Identifying the specific signals that mediate correlation-dependent structural plasticity will be greatly facilitated by exploiting the experimental protocol presented here.

REFERENCES AND NOTES

ACKNOWLEDGMENTS
We thank the laboratory of K. Haas (UCB) for assistance with the MATLAB analysis tool Dyna norm for dynamic morphometric analysis. We also thank M. Meyer for providing us with the tetanus toxin light chain plasmid (SUG-TNT-E897) and D. Freeman for photography of tadpoles. We are grateful to S. Glasgow for advice on statistical analysis and to H. Cline and W. Sossin for useful comments on our manuscript. This work was supported by grants from the Canadian Institutes for Health Research and the Natural Sciences and Engineering Research Council (NSERC) of Canada to E.S.R. and by fellowships from the Deutscher Akademischer Auswärts Dienst and the NSERC CREATE Neuroengineering Training Program to M.M., the Fonds de la recherche en santé du Québec to D.G., and NSERC to P.S. The authors declare that they have no conflicts of interest.

SUPPLEMENTARY MATERIALS
www.sciencemag.org/content/344/6186/904/suppl/DC1
Materials and Methods
Figs. S1 and S2
References
Movies S1 to S4
31 January 2014; accepted 25 April 2014
10.1126/science.1251937

MICROBIAL GENOMICS
Stop codon reassignments in the wild
Natalia N. Ivanova,1* Patrick Schwientek,1* H. James Tripp,1 Christian Rinke,1 Amrita Pati,1 Marcel Huntemann,1 Axel Visel,2,3* Tanja Woyke,1 Nikos C. Kyrpides,1 Edward M. Rubin1,4†

The canonical genetic code is assumed to be deeply conserved across all domains of life with very few exceptions. By scanning 5.6 trillion base pairs of metagenomic data for stop codon reassignment events, we detected recoding in a substantial fraction of the >1700 environmental samples examined. We observed extensive opal and amber stop codon reassignments in bacteriophages and of opal in bacteria. Our data indicate that bacteriophages can infect hosts with a different genetic code and demonstrate phage-host antagonism based on code differences. The abundance and diversity of genetic codes present in environmental organisms should be considered in the design of engineered organisms with altered genetic codes in order to prevent the exchange of genetic information with naturally occurring species.

Since the discovery of the genetic code and protein translation mechanisms (1), a limited number of variations of the standard assignment between unique base triplets (codons) and their encoded amino acids and translational stop signals have been found in bacteria and phages (2–7). Given the apparent ubiquity of the canonical genetic code, the design of genomically recoded organisms with noncanonical codes has been suggested as a means to prevent horizontal gene transfer between bacteria and environmental organisms (8–10). It is also predicted that genomically recoded organisms are immune to infection by viruses, under the assumption that phages and their hosts must share a common genetic code (6). This paradigm is supported by the observation of increased resistance of genomically recoded bacteria to phages with a canonical code (9). Despite these assumptions and accompanying lines of evidence, it remains unclear whether differential and noncanonical codon usage represents an absolute barrier to phage infection and genetic exchange between organisms.

Our knowledge of the diversity of genetic codes and their use by viruses and their hosts is primarily derived from the analysis of cultivated organisms. This is due to our limited access to genome sequences from uncultivated organisms, which are estimated to account for 99% in prokaryotes (11). Advances in single-cell sequencing and metagenome assembly technologies have enabled the reconstruction of genomes of uncultivated bacterial and archaeal lineages (12–14) and the discovery of a previously unknown reassignment of TGA opal stop codons to glycine (4, 5, 14). These initial findings suggest that large-scale systematic studies of uncultivated microorganisms and viruses may reveal the extent and modes of divergence from the canonical genetic code operating in nature.

To explore alternative genetic codes, we carried out a systematic analysis of stop codon reassignments from the canonical TAG amber, TGA opal, and TAA ochre codons in assembled metagenomes and metatranscriptomes from environmental and host-associated samples, single-cell genomes of uncultivated bacteria and archaea, and a collection of viral sequences (Fig. 1A) (15). All sequence data were obtained from the Integrated Microbial Genomes (IMG) database (16). This global collection of sequences comprised 1776 samples from 146 studies, including 750 samples obtained from 17 human body sites (fig. S1) (17). In total, 5.6 terabases of sequence data, including 450 Gb of contiguous sequences (contigs) greater than 1 kb, were analyzed. All samples were classified into human-associated, other host–associated, soil, marine, or freshwater environments according to their metadata (15, 18).

