RESEARCH ARTICLE

METABOLIC GENOMICS

Resource conservation manifests in the genetic code

Liat Shenhav[1,2]* and David Zeevi[1]*†

Nutrient limitation drives competition for resources across organisms. However, much is unknown about how selective pressures resulting from nutrient limitation shape microbial coding sequences. Here, we study this "resource-driven selection" by using metagenomic and single-cell data of marine microbes, alongside environmental measurements. We show that a significant portion of the selection exerted on microbes is explained by the environment and is associated with nitrogen availability. Notably, this resource conservation optimization is encoded in the structure of the standard genetic code, providing robustness against mutations that increase carbon and nitrogen incorporation into protein sequences. This robustness generalizes to codon choices from multiple taxa across all domains of life, including the human genome.

Nitrogen and carbon are major limiting factors in many ecosystems, with recent studies linking their availability to core genomic properties (*1*, *2*). In low-nitrogen environments, there is a strong A+T bias in nucleotide sequences, smaller genome sizes, and a lower incorporation of nitrogen-rich side chains into proteins (*3*–*5*). Opposite trends have been shown for carbon limitation (*2*), and indeed, nitrogen and carbon concentrations are typically inversely correlated (*6*). Thus, recent studies propose a purifying selective pressure associated with resource conservation (*2*, *3*, *5*), which we term "resource-driven selection."

Resource-driven selection postulates that mutations resulting in excess incorporation of nutrients such as nitrogen and carbon are disfavored. However, not all mutations have the same effect on protein sequences, because of constraints imposed by the pattern of codon assignments in the standard genetic code (hereinafter, the structure of the genetic code). The genetic code, common to virtually all of life on earth, can mitigate the effects of mistranslation errors and point mutations (*7*), specifically those leading to radical changes in amino acids. This error minimization is prominent among theories regarding the origin of the genetic code (*8*–*11*), which propose that the code evolved through selection to minimize potential adverse effects of mutations on protein structure and function (*12*–*14*). To quantify code optimality, some theories provide structurally informed amino acid metrics on the basis of hydropathy and stereochemistry [e.g., the polar requirement (PR) scale (*11*) and hydropathy index (*15*)]. To our knowledge, an optimization of nutrient conservation in the genetic code has not been studied thus far.

Results

Widespread purifying selection in the marine environment

To comprehensively characterize how coding sequences of marine microbes are affected by resource availability, we first downloaded 746 samples from the Tara Oceans consortium (n = 136) (*16*), bioGEOTRACES (n = 480) (*17*), and the Hawaii Ocean time series (HOT; n = 68) and Bermuda Atlantic time series (BATS; n = 62) (*17*) (fig. S1A) (*18*). We then devised a computational pipeline that calculates selection metrics from these marine metagenomic samples (fig. S1B). We aligned reads to the Ocean Microbiome Reference Gene Catalog (OM-RGC) (*16*), a database of genes from marine environments that is accompanied by functional information. We searched for single-nucleotide polymorphisms (SNPs) in genes that had sufficient high-quality coverage (fig. S1B) (*18*). Overall, we found 71,921,864 high-confidence SNPs (*18*), in a total of 1,590,843 genes.

To quantify purifying selection on different gene functions, we annotated genes from the OM-RGC database using either KEGG orthology groups (KOs) (*19*) or eggNOG orthologous groups (OGs) (*18*, *20*). Using called SNPs, we calculated for each orthologous group and in each sample the ratio of nonsynonymous polymorphisms to synonymous polymorphisms (pN/pS) (*18*). Across all samples, we found pN/pS ratios to be close to zero with an average of 0.074 in eggNOG OGs [confidence interval (CI), 0.072 to 0.075] (fig. S1C) and 0.079 in KEGG KOs (CI, 0.077 to 0.080) (fig. S1D), indicating purifying selection across the marine environment. To corroborate the validity of calculating selection metrics from metagenomic samples, we compared nonsynonymous mutations leading to "conservative" amino acid substitutions with those leading to "radical" substitutions and found conservative mutations to be significantly more common (permutation test, P < 0.0001) (*21*).

