

HABILITATION THESIS REVIEWER'S REPORT

Masaryk University

Applicant

MGR. EVA BUDINSKA, PH.D.

Habilitation thesis

Integrative Bioinformatics and Computational Modelling in Colorectal Cancer: Unveiling Tumor Heterogeneity through Multi-Omics Data

Reviewer

Prof. MUDr. Ondřej Cinek, Ph.D.

Reviewer's home unit, institution

Charles University in Prague,
Second Faculty of Medicine

Thank you for giving me the opportunity to review the thesis by E. Budinska. The author guides the reader through a series of exciting research projects and papers over the last 15 years, a period when her research focused on colorectal cancer (CRC). Of her rich publication list, 20 papers were selected that well document her long-term commitment to colorectal cancer research, and the rich and important contribution to the field. The thesis is as heterogeneous as the cancer it deals with: it is truly multifaceted, having included numerous numerical and laboratory methods, different cohorts, and various collaborative consortia. It well reflects the daily work of a bioinformatician, statistician, and analyst. Being a computational analyst is - in a way - a thankless job: the author illustrates this well by enumerating her experimental workload across numerous papers: her contribution ranges from 40% to 80%, while her position as an author typically remains a respectable second.

It is a guided tour through an exciting field which is also highly *clinically* relevant. The progress we witness in discerning subtypes helps - or will help - individualise the treatment, with possible positive effects on the overall survival. The identification and classification of hidden subtypes is the main thread winding through the whole text.

I would like to emphasise that I have no doubts that this thesis fulfils and exceeds the criteria imposed on habilitation theses in general, and I see it as a demonstration of a success story of a well-directed scientific career.

The text starts with a chapter on computational methodology. The author communicates her three tools very nicely to the lay reader, putting them into a wider context and describing their utility. The extent is commendably succinct, referring to the repositories for documentation.

The next chapter (3.2.) will draw us into an exciting adventure of supervised and unsupervised classification using data from the European PETACC-3 clinical study. We witness the excitement of the author who participated in finding novel subtypes of the CRC, and even more importantly, validating them in five independent datasets! A reviewer can only congratulate and thank for taking us on the tour through the results. Two minor notes here, one more serious than the other nitpicky one: (a) There is nothing like 'trend towards significance', trend is a statistical term, and here is no such trend. It is a mere *tendency*. If one wants to show something that approached the conventional $P=0.05$, then confidence intervals may be better than mere P values. (b) Figures would deserve titles without having

to refer to the caption. In journals it may be the usual formatting, but here it is utterly annoying to fish somewhere in the very tiny typeface below. For instance, Figures 7 and 9 later in the text are fine in this respect.

I was a bit lost in the short chapter 3.3., as the author does not clearly explain her position on the utility of different animal, xenograft and organoid models, nor provides any future visions, if there are any. Thus, as an outsider to the field, I could not fully appreciate the reason for selecting the two publications into the list of 20 that are presented in this thesis.

Chapter 3.4 deals with an exciting novel field of digital pathology - automated image analysis in an unprecedented depth, noticing parameters different from what a pathologist can systematically investigate. Its potential clinical use is indeed in focus, with a note of the need for prospective simplification of the predictors. The author not only presents the work by herself and colleagues, but also convincingly presents reasons for the integration of expression profiling data with something as 'old' as histopathology. The composition of the tumour as to the cell types, including non-tumour cells, and the heterogeneity across the tumour mass is nicely explained. And - congratulations on having the manuscript (referred to as Ref.#1) now published, with E. Budinska as the senior author.

Chapter 3.6, the last before Conclusions, has four pages, starting with gene expression markers predicting the response to a specific therapy in the PETACC-3 trial; it is not completely clear to me why this particular sub-chapter is not among the classification works in Chapter 3.2. Then a paper of microRNA in faecal material is briefly touched, and yet another work is dedicated three brief paragraphs. In my opinion, this chapter is a bit brief for its intended content, given that the micro RNA in faeces is a rather controversial topic - the laboratory processing is extremely far from being standardised, and the whole method would deserve in-depth discussion, not just three paragraphs. Other chapters provide the reader with much more background information, and also with richer insight into the opinions of the author on the importance of the findings.

As to the final chapter, Conclusions: It reiterates the milestones of the author's journey of CRC research. I admit I miss the overall perspective, even if this would be purely personal. Some overarching view of the 15 years with the CRC research, like - this was a success; that was a dead end and I do not want to see it anymore; this is my favourite method, field of research, this particular question is a topic for future work for the next decade.

The thesis has its limitations, too. First and foremost, the information in Chapter 3.5 *Tumor microenvironment and microbiome* is almost entirely outdated. Its references are mostly truly ancient works with little relevance nowadays. The enormous progress in the field is almost entirely ignored. Where does the text come from? The author herself has contributed to the ground-breaking change of paradigm in the CRC and microbiome! Her excellent Czech review on this topic (no. 15 of her own papers attached) is just referred to but not reflected in Chapter 3.5. The same is true for the recent excellent contribution, the seminal paper by the Segata group in *Nature Medicine*, needless to say that it has been co-authored by E. Budinska. The paper must have been an advanced draft at the time of writing this thesis. Frankly, even the classification into Tumor Microbial Subtypes is not very convincing, given that a validation group is lacking, the discrimination ability of V4 is inherently weak (and no, BLAST queries cannot help). Undefined abbreviations, statements like 'presence of *Incertae Sedis*', and missing reference 54 complete the unfavourable impression of this chapter - this is so much unlike the rest of the thesis. It may be also due to the notoriously superficial review process in the MDPI journal, in which the underlying study was published, otherwise the chapter would be of the high standard seen over the rest of the thesis.

Reviewer's questions for the habilitation thesis defence (number of questions up to the reviewer)

(1) What is your position on the future perspectives of various preclinical models in CRC?

(2) Now when the Segata paper (of which you are a co-author) has been published, do you see any clinical utility of determining the microbiome composition of stool, or biopsies, in CRC patients? If so, what samples and when should be such samples taken?

(3) Where is the place of image processing in the rapidly evolving field of histopathology?

(4) What is your research goal for the next decade?

Conclusion

The habilitation thesis entitled 'Integrative Bioinformatics and Computational Modelling in Colorectal Cancer: Unveiling Tumor Heterogeneity through Multi-Omics Data' by Eva Budinska **fulfils** requirements expected of a habilitation thesis in the field of Environment and Health (*Životní prostředí a zdraví*).

Date: 4/5/2026

Signature: