

Arshan Nasir, Ph.D.

Department of Biosciences, COMSATS University Islamabad, Pakistan
arshan.nasir@gmail.com

EDUCATION

- Ph.D. Informatics** **2012-15**
Illinois Informatics Institute, University of Illinois at Urbana-Champaign, USA
Dissertation: *Structural and Functional Advances in the Evolutionary Studies of Cells and Viruses*
Committee: Dr. Gustavo Caetano-Anollés, Dr. Jay Mittenthal, Dr. Matthew Hudson, Dr. Jian Ma
CGPA: 3.94/4.0
- M.S. Bioinformatics** **2010-12**
Department of Crop Sciences, University of Illinois at Urbana-Champaign, USA
Thesis: *Origin of Viruses Revealed by the Genomic Study of Protein Domain Structures*
Committee: Dr. Gustavo Caetano-Anollés, Dr. Sanda Rodriguez-Zas, Dr. James Whitfield
CGPA: 3.96/4.0
- B.S. Bioinformatics** **2004-08**
COMSATS University Islamabad, Pakistan
Thesis: *HMTV, MMTV, and HERV-K In Silico Analysis: HMTV is Exogenous and Has Close Genetic Relationship with MMTV*
Committee: Dr. Farah Mustafa, Dr. Syed Habib Bokhari
CGPA: 3.73/4.0

EMPLOYMENT

- Assistant Professor in Bioinformatics** **2015-Present**
Department of Biosciences, COMSATS University Islamabad, Pakistan
Profile: <http://ww3.comsats.edu.pk/faculty/FacultyDetails.aspx?UId=625>
- Postdoctoral Research Associate** **2017-18**
Department of Crop Sciences, University of Illinois at Urbana-Champaign, USA

PUBLICATIONS (First Author = 14, *Senior Author = 4)

Articles Featured in Press

1. Malik SS, Azem-E-Zahra S, Kim KM, Caetano-Anollés G, **Nasir A***. 2017. Do viruses exchange genes across superkingdoms of life? **Frontiers in Microbiology** 8:2110.
 - <https://www.sciencedaily.com/releases/2017/12/171206122423.htm> (**Press Release**)
2. **Nasir A**, Caetano-Anollés G. 2015. A phylogenomic data-driven exploration of viral origins and evolution. **Science Advances** 1:1500527

- <http://news.discovery.com/animals/viruses-are-alive-and-are-oldest-living-creatures-150925.htm> (**Press Release**)
- 3. **Nasir A**, Forterre P, Kim KM, Caetano-Anollés G. 2014. The distribution and impact of viral lineages in domains of life. **Frontiers in Microbiology** 5:194
 - <http://blogs.scientificamerican.com/artful-amoeba/uneven-distribution-of-viruses-suggests-surprising-evolutionary-power/> (**Scientific American Blogs 2016**)
- 4. **Nasir A**, Kim KM, Caetano-Anollés G. 2012. Giant viruses coexisted with the cellular ancestors and represent a distinct supergroup along with superkingdoms Archaea, Bacteria and Eukarya. **BMC Evolutionary Biology** 12:156
 - <http://www.sciencedaily.com/releases/2012/09/120913123520.htm> (**Press Release**)

Software and Databases Developed

- 5. Jeong H, Sung S, Kwon T, Seo M, Caetano-Anollés K, Choi SH, Cho S, **Nasir A***, Kim H*. 2016. HGTree: database of horizontally transferred genes determined by tree reconciliation. **Nucleic Acids Research** 44(D1):D610-9
- 6. 5. Oh J, Choi CH, Park MK, Kim BK, Hwang K, Lee SH, Hong SG, **Nasir A**, Cho WS, Kim KM. 2016. CLUSTOM-CLOUD: In-Memory Data Grid-Based software for clustering 16S rRNA sequence data in the cloud environment. **PLoS ONE** 11:e0151064

