Global plant diversity as a reservoir of micronutrients for humanity

Article

Accepted Version

Cantwell Jones, A., Ball, J., Collar, D., Diazgranados, M., Douglas, R., Forest, F., Hawkins, J. ORCID: https://orcid.org/0000-0002-9048-8016, Howes, M.-J., Ulian, T., Vaitla, B. and Pironon, S. (2022) Global plant diversity as a reservoir of micronutrients for humanity. Nature Plants, 8. pp. 225-232. ISSN 2055-0278 doi: https://doi.org/10.1038/s41477-022-01100-6 Available at https://centaur.reading.ac.uk/102463/

To link to this article DOI: http://dx.doi.org/10.1038/s41477-022-01100-6

Publisher: Springer

Global plant diversity as a reservoir of micronutrients for humanity

Aoife Cantwell-Jones*[1], Jenny Ball[2], David Collar[3], Mauricio Diazgranados[2], Ruben Douglas[2], Félix Forest[2], Julie Hawkins[4], Melanie-Jayne R. Howes[2,5], Tiziana Ulian[2], Bapu Vaitla[6], and Samuel Pironon†[2].

1. Science and Solutions for a Changing Planet DTP, and the Department of Life Sciences, Imperial College London, Silwood Park Campus, Ascot, SL5 7PY, UK

2. Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3DS, UK

3. Department of Organismal & Environmental Biology, Christopher Newport University, One Avenue of the Arts, Newport News, VA, 23606, USA

4. School of Biological Sciences, Health and Life Sciences Building, University of Reading, Whiteknights, Reading, Berkshire, RG6 6BX, UK

5. Institute of Pharmaceutical Science, Faculty of Life Sciences & Medicine, King's College London, SE1 9NH, UK

6. Department of Nutrition, Harvard TH Chan School of Public Health, Harvard University, 665 Huntington Ave, Boston, MA, 02115, USA

Corresponding authors: aoife.cantwell-jones15@imperial.ac.uk*; s.pironon@kew.org†

SUMMARY

With more than two billion people suffering from malnutrition and diets homogenising globally, it is vital to identify and conserve nutrient-rich species that may contribute to improving food security and diversifying diets. Of the approximately 390,000 vascular plant species known to science, thousands have been reported to be edible, yet their

nutritional content remains poorly characterised. Here we use phylogenetic information to identify plants with the greatest potential to support strategies alleviating B-vitamin deficiencies. We predict the B-vitamin profiles of >6,400 edible plants lacking nutritional data and identify 1,044 species as promising key sources of B vitamins. Several of these source species should become conservation priorities, as 63 (6%are threatened in the wild and 272 (26%) are absent from seedbanks. Moreover, many of these conservation-priority source species overlap with hotspots of malnutrition, highlighting the need for safeguarding strategies to ensure that edible plant diversity remains a reservoir of nutrition for future generations, particularly in countries needing it most. Although by no means a silver bullet to tackling malnutrition, conserving a diverse portfolio of edible plants, unravelling their nutritional potentials, and promoting their sustainable use are essential strategies to enhance global nutritional resilience.

MAIN

Around two billion people are currently afflicted by micronutrient (vitamin and mineral) deficiencies[1,2]. This "hidden hunger" causes increased susceptibility to infectious disease, reduced cognitive function, blindness and leads to an estimated one million premature deaths yearly[1–3]. Hidden hunger cannot be eradicated by simply increasing calorie intake; diversified, plant-rich diets are needed to provide sufficient nutrients and calories[4–6]. Indeed, driven by a dependence on a few staple crops (often due to the inaccessible, high cost of non-staple foods), micronutrient deficiencies remain highly prevalent in some countries where the daily energy availability exceeds 2,500 kcal[1].

Global plant diversity as a reservoir of micronutrients for humanity

For many people, reducing micronutrient deficiencies could be achieved by expanding diets and tapping into the great diversity of edible plants[6,7]. Increasing consumption of specialty and underutilised crops grown traditionally in local communities, crop landraces little used outside of breeding programmes, and sustainably harvested wild plants could help alleviate global micronutrient deficiencies[1,7,8]. Despite the potential nutritional benefits[9], consuming a diverse range of edible plants is uncommon: humans currently rely on only three crops (rice, maize and wheat) to provide 60% of their plant-derived calories[8,10]. This is partly due to high demand, low prices and large-scale cultivation of a few staple crops, and absent value chains for underutilised species, which together have caused many edible plants to become forgotten and, in some cases, threatened[10–12]. Losing this edible plant diversity would limit options for identifying future food sources and maintaining genetic diversity – both of which will be essential for food security in the face of challenges, such as climate change[13–15]. Interventions promoting the consumption and conservation of a diverse range of species are therefore needed[6,16], but there is limited knowledge of the nutritional profiles of the unknown, but large, number of edible plants[10] (>7,000[17]) or their current threat status (but see[8,18]). This knowledge gap seriously impedes our ability to appreciate how edible plant diversity could contribute to dietary diversity or support existing strategies to tackle specific micronutrient deficiencies. We thus assess the potential of terrestrial angiosperm plant diversity as a reservoir of micronutrients for humanity by 1) predicting the micronutrient profiles of >6,400 edible species and identifying those that could be key micronutrient sources, 2) determining how threatened in situ and protected ex situ micronutrient-source species currently are, and 3) assessing the geographic overlap of threatened-in-situ and not-conserved-ex-situ micronutrient-source species with global malnutrition hotspots.

We focus on six B vitamins – thiamine (B_1), riboflavin (B_2), niacin (nicotinic acid & nicotinamide, B_3), pantothenic acid (B_5), pyridoxine (B_6) and folate (B_9) – that play essential roles as cofactors in metabolism and nerve function, yet are commonly deficient in both developed and developing countries[1]. For example, >40% of the populations in South and South-East Asia are estimated to be deficient in folate[1]. First, using B-vitamin data for 280 nutritionally known (i.e. with data available for at least one B vitamin) angiosperm species found around the world, we tested whether nutrition is predictable using phylogenetic relationships[19], assuming closely related species share similar biochemistries[20]. The selected edible species cover various foods, including fruits, vegetables, nuts, grains, herbs, spices, and pulses.

Results

High and low concentrations of each B vitamin (except pyridoxine) were significantly clustered across the phylogenetic tree (Fig. 1a) when all nutritionally known species were considered. However, the strength of this clustering (phylogenetic signal; measured as Pagel's λ; Supplementary Table 1) varied among the B vitamins, with niacin and folate showing stronger signals ($n \geq 256$; $\lambda \geq 0.302$; $p \leq 0.001$) than thiamine, pantothenic acid and riboflavin ($n \geq 232$; $\lambda \leq 0.293$; $p \leq 0.005$). The lack of signal for pyridoxine ($n = 261$; $\lambda = 0.065$; $p = 1$) may be due to the distinctiveness of different plant tissues (e.g. seeds versus leaves) in terms of nutrition[21], as we found significant signal for pyridoxine when considering only species that have leaves consumed ($n = 54$; $\lambda = 1.01$; $p < 0.001$; Supplementary Table 2). For example, high values of all B vitamins were clustered in Fabales and Poales, and low values, in Rosales and Cucurbitales (Supplementary Table 3). This result is consistent with findings of significant clustering for the top-800 most consumed plants[22], domesticated crops[23] and plants with medicinal use[24] across different taxonomic scales on the angiosperm tree of life.

