EVALUATION REPORT

Of PhD dissertation entitled:

„DIVERSITY AND BIOTECHNOLOGICAL POTENTIAL OF THERMOPHILIC MICROORGANISMS FROM BULGARIAN HOT SPRINGS“

in regard to a competition for acquiring **educational and science degree "Doctor"**

of full-time PhD student **Nikoleta BOTEVA,**

with scientific supervisor: **Prof. Dr. Margarita Kamburova, DSc**

in the scientific field 4.3. ***Biological sciences (****speciality* ***Microbiology)***

REVIEWER: Assoc. Prof. Dr. Svetla Trifonova Danova, DSc

Department of General Microbiology,

Institute of Microbiology "The *Stefan Angelov*" - Bulgarian Academy of Science

FOR THE PROCEDURE:

On the basis of the decision of the Scientific Council of Institute of Microbiology – BAS (*Protocol № 17 /22.04.2021*) and Order № I-41/23.04.2021 issued by the Director of Institute of Microbiology, I have been selected as a member of the Scientific Jury for the above competition. The competition is conducted following the procedure under the ADASRB in force in 2016. I was appointed as a reviewer at the first meeting of the Scientific Jury held on 11.05.2021.

I declare that there is no conflict of interest within the meaning of para. 1 items 3 and 5 of the ADASRB with the evaluated candidate in the competition for acquiring ESD "Doctor".

CANDIDATE INFORMATION:

I have known PhD student Nikoleta Boteva personally since she was a student at UCTM - Master's Program „Biotechnology“, from the course of Probiotics. She is a graduate biotechnologist with a strong research interest in molecular biology of microorganisms from different habitatas. She is fluent in both written and spoken English. She has specialized „*Anaerobic cultivation of thermophilic microorganisms and metagenomic analyzes*“ at the University of Bergen, Department of Biology, Norway under the guidance of Prof. Nils Kåre Birkeland and „*Chromatographic methods for lipids and fatty acids analysis*“ under the guidance of Dr. Anarita Poli and Eduardo Pagnota at the Institute of Biomolecular Chemistry (ICB) (C.N.R.) Italy. She has curiosity and desire to work on the latest scientific problems and she is always looking for a scientific challenge.

DISSERTATION EVALUATION:

1. Relevance and significance of the elaborated problem

The current dissertation submitted to me for evaluation seeks answers to important and relevant scientific problems related to poorly studied thermal ecological niches and their microbiota. Bulgaria is among the countries especially rich in thermal springs whose biodiversity is poorly examined. The study of thermophiles biodiversity is in alliance with their increased ecological and biotechnological interest. The topic is important not only nationally but also internationally especially in seeking new biocatalysts and microbial enzymes. Related explorations enlighten questions of fundamental importance concerning the activity of biological molecules and biocenoses in extreme habitats. Taking into account the limitations of culture-dependent methods for biodiversity characterization, the PhD student has chosen modern metagenomic and molecular genetic analyzes. All this gives me a reason to evaluate the dissertation as relevant and significant, developed in one of Europe's leading laboratories for extremophilic microorganisms at the Institute of Microbiology, BAS, under the supervision of Prof. M. Kamburova.

2. Dissertation volume and structure:

The dissertation comprises 171 standard pages of double-sided text including the dissertation reference list and participation in scientific forums. The generally accepted scheme and recommended dissertation chapters proportion is met, as follows: *Introduction* - 1 page, *Literature review* - 27 pages, *Goals and Tasks* - 1 page, *Methodology* - 14 pages, *Results and Discussion* - 95 pages, *Summary* - 1 page, *Conclusions* - 2 pages, *Contributions* - 2 pages, *References* - 20 pages.

The concise scientific style used in writing the dissertation makes an excellent impression as well as the illustrative material including - 58 coloured figures and 19 tables applied to illustrate the results related to phylogenetic diversity and biotechnological significance of the extremophilic microbiota of Bulgarian thermal springs.

Chapter *Results and Discussion* is in agreement with the assigned tasks and contains six main sections: Metagenomic analysis of microorganisms in samples from three Bulgarian hyperthermal springs and metagenome-assembled genomes reconstruction; Identification of open reading frames containing a significant portion of unknown proteins; Metagenomic analysis of anaerobic cultures from two hot springs enriched with substrates of interest; Cloning, expression and characterization of a thermophilic enzyme (lipase) from unculturable microorganism; Isolation and characterization of extreme thermophilic anaerobes; Synthesis,

purification and characterization of thermostable pectinase from a newly isolated microorganism. The discussion of the obtained results is conducted in a comparative aspect with the results of other authors working in the same field with skilful highlighting the Bulgarian contribution among these studies.

