Bacterial CRISPR abundance jumps up at around 45°C

Xin-Ran Lan, Zhi-Ling Liu, and Deng-Ke Niu*

MOE Key Laboratory for Biodiversity Science and Ecological Engineering and Beijing Key Laboratory of Gene Resource and Molecular Development, College of Life Sciences, Beijing Normal University, Beijing, 100875, China

*Corresponding author: Deng-Ke Niu, dkniu@bnu.edu.cn, dengkeniu@hotmail.com

Abstract

Although functioning for adaptive immunity, CRISPR-Cas systems are present in only 40% of bacterial genomes. In this study, we observed an abrupt transition of bacterial CRISPR abundance at around 45°C. Phylogenetic comparative analysis showed that the CRISPR abundance correlates with growth temperature only in temperature ranges around 45°C. The existence of a saltation point indicates that temperature unlikely determines the thermal distribution of CRISPR-Cas systems directly. We put forward a novel hypothesis referring to the predator effect on immune function previously observed in birds. At around 45°C, an abrupt decrease of cellular predators rapidly decreases bacterial mortality resulting from cellular predator grazing. Consequently, the relative contribution of viral lysis in bacterial mortality increases rapidly with temperature; mutants allocating more resources to the adaptive immunity would be favored. In temperature ranges where the abundance of cellular predators does not change with temperature, the temperature effect on CRISPR abundance would be negligible.

Key words: CRISPR, optimal growth temperature, bacteria, protistan grazing, phylogenetic generalized least squares (PGLS).

CRISPR-Cas is the adaptive immune system of bacteria and archaea. Like the adaptive immune system of jawed vertebrates, it can remember previously encountered pathogens, initiate a rapid response to a second invasion, and eliminate the recurrent invader. However, unlike the ubiquitous presence of adaptive immunity in jawed vertebrates (*1*), the CRISPR-Cas systems are only present in about 40% of bacteria (*2*). The sparse distribution of the CRISPR-Cas systems among bacteria is recognized as one of the major open questions to be solved (*3, 4*). Just after discovering the CRISPR-Cas structures, it was noticed that they are prevalent in the thermophilic archaea and the hyperthermophilic bacteria (*5, 6*). Later extensive scale analyses confirmed the prevalence of

CRISPR-Cas systems in thermophiles and hyperthermophiles and showed a positive correlation between CRISPR abundance and growth temperature (*7-10*). Recently, a computer learning approach with an attempt to control the phylogenetic bias suggests that temperature and oxygen levels are the most influencing ecological factors determining the distribution of the CRISPR-Cas systems (*11*).

By plotting the CRISPR abundance against the optimal growth temperatures (Topt) in a column chart, we see a novel pattern on the thermal distribution of CRISPRs in bacteria (Fig. 1A). An abrupt transition of bacterial CRISPR abundance happens at around 45°C. The CRISPR abundance stabilizes at a low level below 45°C. Moreover, the CRISPR abundance fluctuates at a high level above 45°C. By sliding the windows of each column and calculating the ratio of CRISPR abundance of one column to the next, we confirmed that the transition of CRISPR abundance at 45°C is the most sudden (Fig. 1B).

For a quantitative measure for the relationship between bacterial CRISPR abundance and Topt, we first measured their phylogenetic signals ($\lambda = 0.8$, $p = 3 \times 10^{-44}$ for CRISPR abundance and $\lambda = 0.95$, $p = 4 \times 10^{-197}$ for Topt) and proved the necessity of using phylogenetic comparative methods to control the effects of common ancestors (*12, 13*). Their correlation was examined by the phylogenetic generalized least squares (PGLS) regression, where a significant positive slope corresponds to a significant positive correlation, and a negative correlation indicates the reverse. Globally, bacterial CRISPR abundance is positive correlated with Topt (slope = 0.06, $p = 8 \times 10^{-12}$). However, when the bacteria were classified into three categories according to their Topt, low temperatures ($10 \leq$ Topt ≤ 35°C, $n = 488$), middle temperatures ($35 <$ Topt < 50°C, $n = 97$), and high temperatures ($50 \leq$ Topt ≤ 85°C, $n = 97$), significant correlation was only observed in the middle-temperature bacteria (slope = 0.24, $p = 2 \times 10^{-4}$). To avoid the bias resulting from different sample sizes, we compared the correlations among equal-size (100) samples. Along the Topt axis from low to high, we took overlapping samplings, each time with a step of one sample. PGLS analysis of these datasets showed that significant correlations were observed only in a range of about 20°C near 45°C (Fig. 1C).

