Predicting nutrient content of ray-finned fishes using phylogenetic information

Bapu Vaitla\(^1\), David Collar\(^2\), Matthew R. Smith\(^3\), Samuel S. Myers\(^3,4\), Benjamin L. Rice\(^5\) & Christopher D. Golden\(^1,3\)

Human food and nutrition security is dependent on marine ecosystems threatened by overfishing, climate change, and other processes. The consequences on human nutritional status are uncertain, in part because current methods of analyzing fish nutrient content are expensive. Here, we evaluate the possibility of predicting nutrient content of ray-finned fishes using existing phylogenetic and life history information. We focus on nutrients for which fish are important sources: protein, total fat, omega-3 and omega-6 fatty acids, iron, zinc, vitamin A, vitamin B12, and vitamin D. Our results show that life history traits are weak predictors of species nutrient content, but phylogenetic relatedness is associated with similar nutrient profiles. Further, we develop a method for predicting the nutrient content of 7500+ species based on phylogenetic relationships to species with known nutrient content. Our approach is a cost-effective means for estimating potential changes in human nutrient intake associated with altered access to ray-finned fishes.
Human activity is rapidly transforming marine ecosystems worldwide. Total catch appears to be declining in fisheries around the world, largely due to intense industrial pressure. Climate change-driven increases in sea temperature threaten to alter the abundance and distribution of hundreds of economically and nutritionally important species, and are leading to mass coral bleaching. Ocean acidification, caused by greater absorption of atmospheric carbon dioxide, is inflicting severe damage on biologically rich coral reef habitats. Coastal development, pollution, and other anthropogenic forces are impacting nearly every marine ecosystem in the world.

The consequences of these changes for human health, and especially for nutrient intake, are likely to be severe because fish provide critical nutrients essential to human nutrition, including iron, zinc, vitamin A, vitamin B12, omega-3 and omega-6 fatty acids, and others. In many societies, seafood is the foundation for healthy diets, and its decline presents a significant risk in destabilizing food and nutrition security. The consumption of fish is associated with a wide range of health benefits, including the prevention of various non-communicable diseases and the promotion of cognitive development. Ray-finned fish—the class Actinopterygii—are particularly important to low-income populations, comprising 80.6% of the total global tonnage of subsistence marine capture fish and seafood in 2010.

A lack of information about the nutrient composition of fish species, however, hampers quantification of the nutritional threat to human populations of reduced consumption of wild-harvested fish. Measuring nutrient content is expensive and, as a result, nutrient analyses rarely capture the full breadth of vitamins, minerals, and macronutrients relevant to nutrition. This is exacerbated by the need to collect multiple samples of each species to assess variability across individuals, as well across subpopulations of species. This information gap prevents the design of rational fisheries management strategies and nutritional interventions to optimize public health outcomes in the face of rapidly changing marine conditions.

In this paper, we investigate the possibility that phylogenetic relatedness and life history information explain variation in the nutrient content of key fish species in the most commercially and nutritionally important class, Actinopterygii (ray-finned fishes). To our knowledge, this is the first time that such an approach has been explored. We then use the results of this analysis to develop a method for using shared phylogenetic history as a means of predicting the nutrient content of fish whose nutrient information has not yet been assessed. Such predictions are especially critical in regions of the world with known nutrient deficiencies—which are often the same areas where laboratory capacity is often limited. For fish specifically, recent years have seen an emerging desire among policymakers to design fisheries management and aquaculture development interventions with the specific goal of enhancing nutritional security. We note that these methods, if successful, could be used for a broad range of terrestrial animal and plant species as well—wild and domestic—as well as sub-species, varieties, and breeds.

Our results indicate that life history traits predict species nutrient content only weakly; larger datasets and/or use of missing-data imputation techniques may strengthen these models. We find, however, that phylogenetic relatedness is associated with similar nutrient profiles. This finding allows us to create a method for predicting the nutrient content of 7500+ Actinopterygii species based on phylogenetic relationships to species with known nutrient content.

Results

Interspecific variability in nutrient content and life history. We sourced nutrient content information for 371 species in the class Actinopterygii (see Supplementary Table 1 for full list of species, and see Methods section and refs. 16–28 for data sources). Supplementary Table 2 summarizes the number of species with data, as well as the observed range by nutrient. The dataset includes 26 orders, 126 families, and 279 genera; about 42% of species are in the order Perciformes (perch-like fishes). The set of 371 species represents over half of all global capture fisheries by weight, with all 22 of the world’s most harvested marine fish species included.

We used life history traits and phylogenetic relatedness to predict protein, total fat, omega-3 and omega-6 fatty acids, iron, zinc, vitamin A, vitamin B12, and vitamin D content in Actinopterygii. These nutrients are critical for human nutrition and are generally present in relatively high concentrations in seafood—although, as we describe below, not all species are rich sources of the all the above nutrients. As predictors of nutrient content, we initially considered six continuous and two categorical life history traits describing body size, trophic level, habitat characteristics, and geographic range (Table 1).

Summaries of nutrient content across species indicate that ray-finned fishes make substantial contributions to human nutrition, particularly for low-income coastal populations (Supplementary Tables 2, 3). Of species with available information for a given nutrient, 98% are sources or rich sources of protein, 94% rich sources of vitamin D, and 81% rich sources of vitamin B12. In addition, 13% of species are either sources or rich sources of iron, 14% sources or rich sources of zinc, and 10% sources or rich sources of Vitamin A (see Supplementary Table 1 for data by species and threshold designations of source and rich source by nutrient).

