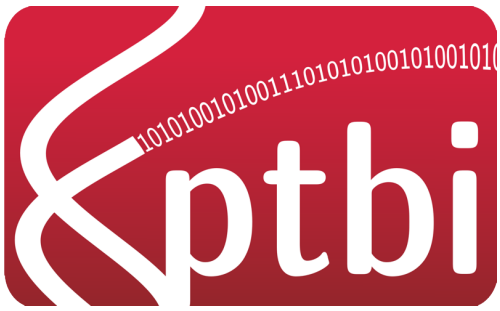
Bioinformatics Fights Viruses
Book of Abstracts

25 June 2020, Toruń, Poland



PROGRAM COMMITEE:

- Prof. Wieslaw Nowak (*Nicolaus Copernicus University, Torun, Poland*)
- Dr. hab. Bartek Wilczynski (*University of Warsaw, Poland*)

LOCAL ORGANIZING COMMITTEE:

- Prof. Wieslaw Nowak (*Nicolaus Copernicus University, Torun, Poland*)
- Dr. Łukasz Pepłowski (*Nicolaus Copernicus University, Torun, Poland*)
- Dr. Katarzyna Walczewska-Szewc (*Nicolaus Copernicus University, Torun, Poland*)
- Dr. Jakub Rydzewski (*Nicolaus Copernicus University, Torun, Poland*)

Contents

Lectures

Network bioinformatics analysis provides insight into drug repurposing for COVID-2019

Xu Li[1]

[1]Director of Data Center, Genenet Pharmaceuticals Co., Ltd., Tianjin, China

The COVID-2019 disease caused by the SARS-CoV-2 virus is a health crisis worldwide. While developing novel drugs and vaccine is long, repurposing existing drugs against COVID-2019 can yield treatments with known preclinical, pharmacokinetic, pharmacodynamic, and toxicity profiles, which can rapidly enter clinical trials. In this study, we present a novel network-based drug repurposing platform to identify candidates for the treatment of COVID-2019. At the time of the initial outbreak, knowledge about SARS-CoV-2 was lacking, but based on its similarity with other viruses, we sought to identify repurposing candidates to be tested rapidly at the clinical or pre-clinical levels. We first analyzed the genome sequence of SARS-CoV-2 and confirmed SARS as the closest virus by genome similarity, followed by MERS and other human coronaviruses. Using text mining and database searches, we obtained 34 COVID-2019-related genes to seed the construction of a molecular network where our module detection and drug prioritization algorithms identified 24 disease-related human pathways, five modules, and 78 drugs to repurpose. Based on clinical knowledge, we re-prioritized 30 potentially repurposable drugs against COVID-2019 (including pseudoephedrine, andrographolide, chloroquine, abacavir, and thalidomide). Our work shows how in silico repurposing analyses can yield testable candidates to accelerate the response to novel disease outbreaks.

Simulation analysis of hot spot residues interactions: A comparative analysis of spike protein Receptor Binding Domain of SARS-CoV 2 and SARS-CoV with Human ACE2

Jayakumar Singh[1]

[1]Department of Biotechnology, Koneru Lakshmaiah Education Foundation, India

The emergence of highly pathogenic novel corona virus in 2019 (SARS-CoV 2) became pandemic and posed a global health emergency. It was well understood that the binding of SARS-CoV (CoV) spike protein (S-Protein) Receptor Binding Domain (RBD) to Angiotensin converting enzyme 2 (ACE 2) receptor initiates the entry of corona virus into the host cells leading to the infection. However, considering the mutations reported in the SARS-CoV 2 (nCoV), the structural changes and the binding interactions of the S-protein RBD of nCoV were not clear. The present study was designed to elucidate the structural changes, hot spot binding residues and their interactions between the nCoV S-protein RBD and ACE2 receptor through various computational applications and molecular dynamic simulations.