Bacterial CRISPR spacer abundance also increases rapidly with Topt (Fig. 1D) around 45°C. However, the increase is gradual. PGLS analysis showed significant correlation between CRISPR spacer abundance and Topt at the middle temperatures (slope = 6.3, $p = 3 \times 10^{-4}$), but not in the low temperatures (slope = –0.10, $p = 0.77$) or the high temperatures (slope = –0.30, $p = 0.85$).

By the way, we also examined the relationship in archaea. Globally, archaeal Topt is positively correlated with CRISPR abundance (n =156, slope = 0.10, $p = 2 \times 10^{-6}$), but not CRISPR spacer abundance (n =156, slope = 0.91, $p = 0.10$). When the 156 archaea were divided into three nearly equal-sized groups according to their Topt. A significant correlation between Topt and CRISPR abundance could be observed in the middle-temperature group ($n = 53$, $41.5 \leq$ Topt ≤ 82.5°C, slope = 0.12, $p = 0.01$), but not in the lower ($n = 52$, $18 \leq$ Topt ≤ 41°C, slope = 0.01, $p = 0.75$) or the higher

one (n = 51, 83 ≤ Topt ≤ 106°C, slope = 0.14, p = 0.14). No abrupt transition of CRISPR abundance was observed in archaea.

Several evolutionary and mechanical links between temperature and CRISPR abundance have been proposed (*9, 14, 15*). However, none of them could explain the abrupt transition at 45°C. Temperature influences many aspects of cellular processes and the physical features of the environment (*16*). No matter linear or nonlinear, the physical effects of temperature increase gradually with temperature. However, the thermal distribution of eukaryotes has an abrupt decline at around 45°C. As temperatures lower than 50°C are widespread on earth, only a few members of each group of eukaryotes can live at higher temperatures (*16-19*). Some bacteria could also kill other bacteria and consume the released nutrients. Except for a few exceptions isolated from hot springs, most predatory bacteria could not grow at temperatures above 45°C (*20, 21*). Here, we propose that the accessibility of cellular predators along the temperature axis might indirectly govern the thermal distribution of the CRISPR-Cas systems (Fig. 2).

The effect of predation risk on allocating the limiting resources to immune function had been shown previously in birds (*22*). The birds captured by cats consistently had smaller spleens than the birds killed by nonpredatory reasons like collisions with windows or cars. When the hosts are exposed to lethal predators, the priority to invest in the immune function would be reduced. Most of the bacterial mortality is caused by cellular predator grazing and viral lysis (*23*). When bacterial mortality mostly comes from predator grazing, allocating the limiting resources to the adaptive immunity does not have priority. By contrast, when bacterial mortality mostly comes from viral lysis, allocating more resources to adaptive immunity would be favored. At around 45°C, with the increase of environmental temperature, the abundance of cellular predators decreases sharply (*16-19*), and thus the grazing pressure on bacteria is abruptly relieved. Bacteria growing at higher temperatures would allocate more resources to the antivirus system. Within the temperature ranges inaccessible to grazing predators or within the temperature ranges that are fully accessible to the grazing predators, temperatures have no effects on the proportion of resources allocated to the antivirus system.

Supplementary Material

Supplementary data are submitted along with the main text.

Acknowledgments

We thank Quan-Guo Zhang and Wen-Hong Deng for helpful discussions and Christine Pourcel and Pierre-Albert Charbit for their technical support.

Funding: This work was supported by the National Natural Science Foundation of China (grant number 31671321).

Author Contributions: DKN conceived the study and wrote the manuscript. XRL and ZLL performed the analysis. All authors read, improved, and approved the final manuscript.

Competing interests: The authors declare no competing interests.

Data and materials availability: All data and code are available as supplementary materials.

