Community richness of amphibian skin bacteria correlates with bioclimate at the global scale

Jordan G. Kueneman1,37, Molly C. Bletz2,37, Valerie J. McKenzie3, C. Guilherme Becker4, Maxwell B. Joseph5, Juan G. Abarca6, Holly Archer3, Ana Lisette Arellano3, Arnaud Bataille7, Matthew Becker8, Lisa K. Belden9, Angelica Crottini10, Robert Geffers11, Célio. F. B. Haddad12, Reid N. Harris13, Whitney M. Holden14, Myra Hughey15, Michael Jarek11, Patrick J. Kearns16, Jacob L. Kerby17, Jos Kielgast18, Atsushi Kurabayashi19,20,21, Ana V. Longo22, Andrew Loudon23,24, Daniel Medina9, José J. Nuñez25, R. G. Bina Perl26, Adrián Pinto-Tomás6,27, Fatitiana C. E. Rabemananjara28, Eria A. Rebollar29, Ariel Rodríguez30, Louise Rollins-Smith31, Christoph C. Tebbe31, Gabriel Vargas Asensio6, Bruce Waldman32, Jennifer B. Walke33, Steven M. Whitfield34, Kelly R. Zamudio35, Ibrahim Zűñiga Chaves36, Douglas C. Woodhams1,2,36* and Miguel Vences36*

Animal-associated microbiomes are integral to host health, yet key biotic and abiotic factors that shape host-associated microbial communities at the global scale remain poorly understood. We investigated global patterns in amphibian skin bacterial communities, incorporating samples from 2,349 individuals representing 205 amphibian species across a broad biogeographic range. We analysed how biotic and abiotic factors correlate with skin microbial communities using multiple statistical approaches. Global amphibian skin bacterial richness was consistently correlated with temperature-associated factors. We found more diverse skin microbiomes in environments with colder winters and less stable thermal conditions compared with environments with warm winters and less annual temperature variation. We used bioinformatically predicted bacterial growth rates, dormancy genes and antibiotic synthesis genes, as well as inferred bacterial thermal growth optima to propose mechanistic hypotheses that may explain the observed patterns. We conclude that temporal and spatial characteristics of the host’s macro-environment mediate microbial diversity.

Microbial symbionts influence animal physiology, evolution and health in a variety of ways1, yet factors governing global-scale patterns in diversity of host-associated microbes are not fully understood. The largest scale study to date found that the primary predictor of microbial diversity was whether the sample was host-associated versus free-living. Furthermore, for host-associated communities, animal versus plant hosts and gut versus skin were the strongest predictors of microbial communities1.

NATURE ECOLOGY & EVOLUTION | VOL 3 | MARCH 2019 | 381–389 | www.nature.com/natecolevol
This underscores the importance of host-association in shaping a unique subset of the earth's microbes; however, most host microbiome studies are from geographically and taxonomically focused studies. A predictive framework for these communities at the global scale is lacking, but vital for understanding ecology and evolution of host-associated microbiomes, detection of dysbiosis or elucidating changes in microbiome functions.

Generalizable rules for predicting free-living microbial taxa are increasingly being explored. For instance, studies have found that geographic ranges of environmental bacteria decrease towards the equator following Rapoport's Rule[64]. In addition, composition and diversity of environmental microbiomes vary with latitude[65-67] and are known to be structured by abiotic factors, such as salinity, pH, temperature, oxygen and nutrients[68-70]. Thus, while consistent patterns have been detected for environmental microbes, it is uncertain if generalizable rules govern composition of naturally occurring host-associated microbiomes. In modern humans, diet and lifestyle are important drivers of gut community similarity[71], and for mammals broadly, diet and host phylogeny are strong predictors of the gut microbiome[72,73].

Amphibian skin is a leading model system to explore host-associated microbial community structure. The skin can be sampled non-destructively, and the need to understand skin microbial ecology is hastened by emerging pandemic diseases[74-76]. In recent decades, amphibian species have been decimated by the invasive fungal pathogens Batrachochytrium dendrobatidis[77,78], and more recently, Batrachochytrium salamandrivorans[79,80]. Previous research on amphibians from aquatic systems has found that amphibian host identity is the strongest predictor of skin-associated bacteria, while developmental life stage and environment are secondary predictors[81-83]. Other studies have found host microhabitat preferences and ecological factors best predict amphibian skin microbiomes[84-87].

At a local scale, amphibian skin microbial diversity varies temporally[88-90] and is reduced when hosts are exposed to habitat destruction, microclimate shifts and captivity[91,92]. Indeed, the research community has generated substantial knowledge on microbial communities of particular amphibians, but typically has focused on small geographic areas, leaving out analysis of climatic variables. While these individual advancements are valuable, large global-scale datasets are needed to evaluate how environmental versus intrinsic factors mediate composition and diversity of amphibian-associated microbiomes. Undeniably, across all investigations into host-associated microbiomes, abiotic effects at the local scale are typically weak, although often statistically detectable, and the influence of climatic variables on microbiomes has rarely been reported[86,87,93-95].

To explore variables that may influence amphibian skin-associated microbial communities at the global scale, we used cutaneous microbiome data from 2,349 post-metamorphic amphibians. We analysed how multiple factors associated with an amphibian's biology, their abiotic and biotic environment, and their biogeography related to these communities. Second, we investigated bacterial richness and composition of the globally distributed, American bullfrog (Lithobates catesbeianus), to separate intrinsic host-related effects from extrinsic environmental effects shaping the skin microbiome. Last, we explored whether our observations agreed with specific, non-mutually exclusive mechanistic hypotheses that could account for the observed diversity patterns: (1) bacterial relative abundance patterns across important bioclimatic predictors will be associated with bacterial thermal growth optima[96], (2) bacteria with faster growth rates, have a competitive advantage over other bacteria and thus may reduce bacterial richness[97-99] and (3) natural environmental fluctuations associated with colder winter temperatures could create opportunities for bacterial turnover and favours dormancy, thus facilitating increased bacterial richness[100] and (4) temperature fluctuations may mediate competitive interactions, such as antibiotic production by microbes, which will influence microbial diversity[101].

To explore these hypotheses, we integrated bacterial community data with information on inferred optimal growth temperatures and quantified predicted functions associated with growth rates, dormancy and antibiotic production. Together, these data reveal global patterns of amphibian skin microbiomes and provide mechanistic insights that deepen our understanding of these communities.