We see a great deal of variation across ray-finned fish species for which we have nutrient information (Supplementary Table 3). In particular, the ranges of minimum and maximum depth, habitat preferences, and latitudinal range are especially wide. Nearly 40% of fish are demersal, but significant fractions are benthopelagic, pelagic, or reef-associated. The chosen set is divided roughly equally into tropical, subtropical, and temperate species. As calculated by trophic level, the mean Actinopterygii species in this dataset is a carnivore, and some are apex predators, but many species are also herbivores, zooplanktivores, detrivores, or omnivores.

Correlations among nutrients and life history variables. We first examined the relationship between nutrient content and life history variables by estimating evolutionary correlations (i.e., the correlations between evolutionary changes inferred using observations among species and their phylogenetic relationships; see Methods section). For this analysis (and all regression models described below), we used the ray-finned fish phylogeny reconstructed in Rabosky et al. Due to sample size restrictions, we exclude omega-3 and omega-6 fatty acids from the correlation matrix below; however, correlations with fatty acids included are given in Supplementary Table 4.

We found the expected associations between pairs of life history variables (top left corner of Fig. 1, lightly shaded). Maximum depth, maximum fish length, and trophic level are moderately associated with each other, suggesting that larger fish tend to live at greater depths and consume a diet higher on the food chain. The a and b parameters describing species’ mass-length scaling relationships are well-correlated to each other, but not to other life history traits.

Patterns are also apparent within the set of nutrients (bottom right corner of Fig. 1, darkly shaded). Fish with more protein tend to contain less Vitamin A and total fat. Fish that are high in total fat also tend to be good sources of Vitamins A, B12, and D. Fish high in iron also tend to be good sources of zinc and all included.
vitamins. Vitamin A and Vitamin D are positively associated, and the latter is correlated with Vitamin B12. Overall, this set of correlations suggests that specific sets of nutrients tend to cluster together. Given that nutritional analyses, due to cost considerations, tend to focus on a small set of nutrients, such observed correlations may be important for roughly inferring levels of unmeasured nutrients.

Finally, we see that life history parameters, controlling for phylogeny, vary in their association to key nutrients (top right of Fig. 1, unshaded). Fish living at greater oceanic depths contain more total fat and Vitamin A, but less zinc. Longer fish contain proportionally less zinc and Vitamin B12, but more Vitamin A; smaller fish, often overlooked in conservation, livelihoods, and nutrition programs, may be richer sources of key micronutrients.

These bivariate associations were tested in multiple regression models fit by phylogenetic least squares (PGLS) in a manner that accounts for phylogenetic signal in species residuals.

These phylogenetic regressions find that life history parameters are either insignificant predictors of nutrient content or have small effects (Table 2). Longer fish at higher trophic levels but closer to the surface contain more protein, although the coefficient magnitudes are relatively small: 100 cm greater length is associated with 0.69 g more protein; a one-unit higher trophic level is associated with 1.29 g more protein per 100 g of fish weight; and 100 meters greater (i.e., deeper) maximum depth is associated with 0.14 g less protein. Maximum depth is also a significant predictor for total fat, omega-3 and omega-6 fatty acids, and vitamin A, but regression coefficients are small in these cases as well. Other relationships are insignificant. Overall, regression models suggest that life history variables are generally poor predictors of fish nutrient content when accounting for phylogenetic dependence in the data.

### Predicting nutrient content of unmeasured species

All nutrients exhibit significant phylogenetic signal; that is, covariance among species in nutrient content tends to be moderately to strongly associated with phylogenetic relatedness. Estimates of Pagel’s λ vary are highest (indicating strong phylogenetic signal) in zinc, vitamin D, and vitamin A, intermediate in total fat, omega-3 and omega-6 fatty acids, and vitamin B12, and weaker in protein and iron (Table 3).

The phylogenetic structuring of nutrient content can be seen in the heatmap of Fig. 1, which shows some clustering of nutrient values among closely related species, although the location of these clusters varies depending on nutrient. For example, various species of the genus *Thunnus* contain high levels of protein, as indicated by the dark red region near the top of the phylogeny in the first column of the heatmap. We also observe such clustering at close phylogenetic distance but across genera; for example, *Engraulis encrasicolus*, *Tricholodon hakoensis*, *Misgurnus anguillicaudatus*, and *Anguilla japonicus* all fall within the top 10% of zinc levels in the database, as indicated by the dark red region near the bottom of the phylogeny in the fourth column of the heatmap; other nearby species also contain relatively large amounts of zinc.