Resource-driven selection apparent across marine microbial genes

On the basis of recent studies (*2*–*5*), we hypothesized that nutrient availability is a central driver of this purifying selection. We thus considered environmental measurements taken alongside each sample, including depth, water temperature, salinity, and concentrations of nitrate, nitrite, oxygen, phosphate, and silicate (fig. S2, A to H) (*18*).

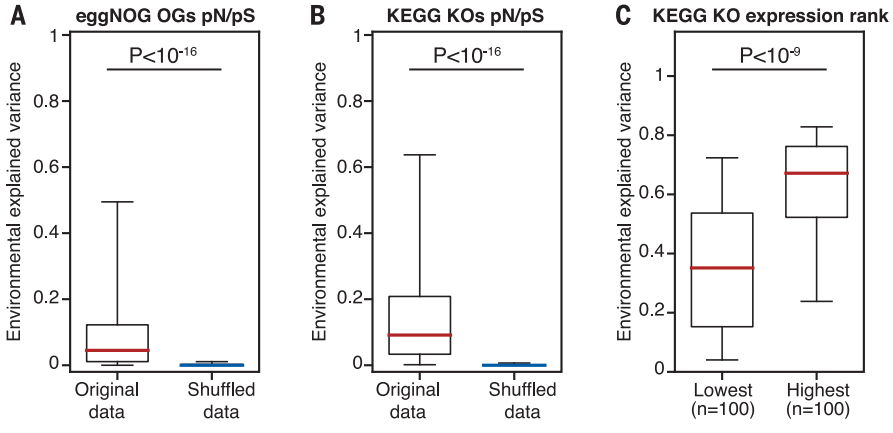

Fig. 1. Analysis of pN/pS ratios reveals resource-driven selection. (**A**) Variance of eggNOG OG pN/pS explained by the environment in the LMM (red) (*18*) compared with the same data with shuffled labels (blue). The box plot lines show median values, the boxes show interquartile ranges, and whiskers show 5th and 95th percentiles. (**B**) Similar data presentation as in (A), for KEGG KO pN/pS. P values were determined with the Wilcoxon signed-rank test. (**C**) Box plot of variance in pN/pS explained by the environment in the LMM in the 100 lowest (left) and 100 highest (right) expressed KEGG KOs. P values were determined with the Mann-Whitney U test.

[1]Center for Studies in Physics and Biology, Rockefeller University, New York, NY, USA. [2]Department of Computer Science, University of California Los Angeles, Los Angeles, CA, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: dzeevi@rockefeller.edu

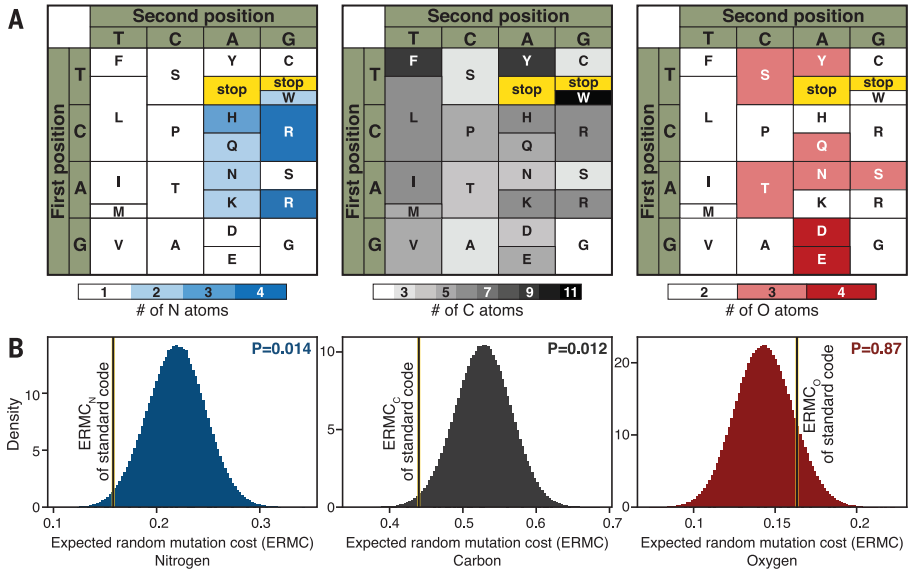

Fig. 2. Resource conservation is facilitated by the genetic code. (**A**) Nitrogen (left), carbon (center), and oxygen (right) content of different amino acids, depicted along with their positions in the standard genetic code. (**B**) Histograms of the ERMC in 1 million random permutations of the genetic code for nitrogen (blue), carbon (black), and oxygen (red). The black and yellow bars mark the ERMCs of the standard genetic code for each of the elements. *P* values were determined with a permutation test.

