

Pressemitteilung

Johannes Gutenberg-Universität Mainz

Kathrin Voigt

17.04.2025

<http://idw-online.de/de/news850901>

Forschungsergebnisse, Wissenschaftliche Publikationen
Biologie, Informationstechnik
überregional



JOHANNES GUTENBERG
UNIVERSITÄT MAINZ

The eukaryotic cell emerged as an evolutionary algorithmic phase transition

Eukaryogenesis happened abruptly, led by the tension of increasing gene length and the constraints on producing longer proteins

An international collaboration between four senior scientists from Mainz, Valencia, Madrid, and Zurich has published groundbreaking research in the journal PNAS, shedding light on the most significant increase in complexity in the history of life's evolution on Earth: the origin of the eukaryotic cell. While the endosymbiotic theory is widely accepted, the billions of years that have passed since the fusion of an Archaea and a Bacteria have resulted in a lack of evolutionary intermediates in the phylogenetic tree until the emergence of the eukaryotic cell. It is a gap in our knowledge, referred to as the black hole at the heart of biology. "The new study is a blend of theoretical and observational approaches that quantitatively understands how the genetic architecture of life was transformed to allow such an increase in complexity", stated Dr. Enrique M. Muro, representative of Johannes Gutenberg University Mainz (JGU) in this project.

Proteins and protein coding genes increase in length

The article in PNAS demonstrates that the distributions of protein lengths and their corresponding genes follow log-normal distributions across the whole tree of life. To do this, 9,913 different proteomes and 33,627 genomes were analyzed. Log-normal distributions typically arise as a result of multiplicative processes. Following Ockham's razor principle, the researchers modeled the evolution of gene length distributions as multiplicative stochastic processes. In fact, they modeled the action of all genetic operators combined in relation to sequence length. Starting from LUCA, i.e., the hypothesized last universal common ancestor from which the three domains of life – the Bacteria, the Archaea, and the Eukarya – originated, the researchers found both theoretically and observationally that the average gene lengths have evolved exponentially over evolutionary time across different species. Furthermore, they discovered a scaling-invariant mechanism of gene growth across the entire tree of life, where the variance directly depends on the mean protein length. By representing all the species captured in the 33,627 genomes, the team was able to observationally verify the predictions and, moreover, show that the average gene length is a very good surrogate for organismal complexity. In a pure exercise of quantitative biology, Dr. Bartolo Luque from the Polytechnic University of Madrid added: "From knowing the average length of protein-coding genes in a species, we can calculate the whole distribution of gene length within that species."

When representing the evolution of the average protein lengths versus their corresponding gene lengths across different species, it is observed that they evolve simultaneously in prokaryotes, because there are almost no non-coding sequences in their genes. However, once the average gene length reaches 1,500 nucleotides, the proteins decouple from the multiplicative process of gene growth, and the average protein length stabilizes after the onset of the eukaryotic cell at about 500 amino acids in a clear threshold, marking the appearance of the eukaryotic cell. From that point onward, and unlike what happens with proteins, the average gene length continues to increase as it did in prokaryotes, due to the presence of non-coding sequences.

Algorithmic phase transition

A critical phenomena analysis then concluded that a phase transition, well studied in the physics of magnetic materials, occurred at a critical gene length of 1,500 nucleotides. This marked eukaryogenesis and divides the evolution of life into two distinct phases: a coding phase (Prokarya) and a non-coding phase (Eukarya). Additionally, characteristic phenomena of these transitions are observed, such as critical slowing down, where the system's dynamics become trapped in many metastable states around the critical point. "This is corroborated in early protists and fungi", said Dr. Fernando Ballesteros from the University of Valencia.

Moreover, "the phase transition was algorithmic," added Professor Jordi Bascompte from the University of Zurich. In the coding phase, in a scenario close to LUCA, with short proteins, increasing the length of proteins and their corresponding genes was computationally simple. However, as the protein lengths grew, the search for longer proteins became unfeasible. This tension caused by genes that grew at the same rate as before while proteins could not be resolved continuously but abruptly with the incorporation of non-coding sequences into the genes. With this innovation, the algorithm for searching for new proteins rapidly reduced its computational complexity, becoming non-linear through the spliceosome and the nucleus, which separated transcription and splicing from translation. This happened at the critical point of phase transition, which this study dates to 2.6 billion years ago.

The study recently published in PNAS not only answers essential questions, but is interdisciplinary, combining computational biology, evolutionary biology, and physics. "It has the potential to interest a wide audience across many disciplines and serve as a foundation for other groups to explore different research avenues, such as energy or information theory", emphasized Dr. Enrique Muro of the Institute of Organismic and Molecular Evolution at Mainz University. The eukaryotic cell, the most significant increase in complexity in the history of life's evolution on Earth, emerged as a phase transition and unlocked the path toward other major transitions – such as multicellularity, sexuality, and sociability – that shaped life on our planet as we know it today.

