PLANT SCIENCES

Leaf metabolic traits reveal hidden dimensions of plant form and function

Tom W. N. Walker[1,2]*, Franziska Schrodt[3], Pierre-Marie Allard[4,5,6], Emmanuel Defossez[1], Vincent E. J. Jassey[7], Meredith C. Schuman[8], Jake M. Alexander[2], Oliver Baines[3,9], Virginie Baldy[10], Richard D. Bardgett[11], Pol Capdevila[12,13], Phyllis D. Coley[14], Nicole M. van Dam[15,16,17], Bruno David[18], Patrice Descombes[2,19,20], María-José Endara[21], Catherine Fernandez[10], Dale Forrister[14], Albert Gargallo-Garriga[22,23,24], Gaëtan Glauser[1], Sue Marr[17,25,26], Steffen Neumann[17,25], Loïc Pellissier[2,19], Kristian Peters[17,25,26], Sergio Rasmann[1], Ute Roessner[27], Roberto Salguero-Gómez[28], Jordi Sardans[22,23], Wolfram Weckwerth[29,30], Jean-Luc Wolfender[5,6], Josep Peñuelas[22,23]

The metabolome is the biochemical basis of plant form and function, but we know little about its macroecological variation across the plant kingdom. Here, we used the plant functional trait concept to interpret leaf metabolome variation among 457 tropical and 339 temperate plant species. Distilling metabolite chemistry into five metabolic functional traits reveals that plants vary on two major axes of leaf metabolic specialization—a leaf chemical defense spectrum and an expression of leaf longevity. Axes are similar for tropical and temperate species, with many trait combinations being viable. However, metabolic traits vary orthogonally to life-history strategies described by widely used functional traits. The metabolome thus expands the functional trait concept by providing additional axes of metabolic specialization for examining plant form and function.

INTRODUCTION

Plants produce a staggering diversity of metabolites—upward of one million throughout the plant kingdom and many thousands within an individual (1) (the metabolome). The plant metabolome is well known to human society, being the source of most cosmetics (2), flavors (3), and medicines (4). Yet, little is known about how the metabolome varies across the plant kingdom. The metabolome is the biochemical basis of physiology, comprising primary metabolites involved in cellular function (e.g., carbohydrates, amino acids, and nucleic acids) and specialized metabolites produced to confront stress (e.g., flavonoids and terpenoids) (1, 5, 6). The metabolome thus encompasses the biochemical mechanisms through which evolutionary and ecological processes shape plant form and function [i.e., morphology and physiology (7)] (8, 9). Plant form and function vary predictably among species because evolutionary and ecological processes generally act to maximize fitness (10–14),

thus limiting the number of morphological or physiological strategies that are viable (7, 15, 16). It follows that the metabolome should also be constrained to a limited number of strategies that maximize fitness (17). However, although cross-species studies are now emerging (18–21), no attempt has been made to characterize how the metabolome varies systematically throughout the plant kingdom. We thus remain unable to place the vast diversity of the plant metabolome into a coherent ecological context.

The plant functional trait concept offers a powerful framework to contextualize the plant metabolome. Plant functional traits are a standardized set of morphological and physiological characteristics that provide universal metrics of plant form and function (22). Plant functional traits have been pivotal in identifying the life-history strategies that govern plant fitness (7, 15, 16); understanding plant impacts on population (23), community (24), and ecosystem (25) processes; and predicting global change effects on ecosystems