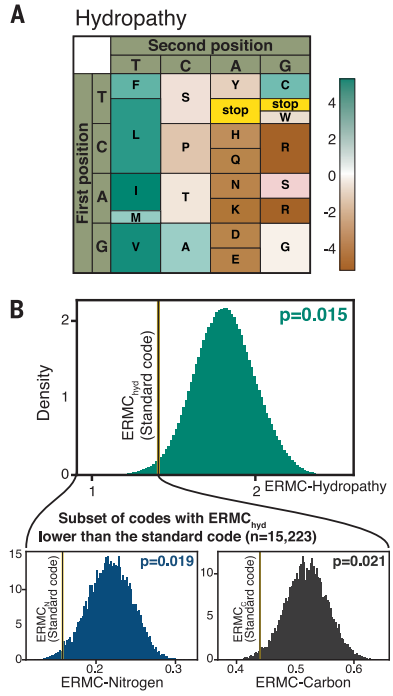

Fig. 3. Optimization for carbon and nitrogen is not confounded by hydropathy. (**A**) Hydropathy of different amino acids, depicted along their positions in the standard genetic code. (**B**, top) Histograms of ERMCs in 1 million random permutations of the genetic code for hydropathy. (Bottom) Histograms of ERMCs for nitrogen (left) and carbon (right) for the subset of hypothetical genetic codes with ERMCs lower than the standard genetic code for hydropathy. *P* values were determined with a permutation test.

These measurements presented consistent correlation patterns with the pN/pS ratios of many KEGG and eggNOG orthologs (fig. S3). However, as they are also correlated with each other (fig. S2I), we cannot accurately estimate their individual effects. We therefore used a linear mixed model (LMM) with variance components (*18*) to estimate the fraction of variance in pN/pS ratios (dependent variable) that is explained by the environment (random effect), while controlling for the correlation structure between the environmental parameters. We term the fraction of variance in pN/pS ratios explained by resource availability the environmental explained variance (EEV) (*18*). Across both KEGG and eggNOG orthologs, a significant fraction of the variance in pN/pS can be attributed to the environment (Mann-Whitney *U* test, $P < 10^{-16}$) (Fig. 1, A and B, and fig. S4, A and B), with nitrate being more strongly correlated with pN/pS ratios than any other environmental parameter (Kolmogorov-Smirnov test, $P < 10^{-30}$ for all comparisons) (fig. S5). Examining typical DNA mutations and amino acid substitutions in nitrate-rich versus nitrate-poor environments, we found that environmental nitrate is associated with specific changes to both DNA and protein sequences, favoring lower nitrogen incorporation into protein sequences when nitrate is scarce (*21*).

This association between environmental measurements and the magnitude of purifying selection is significant even after controlling for potential confounders such as time or effective population size (Mann-Whitney *U* test, $P < 10^{-16}$) (fig. S4, C and D) (*21*), as well as in specific environmental niches (Mann-Whitney *U* test, $P < 10^{-20}$) (fig. S6) (*21*). Additionally, these results were replicated using benchmarking data of assembled genomes from uncultivated single cells from three dominant lineages of the surface ocean (SAR-11, SAR-86, and *Prochlorococcus*) (fig. S4, E to G) (*21*, *22*). These validations demonstrated that the association between selective pressure and environmental conditions is robust to both data type and selection metric and is not confounded by population properties and clade-specific metabolism.

Environmental association is stronger in resource-consuming genes

With nitrate being the environmental factor most strongly associated with pN/pS, we and others hypothesize that mutations that increase the nitrogen requirements of cells are selected against, especially in nitrogen-limited conditions (*3*, *5*, *23*). This implies stronger purifying selection in highly expressed genes (*3*, *4*), where one DNA mutation could translate to thousands of proteins, each consuming more resources (illustrated in fig. S7). We thus used an expression dataset for marine microbial genes (*24*) to rank KEGG KOs by their mean expression (*18*). The 100 most highly expressed KEGG KOs had a significantly higher EEV than the 100 least-expressed ones (Mann-Whitney *U* test, $P < 10^{-9}$) (Fig. 1C and fig. S8). We replicated these results using single-cell data pertaining to specific bacterioplankton lineages (Mann-Whitney *U* test, $P < 10^{-7}$). Additionally, we found that genes encoding extracellular proteins, i.e., for resources excreted from the cell that cannot be recycled, had significantly higher EEV than other gene groups (Mann-Whitney *U* test, $P < 0.05$) (fig. S9) (*18*). This higher EEV for resource-consuming genes further strengthens our results regarding the breadth of resource-driven selection.

