

# Emilio Mastriani

Date of birth: 24 January 1974  
Place of birth: Palermo, Italy  
Citizenship: Italian  
Marital status: married  
Residency:  
Room 1701, Unit 1, 7/F, Jiayuan House, 177 Fudan Street  
Nangang District, 150000 Harbin (China)  
Contacts: GSM: +86 13664502721 QQ.: 3032151181  
E-mail: emiliomastriani@icloud.com  
Web of Science Researcher ID: ABI-7926-2020  
ORCID: 0000-0002-5434-2546



## EDUCATION

**Doctor of Philosophy, Computational Biology**, Selinus University, Bologna, Italy, 2020 November 15

- Unsupervised clustering approach to characterize the CpG Island distribution of the Andes Hantavirus, Prof. Salvatore Fava Ph. D
- Examination fields: Bioinformatics, Statistical methods, Clustering techniques, Virology

**Master of Science, Computer Science, University of Catania**, Catania, Italy, 2000 November

- Covert channel analysis of Linux file system, Prof. V. Cutello
- Part of the thesis was published in the proceedings of ALS Usenix Conference, 2000 October, Atlanta, USA

## Professional Skills Training

- Mathematics for Machine Learning: Linear Algebra. Imperial College London - Verify at [coursera.org/verify/F2RXMZHAFYA6](https://coursera.org/verify/F2RXMZHAFYA6) (15/09/2019–20/10/2019)
- Synapses, Neurons and Brains. Hebrew University of Jerusalem on Coursera, Prof. Idan Segev (10/03/2019–22/05/2019)
- Understanding the Brain: The Neurobiology of Everyday Life. University of Chicago on Coursera (17/04/2017–03/06/2017)
- 9th Advanced School on PARALLEL COMPUTING. CINECA, Bologna, Italy (11/02/2013–15/02/2013)
- Introduction to GPGPU and CUDA programming. CINECA, Bologna, Italy (22/05/2012–24/05/2012)
- International Summer School on Grid Computing. University of Edinburgh and University of Catania, Mariefred, Sweden (08/07/2007–20/07/2007)
- DEISA Training Session. Barcelona Supercomputing Center, Barcellona, Spain (07/03/2007–09/03/2007)
- Advanced Course on Parallel Computing and Grid Computing. Astrophysical Observatory of Catania, Catania, Italy (25/09/2006–29/09/2006)
- Scientific Programming Environment in Linux / Unix. CINECA, Bologna, Italy (19/05/2005–20/05/2005)

## RESEARCH INTERESTS / RESEARCH PROFILE

Nowadays, more than ever, there is an urgent need to use computational and mathematical tools to expand our knowledge in the biological field and especially in virology. Skills in statistics, computer science, genetics, and virology offer an optimal scenario for the use of investigative techniques such as machine and deep learning. Due to this reason, the evolutionary process of SARS-CoV-2 from different countries with similar geographical conditions was investigated with appropriate computational techniques, research that identified the codon under selective pressure and predicted the three-dimensional structure of the associated protein, noting its functional importance. The statistical and computational approaches to analyze the CpG dinucleotides odds ratio allowed us to make a step forward in identifying the Andes (46% of mortality with studies allowed only in Level 4 laboratories) Hantavirus as a unique cluster inside the Orthohantavirus family. This discovery lays the foundations for a hypothesis of its anthroponotic nature. Similarly, the case of the Sabia virus (also associated with the Level 4 laboratories for danger and mortality), by an appropriate computational study could reveal similarities in the composition of the capsid of the Lassa virus, providing concrete pharmacological solutions. And many questions remain open, such as computational models that support the identification of natural hosts for the Andes Hantavirus, for the Sabia virus, and for other RNA viruses of which there is no certainty about the natural reservoir. Future research projects will involve the use of the metric proposed by Lytras and Hunges (SDU) as a future for the Unsupervised clustering to confirm or reject the results obtained. Also, the calculation of the  $\delta^*$ -differences (described by Karlin) could help to identify pathogenic islands in the Andes hantavirus genome. Furthermore, I shall predict possible reservoirs using the dinucleotide study to estimate the "compatibility Host / Virus" in our case *N.Spinosus/H.Andes* and *A.Longipilis/H.Andes*. In conclusion, science is a fantastic journey around the unknown, and the more tools our ship has, the better it will be.