[1]Institute of Biology, University of Neuchâtel, 2000 Neuchâtel, Switzerland. [2]Department of Environmental Systems Science, ETH Zürich, 8092 Zürich, Switzerland. [3]School of Geography, University of Nottingham, Nottingham NG7 2RD, UK. [4]Department of Biology, University of Fribourg, 1700 Fribourg, Switzerland. [5]School of Pharmaceutical Sciences, University of Geneva, 1211 Geneva, Switzerland. [6]Institute of Pharmaceutical Sciences of Western Switzerland, University of Geneva, 1211 Geneva, Switzerland. [7]Laboratoire d'Ecologie Fonctionnelle et Environnement, Université de Toulouse, CNRS, 31062 Toulouse, France. [8]Departments of Geography and Chemistry, University of Zurich, 8057 Zurich, Switzerland. [9]Section for Ecoinformatics and Biodiversity, Department of Biology, Aarhus University, 8000 Aarhus, Denmark. [10]Aix Marseille Université, Avignon Université, CNRS, IRD, IMBE, Marseille, France. [11]Department of Earth and Environmental Sciences, The University of Manchester, Manchester M13 9PT, UK. [12]Departament de Biologia Evolutiva, Ecologia i Ciències Ambientals, Facultat de Biologia, Universitat de Barcelona (UB), Barcelona 08028, Spain. [13]Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona, Barcelona 08028, Spain. [14]School of Biological Sciences, University of Utah, Salt Lake City, UT 84112, USA. [15]Leibniz Institute of Vegetable and Ornamental crops (IGZ), 14979 Großbeeren, Germany. [16]Institute of Biodiversity, Friedrich Schiller University Jena, 07743 Jena, Germany. [17]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, 04103 Leipzig, Germany. [18]Green Mission Pierre Fabre, Institut de Recherche Pierre Fabre, 31562 Toulouse, France. [19]Ecosystems and Landscape Evolution, Swiss Federal Institute for Forest, Snow and Landscape Research (WSL), 8903 Birmensdorf, Switzerland. [20]Musée et Jardins botaniques cantonaux, 1007 Lausanne, Switzerland. [21]Medio Ambiente y Salud (BIOMAS), Facultad de Ingenierías y Ciencias Aplicadas, Universidad de Las Américas, 170124 Quito, Ecuador. [22]CSIC, Global Ecology Unit CREAF-CSIC-UAB, 08193 Bellaterra, Catalonia, Spain. [23]CREAF, 08193 Cerdanyola del Vallès, Catalonia, Spain. [24]Global Change Research Institute, Czech Academy of Sciences, 603 00 Brno, Czech Republic. [25]Leibniz Institute of Plant Biochemistry, Bioinformatics and Scientific Data, 06120 Halle, Germany. [26]Institute of Biology/Geobotany and Botanical Garden, Martin Luther University Halle Wittenberg, 06108 Halle, Germany. [27]Research School of Biology, The Australian National University, 2601 Acton, Australia. [28]Department of Zoology, Oxford University, Oxford OX1 3SZ, UK. [29]Molecular Systems Biology, Department of Functional and Evolutionary Ecology, 1010 University of Vienna, Vienna, Austria. [30]Vienna Metabolomics Center, 1010 University of Vienna, Vienna, Austria.
*Corresponding author. Email: thomas.walker@unine.ch

(26). Plant functional traits thus provide a strong reference point from which to interpret variation in the plant metabolome. However, measurements of the metabolome remain absent from the functional trait concept, partly because they yield complex data describing thousands of metabolites from hundreds of biochemical pathways (17). Advances in chemoinformatics make it now possible to translate metabolite identities into a targeted set of chemical properties (27, 28). Doing so resonates with the functional trait concept—i.e., that quantitative characteristics provide a bridge between identity and function (22)—and is an intuitive way to contextualize the metabolome against plant functional traits. Yet, while chemoinformatics is commonplace in natural products chemistry (29–31), it remains absent from ecology.

Contextualizing the plant metabolome against existing plant functional traits raises two alternative hypotheses. On the one hand, if the metabolome varies colinearly with functional traits, then it offers a biochemical validation of the functional trait concept. This validation is necessary because functional traits are emergent properties of multiple biochemical processes (22, 32), making them proxies for physiology that can be poor predictors of ecological processes (33–35). On the other hand, if the metabolome varies orthogonally to functional traits, then it describes dimensions of plant life-history variation missed by existing functional traits. This orthogonal variation would enhance the explanatory power of the functional trait concept by providing additional axes of trait specialization to examine plant form and function. Either way, integrating measurements of the plant metabolome into the plant functional trait concept will yield insight into how the metabolome varies across the plant kingdom while also offering a high-resolution mechanistic lens with which to advance the functional trait concept itself.