For the five B vitamins demonstrating significant phylogenetic signal (thiamine, riboflavin, niacin, pantothenic acid, and folate), we predicted their concentrations in the nutritionally known species (Supplementary File 1) and the >6,400 nutritionally unknown edible plant species (from 242 angiosperm families) that have been documented to be consumed by humans around the world[17] (Supplementary File 2). For the nutritionally known species, this was done by removing one species at a time from the phylogenetic tree and predicting its B-vitamin concentrations using the strength of the phylogenetic signal for each B vitamin and the B-vitamin concentrations of close relatives[19]. For these five B vitamins, ≥91.4% of nutritionally known species had measured (observed) values within the 95%-confidence intervals of their predicted values (Supplementary Table 4; Extended Data Fig. 1). Further, when comparing the predicted and observed values of nutritionally known species, we found significant relationships for all B vitamins (gls; all $t \geq 6.05$, $p < 0.001$; Supplementary Table 5), and median differences between predicted and observed values for each nutrient were <33% of the standard deviation across species. We then used the results for the nutritionally known species to predict the B-vitamin profiles of the nutritionally unknown species (Fig. 1b; Methods; Extended Data Fig. 2).

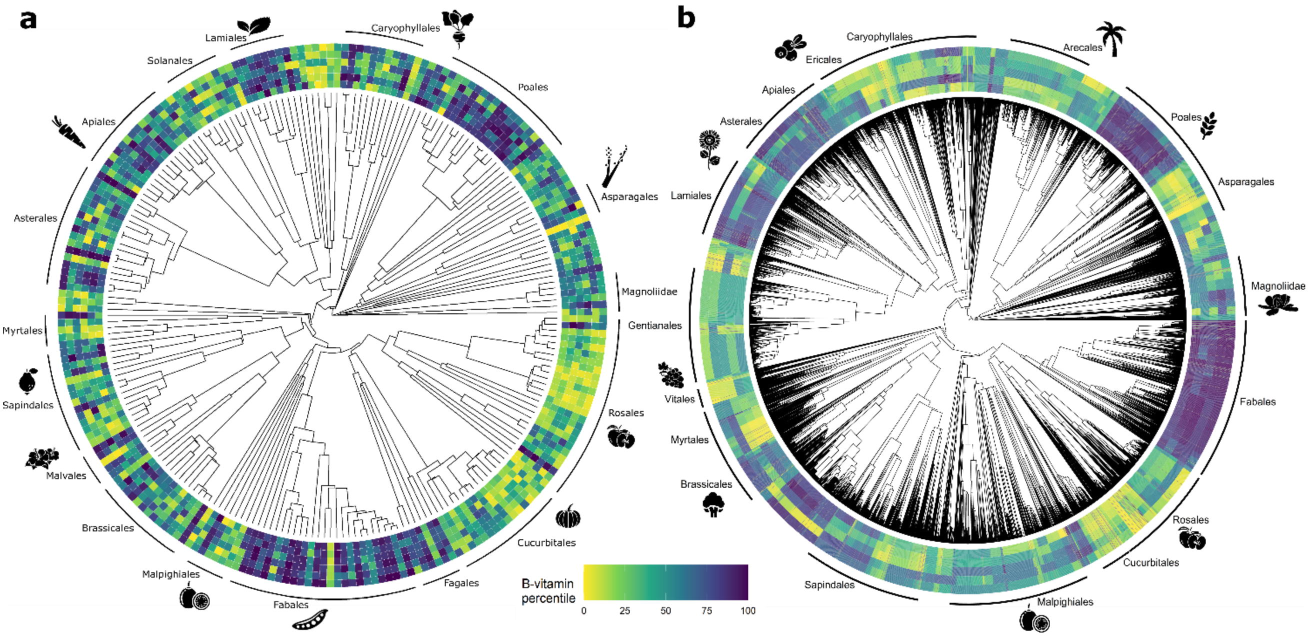

Figure 1: **Phylogenetic trees of a) nutritionally known (*n* = 229) and b) nutritionally known and unknown terrestrial angiosperm edible plant species (*n* = 6,740) with their corresponding observed and predicted B-vitamin profiles, respectively.** From the inside of the ring outwards, the B vitamins are thiamine, riboflavin, niacin, pantothenic acid, (pyridoxine in 1a) and folate. Pyridoxine was omitted from 1b due its lack of phylogenetic signal. B-vitamin values are expressed as their percentile rank, from low (yellow) to high (blue) values**.** Well-represented orders and subclasses are highlighted around the tree.

To identify edible plants with the greatest potential for tackling B-vitamin deficiencies, we focused on those predicted to be major sources (i.e. contribute $\geq$15% of a given B vitamin to the recommended dietary allowance in 100 g of fresh edible plant material[25]; Supplementary Table 6). Additionally, to ensure our list of source species is conservative, we used a second method to identify clades where “source” species are overrepresented (Methods). We therefore focused on edible species predicted to be sources by both approaches (a “conservative scenario”; Supplementary Table 7), but also provide the number of source species predicted by at least one approach as a “best-case scenario”. We found 633 edible species (best-case scenario: 1,745) as predicted sources of thiamine; 25 (608) of riboflavin; 243 (936) of niacin; 0 (686) of pantothenic acid; and 715 (1,314) of folate (Supplementary File 3). This totalled 1,044 (2,475) edible species as sources of at least one B vitamin. In our conservative scenario, no edible plants were sources of pantothenic acid, possibly because pantothenic acid tends to be present in most plants[26]. The low counts of source species for pantothenic acid and riboflavin also reflect that these numbers are likely underestimates, as we used a conservative approach (i.e. all B vitamins have source species in the best-case scenario) and the total number of edible plants is unknown[8]. However, that 1,044 species were predicted to be sources of at least one B vitamin supports other studies suggesting edible plant diversity can contribute significantly to human nutrition[9,27]. These 1,044 source species should be prioritised to have their nutritional profiles chemically validated. Critically, we do not suggest these B-vitamin-source species should replace or are more nutritious than commonly consumed crop staples, yet they could have a range of applications; from informing crop-breeding programmes (e.g. for biofortification[28–30]), to sustainable wild harvest or cultivation by local communities and practitioners[7] or food, drink and nutraceutical companies, to being focal points for policy and NGO interventions[8,10,31].

To gauge opportunities for ensuring long-term preservation and availability of B-vitamin-source species for humanity, we next assessed the threat status of each species in their natural environment[32] (in situ) and whether they are conserved ex situ in seed banks[33,34]. We found 46.1% (*n* = 3,124 species) of edible species had unknown threat status (not assessed or data deficient; Fig. 2), and 14.8% (540) of those assessed were threatened in situ (i.e. possibly threatened, threatened or extinct in the wild), according to the Botanic Gardens Conservation International ThreatSearch database[34]. Overall, 41.8% (2,818) of edible species were not conserved ex situ, with this number rising to 60.9% (329) for species assessed to be threatened in situ. Almost half (49.3%; 1,540) of species without a threat assessment were also missing from ex-situ conservation records. When focusing on source species, figures were less pessimistic but still stark: 34.0% (358) of source species had unknown threat assessment; 9.1% (63) of those assessed were considered threatened (6.0% of all source species); and 26.1% (272) were not conserved ex situ. The percentage of source species not conserved ex situ rose to 44.4% (28) for those threatened in situ; and to 36.0% (129) for those with unknown threat assessment. These source species should be prioritised for future threat-status assessment and in- and ex-situ conservation programmes[35]. These percentages corroborate other studies finding a significant portion of edible plant diversity still at risk of extinction[18,36,37], and should be considered conservative, given a considerable number of edible plant species likely remain undocumented[8] and potentially threatened[4].

