МОЛЕКУЛЯРНАЯ БИОЛОГИЯ, 2023, том 57, № 1, с. 98–100

ГЕНОМИКА. ТРАНСКРИПТОМИКА

УДК 575.852'113

НЕОБЫЧНАЯ ЗАВИСИМОСТЬ МЕЖДУ ЭКСПРЕССИЕЙ ГЕНОВ И ОТРИЦАТЕЛЬНЫМ ОТБОРОМ У ИНФУЗОРИЙ РОДА *Euplotes*[1]

М. А. Молдован[a, *], **С. А. Гайдукова**[b]

[a] *Сколковский институт науки и технологий, Москва, 121205 Россия*

[b] *Факультет биоинженерии и биоинформатики, Московский государственный университет им. М.В. Ломоносова, Москва, 199911 Россия*

**e-mail: mika.moldovan@gmail.com*

Поступила в редакцию 13.09.2022 г.
После доработки 13.09.2022 г.
Принята к публикации 13.09.2022 г.

У большинства исследованных организмов экспрессия генов связана с рядом эволюционных особенностей, относящихся к белоккодирующим последовательностям. В частности, экспрессия генов положительно коррелирует со средней интенсивностью отрицательного отбора и влияет на использование кодонов. Нами изучена связь между экспрессией генов и паттернами отбора у двух видов ресничных простейших рода *Euplotes*. Мы обнаружили, что на использование кодонов влияет экспрессия генов в этих организмах, что указывает на дополнительные эволюционные ограничения на возникновение мутаций в сильно экспрессируемых генах по сравнению с генами, экспрессируемыми с меньшей скоростью. В то же время, на уровне синонимичных и несинонимичных замен мы наблюдаем более сильное ограничение на гены, экспрессируемые с более низкой скоростью, по сравнению с генами с более высокой скоростью экспрессии. Наше исследование дополняет дискуссию об общих закономерностях эволюции и ставит новые вопросы о механизмах контроля экспрессии генов у инфузорий.

Ключевые слова: экспрессия генов, отрицательный отбор, инфузории, ресничные простейшие, *Euplotes*, эволюция белоккодирующих последовательностей

DOI: 10.31857/S002689842301010X, **EDN:** AWYZOB

REFERENCES

1. O'Connor L.J., Schoech A.P., Hormozdiari F., Gazal S., Patterson N., Price A.L. (2019) Extreme polygenicity of complex traits is explained by negative selection. *Am. J. Hum. Genet.* **105**, 456–476.
2. Sella G., Barton N.H. (2019) Thinking About the evolution of complex traits in the era of genome-wide association studies. *Annu. Rev. Genom. Hum. Genet.* **20**, 461–493.
3. Vitsios D., Dhindsa R.S., Middleton L., Gussow A.B., Petrovski S. (2021) Prioritizing non-coding regions based on human genomic constraint and sequence context with deep learning. *Nat. Commun.* **12**, 1504.
4. Adzhubei I.A., Schmidt S., Peshkin L., Ramensky V.E., Gerasimova A., Bork P., Kondrashov A.S., Sunyaev S.R. (2010) A method and server for predicting damaging missense mutations. *Nat. Methods.* **7**, 248–249.
5. Crick F.H.C., Barnett L., Brener S., Watts-Tobin R.J. (1961) General nature of the genetic code for proteins. *Nature.* **192**, 1227–1232.
6. Kimura M. (1983) The neutral theory of molecular evolution. *Cambridge University Press.*
7. Kimura M. (1977) Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. *Nature.* **267**, 275–276.
8. Kryazhimskiy S., Plotkin J.B. (2008) The population genetics of dN/dS. *PLoS Genet.* **4**, e1000304.
9. Yang Z., Bielawski J.P. (2000) Statistical methods for detecting molecular adaptation. *Trends Ecol. Evol.* **15**, 496–503.
10. Zhang J., Yang J.R. (2015) Determinants of the rate of protein sequence evolution. *Nat. Rev. Genet.* **16**, 409–420.
11. Jansen R. (2003) Revisiting the codon adaptation index from a whole-genome perspective: analyzing the relationship between gene expression and codon occurrence in yeast using a variety of models. *Nucleic Acids Res.* **31**, 2242–2251.