Resource conservation as an optimization mechanism in the genetic code

We observed selection against DNA mutations which result in excess incorporation of nutrients, such as nitrogen and carbon, into proteins. However, mutations are constrained by the structure of the genetic code, which minimizes the impact of point mutations on protein structure and function (*12*–*14*). We hypothesized that the genetic code also minimizes the impact of point mutations on nutrient incorporation into proteins. Specifically, the genetic code acts as a buffer between DNA, where mutations occur, and proteins, where resource-driven selection is exerted.

We thus defined, for each element *e* (e.g., carbon, nitrogen), a function quantifying the cost of a single mutation as the added number of amino acid atoms resulting from it.

For example, a missense mutation from codon CCA to CGA results in an amino acid substitution from proline to arginine, with an increase of one carbon and three nitrogen atoms, setting the nitrogen cost of such a mutation to 3 and the carbon cost to 1 (Fig. 2A).

We calculated, for nitrogen, carbon, and oxygen, the expected random mutation cost (ERMC) for the standard genetic code. This calculation considered the abundance of codons, calculated from all marine samples; the transition probability between codons, estimated using the abundance of all single-nucleotide mutations (e.g., for the mutation changing GCA to CCA, we use the abundance of G-to-C transversions); and the cost function, i.e., the number of atoms of each element added after mutation (*18*). For the standard genetic code, and also for codon abundances and mutation rates calculated for marine microbes, we report ERMC values of 0.44, 0.16, and 0.16 for carbon, nitrogen, and oxygen, respectively, corresponding to an average increase of this number of atoms per random mutation (Fig. 2B).

To check if the standard genetic code, along with codon abundances and mutation rates, is indeed robust to resource-consuming mutations, we compared it with other hypothetical codes. We simulated alternative genetic codes by randomizing the first and second position in all codons 1 million times, to create a null distribution of ERMC (*18*). We found that the standard genetic code, common to most life forms, is parsimonious in terms of carbon and nitrogen utilization, given a random mutation. This is marked by a significantly low ERMC for nitrogen ($ERMC_N$ $P = 0.014$) (Fig. 2B) and carbon ($ERMC_C$ $P = 0.012$), but not oxygen ($ERMC_O$ $P = 0.87$).

We compared the extent of robustness against addition of carbon or nitrogen to protein sequences with the robustness against amino acid changes that may affect protein structure and function, as extensively reported (*12*–*14*). We calculated ERMC for changes in hydropathy (*15*) and polar requirement (PR) (*11*), both of which are structurally informed amino acid properties used to determine code error minimization (*18*). We found that these optimization mechanisms are of a similar magnitude of significance as those for nitrogen and carbon conservation ($ERMC_{PR}$ $P = 0.014$; $ERMC_{hyd}$ $P = 0.015$) (fig. S10, C and D) (*18*).

We then devised a hierarchical model to examine the subset of genetic codes (out of 1 million hypothetical codes tested) that have a lower ERMC than the standard code for PR or hydropathy, and we tested whether this subset is also optimized for nitrogen or carbon (*18*). If nutrient optimization is separate from structural optimization, we would expect the standard code to be optimized for carbon and nitrogen, even in comparisons with this subset, with significantly lower $ERMC_N$ and $ERMC_C$ values. Of the 15,223 hypothetical codes that have a hydropathy ERMC ($ERMC_{hyd}$) lower than the standard code, only 270 have a lower $ERMC_N$ ($P = 0.019$) (Fig. 3 and fig. S10E) and only 249 have a lower $ERMC_C$ ($P = 0.021$) (Fig. 3 and fig. S10E). Similarly, of the 13,729 hypothetical codes that have an $ERMC_{PR}$ lower than the standard code, only 83 have a lower $ERMC_N$ ($P = 0.006$) (fig. S10F), and only 442 have a lower $ERMC_C$ ($P = 0.037$) (fig. S10F). This is in contrast with the observed overlap between hydropathy and PR: out of the 15,223 hypothetical codes that have an $ERMC_{hyd}$ lower than the standard code, 6736 have a lower $ERMC_{PR}$ ($P = 0.44$) (fig. S10E). These results indicate that the detected carbon and nitrogen optimization is not confounded by previously reported optimization properties such as hydropathy and PR.