## RESEARCH EXPERIENCE

### Doctoral Thesis, Selinus University, Bologna, Italy, 2020 November 15

- *Unsupervised clustering approach to characterize the CPG Island distribution of the Andes Hantavirus*
- *Ph.D. research project, supervised by Prof. Shu-Lin Liu, Ph.D. A. Rakov and Prof. S. Fava.*
- *Developed in-depth knowledge of literature searching, virology of Orthohantaviruses, calculation of CpG odds ratio along segmented RNA genomes, statistical models for non-normal distribution, analysis of CpG odds ratio in CDS regions, and unsupervised clustering analysis of Orthohantaviruses.*
- *Results reported a statistical significance of the three genomics groups (L, M, S), stated the group of "Small" genome as the more informative (CpG odds ratio), confirmed the CpG odds ratio as the lowest (M, S) compared to others dinucleotides, highlighted as the Andes hantavirus detains the highest CpG odds ratio into CDS and forms an isolated cluster (as also the Tula H.)*

### 2020 November: Research, HMU-UCCSM Centre for Infection and Genomics, Harbin Medical University, Harbin, China

- *Isolating SARS-CoV-2 Strains From Countries in the Same Meridian: Genome Evolutionary Analysis*
- *Research project, conducted with Prof. Shu-Lin Liu, Ph.D. A. Rakov*

- *Developed in-depth knowledge of literature searching, virology of SARS-CoV-2, recombination analysis, phylogenetic tree construction, Mixed Effects Model of Evolution analysis, secondary structures prediction, “disorder” and “transmembrane” protein analyses, ab-initio protein structure prediction, and ontologies of the three-dimensional structure.*
- *JMIR Bioinform Biotech. 2021 Jan 22;2(1):e25995. doi: 10.2196/25995. PMID: 33497425; PMCID: PMC7837406.*

**2020 March: Research**, HMU-UCCSM Centre for Infection and Genomics, Harbin Medical University, Harbin, China

- *HamHeat: A fast and simple package for calculating Hamming distance from multiple sequence data for heatmap visualization*
- *Research project, conducted with Ph.D. A. Rakov, Prof. Dieter M. Schifferli, and Prof. Shu-Lin Liu*
- *Developed in-depth knowledge of literature searching, Hamming distance for multiple sequences, production of the HamHeat package to efficiently calculate Hamming distance for hundreds of aligned protein or DNA sequences of a large number of residues or nucleotides*
- *Pre-print version of the manuscript available at doi.org/10.1101/2020.03.26.009258*

**2020 January: Research**, HMU-UCCSM Centre for Infection and Genomics, Harbin Medical University, Harbin, China

- *Differential Expression Analysis on Schizophrenia Dataset Suggests Pseudogene RNU6-505P as under Selective Pressure*
- *Research project, conducted with Prof. Shu-Lin Liu, Ph.D. A. Rakov*
- *Developed in-depth knowledge of literature searching, genetic of schizophrenia, microarray analysis, correlation analysis, evolutionary analysis, and prediction of the ontological un-functionality in case of mutation*
- *Published version of the manuscript available at doi.org/10.21467/ias.9.1.86-99*

## PUBLICATIONS

1. Mastriani E, Liu SL. Unsupervised Clustering characterizes the CPG dinucleotides distribution of the Andes hantavirus, pre-print: bioRxiv 2021.02.23.432596; doi: <https://doi.org/10.1101/2021.02.23.432596>
2. Mastriani E, Rakov AV, Liu SL. Isolating SARS-CoV-2 Strains From Countries in the Same Meridian: Genome Evolutionary Analysis, JMIR Bioinform Biotech. 2021 Jan 22;2(1):e25995. doi: 10.2196/25995. PMID: 33497425; PMCID: PMC7837406.
3. Emilio Mastriani, Alexey V. Rakov, Shu-Lin Liu. An evolutionary analysis of the SARS-CoV-2 genomes from the countries in the same meridian, pre-print: bioRxiv 2020.11.12.380816; doi: <https://doi.org/10.1101/2020.11.12.380816>
4. Alexey V. Rakov, Dieter M. Schifferli, Shu-Lin Liu, Emilio Mastriani. HamHeat: A fast and simple package for calculating Hamming distance from multiple sequence data for heatmap visualization, pre-print: bioRxiv 2020.03.26.009258; doi: <https://doi.org/10.1101/2020.03.26.009258>
5. Mastriani, E., Rakov, A.V. and Liu, S.-L. 2020. Differential Expression Analysis on Schizophrenia Dataset Suggests Pseudogene RNU6-505P as under Selective Pressure. International Annals of Science. 9, 1 (Jan. 2020), 86-99. DOI: <https://doi.org/10.21467/ias.9.1.86-99>

