Institute
of Experimental
Medicine, CAS

Reviewer's assessment of the PhD thesis "Molecular prognostic and predictive markers in colorectal cancer" by Nazila Navvabi

The thesis focuses at a very important topic of identification of novel molecular markers of colorectal carcinoma (CRC). As the incidence of the disease increases, particularly in younger patients, thorough knowledge of mechanisms involved in CRC initiation and progression is of utmost importance. Suitable prognostic and predictive molecular markers could help to identify persons at greater risk of CRC and, in subjects with the disease, to monitor the response to the therapy.

In the thesis, three types of potential markers are investigated: deregulation of genes involved in alternative splicing; changes in expression of base excision repair glycosylases; and expression of miRNA-140 in relation to sensitivity to oxaliplatin treatment. Most of the data presented in the thesis concerns altered expression of genes encoding factors regulating alternative splicing. The results of this research were published in the article in Cancer Genomics & Proteomics with Nazila Navvabi being the first author. The other two topics were also published in journals with impact factor; Nazila Navvabi contributed as the co-author.

The thesis itself has a standard structure consisting of the introduction, describing in detail various types of diagnostic, prognostic and predictive markers used in CRC. A separate section is devoted to description of mechanisms of alternative splicing in CRC. Further, goal and hypotheses are identified, materials and methods described, and the results interpreted and discussed. There is quite extensive list of cited literature with a total of 256 articles.

Overall, the thesis is well-written and based on solid research published in Q1/Q2 journals. However, while reviewing the text, some questions and comments arose that are summarized below.

General comment:

Most of the experiments presented in the thesis were based on the analyses of mRNA expression. This is acknowledged and correctly stated by the author. However, in Hypothesis 1 (page 49), the text reads:

"... deregulated MBNL activity has been linked to various malignancies, including colorectal cancer. However, no study so far has investigated activity of the all the MBNL paralogs in colorectal cancer. We aimed to examine possible deregulation of all three MBNL paralogs in CRC by comparing expression of their mRNA...". This implies that the author aims to use the mRNA expression data to assess MBNL activity. As the correlation between mRNA expression and protein expression/activity is generally weak, the hypothesis should be better formulated to avoid the confusion. A similar comment concerns Hypothesis 2 and Hypothesis 4 and several other places in the text, including the discussion (see my comment below). The author should have clearly stated that the data is based on mRNA expression which may have certain limitations while interpreting the results.

Minor comments and questions:

Abstract

It would have been optimal if more information on FOXP1 and EPB41L3, as well as tumor-related CD44 variants was provided.

Abbreviations

HOGG1 should read: Human 8-oxoguanine DNA N-glycosylase 1

ELISA should read: Enzyme-linked immunosorbent assay

Introduction

Biomarkers in CRC (2.1.2)

Overall, this chapter would have benefitted from a better structure – e.g. mechanisms leading to genomic instability, or description of various types of cancer biomarkers (diagnostic, predictive, prognostic). Adding paragraphs or text highlights would be very helpful. In addition, cancer biomarkers are further discussed in chapter 2.1.3 and further. In effect, the text is sometimes difficult to follow.

Molecular marker, as defined on page 10, does not necessarily need to be a substance, it could be any parameter that can be easily analyzed.

Diagnostic biomarker (2.1.4)

Tissue biomarkers:

cytokeratins: information on tissue type that is used for diagnosis should be provided

SATB2: high expression levels were found in the lower GI, including appendix, colon and rectum. Given this fact, how could the expression of this protein be used as a specific marker of CRC? The same is true for cadherin 17 that was found to be expressed on the surface epithelium of the duodenum, ileum, appendix and colon, as well as for GPA, that is expressed in stomach, small intestine, colon and rectum, but also in well-differentiated CRC tumors.

Blood biomarkers:

Are changes in genes associated with chromosomal alterations (KRAS, APC) specific for CRC?

Prognostic biomarkers (2.1.5.)

Tissue biomarkers – this section should be better organized into paragraphs for individual types of biomarkers

Cytosine preceding guanine island methylator phenotype: CpG abbreviation – the explanation provided in the thesis, although originating from a published article, is not a standard interpretation; commonly it stands for cytosine – phosphate – guanine.

p53: how about prognostic value of mutations in p53? It is not discussed in the text.

SMAD4: what does „loss of function of chromosome 18 gene function in protein expression" mean?

Blood biomarkers

Concentrations of cfDNA: The sentence "...APC, KRAS, and p53 mutations in the serum..." The formulation is not correct as mutations are found in DNA, not in plasma/serum

Predictive biomarkers (2.1.6.)

Please see my comment in Prognostic biomarkers regarding the structure of the section.

Tissue biomarkers – it should be specified in which tissue the biomarkers are assessed.

PIK3CA: what could be the mechanism of a positive effect of aspirin in adjuvant aspirin therapy in CRC patients?

Materials and methods

This section deals with descriptions of methods of Study 1-3; however it should have been explicitly explained what were the aims of these studies and how do they relate to hypotheses presented in the previous section.

In addition, it should have been clearly stated to which study the description in chapters 4.1 - 4.5 is related to (i.e. Study 1, in agreement with description of methods of other studies).

Results

How would the author explain a lack of correlation between clinical parameters and MBNL genes expression, when significant correlations were detected between clinical parameters and FOXP1 and CD44 (section 5.4)?

Results: Study 2, Study 3

While the results for Study 1 are presented in detail, no figures and/or tables are shown for Study 2 and Study 3. At least some of the data could be presented in a different way than in the text (e.g. WB analyses). This is also true for discussion of the results. What is the reason for such discrepancy?

Discussion

The statement „...we report deregulated expression of the MBNL proteins... "(page 75) is not correct as no MBNL protein expression analysis (WB or other approach) is reported in the thesis.

Technical comments (tables, figures):

Figure 2: The title reads „...number of cases in 2020... ", while according to the original figure the data is presented for 2018.

Figure 3: It should be mentioned in the text, that this figure reports the EU data. In the figure, explanation of colors should be provided.

Figure 4: Quality should be better.

Figure 5 does not illustrate snRNP formation.

Figure 8 – 12: A different font should be used for x-axis description to make it better readable.

Figure 14 and 15 should be larger, including the font.

Table 3-5: Font should be larger.

Table 7: Some groups are very small, which can complicate statistical analysis and make it non-reliable, but this is acknowledged by the author.

Table 8: The numbers should be rounded to a maximum of three decimal points.

Conclusions:

Overall, the thesis is well written, investigates a very important topic of molecular markers in CRC and the results are sound. The author of the thesis proved her ability to conduct high quality research and present the data in journals with impact factor. Thus, I recommend the thesis for the defense.

Prague, 10th May 2023

Pavel Rössner, PhD.