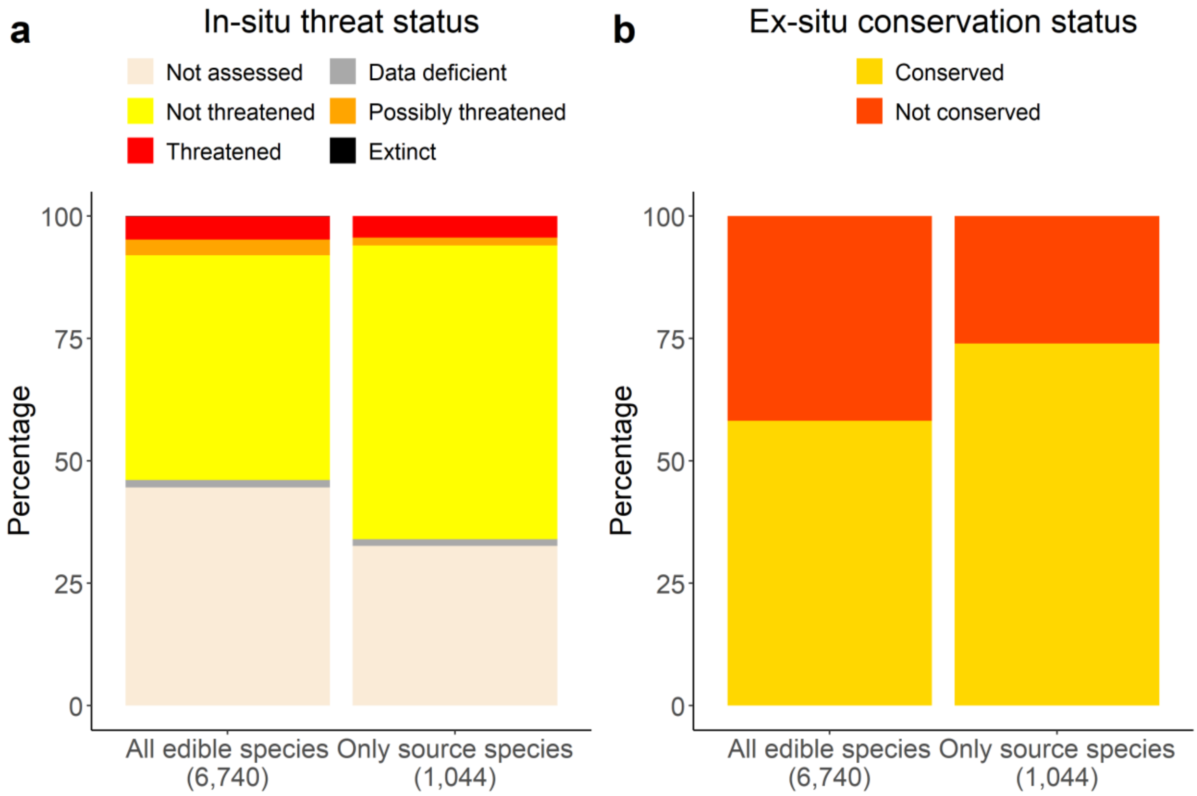

Figure 2: **In-situ threat and ex-situ conservation status of all documented terrestrial edible angiosperm species and predicted B-vitamin source species**. a) “Extinct in the wild”, “Threatened”, “Possibly threatened”, “Not threatened” and “Data deficient” refer to the level of threat each species faces in situ according to the Botanic Gardens Conservation International ThreatSearch database[32]. Some species were absent from this database (“Not assessed”). b) Ex-situ conservation status refers to whether the edible species are recorded as being present in seedbanks – found by searching Genesys[34] and the Millennium Seed Bank Partnership data warehouse[33].

Finally, to explore which countries could benefit most from conserving their edible plant diversity, we mapped the distributions of B-vitamin-source species and investigated regions with high percentages of species threatened in situ or not conserved ex situ (Extended Data Figures 3-5). We additionally examined whether these regions had high prevalence of inadequate intakes of each B vitamin (using the estimated Prevalence of Inadequate Micronutrient Intake Index from Beal et al.[1]). These countries have the most potential to benefit nutritionally from their edible plant diversity (Fig. 3; see Supplementary File 4 for the countries in which each source species is found) but need to take the greatest steps in safeguarding these species for future generations. These analyses may especially apply to low-and middle-income countries, as they are more likely to rely on their native plant diversity during times of hardship[38].

For the 241 countries and island states with thiamine-source species, the mean number of thiamine-source species in each country was 67 (standard deviation (±) = 51), of which 3.2% of species on average were threatened (±3.9%) and 6.1% not conserved ex situ (±5.9%). Higher percentages for both dimensions overlapped in China, South-East Asian countries and Pacific islands (e.g. Sumatera, Jawa, Thailand, Myanmar, & Vanuatu; ≥3% threatened and ≥15% not conserved ex situ; Fig. 3a). For riboflavin, 191 countries had a mean number of 4 (±3) source species, of which on average 4.8% (±11.8%) were threatened and 29.8% (±32.6%) not conserved ex situ. Several Asian and African countries and islands had high percentages of riboflavin-source species not conserved ex situ (e.g. Sumatera, Cambodia, Liberia, Vietnam, Gambia, Sierra Leone & Madagascar; ≥50%), whereas other countries, such as Peru and New Zealand had ≥33% of species threatened in situ. High on both dimensions were North American and Caribbean countries (e.g. El Salvador, Guatemala, Dominican Republic, Puerto Rico, Mexico & United States), with ≥20% of riboflavin-source species threatened and not conserved ex situ (Fig. 3b). Niacin-source species were present in 232 countries and islands (mean number = 22 ±19), where a mean 2.9% (±5.4%) of niacin-source species were threatened and 6.1% (±9.6%) not conserved ex situ. Several countries from sub-Saharan African, South-East Asia and Pacific islands (e.g. Vanuatu, Sumatera, Borneo, Sulawesi, Equatorial Guinea, Jawa, Thailand, & Gabon) had high percentages of niacin-source species not conserved ex situ (≥18%; Fig. 3c). In contrast, some higher latitude countries (e.g. Tadzhikistan, Kuwait, Canada & Lesotho) had ≥10% of species threatened in situ. Finally, folate-source species were present in 231 countries and islands (mean number =

61 ±57), with a mean 4.5% (±7.1%) threatened and 6.9% (±8.6%) not conserved ex situ. Some sub-Saharan countries (e.g. Liberia, Gabon, Republic of Congo, Cameroon, Ivory Coast, & Nigeria) had high percentages (≥19%) of folate-source species not conserved ex situ. In contrast, high percentages of threatened species were concentrated in several European countries (Ireland, Iceland, Norway, Sweden, Denmark, Belarus, Poland & Netherlands; >15%; Fig. 3d). Although these country-level percentages of source species threatened or not conserved ex situ were on average low, conservation efforts for B-vitamin source species could clearly be strengthened in many parts of the world, with South Asian countries being conservation-priority hotspots for thiamine- and niacin-source species, North America for riboflavin, and sub-Saharan Africa for niacin and folate.

