February 14, 2014

Dear Editor,

Please find enclosed the edited manuscript in Word format (file name: 5898-editedTopicHighlight.doc).

Title: Biomarkers in Inflammatory Bowel Diseases: Current Status and Proteomics Identification Strategies

Author: Tue Bennike, Svend Birkelund, Allan Stensballe, Vibeke Andersen

Name of Journal: *World Journal of Gastroenterology*

ESPS Manuscript NO: 5898

The manuscript has been improved according to the suggestions of reviewers:

1. Format has been updated
The layout of the first page has been updated to meet the requirements. The reference list has been updated with PMIDs, besides the DOI. We have furthermore added three acknowledgements (Lau Sennels, Gunna Christiansen, Kathrine B. Christensen) for inputs and proof-reading, after the conclusion chapter (page 21).
Overall – Changed "post-translational" to "posttranslational"
Overall – Changed greek letters (α, …) to alpha, beta…
Page 1 – changed format and added ESPS Manuscript NO.

2. Revision has been made according to the suggestions of the reviewers:

(01429020)
Dear reviewer,

First of all thank you for agreeing to review our paper, and your comments! We have included all corrections.

a. The incorporation of tables is a great idea, and have added to the quality of the paper. Two new tables have been added to emphasize the feasibility and perspectives in using proteomics for identifying novel biomarker candidates, and limitations of current strategies. As you suggested, Table 1 lists the mentioned commonly used biomarkers. We considered constructing Table 2 with all identified proteomic biomarker candidates; however, such a table would include several hundreds of biomarkers and most studies cannot be directly compared due to differences in investigated disease groups etc. which complicates the table even more. As our main purpose of including the table is to make the message clearer, we have made the table as an overview of the mentioned articles, with the main findings and perspectives. As such it is our hope that, the table will help better explain and clarify that proteomics is feasible for obtaining new biomarker candidates, and provide the reader with a better overview of novel and possible coming biomarkers.

Page 8 – Added "(Table 1)"
Page 12 – Added "(Table 2)"

b. You make a fair point, that we, unintentionally, are disappointing the reader in the text. Our motivation for including these statements regarding the poor hopes of finding novel biomarkers on the DNA or RNA level was to put emphasis on the advantages of protein studies. To improve the message, we have implemented the following changes:

Page 6 – Deleted "A likely explanation is that in many cases genomic information alone is not sufficient to describe the complex interactions necessary to fully understand disease etiologies, and in terms of diagnosis, genetics only provide risk factors. Transcriptomics is an extension of genomics, where the transcriptome is sequenced (mRNA characterization), which allows for identifying upregulated genes. However, the protein abundance in the cell does not always correlate with the transcript level, as for example protein turn-over i.e. synthesis and degradation rates is not taken into account[24,33]. This is a potential issue when searching for biomarkers later to be detected in the serum or stools using other methods than transcriptomics."
Page 6 – Deleted section name "Proteomics-driven Biomarker Discovery" and changed the section text to "As mentioned, proteins represent an obvious target for biomarker discovery studies, and as PTMs dramatically increases the diversity of the mature proteins, they represents a promising area for IBD biomarker studies. PTMs are introduced after translation of the RNA transcripts (Figure 1), hence analyzing DNA and RNA transcripts does not directly provide information about the PTMs. A key technique capable of measuring…"
Page 7 – Added "Proteomics is the large-scale identification of proteins, and can often cover the study of all expressed proteins by an organism (the proteome)"
Page 7 – Changed "Proteomics has been limited mainly by the speed and sensitivity of the mass spectrometers." to "Previously, proteomics has been limited mainly by the speed and sensitivity of the mass spectrometers."
Page 7 – Changed "Such biomarkers could be of great diagnostic and prognostic value, and provide a better understanding of the involved disease pathways" to "Knowledge of disease mechanisms in general can lead to improved future development of preventive and treatment strategies. Thus, the clinical use of a panel of biomarkers represents a diagnostic and prognostic tool of potentially great value.."

c. We have not commented on the different expensiveness of the genomics, transcriptomics and proteomic studies, as it is our hope that these techniques only will be needed in the biomarker discovery phase. As soon as promising biomarker candidates have been identified using these expensive techniques, the biomarkers can hopefully be included in a (cheaper) diagnostic tool such as ELISA or protein chips. However, I agree that this was not explicit in the text, so we have implemented the following:

Page 7 – added "Antibodies to identified biomarkers for CD and UC found by proteomics can be generated for development of immunoassays and immunohistochemistry for evaluating the markers clinical use in routine tests, less expensive than sequencing genomes, transcriptomes or MS driven proteomics.".