Results
Bioclimatic correlates with richness of amphibian skin microbiomes. We built linear mixed models (LMMs) for bacterial richness (number of bacterial sub-operational taxonomic units, sOTUs) and evenness (Simpson's E), from a combination of biotic and abiotic factors including subsets of least-correlated bioclimatic predictors (Supplementary Table 5). Our preferred LMM (Fig. 1a and Supplementary Table 7) based on lowest Akaike Information Criterion (AICc) value included five bioclimatic variables, as well as amphibian species richness, latitude and elevation while controlling for four random factors: sequencing centre, host habitat class, collection habitat and host phylogeny (as amphibian family). The biotic variables, host phylogeny and microhabitat, were not included as fixed factors due to their inconsistent and possibly site-driven effects (see below). The highest coefficient value corresponded to minimum temperature of the coldest month (hereafter referred to as Bio6, see Supplementary Table 7). A partial effect analysis (Fig. 1b) revealed that bacterial richness negatively related to Bio6, as it did in an independent analysis of this variable (Pearson's r = −0.301; P < 0.001). In the multivariate context, that is, controlling for the very strong Bio6 effect, richness is predicted to increase with mean temperature of driest quarter (Bio9), and to decrease with latitude and altitude (Fig. 1b). These variables, however, have inverse relationships when analysed independently; richness decreased with Bio9 (r = −0.222; P < 0.001) and increased with latitude (r = 0.175; P < 0.001) and elevation (r = 0.188; P < 0.001).

Alternative models included mean annual temperature range (Bio7), which showed a positive correlation with bacterial richness (Supplementary Tables 7 and 8). Given the high correlation between these and many other bioclimatic variables (Supplementary Table 5), these results suggest that richness of amphibian skin microbiomes is higher in more seasonal environments with colder winter temperatures. The preferred LMM for Simpson's E had very low R² values for all variables (Supplementary Table 10), confirming that evenness was not strongly influenced by any of the predictors included in our study.

Path analyses provided additional support to our central finding that Bio6 had a strong effect on richness, indicating that (1) elevation strongly influenced Bio6, but only had weak direct effects on bacterial richness and (2) Bio6 influenced host richness and host phylogeny, but these two predictors had comparatively weak direct effects on bacterial richness (Fig. 1c,d).

Bioclimatic explains abundance of bacterial taxa in the amphibian microbiome. Bacterial community similarity based on sOTUs was only marginally influenced by bioclimate (Table 1), but at phylum level, the relative abundance of Proteobacteria increased, and that of several other bacterial phyla decreased with Bio6 (Fig. 2a). We used binomial mixed models to evaluate the effect of bioclimate on the 27 most abundant bacterial genera (greater than 0.5% relative abundance). Eighteen of these genera were negatively correlated with mean temperature of driest quarter (Bio9), and to decrease with latitude and altitude (Fig. 1b). These variables, however, have inverse relationships when analysed independently; richness decreased with Bio9 (r = −0.222; P < 0.001) and increased with latitude (r = 0.175; P < 0.001) and elevation (r = 0.188; P < 0.001).

Alternative models included mean annual temperature range (Bio7), which showed a positive correlation with bacterial richness (Supplementary Tables 7 and 8). Given the high correlation between these and many other bioclimatic variables (Supplementary Table 5), these results suggest that richness of amphibian skin microbiomes is higher in more seasonal environments with colder winter temperatures. The preferred LMM for Simpson's E had very low R² values for all variables (Supplementary Table 10), confirming that evenness was not strongly influenced by any of the predictors included in our study.

Path analyses provided additional support to our central finding that Bio6 had a strong effect on richness, indicating that (1) elevation strongly influenced Bio6, but only had weak direct effects on bacterial richness and (2) Bio6 influenced host richness and host phylogeny, but these two predictors had comparatively weak direct effects on bacterial richness (Fig. 1c,d).

Bioclimatic explains abundance of bacterial taxa in the amphibian microbiome. Bacterial community similarity based on sOTUs was only marginally influenced by bioclimate (Table 1), but at phylum level, the relative abundance of Proteobacteria increased, and that of several other bacterial phyla decreased with Bio6 (Fig. 2a). We used binomial mixed models to evaluate the effect of bioclimate on the 27 most abundant bacterial genera (greater than 0.5% relative abundance). Eighteen of these genera were negatively correlated with Bio6 (that is, increased in relative abundance with colder temperatures) (Fig. 2b). The standard deviation in slopes among genera was estimated to be 0.7774, and the main effect of Bio6 overlapped zero (maximum likelihood estimate: −0.4921, s.e.m. = 0.3114, Z = −1.580, P = 0.114). The effects of Bio6 also varied with latitude, with an estimated standard deviation of 2.910 (Supplementary Fig. 5). We found that the relative abundance of many bacterial genera produces...
Influences of host phylogeny and microhabitat on microbiome richness. Biotic factors, included in LMMs as random factors, contributed to explaining richness of amphibian skin-associated microbiomes, but revealed only limited globally applicable patterns. This is apparent from the strong effect of host microhabitat preference (although lower in coefficient value than bioclimatic factors) when included as a fixed effect (Supplementary Table 13). A detailed analysis of host microhabitat preference suggested this was caused by idiosyncratic effects in different geographical regions (Fig. 3; see Supplementary Fig. 6 for a more fine-scale categorization). For example, aquatic frogs had low sOTU richness in the USA but high values in Panama. As one moderately consistent pattern, arboreal amphibians in five countries had on average lower bacterial richness than terrestrial amphibians, and these differences were statistically significant in three countries (Brazil, Madagascar, Panama; $P <$ 0.05 in false discovery rate-corrected Wilcoxon $U$-tests; Fig. 3).

