**University of Leicester**

**Future 50 PhD Scholarship**

|  |  |
| --- | --- |
| **Project Reference** | CMS Morozov |

|  |  |  |  |
| --- | --- | --- | --- |
| **First Supervisor** | Dr Andrew Morozov | | |
| **School/Department** | Computing and Mathematical Sciences, | | |
| **Email** | [am379@leicester.ac.uk](mailto:am379@leicester.ac.uk) | **Telephone Ext** | 5054 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Second Supervisor** | Professor Chris Bayliss | | |
| **School/Department** | Department of Genetics and Genome Biology, CLS | | |
| **Email** | cdb12@leicester.ac.uk | **Telephone Ext** | 3465 |

|  |  |
| --- | --- |
| **Additional Supervisor** |  |

**Section 2 – *Project Information***

|  |  |  |
| --- | --- | --- |
| **Project Title** | **Mathematical modelling of CRISPR-Cas systems using big data** | |
| **Project Highlights:** | 1. | Mathematical modelling of CRISPR-Cas systems will be backed up by extensive data analysis of large sets of large size genomic data (*Data Analysis and Computational Modelling*). |
| 2. | We will computationally address the fundamental scientific hypothesis that the CRISPR-Cas system impedes control of bacteria by their natural enemies – phages (*Life Science Interface*). |
| 3. | Results of mathematical and computational modelling will contribute to improve the efficiency of phage therapy in medicine, agriculture, and fighting pathogens in various environments (*Environment*). |
| **Project Summary** | | |
| This project is *at interface between mathematics, data analysis and life sciences*. Bacteriophages/phages are bacterial viruses, they regulate all aspects of evolution and dynamics of prokaryotes, including highly pathogenic bacteria. Defence mechanisms of prokaryotes can largely reduce their control by phages, and CRISPR-Cas system is considered to be a major mechanism of adaptive immunity. The project will provide the basis for *the mathematical and computational modelling of functioning CRISPR-Cas systems*. The CRISPR-Cas system learns from previous phage infections and integrates small pieces from phage genomes (spacers) into the microbial genome; the resulting library of spacers in CRISPR arrays is then compared with the DNA of potential invaders. The PhD project will include both modelling and bioinformatics (big data) to explore the statistical distribution of spacers in bacterial genomes as well as the connection between the statistical distribution of spacers and efficiency of CRISPR-Cas systems to resist natural enemies. CRISPR-Cas9 systems have recently received much attention: the *Nobel Prize in Chemistry in 2020* was awarded for the application of CRISPR-Cas9 in gene editing. Surprisingly, the role CRISPR in regulation of phage control of bacteria in natural and artificial environment across different time and space scales in systems with strong non-linear feedbacks is still largely unclear. Mathematical modelling backed up by analysis of large genomic data sets should highly contribute to understanding the above questions. Our recent mathematical model shows that statistical distribution of spacers in microbial genomes generally follows a power law distribution and this explains the rareness of all-resistant super microbes. However, this model largely ignores the diversity of CRISPR spacers, dynamical feedbacks from the environment, and evolutionary aspects of the problem. Using advanced mathematical models of CRISPR and validating those models from updated genomic data would allow us to computationally test the central hypothesis on the predominant role of CRISPR as an important obstacle of the phage control of pathogens. Our final goal is to use the results of mathematical modelling to improve the efficiency of phage therapy in medicine, agriculture and the environment. **The PhD student will receive training in data analytics (big data), modelling and genetics.** | | |

**References**

[1] Pavlova, Y.S., Paez-Espino, D., **Morozov, AYu**, Belalov, IS. 2021 Searching for fat tails in CRISPR-Cas systems: Data analysis and mathematical modeling. PLoS Computational Biology 17 (3), e1008841

[2] Sandhu SK, Bayliss CD, **Morozov AYu**. How does feedback from phage infections influence the evolution of phase variation in Campylobacter?, 2021. PLoS Computational Biology. 2021 17(6): e1009067.

[3] Egilmez, H.I., **Morozov, A.Yu.** Galyov, E.E., 2021. Modelling the spatiotemporal complexity of interactions between pathogenic bacteria and a phage with a temperature-dependent life cycle switch. Scientific Reports, 11(1), 1-13.

[4] Egilmez, H.I., **Morozov**, **A**. **Yu**., Clokie, M.R.J., Shan, J., Letarov, A., **Galyov E.E**., 2018. Temperature-dependent virus lifecycle choices may reveal and predict facets of the biology of opportunistic pathogenic bacteria. Scientific Reports, 8 (1), 9642.