Here, we performed an assessment of the leaf metabolomes of 457 tropical and 339 temperate plant species to characterize variation in the metabolome and test whether it validates, or expands upon, existing plant functional traits. First, we examined the chemical properties of the whole set of identified metabolites ($N = 4292$) to derive a minimum number of "metabolic functional traits" that capture chemical variation in the leaf metabolome. We then characterized how metabolic traits vary among species to identify major axes of leaf metabolic specialization in the plant kingdom. Last, we combined metabolic traits with eight widely used functional traits (plant height, seed mass, stem-specific density, leaf area, specific leaf area (SLA) and leaf carbon, and nitrogen and phosphorus concentrations; hereafter "classical traits") (7) to determine whether axes of metabolic specialization are colinear with, or orthogonal to, plant life-history strategies described by classical functional traits.

RESULTS

Five chemical properties capture chemical variation in the leaf metabolome

We expressed the leaf metabolomes of tropical (36) and temperate (37) plant species as the presence or absence of annotated metabolic features (i.e., metabolites) detected through liquid chromatography–mass spectrometry (LC-MS) (tropical, $N = 3356$; temperate, $N = 2227$). We then characterized the chemistry of the leaf metabolome by examining variation in the chemical properties of all identified metabolites from both datasets ($N = 4292$) (31). We focused on 21 properties describing metabolite constitution, geometry, or topology that have relationships with (bio)chemical activity (table S1). Pairwise correlations among chemical properties (Fig. 1A, top) reveal that properties separate into five clusters, each representing a distinct facet of leaf metabolite chemistry. The largest cluster (Fig. 1A, green) describes overall size and structural complexity, comprising total numbers of atoms and bonds, molecular weight, a predictor of polarity based on the proportion of noncarbon atoms [M log P; (38)], and related topological parameters [i.e., Wiener numbers (39) and eccentric connectivity (40)]. A second cluster (Fig. 1A, purple) describes bond conjugation and reactivity, comprising numbers of aromatic atoms and bonds, numbers of atoms in the largest conjugated (pi-)system, and a derived complexity index indicative of promiscuous activity in drugs [f_{MF}; (41)]. A third cluster (Fig. 1A, yellow) describes polar intermolecular (noncovalent) forces, comprising total and mass-specific polar surface areas and numbers of hydrogen (H) bond donors or acceptors (42, 43). These forces are important for the selective binding of small molecules to DNA (44), RNA (45), and proteins (46). A fourth cluster (Fig. 1A, blue) comprises two predictors of metabolite polarity based on summed properties of atoms and bonds [X log P (47) and A log P (48)], where lower values indicate higher polarity. A final cluster (Fig. 1A, red) describes carbon bond saturation via the relative occurrence of sp^3 (saturated, three-dimensional) and sp^2 (unsaturated, planar) hybridization (49). Overall, these clusters indicate that the leaf metabolite chemistry is highly structured and can be described by five distinct facets: size/complexity, bond conjugation/reactivity, contributions to polar intermolecular forces, general polarity, and carbon bond saturation.