---

[1] Статья представлена авторами на английском языке.
Статья рекомендована к печати М.С. Гельфандом.

12. Sharp P.M., Li W.-H. (1987) The codon adaptation index-a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* **15**, 1281–1295.
13. Seplyarskiy V.B., Soldatov R.A., Koch E., McGinty R.J., Goldmann J.M., Hernandez R.D., Barnes K., Correa A., Burchard E.G., Ellinor P.T., McGarvey S.T., Mitchell B.D., Vasan R.S., Redline S., Silverman E., Weiss S.T., Arnett D.K., Blangero J., Boerwinkle E., He J., Montgomery C., Rao D.C., Rotter J.I., Taylor K.D., Brody J.A., Chen Y.-D.I., de Las Fuentes L., Hwu C.-M., Rich S.S., Manichaikul A.W., Mychaleckyj J.C., Palmer N.D., Smith J.A., Kardia S.L.R., Peyser P.A., Bielak L.F., D O'Connor T., Emery L.S., NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium, TOPMed Population Genetics Working Group, Gilissen C., Wong W.S.W., Kharchenko P.V., Sunyaev S.R. (2021) Population sequencing data reveal a compendium of mutational processes in the human germ line. *Science*. **373**, 1030–1035.
14. Pál C., Papp B., Hurst L.D. (2001) Highly expressed genes in yeast evolve slowly. *Genetics*. **158**, 927–931.
15. Rocha E.P.C., Danchin A. (2004) An analysis of determinants of amino acids substitution rates in bacterial proteins. *Mol. Biol. Evol.* **21**, 108–116.
16. Subramanian S., Kumar S. (2004) Gene expression intensity shapes evolutionary rates of the proteins encoded by the vertebrate genome. *Genetics*. **168**, 373–381.
17. Tuller T., Kupiec M., Ruppin E. (2008) Evolutionary rate and gene expression across different brain regions. *Genome Biol.* **9**, R142.
18. Marais G., Domazet-Lošo T., Tautz D., Charlesworth B. (2004) Correlated evolution of synonymous and nonsynonymous sites in drosophila. *J. Mol. Evol.* **59**, 771–779.
19. Hodgins K.A., Yeaman S., Nurkowski K.A., Rieseberg L.H., Aitken S.N. (2016) Expression divergence is correlated with sequence evolution but not positive selection in conifers. *Mol. Biol. Evol.* **33**, 1502–1516.
20. Popescu C.E., Borza T., Bielawski J.P., Lee R.W. (2006) Evolutionary rates and expression level in *Chlamydomonas*. *Genetics*. **172**, 1567–1576.
21. Gout J.-F., Kahn D., Duret L. (2010) The relationship among gene expression, the evolution of gene dosage, and the rate of protein evolution. *PLoS Genet.* **6**, e1000944.
22. Cherry J.L. (2010) Expression level, evolutionary rate, and the cost of expression. *Genome Biol. Evol.* **2**, 757–769.
23. Wei C., Chen Y.-M., Chen Y., Qian W. (2021) The missing expression level–evolutionary rate anticorrelation in viruses does not support protein function as a main constraint on sequence evolution. *Genome Biol. Evol.* **13(4)**, evab049.
24. Feyertag F., Berninsone P.M., Alvarez-Ponce D. (2016) Secreted proteins defy the expression level–evolutionary rate anticorrelation. *Mol. Biol. Evol.* **34(3)**, 692–706.
25. Feyertag F., Berninsone P.M., Alvarez-Ponce D. (2019) N-glycoproteins exhibit a positive expression level–evolutionary rate correlation. *J. Evol. Biol.* **32**, 390–394.
26. Rzeszutek I., Maurer-Alcalá X.X., Nowacki M. (2020) Programmed genome rearrangements in ciliates. *Cell. Mol. Life Sci.* **77**, 4615–4629.
27. Prescott D.M. (1994) The DNA of ciliated protozoa. *Microbiol. Rev.* **58**, 233–267.
28. Mozzicafreddo M., Pucciarelli S., Swart E.C., Piersanti A., Emmerich C., Migliorelli G., Ballarini P., Miceli C. (2021) The macronuclear genome of the Antarctic psychrophilic marine ciliate *Euplotes focardii* reveals new insights on molecular cold adaptation. *Sci. Rep.* **11**(1), 18782.