d. In regards to your final comment, we agree. Unfortunately, we have not been able to come up with any biomarkers highly usable for clinical practice to differentiate CD and UC. As the review concludes, these have yet to be identified and are likely present e.g. in forms as disease specific PTMs. However, it is our hope that we through the review have been successful in demonstrating that the proteomics driven biomarker identification is a feasible approach for identifying new and hopefully more specific and sensitive, biomarkers than currently available. Besides

emphasizing the feasibility of the proteomics approach, and highlighting the likely presence of IBD biomarkers, we also hope that we successfully conveyed the message that posttranslational modifications and global protein profiling studies, represents two new promising targets for biomarker studies, made available by the recent technological development.

Again, thank you for thorough review of our paper. Your comments have definitely improved the quality of the paper!

Sincerely,
Tue Bennike

(00504581)
Dear reviewer,

First of all, thank you for agreeing to review our paper, and your comments! We have included all corrections.

a. Regarding the readability, you are right that the paper is not easily read for people not working directly in the field. This was of course not the intention and we have made several corrections to ease the readability, by including more explanations for the terminology and methodology. We have, furthermore, included a figure 2, to better explain the difference between the techniques MALDI/SELDI TOF MS and LC-ESI MS Thank you for this important observation! We hope you agree that the paper now is more easily read. Improvements include:

Figure 2 legend - Two commonly used MS techniques. A: MALDI/SELDI TOF MS, where the peptide or protein sample is dried on a target plate. Subsequently, a laser is used to evaporate the dried sample, and the generated gas phase ions are analyzed by the mass spectrometer. B: LC-ESI MS, where the liquid peptide (or protein) sample is separated on a LC column, and sequentially eluted often over several hours. The eluted peptides are injected directly into the mass spectrometer by ESI and analyzed.

Page 2 – Added "protein"
Page 2 – Changed **"**exposure to**"** to "treatment with"
Page 2 – Added "(determination and quantification of the complete protein content)"
Page 2 – Changed **"**Understanding of" to "Knowledge of"
Page 2 – Changed "The intestinal tissue remains an obvious place to search for novel biomarkers, which later can be screened for in blood, urine or stool" to "The intestinal tissue remains an obvious place to search for novel biomarkers, which blood, urine or stool later can be screened for"
Page 2 – Changed "accurately " to "accurate"
Page 3 – Changed ""have" to "has"
Page 5 – Changed **"**quality of life" to "life quality"
Page 5 – Changed "due to" to "measured on"
Page 5 – Changed "correct diagnosis" to "better diagnosis"
Page 5 – Deleted "or specific proteoforms"
Page 5 – Added ", including the proteins"
Page 5 – Deleted "and cellular behavior"

