COBISS: 1.22

INTERVIEW WITH ROK KOTANJŠEK AND NINA GUNDE-CIMERMAN

WELCOME TO THE -OMICS ERA OF THE 21ST CENTURY: WILL *PROTEUS ANGUINUS* FINALLY REVEAL ALL ITS MYSTERIES?

cunducted by Janez MULEC

There are few milestones in speleobiology, and *Proteus anguinus*, commonly named the olm, proteus or "human fish", is undoubtedly part of its (hi)story. The animal was first mentioned in 1689 by Johann Weikhard von Valvasor (Janez Vajkard Valvasor) in his *The Glory of the Duchy of Carniola* (*Slava Vojvodine Kranjske*). The scientific name *Proteus anguinus* was given by Josephus Nicolaus Laurenti in 1768. The animal gained the interest of the scientific community later in the 19th century, culminating in 1859 in Charles Darwin's famous monograph *On the Origin of Species* as an example of evolutionary reduction of body structures through disuse. The study of this remarkable animal intensified in the 20th century. Its longevity, its general peculiarity and the mystery governing the underground world of the Dinaric karst triggered more systematic research. Much interesting data was obtained from study of proteus in captivity in various laboratories worldwide, including Slovenia, France, England, Belgium, Hungary, Germany and Italy.

In the 21st century, the year 2019 represents another important milestone in the research of proteus. The draft genome of *Proteus anguinus* was first publicly presented on 25 November 2019 in Ljubljana, Slovenia. The event coincided with 330 years since proteus was introduced to the scientific community by Valvasor, 160 years since its popularization by Darwin, and the 100th anniversary of Ljubljana University. This is undoubtedly an achievement with many superlatives, starting from excellent collaboration among different international laboratories and the very short period – one year – used to finish the sequencing of the huge proteus genome, which is 15-times larger in size than the human genome.

On this occasion we spoke with the two most important players in the story, prof. Nina Gunde-Cimerman and prof. Rok Kostanjšek, from the University of Ljubljana, Biotechnical Faculty. Prof. Nina Gunde-Cimerman directs research of extremophiles adapted to extreme conditions, in particular fungi, at the Biotechnical Faculty. Prof. Rok Kostanjšek leads a research team that continues the investigation of proteus begun by his predecessors, prof. Lili Istenič and prof. Boris Bulog.

- *Sixty years of continuous research at the Biotechnical Faculty have offered us many interesting findings. And truly, this relatively long period has also revealed many gaps in our knowledge about proteus, its ecology, evolution and physiology. If you could name the five most urgent scientific questions related to proteus, what would they be?*

Rok Kostanjšek: From the perspective of proteus conservation, the most urgent questions would relate to the exact number, size and structure of proteus populations. The answers to these questions are crucial for assessing the vulnerability of the populations and setting conservation priorities for proteus.

Despite numerous known localities for proteus, these only represent cave entrances, springs and accessible parts of underground water systems, leaving our knowledge of the actual geographical distribution of proteus far from complete. The exact range of populations and aquifers that supply water to proteus habitats is another open question that is essential for targeted conservation measures for this charismatic amphibian.

If we focus on proteus biology, one of the main questions would be the nature and origin of the mechanisms behind its adaptations to the underground environment

DOI: 10.3986/ac.v49i1.8009

which, despite independent transitions of proteus ancestors from the surface to the caves, led to the same set of troglomorphic features (such as loss of vision, depigmentation, elongation of body parts and development of specialized senses).

Since proteus males and females are generally indistinguishable by external morphology, sex determination by non-destructive methods is another open question in proteus biology, the clarification of which is essential for understanding proteus' population genetics and establishing a sustainable breeding program. Then there are open questions related to mechanisms behind its ability to survive extreme starvation, tolerance to anoxic conditions and extraordinary ability to regenerate, which might improve our knowledge of molecular mechanisms applicable in human medicine. Last but not least, there are also open questions of evolutionary mechanisms leading to extreme genome size in proteus and its relatives, which may provide answers to the broader biological question of non-coding parts of genomes in general.

Proteus lives in the karst underground of the Dinaric Alps from Italy on the north to Montenegro on the south. Waters in karst are generally very vulnerable to human intervention, particularly if they face pollution. It is known that proteus tolerates lowered oxygen concentrations in water relatively well in the short term. In the long term, many heavy metals are bio-accumulated in its various tissues. But how does proteus respond to microbial attack from human-mediated pollution or intervention?