We used a statistic of increased coding potential under alternate genetic codes as calculated by ab initio gene finder Prodigal (19), which was selected for its low rate of false-positive predictions (15). Contigs showing significantly higher coding potential when annotated with modified translation tables were forwarded to filtering and quality control to confirm stop codon reassignment through multiple sequence alignments to known homologs from the National Center for Biotechnology Information protein database (Fig. 1A) (15).

By applying this approach to 450 Gb of contigs larger than 1 kb in size, we identified 31,415 contigs with evidence of stop codon reassignment, adding up to a total of 198 Mb of recoded DNA (Fig. 1A). No recoding was observed in the metatranscriptome data. Varying ratios of reassigned to total contigs were observed in samples from terrestrial and aquatic environments and from human mouth, throat, and stool microbiomes.

1Department of Energy Joint Genome Institute (DOE JGI), Walnut Creek, CA 94598, USA. 2Genomics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. 3School of Natural Sciences, University of California, Merced, CA 95343, USA. 4These authors contributed equally to this work. *Corresponding authors. E-mail: emrubin@lbl.gov

*Present address: Berkeley National Laboratory, Berkeley, CA 94720, USA.
†Present address: California Institute of Technology, Pasadena, CA 91125, USA.
The greatest reassignment ratio was in a groundwater sample from a sulfidic aquifer, where 10.4% of all the assembled contigs displayed evidence that one of the three stop codons had been reassigned (table S2). High ratios of contig recoding were also detected in human oral microbiome (table S2).

Reassignment of all three stop codons was found but with different preferences by domain and habitat (Fig. 2). We observed distinct patterns of stop codon reassignment in the three domains of life, with bacteria showing only opal reassignments, ochre reassignments restricted to eukaryotes, and archaea devoid of codon reassignments (15). Among viruses, we found both amber and opal reassignments. These observations are restricted to DNA viruses only because of the scarcity of sequence information for RNA viruses of bacteria and archaea (Fig. 2) (15). Metagenomes of human body sites showed a high rate of reassignments compared with most other sampling sites. Only 10% of all contigs examined in this study originated from human body sites, but they represented 51% of all contigs with codon reassignments. The majority of the remaining stop reassignments were found in freshwater environments (44%), representing 13% (56.0 Gb) of all examined metagenomes. In contrast, marine samples contributed only 4% of recorded sequence, although they represent 48% (211.6 Gb) of the total data set (15). This suggests that codon reassignments are more abundant in freshwater than in marine samples.

Among bacteria, previous reports of recoding were restricted to the reassignment of opal stop codons (3–5, 20). Despite our extensive sampling of bacterial sequences, we also observed reassignment exclusively for opal codons. Opal reassignments to Trp have been previously observed in *Mollicutes* (20) and *Candidatus Hodgkinia cicadiciola* (3), and *opal* reassignments to Gly have been observed in uncultivated representatives of candidate phyla SR1 and Gracilibacteria (4, 5). Our extensive survey suggests that *opal* reassignment in bacteria is likely limited to the same specific lineages (Fig. 3). The multiple SR1 and Gracilibacteria sequences enabled us to explore the evolutionary origin of stop codon reassignment in these closely related, uncultivated bacterial lineages. A maximum likelihood phylogenetic tree revealed that *opal* reassignment occurred in the last common ancestor of these sister lineages after its separation from the Peregrines (PERs) group and before the divergence of SR1 and Gracilibacteria (Fig. 3). The same phylogenetic analysis performed for the *opal* reassigned members of the class *Mollicutes* indicates a single reassignment event within the last common ancestor of the *Mycoplasmatales* and Entomoplasmatales.

Although the average GC content of the entire data set was 55%, recoded bacterial sequences had an average GC content of 32%, consistent with previous studies (21). In recoded contigs, we observed a shift to low-GC synonymous codons and/or a shift to low-GC nonsynonymous codons for amino acids with similar chemical properties, supporting the hypothesis that changes in the genomic GC content correlate with and may drive...
reassignment of stop codons (figs. S3 and S4). In addition, low-GC organisms used the ochre stop codon (TAA) to a higher extent (84% in organisms with GC content <32% versus 41% in organisms with GC content >64%) than amber (TAG) and opal (TGA). In extreme cases of low GC content, nearly all genes terminate in an ochre stop codon.