Page 5 – Chaged "for" to "in" twice
Page 5 – Changed "Each gene can give rise to several transcripts by alternating promoters, differential transcription terminations and alternative splicing events or alleles, resulting in a total of roughly 100,000 different mRNA transcripts. After translation of the mRNA code into protein…" to "During protein synthesis, the DNA code is first transcribed into different RNA transcripts. Each gene can give rise to several RNA transcripts resulting in a total of roughly 100,000 different RNA transcripts (Figure 1), which in turn are translated into 100,000 different proteins. After translation, …"
Page 5 – Added "More than 200 distinct biologically relevant PTMs have been identified[27], so each RNA transcript can be more than 200 different proteoforms.".
Page 6 – Added "and the final mature protein products are termed proteoforms".
Page 6 – Added "which constitutes the human proteome (all different proteins present)."
Page 6 – Added "When searching for biomarkers, it is possible to analyze the target sample on the DNA level, the RNA transcript level or the protein level. Techniques for studying an organisms DNA code (genome) or RNA transcripts (transcriptome) have the advantage that entire genomes and transcriptomes can be sequenced and studied"
Page 6 – Deleted "using these techniques."
Page 6 – Changed "Using genomics techniques" to "Using genomic sequencing techniques"
Page 6 – Added "cellular"
Page 6 – Added "pathways involved in"
Page 6 – Changed "peptides derived from proteins from which the molecular weight" to "peptides derived from proteins. From the m/z the molecular weight"
Page 6 – Added "Proteins represent an obvious target for biomarker discovery studies, and as PTMs dramatically increases the diversity and in many cases function of the mature proteins, they represent a promising area for IBD biomarker studies. PTMs are introduced after translation of the RNA transcripts (Figure 1), hence analyzing DNA and RNA transcripts does not directly provide information about the PTMs."
Page 7 – Added "Proteomics is the large-scale identification of proteins, and can often cover the study of all expressed proteins by an organism (the proteome)."
Page 7 – Changed "The technique is based on…" to "The bottom-up mass spectrometry (MS) strategy is based on…"
Page 7 – Added ", which have been enzymatically cleaved into minor peptides"
Page 7 – Changed "The technique is based on measuring the mass-to-charge ratio (m/z) of peptides derived from proteins, from which the molecular weight of the peptide can be calculated[25]. Each peptide is, furthermore, isolated and collided with a gas, from which the peptide fragment m/z's are measured. The proteins in the sample are subsequently identified by searching the peptide masses and fragment m/z's signals against an in silico generated database, inferred from a reference database of protein sequences." to "From the measured m/z's the molecular weight of the intact peptides can be calculated[25]. In addition to calculating the intact masses, the peptides are collided with an inert gas which fragments the peptides, and the fragment m/z's are measured. The proteins in the sample are subsequently identified by searching the peptide masses and fragment m/z's against an *in silico* generated database, inferred from a reference database of protein sequences. By matching the *in silico* calculated peptide masses and fragment m/z's to the measured, the peptides and hence the proteins, are identified."
Page 7 – Deleted "In addition, more than 10,000 proteins can be identified from HeLa cell lines, constituting 94% of the transcriptome of the cell line."
Page 7 – Changed Serum Biomarkers section, by incorporating a table and putting many facts from the text here "ASCA is an antibody with affinity for sequences in the cell wall of the yeast Saccharomyces cerevisiae. 39-79% of the CD patients tested positive, whereas only 5- 15% of the UC patients tested positive. ASCA has, furthermore, been found in 14-18% of a healthy control group, so the specificity and sensitivity for CD patients is relatively low, limiting the diagnostic value of the marker…"to "ASCA is an antibody with affinity for antigens in the cell wall of the

yeast Saccharomyces cerevisiae. In comparison to UC patients, CD patients are often positive for ASCA (Table 1)[41–43]. However, a substantial amount of healthy controls are also positive for ASCA positive[44], indicating that specificity and sensitivity for CD patients are relatively low; limiting the diagnostic value of the marker in differentiating CD from UC.
ANCAs are antibodies with affinity for neutrophil granules. The antibodies have been found in a variety of immune conditions, including Wegener's granulomatosis and rheumatoid arthritis (RA)[4]. When staining for ANCA, different patterns have been observed for UC and CD patients using immunofluorescence microscopy (Table 1), and mainly UC patients display perinuclear ANCA (pANCA) staining compared to CD patients [41,42,45]. Nonetheless, like the case of ASCA, a substantial amount of healthy controls are pANCA positive [44]"
Page 9 – Changed "IBD Diagnosis and Known Biomarkers" to "Diagnosis of Inflammatory Bowel Disease and Known Biomarkers"
Page 9 – Changed "IBD" to "the IBDs"
Page 9 – Changed "non are" to "no single biomarker is"
Page 9 – Changed "Serum Biomarkers" to "Antibodies and Serum Biomarkers"
Page 10 – Added "Lastly,"
Page 10 – Changed "TNF)α and IL-1β" to "TNF)-alpha and IL-1-beta"
Page 10 – Added "an"
Page 10 – Changed "not enough to separate the two groups" to "insufficient to differentiate CD patients from UC patients"
Page 10 – Added "are a"
Page 10 – Changed "from which the biomarkers are likely to originate" to "both from which potential biomarkers are likely to originate"
Page 10 – Changed "usually UC" to "UC usually"
Page 10 – Changed "markers" to "biomarkers"
Page 10 – Added "may affect any part of the gastrointestinal tract"
Page 10 – Changed "launched to amongst others understand the medical significance" to "launched to unravel the medical significance"
Page 11 – Changed "The stool, therefore, seems like a promising place to identify novel biomarkers with high sensitivity and specificity and which may explain the disease etiologies." to "Novel biomarkers with high sensitivity and specificity may, therefore, be identified from stools"
Page 11 – Added "for IBD screening"
Page 11 – Added "(Table 1)"
Page 11 – Deleted headlines "Calprotein" and "Lactoferrin"
Page 11 – Changed "evaluations" to "examination"
Page 11 – Changed "are" to "is"
Page 11 – Added "disease"
Page 12 – Changed "*In vitro* purified colon epithelial cells from patients" to "Human adenocarcinoma cells were in vitro"
Page 12 – Changed "was" to "is"
Page 12 – Added "intensities"
Page 12 – Changed "The proteins were first separated using two-dimensional polyacrylamide gel electrophoresis (2D-PAGE). By staining all protein in the gels, different samples (gels) can be compared and differentiating protein spots can be visually identified. Spots of interests was cut from the gel, the proteins were digested to specific peptides using tryptic in-gel digestion, and the proteins were finally identified using matrix-assisted laser desorption/ionization–time of flight (MALDI-TOF) MS. Some proteins were furthermore validated by western blotting." to ". Spots of interests was cut from the gel with a knife and the proteins were enzymatically digested to specific peptides using the protease trypsin (in-gel digestion). The digestion of the proteins is an essential step for protein identification, as no MS technique currently exist that can identify thousands of intact proteins in a complex sample. This is only possible when using digested proteins, i.e. peptides. The proteins were identified based on the peptides using MS, with the