Nina Gunde-Cimerman: Proteus is an amphibian, and amphibians, in general, are globally endangered. According to the global assessment, 43% of amphibian species are in decline, while an additional 32% are threatened. The most important factors contributing to the decline are eradication, fragmentation and pollution of their natural habitats, climate change, and perturbation of natural infection cycles with microbial pathogens. Proteus has thin, highly vascularized skin and lacks protective pigments, which makes it highly sensitive to pollution and infectious microbes. Proteus is thus currently classified as "vulnerable" by the International Union for Conservation of Nature (IUCN). Because of the inaccessibility of its natural habitat, scientific studies on proteus microbial pathogens depend mainly on animals kept in captivity under controlled conditions simulating the natural environment. However, suboptimal artificial conditions in captivity can generate stress and decrease the immune response; therefore these studies can't be completely correlated to proteus conditions in the natural environment. So far it was reported that proteus in captivity acquired infections with amoebae, bacteria *Aeromonas hydrophila*, oomycete water molds *Saprolegnia* sp. and black yeast *Exophiala* sp. Water was the most probable source of infection and weakened immune status the reason.

Currently, the most significant and well-described pathogens of amphibians in nature are the chytrid fungi *Batrachochytrium dendrobatidis* (Bd) and *B. salamandrivorans* (Bsal) that cause damage of the skin and disrupt vital functions. These fungi have been causing amphibian decline globally over the last decade. Until we started a three-year, still-ongoing study on the presence of these fungi and other pathogenic microbes, including ranaviruses, in the natural habitat of proteus, it was not known if these infections had spread as well in this geographic area. So far we have not yet confirmed their presence in Slovenia, although we did a thorough sampling of the waters in the natural habitat of proteus. This is definitely good news; however the situation may rapidly change.

We can read in many journals, from popular to scientific, that numerous species are threatened with extinction today. Based on IUCN's (International Union for Conservation of Nature) estimation, this means more than 30,000 species, with 43% of amphibians as the most threatened taxonomic group. What we can say about proteus in this respect; how vital is its population in Slovenia and the wider region?

Rok Kostanjšek: In view of the constant drainage of surface pollutants into underground waters, their accumulation in underground food chains and the inability of proteus to retreat to surrogate habitats, all populations deserve considerable conservation attention. At the same time, the inaccessibility of underground habitats and lack of information on the size and structure of proteus populations make it very difficult to assess their vitality and vulnerability. Nevertheless, due to the small genetic pool, the lowest vitality is to be expected in the smallest and most highly localized populations, such as the black subspecies in Bela krajina or the population in Vir pri Stični, both of which should therefore be considered a conservation priority.

Any deeper ecological study cannot be imagined nowadays without extensive molecular data. Scientists are not satisfied only with analysis of short segments of DNA. Whole genome sequencing is an increasingly *adopted molecular technique, but few higher animals have been "lucky" enough to be sequenced so far. Not many laboratories and facilities can provide such a service. How did you organize the consortium for sequencing the "genome of human fish"?*

Nina Gunde-Cimerman: Many things in life happen by coincidence. In October 2017 I was invited to give a talk at the 12th Annual Meeting of the International Conference on Genomics (ICG-12), Shenzhen, China. The topic of my talk was related to the focus of my research group at the Biology Department, Biotechnical Faculty, University of Ljubljana, that is, extremophilic fungi that not only populate different extreme environments but also represent a health threat due to their opportunistic pathogenic potential. The theme and the talk aroused quite a lot of interest and consequently I was invited by BGI, Shenzen, to discuss a possible collaboration. I was very thankful for this opportunity, because BGI Shenzen represents one of the leading global centres for genome sequencing, known for their professionalism, speed and interest in inspiring projects. Not only do they have up to date technology, they also develop it themselves, resulting recently in the DNBSEQ-T7 Genetic Sequencer, that can extremely quickly sequence long DNA molecules and has an incredible capacity for generating big data. For example they have sequenced in the past some quite iconic plants and animals, i.e. millet, rice, the giant panda, 40 types of silkworm, the SARS virus, even a 4,000-year-old man, the Alaskan Malamute, etc. Although my initial intention was to discuss the possibility of sequencing the extremophilic fungi, which became a reality and on which we are now collaborating, I decided, on the spur of the moment, to present to BGI the most iconic Slovenian animal, proteus. Since proteus has a huge genome— at the time of my presenting the idea, it was estimated to be 15 times bigger than the human genome— this daring idea required quite a bit of convincing and clarification. To make a long story short, with the help of colleagues from Ljubljana, Slovenia, and Aarhus University, Denmark, we succeeded and embarked on this prestigious project together.