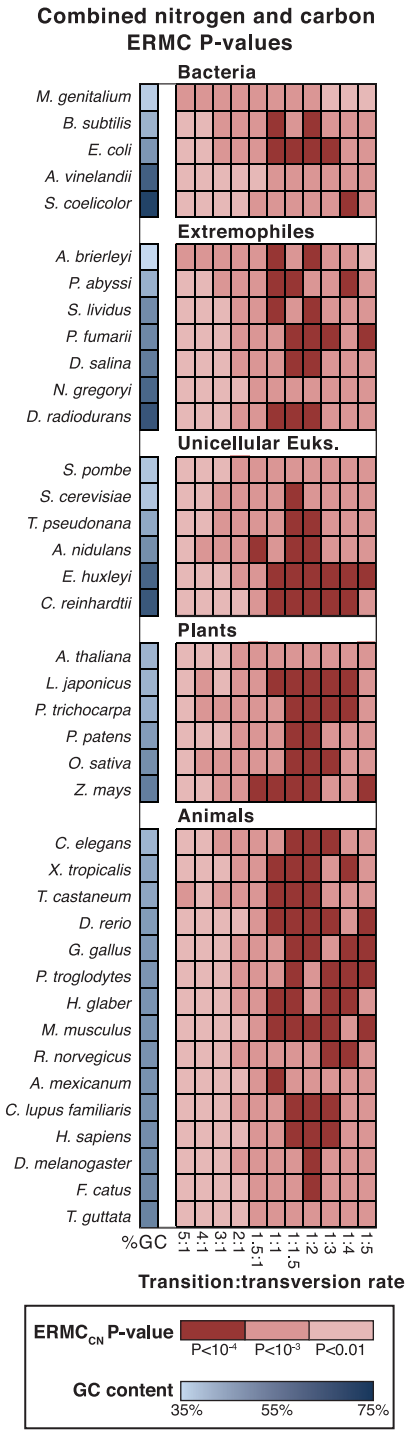

Fig. 4. The genetic code is optimized for resource conservation across organisms. Heatmap of $ERMC_{CN}$ *P* values across 39 organisms and 11 transition:transversion rates. Organisms in each group are ordered by the GC content of their coding sequences.

Remarkably, only 128 out of 1 million randomized genetic codes were better than the standard code in conservation of nitrogen and carbon together ($ERMC_{CN}$ $P = 1.3 \times 10^{-4}$) (fig. S10, G and H). This number is significantly smaller than the number of hypothetical codes expected to have both a lower $ERMC_C$ and $ERMC_N$ (chi-square test of independence, $P = 0.0013$) (table S1). This is possibly driven by a small overlap between the positions of high-nitrogen and high-carbon amino acids. This property of the standard code potentially enables concurrent optimization for both carbon and nitrogen. These results highlight a new optimization principle of the genetic code that is of similar magnitude—and independent of—previously proposed principles.

The genetic code facilitates resource conservation across kingdoms

To show that the resource robustness of the genetic code was not limited to our dataset, we calculated the ERMC of 187 strains of marine microbes in the genera *Prochlorococcus* and *Synechococcus*. We computed codon abundances and mutation rates using published protein-coding sequences (*1*) and the accepted transition:transversion rate of 2:1 (*18*, *25*). We identified significant conservation of carbon, nitrogen, and both elements combined ($ERMC_C$ mean, $P = 0.013$ and $P = 0.020$; $ERMC_N$ mean, $P = 0.049$ and $P = 0.032$; $ERMC_{CN}$ $P = 0.0004$, $P = 0.0007$ for *Prochlorococcus* and *Synechococcus*, respectively) (fig. S11A).