Phylogenomics and Computational Biology of Viruses and Microorganisms

- 7. Da Cunha V, Gaia M, **Nasir A**, Forterre P. 2018. Asgard archaea do not close the debate about the universal tree of life topology. **PLOS Genetics**. 14: e1007215.
- 8. Jeong H, **Nasir A***. 2017. A preliminary list of horizontally transferred genes in prokaryotes determined by tree reconstruction and reconciliation. **Frontiers in Genetics** 8:112.
- 9. **Nasir A**, Kim KM, Caetano-Anollés G. 2017. Phylogenetic tracings of proteome size support the gradual accretion of protein structural domains and the early origin of viruses from primordial cells. **Frontiers in Microbiology** 8: 1178.
- 10. **Nasir A***, Kim KM, Caetano-Anollés G*. 2017. Long-term evolution of viruses: A Janus-faced balance. **Bioessays** 16: doi: 10.1002/bies.201700026.
- 11. Da Cunha V, Gaia M, Gadelle D, **Nasir A**, Forterre P. 2017. Lokiarchaea are close relatives of Euryarchaeota, not bridging the gap between prokaryotes and eukaryotes. **PLoS Genetics** 13:e1006810
- 12. **Nasir A**, Caetano-Anollés G. 2017. Identification of capsid/coat related protein folds and their utility for virus classification. **Frontiers in Microbiology** 8:380.

13. **Nasir A**, Kim KM, Da Cunha V, Caetano-Anollés G. 2016. Arguments reinforcing the three-domain view of diversified cellular life. **Archaea**. 2016:1851865
14. **Nasir A**, Kim KM, Caetano-Anollés. 2015. Lokiarchaeota: eukaryote-like missing links from the microbial dark matter? **Trends in Microbiology** 23:448-50
15. **Nasir A**, Sun FJ, Kim KM, Caetano-Anollés G. 2015. Untangling the origin of viruses and their impact on cellular evolution. **Annals of the New York Academy of Science** 1341:61-74
16. **Nasir A**, Kim KM, Caetano-Anollés G. 2014. Global patterns of protein domain gain and loss in superkingdoms. **PLoS Computational Biology** 10: e1003452
17. **Nasir A**, Kim KM, Caetano-Anollés G. 2014. A phylogenomic census of molecular functions identifies modern thermophilic archaea as the most ancient form of cellular life. **Archaea** 2014:706468
18. Kim KM, **Nasir A**, Hwang K, Caetano-Anollés G. 2014. A tree of cellular life inferred from a genomic census of molecular functions. **Journal of Molecular Evolution** 79: 240-62
19. Kim KM, **Nasir A**, Caetano-Anollés G. 2014. The importance of using realistic evolutionary models for retrodicting proteomes. **Biochimie** 99:129-137
20. Caetano-Anollés G, **Nasir A**, Zhou K, Caetano-Anollés D, Mittenthal JE, Sun FJ, Kim KM. 2014. Archaea: the first domain of diversified life. **Archaea** 2014: 590214
21. **Nasir A**, Caetano-Anollés G. 2013. Comparative analysis of proteomes and functionomes provides insights into origins of cellular diversification. **Archaea** 2013:648746
22. Yafremava LS, Wielgos M, Thomas S, **Nasir A**, Wang M, Mittenthal JE, Caetano-Anollés, G. 2013. A general framework of persistence strategies for biological systems helps explain domains of life. **Frontiers in Genetics** 4:16
23. **Nasir A**, Kim KM, Caetano-Anollés G. 2012. Viral evolution: primordial cellular origins and late adaptation to parasitism. **Mobile Genetic Elements** 2:247-252
24. Caetano-Anollés G, **Nasir A**. 2012. Benefits of using molecular structure and abundance in phylogenomic analysis. **Frontiers in Genetics** 3:172
25. **Nasir A**, Naeem A, Khan MJ, Nicora DH, Caetano-Anollés G. 2011. Annotation of protein domains reveals remarkable conservation in the functional make up of proteomes across superkingdoms. **Genes** 2:869-911