- *The genome of proteus is 15-times larger than the human genome and is one of the biggest sequenced animal genomes so far. It is hard to imagine such a huge amount of data; even more what kind of computing power must be provided to process it? Being now in the initial phase of bioinformatics analysis, can you share with us any interesting preliminary findings on proteus' genome or its transcriptome, which gives vital information about genes which are "turned on" at certain times and in certain tissues?*

Nina Gunde-Cimerman: The challenges of sequencing what is at present the largest known animal genome were huge; however, they have recently been overcome by the combined efforts of partners from Slovenia (University of Ljubljana), Denmark (Aarhus University) and China ("BGI Research", "MGI Tech" and "Lars Bolund Institute of Regenerative Medicine"). The combined expertise and technological capabilities led to the first successful reading of data. Multiple readings were needed, generating over 7 Tbp of information, which was accompanied by multiple readings of transcriptomic data from fourteen different proteus tissues. Initially we have to perform bioinformatic analysis just to organize this vast amount of sequencing data into an assembled genome; only then can we address several highly important biological topics. Initially the focus will be on description of the genome structure, annotation and assembly of the genomic data and analysis of the non-coding part of the genome. We will try to illuminate the processes that generated its enormous size. Other important phases after this initial and most needed one will be identification of the mechanisms behind the loss of pigmentation in proteus and identification of the potential mechanisms associated with the longevity of proteus. These and other results that we hope for will provide essential information needed for focused conservation efforts on proteus, resulting in establishment of a sustainable breeding program. Additionally, they might provide new clues about proteus longevity and other biomedically relevant information.

- *There are probably many other questions you hope to answer from this project, including for conservation biology. For example, we know that is hard to determine sex of these animals, but once you know this information you could mate animals in a "controlled manner" …*

Rok Kostanjšek: Since the sex of the proteus is generally not distinguishable by external morphology or even by cytogenetic analysis, the identification of appropriate sex-related genetic markers in the proteus genome is considered one of the priorities of the "Proteus Genome Project". Unfortunately, sex determination is not the only obstacle in the development of a controlled breeding program. Other difficulties are the late sexual maturity of the females, which occurs at the earliest at the age of twelve to fifteen years, the long reproductive cycles of five to seven years needed to provide the appropriate breeding conditions, frequent developmental abnormalities and egg infections. Despite several successful breeding attempts and its immense potential for research and conservation of the species, the establishment of sustainable breeding of proteus remains challenging, to say the least.

All of the topics we have discussed till now were related to white proteus, but in the SE part of Slovenia, Bela krajina, there also exists a "black version" of proteus (Proteus anguinus parkelj). Its discovery in 1986 came as a big surprise, but it seems that genetically it is not very distant from its white counterpart. Is it possible that the very large genome could be the reason for this "phenotypic" difference and that proteus hides deep in its genome more than we can see at present?

Rok Kostanjšek: It is more likely that any potential surprise hidden in the proteus genome can be ascribed to over 100 million years of independent evolution than to the genome size itself. The latter also most likely does not "hold the key" to the morphological differences between black and white subspecies. The black subspecies entered the caves no later than some lineages of the white subspecies. At the same time, the development of pigmentation and eyes begins normally in the larvae of both subspecies. It is therefore more likely that the morphological differences between the subspecies are due to differences in the mechanisms controlling the course of development of troglomorphic features.

Proteus is an excellent indicator of environmental pollution and a top predator in the karst underground food chain, but is also a very important cultural symbol for Slovenia and the Classical Karst. Its mystery is slowly being unravelled. What is next on your agenda?

Nina Gunde-Cimerman: At present we are quite overwhelmed with tasks at hand. We have to dig deep into the mysteries hidden in the enormous proteus genome and transcriptomes, and at the same time obtain more insights into potential threats it encounters in nature. The world is changing in front of our eyes, conditions for many animals (and humans) are becoming more and more extreme and are threatening the existence of entire biomes. One such potentially threatened biome is karstic waters, not only important as a habitat for proteus but also as a supply of Slovenian drinking water. Another threatened biome is polar areas, exposed to rapid melting and changes that we have now only started to fully acknowledge. Triglav glacier has almost disappeared, and maybe we should do something fast before it is just a distant remembrance.

Janez Mulec