Although our pipeline for alternative genetic code detection was initially developed for prokaryotic genes, it can identify eukaryotic sequences with reassigned stop codons. In agreement with previous reports (2), among eukaryotic sequences we observed recoding for opal reassignment in mitochondrial sequences and ochre and amber reassignment in nuclear sequences. Nuclear sequences with stop codon reassignments appear to belong to the representatives of Ciliophora (table S3), and ochre reassignments were found exclusively in freshwater samples (Fig. 2).

We identified 19 complete and nearly complete DNA phages with amber stop codon reassignments in 177 out of 784 human microbiome samples (15). Previous reports of alternative genetic codes in DNA viruses are limited to opal reassignment to Trp in Mycoplasma phages (7, 22). In our study, we identified two phages with opal reassignment to Gly, a phage with amber reassignment to Ser, and 14 phages with amber reassignment to Gln (table S8)—a code previously observed only in nuclear genes of eukaryotes (23). These reassignments are supported by protein alignments (figs. S6 to S8), as well as by the presence of amber-recognizing Gln-tRNA_{AUA} and Ser-tRNA_{AUA} in the phage contigs (figs. S9 and S10). We infer from the genome structure of recoded DNA phages that they belong to the order of Caudovirales. None of the amber-reassigned DNA phages was embedded into recognizable bacterial sequences, and no integrases were detected in the phage genomes, suggesting that these DNA phages are lytic.

Because phages largely depend on the translational machinery of their hosts, it has been suggested that they must use the same genetic code (6, 9). Evidence supporting the matching usage of genetic codes between an opal-reassigned phage and its host was obtained by looking for footprints of phage infections in phage-derived spacers of the CRISPR (clustered regularly interspaced short palindromic repeat) adaptive immune system of bacteria (24). Out of 26 unique spacers found on CRISPR-harboring contigs with opal reassignment,

**Fig. 3.** A maximum likelihood phylogenetic tree of bacterial stop codon reassigned sequences, based on concatenated alignments of protein-coding marker genes. The arrow at the root of the tree points to the outgroup (Terrabacteria). The tree shows the recoded taxonomic groups Mycoplasmatales and Entomoplasmatales (opal to Trp), SR1 (opal to Gly), and Gracilibacteria (opal to Gly) along with non-recoded reference phyla. The highly reduced alphaproteobacterial Candidatus Hodgkinia cicadicola genome was not included. The red circles denote two reassignment events. PVC, Planctomycetes, Verrucomicrobia, and Chlamydiae; FCB, Fibrobacteres, Chlorobi, and Bacteroidetes. Sequences published in (4, 5, 14).
two spacers had an exact full-length match in the sequence of opal-recoded phages (table S9). Alignment of protein-coding genes on both contigs confirmed that they have opal to Gly reassignment.

The observation of amber reassignments in phages raises questions about the genetic code of their target hosts, given the apparent absence of amber-recoded bacterial genomes from environments in which amber-recoded phages were present (Fig. 2). This raises the possibility that genetic code differences between phages and hosts do not constitute an obligate barrier to phage infection. By analyzing 29,037 spacers found in CRISPR elements from 553 human oral and stool samples, we identified five spacers (each 33 to 37 bp long) that were identical to sequences from three different amber-recoded phage genomes (table S9). The contigs containing the CRISPR spacers also included bacterial genes with the full complement of canonical stop codons. The identified bacterial genes were nearly identical to genes from two Prevotella strains that were isolated from human airways and subgingival plaque and shown to have a standard genetic code. These data suggest that amber-recoded phages can infect hosts with different genetic codes, in this case the standard code.

To gain further insight into mechanisms that may enable amber-recoded phages to infect hosts with different genetic codes, we examined the assembled genomes of amber-recoded phages. In several of these phage genomes, we identified genes for peptide chain release factor 2 (RF-2), which terminates translation at ochre and opal stop codons. Sequenced isolate phages lack genes for release factors, apparently harnessing the host-encoded release factors. The presence of RF-2 in a phage genome suggests that the phage may infect a host lacking RF-2; a hallmark of opal-recoded bacterial genomes (3, 25). Consistent with this possibility, the human oral cavity environments where amber-recoded RF-2-containing phages were detected lack amber-recoded bacteria but are enriched for opal-recoded bacteria. A further atypical feature noted in the genome of one of these phages (phage 2) is a bimodal pattern of amber reassignment across the genome (Fig. 4A). Initial annotation of this phage genome suggests that it is a lytic phage broadly related to T4 (fig. S11), in which amber has been reassigned to code for Gln.