3. Literature review and research goals and tasks framework

The Literature review is written in an objective and targeted manner showing the information from the last years. Extremophilic and thermophilic microorganisms are considered consistently – the published data reporting for their biodiversity, metabolic strategies and the molecular basis of thermophilicity and thermostability. The critical analysis and the way the new methods and approaches are presented make an excellent impression. Nikoleta Boteva has selected and incorporated the most important elements from the information relevant to the metagenomic analysis – taxonomic and functional profiling of microbial communities and metagenomic approaches for new thermophilic biocatalysts discovery concluding with the biotechnological significance of the thermophiles. I suggest including the achievements of Bulgarian research groups working on extremophilic microorganisms to underline the need of studying the unexplored Bulgarian hot springs. Despite this recommendation, the scope of the literature review indicates a deep knowledge and understanding of the researched scientific problems. The literature review contains a short analysis of the phylogenesis of mesophiles and thermophiles in the light of the newest theories of life origin and evolution. Thus the need for exploring biodiversity using the metagenomic approach is addressed. This is only a part of an ambitious goal: "*to elucidate the phylogenetic biodiversity and biotechnological potential of thermophilic microorganisms in Bulgarian hot springs*". The goal and the 8 tasks are very clearly defined and illustrated with an original scheme showing the adopted approaches for achieving the purpose. To accomplish the goal 8 tasks are provided:

1. Taxonomic and functional metagenomic analysis of DNA purified from Bulgarian hot springs sediments and water to characterize the microbial communities.

2. Metagenome-assembled genomes reconstruction and elucidation of the thermophilic participants in the nitrogen cycle.

3. Metagenomic analysis of anaerobic extremophilic enrichment cultures.

4. Cloning, expression, purification and characterization of a thermophilic lipase from a reconstructed genome of an unculturable thermophile.

5. Isolation of extremely thermophilic aerobes.

6. Isolation of extreme and/or hyperthermophilic anaerobes.

7. Screening for carbohydrate-degrading enzyme producers.

8. Purification and characterization of a thermophilic enzyme with applied potential from a culturable thermophile.

I have to point that many of the tasks are so comprehensive and complex that they could be considered separately as distinguishing research sufficient for a PhD thesis. The research scope is a scientific challenge even for a big research group considering the financial time frame of 3 years. Even though the research is not restricted to one object. There are collected and compared sediment and water samples from Rupi basin (the basin forms three spatially distant hot springs), from the spring located near Vlasa (86°C), central Bulgaria, the hot spring Mizinka in Velingrad (88°C), central Bulgaria, the spring near Levunovo village (83°C), southwest Bulgaria.

The PhD student does not restrict the research to characterisation the inhabitants of geothermal springs but goes further and perform cultivation and identification of aerobic and anaerobic thermophiles, and cloning and expression of lipase derived from the reconstructed genome of uncultivated thermophilic microorganism comparing the enzyme with ones isolated from culturable bacterias of the same group.

4. Evaluation of the applied methods:

The ambitious research program is accomplished with a vast array of methods including culturable microbiological, molecular-genetic and recent bioinformatic tools. The methods are systemically presented in chapter *Materials* covering entirely the work stages. I would recommend for future research the methods to be thematically grouped - microbiological, genetic, bioinformatics. This notice certainly does not undervalue the complexity and complication of the work. The inclusion of metagenomic analysis of shotgun sequencing data obtained with MiSeq Illumina paired ends platform is relevant as well as the processing of the sequences: sequences quality assurance (FastQC), trimming, adapter removing (Trimmomatic), de novo assembly (CLC Genomics Workbench). The metagenomic DNA extraction is not an easy task itself taking into account the low microbial number in hot spring sediment and water samples. This proves an excellent theoretical and experimental preparation and skills combining microbiological methods, molecular genetics and bioinformatics programs, which are essential for modern polyphase taxonomy. The selection and application of such a diverse panel of methodical approaches demonstrate that the PhD student is a developed molecular biologist and microbiologist.

Nikoleta Boteva has learned and applied the Hungate technique for anaerobic cultivation, which is a new method for the department. Therefore she managed to cope with and overcome difficulties and to adapt the methods in a way enabling fulfilling a particular experimental task.