To explore whether this nutrient conservation optimization in the genetic code extends across organisms, we performed a similar calculation using codon abundances from 39 organisms across all domains of life, including all human protein-coding sequences and a range of transition:transversion rates (*18*). Similarly to marine microbes, we found that the genetic code is optimized in terms of resource utilization for all tested organisms, and it is marked by a significantly lower ERMC for nitrogen and carbon combined, across all transition:transversion rates ($P < 0.01$) (Fig. 4). Moreover, we found significant optimization even in the theoretical case where all codon

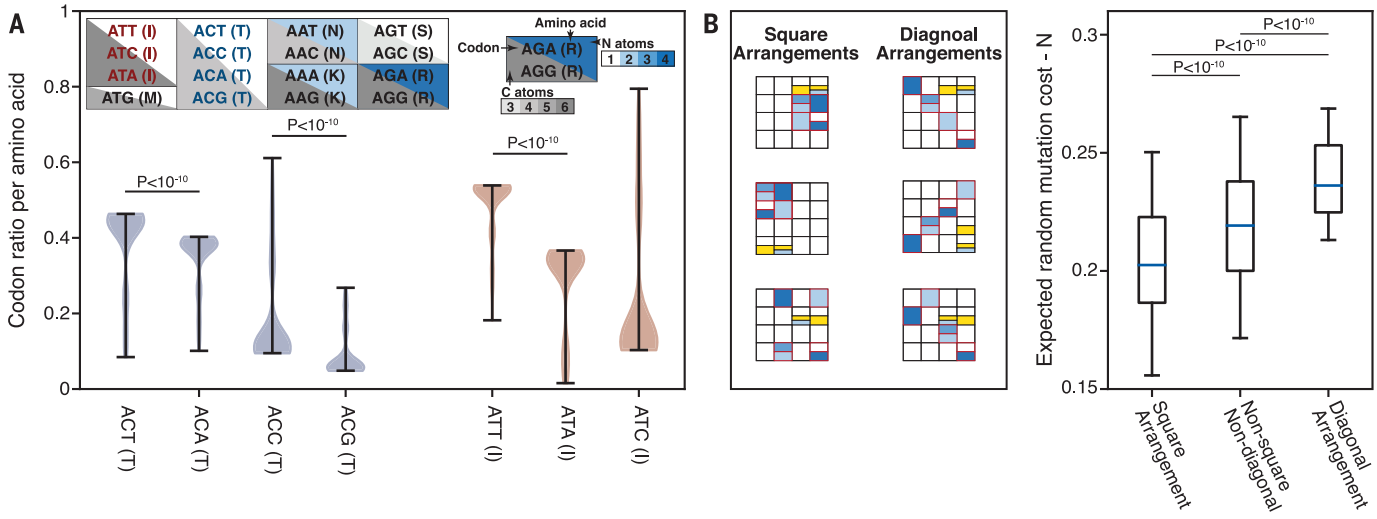

Fig. 5. Structural properties and codon usage bias underlying optimality in the genetic code. (**A**) Violin plot of codon usage among 187 species of *Prochlorococcus* and *Synechococcus*, showing significant preference of threonine codons ACT and ACC compared with ACA and ACG and of isoleucine codon ATT compared with ACA. *P* values were determined with the Wilcoxon signed-rank test. (**B**) $ERMC_N$ values for square arrangements (left) and diagonal arrangements (right) (*18*), compared with all other arrangements (center) out of 10,000 hypothetical arrangements. The box plot lines show median values, the boxes show interquartile ranges, and whiskers show 5th and 95th percentiles. *P* values were determined with the Mann-Whitney *U* test.

abundances are the same ($P < 0.01$) (fig. S11B). The codon abundances of a great majority of organisms also demonstrate significantly lower ERMC values for nitrogen (fig. S11C) and carbon (fig. S11D), for a wide range of transition:transversion rates. These results indicate that resource optimization in the genetic code transcends taxonomy, codon choices, and mutation rates.