Given our findings that nutrient content is weakly predicted by life history but tends to be phylogenetically structured, we developed a procedure for predicting nutrient content in unmeasured species based on their phylogenetic relationships to measured species and empirical estimates of phylogenetic signal (see Methods for details). We used a jacknifing procedure to determine that this method has acceptable accuracy. Observed nutrient values fell within the prediction 95% confidence intervals for at least 89.7% of cases across all variables (see Fig. 2, as well as Supplementary Fig. 1 for residual plots) and median deviations between predicted and observed values were within 40% of the among-species standard deviation. We note that the method can lead to wide confidence intervals when long phylogenetic branches separate unmeasured from measured species. This prediction method also compares favorably to predictions based on life history regression models; average accuracy and median

---

**Table 1 Nutrient and life history traits used in the analysis**

| Nutrients | Description |
|-----------|-------------|
| Protein   | “Serves as the major structural component of all cells of the body, and functions as enzymes, in membranes, as transport carriers, and as some hormones” |
| Total fat | “Energy source and, when found in foods, is a source of n-6 and n-3 polyunsaturated fatty acids” |
| Omega-6 polyunsaturated fatty acids (linoleic acid) | “Essential component of structural membrane lipids, involved with cell signaling, and precursor of eicosanoids. Required for normal skin function” |
| Omega-3 polyunsaturated fatty acids (α-linoleic acid) | “Involved with neurological development and growth. Precursor of eicosanoids” |
| Iron | “Component of hemoglobin and numerous enzymes; prevents microcytic hypochromic anemia” |
| Zinc | “Component of multiple enzymes and proteins; involved in the regulation of gene expression” |
| Vitamin A | “Required for normal vision, gene expression, reproduction, embryonic development and immune function” |
| Vitamin B12 | “Coenzyme in nucleic acid metabolism; prevents megaloblastic anemia” |
| Vitamin D | “Maintains serum calcium and phosphorus concentrations, and, in turn, bone health” |

| Life history traits | Description |
|---------------------|-------------|
| Maximum length | Largest value ever reported for a given species, in cm (may not correspond to maximum length for subpopulation) |
| Trophic level | Weighted mean of the trophic level of the organisms that form the diet of the species; primary producers are assigned a value of one |
| Habitat | Categorical; bathydemersal, benthopelagic, demersal, pelagic, pelagic-neritic, pelagic-oceanic, reef-associated |
| Latitudinal range | Categorical; tropical, subtropical, temperate, boreal/austral, polar |
| Minimum depth | Minimum distance (m) below sea surface where species generally live |
| Maximum depth | Maximum distance (m) below sea surface where species generally live |
| a length-weight parameter | Empirically determined scalar parameter of the function $W = aL^b$ |
| b length-weight parameter | Empirically determined exponential parameter of the function $W = aL^b$ |

*Refs. 35–48. All nutrient function descriptions taken from National Academies (2017); see ref. 50 for more details. See FishBase glossary at www.fishbase.org for detailed descriptions of life history traits.*
differences are lower in the phylogenetic signal-only method (Table 4). Moreover, the phylogenetic prediction method can be readily applied to any species whose phylogenetic relationship to measured species has been estimated. Regression-based prediction requires this information in addition to information on species life history, although imputation methods may help to predict missing observations from a combination of other life history predictors and phylogenetic information19.

Because our phylogenetic method performed well in validation tests, we used it to estimate missing nutrient values for all commercially valuable species in our original 371 species, as well as for all 7500+ unmeasured species in the Actinopterygii phylogeny (see Supplementary Data Files 2–10)14. Moreover, we used this approach to identify ray-finned fish families representing promising sources of key nutrients (Table 5). Some of the families we identify contain no or relatively few measured species, but our phylogenetic method predicts their potential to contain species rich in specific nutrients.

**Discussion**

At least one-tenth of the global population is vulnerable to future deficiencies in micronutrient intake linked to the degradation of marine ecosystems7. Unsustainable fishing practices, the effects of climate change on sea temperature and dissolved oxygen content, and pollution are all major threats to a wide range of commonly consumed fish stocks. Quantifying the exact magnitude of these risks depends on knowing the nutrient content of key species—information that is extremely expensive to obtain, given the cost of directly evaluating nutrient composition data in a laboratory setting. In this study, we develop a modeling framework for using phylogenetic and life history information to predict nutrient content in fish species of the economically and nutritionally important class Actinopterygii (ray-finned fishes). We focus on nine nutrients—protein, total fat, omega-3 and omega-6 fatty acids, iron, zinc, vitamin A, vitamin D, and vitamin B12. Fish can be important sources of each of these, especially among coastal populations in the developing world.

We find that most nutrients exhibit substantial phylogenetic signal—i.e., covariance among species in nutrient values is proportional to their shared evolutionary history—and a model based solely on phylogenetic relationships and empirical estimates of phylogenetic signal provide reasonable predictions of species nutrient content. This phylogenetic signal-based model provides better predictive ability than the multiple regression
Table 2 Phylogenetic generalized least squares models predicting nutrient content