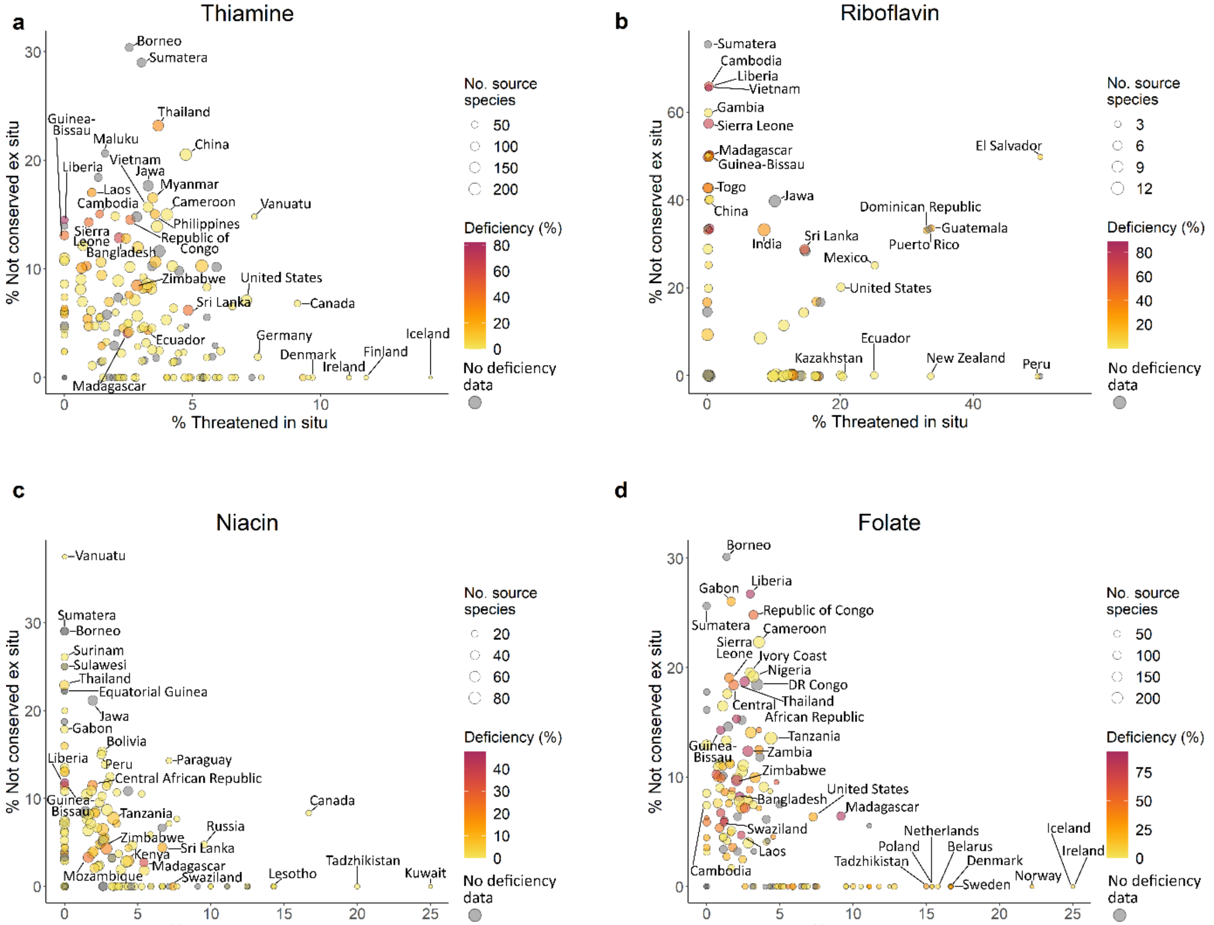

Figure 3: **Percentages of B-vitamin source species that are globally threatened in situ and not conserved ex situ in 244 countries and island states.** Each data point represents a

country or island, with the size indicating the number of B-vitamin-source species found in that country, and colour highlighting the prevalence of inadequate intake (deficiency) in the respective B vitamin[1]. Countries and islands with an area $\leq$6,000 km^2 are not shown in these plots. “Threatened species” refers to species that are possibly threatened, threatened or extinct in the wild globally, according to the Botanic Gardens Conservation International ThreatSearch database[32]. Species “not conserved ex situ” refers to those absent from Genesys[34] and the Millennium Seed Bank Partnership databases[33]. The data in these scatterplots (including small countries and islands) are available in map and table format in Extended Data Figures 3-5 and Supplementary File 4.

Several of these conservation-priority hotspots additionally had high prevalence of B-vitamin deficiencies (Supplementary File 4). Thailand is of conservation priority for thiamine-source species, and 33.0% of the population suffered from thiamine deficiency. Guatemala (riboflavin conservation priority) had prevalence of riboflavin deficiency $\geq$26.6%, whereas the Republic of Congo and Liberia (folate conservation priorities) both had $\geq$54.2% people suffering from folate deficiency. Overall, this overlap of conservation-priority and malnutrition hotspots underlines the need for improved safeguarding strategies in these countries, to ensure edible plant diversity remains a reservoir of nutrition for future generations.

Despite this need to improve safeguarding strategies, the high numbers of thiamine- and folate-source species across many countries highlight some potential for increased, sustainable consumption of edible plant diversity in targeted health interventions. Identifying which edible species would be most appropriate for such interventions would however require further work, such as choosing species with optimal growth rates, pest resistance and climate resilience[4,14], or matching cultural taste preferences[39,40]. Moreover, it would be imperative to undertake this research using a participatory approach, involving scientists, practitioners, local communities, policy makers and other relevant stakeholders[39,41]. This will be important for gauging the potential productivity and demand of specific species, both of which would determine pricing – with low prices needed to reach people most afflicted by B-vitamin deficiencies[2]. Such work is currently conducted by organisations such as CGIAR (formerly Consultative Group on International Agricultural Research) and Bioversity International, which, for example, have improved value chains in Guatemala for the nutritious edible plants chaya (*Cnidoscolus aconitifolius*) and tepary bean (*Phaseolus acutifolius*)[42]. Our approach represents a first step to complement the activities of such organisations, and we appreciate that a large amount of work would still be required at multiple scales for edible plant diversity to be used to improve B-vitamin intakes.

Although this study provides a first insight into the potential of edible plant diversity to alleviate B-vitamin deficiencies, it comes with a few limitations. Firstly, we could not predict the pyridoxine concentration in nutritionally unknown species, due to its lack of phylogenetic signal. Secondly, the limitations of the prediction method (see Vaitla, Collar, et al.[19]) imply the predictions are likely to be conservative, i.e. when comparing predicted and observed B-vitamin concentrations for nutritionally known species (Extended Data Fig. 1), there was a tendency to underestimate their concentrations, and our best-case scenario predicted a greater number of source species. Further, the large nutritional discrepancies arising from consuming different plant tissues (e.g. seeds versus leaves) cannot be accurately estimated. In the future, knowing these plant tissues will also be important for any health intervention using edible plant diversity, as it will determine the feasibility of including source species into existing diets[7]. Thirdly, nutritional profiles of plants can vary spatiotemporally, depending on, for example, soil and climate conditions, thus our predictions may not represent the full variation of nutrient concentrations observed for a given species[43]. Finally, the nutritional profile of the raw edible plant may not equate to oral bioavailability in humans. Edible plants often contain “anti-nutrients”, such as protease-inhibitors, oxalates and phytates, which may reduce absorption of some micronutrients[29], and the B-vitamin profile of plants may change during processing[44]. In addition, it is necessary to understand how to safely prepare species for consumption, e.g. 23.9% of source species have additional uses as vertebrate or invertebrate poisons, though have records of safe use[17]. These limitations underline that our predictions need to be confirmed with chemical analyses performed across edible species from multiple geographic regions, combined with better understanding of the bioaccessibility of their B vitamins.

Conclusions

Overall, we find that closely related edible plant species share similar B-vitamin profiles (for thiamine, riboflavin, niacin, pantothenic acid and folate), enabling the prediction of B-vitamin profiles for >6,400 nutritionally unknown edible species, which resulted in 1,044 species being identified as potential B-vitamin sources. Despite the potential of these 1,044 species to contribute to future nutritional resilience, many are threatened in their natural environments and not yet preserved in seedbanks, including in some of the countries most afflicted by B-vitamin deficiencies. Although tackling micronutrient deficiencies requires a multi-faceted approach across multiple scales of governance, our study offers the means to concentrate efforts on species most likely to act as future reservoirs of B vitamins. Edible plant diversity has the potential to complement existing strategies to improve nutrition, provided we make it accessible to those who need it most, use it sustainably and guarantee its long-term conservation.