This is seen and from the selected methodical approaches applied to characterize important biologically active enzymes, especially for cloning and expression of a thermophilic lipase from a reconstructed genome of an uncultivated thermophile. I acknowledge the fact that she applied in Bulgaria what she has been learned during two specializations abroad - in Norway and Italy.

5. Evaluation of the obtained results:

Chapter "*Results and Discussion"* is the largest and follows the experimental tasks describing the obtained results which include: (1) Metagenomic analysis of the microbiota in three Bulgarian hyperthermal springs - sediment samples from Rupi II and Vlasa and a water sample from Rupi I; (2) along with draft genomes reconstruction; (3) Identification of open reading frames containing a significant amount of unknown proteins; (4) Metagenomic analysis of anaerobic cultures from two springs enriched with desired substrates; (5) selection of a gene encoding thermophilic lipase from a reconstructed genome of unculturable thermophile which is (6) successfully cloned, expressed, purified and characterized; (7) Isolation and characterisation of extreme-thermophilic anaerobes (8) Synthesis, purification and characterization of thermostable pectinase from a newly isolated anaerobic microorganism from genus *Caldicellulosiruptor*.

The entire research is planed well. Already in the first stage of the research after analysing the differences in the hot springs chemical composition the PhD student has chosen to analyze sediment samples from Rupi II and Vlasa and a water sample from Rupi I. The adopted strategy of metagenomic shotgun sequencing and not metagenomic 16S rRNA amplicon intends to obtain information about both the taxonomic structure and functioning of the thermophilic communities and their biotechnological potential. A good impression is derived from the taxonomic structure characterisation based at first on the raw reads analyzed with MG Rast using the RefSeq database. The microbial diversity in three geothermal springs is annotated on a phylum and genus level distinguishing common and unique microbial genera with relative abundance above 1%. Nikoleta Boteva has analyzed the biodiversity and concludes that despite the low organic content in hot springs the identification of two different dominating metabolic types (the lithotrophic *Aquifex*, *Hydrogenobacter* and the heterotrophic *Thermus*) demonstrate the metabolic pool of organotrophic synthesis by the lithotrophs enabling the growth of the heterotrophs. For the first time is obtained information about dominating genera of main microbial phyla and is shown a low representation of archaea similar to thermal habitats in India. The successful analysis of the metagenomic sequences and handling with different

databases for sequences annotation is the basis for identifying different species of the dominant genera. For the first time is reported (although minimally presented) the presence of photosynthetic bacteria. Photosynthetic processes are not reported at temperatures above 75°C. The data concerning the archaeal taxonomic profiles and the thermophilic viruses are new and valuable which is a prerequisite for further analysis. With these and other detailed analysis, the PhD student presents a broad and scientifically sustained picture of great biodiversity, a high level of novelty and applied potential, which I appreciate. The same is valid for the metabolic reconstruction of the shotgun sequences performed with MG Rast and the KEGG database. Furthermore, the functional analysis was performed on the assembled metagenomes aiming the reported functional characteristic to refer to complete ORFs which could be cloned. This also proves that the analysis and the tasks follow a logical order. Through the application of the Anvi'o program, Nikoleta Boteva has reconstructed numerous genomes from uncultivated thermophiles from Rupi I, Rupi II and Vlasa. The higher number of draft genomes are recovered from Rupi II - 18 comprising mostly uncultivated bacteria and 4 genomes are referred to as uncultivated archaea.

The PhD student's work provides very useful information about the poorly studied prokaryotic component in these extreme ecological niches driving the biogeochemical cycle of nitrogen, sulfur or carbon. This is illustrated with original schemes as figures 34 and 36 (concerning the nitrogen and carbon cycle respectively). Different metabolic types are revealed comprising heterotrophs and chemolithotrophs, including methanogenic, acetogenic, hydrogenotrophic, nitrifying (including ammonia oxidizers and nitrite oxidizers), sulfur-oxidizing and oxidising aromatic compounds. The prokaryotic component driving the biogeochemical cycle of nitrogen, sulfur or carbon is still unclear. The presented work revealed genome-based characteristics defining the physiological advantages of thermophiles and archaeal representatives, allowing their abundance and ecological success, as well as in such extreme habitats. As though, to verify the reliability of obtained sequences information the PhD student complements the work program with a *Metagenomic analysis of anaerobic cultures* enriched with xylan, from hot springs Levunovo and Mizinka.