| Nutrient          | (1) Protein | (2) Total fat | (3) Omega-3 | (4) Omega-6 | (5) Iron | (6) Zinc | (7) Vitamin A | (8) Vitamin B12 | (9) Vitamin D |
|-------------------|-------------|---------------|-------------|-------------|---------|---------|---------------|----------------|--------------|
| Intercept         | 14.27***    | 19.8***       | 0.70        | 0.21        | 0.90*** | 0.71*** | 4.11***       | 2.60***        | 2.48***      |
| (1.06)            | (4.22)      | (1.80)       | (1.83)      | (4.67)      | (3.92)  | (3.47)  | (4.81)        | (3.59)         |             |
| Max length        | 0.0069*     | −0.0003      | −0.0006     | −0.0001     | 0.0002  | −0.0006 | 0.0016        | −0.0004       | 0.0018       |
| (2.52)            | (−0.36)     | (−1.04)      | (−0.76)     | (0.63)      | (−1.82) | (0.90)  | (−0.52)       | (1.79)         |             |
| Trophic level     | 1.29***     | −0.17        | −0.0262     | 0.0196      | −0.07   | 0.01    | −0.43         | −0.24          | −0.21        |
| (3.44)            | (−1.54)     | (−0.25)      | (−0.62)     | (−1.21)     | (−1.16) | (−1.39) | (−1.59)       | (−1.19)        |             |
| Max depth         | −0.0014***  | 0.0004***    | 0.0002**    | 0.0001**    | 0.00001 | 0.0001**| 0.0001**      | 0.00001        | 0.00001      |
| (−3.99)           | (3.40)      | (3.21)       | (2.87)      | (1.67)      | (−1.30) | (5.22)  | (−1.32)       | (−0.61)        |             |
| N                 | 183         | 183          | 81          | 81          | 179     | 111     | 101           | 89             | 93           |
| Pagel’s λ         | 0.000       | 0.627        | 0.672       | 0.637       | 0.165   | 0.607   | 0.725         | 0.480          | 0.784        |
| Log-likelihood    | 34.99       | 53.89        | 15.31       | 9.83        | 9.34    | 19.82   | 33.82         | 25.86          | 32.29        |
| p-Value           | <0.0001     | <0.0001      | 0.0018      | 0.0180      | 0.0219  | 0.0002  | <0.0001       | <0.0001        | <0.0001      |

T-statistics in parentheses. All outcome variables except protein are logged
**p < 0.001 *p < 0.01 p < 0.05

Table 3 Unconditional phylogenetic signal (Pagel’s λ) of each nutrient across subsets of Actinopterygii

| Nutrient              | N   | λ    | p (λ = 0, likelihood ratio test) |
|-----------------------|-----|------|-------------------------------|
| Protein               | 270 | 0.4188 <0.0001 |
| Total fat             | 267 | 0.6627 <0.0001 |
| Omega-3 fatty acids   | 89  | 0.6036 0.0014 |
| Omega-6 fatty acids   | 89  | 0.5366 0.1782 |
| Iron                  | 254 | 0.3461 0.0008 |
| Zinc                  | 146 | 0.8538 <0.0001 |
| Vitamin A             | 122 | 0.7444 1.0000 |
| Vitamin B12           | 102 | 0.5080 0.0008 |
| Vitamin D             | 103 | 0.7640 <0.0001 |

model (Table 4), although the latter accounts for both phylogenetic signal and life history trait values. This seemingly counterintuitive result can be explained by two factors. First, the sample size for estimating parameters of phylogenetic signal-based model is greater than that of phylogenetic regression model for all nutrients (see Tables 2, 3). Some of the increased predictive capacity of the phylogenetic signal model may be a consequence of better parameter estimation associated with a larger sample of species. Second, even though some life history variables (such as maximum depth) are significant predictors of some nutrients, the magnitude of effect is relatively small (Table 2). Therefore, inclusion of life history variables in nutrient predictions, as in the phylogenetic regression models, does not seem to offset the loss of sample size imposed by the additional requirement of having data on species life history. Note that we do not intend for our results to suggest that life history is unimportant in determining fish nutrient content, but rather that based on available data, incorporation of life history information does not improve predictions of species nutrient values. With the addition of more life history data and/or use of missing-data imputation techniques, phylogenetic regression models may ultimately yield better predictions than the phylogenetic signal model.

In addition, the method developed in this paper provides the basis for targeting new potential sources of key nutrients. The tendency for variation in nutrients to be phylogenetically structured means that groups of species that are closely related to nutrient-rich species are also reasonably likely to be nutrient-rich. As noted above, we explored this possibility by looking across predicted and observed nutrient values for more than 7000 ray-finned fish species, and identified families that are likely to contain species rich in key nutrients. The nutritional quality of the members of some of these families is well known; scombrids, carangids, salmonids, and clupeids make large contributions to existing fisheries. However, we also identify as potentially nutrient-rich several fish families for which species nutrient content of their species is mostly unknown (Table 5). For example, our predictions suggest that exocoetids (flying fish), sphyraenids (barracuda), and centropomids (snook) may be high in protein even though protein content has been measured for no more than one species in any of these families. In addition, even though no species of muraenid, ophichthid, congrid, or channichthyid has been assayed for total fat content, their close phylogenetic relationships to species high in fats lead us to suggest that they may be good sources of this nutrient. To be clear, we recommend that our predictions be used to inform the selection of potentially nutrient-rich species for nutrient content measurement, and we caution against use of our predictions as strong statements on which species should be exploited for nutrition.

Moreover, our identification of nutrient-rich families is limited in some important ways. First, we could only predict nutrient content in species represented in the Rabosky et al. phylogeny, and so some families may be over- or under-represented in our predictions14. For example, some nutrient-rich families may not appear in Table 5 if they have relatively few species in the phylogeny. Conversely, some families may appear to contain a large number of nutrient-rich species because they are well sampled in the phylogeny and because they contain one or more species known to be high in particular nutrients or are closely related to species of high nutritional quality. In fact, our phylogenetic prediction method is unable to make distinctions between unmeasured species of the same phylogenetic distance from measured relatives, though new nutritional data would likely help resolve this issue. Finally, we acknowledge that our predictions are contingent on Rabosky et al.’s estimation of phylogenetic relationships among these ray-finned fishes14, whose relation to the true phylogeny is unknown. Future work will refine these predictions as increasing phylogenetic information becomes available. Despite these limitations, we view our approach as a first step toward more precise estimates of the consequences of marine ecosystem transformation.

We identify several priorities in taking the conclusions of this research forward. First, improving the phylogenies of other classes important for human nutrition, including other marine seafood species as well as freshwater fish, is essential. Second, increasing the size of the nutrient validation sample—that is,
**Fig. 2** Phylogenetic relationships and nutrient content, expressed in percentiles. Eighty-four of the most commercially and nutritionally important of the 371 species in the database described above are shown. The time-scale is shown as millions of years ago (mya). The color scales indicate percentiles of nutrient content; see Supplementary Table 1 for ranges of each variable. Common names and pictorial representations of select groups of fishes are shown. All images are taken from Wikipedia, and are Public Domain images.
measuring the nutrient content of more species directly—would enable a better utilization of the large amount of life history information available. It is possible, perhaps even likely, that the weak associations between life history variables and nutrient content in this study is due to an overly restrictive validation sample. We recommend expanding the sample by prioritizing nutrient assays in species that belong to families identified as potentially nutrient-rich, as presented in Table 5. Third—approaching the small sample problem from the opposite direction—use of techniques like that of Thorson et al. (2017) to predict life history parameters would allow more existing nutrient information to be utilized.

We utilize several methods in this paper: evolutionary correlations, Brownian-motion models of trait evolution, and phylogenetic least squares regression. We believe that the availability of data is the key determinant of which of these methods is preferable in predicting fish nutrient content. In the absence of life history information, evolutionary correlations and phylogenetic signal-based models provide a rough sense of which nutrients are likely to be found together in Actinopterygii. As phylogenies expand and nutrient and life history datasets grow, life history regression models will likely improve on the predictive power of these methods. Overall, we believe that the combination of phylogenetic and life history information holds great potential for inexpensively inferring the nutrient content of a wide range of wild foods, and thereby quantifying the impacts of ecosystem transformation on human food and nutrition security.

### Methods

**Data sources.** We utilized fish nutrient information from food composition tables from Argentina, Bangladesh, Cambodia, Canada, Chile, Denmark, Japan, Finland, Gambia, Greece, Italy, Japan, Malaysia, South Korea, Peru, Turkey, and the United States; 89% of the species we examined have entries in either the Korea, Japan, Bangladesh, or Argentina tables. The life history data come from...
Table 5 Ray-finned fish families identified as potentially nutrient-rich based on phylogenetic prediction and observation of nutrient content

| Nutrient | Family name        | Types of fishes | Number of total species in family | Number of measured species in family | Number of nutrient-rich measured species in family |
|----------|--------------------|-----------------|-----------------------------------|--------------------------------------|-----------------------------------------------|
| Protein  | Serranidae         | Sea basses      | 111                               | 7                                    | 2                                             |
|          | Carangidae         | Jacks, pompanos | 80                                | 12                                   | 10                                            |
|          | Scombridae         | Tunas, mackerels| 37                                | 22                                   | 18                                            |
|          | Exocoetidae        | Flying fishes   | 32                                | 1                                    | 1                                             |
|          | Sphyraenidae       | Barracudas      | 14                                | 1                                    | 1                                             |
|          | Centropomidae      | Snooks          | 12                                | 0                                    | 0                                             |
|          | Salmonidae         | Salmon, trout   | 12                                | 12                                   | 7                                             |
|          | Istiophoridae      | Billfishes      | 8                                 | 1                                    | 1                                             |
|          | Muraenidae         | Moray eels      | 56                                | 0                                    | 0                                             |
|          | Clupeidae          | Herrings        | 49                                | 10                                   | 8                                             |
|          | Nototheniidae      | Cod icefishes   | 28                                | 1                                    | 1                                             |
|          | Engraulidae        | Anchovies       | 22                                | 6                                    | 4                                             |
|          | Ophichthidae       | Snake eels      | 20                                | 0                                    | 0                                             |
|          | Scombridae         | Tunas, mackerels| 19                                | 21                                   | 10                                            |
|          | Anguillidae        | Freshwater eels | 15                                | 1                                    | 1                                             |
|          | Carangidae         | Jacks, pompanos | 14                                | 12                                   | 6                                             |
|          | Channichthyidae    | Crocodile icefishes | 13                              | 0                                    | 0                                             |
|          | Congridae          | Garden eels     | 13                                | 0                                    | 0                                             |
|          | Salmonidae         | Salmon, trout   | 12                                | 12                                   | 5                                             |
|          | Centr._pholidae    | Medusafishes    | 10                                | 2                                    | 2                                             |
| Total fat| Cyprinidae         | Minnows, carp   | 118                               | 20                                   | 9                                             |
|          | Cobitidae          | Loaches         | 102                               | 2                                    | 2                                             |
|          | Nemachelidae       | Stone loaches   | 36                                | 0                                    | 0                                             |
|          | Balitoridae        | River loaches   | 34                                | 0                                    | 0                                             |
|          | Catostomidae       | Suckers         | 10                                | 0                                    | 0                                             |
|          | Scombridae         | Tunas, mackerels| 10                                | 19                                   | 10                                            |
|          | Cyprinidae         | Minnows, carp   | 162                               | 14                                   | 6                                             |
|          | Cobitidae          | Loaches         | 102                               | 1                                    | 1                                             |
|          | Nemachelidae       | Stone loaches   | 36                                | 0                                    | 0                                             |
|          | Balitoridae        | River loaches   | 34                                | 0                                    | 0                                             |
| Iron     | Cyprinidae         | Minnows, carp   | 162                               | 14                                   | 6                                             |
|          | Cobitidae          | Loaches         | 102                               | 1                                    | 1                                             |
|          | Nemachelidae       | Stone loaches   | 36                                | 0                                    | 0                                             |
|          | Balitoridae        | River loaches   | 34                                | 0                                    | 0                                             |
| Zinc     | Cyprinidae         | Minnows, carp   | 162                               | 14                                   | 6                                             |
|          | Cobitidae          | Loaches         | 102                               | 1                                    | 1                                             |
|          | Nemachelidae       | Stone loaches   | 36                                | 0                                    | 0                                             |
|          | Balitoridae        | River loaches   | 34                                | 0                                    | 0                                             |
| Vitamin A| Muraenidae         | Moray eels      | 56                                | 0                                    | 0                                             |
|          | Cyprinidae         | Minnows, carp   | 44                                | 6                                    | 1                                             |
|          | Cottidae           | Sculpins        | 32                                | 1                                    | 1                                             |
|          | Nototheniidae      | Cod icefishes   | 28                                | 1                                    | 1                                             |
|          | Mastacembelida     | Spiny eels      | 21                                | 1                                    | 1                                             |
|          | Ophichthidae       | Snake eels      | 20                                | 0                                    | 0                                             |
|          | Anguillidae        | Freshwater eels | 15                                | 1                                    | 1                                             |
|          | Abyssocottidae     | Deep water sculpins | 14                              | 0                                    | 0                                             |
|          | Channidae          | Snakeheads      | 14                                | 1                                    | 1                                             |
|          | Channichthyidae    | Crocodile icefishes | 13                              | 0                                    | 0                                             |
|          | Congridae          | Garden eels     | 13                                | 0                                    | 0                                             |
|          | Centr._pholidae    | Medusafishes    | 10                                | 2                                    | 2                                             |
| Vitamin B12| Cyprinidae      | Minnows, carp   | 136                               | 4                                    | 2                                             |
|          | Cobitidae          | Loaches         | 57                                | 1                                    | 1                                             |
|          | Clupeidae          | Herrings        | 54                                | 4                                    | 4                                             |
|          | Carangidae         | Jacks, pompanos | 25                                | 8                                    | 5                                             |
|          | Engraulidae        | Anchovies       | 23                                | 2                                    | 1                                             |
|          | Salmonidae         | Salmon, trout   | 21                                | 6                                    | 4                                             |
|          | Osmeridae          | Smelts          | 13                                | 2                                    | 2                                             |
| Vitamin D| Carangidae         | Jacks, pompanos | 60                                | 8                                    | 5                                             |
|          | Muraenidae         | Moray eels      | 56                                | 0                                    | 0                                             |
|          | Clupeidae          | Herrings        | 54                                | 6                                    | 6                                             |
|          | Salmonidae         | Salmon, trout   | 30                                | 8                                    | 7                                             |
|          | Nototheniidae      | Cod icefishes   | 28                                | 1                                    | 1                                             |
|          | Cyprinidae         | Minnows, carp   | 20                                | 3                                    | 1                                             |
|          | Anguillidae        | Freshwater eels | 15                                | 1                                    | 1                                             |
|          | Channichthyidae    | Crocodile icefishes | 13                              | 0                                    | 0                                             |
|          | Sphyraenidae       | Barracudas      | 11                                | 1                                    | 1                                             |

Final column lists number of species in each family identified as potentially nutrient-rich, based on phylogenetic prediction and observation of nutrient content. Species are considered nutrient-rich if their nutrient content rank within the 500 largest values out of more than 7000 species from the Rabosky et al. phylogeny with predicted or known nutrient content. Values in parentheses indicate the number of measured species that are nutrient-rich over the total number of species measured for that family.
FishBase, a publicly accessible database containing taxonomic, biological, ecological, life history, and human use information on finishes. We utilized the species-level, tree-calibrated phylogeny for Actinopterygii assembled by Rabosky et al. Although other large-scale species-level phylogenetic trees are also available for this group, we chose the Rabosky et al. phylogeny because it maximizes overlap with nutrient content data. Rabosky et al.’s reconstruction of the phylogeny was based on maximum likelihood phylogenetic analysis of 12 genes with subsequent smoothing of among-lineage substitution rate heterogeneity and divergence date estimation using fossil calibrations. The original phylogeny includes over 7500 ray-finned fish species. As described in more detail further below, we used this phylogeny in combination with life history predictor variables and estimates of phylogenetic signal (i.e., the tendency for phenotypic similarity among species to be proportional to their time of shared evolution) to predict nutrient content in species lacking such information.

Evolutionary correlations. We first explored bivariate evolutionary correlations between all pairwise combinations of life history and nutrient variables. Here, associations between evolutionary changes in pairs of variables were estimated using phylogenetic least squares (PGLS) as implemented in the R package ‘geiger’. Because species with a recent common ancestor are expected to have more similar trait values, the assumptions of ordinary least squares (OLS) are violated. PGLS accounts for phylogenetic non-independence using shared ancestry as inverse weights on the elements of the residual variance-covariance matrix used in the OLS approximation, would indicate trait evolution completely independent of phylogeny, λ = 1 indicates that shared evolutionary history (i.e., shared phylogenetic branch lengths) predict phenotypic similarity among species, and intermediate 0 < λ < 1 discounts the phylogenetic dependence of trait values among species. The best-fit inverse matrix V−1 is then used to estimate the predictor coefficient β of the phylogenetic prediction of nutrient content. For each nutrient, we removed one measured species from the dataset and applied the method to predict the species’ nutrient value and calculate its prediction interval. We then determined whether the prediction interval contained the measured value. If it does not, we label that trial as an error, and then calculated the error rate for each nutrient over all measured species. We also calculated the median percent deviation between predicted and measured values as a proportion of the standard deviation of the sample of all species nutrient values. We then compared error rates and accuracy of phylogenetic-only predictions to predictions based on the best-fit multiple regression models. These latter predictions were obtained following the method of Garland and Ives and were restricted to the sample of species for which life history data were available.

Validity. We used a jackknifing approach to validate the method for phylogenetic prediction of nutrient content. For each nutrient, we removed one measured species from the dataset and applied the method to predict the species’ nutrient value and calculate its prediction interval. We then determined whether the prediction interval contained the measured value. If it does not, we label that trial as an error, and then calculated the error rate for each nutrient over all measured species. We also calculated the median percent deviation between predicted and measured values as a proportion of the standard deviation of the sample of all species nutrient values. We then compared error rates and accuracy of phylogenetic-only predictions to predictions based on the best-fit multiple regression models. These latter predictions were obtained following the method of Garland and Ives and were restricted to the sample of species for which life history data were available.

Code availability. All code used in the analysis is made available as Supplementary Data 11–15. SD1 is the script for evolutionary correlations. SD12 is the script for phylogenetic least squares. SD13 is the script for estimating the phylogenetic signal of nutrient variables. SD14 is the script for validating the predictions under the lambda model. SD15 is the script for validating predictions under the lambda plus phylogenetic regression model.

Data availability. We declare all data used in the above analysis, tables, and figures to be available within the paper and in the supplementary information files. Supplementary Data 1 is the fish life history and nutrient content database. Phylogenetically predicted nutrient values are labeled as Supplementary Data 2–10.

Received: 22 November 2017 Accepted: 24 August 2018
Published online: 25 September 2018

References

1. Pauly, D. & Zeller, D. Catch reconstructions reveal that global marine fisheries catches are higher than reported and declining. Nat. Commun. 7, 10244 (2016).
2. Cheung, W. W. L. et al. Large-scale redistribution of maximum fisheries catch potential in the global ocean under climate change. Glob. Change Biol. 16, 24–35 (2009).
3. Cheung, W. W. L., Watson, R. & Pauly, D. Signature of ocean warming in global fisheries catch. Nature 497, 365–368 (2013).
4. Perry, A. L., Low, P. J., Ellis, J. R. & Reynolds, J. D. Climate change and distribution shifts in marine fishes. Science 308, 1912–1915 (2005).
5. Pinsky, M. L., Worm, B., Fogarty, M. J., Sarniento, J. L. & Levin, S. A. Marine taxa track local climate velocities. Science 341, 1239–1242 (2013).
6. Donner, S. D., Rickel, G. J. & Heron, S. F. A new, high-resolution global mass coral bleaching database. PLoS ONE 12, e0175490 (2017).
7. Hoegh-Guldberg, O. et al. Coral reefs under rapid climate change and ocean acidification. Science 318, 1737–1742 (2007).
8. Halpern, B. S. et al. The global map of human impact on marine ecosystems. Science 319, 948–952 (2008).
9. Golden, C. D. et al. Fall in fish catch threatens human health. Nature 534, 317–320 (2016).
10. Thilsted, S. H. et al. Sustaining healthy diets: The role of capture fisheries and aquaculture for improving nutrition in the post-2015 era. Food Policy 61, 126–131 (2016).
11. Benn, C. et al. Contribution of fisheries and aquaculture to food security and poverty reduction: assessing the current evidence. World Dev. 79, 177–196 (2016).
12. Casner-Montemayor, A. M., Pauly, D., Weatherdon, L. V. & Ota, Y. A global estimate of seafood consumption by Coastal Indigenous Peoples. PLoS ONE 11, e0166681 (2016).
13. Mozaffarian, D. & Rimm, E. B. Fish intake, contaminants, and human health: evaluating the risks and benefits. JAMA 296, 1885–1899 (2006).
14. Lund, E. K. Health benefits of seafood; is it just the fatty acids? Food Chem. 140, 413–420 (2013).
15. Pauly D. & Zeller D. (eds). Sea Around Us Concepts, Design And Data (2018) www.seaaroundus.org. Accessed 10 April 2018.

16. Revell, L. J. Phylogenetic signal and linear regression on species data. Methods Ecol. Evol. 1, 319–329 (2010).

17. Pagel, M. Inferring evolutionary processes from phylogenies. Zool. Scr. 26, 331–348 (1997).

18. Pagel, M. Inferring the historical patterns of biological evolution. Nature 401, 877–880 (1999).

19. Thorson, J. T., Munch, S. B., Cope, J. M. & Gao, J. Predicting life history parameters for all fishes worldwide. Ecol. Appl. 27, 2262–2276 (2017).