Despite a very strong effect of main host clades (families) in response screening (Supplementary Table 4), host phylogeny based on a nMDS proxy, was not a top predictor of bacterial richness when

Table 1 | Permutational multivariate analysis of variance (PERMANOVA) models of beta diversity, showing influence of selected predictors on weighted UniFrac distances among amphibian cutaneous microbiomes

| d.f. | Model A | Model B |
|------|---------|---------|
|      | Sum of sqs | $R^2$ | $F$ | Sum of sqs | $R^2$ | $F$ |
| Bio4 (temp. seasonality) | 1 | 1.06 | 0.00278 | 8.019 | 0.994 | 0.0035 | 7.460 |
| Bio6 (min. temp. coldest month) | 1 | 1.81 | 0.00475 | 13.695 | 1.219 | 0.00429 | 9.147 |
| Bio7 (annual temp. range) | 1 | 1.03 | 0.0027 | 7.782 | 1.114 | 0.00392 | 9.147 |
| Bio9 (mean temp. driest quarter) | 1 | 0.84 | 0.0022 | 6.323 | 0.461 | 0.00162 | 3.462 |
| Bio12 (annual precip.) | 1 | 2.62 | 0.00688 | 19.822 | 2.723 | 0.00957 | 20.435 |
| Bio14 (precip. driest month) | 1 | 4.15 | 0.01089 | 31.375 | 3.922 | 0.01379 | 29.437 |
| Bio18 (precip. warmest quarter) | 1 | 1.42 | 0.00372 | 10.714 | 1.86 | 0.00654 | 13.960 |
| Amphibian richness | 1 | 0.66 | 0.00175 | 5.029 | 0.453 | 0.00159 | 3.402 |
| Elevation | 1 | 1.44 | 0.00379 | 10.915 | 1.487 | 0.00523 | 11.162 |
| Latitude (absolute value) | 1 | 1.59 | 0.00417 | 12.008 | 1.099 | 0.00387 | 8.251 |
| Habitat class | 4 | 7.01 | 0.01842 | 13.268 | 5.775 | 0.02031 | 10.835 |
| Amphibian family/host phylogeny nMDS | 26/1 | 17.99 | 0.04727 | 5.237 | 1.4 | 0.00492 | 10.504 |

All continuous variables were rescaled before analysis. Models that differ by the host phylogeny proxy included: amphibian family in model A, host phylogeny non-metric multidimensional scaling (nMDS) variable in model B. PERMANOVA-based $P$-values of $F$ are $<0.001$ for both models and all predictors.
Between-site (other frogs) Within-site (bullfrogs) Between-site (bullfrogs) Within-site (bullfrogs)

Fig. 2 | Factors influencing skin microbiome composition. a, Relative abundances of the top 10 bacterial phyla along the gradient of minimum temperature of coldest month (Bio6). Proteobacteria, Bacteroidetes and Verrucomicrobia exhibited the strongest correlations with bioclimate ($r$ values are from Kendall–Tau correlation tests; all $P < 0.001$). Grey shading represents 95% confidence intervals. b, Effect estimates of Bio6 on dominant bacterial genera. Each maximum likelihood estimate is displayed as a point, and the dashed line indicates no effect. Bacterial genera responded differentially, most being more abundant with cold winters (negative values; blue dots) and some being less abundant (positive values; red dots). c, Boxplot of weighted Unifrac distances among bullfrogs and other sympatric amphibians from four countries (Brazil, Japan, South Korea and United States). Bullfrog microbiomes are significantly more similar to those of sympatric anurans than to those of bullfrogs from other locations (horizontal lines above the boxplots indicate significant comparisons; Monte Carlo approximation, Bonferroni-adjusted, all comparisons $**P < 0.001$). Boxplots display the first quartile, median, third quartile and maximum values along with outlier data points.

Bullfrogs mirror native amphibians in microbiome richness and beta diversity. The American bullfrog is globally distributed, allowing for the unique opportunity to explore skin microbiomes across disparate biogeographic regions and to compare bullfrogs to other co-occurring species in dissimilar regions. For this purpose, we collected 139 American bullfrog samples from Brazil, Japan, South Korea and the USA. Similar to our findings with the full dataset, American bullfrogs had higher bacterial richness in localities with lower minimum temperature of the coldest month (Bio6) (Fig. 1a). It is important to note, however, that bullfrog-specific data did not span the full range of the Bio6 gradient.

To examine host effects on patterns of beta diversity we calculated unweighted and weighted Unifrac distances between microbiomes of American bullfrogs and other sympatric amphibians and compared these to distances between allopatric populations of American bullfrogs. Comparisons of both distance metrics showed the same pattern. We found that pairwise distances among bullfrogs and other sympatric amphibians were smaller than pairwise distances among bullfrogs from different sites (Weighted Unifrac, Monte-Carlo approximation; $Z = 11.85, P < 0.001$ (Fig. 2c)). Additionally, pairwise distances among allopatric bullfrogs were only marginally different from pairwise distances among allopatric non-bullfrog samples. ($Z = 3.07, P < 0.001$; Fig. 2c). Analysis of core communities revealed that no sOTUs were shared among American bullfrogs across continents at $\geq 70\%$.

Indeed, across the full dataset no sOTUs were shared amongst 90 or 100% of the samples, and only one sOTU was shared among 80% of samples (a *Klebsiella* sp.). For the 27 most abundant bacterial genera (Fig. 2b), the effect of Bio6 in binomial mixed models was marginally correlated between the dataset comprising all host taxa and the bullfrog dataset (Pearson's product-moment correlation $-0.368, 95\%$ confidence interval of $-0.656$ and 0.014, $P = 0.0592$; Supplementary Fig. 11). While controlling for species-specific effects, these bullfrog-specific results support that bioclimatic and site-specific factors best explain variation in amphibian skin-associated microbial diversity.

Bacterial genes may explain correlation of bacterial richness with bioclimate. We used Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt) to explore potential hypotheses that could explain observed patterns of bacterial richness across bioclimatic gradients. This tool predicts the functional profile of the entire microbiome through matching 16S amplicon sequences to known bacterial genome data. It is therefore important to note the intrinsic limitations of this predictive framework. On average, 81% ($\pm 18\%$ s.d.) of the community was mapped to the Greengenes database required for PICRUSt analyses. Furthermore, amphibian skin microbiomes had suitable Nearest Sequence Taxon Index values (see Methods) validating their use in PICRUSt analyses. Using these data, we analysed average predicted bacterial genera (Fig. 2b), the effect of Bio6 in binomial mixed models was marginally correlated between the dataset comprising all host taxa and the bullfrog dataset (Pearson's product-moment correlation $-0.368, 95\%$ confidence interval of $-0.656$ and 0.014, $P = 0.0592$; Supplementary Fig. 11). While controlling for species-specific effects, these bullfrog-specific results support that bioclimatic and site-specific factors best explain variation in amphibian skin-associated microbial diversity.