6. Liu H, Zeng Z, Afshar M, Lin C, Wang S, Yang H, Liu S, Kelemen LE, Xu W, Ma W, Xiang Q, Mastriani E, Wang P, Wang J, Liu SL, Johnston RN, Köbel M. Overexpression of IGF2BP3 as a Potential Oncogene in Ovarian Clear Cell Carcinoma. *Front Oncol.* 2020 Jan 30; 9:1570. doi:10.3389/fonc.2019.01570. PMID: 32083017; PMCID: PMC7002550.
7. Rakov AV, Mastriani E, Liu SL, Schifferli DM. Association of Salmonella virulence factor alleles with intestinal and invasive serovars. *BMC Genomics.* 2019 May 28;20(1):429. doi: 10.1186/s12864-019-5809-8. PMID: 31138114; PMCID: PMC6540521.
8. Mastriani E, Zhai R, Zhu S. Microarray-Based MicroRNA Expression Data Analysis with Bioconductor. *Methods Mol Biol.* 2018; 1751:127-138. doi: 10.1007/978-1-4939-7710-9\_9. PMID: 29508294.
9. Liu H, Zeng Z, Wang S, Li T, Mastriani E, Li QH, Bao HX, Zhou YJ, Wang X, Liu Y, Liu W, Hu S, Gao S, Yu M, Qi Y, Shen Z, Wang H, Gao T, Dong L, Johnston RN, Liu SL. Main components of pomegranate, ellagic acid and luteolin, inhibit metastasis of ovarian cancer by down-regulating MMP2 and MMP9. *Cancer Biol Ther.* 2017 Dec 2;18(12):990-999. doi: 10.1080/15384047.2017.1394542. Epub 2017 Nov 27. PMID: 29173024; PMCID: PMC5718784.
10. Tang L, Mastriani E, Zhou YJ, Zhu S, Fang X, Liu YP, Liu WQ, Li YG, Johnston RN, Guo Z, Liu GR, Liu SL. Differential degeneration of the ACTAGT sequence among Salmonella: a reflection of distinct nucleotide amelioration patterns during bacterial divergence. *Sci Rep.* 2017 Sep 8;7(1):10985. doi: 10.1038/s41598-017-11226-9. PMID: 28887484; PMCID: PMC5591236.
11. Liu H, Liu J, Wang S, Zeng Z, Li T, Liu Y, Mastriani E, Li QH, Bao HX, Zhou YJ, Wang X, Hu S, Gao S, Qi Y, Shen Z, Wang H, Yu M, Gao T, Johnston RN, Liu SL. Enterolactone has stronger effects than enterodiol on ovarian cancer. *J Ovarian Res.* 2017 Jul 24;10(1):49. doi: 10.1186/s13048-017-0346-z. PMID: 28738876; PMCID: PMC5525236.
12. Dr. Le Tang, Dr. Songling Zhu, Dr. Emilio Mastriani, Miss Xin Fang, Dr. Yu-Jie Zhou, Dr. Yong-Guo Li, Prof. Randal Johnston, Dr. Zheng Guo, Ms. Gui-Rong Liu. (2017) "Conserved intergenic sequences revealed by CTAG-profiling in Salmonella: thermodynamic modeling for function prediction", *Scientific Reports*, Accepted 23 January 2017
13. Huidi Liu, Emilio Mastriani, Zi-Qiao Yan, Si-Yuan Yin, Zheng Zeng, Hong Wang, Qing-Hai Li, Hong-Yu Liu, Xiaoyu Wang, Hong-Xia Bao, Yu-Jie Zhou, Jun-Jie Kou, Dongsheng Li, Ting Li, Jianrui Liu, Yongfang Liu, Lin Yin, Li Qiu, Liling Gong & Shu-Lin Liu (2016) "SOX7 co-regulates Wnt/ $\beta$ -catenin signaling with Axin-2: both expressed at low levels in breast cancer", *Scientific Reports* 6, Article number: 26136 (2016) doi:10.1038/srep26136
14. Pappalardo, F., Calonaci, C., Pennisi, M., Mastriani, E., Motta, S. (2009) "HAMFAST: Fast Hamming Distance Computation", *IEEE World Congress on Computer Science and Information Engineering (CSIE 2009)*, 1, 569-572.
15. Pappalardo, F., Pennisi, M., Mastriani, E., Motta, S. (2009) "Grid-based atherosclerosis simulations", *Final Workshop of Grid Projects, PON Ricerca 2000-2006*, no. 1575, ISBN: ISBN: 978-88-95892-02-3, 403-410.
16. Pennisi, M., Pappalardo, F., Mastriani, E., Chiacchio, F., Motta, S. (2009) "A biological optimization problem on the Grid", *Final Workshop of Grid Projects, PON Ricerca 2000-2006*, no. 1575, ISBN: 978-88-95892-02-3, 403-410.
17. Alemani, D., Pappalardo, F., Pennisi, M., Mastriani, E., Motta, S. (2009) "Grid-based cancer growth simulations", *Final Workshop of Grid Projects, PON Ricerca 2000-2006*, Avviso 1575, pp 411-417, ISBN: 978-88-95892-02-3.
18. A. Pedretti, A. Lombardo, G. Vistoli, E. Mastriani, F. Pappalardo, S. Motta (2009), "Porting