Images:

https://download.uni-mainz.de/presse/10_iome_eukaryoten.01.jpg

Dr. Enrique Muro of Mainz University showing one of the pieces of evidence of the phase transition using the research team's work repository on GitHub
photo/©: Juan Mac Donagh

https://download.uni-mainz.de/presse/10_iome_eukaryoten.02.jpg

Evolution of protein-coding gene length distributions across evolutionary time
ill./©: Fernando J. Ballesteros

Related links:

- <https://cbdm-01.zdv.uni-mainz.de/~muro/> – Genome and Proteome Evolution research group of Dr. Enrique Muro at the JGU Institute of Organismic and Molecular Evolution (IOME)
- <https://cbdm.uni-mainz.de/> – Computational Biology and Data Mining research group of Professor Miguel Andrade at the JGU Institute of Organismic and Molecular Evolution (IOME)
- <https://iome.biology.uni-mainz.de/> – Institute of Organismic and Molecular Evolution (IOME) at JGU
- <https://www.blogs.uni-mainz.de/fb10-biologie-eng/> – Faculty of Biology at JGU
- <https://iqcb.uni-mainz.de/> – Institute of Quantitative and Computational Biosciences (IQCB) at JGU

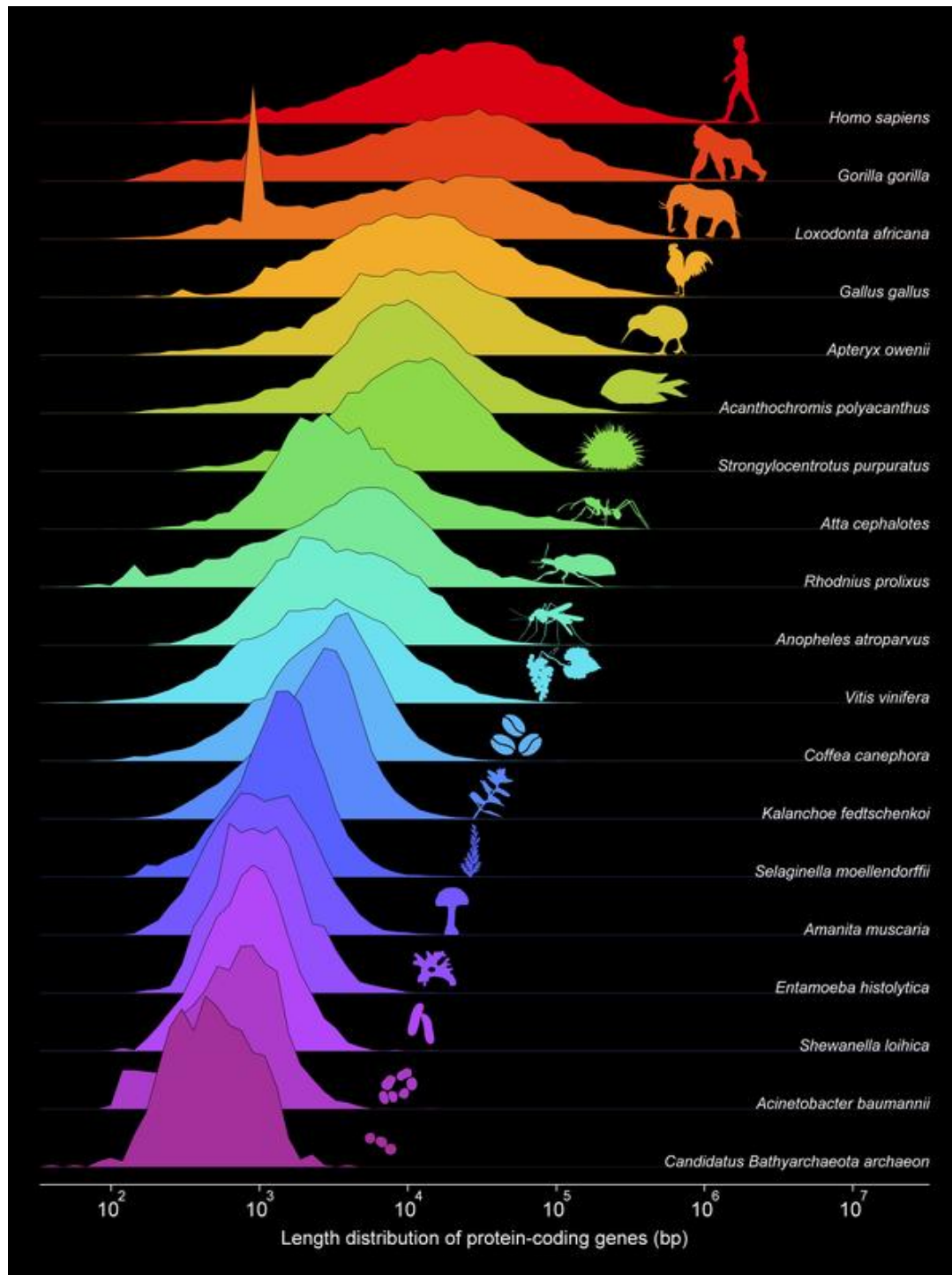
wissenschaftliche Ansprechpartner:

Dr. Enrique M. Muro
Principal Investigator (tenured)

Institute of Organismic and Molecular Evolution (IOME)
Johannes Gutenberg University Mainz
55099 Mainz, GERMANY
phone: +49 6131 39-21581
e-mail: muro@uni-mainz.de
<https://cbdm-01.zdv.uni-mainz.de/~muro/>

Originalpublikation:

Enrique M Muro, Fernando J Ballesteros, Bartolo Luque, and Jordi Bascompte, The emergence of eukaryotes as an evolutionary algorithmic phase transition, Proceedings of the National Academy of Sciences (PNAS) 122 (13): e2422968122, 27 March 2025,
DOI: [10.1073/pnas.2422968122](https://doi.org/10.1073/pnas.2422968122)
<https://www.pnas.org/doi/10.1073/pnas.2422968122>



Evolution of protein-coding gene length distributions across evolutionary time
ill./©: Fernando J. Ballesteros



Dr. Enrique Muro of Mainz University showing one of the pieces of evidence of the phase transition using the research team's work repository on GitHub
photo/©: Juan Mac Donagh