20. Universidad Nacional de Luján. Tabla de composición de alimentos (2010) http://www.unlu.edu.ar/~argofood/Tablas/Tabla.htm. Accessed 25 July 2017.

21. Shaheen, N. et al. Food Composition Tables for Bangladesh. (University of Dhaka, Dhaka, 2013).

22. Seyha, S. & Kuong, K. Food Composition Table for Cambodia (Ministry of Agriculture, Forestry and Fisheries of Cambodia, Cambodia, 2013).

23. Centre for Indigenous Peoples’ Nutrition and Environment (McGill University). Traditional Food Composition Nutribase (2005) https://www.mcgill.ca/cine/files/cine/Traditional_Food_Composition_Nutribase.pdf. Accessed 25 July 2017.

24. Rittensober, D., Stadlmayr, B., Nowak, V., Du, J. & Charrondiere, U. Report for the development of the FAO/INFOODS user database for fish and shellfish (uFish)—challenges and possible solutions. Food Chem. 193, 112–120 (2016).

25. Pryme, C. J. & Paul, A. A. Food Composition Table for Use in The Gambia (Institute of Human Nutrition Research, Cambridge, UK, 2011).

26. Office for Resources, Policy Division Science and Technology Policy Bureau, Japan. National Tables of Food Composition in Japan, Seventh Revised edn (2015) http://www.mext.go.jp/en/policy/science_technology/policy/title01/detail01/1374030.htm. Accessed 25 July 2017.

27. Kim, J. C. et al. 8th Revision Standard Food Composition Table. (Department of Agriculture, N.A.A, R.D.A, Korea, 2013).

28. Korkalo, L., Hauta-Alus, H. & Mutanen, M. Food Composition Tables for Mozambique, Version 2 (Department of Food and Environmental Sciences, University of Helsinki, Finland, 2011).

29. Food and Agricultural Organization (FAO) of the United Nations. FishStat: Software for Fishery Statistical Time Series. In: FAO Fisheries and Aquaculture Department [online] (2017) http://www.fao.org/fishery/statistics/software/fishstat/en. Accessed 13 July 2017.

30. Rabosky, D. L. et al. Rates of speciation and morphological evolution are correlated of interspecies diversity worldwide. Philos. Trans. R. Soc. Lond. B Biol. Sci. 366, 119–157 (1989).

31. Grafen, A. The phylogenetic regression. Philos. Trans. R. Soc. Lond. B Biol. Sci. 326, 119–157 (1989).

32. Martins, E. P. & Hansen, T. F. Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data. Am. Nat. 149, 646–667 (1997).

33. Freckleton, R. P., Harvey, P. H. & Pagel, M. Phylogenetic analysis and comparative data: a test and review of evidence. Am. Nat. 160, 712–726 (2002).

34. Svakumaran, S., Martell, S. and Huffman, L. The Concise New Zealand Food Composition Tables, 9th edn. (The New Zealand Institute for Plant & Food Research Limited and Ministry of Health, New Zealand, 2012).

35. US Department of Agriculture, Agricultural Research Service, Nutrient Data Laboratory. USDA National Nutrient Database for Standard Reference, Release 28 (Slightly revised) (2016) http://www.ars.usda.gov/ba/bhnrc/nrdl. Accessed 25 July 2017.

36. Stadlmayr, B. et al. Table de composition des aliments d’Afrique de l’Ouest, Food and Agriculture Organization of the United Nations (2012) http://www. fao.org/docrep/015/i2698b/i2698b00.htm. Accessed 25 July 2017.

37. Dignan, C., Burlingame, B., Kumar, S., and Aalbersberg, W. The Pacific Islands Food Composition Tables, Second edition. Food and Agriculture Organization of the United Nations (2004) http://www.fao.org/docrep/007/y5432e/y5432e00.htm. Accessed 25 July 2017.

38. Groves, R. & Pauly, D., eds. FishBase (2016) www.fishbase.org. Accessed 13 July 2017.

39. Near, T. J. et al. Resolution of ray-finned fish phylogeny and timing of diversification. Proc. Natl Acad. Sci. 109, 13698–13703 (2013).

40. Betancur-R. R. et al. The tree of life and a new classification of bony fishes. Philos. Trans. R. Soc. B Biol. Sci. 368, 20130011 (2013).

41. Revell, L. J. Testing quantitative genetic hypotheses about the evolutionary rate matrix for continuous characters. Evol. Ecol. Res. 10, 311–321 (2008).

42. Revell, L. J. & Harmon, L. J. Testing quantitative genetic hypotheses about the evolutionary rate matrix for continuous characters. Evol. Ecol. Res. 10, 311–321 (2008).

43. Revell, L. J. & Collar, D. C. Phylogenetic analysis of the evolutionary correlation using likelihood. Evolution 63, 1990–1998 (2009).

44. Pennell, M. W. et al. Geiger v2.0: an expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. Bioinformatics 30, 2216–2218 (2014).

45. R Core Team. R: A language and environment for statistical computing. 2018 (R Foundation for Statistical Computing, Vienna, Austria, 2017) http://www. r-project.org/