This phage also contains a noncanonical Glutamine-encoding tRNA<sup>CUA</sup>, but closer examination of amber distribution across its genome reveals two large domains with distinct gene content and codon usage. The low-amber (LA) domain contains genes often found in early-stage phage infection, such as DNA polymerase. The LA domain also contains

---

**Fig. 4.** Phage infections across genetic code boundaries. (A) Genome of phage 2. The phage genome is broadly divided into two domains with strong bias in codon utilization as well as strand preference. (B) Model of infection of opal-recoded hosts by amber–Gln-recoded phages.
the RF-2 gene required for normal translation of amber-recoded genes. Open reading frames in the LA domain are almost entirely devoid of in-frame amber codons and instead rely nearly exclusively on canonical glutamine codons to encode for glutamine (Fig. 4A). In contrast, the high-amber (HA) domain with frequent in-frame amber codons contains genes often associated with late stages of phage infection, such as packaging and assembly components (e.g., predicted tail fiber protein, minor tail protein, tail tape measure protein, or tail-associated lysozyme) (Fig. 4A). This distinct codon utilization, combined with the presence of RF-2 and a Gln-tRNA_{GUA} in the amber-recoded phage, suggests that the amber-recoded phage actively interferes with the translation of its presumed opal-recoded host through a proposed mode of phage-host antagonism (Fig. 4B). In this model, upon initial phage infection abundant host-derived RF-1 (the releasing factor that terminates polypeptide chain elongation at amber codons) interferes with the translation of amber-containing phage HA domain genes, so they are initially not expressed. In contrast, critical amber-free phage LA domain genes can be normally translated initially not expressed. In contrast, critical containing phage HA domain genes, so they are codons) interferes with the translation of 4B). In this model, upon initial phage infection (Fig. 2. R. D. Knight, S. J. Freeland, L. F. Landweber, et al. 5. C. Rinke et al. 6. L. A. Shackelton, E. C. Holmes, et al.

CANCER GENOMICS

Activating hotspot L205R mutation in PRKACA and adrenal Cushing’s syndrome

Yanan Cao,1 Minghui He,2 Zhibo Gao,2 Ying Peng,1 Yanli Li,1 Lin Li,2 Weiwei Zhou,1 Xiangchun Li,2 Xu Zhong,3 Yiming Lei,2 Tingwei Su,1 Hang Wang,2 Yiran Jiang,3 Lin Yang,2 Wei Wei,1 Xu Yang,2 Xiuli Jiang,1 Li Liu,4 Juan He,5 Junna Ye,1 Qing Wei,2 Yirong Li,2 Weiqing Wang,1 Jun Wang,5,6,7,† Guang Ning1,8

Adrenal Cushing’s syndrome is caused by excess production of glucocorticoid from adrenocortical tumors and hyperplasias, which leads to metabolic disorders. We performed whole-exome sequencing of 49 blood-tumor pairs and RNA sequencing of 44 tumors from cortisol-producing adrenocortical adenomas (ACAs), adrenocorticotropic hormone–independent macronodular adrenocortical hyperplasias (AIMAHs), and adrenocortical oncocytomas (ADOs). We identified a hotspot in the PRKACA gene with a L205R mutation in 69.2% (27 out of 39) of ACAs and validated in 65.5% of a total of 87 ACAs. Our data revealed that the activating L205R mutation, which locates in the P+1 loop of the protein kinase A (PKA) catalytic subunit, promoted PKA substrate phosphorylation and target gene expression. Moreover, we discovered the recurrently mutated gene DOT1L in AIMAHs and CLASP2 in ADOs. Collectively, these data highlight potential functionally mutated genes in adrenal Cushing’s syndrome.

Cushing’s syndrome is caused by excessive glucocorticoid production, which may lead to a series of metabolic disorders such as obesity, glucose intolerance, and hypertension. Adrenal Cushing’s syndrome results from autonomous production of cortisol from adrenocortical tumors (ACTs), which are most common in adult females. Cortisol-producing ACTs include benign adrenocortical adenoma (ACA), malignant adrenocortical carcinoma (ACC), and rare forms of bilateral adrenal hyperplasia (BAH) and adrenocortical oncocytoma (ADO). BAHs consist of macronodular and micronodular hyperplasias, such as adrenocorticotropin-independent macronodular adrenocortical hyperplasia (AIMAH) and primary pigment nodular adrenocortical disease (PPNAD).

Several genetic alterations have been described in inherited or sporadic BAHs. The inactivating mutations in PRKARIA (cyclic adenosine monophosphate (cAMP)–dependent protein kinase regulatory subunit type 1α) are dominant in

REFERENCES AND NOTES