METHODS

Nutritional data for the nutritionally known edible plant species

B-vitamin data (thiamine (B_1), riboflavin (B_2), niacin (B_3), pantothenic acid (B_5), pyridoxine (B_6) and folate (B_9)) on 321 edible species and varieties were collected from eight online national and regional food-composition databases: United Kingdom[45]; United States[46]; New

Zealand[47]; Japan[48]; West Africa[49]; Malawi[50]; India[51]; and the Pacific Islands[52] (see Supplementary Table 8 for the list of species included). These food-composition databases were selected to get a global cover of nutritionally known species, using only values for which B-vitamin data on 100 g of the unprocessed, fresh edible portion of the plant were available. We ensured all B-vitamin concentrations were in consistent units of measurement across the databases and have presented data based on accepted analytical methods for assessing B-vitamin composition (see Supplementary Table 9 for a summary of the analytical methods used by the different food-composition databases). In 27 cases where multiple entries for a given species existed (due to multiple tissues being consumed: e.g. garden cress leaves or seeds; or multiple varieties existing: e.g. broccoli, cabbage, cauliflower and Brussels sprouts for *Brassica oleracea* L.), we selected either the most common variety or tissue consumed, or selected one entry randomly, leading to 41 entries being removed (Supplementary Table 8). Additionally, some entries for nutritionally known "species" comprised multiple species within a genus (*n* = 6; e.g. *Dioscorea* "yam" comprises several species), but only one mean B-vitamin profile was available. In these cases, one species was chosen to represent the genus, by choosing the most commonly eaten species or, when the species could not be differentiated using this criterion, a representative species was chosen from the genus randomly (Supplementary Table 8). After these removal steps, we were left with 280 nutritionally known species (Supplementary Table 8).

Species were grouped based on the part consumed, following Ray, Ray & Sreevidya[53]: seeds and grains (*n* = 77; including legumes); leaves and leafy shoots (56); flowers (4); fleshy fruits (110; including berries); underground organs (30; including true roots and underground storage organs, such as bulbs, tubers and rhizomes); and other (3; e.g. bulbils and petioles). These groups were pooled for all further analyses (Supplementary Methods 1, Supplementary Table 2). All further work was performed using R (v. 3.6.3 and v. 4.0.2[54]).

Edible plant species

Data on edible plants came from the *World Checklist of Useful Plant Species* database[17], a compilation of 13 datasets representing 40,292 species categorised by their uses. Here, only "human food" (viz. recorded use of this plant being consumed; *n* = 7,039) terrestrial angiosperms were used, although we acknowledge many non-terrestrial and/or non-angiosperm plant species also have the potential to be nutritious[55]. Note that the database does not specify the part of the plant consumed. This list was divided into "nutritionally known" (i.e. were present in the food-composition databases) and "nutritionally unknown" species. As each B vitamin varied in data availability, the number of nutritionally known species varied between 232 (for pantothenic acid) and 280 (for thiamine), and nutritionally unknown species, between 6,460-6,508.

Phylogenetic inference

To construct the phylogenetic tree, a dated, species-level, backbone phylogenetic tree for Spermatophyta[56], as modified by Jin & Qian[57] ("GBOTB.extended" tree) was pruned to contain only the desired species. The original phylogenetic tree by Smith and Brown[56], which contains >79,000 terminal taxa, was constructed using maximum likelihood by mining the available molecular data for Spermatophyta from GenBank[58] and dated according to Magallón et al.[59]. To consolidate the nomenclature between the Smith and Brown[56] backbone phylogenetic tree and the edible plant species list, species names were homogenised against the *World Checklist of Vascular Plants*[60] using the package "Taxonstand"[61]. Some of the edible plant species were absent from the backbone phylogenetic tree (*n* = 3,130 species). Using the *bind.relative* function in the V.Phylomaker package[57], missing species with at least one congeneric species on the tree were attached randomly to a congeneric (*n* = 2,754), and missing species without a congeneric species (*n* = 376) were added randomly to another member of their taxonomic family. We repeated this random addition of species 100 times, to create 101 replicate trees. The resulting set of edible plant trees had 6,740 terminal taxa from 242 angiosperm families. All figures of phylogenetic trees were created using “ggtree”[62] and “ggplot2”[63].

Testing for phylogenetic signal for B vitamins among nutritionally known species

To test for phylogenetic signal among the nutritionally known species, Pagel’s λ was calculated for each B vitamin[64] with its likelihood-ratio test (LRT) using the “phytools” package[65]. Pagel’s λ finds the transformation of the branch lengths that best predicts the trait distribution on the phylogeny expected under a random-walk, viz. Brownian-motion model[66]. Pagel’s λ was chosen over other indices of signal (e.g. Blomberg’s *K*), as it is more robust to missing phylogenetic information[67,68]. It normally ranges between 0 and 1 (though >1 is possible), with 0 representing no phylogenetic signal (closely related species do not share similar nutrition; branch-length transformation results in a “star” phylogeny) and 1 representing closely related species being as similar nutritionally as would be expected under Brownian-motion evolution (no branch-length transformation required[64]). To assess whether the random placement of missing species could have affected measures of signal, we calculated phylogenetic signal for the 101 replicate trees (Supplementary Table 10).

Predicting the B-vitamin profiles of edible plants

To estimate the B-vitamin profiles of the nutritionally unknown species, we followed the method of Vaitla, Collar, et al.[19]. This method assumes a Brownian-motion model of evolution, allowing the B-vitamin profile of a nutritionally unknown edible species to be approximated as the estimated state for the most recent common ancestor (MRCA) between the nutritionally unknown species and its most closely related nutritionally known species[19]. This is because, under Brownian motion, the concentration of a given B vitamin is not expected to change along any branch of the phylogenetic tree[64] – that is, the change in B-vitamin concentration along a branch has an expected mean of zero and an unknown, constant variance[64] - σ^2 (estimated using the *fitContinuous* function from “geiger” R package[69]).

Therefore, for each B vitamin, the edible plant tree was transformed using the lambda values calculated for the nutritionally known species above (Extended Data Fig. 2). Each nutritionally unknown species then took the value of the estimated state for the MRCA between it and its nearest nutritionally known species, where the state for this internal node was taken as the branch-length-weighted mean of the states at the nodes immediately shallower and deeper to it. The uncertainty of each estimate was calculated as $t \times \sigma^2$, where *t* is the branch length of the nutritionally unknown species to its MRCA with a nutritionally known species[64]. Uncertainty therefore increases with the time elapsed since the divergence between the nutritionally unknown edible species and its most closely related nutritionally known species. We constructed 95%-confidence intervals around each predicted value[19] as $\pm 1.96 \times \sqrt{t \times \sigma^2}$.

To validate this prediction method, we used jackknifing[19]. For each B vitamin, one nutritionally known species was removed from the dataset, and its value and 95%-confidence intervals estimated. We then determined if the measured value fell within its estimated confidence intervals (error rate). For each B vitamin, we calculated the median percentage deviation (standardised by the standard deviation) between predicted and measured (nutritionally known) values[19]. Finally, to examine the strength of the relationship between the predicted and observed values, we used generalised least squares models (gls; “nlme” package[70]). As the residual model variance increased with observed values, a model variance structure was included: either as an exponent of the variance covariate (for thiamine, niacin and pantothenic acid) or as a constant plus power of the variance covariate (for riboflavin and folate). This was based on which gave the lowest model Akaike Information Criterion (AIC).

Pyridoxine was not modelled, as its lack of significant signal precluded it from being predicted.