Based on the sequence alignment, a total of 58 residues were found mutated in nCoV S-protein RBD and 28 mutated residues were identified in the binding motif. These mutations led to the structural changes in the nCoV S-protein RBD 3d structure with 4 helices, 10 sheets and intermittent loops while CoV RBD was found only with 2 helices and 5 sheets. With these structural changes, the nCoV RBD was found binding to ACE2 receptor with 11 hydrogen bonds and 1 salt bridge. The hot spot amino acids involved in the binding identified by interaction analysis includes Glu 35, Tyr 83, Asp 38, Lys 31, Glu 37, His 34, Lys 353 and Asp 30 amino acid residues of ACE2 receptor and Gln 493, Gln 498, Asn 487, Tyr 505, Lys 417, Thr 500, Tyr 489, Asn 501, Tyr 453 and Ala 475 residues in nCoV S-protein RBD. Interestingly, the two mutated residues Gln 493 and Asn 501 were found having highest Hydrogen bonding interaction percentage along with other hot spot residues and thereby binding to ACE 2 receptor with higher stability and rigidity as evident from the RMSD and RMSF data of the 100ns simulation. Concluding, the hotspots information could help in designing blockers for inhibiting the binding of nCoV spike protein RBD to human ACE2 receptor.

The In Silico Prediction of the siRNA Design for Deterring the SARS-CoV-2 Virus: a Structural Bioinformatics approach

Arli Aditya Parikesit[1]

[1]Department of Bioinformatics, School of Life Sciences, Indonesia International Institute for Life Sciences, Jl. Pulomas Barat Kav.88 Jakarta 13210 Indonesia

Nowadays, researchers all over the world contribute to solving scientific problems related to the SARS-CoV-2 pandemic. There are thousands of viral genomes sequenced from patients from a variety of countries and regions. Additionally, we know the full genomes of other betacoronaviruses. A combination of all the data sources presents an excellent opportunity to construct viable sequential alignments. These are especially useful as meaningful constraints for RNA structural analyses. In the case of the SARS-CoV-2 genome, its 5' and 3' UTRs fold into RNA motifs, which play a vital role in the virus replication process. Therefore, they are promising candidates for drug targets.

The RNA-Puzzles initiative unites researchers focused on RNA 3D structure prediction. Several groups of bioinformatics and structural biology experts participate in RNA-Puzzles contests directed on the modeling of different RNA 3D folds. With the global pandemic spreading, the RNA-Puzzles community agreed to share efforts. The participants modeled 5' and 3' UTRs, jointly generating more than a hundred of 3D models.

We present the results of comparative analyses performed on this dataset. The high number of 3D models combined with their lengths and complexity makes such studies challenging to implement. Therefore, we created specialized tools and computational pipelines, now available also to the public. We present our methodology of how to select promising RNA 3D models among a set of candidates. Then, we showcase its application on RNA 3D models of 5' and 3' UTRs of the SARS-CoV-2 genome.

Shotgun Transcriptome and Isothermal Profiling of SARS-CoV-2 Infection Reveals Unique Host Responses, Viral Diversification, and Drug Interactions

Christopher E. Mason[1]

[1]Department of Physiology and Biophysics Weill Cornell Medicine, WorldQuant (WQ) Init. for Quantitative Prediction Information Society Project, Yale Law School, USA

Testing, quarantine, contact tracing and social distancing are the major weapons to fight the spread of COVIS 19 until now. Based on microsimulations of the MOCOS model we investigate how the epidemic phase transition depends on the various parameters of the aforementioned counterstrategies. We discuss some methods to estimate the testing and tracking efficiency in Poland. Finally we present results about the potential impact of the recently introduced ProteGo-Safe contact tracking app.

MetaSUB efforts in SARS-CoV-2 fight

Paweł P Łabaj[1]

[1]Malopolska Center of Biotechnology, Jagiellonian University, Cracow, Poland

The MetaSUB Research Consortium involves scientists and researchers with varied areas of expertise worldwide to undertake challenges and seek the unknown. It addresses a wide variety of scientific questions and exploration of different fields with main aim to build a molecular profile of cities around the globe to improve their design, functionality, and impact on health. With the emergence of the COVID-19 pandemic in 2020 the MetaSUB International Consortium banded together to sample urban environments in cities around the globe before, during and after the pandemic.

The collected samples are going to be testes with use of RT-qPCR SARS-CoV-2 specific tests as well as with used LAMP based test. In addition, Whole Metagenomic Sequencing will be performed. In so far completed analysis of pilot NY WMS data where 86 samples were collected from handrails, kiosks, and floors in Grand Central and Times Square subway stations between March 6 and 13th, 2020 we did not observe significant counts or proportions.

Interestingly in Krakow, where samples were collected between March 17th and June 6th from 20 locations 22 samples among 172 were positives when tested with RT-qPCR tests. In addition, in parallel collection of samples form 6 objects within area of main train station 1 sample has been found positive with use of LAMP based tests.