- of a Docking Application on PI2S2 Grid infrastructure”, Final Workshop of Grid Projects, “PON Ricerca 2000-2006, Avviso 1575”, pp 468-473, ISBN: 978-88-95892-02-3.
19. Mastriani,E., Pappalardo,F., Motta,S. (2008) “PI2S2-Immunogrid Integration and Bioinformatics large scale simulations”, Grid Open Day, Engineer Faculty, University of Catania, May, 9th, Catania, Italy
  20. Pennisi M., Catanuto R., Mastriani E., Cincotti A., Pappalardo F., Motta S. (2008) “Optimization of Vaccine Schedules using Simulated Annealing”, Lecture Notes in Computer Science.
  21. Pappalardo,F., Gullo,F., Catanuto,R., Mastriani,E., Pennisi,M., Musumeci,S., Motta,S. (2008) “Agent based modeling of humoral response to atherogenesis”, Computational Intelligence Methods for Bioinformatics and Biostatistics (CIBB 2008), Lecture Notes in Bioinformatics, 5488, Springer . ISBN: 978-88-903537-1-0
  22. F. Pappalardo, S. Motta, P.-L. Lollini, E. Mastriani, M. Pennisi, (2007) “Towards a personalized schedule with Triplex vaccine”, Lecture Notes in Computer Science 4578, 620—626.
  23. F. Pappalardo, S. Motta, P.-L. Lollini, E. Mastriani. (2007) “Analysis of vaccine's schedules using models”, Cellular Immunology, 244, 137-140.
  24. Mastriani,E., MD.Halling-Brown, A. Shepherd, and S. Motta, (2007) “ImmunoGridPI2S2 integration Architecture”, Proceedings of the Symposium “GRID Open Days at the University of Palermo”, ISBN 978-88-95892-00-9, 127- 131.
  25. Mastriani,E., Pappalardo,F., Pennisi,M., Catanuto,R., Motta,S. (2007) “Tumor associated antigen variation - in silico experiments using the sicilian grid”, Proceedings of the Symposium “GRID Open Days at the University of Palermo”, ISBN 978-88-95892-00-9, 127-131.
  26. F. Pappalardo, S. Motta, P.-L. Lollini, E. Mastriani. (2007) “Analysis of vaccine's schedules using models”, Cellular Immunology, 244, 137-140. [IF 1.808]
  27. F. Pappalardo, E. Mastriani, P.-L. Lollini, S. Motta. (2006) “Analysis of vaccine's schedules using models”, Proceedings of International Congress of Immunogenomics and Immunomics (BCII 2006), October 8-12, 2006 – Budapest, Hungary, pp. 63-64.
  28. F. Pappalardo, E. Mastriani, P.-L. Lollini, S. Motta. (2006) “Analysis of vaccine's schedules using models”, Proceedings of International Congress of Immunogenomics and Immunomics 63-64.
  29. F. Pappalardo, P.-L. Lollini, S. Motta, E. Mastriani. (2006) “Model Deconstruction of an immunoprevention Vaccine”, Lecture Notes in Computer Science, 350—353.
  30. F. Pappalardo, S. Motta, P.-L. Lollini, E. Mastriani, M. Pennisi. (2006) “The Stabilization Effect of the Triplex Vaccine”, Proceedings of the 7<sup>th</sup> International FLINS Conference, Applied Artificial Intelligence, World Scientific, 587—592.
  31. F. Pappalardo, E. Mastriani, P.-L. Lollini, S. Motta. (2006) “Genetic Algorithm against Cancer”, Lecture Notes in Computer Science, 3849 223-- 228.
  32. F. Pappalardo, S. Motta, P.-L. Lollini, E. Mastriani and M.Pennisi (2006) “The Stabilization Effect of the Triplex Vaccine”, in proceedings of the Third International Meeting on Computational Intelligence Methods for Bioinformatics and Biostatistics , World Scientific
  33. F. Pappalardo, P.-L. Lollini, E. Mastriani and S. Motta. (2006) “Analysis of vaccine’s schedules using models”, in proceedings of IEEE World Congress on Computational Intelligence, s.s. Neural Networks Applications to Bioinformatics, IEEE press
  34. F. Pappalardo, P.-L. Lollini, E. Mastriani and S. Motta. (2006) “Model Deconstruction of an Immunoprevention Vaccine”, Lecture Notes in Computer Science

