УДК 616.9

Vladimir L. Motin

SYSTEMS BIOLOGY OF *YERSINIA PESTIS*

University of Texas Medical Branch, Departments of Pathology and Microbiology & Immunology, Galveston, Texas, USA

Omics technologies developed in recent years significantly changed our approach in study molecular biology of bacterial pathogens. The research now can be conducted at a whole-genome level enabling us to simultaneously measure variable components, such as proteins, nucleic acids, metabolites, cell lipid composition, *etc*. This complex work requires a constant interaction of multi-disciplinary experts to integrate biology, instrumentation and computational science. A multi-omics research (transcriptomics, proteomics, metabolomics) was applied to *Y. pestis*-centric system biology study. Integrated omics datasets allowed to refine genome annotations, discover novel putative virulence factors, build regulatory networks, and create an interactive computational metabolic model.

Key words: *Yersinia*, systems biology, plague, omics research, modeling

В.Л.Мотин

Системная биология *Yersinia pestis*

Медицинский факультет Университета штата Техас, Галвестон, США

Разработанные в последние годы методы -омик (протеомика, геномика, транскриптомика и др.) существенно изменили подход к изучению молекулярной биологии бактериальных патогенов. В настоящее время исследования могут проводиться на уровне полного генома, что дает возможность одновременно оценивать различные компоненты (белки, нуклеиновые кислоты, метаболиты, липиды и т.д.). Эта работа требует постоянного взаимодействия специалистов по многим дисциплинам для интеграции биологии, инструментального анализа и компьютерных наук. Исследования на основе мультиомик были осуществлены при анализе системной биологии *Y. pestis*. Полученные обобщенные наборы данных позволили внести уточнения в аннотации геномов, выявить новые потенциальные факторы вирулентности, построить регуляторные сети и создать интерактивную компьютерную модель метаболизма.

Ключевые слова: *Yersinia*, системная биология, чума, исследования на основе -омик, моделирование

The systems biology strategy employs integrative experimental and computational approaches to analyze data obtained in different "omics" experiments. Ideally, biological samples should be produced for the simultaneous assay by transcriptomic, proteomic, metabolomic and lipidomic methods, eventually resulting in the development of a predictive computational model. As a part of the project named Systems Biology of Enteropathogens funded by the U.S. National Institutes of Health, we applied a system biology approach to study *Yersinia* spp., particularly the *Y. pestis* microbe, the causative agent of plague. The project was limited to studying the microorganism *in vitro*, rather than evaluating a host-pathogen interaction. The samples produced at V.L. Motin's laboratory at the University of Texas Medical Branch (UTMB) were processed to the transcriptome at the J. Craig Venter Institute (JCVI), and to the proteome, metabolome and lipidome assays at the Pacific Northwest National Laboratory (PNNL). The "omics" data were used by the computational experts at the University of California, San Diego (UCSD) to build predictive models of biological and regulatory networks in *Yersinia*, particularly those related to pathogenicity. In parallel, a similar approach was applied to *Salmonella*, another member of the family *Enterobacteriaceae*, in collaboration with the Oregon Health and Sciences University (OHSU) that allowed us to coordinate both studies of these medically important pathogens. The experimental protocols, rough data and results obtained in these studies are publicly available at the Center for Systems Biology of Enteropathogens web site: http://www.sysbep.org/.

Here I present some findings obtained during the course of the *Yersinia* part of the project. The following *Yersinia* strains were used: fully virulent *Y. pestis* CO92, biovar Orientalis, isolated from the human case of primary pneumonic plague in the U.S. [1]; strain *Y. pestis* Pestoides F with limited virulence, belonging to the subspecies *caucasica*, which likely originated from a Transcaucasian highland plague focus [2, 3]; and pathogenic *Y. pseudotuberculosis* PB1, serovar O1b [4].

Omics for Genome Annotation Refinement

A tremendous growth in the number of completely sequenced genomes resulted in an almost exclusive automated annotation of the generated sequences. Therefore, a correct and complete accounting for the open reading frames (ORFs) of the reference organisms in GenBank should help us to reduce errors of further computational annotations. In this respect a so-called proteogenomic approach based on analysis of the transcriptome and proteome of the organism serves a purpose of correct identification of ORFs. We applied this approach to refine annotation of the two above-mentioned *Y. pestis* strains CO92 and Pestoides F, as well as *Y. pseudotuberculosis* PB1. To perform a transcriptome analysis,

a universal *Yersinia* microarray chip was designed that consisted of 70-mer oligonucleotides corresponding to the putative ORFs of seven different *Yersinia* genomes. Proteome data were obtained by reverse-phase nano-capillary LC-MS/MS analyses. Bacteria were grown in a chemically- defined medium modified by R. Brubaker [5] at 26 and 37 °C, and samples for isolation of RNA and a global protein digest were taken at different time points. Analysis of transcripts and peptides produced by bacteria during applied experimental conditions allowed us to find 96 previously undocumented annotation errors across these three *Yersinia* strains. Summary of genome annotation refinements included such features as discovery of novel ORFs, extended start sites, identification of frameshifts, and translated "pseudogenes." Among the genes that were missed or misinterpreted during previous annotations in the GenBank conducted by different research groups were essential ribosomal protein, several virulence-associated factors, a transcriptional regulator, and many hypothetical proteins. Details of this study can be found in our recent publication [6].

