/IAI.01931-14](https://doi.org/10.1128/IAI.01931-14)
- [169] Ota H, Nakayama J, Momose M, et al. *Helicobacter pylori* infection produces reversible glycosylation changes to gastric mucins. Virchows Arch. 1998;433(5):419–426. doi: [10.1007/s004280050269](https://doi.org/10.1007/s004280050269)
- [170] Magalhaes A, Marcos-Pinto R, Nairn AV, et al. *Helicobacter pylori* chronic infection and mucosal inflammation switches the human gastric glycosylation pathways. Biochim Biophys Acta. 2015;1852(9):1928–1939. doi: [10.1016/j.bbadis.2015.07.001](https://doi.org/10.1016/j.bbadis.2015.07.001)
- [171] Jin C, Kenny DT, Skoog EC, et al. Structural diversity of human gastric mucin glycans. Mol Cell Proteomics. 2017;16(5):743–758. doi: [10.1074/mcp.M117.067983](https://doi.org/10.1074/mcp.M117.067983)
- [172] Chevalier Curt M J, Lecoïnte K, Mihalache A, et al. Alteration or adaptation, the two roads for human gastric mucin glycosylation infected by *Helicobacter pylori*. Glycobiology. 2015;25(6):617–631.
- [173] Machado AMD, Desler C, Bøggild S, et al. *Helicobacter pylori* infection affects mitochondrial function and DNA repair, thus, mediating genetic instability in gastric cells. Mech Ageing Dev. 2013;134(10):460–466. doi: [10.1016/j.mad.2013.08.004](https://doi.org/10.1016/j.mad.2013.08.004)
- [174] Holland RL, Bosi KD, Seeger AY, et al. Restoration of mitochondrial structure and function within *Helicobacter pylori* VacA intoxicated cells. Adv Microbiol. 2023;13(8):399–419.
- [175] Winter JA, Letley DP, Cook KW, et al. A role for the vacuolating cytotoxin, VacA, in colonization and *Helicobacter pylori*-induced metaplasia in the stomach. J Infect Dis. 2014;210(6):954–963.
- [176] Correa P, Piazuelo MB. The gastric precancerous cascade. J Dig Dis. 2012;13(1):2–9. doi: [10.1111/j.1751-2980.2011.00550.x](https://doi.org/10.1111/j.1751-2980.2011.00550.x)
- [177] Barrozo RM, Cooke CL, Hansen LM, et al. Functional plasticity in the type IV secretion system of *Helicobacter pylori*. PLOS Pathog. 2013;9(2):e1003189. doi: [10.1371/journal.ppat.1003189](https://doi.org/10.1371/journal.ppat.1003189)
- [178] Petersen CP, Mills JC, Goldenring JR. Murine models of gastric corpus preneoplasia. Cell Mol Gastroenterol Hepatol. 2017;3(1):11–26. doi: [10.1016/j.jcmgh.2016.11.001](https://doi.org/10.1016/j.jcmgh.2016.11.001)
- [179] Sheh A, Fox JG. The role of the gastrointestinal microbiome in *Helicobacter pylori* pathogenesis. Gut Microbes. 2013;4(6):505–531. doi: [10.4161/gmic.26205](https://doi.org/10.4161/gmic.26205)
- [180] Lee CYQ, Chan YT, Cheok YY, et al. *Helicobacter pylori* infection elicits type I interferon response in human monocytes via toll-like receptor 8 signaling. J Immunol Res. 2022;2022(1):3861518.
- [181] Song Y, Milon B, Ott S, et al. A comparative analysis of library prep approaches for sequencing low input transcriptome samples. BMC Genomics. 2018;19(1):696. doi: [10.1186/s12864-018-5066-2](https://doi.org/10.1186/s12864-018-5066-2)
- [182] Bockerstett KA, Wong CF, Koehm S, et al. Molecular characterization of gastric epithelial cells using flow cytometry. Int J Mol Sci. 2018;19(4):1096. doi: [10.3390/ijms19041096](https://doi.org/10.3390/ijms19041096)
- [183] Bockerstett KA, Lewis SA, Noto CN, et al. Single-cell transcriptional analyses identify lineage-specific epithelial responses to inflammation and metaplastic development in the gastric corpus. Gastroenterology. 2020;159(6):2116–29 e4. doi: [10.1053/j.gastro.2020.08.027](https://doi.org/10.1053/j.gastro.2020.08.027)
- [184] Tian L, Chen F, Macosko EZ. The expanding vistas of spatial transcriptomics. Nat Biotechnol. 2023;41(6):773–782. doi: [10.1038/s41587-022-01448-2](https://doi.org/10.1038/s41587-022-01448-2)
- [185] Erhardsson M, Santos L, Benktander J, et al. Rna-seq FastQ files for the article “The mouse gastric surface epithelial cell and its response to early *Helicobacter pylori* infection” available at the European Nucleotide Archive (ENA), accession number PRJEB70775 [Internet]. Cambridge, UK: EMBL-EBI; 2025 [cited 2026 Jan 2]. Available from: <https://www.ebi.ac.uk/ena/browser/view/PRJEB70775>
- [186] Erhardsson M, Santos L, Benktander J, et al. Ngi data delivery for the article “The mouse gastric surface epithelial cell and its response to early *Helicobacter pylori* infection”

- available at Svensk Nationell Datatjänst (SND) [Internet]. Göteborg, Sweden: SND; 2023 Available from [cited 2026 Jan 2]. doi: [10.5878/ejy9-2895](https://doi.org/10.5878/ejy9-2895)
- [187] Erhardsson M, Santos L, Benktander J, et al. Supplemental files and code for the article “The mouse gastric surface epithelial cell and its response to early *Helicobacter pylori* infection” available at GitHub repository mattias-erhardsson/lmpc-infection-rnaseq [Internet]. San Francisco USA: GitHub; 2025 Available from [cited 2026 Jan 2]. Available from: <https://github.com/mattias-erhardsson/lmpc-infection-rnaseq>
- [188] Erhardsson M, Santos L, Benktander J, et al. GitHub repository for the article “The mouse gastric surface epithelial cell and its response to early *Helicobacter pylori* infection” archived at Zenodo [Internet]. CERN, Switzerland: Zenodo; 2025 [cited 2026 Jan 2]. doi: [10.5281/zenodo.18109350](https://doi.org/10.5281/zenodo.18109350)
- [189] Erhardsson M, Santos L, Benktander J, et al. Glycomics data for the article “The mouse gastric surface epithelial cell and its response to early *Helicobacter pylori* infection” available at GlycoPost [Internet]. Tokyo, Japan: GlycoPost; 2024 [cited 2026 Jan 2]. doi: [10.50821/GLYCOPOST-GPST000464](https://doi.org/10.50821/GLYCOPOST-GPST000464)