The project aims to characterize the change in the urban metagenome and possibly isolate the presence of the SARS-CoV-2 virus. It will be possible once all samples from all involved cities will be processed and analysed in context of pandemic related meta-data.

Dealing with insufficient primer specificity

Michal Kowalski[1]

[1]Bioinformatics Research Group @ MCB UJ

SARS-CoV-2 virus has exploded as pandemic, media-trend and very serious threat to civilization. Its robustness in mutagenic potential and severe complications for people who had been contracted with it created a strong need for fast and reliable diagnostic methods. Unfortunately some of proposed primers for RT-qPCR diagnostics had been working very poorly or not been working at all. To make things worst, SARS-CoV-2 virus as a ssRNA(+) virus is highly unstable and it leads to stochastic mutagenic events (fast evolution). We propose a free of charge, interdisciplinary platform for development of diagnostic tests all around the world, to speed up wet lab inference with in silico global primer benchmarks and design tools. This platform is being developed to fit needs of wet lab scientists and bioinformaticians. Proposed pyprimer platform is highly efficient, human-readable, aware of computational resources, open source licensed and almost fully implemented in Python (so also ready for further, easy development). In nearest future pyprimer will incorporate newest data science techniques for the task of primer design and also will be publicly available within PyPI, Anaconda Cloud, DockerHub and GitHub not only as plain python library, but also with GUI, WebApp and as a contenerized version for fast deploy.

Age, inflammation and disease location are critical determinants of intestinal expression of SARS-CoV-2 receptor ACE2 and TMPRSS2 in inflammatory bowel disease

Jan K. Nowak[1]

[1]Translational Gastroenterology Unit, University of Oxford, Oxford, UK

Two proteins are known to play a key role in coronavirus disease 2019 (COVID-19): angiotensin-converting enzyme 2 (ACE2) which allows severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) cellular entry, and transmembrane serine protease 2 (TMPRSS2) that primes the viral spike protein. Inflammatory bowel diseases (IBD) are chronic and debilitating conditions affecting 1% of the population in the most developed countries. We aimed to investigate the correlates of ACE2 and TMPRSS2 expression in the intestine of patients with IBD and controls in the European study IBD-Character. The analysis of transcriptomic data (Agilent SurePrint G3 v2) from patients (n=138) and controls (n=154) revealed that age, inflammation and disease location determine the intestinal expression of ACE2 and TMPRSS2 in IBD. In UC the mucosal expression of ACE2 was 70% greater than in controls ($p = 2.1 \cdot 10^{-11}$), 50% higher in inflamed vs non-inflamed tissue ($p = 6.3 \cdot 10^{-5}$), and correlated with Mayo endoscopic subscore (rho=0.43, $p = 3.2 \cdot 10^{-5}$). TMPRSS2 associated with colonic inflammation (p=0.0179), UC extent (150% difference, p=0.0002), and male sex (p=0.03); it was also greater in both UC (by 30%, p=0.023) and CD ileum compared with controls (70% difference, p=0.023). ACE2 expression in CD ileum was reduced by 60% vs controls (p=0.0175) possibly revealing the loss of epithelium or novel CD pathophysiology. These results provide mechanistic insight into the putative causes of a more severe COVID-19 course in patients with active IBD.

CIRCA COVID-19 CXR-based diagnosis

Joanna Polańska[1]

[1]Silesian University of Technology, Gliwice, Poland

COVID-19 pandemic disease is caused by SARS-CoV-2 coronavirus, which originated in China in Wuhan at the end of the year 2019. Because of the growing number of patients, the health system is highly overloaded, even in developed countries. In addition, the number of coronavirus tests performed is limited. Before 2019, the disease was not identified in humans; therefore, methods of better detection, treatment and prevention are intensively sought. Deep learning techniques can help in faster detection of COVID-19 cases since their usefulness has already been reported in various fields of medicine. We aimed to develop an efficient and robust machine learning based system for automated chest X-ray analysis in support of COVID-19 diagnosis. The proposed solution is a two-step combination of U-Net and expansive NN architectures allowing to distinguish COVID-19 patients with the sensitivity of 92% and PPV 94%. Combination of two deep learning networks, working in the sequential mode, permits to perform both lung segmentation and classification tasks efficiently, independently of the image quality. It also leads to the single network to be of lighter architecture with fewer parameters to tune and smaller datasets to be used at the training stage. The CIRCA implementation for chest X-Ray analysis is publicly available at https://covid.aei.polsl.pl.