Multi-Omics Strategy to Unraveling Yersinia Pathogenesis

Y. pestis, the agent of plague, an acute and highly lethal infection, is a recently evolved clone of *Y. pseudoturbeculosis*, the enteropathogen causing a limited infection in mammals and birds [7]. Both organisms share ~97% homology at the nucleotide level, as determined by whole-genome sequence analysis [8]. By making a comprehensive omics-driven comparative analysis of *Y. pestis* and *Y. pseudoturberculosis*, it might be possible to gain an understanding as to how an acute and highly lethal bacterial pathogen, such as *Y. pestis*, has evolved from its less virulent progenitor. Furthermore, a cross-comparison of epidemic and non-epidemic biovars of *Y. pestis* should provide insight into the mechanism underlying a virulence-restricted phenotype of strains, such as Pestoides F. One of the omics approaches for this type of comparative studies can be the use of defined *in vitro* conditions mimicking infection, such as transition from the ambient to the mammalian body temperature, acidified pH, oxidative stress, *etc.* that allow us to address global changes in organisms at different time points. We were the first to describe whole-genome temporal transcriptional changes in *Y. pestis* during a temperature shift from 26 to 37 °C. We found that 235 and 274 of the chromosomal genes were up- and down-regulated, respectively, and many plasmid genes were thermoregulated as well [9]. However, other omics approaches, for example proteomic and metabolomic analyses, were not simultaneously performed with the grown bacteria obtained for the transcriptomic experiments. Therefore, we conducted a temporal multi-omic cross-comparison of CO92, Pestoides F and PB1 strains by using a universal *Yersinia* microarray chip for gene expression and mass spectrometric methods for evaluation of protein and metabolite levels. This analysis revealed high gene and protein expression levels of conserved major virulence factors between two *Yersinia* species, as well as conserved post-transcriptional control of metabolism. Importantly, statistical inference modeling methods applied to the multi-omics integration of the network predicted novel virulence factors, including many genes annotated as having an unknown function. The role of these putative virulence factors in *Yersinia* pathogenesis is awaiting experimental confirmation. The results of the study were recently published [10].

Genome-Scale Metabolic Network Reconstruction

A first attempt to construct a metabolic network and corresponding mathematical model for *Y. pestis* was made based on annotation of the microtus biovar strain 91001 that is avirulent to humans. However, the construction of this model was not supported by the actual experiments on detection of metabolite spectra [11]. As a part of our systems biology project, we built a broader metabolic model for the virulent *Y. pestis* CO92. Importantly, during construction of the metabolic network, we used metabolite identification and uptake/secretion measurements in samples of supernatants of *Y. pestis* grown at 26 °C on different carbon sources. Time points were taken every two hours for the ten hours of the growth period. The models contained 815 genes, 678 proteins, 963 unique metabolites, and 1678 reactions. This model serves as a base for our current efforts to integrate other "omics" data sets that we produced for this pathogen to fill gaps for model stimulations to occur. Another important result of this study was a prediction of essential genes of *Y. pestis* that might be valuable targets for development of antimicrobials. The list of the essential genes as well as experimental details can be found in our recent publication [12].

Conclusions and Future Directions

My personal lesson that I learned participating in this large systems biology project is that these days the success of the research depends a great deal on proper interaction provided by multi-disciplinary experts and sophisticated technological and computational approaches. The research at the whole-genome level became tremendously complicated, and a massive amount of data produced during the "omics" experiments should be analyzed and re-analyzed again and again. Next-generation sequencing technology will be very instrumental not only for the massive genomic sequencing, but for the expression studies (RNA-seq technology) as well, likely replacing microarrays for most applications. In addition, RNA-seq is very helpful in identifying small (non-coding) RNAs (sRNAs) that are critical regulators of the cellular network. The use of RNomics has allowed several research groups to identify more than a hundred abundant sRNAs in *Yersinia* spp. that play important roles in physiology and pathogenicity. And last, but not least, we always should remember that above all an interaction

of a pathogen and a mammal (as well as an arthropod vector and/or microbiota) is a key point of the microbial pathogenesis research. Thus, systems biology projects should address both sides of "omics" – the microbe and its host.