Resource conservation may bias codon usage

We examined all amino acids encoded by codons with adenine in the first position, focusing on codon usage of the amino acid threonine. We note that a C-to-G transversion in the second position for codons ACT and ACC yields serine (AGT and AGC, respectively), but the same mutation for codons ACA and ACG yields arginine (AGA and AGG) (Fig. 5A, inset). Arginine has higher carbon and nitrogen contents than serine. We thus hypothesized that for a cell to conserve nutrients in case of a random mutation, codons ACT and ACC should have a higher abundance than codons ACA and ACG, respectively, given a known genomic GC bias. We examined codon usage in 187 *Prochlorococcus* and *Synechococcus* strains and found significantly higher use of ACT than of ACA (Wilcoxon signed-rank test, $P < 10^{-20}$) (Fig. 5A) and significantly higher use of ACC than of ACG (Wilcoxon signed-rank test, $P < 10^{-20}$) (Fig. 5A). Similarly, the isoleucine codon ATT had a higher abundance than ATA (Wilcoxon signed-rank test, $P < 10^{-20}$) (Fig. 5A). These results point to resource conservation as a central driving force in guiding codon usage and thereby affecting not only protein sequence but also cellular translation efficiency.

Structural principles drive optimization in the genetic code

Codons of the nitrogen-rich amino acids histidine, glutamine, asparagine, lysine, and arginine span only two nucleotides in their first position and two in their second position. We define this organization to be a "square" arrangement and hypothesize that it amplifies nitrogen conservation (Fig. 5B) (*18*). Specifically, in the square arrangement, some codons require at least two mutations to increase the number of nitrogen atoms (e.g., those coding for alanine and valine). This is in contrast to other hypothetical arrangements, including a "diagonal" one in which nitrogen-rich amino acid codons span all possible nucleotides in the first and second positions (Fig. 5B) (*18*). On the basis of these results, it seems that the diagonal arrangement would be nutrient-wasteful, as a single mutation in the first or second codon position could increase the nitrogen content of a protein. We generated 10,000 hypothetical codes, with 220 arrangements embodying a square structure and 127 embodying a diagonal one. Compared with all other possible arrangements, the square arrangements present a significantly lower $ERMC_N$, whereas diagonal arrangements exhibit a significantly higher $ERMC_N$ (Mann-Whitney *U* test, $P < 10^{-10}$ for both) (Fig. 5B). This indicates that resource optimization in the standard code is driven by structural principles, perhaps underlying the optimization observed across kingdoms.

Discussion

Here, we characterized and quantified the selective forces exerted by nutrient availability on protein-coding genes in marine environments. We provide a data-driven, population-level perspective and show that resource-driven selection is a ubiquitous force. We further show that a significant portion of DNA mutations may not result in increased nutrient incorporation into protein sequences, owing to the pattern of codon assignments in the genetic code.

In light of these results, we hypothesize that resource-driven selection is equally exerted on all parts of the protein-coding gene. This sets it apart from selection to maintain the structural integrity of a protein or the function of its active site, both of which occur predominantly in structurally important regions (*26*). Thus, accounting for resource-driven selection may improve the identification of alternative translation start sites, alternatively spliced introns, or readthrough stop codons, as intermittently translated regions of the protein may be under weaker resource-driven selection than are constitutively translated ones (*27*).

Our results showing that the genetic code optimizes nutrient conservation are in line with theories sporting an early fixation of an optimal genetic code, suggesting selection for error minimization (*10*, *28*). Nevertheless, the genetic code is also near-immutable, as evident in heterotrophic eukaryotes, which as net nitrogen producers still harbor a nitrogen-conservative code. This implies that the genetic code can be viewed as a buffer between the evolutionary forces of mutation and selection, the former occurring in DNA sequences and the latter predominantly in proteins. In the case of nutrient conservation, many DNA mutations do not result in the incorporation

of additional nutrients into proteins and are thus not selected against, which may allow more "freedom" for fitness gradients to explore the mutation space.