Identification of SARS-CoV-2 virus sequences types based on interpretable machine learning models

Mateusz Kudla[1]

[1]University of Applied Sciences Mittweida, Saxon Institute for Computational Intelligence and Machine Learning, Mittweida, Germany

Group testing for SARS-CoV-2

Szymon Toruńczyk[1]

[1]Institute of Informatics, University of Warsaw, Poland

SEIR-based modelling of COVID-19 in Poland

Krzysztof Gogolewski[1], Magdalena Rosinska[2], Daniel Rabczenko[3], Ewa Szczurek[1], Blazej Miasojedow[4], and Anna Gambin[1]

[1]Institute of Informatics, University of Warsaw
[2]Department of Epidemiology and Surveillance of Infectious Diseases, National Institute of Public Health - National Institute of Hygiene
[3]Department of Monitoring and Analysis of the Population Health State, National Institute of Public Health – National Institute of Hygiene
[4]Institute of Applied Mathematics and Mechanics, University of Warsaw

The worldwide epidemics of COVID-19 has become a leading topic of research in a variety of scientific fields. It is of high interest to understand the dynamics of the virus spread which naturally depends on dozens of factors that should be taken into account and can be approached from various perspectives.

In our work, we extend epidemiological SEIR models commonly used these days in the literature. Using the data provided by national authorities concerning the daily number of new cases, deaths, recoveries, and performed tests we try to provide some insight into the short- and mid-term dynamics of the disease in Poland. Additionally, we aim to determine the influence of the introduction or release of the social restrictions on the dynamics of the epidemics. Finally, we try to deduce how the testing strategies influences the virus spread in the society.

In summary, our research aims to quickly describe the prognostics as well as the differences in the dynamics and evaluate the possible influence of changes introduced to the social politics of Poland regarding COVID-19.

Estimate of Covid-19 prevalence using imperfect data

Krzysztof Mnich[1]

[1]University of Białystok, Poland

The impact of testing, contact tracing and quarantine - insights from the MOCOS microsimulation model

Tyll Krueger[1]

[1]Wroclaw University of Science and Technology, Department of Electronics, Wroclaw, Poland

Testing, quarantine, contact tracing and social distancing are the major weapons to fight the spread of COVIS -19 until now. Based on microsimulations of the MOCOS model we investigate how the epidemic phase transition depends on the various parameters of the aforementioned counterstrategies. We discuss some methods to estimate the testing and tracking efficiency in Poland. Finally we present results about the potential impact of the recently introduced ProteGo-Safe contact tracking app.

SARS-CoV-2 structure modeling within RNA-Puzzles initiative

Tomasz Zok[1], Maciej Antczak[1,2], Mariusz Popenda[2], Joanna Sarzynska[2], Marcin Zablocki[1], Ryszard W. Adamiak[1,2], and Marta Szachniuk[1,2]

[1]Institute of Computing Science, Poznan University of Technology
[2]Institute of Bioorganic Chemistry, Polish Academy of Sciences

Nowadays, researchers all over the world contribute to solving scientific problems related to the SARS-CoV-2 pandemic. There are thousands of viral genomes sequenced from patients from a variety of countries and regions. Additionally, we know the full genomes of other betacoronaviruses. A combination of all the data sources presents an excellent opportunity to construct viable sequential alignments. These are especially useful as meaningful constraints for RNA structural analyses. In the case of the SARS-CoV-2 genome, its 5' and 3' UTRs fold into RNA motifs, which play a vital role in the virus replication process. Therefore, they are promising candidates for drug targets.

The RNA-Puzzles initiative unites researchers focused on RNA 3D structure prediction. Several groups of bioinformatics and structural biology experts participate in RNA-Puzzles contests directed on the modeling of different RNA 3D folds. With the global pandemic spreading, the RNA-Puzzles community agreed to share efforts. The participants modeled 5' and 3' UTRs, jointly generating more than a hundred of 3D models.

We present the results of comparative analyses performed on this dataset. The high number of 3D models combined with their lengths and complexity makes such studies challenging to implement. Therefore, we created specialized tools and computational pipelines, now available also to the public. We present our methodology of how to select promising RNA 3D models among a set of candidates. Then, we showcase its application on RNA 3D models of 5' and 3' UTRs of the SARS-CoV-2 genome.