Genome Sequencing Projects

26. Kim KM, Choe H, Kim BK, **Nasir A**. 2017. Complete genome of a metabolically-diverse marine bacterium *Shewanella japonica* KCTC 22435^T. **Marine Genomics** S1874-7787(17)30117-4.
27. Choe H, Lee SH, Kim SG, Park DS, **Nasir A**, Kim KM. 2016. Complete genome of *Pseudoalteromonas phenolica* KCTC 12086 T (= O-BC30 T), a marine bacterium producing polybrominated aromatic compounds. **Journal of Biotechnology** 218:23-24
28. Lee SH, Choe H, Bae KS, Park DS, **Nasir A**, Kim KM. 2016. Complete genome of *Streptomyces hygroscopicus* subsp. *limoneus* KCTC 1717 (= KCCM 11405), a soil bacterium producing validamycin and diverse secondary metabolites. **Journal of Biotechnology** 219:1-2
29. Lee SH, Choe H, Kim SG, Park DS, **Nasir A**, Kim BK, Kim KM. 2016. Complete genome of brown algal polysaccharides-degrading *Pseudoalteromonas issachenkonii* KCTC 12958 T (= KMM 3549 T). **Journal of Biotechnology** 219:86-87
30. Lee SH, Choe H, Kim SG, Park DS, **Nasir A**, Kim BK, Kim KM. 2016. Complete genome of biodegradable plastics-decomposing *Roseateles depolymerans* KCTC 42856 T (= 61A T). **Journal of Biotechnology** 220:47-8
31. Lee SH, Choe H, **Nasir A**, Park DS, Kim KM. 2016. Complete genome sequence of nitrilotriacetate-degrading *Aminobacter aminovorans* KuCTC 2477T. **Genome Announcements** 4: e01363-16.
32. Choe H, Kim S, Oh J, **Nasir A**, Kim BK, Kim, KM. 2015. Complete genome of *Kangiella geojedonensis* KCTC 23420 T, putative evidence for recent genome reduction in marine environments. **Marine Genomics** 24: 215-217.
33. Lee SH, Choe H, Kim BK, **Nasir A**, Kim KM. 2015. Complete genome of the marine bacterium *Wenzhouxiangella marina* KCTC 42284(T). **Marine Genomics** 3:277-80.

BOOK CHAPTERS

1. Caetano-Anollés G, Minhas BF, Aziz F, Mughal F, Shahzad K, Tal G, Mittenthal JE, Caetano-Anollés D, Koc I, **Nasir A**, Caetano-Anollés K, Kim KM. 2017. The Compressed Vocabulary of the Proteins of Archaea. In: Witzany G. (eds). **Biocommunication of Archaea**. Springer, Cham pp. 147-174.
2. Hameed H, **Nasir A***. 2017. Woese's three domains of cellular life. In: Vonk J, Shackelford T. (eds.). **Encyclopedia of Animal Cognition and Behavior**. Springer International Publishing. pp. 1-3.

GRANTS & FUNDING

uBiome, USA (PI, \$76,000 worth research costs)

2016-18

Characterizing the gastrointestinal, oral, and skin microbiota in healthy Pakistani adults: A metagenomic and microbiome-centric approach towards next generation antimicrobial therapies.

Higher Education Commission, Pakistan (Co-PI, \$5,000) **2016-18**
Functional metagenomics approach to characterize antibiotic resistance genes in soil microbiota.

Higher Education Commission, Pakistan (PI, \$4,500) **2015-17**
The concept of superorganisms: A case study of humans and their microbiomes.

Higher Education Commission, Pakistan (Co-PI, \$4,900) **2015-17**
Identification of genes with conserved *cis*-acting polyadenylation elements across various eukaryotic species.

ACADEMIC HONORS & AWARDS

Best Research Paper Award **2017**
Jeong et al. *Nucleic Acids Research* 44:D610-9 awarded best Research Paper by the Higher Education Commission (HEC), Pakistan. The competition involved all researchers and all universities in Pakistan (52 awardees out of 747 shortlisted across all country).

Dissertation Completion Fellowship **2014**
Awarded by the Graduate College of the University of Illinois at Urbana-Champaign to outstanding PhD students in their final year to support dissertation completion and defense. Awardees are selected after a campus-wide competition involving all departments and colleges. In the award year, 29 out of 76 shortlisted students from all campuses and colleges were awarded.

Chateaubriand Fellowship **2013**
A grant offered by the Embassy of France in the United States. It supports outstanding PhD students from US universities who wish to conduct research in France for a period ranging from 4-9 months.
<https://www.chateaubriand-fellowship.org/-Science-Technology-Engineering-Math-Health-.html>

BioMed Central Research Award **2012**
Competition involved ~20,000 articles published by >200 open access journals in 2012. Our manuscript (Nasir et al. 2012. *BMC Evol Biol.* 12:156) was one of the two runners-up (out of ~20,000 articles that were considered).

International Seeds Grant Program **2012**
One-month research internship at the Korea Research Institute of Bioscience and Biotechnology, Korea

List of Teachers Ranked as Excellent by Their Students **2012**
Human Anatomy and Physiology II (MCB 247)

NSF Catalyzing New International Collaborations Award **2011**

Fully-funded three-month research internship at the Korea Research Institute of Bioscience and Biotechnology, Korea

Faculty Development Program Fellowship

2010-12

Awarded by COMSATS University Islamabad, Pakistan to pursue MS at the University of Illinois at Urbana-Champaign