Identifying edible plants that are sources of B vitamins

"Sources" of a B vitamin were defined as contributing $\geq$15% towards recommended dietary allowances[25,27] for active females (31-50 years; RDAs; Supplementary Table 6) per 100 g of fresh edible plant material consumed. RDAs for females were chosen, as females are often more likely to suffer from vitamin deficiencies[71].

To validate which edible plants are sources of each B vitamin, we additionally predicted source species using a second approach that is analogous to the "nodesig" function in PHYLOCOM[72,73] (as in Saslis-Lagoudakis et al.[24]). This second approach uses the phylogenetic tree of nutritionally known species to find taxonomic groups (nodes) that have significantly more descendants that are sources of the vitamin than would be expected from sampling the tips of the tree randomly ("hot nodes"). We then identified the corresponding hot nodes on the edible plant tree – nutritionally unknown plant descendants from these nodes were labelled as being predicted sources. To validate this second approach, we predicted if a given nutritionally known species would be identified as a "source", by removing one species at a time from the phylogenetic tree and predicting the remaining species. As the results of this approach are binary – a species is (or is not) a source – we examined the proportion of nutritionally known species that were correctly identified as being a source (or not), and, among those incorrectly identified, if they were false positives or negatives. Across the five B vitamins, $\geq$75.0% of nutritionally known species were correctly identified. For all B vitamins, except niacin, the percentage of false negatives was higher than that of false positives (ranging for the B vitamins between 0.750-19.3% vs. 0.571-11.8%, respectively). More species were identified as being sources than when using the first approach by Vaitla, Collar, et al.[19], suggesting this approach is likely to give conservative estimates of the number of edible plants that are sources of each nutrient (Supplementary Table 7). Edible species were therefore identified as “sources” of each B vitamin if they were predicted by both the method of Vaitla, Collar, et al.[19] and this second approach.

Conservation status of edible plant diversity

Edible-plant threat status (in situ) was assessed by searching the Botanic Gardens Conservation International (BGCI) ThreatSearch database[32]. ThreatSearch is the most comprehensive database for conservation assessments and contains assessments performed by the International Union for the Conservation of Nature, among other sources[32]. Species absent from the database were labelled as “not assessed”. We selected only global assessments, and for species with multiple entries available, we chose the most recent entry. When entries could not be differentiated using these criteria and each entry was different, we kept all entries (n = 40 species). “Interpreted conservation status” was used. To ascertain whether the edible plants have been conserved ex situ (i.e. have seedbank records), we searched the Millennium Seed Bank Partnership[33] and Genesys Global Portal on Plant Genetic Resources[34].

Geographic distribution of source edible plants

Presence-absence data at level 3 (country to sub-country) of the *World Geographical Scheme for Recording Plant Distribution*[74] were obtained from the *World Checklist of Vascular Plants*[60] for all edible plants, which we used to map the number of native and introduced source species for each B vitamin. We then calculated for each country and sub-country the percentage of source species that are threatened in situ (possibly threatened, threatened or extinct in the wild[32]) or not conserved ex situ, by dividing the number of threatened or not-conserved-ex-situ source species by the total number of source species in that country/sub-country for each B vitamin.

To explore the countries that would benefit most from improving conservation of their edible plant diversity, we visually assessed whether countries with the highest prevalence of inadequate B-vitamin intake also had the highest percentages of source species threatened or not conserved ex situ for each B vitamin. Estimates of the prevalence of inadequate intake were taken from Beal et al.[1], who combine 1) food composition databases, 2) FAO agricultural production, import and export data, and 3) population-weighted Estimated Average Requirements, to calculate the average Prevalence of Inadequate Micronutrient Intake Index for each country between 1961 and 2011 (we used the most recent estimate available: 2011). We reconciled the plant distribution and inadequate-intake data to the same spatial resolution (viz. country level) and produced maps showing the co-distributions of threatened and not-conserved-ex-situ source species and inadequate intake using the “biscale” package[75], with the numerical range for each colour determined using the Fisher natural breaks classification method[76] (Extended Data Figures 3-5; Supplementary File 4).

Data availability: All data used were accessed from publicly available databases and are indicated in the Methods. B-vitamin predictions are available in Supplementary Files 1 and 2. Species predicted to be B-vitamin sources under the conservative and “best-case” scenarios are listed in Supplementary File 3, with the names of the countries in which each consensus source species is found. Supplementary File 4 contains, for each country, the number of source species, percentages of source species that are threatened in situ and not conserved ex situ, and the prevalence of each B-vitamin deficiency. A list of the nutritionally known edible species used is available in Supplementary Table 8.

Code availability: R scripts for testing for phylogenetic signal, performing the predictions for the nutritionally known and unknown species are available in Supplementary Files 5-8.

Acknowledgements: ACJ was funded by the Basil Furneaux Memorial Fund, Imperial College London, and the Natural Environment Research Council [NE/S007415/1]. BV received funding from the Wellcome Trust Our Plant, Our Health programme (Grant number: 106864MA). SP received funding from Royal Botanic Gardens, Kew pilot study fund (reference: 11492-100). We would like to thank M. Soto Gomez, the Gill and Graystock lab group (Imperial College London), and the Science Directorate (Royal Botanic Gardens, Kew) for their feedback on project ideas. We would also like to thank M. Bidartondo for his comments on this manuscript; N. Black, R. Govaerts, I. Ondo, R. Turner for help accessing the World Checklist of Vascular Plants and Useful Plants data; D. Satori and T. Cossu for help accessing conservation data; F. Fletcher for help assembling the nutritional data; K. Kam for his help with code; and P. Jones for proofreading.

Author contributions: SP conceived the project. ACJ and SP designed the study. ACJ and JB collected the nutritional data. MD and TU compiled and provided the list of edible plant species. DC and BV developed and provided the code for the nutrient predictions. SP provided the data on plant distributions. ACJ performed analyses with guidance from SP, FF, JH, MJH and RD. ACJ wrote the manuscript with guidance from SP. All authors provided feedback on the manuscript and gave authorisation for publication.

ETHICS DECLARATION:

Competing interests: No competing interests.

MAIN REFERENCES

1. Beal, T., Massiot, E., Arsenault, J. E., Smith, M. R. & Hijmans, R. J. Global trends in dietary micronutrient supplies and estimated prevalence of inadequate intakes. *PLoS*

One **12**, e0175554 (2017).

2. FAO, IFAD, UNICEF, WFP & WHO. *The State of Food Security and Nutrition in the World 2020. Transforming food systems for affordable healthy diets.* https://doi.org/10.4060/ca9692en (2020).

3. Gernand, A. D., Schulze, K. J., Stewart, C. P., West, K. P. & Christian, P. Micronutrient deficiencies in pregnancy worldwide: health effects and prevention. *Nat. Rev. Endocrinol.* **12**, 274–289 (2016).

4. Pilling, D., Bélanger, J. & Hoffmann, I. Declining biodiversity for food and agriculture needs urgent global action. *Nat. Food* **1**, 144–147 (2020).

5. Nelson, G. *et al.* Income growth and climate change effects on global nutrition security to mid-century. *Nat. Sustain.* **1**, 773–781 (2018).

6. Lachat, C. *et al.* Dietary species richness as a measure of food biodiversity and nutritional quality of diets. *Proc. Natl. Acad. Sci.* **115**, 127–132 (2018).

7. Siddique, K. H. M., Li, X. & Gruber, K. Rediscovering Asia’s forgotten crops to fight chronic and hidden hunger. *Nat. Plants.* **7**, 116–122 (2021).

8. Ulian, T. *et al.* Unlocking plant resources to support food security and promote sustainable agriculture. *Plants, People, Planet* **2**, 421–445 (2020).