Email: tomasz.zok@cs.put.poznan.pl

Towards computational elucidation of SARS-CoV-2 virus entry point mechanostability

Katarzyna Walczewska-Szewc[1,2], Lukasz Peplowski[1,2], Beata Niklas[1], M. Lewandowski[1], Julia Berdychowska[1], and Wieslaw Nowak[1,2]

[1]Faculty of Physics, Astronomy and Informatics, Nicolaus Copernicus University, Grudziądzka 5, 87-100 Toruń, Poland

[2]Centre for Modern Interdisciplinary Technologies, Nicolaus Copernicus University, Wileńska 4, 87-100 Toruń, Poland

Before replication viruses attach and enter host cells. SARS-CoV-2 virus has multiple entry points but the major route starts from an attachment of its spike protein S (pS) to Angiotensin-Converting Enzyme 2 (ACE2) present in lungs, intestines, arteries etc. The probability of successful infection depends on mechanical stability of the pS-ACE2 interface and may be lowered by using selective drugs. Based on a structure of pS-ACE2 complex1 (6M0J), a series of our own 100 ns trajectories (GROMACS, CHARMM37) and following recent MD simulations we identified main "force clamps" responsible for the virus attachment to the membrane protein ACE2. To reduce computational efforts, before more extensive nanomechanical studies of this system are undertaken, we developed a simplified model of the SARS-CoV-2 virus entry point. Our preliminary data show that dynamical parameters of the interface are quite well reproduced in the model. Docking (Glide, Accelrys Inc.) of SSAA09E, nicotine, chloroquine, dexamethasone, remdesivir – compounds discussed in the context of COVID-19 - shows that our small system provides a reliable starting point for screening for anti-virial drugs limiting the virus entry into a cell. Thus, this study allowed us to build a minimal test system which is able to reconstruct the behavior of the pS-ACE2 complex interface. This is the first step of a larger project aiming for design of an automatic, structural-bioinformatics tool facilitating the anti-COVID-19 drug development.

Acknowledgement: This project is funded by IDUB N. Copernicus ANTICO and #MEMOBIT grants. ICNT UMK computer facilities are acknowledged.

Email: kszewc@fizyka.umk.pl

Genome-wide mapping of therapeutically-relevant SARS-CoV-2 RNA structures

Ilaria Manfredonia[1], Chandran Nithin[2], Almudena Ponce-Salvatierra[2], Pritha Ghosh[2], Tomasz K. Wirecki[2], Tycho Marinus[2], Eric J. Snider[3], Natacha S. Ogando[3], Martijn J. van Hemert[3], Janusz M. Bujnicki[2], and Danny Incarnato[1]

[1]Department of Molecular Genetics, Groningen Biomolecular Sciences and Biotechnology Institute (GBB), University of Groningen, Nijenborgh 7, 9747 AG, Groningen, the Netherlands
[2]Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology, ul. Ks. Trojdena 4, 02-109 Warsaw, Poland
[3]Molecular Virology Laboratory, Department of Medical Microbiology, Leiden University Medical Center, Leiden, the Netherlands

SARS-CoV-2 is a betacoronavirus with a linear single-stranded, positive-sense RNA genome of 30 kb, whose outbreak caused the still ongoing COVID-19 pandemic. The ability of coronaviruses to rapidly evolve, adapt, and cross species barriers makes the development of effective and durable therapeutic strategies a challenging and urgent need. As for other RNA viruses, genomic RNA structures are expected to play crucial roles in several steps of the coronavirus replication cycle. Despite this, only a handful of functionally conserved structural elements within coronavirus RNA genomes have been identified to date. We performed RNA structure probing by SHAPE-MaP to obtain a single-base resolution secondary structure map of the full SARS-CoV-2 coronavirus genome. The SHAPE-MaP probing data recapitulate the previously described coronavirus RNA elements (5' UTR, ribosomal frameshifting element, and 3' UTR), and reveal new structures. Secondary structure-restrained 3D modeling of highly-structured regions across the SARS-CoV-2 genome allowed for the identification of several putative druggable pockets. Furthermore, 8% of the identified structure elements show significant covariation among SARS-CoV-2 and other coronaviruses, hinting at their functionally-conserved role. In addition, we identify a set of persistently single-stranded regions having high sequence conservation, suitable for the development of antisense oligonucleotide therapeutics. Collectively, our work lays the foundation for the development of innovative RNA-targeted therapeutic strategies to fight SARS-related infections.