Institute Silver and Campus Bronze medals

2009

Awarded by COMSATS University Islamabad, Pakistan for undergraduate CGPA > 3.50

ACADEMIC SERVICE

Guest Associate Editor

2018

Special Topic on “**Viruses, Genetic Exchange, and Tree of Life**” alongside Professors Gustavo Caetano-Anollés and Jean-Michel Claverie be published simultaneously in *Frontiers in Microbiology*, *Frontiers in Genetics*, and *Frontiers in Plant Science*

<https://www.frontiersin.org/research-topics/7867/viruses-genetic-exchange-and-the-tree-of-life>

Managing Editor

2018

Lifetime Issue on *Origin and Evolution of the Viruses* in *Frontiers in Biosciences*

http://www.bioscience.org/special-issue-details?editor_id=1569

Papers Reviewed

Molecular Biology and Evolution, *PLoS Genetics*, *PLoS ONE*, *Evolution*, *Journal of Molecular Evolution*, *Virology*, and *(Nature) Scientific Reports*

PRESS COVERAGE (SELECTED)

Popular Science: <https://www.popsci.com/new-evidence-that-viruses-are-alive>

Scientific American: <http://blogs.scientificamerican.com/artful-amoeba/uneven-distribution-of-viruses-suggests-surprising-evolutionary-power/>

COSMOS (Australia): <https://cosmosmagazine.com/biology/what-came-first-cells-or-viruses>

Der Spiegel (Germany): <http://www.spiegel.de/wissenschaft/natur/grosse-viren-sind-doch-lebewesen-a-856204.html>

Science and Life (Russia): <https://www.nkj.ru/news/27078/>

Wired (UK): <http://www.wired.co.uk/article/viruses-are-living-things-that-share-common-ancestry-with-cells>

La Scienze (Italy):

http://www.lescienze.it/news/2015/09/30/news/virus_esseri_viventi_ripiegamento_proteine-2782780/

Small Things Considered: <http://schaechter.asmblog.org/schaechter/2016/01/pictures-considered-33-viral-abundance-and-diversity-.html>

Science Daily: <http://www.sciencedaily.com/releases/2012/09/120913123520.htm>

Chicago Tonight on WTTW: <https://chicagotonight.wttw.com/2012/10/16/scientific-chicago-neil-shubin>

RESEARCH TRAINING

Graduate Exchange Student

Spring 2014 & 15

Institute of Genetics and Microbiology, University of Paris-Sud, Paris, France

Advisor: Prof. Patrick Forterre

Project: Generating an in-depth understanding of virus host preferences (Nasir et al. *Frontiers in Microbiology* 2014, 5: 194)

Graduate Exchange Student

Summer 2011 & Fall 2012

Korea Bioinformation Center, Korea Research Institute of Bioscience and Biotechnology

Advisor: Dr. Kyung Mo Kim

Project: A phylogenomic exercise to calculate protein domain gains versus losses on the branches of a tree of life (Nasir et al. *PLoS Computational Biology* 2014, 10: e1003452)

Undergraduate Exchange Student

Fall 2009

Department of Crop Science, University of Illinois at Urbana-Champaign

Advisor: Prof. Gustavo Caetano-Anollés

Project description: Algorithm development to speed up the process of assigning structures to protein sequences using Python under UNIX environment

TEACHING

Instructor, Introduction to Bioinformatics (BIO 310)

2015-present

Department of Biosciences, COMSATS University Islamabad

- Classes involve both lecture and laboratory components. Course contents included bioinformatics sequence and structure analysis, phylogenetics, and metagenome annotation

Teaching Assistant, Human Anatomy and Physiology II

2012

School of Molecular and Cellular Biology, University of Illinois at Urbana-Champaign

- Taught human anatomy using anatomical models, cadavers, and power points to undergraduates

MENTORING

Class Counselor**2015-17**

Department of Biosciences, COMSATS University Islamabad

- Provided career counseling to freshman and junior level undergraduate students enrolled into the BS Bioinformatics program at the COMSATS Institute

Team Leader, Summer Research Opportunities Program**Summer 2013 & 14**

The Graduate College, University of Illinois at Urbana-Champaign

- Assisted 10-12 undergraduate students in successful completion of their summer research projects. Served as conduit between the faculty mentors and students. Designed and organized weekly research team meetings to hold discussions on various topics in science. Weekly class on teaching research writing skills.