The strong collinearity among chemical properties (Fig. 1A, top) suggests that the five facets of leaf chemistry can be represented by a small number of discriminant variables. Indeed, we observe that a principal components analysis (PCA) performed on a subset of five chemical properties, one from each cluster (Fig. 1C), is equivalent to a PCA performed on all 21 properties (PC1: $r_{4290} = 0.97$; PC2: $r_{4290} = 0.99$; PC3: $r_{4290} = 0.93$; $P < 0.001$ in all cases; Fig. 1B, left). We selected molecular weight, H-bond acceptor count, aromatic atom count, a-polarity (X log P), and carbon hybridization (sp^3:sp^2) as five properties that are both general structural characteristics of leaf chemistry and intuitive for nonspecialists to interpret (Fig. 1C and table S1). However, most combinations of properties covering all facets of leaf chemistry also yield equivalent PCAs (fig. S5). Pairwise correlations between all properties and subset principal component (PC) scores further confirm that selected properties capture observed leaf chemical variation, in that correlations are consistent within each cluster (Fig. 1A, bottom). Selected properties also describe biologically relevant chemical space according to an existing framework that classifies natural products based on size, then aromaticity, and then lipophilicity (PC1: $r_{4290} = 0.99$; PC2: $r_{4290} = 0.87$; $r_{4290} = 0.81$; $P < 0.001$ in all cases; Fig. 1B, right) (31). The first axis of the subset PCA (38% variance) separates heavy metabolites with many H-bond acceptors from light metabolites with fewer H-bond acceptors. Alignment between metabolite size and numbers of H-bond acceptors is expected because larger molecules usually engage in stronger intermolecular interactions (50). The second axis (37% variance) separates reactive metabolites with more aromatic carbon bonds from unreactive metabolites with more saturated carbon bonds. This axis is also intuitive because conjugated bond systems are unsaturated, reactive, and planar, whereas saturated bonds, such as those abundant in fatty acids, are unreactive and three-
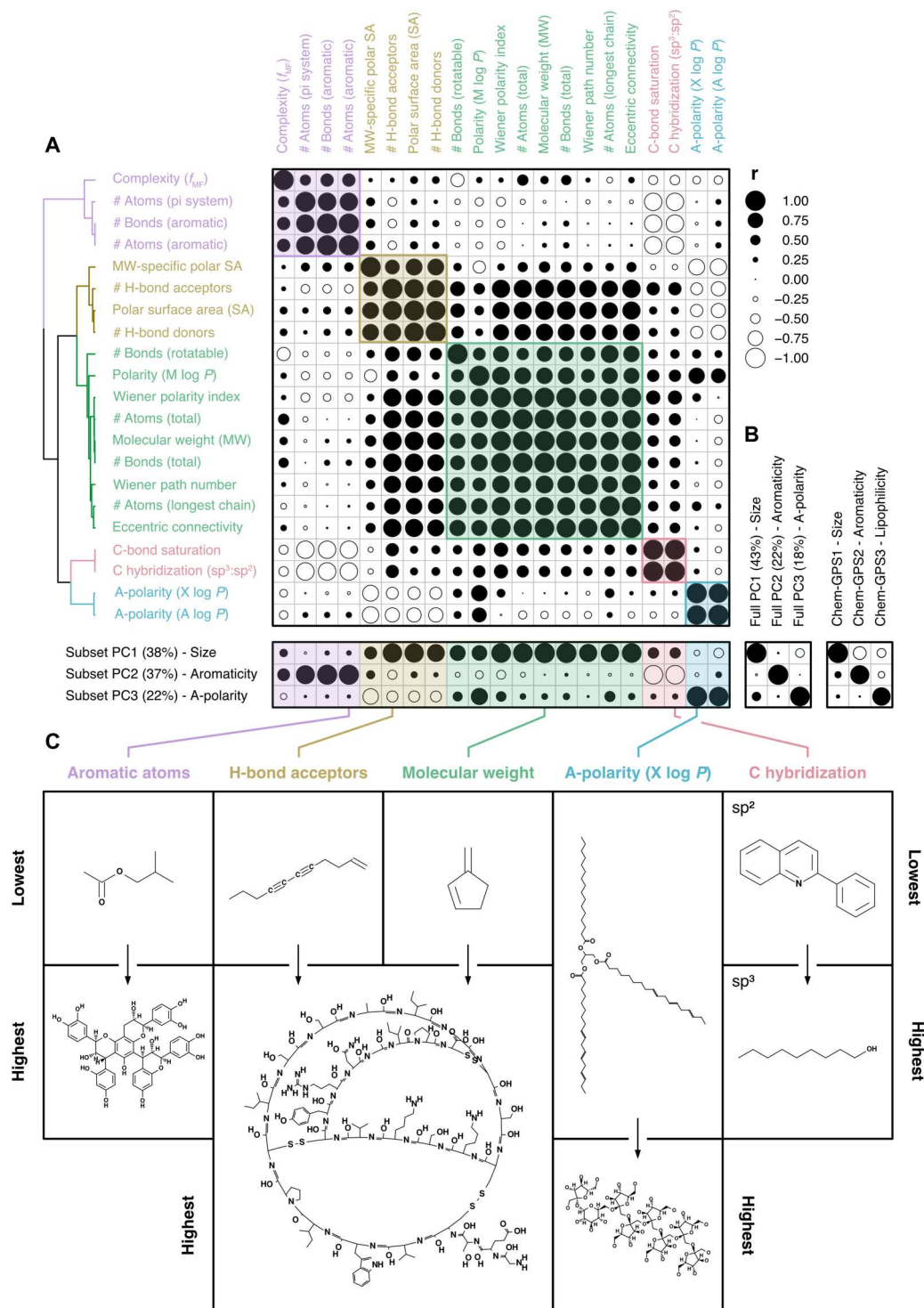
Fig. 1. Variation in leaf metabolite chemistry can be captured by five chemical properties. (**A**) Coefficients for Pearson correlations among 21 chemical properties (table S1) of identified annotated metabolites detected in tropical and temperate datasets ($N = 4292$). Properties are ordered on the basis of correlation similarity (hierarchical clustering; Ward), with colors separating distinct clusters of properties (maximal silhouette width: $K = 5$). The bottom panel shows correlations between chemical properties and scores from the first three axes of a PCA performed on a subset of five properties, as well as (**B**) between subset PC scores and either scores from a PCA performed on all 21 properties (left) or metabolite positions on the first three axes of ChemGPS-NP (*31*) space (right panel). Circle sizes indicate coefficient strengths, while fills discriminate between positive (black) and negative (white) correlations. (**C**) Chemical properties selected to represent each cluster, alongside chemical graphs of metabolites with the lowest and highest values for each trait.