**Factors influencing skin microbiome composition.** Relative abundances of the top 10 bacterial phyla along the gradient of minimum temperature of coldest month (Bio6). Proteobacteria, Bacteroidetes and Verrucomicrobia exhibited the strongest correlations with bioclimate ($r$ values are from Kendall–Tau correlation tests; all $P < 0.001$). Grey shading represents 95% confidence intervals. Effect estimates of Bio6 on dominant bacterial genera. Each maximum likelihood estimate is displayed as a point, and the dashed line indicates no effect. Bacterial genera responded differentially, most being more abundant with cold winters (negative values; blue dots) and some being less abundant (positive values; red dots). Boxplot of weighted Unifrac distances among bullfrogs and other sympatric amphibians from four countries (Brazil, Japan, South Korea and United States). Bullfrog microbiomes are significantly more similar to those of sympatric anurans than to those of bullfrogs from other locations (horizontal lines above the boxplots indicate significant comparisons; Monte Carlo approximation, Bonferroni-adjusted, all comparisons $**P < 0.001$). Boxplots display the first quartile, median, third quartile and maximum values along with outlier data points.
Fig. 3 | sOTU richness of skin bacterial communities among amphibians occupying different microhabitats, examined for six different countries. Boxplots show mean, standard deviation and range (outliers as grey dots). Significant differences are highlighted by asterisks (pairwise false discovery rate-corrected Wilcoxon U-tests per country, Bonferroni-corrected over five country comparisons: *P < 0.05). The boxplot displays the first quartile, median, third quartile and maximum values, along with outlier data points. For simplified representation, from the original five-category classification semiaquatic was merged with aquatic, and scansion with terrestrial, reflecting the main microhabitat of the respective species (full graph in Supplementary Fig. 6).

Fig. 4 | Correlations of gene copy number and predicted microbial functions with bioclimate. a. Average predicted 16S rRNA gene copy number increases with Bio6 (Kendall rank correlation, $r = 0.22, P < 0.0001$). b. Relative abundance of dormancy-associated gene pathways (summarized for sporulation, toxin/antitoxin and resuscitation) decreases with Bio6 (Kendall rank correlation, $r = −0.27, P < 0.0001$). Average copy number and relative gene abundances are calculated per microbial community, and values are represented as medians per sampling site. Grey shading represents 95% confidence intervals.

Discussion

This study expands on previous research by examining macro-ecological patterns of amphibian skin bacteria. Our data revealed that temperature-associated factors—in particular, cold winter temperatures and seasonality—consistently correlate with bacterial richness and to a lesser extent with bacterial composition on amphibian skin at the global scale. Our results reflect an inverse latitudinal richness-effect given that a simple regression analysis indicated decreasing richness at lower latitudes. This result probably occurs because temperature-related bioclimatic variables, such as Bio6, and latitude were highly correlated with each other across sampling localities. This pattern is in contrast with what is observed for most free-living macro-eukaryotes, including amphibians, but mirrors findings of bacterial communities from other environments. Only rarely have previous studies on bacterial communities found a conventional latitudinal effect (higher diversity at lower latitudes). For amphibians, we expect that most skin bacteria are environmentally acquired. Thus, the observed inverse latitudinal richness-effect could be a function of diversity patterns of environmental substrates. While there are several limitations to our design, such as lack of environmental substrate samples and variability in sampling date, our data provide evidence for a skin-associated diversity gradient in part explained by the latitude-associated temperature regime. This finding is further supported by analyses of American bullfrogs, which are globally distributed.

Our study also demonstrates that amphibian microhabitat usage influences skin bacterial richness, and that bacterial composition differs among coarse host taxonomic categories; that is, amphibian families. While at the local scale, previous studies have demonstrated host-specific patterns in amphibian skin bacteria, the nMDS proxy for amphibian phylogeny herein was less consistently associated with bacterial richness and composition. A low phylogenetic effect on microbiome composition was also found in an in-depth study of amphibian fauna sampled across Madagascar. Amphibian skin physiology and secretions are partly conserved across the amphibian tree of life. For instance, multiple unrelated amphibian families have been defined on the basis of important ecomorphological traits such as arboreality (Hyliidae, Rhacophoridae, Hyperoliidae) that may influence microbiome characteristics of constituent species.

Congruent with other studies, we found differences in microbial richness among frogs occupying different microhabitats, but these were only partly consistent across this dataset (Fig. 3). For various countries and latitudes, our data suggest that skin microorganisms

amino acid metabolism and biosynthesis of secondary metabolites. All the aforementioned gene pathways are well studied.

Bacterial taxa are known to code 1–15 ribosomal (rRNA) operons in their genome. Commonly referred to as rRNA copy number, this operon number is a robust and well-studied trait of bacteria that relates to bacterial growth rate and efficiency. Previously, rRNA copy number has been identified as an important variable explaining community composition of amphibian skin bacteria during amphibian development. We found that average predicted rRNA copy number was positively correlated with Bio6 (that is, greater in warmer climates, Kendall rank correlation, $r = 0.21, P < 0.0001$), and that dormancy-associated functional-gene abundance was negatively correlated with Bio6 (that is greater in colder climates, Kendall rank correlation, $r = −0.27, P < 0.0001$; Fig. 4). Last, gene abundance of antibiotic synthesis pathways was also negatively correlated with Bio6 (Kendall rank, $r = −0.23, P < 0.0001$; Supplementary Fig. 8).
in arboreal hosts are less rich than in terrestrial species, whereas microbiome richness on aquatic hosts varied substantially, as previously seen in Central American amphibians. It is possible that rinsing transient microbes off arboreal species is more efficacious than rinsing microbes off terrestrial species, and/or the higher diversity on terrestrial species reflects higher diversity of the soil environment that they inhabit. Alternatively, amphibian ecology could influence skin-shedding rates, secretion of skin defence compounds and skin structure, which in turn affect bacterial richness. These additional factors may correlate with species identity or major amphibian clades, such as families. In light of these independent effects of microhabitat, host species and phylogeny on the amphibian cutaneous microbiome, it is remarkable that the effects of bioclimate are relatively consistent at the global scale.

We found only a single prevalent bacterium (Klebsiella) in 80% of our samples and no core community for American bullfrogs across continents. This finding supports the hypothesis that amphibian skin microbiomes are strongly influenced by local bacterial source communities and abiotic conditions, including temperature. Amphibians can actively thermoregulate to some degree, such as sitting on a warm rock; however, this capacity is unlikely to outweigh the thermal environment on a global scale.

Amphibian tissues provide a rich source of resources for microbes. However, as ectotherms, they do not offer protection from seasonal temperature changes. Our results suggest that skin-associated bacteria of ectotherms are under environmental selection and that in cooler climates they are selected on to withstand temperatures outside of their growth range (for example, dormancy in cold climates). We hypothesized that natural environmental fluctuations associated with cold winter temperatures could favour dormancy and promote higher diversity by allowing unique bacterial taxa to become active during different times of the year and by allowing bacteria to take advantage of continual microbial turnover. Dormancy, a characteristic of many bacteria, has been comparatively well studied and is highlighted as a factor influencing bacterial biogeography. In our hypotheses about why we observe an inverse latitudinal diversity gradient, we assert that a combination of bacterial characteristics in these communities may explain this pattern. First, higher predicted rRNA copy numbers, which signifies fast growth, were found in warm, stable thermal environments, suggesting that taxa in these environments are able to out-grow and potentially exclude other bacterial taxa (Fig. 4). This result is linked to a reduction in the richness of bacteria on the skin of amphibians in these environments. Second, we hypothesized that a periodic resurgence from the microbial seed bank on amphibian skin, facilitated through dormancy, may bring about higher richness on amphibians inhabiting regions with seasonal thermal changes. In support of this hypothesis, we found that dominant bacterial genera and abundance of dormancy genes found on amphibian skin were non-randomly distributed across a temperature (Bio6) gradient (Fig. 4). Indeed, dormant or slow-growing bacteria are more prevalent in environments with seasonal temperature variation, which probably affects nutrient or growth conditions. This hypothesis is consistent with our finding that bacteria with lower thermal growth optima are more abundant on amphibians in regions with colder winter temperatures (see Supplementary Results). Importantly, estimating thermal optima of bacterial genera within our dataset from databases of bacterial thermal optima is not a direct comparison, and thus does not match the exact conditions of amphibian skin. Further details on these limitations are discussed in the Supplementary Information. Alternatively, or as a compounding effect, chemical disturbances (for example, antibiotic synthesis) in these cooler environments may also play a role in shaping global diversity patterns of bacterial symbionts of ectotherms. We hypothesized that moderate disturbance via chemical antibiotic production from microbiomes may create a more heterogeneous landscape, facilitating open niches, niche specialization and ultimately greater microbial diversity. Again, we found that predicted antibiotic synthesis gene abundances found on amphibian skin were non-randomly distributed across the Bio6 gradient (Supplementary Fig. 8). Direct measurements of functional genes are required to confirm our results obtained from PICRUSt predicted gene functions.

Future studies may extend these findings in a variety of ways. As new datasets of both endotherms and ectotherms become available, a meta-analysis including both groups could provide greater insight into either the generality or specificity of our findings. For example, a strong effect of bioclimate on skin microbiomes of endotherms is unlikely given that temperature fluctuations in cold environments are less extreme for microbiomes associated with most endothermic animals, underscoring the importance of studying a broad diversity of animal taxa to understand global host-associated diversity patterns. Future studies could also explore the influence of bioclimate on host-associated microbial communities both within and across vertebrate and invertebrate groups. As microbial genomic databases become better equipped and sequencing of bacterial gene content from whole communities becomes more commonplace, future work could address hypotheses related to ours with sample-specific microbial genomic data. Experimental translocation and temperature manipulations of amphibians could also test for selection of amphibian-associated microbial phenotypes and genotypes.

Our results indicate that amphibian skin bacterial composition changes across a bioclimatic gradient, and that bacterial richness per host individual decreases towards warmer, more stable thermal environments. However, due to the compositional nature of sequence data, we acknowledge that changes in abundance of specific bacterial taxa across bioclimatic gradients are influenced by changes of other bacterial taxa. For this reason, our results focus primarily on the global richness patterns across climatic gradients. Future sequencing projects could include a DNA spike-in during sequencing, which enables better estimation of absolute microbial abundances for among-sample comparisons.

Bioclimatic variables, in particular, minimum temperature of the coldest month and seasonal temperature variation, consistently correlated with cutaneous microbiomes at the global scale. The importance of this aspect of bioclimate in shaping host-associated microbiomes was previously unknown. Our data help explain fundamental questions of microbial biogeographical diversity and offer insights into how climatic variation may affect host microbiomes. In the face of rapid environmental change around the globe, climatic changes may alter host microbiomes, which, in turn, could have consequences on maintenance of host health and selection and evolution of amphibians.

**Methods**

**Summary of the metanalysis and newly sampled amphibians.** We assembled samples from 2,349 individual post-metamorphic amphibians, comprising 27 amphibian families (205 species) collected across 13 countries (five continents) (including 538 samples newly sequenced for this study). A summary of amphibian sampling effort across continents is provided in Supplementary Table 1 and Supplementary Fig. 1. All amphibians were swabbed using sterile swabs and DNA directly extracted from these. The V4 region of the 16S rRNA gene was amplified with barcoded primers (515F–806R) and sequenced on Illumina MiSeq platforms (details in the Supplementary Information). Raw sequence data was compiled from newly sequenced datasets and from published studies (Supplementary Table 1).

Sequences were quality filtered and further analysed in Quantitative Insights into Microbial Ecology (QIIME) and deblur. sOTUs were determined using the deblur workflow (https://github.com/biocore/deblur). After filtering and decontamination procedures, the dataset comprised 45,932,673 reads and an average of 19,554 reads per sample. Samples were subsequently rarefied at 2,500 reads per sample and had an average of 277 sOTUs.

For analyses, these samples were subdivided into different subsets: (1) the full dataset with samples from all 2,349 amphibian individuals; and (2) a dataset of 1,801 individuals for which host phylogenetic data were available (and excluding American bullfrogs). Additionally, an American bullfrog dataset, comprising
139 samples of this species ranging from 29.5°S–42.4°N latitude, was analysed separately to control for amphibian species identity while exploring biogeographic predictors of skin bacterial communities. We also compiled a standardized dataset with 828 individuals, representing a more even sampling of 7–10 samples per host species, for calculating core communities.

Richness analysis of amphibian skin bacterial communities. To analyse the correlation of abiotic and biotic factors with alpha diversity of amphibian skin microbiomes we used QIIME to calculate two response variables, number of observed OTUs and sOTUs representing community richness and Simpson's R estimating community evenness. For abiotic predictors, we chose 19 bioclimatic variables as well as absolute latitude and elevation. In addition, we included biotic variables, such as amphibian species richness and host phylogeny. Furthermore, we controlled for the effect of selected categorical variables including sequencing centre, host microhabitat preference and collection habitat of hosts. A full discussion of predictor variables is provided in the Supplementary Methods (also see Supplementary Table 3). Bioclimatic data was extracted from 1-km spatial resolution climate surfaces for global land areas46. Host richness was approximated by extracting amphibian richness data from available maps at 10x10 km resolution71. Host phylogeny was alternatively represented by (1) a categorical variable using amphibian family as a taxonomic proxy, or (2) evolutionary divergences, a variable obtained by nMDS. Evolutionary divergences among host species were calculated as patristic phylogenetic distances from an ultrametric time-tree recovered from the timetree.org database68. To include host phylogeny in models we created a nMDS proxy for host phylogeny, constrained to one dimension, on the patristic distance matrix (Kruskal's Stress 1 = 0.0323). nMDS response (one-dimensional-ordination explaining 20% of the variation) showed no outliers, such that closely related families have similar values (Supplementary Fig. 12).

To assess the effect of all predictors on richness and evenness of bacterial communities (number of sOTUs), we first used Response Screening adjusted for false discovery rate in JMP 13.0 (SAS Institute). Most bioclimatic predictor variables were strongly correlated, and we therefore applied various strategies to compile sets of variables that were strong predictors of the data, least-correlated, or potentially biologically informative on the basis of a priori assumptions. The selected subsets of variables were implemented in alternative LMMs in R using the lme4 package72,73 and then evaluated on the basis on AICc, R² and variance inflation factor values. Coefficient values for all predictors were obtained through the fixef() function in the nlme package. The sjPlot package was used to created model estimate and effect plots (Fig. 1 and Supplementary Fig. 2). All statistical significance is reported using a two-tailed approach. See Supplementary Information for a full description of model selection procedures, predictor variables included in each model and coefficients and variance inflation factor values for each variable.

To understand directionality and confirm the relative strength of predictors selected in our LMM model selection procedures, we built six ecologically meaningful path models that included a combination of variables directly or indirectly affecting sOTU richness; Bio6, Bio67 and elevation, as well as two biotic variables, host phylogeny and amphibian species richness (Fig. 1d,e and Supplementary Fig. 3). These confirmatory models were primarily designed to better understand the interactions of the predictor variable found to be most influential (Bio6) with biotic predictors; that is, host richness and host phylogeny. Path models with the response variable Bio6 averaged by collection site were also performed (Supplementary Fig. 4). From a correlation matrix, estimates of standardized path coefficients with their associated standard errors were derived by maximum Wishart likelihood (500 iterations), allowing the identification of significant paths. Latent (unmeasured ‘υ’) variables, corresponding to variance attributed to error and any unmeasured predictors, were estimated for each response variable in the model.

Compositional analysis of amphibian skin microbiomes. A phylum-level taxonomic summary of amphibian skin microbiomes by host species (n = 205 species) within each country is provided in Supplementary Table 2. We used QIIME to calculate beta diversity as weighted Unifrac distances. Factors driving patterns of beta diversity were investigated with PERMANOVA* estimating Pseudo-F and P values with marginal effects.

To understand differentiation of American bullfrogs across sites relative to their differentiation from sympatric amphibians we selected samples from only those locations from which bullfrog data were available. We used the make_distance_boxplots.py script in QIIME to calculate pairwise Unifrac distances among different categories of interest, and used Monte-Carlo approximations in R17, adjusted for false discovery rate, to identify significant differences among these categories. The relative abundance of each of 27 most abundant bacterial genera (overall relative abundance >0.5%) across Bio6 was analysed by three alternative generalised linear mixed-effects models that differed in random effect structures (Supplementary Information). Models were fit using a binomial likelihood with the glmr function in the lme4 package and chosen by performance based on AICc. We accounted for spatial differences via a random intercept for binned longitude and latitude and allowed the effects of minimum temperature of coldest month (Bio6) to vary among frog genera and latitude via a random slope. Due to the compositional nature of the data, the observation that some taxa decrease along the climatic gradient, while others increase, is just one of many potential underlying dynamic processes that could yield such biotic responses. Therefore, it is not possible to distinguish whether there are true changes to the community in both directions, or whether a few taxa are changing substantially in one direction and influencing the proportional abundance of another taxa selection that may appear to change in the opposite direction.

Predicted thermal optima, bacterial growth rates, dormancy genes and antibiotic synthesis genes. Kendall–Tau rank correlations were run on the relative abundance of all bacterial genera with a representation greater than 0.1% and the strongest predictor variable of bacterial richness, Bio6. For these genera, information on thermal optima of bacteria74 was obtained from available databases. The thermal optima for bacterial species in the studied amphibian microbiomes are not known directly but were predicted from data of other species in the same genera, studied under laboratory conditions. For a given species, all isolates from a given database were first averaged together to provide one temperature per species. All averaged species temperatures for the genera of interest were then extracted and used for analysis (see Supplementary Methods). This procedure was implemented to minimize over-representation of particular species within a given genus. Mann–Whitney U-tests were then used to compare thermal optima and the variance in the thermal optima between genera that were positively and negatively correlated with Bio6.

Accessed PICRUST to estimate bacterial gene function, see Supplementary Methods for details of sOTU clustering and sample selection. In PICRUST, we normalized the dataset by predicted rRNA copy number and then predicted the metageneome of each sample to investigate functional abundance of dormancy and antibiotic synthesis pathways. Dormancy analyses included all KOs contributing to sporulation, toxins and antibiotics and resuscitation75. Antibiotic KOs were extracted from KEGGS antibiotic synthesis category. Kendall–Tau correlations were used to explore the relationship between these functions and our main predictor, Bio6. In addition to these functional abundance analyses, we directly explored the average predicted rRNA copy number within a microbiome and how it correlates with Bio6. Predicted rRNA copy number is frequently used to estimate bacterial growth rates76. Amphibian samples in this study had sufficient Nearest Neighbour Taxon Index (NSTI) values for analyses (mean = 0.069 ± 0.034, median = 0.061 ± 0.040). In context and according to a previous study, human-associated samples had the lowest (best) NSTI values (0.03 ± 0.2), whereas mammalian guts and soil samples had much higher (worse) NSTI values, (0.14 ± 0.06 and 0.17 ± 0.02), respectively. Importantly, NSTI values of 0.1 for 16S RNA marker gene surveys and shotgun metagenomes still resulted in an Spearman ρ of roughly 0.8 and was considered an accurate gene category assignment77. Furthermore, NSTI is an aggregate measure based off of branch length and does not correspond to sequence similarity.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
A full description of data analyses is provided in the Supplementary Information. Data for all newly sequenced samples is available on the Short Read Archive (Bioproject PRJNA474496). All figures include associated raw data and there are no restrictions on data availability. Correspondence and requests for materials should be addressed to M.V. or D.C.W.

Received: 19 February 2018; Accepted: 6 January 2019; Published online: 18 February 2019

References
1. Thompson, L. et al. A communal catalogue reveals Earth’s multiscale microbial diversity. Nature 531, 457–463 (2016).
2. Hird, S. M. Evolutionary biology needs wild microorganisms. Front. Microbiol. 8, 1–10 (2017).
3. Jiménez, R. R. & Sommer, S. The amphibian microbiome: natural range of variation, pathogenic dysbiosis, and role in conservation. Biodivers. Conserv. 26, 763–786 (2016).
4. Lear, G. et al. Following Rapoport’s Rule: the geographic range and genome size of bacterial taxa decline at warmer latitudes. Environ. Microbiol. 8, 3152–3162 (2017).
5. Amend, A. S. et al. Macroecological patterns of marine bacteria on a global scale. J. Biogeogr. 40, 800–811 (2013).
6. Baldwin, A. J. et al. Microbial diversity in a pacific ocean transect from the arctic to antarctic circles. Appl. Environ. Microbiol. 41, 91–102 (2005).
7. Fuhrman, J. A. et al. A latitudinal diversity gradient in planktonic marine bacteria. Proc. Natl Acad. Sci. USA 105, 7774–7778 (2008).
8. Tedersoo, L. & Nara, K. General latitudinal gradient of biodiversity is reversed in ectomycorrhizal fungi. New Phytol. 185, 351–354 (2010).
39. Webster, N. S. et al. Host-associated coral reef microbes respond to the cumulative pressures of ocean warming and ocean acidification. *Sci. Rep.* 6, 19324 (2016).
40. Olken, P. A., Morrow, K. M., Willis, B. & Bourne, D. Implications of ocean acidification for marine microorganisms from the free-living to the host-associated. *Front. Mar. Sci.* 4, 47 (2017).
41. Longo, A. V. & Zamudio, K. R. Temperature variation, bacterial diversity and fungal infection dynamics in the amphibian skin. *Mol. Ecol.* 18, 4787–4797 (2017).
42. Longo, A. V. & Zamudio, K. R. Environmental fluctuations and host skin bacteria shift survival advantage between frogs and their fungal pathogen. *ISME J.* 11, 349–361 (2017).
43. Novakova, E. et al. Mosquito microbiome dynamics, a background for prevalence and seasonality of West Nile virus. *Front. Microbiol.* 8, 526 (2017).
44. Corkney, R. et al. The biosignature kinetic spectrum for temperature. *PLoS ONE* 11, 1–29 (2016).
45. Roller, B. R., Stoddard, S. F. & Schmidt, T. M. Exploiting rRNA operon copy number to investigate bacterial reproductive strategies. *Nat. Microbiol.* 18, 386–392 (2015).
46. Mengert, D. R. et al. Decreases in average bacterial community rRNA operon copy number during succession. *ISME J.* 10, 1147–1156 (2016).
47. Kerns, P. & Shade, A. Trait-based patterns of microbial dynamics in dormancy potential and heterotrophic strategy: case studies of resource-based and post-press succession. *ISME J.* 12, 2575–2581 (2018).
48. Chzar, T., Hoekstra, R. F. & Pagie, L. Chemical warfare between microbes. *PNAS* 99, 786–790 (2002).
49. Lennon, J. T. & Jones, S. E. Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nat. Rev. Microbiol.* 9, 119–130 (2011).
50. Langille, M. G. et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* 31, 814–821 (2013).
51. Mittelbach, G. G. et al. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecol. Lett.* 10, 315–331 (2007).
52. Wiens, J. J. Global patterns of diversification and species richness in amphibians. *Am. Nat.* 170, 586–5106 (2007).
53. Prest, T. L., Kimball, A. K., Kueneman, J. G. & McKenzie, V. J. Host-associated bacterial community succession during amphibian development. *Mol. Ecol.* 27, 1992–2006 (2018).
54. Bahram, M. et al. Structure and function of the global topsoil microbiome. *Nature* 560, 233–237 (2018).
55. Staddon, W. J., Trevers, J. T., Duchesne, L. C. & Colombo, C. A. Soil microbial diversity and community structure across a climatic gradient in western Canada. *Biodivers. Conserv.* 7, 1081–1092 (1998).
56. Louden, A. H. et al. Microbial community dynamics and effect of environmental microbial reservoirs on red-backed salamanders (*Plethodon cinereus*). *ISME J.* 8, 830–840 (2014).
57. Walke, J. B. et al. Amphibian skin may select for rare environmental microbes. *Cell* 155, 472–473 (2014).
58. Bleden, L. K. et al. Panamanian frog species host unique skin bacterial communities. *Front. Microbiol.* 6, 1171 (2015).
59. Sabino-Pinto, J. et al. Composition of the cutaneous bacterial community in amphibians: effects of captivity, host species, and body region. *Microb. Ecol.* 72, 460–469 (2016).
60. Rebollar, A. A. et al. Skin bacterial diversity of Panamanian frogs is associated with host susceptibility and presence of *Batrachochytrium dendrobatidis*. *ISME J.* 10, 1682–1695 (2016).
61. Bagley, S. T. Habitat association of *Klebsiella species*. *Infect. Control.* 6, 52–58 (1985).
62. Loevy, K. J., Fisk, M. C. & Lennon, J. T. Microscale insight into microbial seed banks. *Front. Microbiol.* 7, 2040 (2017).
63. Stubbendieck, R. M., Vargas-Bautista, C. & Straight, P. D. Bacterial communities: interactions to scale. *Front. Microbiol.* 7, 1–19 (2016).
64. Fick, S. E. & Hijmans, R. J. *WorldClim 2*: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* 37, 4302–4315 (2017).
65. Caporaso, J. G. et al. *QIIME* allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336 (2010).
66. Amir, A. et al. Deblur rapidly resolves single-nucleotide community sequence variation. *PLoS Biol.* 12, e1005796 (2016).
67. Wheeler, D. L. et al. The NCBI SRA to represent the current state of knowledge. *Scientific Data* 4, 170118 (2017).
68. Hedges, S. B., Dudley, J. & Kumar, S. TimeTree: a public knowledge-base of life’s diversity. *Proc. Natl Acad. Sci. USA* 107, 19054 (2010).
69. Pace, M. L. *Microbial diversity*. *Science* 332, pii: 20170582 (2017).
70. R Core Team, R. *R": A Language and Environment for Statistical Computing (R Core Team, 2016).
71. Bates, D. M., Machler, M., Bolker, B. & Walker, S. C. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48 (2017).
72. Lidecke, D. *Data visualization for statistics in social science*. R package https://doi.org/10.5281/zenodo.1308157 (2017).

---

388

**ARTICLES**

**NATURE ECOLOGY & EVOLUTION** | VOL 3 | MARCH 2019 | 381–389 | www.nature.com/natecolevol
Acknowledgements
This study was supported by grants of National Science Foundation (DEB-1146284 to V.J.M.; IOS-1121758 to L.R.-S.; DEB-1310036 to A.V.L.), Templeton Foundation to V.J.M., Deutsche Forschungsgemeinschaft (DFG) to M.V. (VE247/9-1), CAPES to M.V. and C.F.B.H., FAPESP (2013/50741-7) and CNPq to C.F.B.H., Simons Foundation (429440, WTW) to J.G.K., Deutscher Akademischer Austauschdienst (DAAD) to M.C.B., University of Costa Rica (Project 801-B2-029) and Costa Rican Ministry of Science and Technology (849-PINN-2015-1) to J.G.A., the Portuguese National Funds through FCT (Exploratory Research Project: IF/00209/2014/CP1256/CT0011) to A.C. and the National Research Foundation of Korea (2015R1D1A1A01057282) to B.W. We are indebted to M. Kondermann for her assistance with laboratory work, A. Borzee, R. Kakehashi and T. Kosch for their support in sample collection, and the Organization for Tropical Studies, La Selva Biological Station and G. Alvarado for field support in Costa Rica. All new sampling was done with the appropriate permits of national authorities, where required; in Costa Rica, of the Institutional Biodiversity Commission of the University of Costa Rica (Resolution 371 014) and the Costa Rican Ministry of Environment and Energy (Resolution 091-2012-SINAC).

Author contributions
J.G.K., M.C.B., V.J.M., D.C.W. and M.V. conceived the study, coordinated the analyses and wrote the manuscript. J.G.K., M.C.B., D.C.W., G.B. and M.J. designed and performed data analysis. J.G.A., A.B., M.B., I.B., A.C., C.F.B.H., R.N.H., W.H., M.H., L.R.S., J.K., J.K., A.K., A.L., A.H.L., D.M., J.J.N., B.G.B.P., A.P.T., F.C.E.R., E.A.R., A.R., I.R.S., G.V.A., B.W., J.B.W., S.M.W., K.Z., I.Z.C. contributed materials and data. H.A., L.A., R.G. and M.J. performed laboratory work. P.J.K., R.S. and C.J.C. contributed to data analysis. All authors contributed to the development and revision of the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41559-019-0798-1.

Correspondence and requests for materials should be addressed to D.C.W. or M.V.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
- Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted
  Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated
- Clearly defined error bars
  State explicitly what error bars represent (e.g. SD, SE, CI)

Software and code

Policy information about availability of computer code

Data collection: No code or software was used for data collection.

Data analysis: We used Qiime version 1.9, Deblur, R version 3.3.2, Picrust version 1.1.3, SPSS v24, JMP 13.0 and Systat 13.2

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

A full description of data analyses is provided in the electronic supplementary material.

Data for all newly sequenced samples will be available on the Short Read Archive (Bioproject number will be added upon manuscript acceptance). Note that we have not yet submitted all of the data as we cannot exclude the reviewers requesting inclusion or exclusion of some data, and thus will upload the final data once
reviewer comments and editorial decision are received. The raw data are in principle not relevant for manuscript review, but we would of course make them available to reviewers if requested. Note that we also submit a detailed table with class-level OTU distribution per species which provides very detailed additional information. All figures include information on associated raw data and there is no restriction on data availability.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences

Study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

In this analysis we used all available data within the scope of this project. For new sequences, we sequenced all available samples from a certain locality and species. Sample sizes within groups/species are based on the number of samples needed in previous studies to partition microbial sample variation.

Data exclusions

We did no exclude any samples that met our sample/sequence criteria. Bacterial sequences identified as potential contaminants were excluded, and details of these exclusions are reported in the manuscript (all sequences are available in the raw data that have been made available through SRA - Bioproject number is provided).

Replication

NA. All data was included in this meta-analysis.

Randomization

NA. All data was included in this meta-analysis and no experimental procedures were conducted.

Blinding

NA. All data was included in this meta-analysis.

Materials & experimental systems

Policy information about availability of materials

n/a Involved in the study

Unique materials

Antibodies

Eukaryotic cell lines

Research animals

Human research participants

Research animals

Policy information about studies involving animals. ARRIVE guidelines recommended for reporting animal research

Animals/animal-derived materials A table of this data is provided in the supplement.

Method-specific reporting

n/a Involved in the study

ChIP-seq

Flow cytometry

Magnetic resonance imaging