References

1. V. Müller, R. J. de Boer, S. Bonhoeffer, E. Szathmáry, An evolutionary perspective on the systems of adaptive immunity. *Biol. Rev.* **93**, 505-528 (2018).
2. K. S. Makarova *et al.*, Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* **18**, 67-83 (2020).
3. E. V. Koonin, Open questions: CRISPR biology. *BMC Biol.* **16**, 95 (2018).
4. H. Ledford, Five big mysteries about CRISPR's origins. *Nature* **541**, 280-282 (2017).
5. R. Jansen, J. D. A. v. Embden, W. Gaastra, L. M. Schouls, Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* **43**, 1565-1575 (2002).
6. K. S. Makarova, L. Aravind, N. V. Grishin, I. B. Rogozin, E. V. Koonin, A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Res.* **30**, 482-496 (2002).
7. K. S. Makarova, Y. I. Wolf, S. Snir, E. V. Koonin, Defense islands in bacterial and archaeal genomes and prediction of novel defense systems. *J. Bacteriol.* **193**, 6039-6056 (2011).
8. R. E. Anderson, W. J. Brazelton, J. A. Baross, Using CRISPRs as a metagenomic tool to identify microbial hosts of a diffuse flow hydrothermal vent viral assemblage. *Fems Microbiology Ecology* **77**, 120-133 (2011).
9. A. D. Weinberger, Y. I. Wolf, A. E. Lobkovsky, M. S. Gilmore, E. V. Koonin, Viral diversity threshold for adaptive immunity in prokaryotes. *mBio* **3**, e00456-00412 (2012).
10. U. Gophna *et al.*, No evidence of inhibition of horizontal gene transfer by CRISPR–Cas on evolutionary timescales. *ISME J.* **9**, 2021-2027 (2015).
11. J. L. Weissman, R. M. R. Laljani, W. F. Fagan, P. L. F. Johnson, Visualization and prediction of CRISPR incidence in microbial trait-space to identify drivers of antiviral immune strategy. *ISME J.* **13**, 2589-2602 (2019).
12. J. Felsenstein, Phylogenies and the comparative method. *Am. Nat.* **125**, 1-15 (1985).
13. M. R. E. Symonds, S. P. Blomberg, in *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology: Concepts and Practice,* L. Z. Garamszegi, Ed. (Springer Berlin Heidelberg, Berlin, Heidelberg, 2014), pp. 105-130.
14. N. M. Høyland-Kroghsbo, K. A. Muñoz, B. L. Bassler, Temperature, by controlling growth rate, regulates CRISPR-Cas activity in *Pseudomonas aeruginosa*. *mBio* **9**, e02184-02118 (2018).
15. J. Iranzo, A. E. Lobkovsky, Y. I. Wolf, E. V. Koonin, Evolutionary dynamics of the prokaryotic adaptive immunity system CRISPR-Cas in an explicit ecological context. *J. Bacteriol.* **195**, 3834-3844 (2013).
16. A. Clarke, The thermal limits to life on Earth. *Int J Astrobiol* **13**, 141-154 (2014).
17. T. D. Brock, in *Thermophiles Biodiversity, Ecology, and Evolution,* A.-L. Reysenbach, M. Voytek, R. Mancinelli, Eds. (Springer US, Boston, MA, 2001), pp. 1-9.
18. T. D. Brock, Life at high temperatures. *Science* **230**, 132-138 (1985).

19. M. R. Tansey, T. D. Brock, The upper temperature limit for eukaryotic organisms. *Proc. Natl. Acad. Sci. USA* **69**, 2426-2428 (1972).
20. H. Reichenbach, The ecology of the myxobacteria. *Environ Microbiol* **1**, 15-21 (1999).
21. H. N. Williams, H. Chen, Environmental regulation of the distribution and ecology of *Bdellovibrio* and like organisms. *Front. Microbiol.* **11**, 19 (2020).
22. A. P. Møller, J. Erritzøe, Predation against birds with low immunocompetence. *Oecologia* **122**, 500-504 (2000).
23. H. Takasu, T. Kunihiro, S.-i. Nakano, Protistan grazing and viral lysis losses of bacterial carbon production in a large mesotrophic lake (Lake Biwa). *Limnology* **15**, 257-270 (2014).