technique called matrix-assisted laser desorption/ionization–time of flight (MALDI-TOF) MS (Figure 2A). MALDI-TOF MS is a sensitive technique, but it involves placing a few drops of the sample on a plate which is left to dry prior to analysis. During analysis a laser is used to evaporate small spots from the dried droplet and ions in the produced gas are analyzed by MS. In the study…"

Page 12 – Changed "interferon-γ, IL-1β and IL-6 (TNF-α" to "interferon-gamma, IL-1-beta and IL-6 (TNF-alpha"

Page 12 – Deleted "-group"

Page 12 – Changed "digested to specific peptides using tryptic in-gel digestion" to "enzymatically digested to specific peptides using the protease trypsin (in-gel digestion)"

Page 13 – Changed "Some proteins were furthermore validated by western blotting" to "Subsequently, human epithelial cells were isolated from UC and CD patients. Based on the findings, the samples were analyzed for the enzyme indoleamine-2.3-dioxygene using antibodies by western blotting"

Page 13 – Added "commonly"

Page 13 – Added "Subsequently, human epithelial cells were isolated from UC patients and CD patients. Based on the findings, the samples were analyzed for the enzyme indoleamine-2.3-dioxygene using antibodies by western blotting."

Page 13 – Deleted "now"

Page 13 – Deleted "Obviously, serum markers do not have this consideration, and the material is easier to obtain."

Page 13 – Changed "This information might be highly important when it comes to explaining disease etiologies." To "Information regarding non-changing proteins might prove equally important as changing proteins for studies seeking to describe disease etiologies"

Page 13 – Added "MALDI-TOF MS can also be conducted using intact proteins without prior enzymatic protein digestion. A variant of MALDI-TOF MS is to spot the protein mixture on a modified surface, to which the intact proteins bind and subsequently the intact masses of the proteins can be obtained by MS. This technique is called surface-enhanced laser desorption/ionization time of flight mass spectrometry (SELDI-TOF MS) (Figure 2A). However, when studying intact proteins using MALDI-TOF MS or SELDI-TOF MS, one usually does not obtain identification of the detected signals. Electrospray ionization (ESI) remains the only MS technique for identifying and quantifying several thousands of proteins in a high-throughput manner (Figure 2B). ESI involves spraying the digested proteins directly into the MS. By incorporating liquid chromatography (LC) with columns prior to the ESI process, the peptides can be separated and sequentially eluted over several hours. This gives the MS systems enough time to analyze a large proportion of the eluted peptides which subsequently can be identified. In this way, large-scale proteomic studies …"

Page 13 – Added "Several of the unidentified signals were subsequently identified by using MALDI MS/MS, western blotting, and ELISA assay"

Page 14 – Added "post infection"

Page 14 – Deleted "*in-situ*"

Page 15 – Deleted "of the four"

Page 15 – Changed "By characterizing the protein profiles with SELDI-TOF MS" to "By characterizing the serum only by the m/z signals and not identified proteins with SELDI-TOF MS"

Page 15 – Added "were identified:"

Page 15 – Deleted "protein profiles of"

Page 16 – Changed "Subjects" to "individuals"

Page 16 – Changed "tryptic digested" to "separated only in one dimension in contrast to 2D-PAGE, which allowed the entire visualized gel lane to be cut into pieces and digested with trypsin"

Page 16 – Added "The feasibility of the biomarker candidates remains to be validated."

