WESTFÄLISCHE
WILHELMS-UNIVERSITÄT
MÜNSTER

medizinische
fakultät
Westfälische
Wilhelms-Universität Münster

Institut für Bioinformatik, Niels-Stensen-Straße 14 | 48149 Münster

Wojciech Makalowski, Ph.D.
Professor and Head
Institute of Bioinformatics
Niels-Stensen-Straße 14
48149 Münster
Germany

Tel. +49 (0) 251 83-5 30 04

wojmak@uni-muenster.de
http://www.bioinformatics.uni-muenster.de

Rada Naukowa IChB PAN
Poznań, Poland

February 25, 2020

Evaluation of the dissertation *Historia biologiczna populacji ludzkich zamieszkujących obszar współczesnej Polski w pierwszych wiekach naszej ery – interdyscyplinarne badania archeogenomiczne* by Ireneusz Stolarek

Understanding our past has always been one of the major intellectual activities of the humankind. Since appearance of modern humans somewhere in Africa around 300,000 years ago, humans populated almost all corners of the Earth and different climatic zones. However, over time different populations mixed together and replaced one each other in any given location. Consequently, it is not straight forward to guess who exactly were people who lived in today's Poland several thousand or even hundred years ago. However, this became a little bit easier with the advent of ancient DNA sequencing technology, which led to a new discipline *archeogenomics.*

The core of the thesis consists of three papers published within last three years. Two papers were published in the journal *Scientific Reports* and one in *GigaScience.* Both journals have high reputation and are classified in the first quartile in their respective categories of scientific journals according to InCites Journal Citation Reports ranking (accessed on February 20, 2020). In all three cases samples came from archeological sites in Poland. First two publications are very similar and complementary. They describe, based on full mitochondrial genomes, genetic structure of two populations living in the South Baltic region, one in the current Western Poland and the other in Eastern Poland. Both publications are multi-disciplinary and multi-authored. In both cases, Mr. Stolarek's contribution

was significant based on notes in the papers: "I.S. conceived the study, participated in the study design, analyzed the data, created figures, discussed the results, and wrote the manuscript" and "I.S. participated in the study design, analyzed the data, created figures, discussed the results, and wrote the manuscript." Out of these different tusks, data analyses are, in my opinion, the most important contribution to both projects. A small digression here: while the thesis contains 26 pages of co-authors contribution to the papers included in the thesis, there's not a single paragraph describing PhD candidate's contribution to the papers, which obviously is more relevant to the presented thesis. Nevertheless, applied methodology is a state of the art and adequate to the presented research and enabled proper discussion and conclusions in the presented articles.

The third published paper, "Comprehensive analysis of microorganisms accompanying human archaeological remains," is slightly different but interestingly is based partially on the same archeological material. This time, the authors didn't analyzed human genetic material but microbiome of the human remains. In this case, Mr. Stolarek's contribution is less than in previous two papers but nevertheless important one. According to declaration in the paper he performed statistical analysis, including visualization of the results. By applying metagenomic analysis, the authors were able to demonstrate that most of detected bacteria were recent contaminations of samples due to different, not optimal, storage conditions. However, careful analysis allowed identification ancient species, including potential contamination.

The three papers in the thesis are preceded by an abstract in both Polish and English and extensive introduction, which strangely is called "OPIS WYNIKÓW PRACY DOKTORSKIEJ." Strangely, because out of 29 pages in this section only 3 are dedicated to presentation of the results. Nevertheless, the candidate quite nicely introduced the readers to the subject of his thesis. This is by no means an easy tasks as his thesis is based on multidisciplinary approach and many topics had to be introduced in order to understand and appreciate his work. Here we are introduced to archeogenomics, basic genetic terminology, ancient DNA, origins and early migrations of *Homo sapiens*, early history of modern humans in Europe, and very short history of archeogenomics in Poland. This part is complemented with goals of the presented work and mentioned before summary of the published papers.

I found this part of the dissertation very interesting. Unfortunately, my enthusiasm was hinder by surprising sloppiness in writing of this part. This is in strike contrast with very precise and clear writing demonstrated in published papers. Would the language be an issue? I presume not, most likely there was more intense correction and editing effort applied to the manuscripts submitted to the international journals than to the presented thesis. Here are few examples.

1) References are given in two different styles, first few pages employ numbered style and later "name and date" style is used. Unfortunately, because of that the latter references are not listed in the "Bibliography" section. Moreover, a significant part of this part lack of references all together. This applies to the following sections: AMH w Europie, Neolit – powstanie i rozwój rolnictwa, and Późny Neolit i Epoka Brązu. My guess is that these parts were written based on a textbook, which is fine but it should be referred to.

2) Figure 2 has different styles of numbers on its left and right site.

3) I'm not a big fun of mixing languages in a single document, therefore I was surprised to see that in Polish part of the thesis suddenly an English acronym (yta) was used. I think it brings unnecessary dissonance to the text.

4) Some figures, including Fig. 2 are difficult to read. The font applied is too small.

5) In *Archeogenomika w Polsce* section the "oldest" DNA from Poland is mentioned. It would be interesting to know how old is this DNA.

6) In the summary of results we learn that DNA was sequenced using NGS method. However, we don't learn here what NGS method was employed. As a matter of fact it would be nice to have a short introduction to NGS and how it change archeological studies.

7) In general, it is quite difficult to navigate this dissertation because of lack of page numbering. Although, there is a Table of Content, since pages throughout the thesis are not numbered, it is useless.

Finally, I have a couple of questions, points for discussion that I hope we can converse about during the defense. In the section "Kopalny DNA" we can read "Pierwsze tego typu badania przeprowadzono w roku 1988. Analizie poddano DNA wyizolowany ze szczątków ludzkich datowanych na około 7 tya i pochodzących z Florydy. Stwierdzono, że zidentyfikowany haplotyp mtDNA nie występuje współcześnie w puli genowej rdzennych Amerykanów, jest natomiast obecny u współczesnych Europejczyków, choć stosunkowo rzadko." I would like to know candidate's opinion why in pre-columbian sample there was an European not native American haplotype present? Another topic I would to discuss is taxonomical rank of Neanderthals and Denisovans. Mr. Stolarek considers them as a subspecies of *Homo sapiens,* however it is quite common in the literature, including papers reporting original discoveries, to treat them as separate species.

In summary, despite some shortcomings, we are presented with the very good thesis, which qualifies for a PhD degree fulfillment.

Sincerely,

Wojciech Makałowski

Prof. Dr. Wojciech Makalowski

Professor and Head