29. Lobanov A.V., Heaphy S.M., Turanov A.A., Gerashchenko M.V., Pucciarelli S., Devaraj R.R., Xie F., Petyuk V.A., Smith R.D., Klobutcher L.A., Atkins J.F., Miceli C., Hatfield D.L., Baranov P.V., Gladyshev V.N. (2016) Position-dependent termination and widespread obligatory frameshifting in *Euplotes* translation. *Nat. Struct. Mol. Biol.* **24**, 61–68.
30. Grabherr M.G., Haas B.J., Yassour M., Levin J.Z., Thompson D.A., Amit I., Adiconis X., Fan L., Raychowdhury R., Zeng Q., Chen Z., Mauceli E., Hacohen N., Gnirke A., Rhind N., di Palma F., Birren B.W., Nusbaum C., Lindblad-Toh K., Friedman N., Regev A. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652.
31. Bray N.L., Pimentel H., Melsted P., Pachter L. (2016) Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol.* **34**, 525–527.
32. Altschul S.F., Gish W., Miller W., Myers E.W., Lipman D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.
33. Abascal F., Zardoya R., Telford M.J. (2010) TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Res.* **38**, W7–13.
34. Nei M., Gojobori T. (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **5**, 418-426.
35. Nielsen R., Yang Z. (1998) Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics*. **148**, 929–936.
36. Yang Z. (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591.
37. Dana A., Tuller T. (2014) The effect of tRNA levels on decoding times of mRNA codons. *Nucleic Acids Res.* **42**, 9171–9181.
38. Zhou Z., Dang Y., Zhou M., Li L., Yu C., Fu J., Chen S., Liu Y. (2016) Codon usage is an important determinant of gene expression levels largely through its effects on transcription. *Proc. Natl. Acad. Sci. USA.* **113**(41), E6117-E6125.
39. Akashi H., Gojobori T. (2002) Metabolic efficiency and amino acid composition in the proteomes of *Escherichia coli* and *Bacillus subtilis*. *Proc. Natl. Acad. Sci. USA.* **99**, 3695–3700.
40. Lawrence J.R., Snyder R.A. (1998) Feeding behaviour and grazing impacts of a *Euplotes* sp. on attached bacteria. *Can. J. Microbiol.* **44**, 623–629.
41. Wiackowski K., Fyda J., Ciecko A. (2004) The behaviour of an omnivorous protozoan affects the extent of induced morphological defence in a protozoan prey. *Freshwater Biol.* **49**, 801–809.

Unusual Dependence between Gene Expression and Negative Selection in *Euplotes*

M. A. Moldovan[1, *] and S. A. Gaydukova[2]
[1]Skolkovo Institute of Science and Technology, Moscow, 121205 Russia
[2]Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University, Moscow, 199911 Russia
**e-mail: mika.moldovan@gmail.com*

In most of the studied organisms, gene expression is associated with a number of evolutionary features pertaining to the protein-coding sequences. In particular, gene expression positively correlates with the average intensity of negative selection and influences codon usage. Here, we study the connection between gene expression and selection patterns in two species of ciliate protists of the genus *Euplotes*. We find that codon usage is influenced by gene expression in these organisms, pointing at additional evolutionary constraints on mutations in heavily expressed genes relative to the genes expressed at lower rates. At the same time, at the level of synonymous vs. non-synonymous substitutions we observe a stronger constraint on the genes expressed at lower rates relative to those with higher rates of expression. Our study adds to the discussion about the general evolutionary patterns and opens new questions about the mechanisms of control of gene expression in ciliates.

Keywords: gene expression, negative selection, infusoria, ciliates, *Euplotes*, coding constraint