THESIS SUPERVISION

MS Thesis Supervision**Fall 2016, 17**

Department of Biosciences, COMSATS University Islamabad

- Maria Batool, *Characterization of the gastrointestinal and oral microbiota in Pakistani adults*
- Bushra Arif, *What can protein structure tell us about the host-microbiota relationships?*

BS Thesis Supervision**2014-16**

Department of Biosciences, COMSATS University Islamabad

- Almeerah Nadeem, *Comparative analysis of the quality of sequenced genome assemblies in archaea*
- Sadia Farooq, *Comparative analysis of the quality of sequenced genome assemblies in bacteria*
- Momal Aziz, *Protein Alignment Inspector: A software to evaluate protein domain make-up prior to phylogenetic reconstruction*
- Hira Hameed, *A mapping of viral protein folds onto a phylogeny derived from protein structure and abundance*
- Syeda Azem-e-Zahra and Shahana Seher Malik, *A bioinformatics exploration of the viral genomic dark matter*
- Rahma Shahid, *The greater genomic relationships of Archaea: reconstructing the archaeal hologenome*
- Aisha Shafqat and Shehneez Zulfiqar, *Virus-host interactions: A case-study of human viruses*

UNIVERSITY SERVICE

Department of Biosciences, COMSATS University Islamabad

- Co-developed '**Biological Data Analysis**' course for undergraduate and graduate students enrolled into the various MS and BS level programs offered by the Department of Biosciences.

- Member of *departmental Event Management Committee*. Organized seminars, departmental tours, and workshops along with other committee members.
- *Orientation Day Committee*
- Member of *Departmental Undergraduate Program Committee*
- Member of *Students' Research Quality Committee*

SEMINARS & PRESENTATIONS

Invited Talks

“What can informatics tell us about viral evolution?” Department of Biological Sciences, Auburn University, Auburn, AL, March 3, 2016 **2016**

“Structural and Functional Advances in the Evolutionary Studies of Cells and Viruses”, Lahore University of Management Sciences, Lahore, Pakistan, November 18, 2016 **2015**

“What can informatics tell us about viral evolution?” 13th International Conference on Frontiers of Information Technology Conference, Islamabad, December 14-16, Pakistan **2015**

Poster Presentations

“A Data-Driven Approach to Study Virus Origins and Evolution” 4th World Congress on Virology, San Antonio, TX, October 6-8, 2014

“Origin and Evolution of the Viral Supergroup”, SMBE 2013, Annual Society for Molecular Biology and Evolution (SMBE) meeting, Chicago, IL, July 7-11, 2013

“Annotation of protein domains reveals remarkable conservation in the functional make up of proteomes”, IEEE International Conference on Bioinformatics and Biomedicine, Atlanta, GA, November 12-15, 2011

Panelist

Modern Trends in Bioscience and Bioinformatics, COMSATS Institute of Information Technology, Islamabad, Pakistan, March 22nd, 2016

Participant

International Conference on *Frontiers of Information Technology*, COMSATS Institute of Information Technology, Abbottabad, Pakistan, December 16-18, 2009

9th Biennial PSBMB Conference on *Advances in Biochemistry and Molecular Biology*, PMAS Arid Agriculture University, Rawalpindi, Pakistan, December 17-20, 2008

ACADEMIC & TECHNICAL SKILLS

Coursework Completed

2010-15

University of Illinois at Urbana-Champaign

- Applied Statistical Methods I (CPSC 440), Applied Statistical Methods II (CPSC 542), Regression Analysis (CPSC 541), Advanced Data Analysis (STAT 448), Database Systems (CS 411), Statistical Genomics (ANSC 545), Bioinformatics and Systems Biology (CPSC 567), Perl & Unix for Bioinformatics (CPSC 565), Computational Cancer Biology (BIOE 598), Genetics of Higher Organisms (CPSC 452), Applied Bioinformatics (CPSC 569), Virology and Viral Pathogenesis (MCB 433)

Programming Languages

- SAS, SPSS, Python, R, MySQL, PAUP*

REFERENCES

Dr. Gustavo Caetano-Anollés, Professor of Bioinformatics, Department of Crop Science, UIUC
Ph: +1-217-333-8172, *Email:* gca@illinois.edu

Dr. Patrick Forterre, Professor, Pasteur Institute, Paris, France
Ph: +33 01 45 68 87 91, *Email:* patrick.forterre@pasteur.fr

Dr. Jay E Mittenthal, Associate Professor of Cell and Developmental Biology, Emeritus, The School of Molecular and Cellular Biology, UIUC
Ph: +1-217- 333-4846, *Email:* mitten@life.illinois.edu

Dr. Kyung Mo Kim, Senior Research Scientist, Division of Polar Life Sciences, Korea Polar Research Institute, Republic of Korea
Ph: +82-10-5022-8992, *Email:* kmkim@kopri.re.kr