9. Powell, B. *et al.* Improving diets with wild and cultivated biodiversity from across the landscape. *Food Secur.* **7**, 535–554 (2015).

10. Hunter, D. *et al.* The potential of neglected and underutilized species for improving diets and nutrition. *Planta.* **250**, 709–729 (2019).

11. Khoury, C. K. *et al.* Increasing homogeneity in global food supplies and the

implications for food security. *Proc. Natl. Acad. Sci.* **111**, 4001–4006 (2014).

12. Magrach, A. & Sanz, M. J. Environmental and social consequences of the increase in the demand for 'superfoods' world-wide. *People Nat.* **2**, 267–278 (2020).

13. Díaz, S. *et al.* Pervasive human-driven decline of life on Earth points to the need for transformative change. *Science.* **366**, eaax3100 (2019).

14. Pironon, S. *et al.* Potential adaptive strategies for 29 sub-Saharan crops under future climate change. *Nat. Clim. Chang.* **9**, 758–763 (2019).

15. Jones, S. K. *et al.* Agrobiodiversity Index scores show agrobiodiversity is underutilized in national food systems. *Nat. Food* **2**, 712–723 (2021).

16. Barrett, C. B. *et al.* Bundling innovations to transform agri-food systems. *Nat. Sustain.* **3**, 974–976 (2020).

17. Diazgranados, M. *et al.* World checklist of useful plant species. *R. Bot. Gard. Kew, London* (2020).

18. Castañeda-Álvarez, N. P. *et al.* Global conservation priorities for crop wild relatives. *Nat. plants* **2**, 1–6 (2016).

19. Vaitla, B. *et al.* Predicting nutrient content of ray-finned fishes using phylogenetic information. *Nat. Commun.* **9**, 1–10 (2018).

20. Agrawal, A. A., Salminen, J. & Fishbein, M. Phylogenetic trends in phenolic metabolism of milkweeds (*Asclepias*): evidence for escalation. *Evol. Int. J. Org. Evol.* **63**, 663–673 (2009).

21. Albuquerque, T. G., Nunes, M. A., Bessada, S. M. F., Costa, H. S. & Oliveira, M. B. P. P. Biologically active and health promoting food components of nuts, oilseeds,

fruits, vegetables, cereals, and legumes. in *Chemical Analysis of Food* 609–656 (Elsevier, 2020).

22. Şerban, P., Wilson, J. R. U., Vamosi, J. C. & Richardson, D. M. Plant diversity in the human diet: weak phylogenetic signal indicates breadth. *Bioscience* **58**, 151–159 (2008).

23. Dempewolf, H., Rieseberg, L. H. & Cronk, Q. C. Crop domestication in the Compositae: a family-wide trait assessment. *Genet. Resour. Crop Evol.* **55**, 1141–1157 (2008).

24. Saslis-Lagoudakis, C. H. *et al.* The use of phylogeny to interpret cross-cultural patterns in plant use and guide medicinal plant discovery: an example from *Pterocarpus* (Leguminosae). *PLoS One* **6**, e22275 (2011).

25. Meyers, L. D., Hellwig, J. P. & Otten, J. J. *Dietary reference intakes: the essential guide to nutrient requirements*. (National Academies Press, 2006).

26. Miller, J. W. & Rucker, R. B. *Present Knowledge In Nutrition.* 273-287. (Academic Press, 2020).

27. Pinela, J., Carvalho, A. M. & Ferreira, I. C. F. R. Wild edible plants: Nutritional and toxicological characteristics, retrieval strategies and importance for today's society. *Food Chem. Toxicol.* **110**, 165–188 (2017).

28. Gruber, K. Agrobiodiversity: The living library. *Nat. 2017 5447651* **544**, S8–S10 (2017).

29. Caproni, L., Raggi, L., Talsma, E. F., Wenzl, P. & Negri, V. European landrace diversity for common bean biofortification: a genome-wide association study. *Sci. Rep.* **10**, 1–13 (2020).

30. Dwivedi, S. L. *et al.* Diversifying food systems in the pursuit of sustainable food production and healthy diets. *Trends Plant Sci.* **22**, 842–856 (2017).

31. Borelli, T. *et al.* Local Solutions for Sustainable Food Systems: The Contribution of Orphan Crops and Wild Edible Species. *Agron.* **10**, 231 (2020).

32. BGCI. ThreatSearch online database. ThreatSearch online database. Botanic Gardens Conservation International. www.bgci.org/threat_search.php (2020).

33. RBG Kew's MSB Partnership. https://www.kew.org/science/our-science/projects/banking-the-worlds-seeds.

34. Global Crop Diversity Trust. Genesys global portal of Plant Genetic Resources for Food and Agriculture. https://www.genesys-pgr.org (2018).

35. Lughadha, E. N. *et al.* Extinction risk and threats to plants and fungi. *Plants, People, Planet* **2**, 389–408 (2020).

36. Vincent, H. *et al.* Modeling of crop wild relative species identifies areas globally for in situ conservation. *Commun. Biol.* **2**, 1–8 (2019).

37. Khoury, C. K. *et al.* Comprehensiveness of conservation of useful wild plants: An operational indicator for biodiversity and sustainable development targets. *Ecol. Indic.* **98**, 420–429 (2019).

38. Quave, C. L. & Pieroni, A. A reservoir of ethnobotanical knowledge informs resilient food security and health strategies in the Balkans. *Nat. Plants* **1**, 1–6 (2015).

39. de Medeiros, P. M. *et al.* Local knowledge as a tool for prospecting wild food plants: experiences in northeastern Brazil. *Sci. Rep.* **11**, 1–14 (2021).

40. Sogbohossou, E. O. D. *et al.* A roadmap for breeding orphan leafy vegetable species: a

case study of *Gynandropsis gynandra* (Cleomaceae). *Hortic. Res.* **5**, 1–15 (2018).

41. Pascual, U. *et al.* Biodiversity and the challenge of pluralism. *Nat. Sustain.* **4**, 1–6 (2021).

42. Amaya, N., Meldrum, G. & Padulosi, S. Promoting chaya and tepary bean to improve diet quality, climate resilience, and incomes in Guatemala. https://cgspace.cgiar.org/bitstream/handle/10568/109363/Guatemala%20Impact%20Brief%20A4.pdf (2020).

43. Davis, D. R., Epp, M. D. & Riordan, H. D. Changes in USDA Food Composition Data for 43 Garden Crops, 1950 to 1999. *J. Am. Coll. Nutr.* **23**, 669–682 (2004).

44. Hotz, C. & Gibson, R. S. Traditional food-processing and preparation practices to enhance the bioavailability of micronutrients in plant-based diets. *J. Nutr.* **137**, 1097–1100 (2007).

METHODS REFERENCES

45. McCance, R. A. & Widdowson, E. M. *McCance and Widdowson's the Composition of Foods*. (Royal Society of Chemistry, 2014).

46. U.S. Department of Agriculture. FoodData Central. https://fdc.nal.usda.gov/fdc-app.html#/food-search (2019).

47. Sivakumaran, S., Huffman, L. & Sivakumaran, S. The New Zealand food composition database: a useful tool for assessing New Zealanders' nutrient intake. *Food Chem.* **238**, 101–110 (2018).

48. MEXT. Standards tables of food composition in Japan. https://www.mext.go.jp/en/policy/science_technology/policy/title01/detail01/1374030.htm (2015).

49. Vincent, A. *et al.* FAO/INFOODS Food Composition Table for Western Africa (2019) User Guide & Condensed Food Composition Table. *Table Compos. des Aliment. FAO/INFOODS Pour L'afrique L'ouest* (2019).

50. MAFOODS. *Malawian Food Composition Table*. (2019).

51. Longvah, T., Anantan, I., Bhaskarachary, K., Venkaiah, K. & Longvah, T. *Indian food composition tables*. (National Institute of Nutrition, Indian Council of Medical Research Hyderabad, 2017).

52. Dignan, C., Burlingame, B., Kumar, S. & Aalbersberg, W. *The Pacific Islands food composition tables*. (2004).

53. Ray, A., Ray, R. & Sreevidya, E. A. How many wild edible plants do we eat—Their diversity, use, and implications for sustainable food system: An exploratory analysis in India. *Front. Sustain. Food Syst.* **4**, 56 (2020).

54. R Core Team. R: A language and environment for statistical computing. (2020).

55. Koyande, A. K. *et al.* Microalgae: A potential alternative to health supplementation for humans. *Food Sci. Hum. Wellness* **8**, 16–24 (2019).

56. Smith, S. A. & Brown, J. W. Constructing a broadly inclusive seed plant phylogeny. *Am. J. Bot.* **105**, 302–314 (2018).

57. Jin, Y. & Qian, H. V. PhyloMaker: an R package that can generate very large phylogenies for vascular plants. *Ecography.* **42**, 1353–1359 (2019).

58. Sayers, E. W. *et al.* GenBank. *Nucleic Acids Res.* **47**, D94–D99 (2019).

59. Magallón, S., Gómez-Acevedo, S., Sánchez-Reyes, L. L. & Hernández-Hernández, T. A metacalibrated time-tree documents the early rise of flowering plant phylogenetic

diversity. *New Phytol.* **207**, 437–453 (2015).

60. Govaerts, R., Nic Lughadha, E., Black, N., Turner, R. & Paton, A. The World Checklist of Vascular Plants, a continuously updated resource for exploring global plant diversity. *Sci. Data* **8**, 1–10 (2021).

61. Cayuela, L., Granzow-de la Cerda, Í., Albuquerque, F. S. & Golicher, D. J. Taxonstand: An R package for species names standardisation in vegetation databases. *Methods Ecol. Evol.* **3**, 1078–1083 (2012).

62. Yu, G., Smith, D. K., Zhu, H., Guan, Y. & Lam, T. T. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* **8**, 28–36 (2017).

63. Wickham, H. *et al.* Package “ggplot2”, Create Elegant Data Visualisations Using the Grammar of Graphics, version 3.1. 1. (2019).

64. Pagel, M. Inferring the historical patterns of biological evolution. *Nature* **401**, 877–884 (1999).

65. Revell, L. J. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217–223 (2012).

66. Swenson, N. G. *Functional and phylogenetic ecology in R.* (Springer, 2014).

67. Münkemüller, T. *et al.* How to measure and test phylogenetic signal. *Methods Ecol. Evol.* **3**, 743–756 (2012).

68. Molina-Venegas, R. & Rodríguez, M. Á. Revisiting phylogenetic signal; strong or negligible impacts of polytomies and branch length information? *BMC Evol. Biol.* **17**, 1–10 (2017).

69. Harmon, L. J., Weir, J. T., Brock, C. D., Glor, R. E. & Challenger, W. GEIGER: investigating evolutionary radiations. *Bioinformatics* **24**, 129–131 (2008).

70. Pinheiro, J. *et al.* Package ‘nlme’. *Linear nonlinear Mix. Eff. Model. version* **3**, (2017).

71. Bird, J. K., Murphy, R. A., Ciappio, E. D. & McBurney, M. I. Risk of deficiency in multiple concurrent micronutrients in children and adults in the United States. *Nutrients* **9**, 655 (2017).

72. Webb, C. O., Ackerly, D. D. & Kembel, S. W. Phylocom: software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics* **24**, 2098-2100 (2008).

73. Abellán, P., Carrete, M., Anadón, J. D., Cardador, L. & Tella, J. L. Non-random patterns and temporal trends (1912–2012) in the transport, introduction and establishment of exotic birds in Spain and Portugal. *Divers. Distrib.* **22**, 263–273 (2016).

74. Brummitt, R. K., Pando, F., Hollis, S. & Brummitt, N. *World Geographical Scheme for Recording Plant Distributions*. (2001).

75. Prener, C., Grossenbacher, T. & Zehr, A. biscale: Tools and palettes for bivariate thematic mapping. (2020).

76. Fisher, W. D. On grouping for maximum homogeneity. *J. Am. Stat. Assoc.* **53**, 789–798 (1958).

ADDITIONAL INFORMATION:

Supplementary Information is available for this paper:

Extended Data:

- Extended Data Fig. 1: Predicted versus observed values (and their relationships) for B vitamins in nutritionally known species.
- Extended Data Fig. 2: Summary of the method used to estimate the B-vitamin profiles of nutritionally unknown species.
- Extended Data Fig. 3: Overlap of the number of edible plant species that are sources of each B vitamin per region with the prevalence of inadequate B-vitamin intake.
- Extended Data Fig. 4: Overlap of the percentage of threatened-in-situ B-vitamin-source species per region with the prevalence of inadequate B-vitamin intake.
- Extended Data Fig. 5: Overlap of the percentage of B-vitamin-source species that are not conserved ex situ per region with the prevalence of inadequate B-vitamin intake.

Supplementary:

- Supplementary Table 1: Phylogenetic signal of each B vitamin across nutritionally known species.
- Supplementary Table 2: Strength of phylogenetic signal (Pagel's λ) when the nutritionally known species were split by their tissue consumed.
- Supplementary Table 3: Nutrient summary of the two B-vitamin-"richest" and "-poorest" orders within nutritionally known edible plants.

- Supplementary Table 4: Performance of predicting B-vitamin values of nutritionally known species based on phylogenetic-signal strength and the values of close relatives.
- Supplementary Table 5: Examining the strength of the relationship between the predicted and observed values for nutritionally known species.
- Supplementary Table 6: Recommended dietary allowances for B vitamins, split by sex and life stages.
- Supplementary Table 7: Number of edible plant species predicted to be sources (contribute ≥15% to recommended dietary allowances) of each B vitamin using the different prediction methods.
- Supplementary Table 8: List of nutritionally known species and their source databases.
- Supplementary Table 9: Summary of the analytical methods used by food-composition databases to extract and measure B vitamins.
- Supplementary Table 10: Variation in phylogenetic signal of each B vitamin across nutritionally known species, depending on the placement of the missing species.
- Supplementary Methods 1: Investigating how the presence of different plant tissues affects phylogenetic signal.

Supplementary File 1: Predicted values for nutritionally known species (using jackknifing) for thiamine, riboflavin, niacin, pantothenic acid and folate. Following the approach of Vaitla, Collar, et al. (2018).

Supplementary File 2: Edible plant predictions (nutritionally unknown species) for thiamine, riboflavin, niacin, pantothenic acid and folate. Following the approach of Vaitla, Collar, et al. (2018).

Supplementary File 3: Consensus and best-case B-vitamin source species. These plants were either predicted to be B-vitamin sources by both prediction methods or at least one prediction method, respectively. The countries in which the consensus source species are found are also given.

Supplementary File 4: For each country, the number of source species, the percentages of source species that are threatened in situ and not conserved ex situ, and the prevalence of deficiency of thiamine, riboflavin, niacin and folate.

Supplementary File 5: R script for measuring phylogenetic signal of B vitamins in nutritionally known species.

Supplementary File 6: R script for predicting the B-vitamin concentrations for nutritionally unknown species following the method of Vaitla, Collar et al. (2018).

Supplementary File 7: R script for validating the prediction method of Vaitla, Collar et al. (2018) by predicting the B-vitamin concentrations of nutritionally known species.

Supplementary File 8: R script for predicting B-vitamin-source species using the “hot-node” approach.