Acknowledgements

I would like to acknowledge the contributions of key members of the Center for Systems Biology of Enteropathogens: Drs. C. Ansong, T.O. Metz, K. McAteer, J.E. McDermott, and J.N. Adkins (PNNL); Drs. M.B. Jones and S.N. Peterson (JCVI); Drs. D.E. Hyduke, B. Schmidt, P. Charusanti, and B.O. Palson (UCSD); Dr. F. Hefron (OHSU). This program was supported by NIH/NIAID award #Y1-AI-8401.

References

1. *Doll J.M., Zeitz P.S., Ettestad P., Bucholtz A.L., Davis T., Gage K.* Cat-transmitted fatal pneumonic plague in a person who traveled from Colorado to Arizona. Am. J. Trop. Med. Hyg. 1994; 51:109–14.

2. *Motin V.L., Georgescu A.M., Elliott J.M., Hu P., Worsham P.L., Ott L.L., Slezak T.R., Sokhansanj B.A., Regala W.M., Brubaker R.R., Garcia E* .Genentic variability of *Yersinia pestis* isolates as predicted by PCR-based IS*100* genotyping and analysis of structural genes endocing glycerol-3-phosphate dehydrogenase (*glpD*). J. Bacteriol. 2002; 184:1019–27.

3. *Onishenko G.G., Kutyrev V.V.,* editors. Natural plague foci in the Caucasis, Caspian Sea Region, Middle Asia and Siberia. Moscow: Medicine; 2004. 192 p.

4. Burrows T.W., Bacon G.W. V and W antigens in strains of *Pasteurella pseudotuberculosis*. Brit. J. Exp. Pathol. 1960; 41:38–44.

5. *Fowler J.M., Brubaker R.R.* Physiological basis of the low calcium response in *Yersinia pestis*. Infect. Immun. 1994; 62: 5234–41.

6. *Schrimpe-Rutledge A.C., Jones M.B., Chauhan S., Purvine S.O., Sanford J.A., Monroe M.E., Brewer H.M., Payne S.H., Ansong C., Frank B.C., Smith R.D., Peterson S.N., Motin V.L., Adkins J.N.* Comparative omics-driven genome annotation refinement: application across *Yersiniae*. PLoS One. 2012; 7(3):e33903.

7. *Achtman M., Zurth K., Morelli G., Torrea G., Guiyoule A., Carniel E. Yersinia pestis*, the cause of plague, is a recently emerged clone of *Yersinia pseudotuberculosis*. Proc. Natl Acad. Sci. USA. 1999; 96:14043–8.

8. *Chain P.S., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O., Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L., Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C., Simonet M., Chenal-Francisque V., Souza B., Dacheux D., Elliott J.M., Derbise A., Hauser L.J., Garcia E.* Insights into the evolution of *Yersinia pestis* through whole-genome comparison with *Yersinia pseudotuberculosis*. Proc. Natl Acad. Sci. USA. 2004; 101:13826–31.

9. *Motin V.L., Georgescu A.M., Fitch J.P., Gu P.P., Nelson D.O., Mabery S.L., Garnham J.B., Sokhansanj B.A., Ott L.L., Coleman M.A., Elliott J.M., Kegelmeyer L.M., Wyrobek A.J., Slezak T.R., Brubaker R.R., Garcia E.* Temporal global changes in gene expression during temperature transition in *Yersinia pestis*. J. Bacteriol. 2004; 186:6298–6305.

10. *Ansong C., Schrimpe-Rutledge A.C., Mitchell H.D., Chauhan S., Jones M.B., Kim Y.M., McAteer K., Deatherage Kaiser B.L., Dubois J.L., Brewer H.M., Frank B.C., McDermott J.E., Metz T.O., Peterson S.N., Smith R.D., Motin V.L., Adkins J.N.* A multi-omic systems approach to elucidating *Yersinia* virulence mechanisms. Mol. Biosyst. 2013; 9:44–54.

11. *Navid A., Almaas E.* Genome-scale reconstruction of the metabolic network in *Yersinia pestis*, strain 91001. Mol. Biosyst. 2009; 5:368–75.

12. *Charusanti P., Chauhan S., McAteer K., Lerman J.A., Hyduke D.R., Motin V.L., Ansong C., Adkins J.N., Palsson B.O.* An experimentally-supported genome-scale metabolic network reconstruction for *Yersinia pestis* CO92. BMC Syst. Biol. 2011; 5:163.

Authors:

Motin Vladimir L. University of Texas Medical Branch, Departments of Pathology and Microbiology & Immunology. Galveston, Texas, USA. E-mail: vlmotin@utmb.edu

Об авторах:

Мотин В.Л. Медицинский факультет Университета штата Техас. Галвестон, США. E-mail: vlmotin@utmb.edu

Поступила 21.08.13.