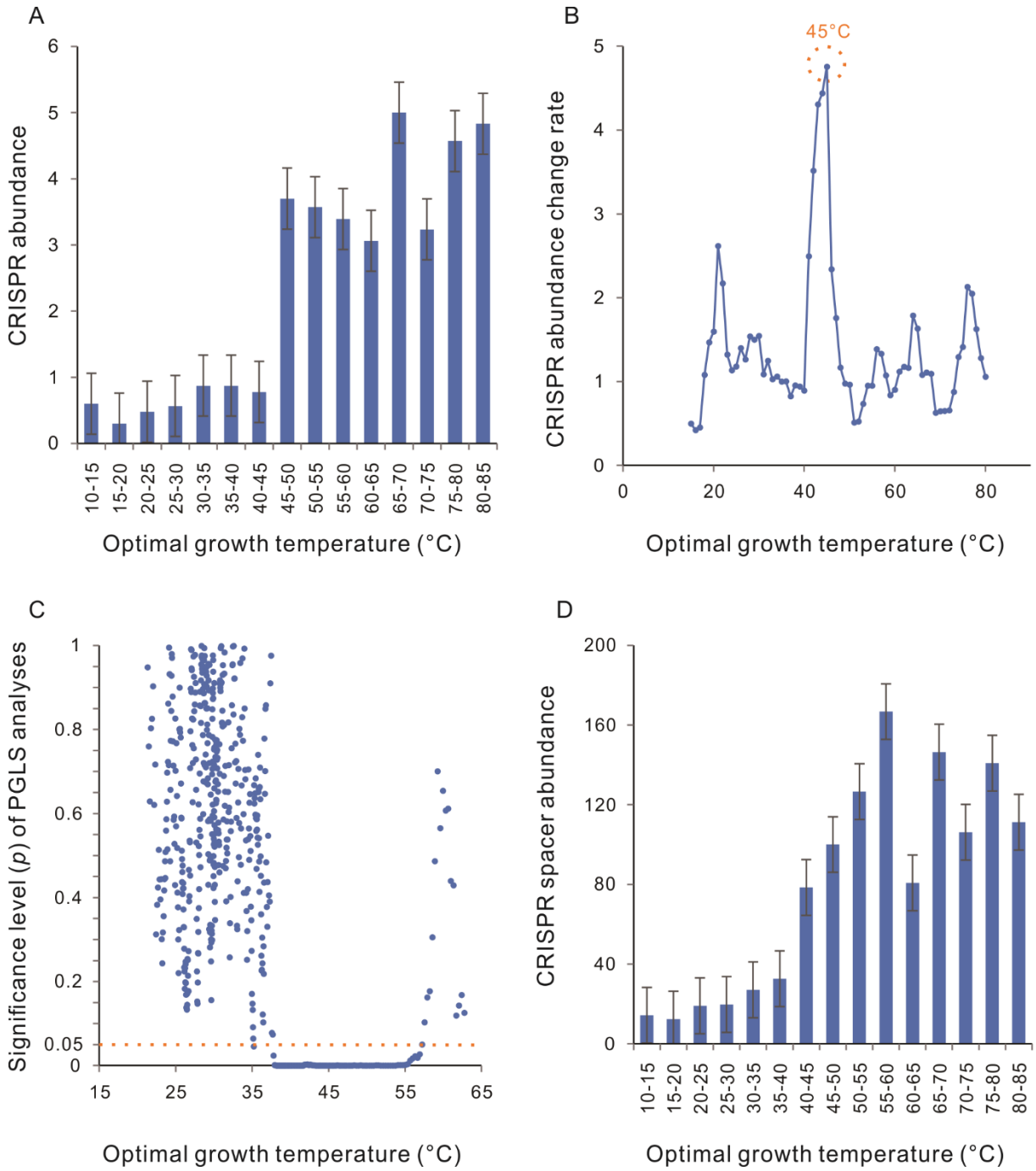

Fig. 1. Distinctive relationships of bacterial CRISPR abundance with different optimal growth temperature (Topt) ranges. (A) 45°C is the saltation point of the thermal distribution of bacterial CRISPR abundance. (B) The change rate of bacterial CRISPR abundances along the axis of Topt, which was calculated from the ratio of bacterial CRISPR abundance of each column to the next. The columns not shown in Fig. 1A were obtained by sliding one degree for each time. (C) Phylogenetic generalized least squares (PGLS) analysis showed a significant positive correlation between bacterial CRISPR abundances and Topt only around 45°C. The 682 bacteria were aligned along the Topt axis. One hundred neighboring samples were taken in each round of PGLS analysis. (D) The abundance of CRISPR spacer also increases rapidly with Topt around 45°C, but no saltation point has been observed. In (A) and (D), the average values and the standard errors are presented.

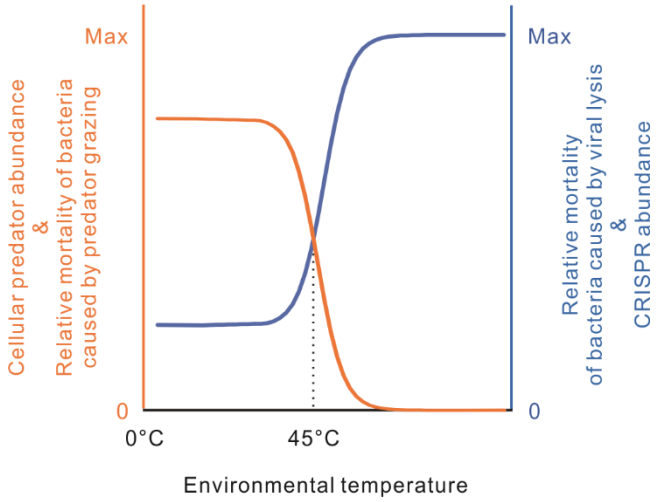

Fig. 2. A model for the thermal distribution of bacterial CRISPR-Cas systems. At low temperatures, the resource allocation to CRISPR-Cas systems is limited by the grazing pressure of cellular predators. At temperatures around 45°C, the abundance of cellular predators decreases abruptly with environmental temperature, so the available resources for the immune function increase rapidly. When cellular predators almost disappear at high temperatures, most mortality results from viral lysis; most bacteria have evolved the highest antivirus capacity.