REFERENCES AND NOTES

1. P. M. Berube, A. Rasmussen, R. Braakman, R. Stepanauskas, S. W. Chisholm, *eLife* **8**, e41043 (2019).
2. D. R. Mende *et al*., *Nat. Microbiol.* **2**, 1367–1373 (2017).
3. J. J. Grzymski, A. M. Dussaq, *ISME J.* **6**, 71–80 (2012).
4. H. Akashi, T. Gojobori, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 3695–3700 (2002).
5. F. L. Hellweger, Y. Huang, H. Luo, *ISME J.* **12**, 1180–1187 (2018).
6. E. R. M. Druffel, S. Griffin, A. I. Coppola, B. D. Walker, *Geophys. Res. Lett.* **43**, 5279–5286 (2016).
7. R. T. Hinegardner, J. Engelberg, *Science* **142**, 1083–1085 (1963).
8. G. Gamow, *Nature* **173**, 318 (1954).
9. J. T. Wong, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 1909–1912 (1975).
10. S. J. Freeland, L. D. Hurst, *J. Mol. Evol.* **47**, 238–248 (1998).
11. C. R. Woese, D. H. Dugre, W. C. Saxinger, S. A. Dugre, *Proc. Natl. Acad. Sci. U.S.A.* **55**, 966–974 (1966).
12. S. E. Massey, *J. Mol. Evol.* **67**, 510–516 (2008).
13. A. S. Novozhilov, E. V. Koonin, *Biol. Direct* **4**, 44 (2009).
14. N. Torabi, H. Goodarzi, H. Shateri Najafabadi, *J. Theor. Biol.* **244**, 737–744 (2007).
15. D. Haig, L. D. Hurst, *J. Mol. Evol.* **33**, 412–417 (1991).
16. S. Sunagawa *et al*., *Science* **348**, 1261359 (2015).
17. S. J. Biller *et al*., *Sci. Data* **5**, 180176 (2018).
18. Materials and methods are available as supplementary materials.
19. M. Kanehisa, S. Goto, *Nucleic Acids Res.* **28**, 27–30 (2000).
20. J. Huerta-Cepas *et al*., *Mol. Biol. Evol.* **34**, 2115–2122 (2017).
21. Supplementary text is available as supplementary materials.
22. M. G. Pachiadaki *et al*., *Cell* **179**, 1623–1635.e11 (2019).
23. A. Dufresne, L. Garczarek, F. Partensky, *Genome Biol.* **6**, R14 (2005).
24. B. C. Kolody *et al*., *ISME J.* **13**, 2817–2833 (2019).
25. E. Urbach, D. J. Scanlan, D. L. Distel, J. B. Waterbury, S. W. Chisholm, *J. Mol. Evol.* **46**, 188–201 (1998).
26. J. Liu, Y. Zhang, X. Lei, Z. Zhang, *Genome Biol.* **9**, R69 (2008).
27. R. W. Read *et al*., *ISME J.* **11**, 2267–2278 (2017).
28. S. J. Freeland, R. D. Knight, L. F. Landweber, L. D. Hurst, *Mol. Biol. Evol.* **17**, 511–518 (2000).
29. L. Shenhav, D. Zeevi, Resource conservation manifests in the genetic code, Zenodo (2020); doi: 5281/zenodo.4001183.

ACKNOWLEDGMENTS

We thank T. Korem and N. D. Rubinstein for discussions, suggestions, and help in analyses. We thank J. P. Eckmann, O. X. Cordero, I. Mizrahi, and N. Antonovsky for their suggestions. **Funding:** D.Z. was funded by the James S. McDonnell Foundation. **Author contributions:** L.S. and D.Z. conceived the project, performed analyses, and wrote the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the manuscript, supplementary materials, and referenced datasets. Code is available at Zenodo (*29*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6517/683/suppl/DC1
Materials and Methods
Supplementary text
Figs. S1 to S16
Tables S1 to S4
References (*30*–*56*)
MDAR Reproducibility Checklist

23 October 2019; resubmitted 4 May 2020
Accepted 11 September 2020
10.1126/science.aaz9642

Resource conservation manifests in the genetic code

Liat Shenhav and David Zeevi

Science **370** (6517), 683-687.
DOI: 10.1126/science.aaz9642

Maximizing elements for your genome

The accumulation of mutations is typically limited by selective parameters. One such parameter is the elements needed to build proteins and molecules to maintain cells. Examining the underlying carbon, oxygen, and nitrogen content for different amino acids, Shenhav *et al.* examined the selective pressure resulting from nutrient limitation (see the Perspective by Polz and Cordero). The authors identified "resource-driven" selection as a purifying selective force associated with environmental nutrient availability, particularly nitrogen, and determined the impact of mutations on the organismal nutritional budget. From this constraint, the authors have proposed that the structure of the genetic code across organisms reflects the mutational impact on elemental resources.

Science, this issue p. 683; see also p. 655