Manuscript No.: 33590

Response to Reviewer 03476682

Thank you for your comments.
We have revised the manuscript accordingly. Because there are no specific corrections we have tried to outline our changes in broader terms (we have used the Markup in Word to show changes).
3. Comments

1. *Please summaries key results with reference to study objectives. (b)The overall manuscript is too long*.

Our response:
(a) The study objectives:
(i) to identify genes (and processes) regulated upon *H. pylori* infection in a novel cell line that is polyclonal in nature and produces an adherent mucous layer. The E12 cell line supports *H. pylori* infection where the parental HT29 cell line does not. The adherent mucus provides an environment similar to that present in the stomach. Infection is through host-pathogen interaction within the mucus, not by invasion.
(ii) to identify key glycosylation-related events involved in these host-pathogen interactions. Initial contact between host and pathogen is most likely through the most distal glycans on both host and pathogen glycoconjugates.

The key results:
Microarray analysis identified 276 genes that were significantly differentially expressed upon *H. pylori* infection. Six genes were involved in glycosylation-related processes. Both MUC20 and REG4 are genes of interest in this model system. Gene ontology analysis was consistent with previous studies on *H. pylori* infection which suggests both shorter and simpler glycan structures. Some lectins (e.g. LGAL series), enzymes regulating activated sugar levels and transporters of these same metabolites appear important in the host response.

(b) We have shortened the Introduction, Results and Discussion sections. Please see changes in the Markup. To summarise:

Introduction: Shortened by approx. 140 words (9 lines)
Materials & Methods: Generally untouched – few minor linguistic revisions and simplifications
Results: Shortened by approx. 200 words (17 lines)
Discussion: Shortened by approx. 160 words (15 lines)

2. *Please discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias.*

Our response:
The conclusions from this study are limited by the use of a single cell line (E12) and a single *H. pylori* strain (26695). We would argue that the use of a single cell line is justified based on the novel features of this cell line. If the argument is that by using several cell lines reproducibility

across the lines would support our findings, we would argue that we do not expect reproducibility because the E12 cell line is quite unique with both its adherent mucous layer and its polyclonal nature.

Infection of E12 cells with other *H. pylori* strains would be very interesting. This could have been investigated using qRT-PCR with a few candidate genes. Microarray analysis would have been too costly. RNA-Seq or similar might have been possible. Other *H. pylori* strains will almost certainly produce different expression profiles in host cells and possibly a large enough analysis could differentiate common mechanisms across strains from unique features for individual strains. Strain 26695 is phenotypically $BabA^-$ so would be expected to produce a different response in host cells to a $BabA^+$ strain, though expression changes downstream of the adhesin-receptor mechanism might be seen to converge. In the same way a $TLR2^-$ host will differ from a $TLR2^+$ host. Each report in the literature adds a little to the story and may lead to a stage where all host variables or all strain variables could be investigated exhaustible and simultaneously.

The paper would be strengthened by the addition of supporting protein and glycan analysis. Transcript levels of glycosyltransferases and related enzymes may not directly translate into protein levels, though it is generally accepted that there is good correlation. Enzyme levels might not directly translate into potential glycan structures. It would be possible to use antibodies to Lewis structures, for example, to confirm predicted changes in these glycan structures. It would also be possible to use lectin arrays to identify predicted glycan changes. In addition, it would be possible to release glycans from glycoconjugates and directly determine structures by mass spectrometry and other methods. Whether the sensitivity of these methods is sufficient to detect subtle changes in the described model could be an issue.

Fold changes in gene transcript levels between uninfected and infected cells were small. Microarray analysis is not a very precise technique: it has the advantage of providing global expression changes but it requires a more robust technique such as qRT-PCR to provide accurate fold changes. In the current study we were somewhat limited in our conclusions by the variable behaviour of the host cells even when uninfected. It is possible that the cells are unstable because of the high levels of mucus that they are secreting or it may be that the polyclonal population changes in cell type composition (though not permanently). This variability, whatever the reason, prevented several apparently regulated genes showing statistical significance.

This is not an exhaustive list but in our opinion these are the major limitations. Most of these issues have been alluded to in the manuscript. In order to keep the manuscript as short as possible, we feel that it would not be appropriate to expand these issues in the paper.

4. Minor linguistic revision is necessary.
Small changes have been made in each section (see MarkUp).

**Additional note to Reviewer and Editor:**
Please note that during this revision the reference list was reduced from 73 to 69. Updating this in Endnote means that the first author's name is no longer in **BOLD**. Since this has to be manually changed I would like to leave this until I have responded to ALL reviewers' comments.