Page 16 – Added "for IBD"
Page 16 – Chaged "analyzed" to "separated"
Page 16 – Deleted "the protein profiles of"
Page 17 – Changed "being functionally related" to "and were found to be functionally related"
Page 17 – Deleted "the protein profiles of"
Page 17 – Added "MS"
Page 17 – Changed "The layer keeps" to "The mucosa prevents"
Page 17 – Added "and MALDI-TOF MS."
Page 17 – Added "bacterial ribosomal RNA genes were analyzed by oligonucleotide fingerprinting and"
Page 17 – Changed "meta-proteome" to "combined proteome"
Page 17 - Added "presenting new possibilities for diagnosis and therapy."
Page 18– Deleted "the first approved anti-TNFα agent accepted for IBD treatment"
Page 18 – Changed "is" to "was"
Page 18 – Deleted "the protein profiles of"
Page 18 – Added "MS"
Page 18 – Added "Interestingly, the group was unable to confirm the findings by Meuwis MA et al.[71], that PF4 could be a biomarker for infliximab response, emphasizing that the biomarker candidates need further validation."
Page 18 – Changed "Many studies have, as apparent" to "As apparent, many studies have"
Page 19 – Added "(Table 2)"
Page 19 – Added "and all biomarker candidates identified so far lacks follow-up validation studies"
Page 19 – Deleted "the first"
Page 19 – Deleted "In addition, "
Page 19 – Changed "diagnosis and prognosis" to "diagnosis, prognosis and identify novel targets for therapy"
Page 19 – Changed "that such biomarkers" to "the presence of biomarkers"
Page 19 – Changed "Posttranslational Modifications in IBDs" to "Posttranslational Modifications as Biomarkers"
Page 19 – Added "and characterize"
Page 19 – Changed "differentiated CD from UC, demonstrating the presence of protein biomarkers." To "differentiated CD patients from UC patients. However, only based on unidentified m/z signals and not using identified protein or peptide biomarkers, from which the disease etiologies might be better explained. Nonetheless, these studies demonstrate the presence of usable biomarkers yet to be identified"
Page 19 – Changed "In addition, when biomarkers are identified it opens for the possibility of constructing ELISA tests and protein array chips, where antibodies are used to detect the abundance of one or more antigens" to "Identified biomarkers hold the potential for designing diagnostic ELISA tests and protein array chips, where antibodies are used to detect the abundance of one or more antigens"
Page 19 – Added "and characterize"
Page 20 – Deleted "substoichiometric,"
Page 20 – Deleted "Projects seeking to identify novel diagnostic and prognostic biomarkers in the intestinal tissue should consider the state of inflammation of the tissue, depending on what marker is being investigated."
Page 20 – Added "Several PTMs are known to be involved in the inflammatory responses, and PTMs could be involved in the IBD disease etiologies"
Page 21 – Changed "Today, the critical diagnosis of UC and CD patients remains troublesome" to "The diagnosis of UC and CD patients remains difficult, especially in the early stages of the diseases, and early and accurate diagnosis of IBD-patients is crucial."
Page 21 – Changed "However, none of the identified biomarkers have been implemented in

clinical daily use, and the diagnosis remains a combination of inflammation biomarkers and histological evaluation based on colonscopy, amongst others." to "However, none of the identified biomarkers have been implemented in clinical daily use, and the diagnosis is based on a combination of disease history, colonscopy inflammation biomarkers and histological evaluation."

Page 21 – Deleted "and therapeutic targets"

Page 21 – Changed "tremendous" to "immense"

Page 21 – Changed "utilized" to "used"

Page 21 – Added "Besides protein abundances, PTMs represent promising targets for biomarker discovery studies."

Page 21 – Added "and prognostic"

Page 21 – Changed "Besides aiding physicians in making a correct diagnosis and treatment strategy as well as function as a prognostic marker, such knowledge might also help identify disease pathways and ultimately lead to a cure" to "Besides aiding physicians in making a correct diagnosis and treatment strategy, knowledge of disease specific proteins and PTMs might identify disease pathways and new targets for therapeutic, leading to improved pharmaceutical drugs"

Page 21 – Added "ultimately leading to improved treatment strategies"

Again, thank you for thorough review of our paper. Your comments have definitely improved the quality of the paper!

Sincerely,
Tue Bennike

3 References and typesetting were corrected

In addition to the DOIs, PMIDs have been added to all references when available.

Thank you again for publishing our manuscript in the *World Journal of Gastroenterology!*

Sincerely,

Tue Bennike, MS
Department of Health Science and Technology
Aalborg University
Fredrik Bajers Vej 3B
9220 Aalborg East, Denmark
Fax: +45-9815-4008
Phone: +45-2613-9003
E-mail: tbe@hst.aau.dk