35. S. Motta, F. Pappalardo, E. Mastriani, P.-L. Lollini. (2006) "Cancer immunoprevention: what can we learn from in silico models?", Proceedings of International Congress of Immunogenomics and Immunomics 70.
36. F. Pappalardo, E. Mastriani, P.-L. Lollini, S. Motta. (2006) "Analysis of vaccine's schedules using models", Proceedings of International Congress of Immunogenomics and Immunomics 63-64.
37. F. Pappalardo, P.-L. Lollini, E. Mastriani and S. Motta (2005) "Genetic Algorithm against Cancer", in Proceedings of 2-th International Meeting on Computational Intelligence Methods for Bioinformatics and Biostatistics, Crema, IT, Italy, Lecture Notes in Computer Science, 3849 223–228
38. F. Pappalardo, P.-L. Lollini, S. Motta, E. Mastriani (2005) "Maths against Cancer", in Proceedings of 5-th International ISAAC Congress, July 25-30, Catania, Sicily, Italy
39. Motta, S., Lollini, P.-L., Pappalardo, F., Mastriani, E. (2005) "Modelling can help: an analysis of the effects of an immunoprevention vaccine", 4th World Congress on cellular and Molecular Biology, (2005), 136, ISSN:0145-5680
40. V. Cutello and F. Pappalardo, E. Mastriani (2002) "An evolutionary algorithm for the Minimum Hitting Set Problem", in Proceedings of the 3-rd WSEAS International Conference on Evolutionary Computation, Interlaken, Switzerland, pp. 277–281, [www.wseas.org](http://www.wseas.org)
41. V. Cutello and F. Pappalardo, E. Mastriani (2002) "An evolutionary algorithm for the T-constrained variation of the Minimum Hitting Set Problem", in Proceedings of the 2002 IEEE World Congress on Computational Intelligence, Honolulu, HI, USA, May 12-17, Vol. 1, pp. 366–371, [www.ieee.org](http://www.ieee.org).

#### **REVIEWER FOR INTERNATIONAL JOURNALS**

- *BMC Immunology*, ISSN/eISSN 1471-2172
- *JMIR Medical Informatics*, ISSN/eISSN 2291-9694
- *PLOS ONE*, ISSN/eISSN 1932-6203

#### **AREAS OF TEACHING EXPERTISE**

- *Bioinformatics*
- *Computational virology*
- *Machine learning*
- *Computer science*
- *Mathematics*

#### **CONFERENCE ORGANIZATION**

- *Organization of the training event "Grid School", September 21 to 30, 2009, Department of Engineering, University of Catania, Catania, Italy*
- *Organization of the training event "A day with the Grid", 21 October 008, University of Catania, Catania, Italy*

#### **TEACHING EXPERIENCE**

- *Active participation in training "Grid Update", 4-14 September 2009, National Institute of Nuclear Physics, Catania, Italy*

- *Training Course on Parallel Computing High Performance, 29 September-6 October 2008, University of Catania, Catania, Italy*
- *Tutorial on numerical methods for systems of high-performance parallel computing, 27-29 November 2007, Dept. of Physics, University of Catania, Italy*
- *Training Project GARR-B Course for network specialists, 16th November 2002 - 6 May 2003, CEA, Catania, Italy*

## **REFEREES**

- PhD, Shu-Lin Liu, Full Professor, Systemomics Center, College of Pharmacy, Genomics Research Center, State-Province Key Laboratories of Biomedicine-Pharmaceutics of China. Email: slliu@hrbmu.edu.cn
- PhD, Alexey V. Rakov, Senior Researcher, Institute of Epidemiology and Microbiology, 1 Selskaya Street, Vladivostok, Russia. Email: rakovalexey@gmail.com
- PhD, Francesco Pappalardo, Associate Professor of Computer Science, Teacher Referee for CdLM in Chemistry and Pharmaceutical Techniques, Department of Drug Sciences. Email: francesco.pappalardo@unict.it