The selected approach for microbial community enrichment using media with defined substrate aims to influence the thermophiles composition to isolate promising microbial producer. Thus, unlike the current understanding, is revealed higher biodiversity in the culture derived from the hotter spring Mizinka including bacterial and archaeal representatives. The genetic determinants of the revealed dominating glycosyl transferase families GT4 (37 ORFs from Levunovo and 39 ORFs from Mizinka) and GT2 (22 ORFs from Levunovo and 24 ORFs

from Mizinka) are characterized. These families combine various enzymes involved in the biosynthesis of sucrose, digalactosyl, diacylglycerol, trehalose, cellulose, chitin, hyaluronan and others. From the family of glycosyl hydrolases are revealed representatives of GH13, GH4, GH3, GH57, GH3, GH2 and unknown ones. The obtained results define the next work stage – analysing the draft genome of uncultivated thermophile *Candidatus* Calescibacterium (R2.1-Calescibacterium) reconstructed from Rupi II metagenome. Among the inspected genes with applied potential, the PhD student selects and successfully clone, express and characterize a thermophilic lipase (LipR) from this uncultivated bacteria. At this stage, she skillfully combines genetic analysis with biochemical and molecular ones. She proves the phylogenetic affiliation of this enzyme to alpha/beta fold hydrolase and besides the characterization of the heterologously expressed enzyme Nikoleta Boteva concludes that the extremophilic lifestyle has directed the evolution of the proteins to some common, structurally adaptive strategies that provide resistance to denaturing agents.

The last part of the experimental work determines the biotechnological potential of cultivated aerobic and anaerobic thermophiles. The PhD student has isolated and identified (based on 16S rRNA gene analysis) 17 strains from genus *Geobacillus* and 5 anaerobic strains belonging to genera *Caldicellulosiruptor*, *Thermotoga*, *Thermovorax*. Some of the isolated strains have 16S rRNA gene identity to the closest neighbour below the threshold for species affiliation. Here I would like to ask the PhD student:

Does she think there are new taxa at the studied Bulgarian hot springs among those described and identified by herself? And if so, when this will be published?

6. Evaluation of the conclusions, contributions and scientometric indicators according to the requirements of the ADASRB

The presented dissertation summarizes a huge volume and thematic diversity of research methods and diverse results from various genomic, biochemical, taxonomic and microbiological studies of unexplored on this scale Bulgarian thermal springs and their microbiota. Despite this fact, the PhD student has successfully summarized the work in 12 clearly and precisely formulated conclusions. They are a consequence of the results obtained from the performed 8 experimental tasks. In the summary, PhD student describes the perspectives and possibilities for further projects and research developments.

There are undoubtedly many original contributions in the work, both scientific and methodological. The contributions are objectively divided into scientific and scientific-applied and the original ones are clearly visible among them. Some of the protocols of the experiments

with success can serve as new algorithms for characterizing microbial biodiversity and biotechnological potential not only of the microbiota from extreme habitats.

However, only a small part of these new and interesting achievements have found a place in publications related to the dissertation. These are two articles in periodical scientific journals with Impact Factor and one chapter of a book in an international edition (Springer), which fully fulfils the requirements of the Act. I would like to recommend the other results to be published after the defence because they are extremely interesting and are a prerequisite for characterizing these habitats.

CONCLUSION

In conclusion, I want to emphasize that the material is sufficient for acquiring ESD "Doctor". The topic is relevant, the PhD student has mastered modern methods, the experiments are set methodically correct, the results are reliable and are a solid basis for further scientific and applied projects.

At the risk of repeating myself, I will emphasize that the work far exceeds what is needed to obtain an ESD "Doctor" and proves high professional training and scientific knowledge and skills in the field of modern microbiology of the candidate. The experience gained in the course of work in terms of cultivation and isolation of anaerobic thermophiles is a prerequisite for initiating projects searching for new anaerobic producers of enzymes with applied potential. As the doctoral student assessed by herself, the data obtained from the metagenomic analysis can serve as a basis for creating more selective laboratory conditions searching the access to unculturable microorganisms.

Based on the above, I can confidently say that the peer-reviewed dissertation is an original scientific work, with theoretical and applied significance. It meets and exceeds all conditions of the Act for the Development of the Academic Staff in the Republic of Bulgaria.

All this gives me a reason for an excellent evaluation of the dissertation, and I propose to the respected scientific jury to award the **educational and scientific degree "DOCTOR" to NIKOLETA BOTEVA**, in **Professional field 4.3. Biological sciences** (speciality Microbiology).

Sofia, 13.06.2021 Prof. Svetla Danova, DSc.