dimensional (51). The third axis (18% variance) separates nonpolar metabolites from polar metabolites (29). Our findings thus show that leaf chemical variation can be represented by five chemical properties, one from each facet of leaf chemistry.

Five chemical properties provide more information than metabolite family

The five selected chemical properties not only represent leaf metabolite chemistry (Fig. 1), but also reflect known differences among metabolite families. Using the subset PCA to interpret variation among metabolite families (Fig. 2), we observe that, relative to other identified metabolites, peptides are large, display average bond saturation, and are moderately polar; carbohydrates are of average size, display high bond saturation, and are highly polar; and fatty acids are smaller, display high bond saturation, and are of average polarity. Lignans, alkaloids, coumarins, and flavonoids, families of specialized metabolites, display low bond saturation due to the presence of many aromatic bonds, with alkaloids and coumarins also being relatively small. Terpenoids, a family of abundant specialized metabolites ($N = 2306$), vary in size, bond saturation, and polarity, echoing their diversity across the plant kingdom (52). Although chemical properties generally differentiate among metabolite families, all families have variable and overlapping chemistries (Fig. 2, marginal boxplots). We thus suggest that chemical properties, quantitative metrics of the chemistry that underpins metabolite composition and activity, provide more precise information about metabolite function than metabolite family identity. Drug discovery and medicinal chemistry rely on many of the same chemical properties to screen for bioactive molecules (29–31). Such a perspective resonates with the functional trait concept by reaching beyond taxonomy to describe function (22) and provides a strong foundation for a novel set of metabolic functional traits.

Five metabolic functional traits describe two axes of leaf metabolic specialization

We converted selected properties into plant-level metabolic functional traits to examine how the leaf metabolome varies among tropical and temperate plant species. Tropical and temperate datasets were analyzed separately due to differences in data structure (see Materials and Methods). We first compared the volume and "lumpiness" (i.e., aggregation around values) of multidimensional space occupied by metabolic traits to four null models to test whether species converge on a restricted set of interrelated trait combinations (7). Observed hypervolumes of metabolic traits are 98.1 to 99.8% smaller and 71.0 to 94.7% lumpier than those of three null models assuming no trait covariance (tables S2 and S3; $P < 0.001$ in all cases). However, observed hypervolumes are larger (tropical: $P = 0.002$; temperate: $P = 0.047$) than that of the null model assuming trait covariance and selection against extreme values while also being similarly lumpy (tropical: $P = 0.662$; temperate: $P = 0.858$). By comparison, observed hypervolumes of eight classical traits are smaller than null models assuming no trait covariance (table S2; $P < 0.001$ in all cases), either the same size as (tropical species; $P = 0.154$) or smaller than (temperate species; $P = 0.018$) that of the null model assuming selection against extreme values, and lumpier than all null models (table S3; $P < 0.001$ in all cases). Collectively, these findings yield three insights. First, metabolic functional traits are interrelated, so plants are constrained to a limited number of viable metabolic trait combinations. Second, while extreme metabolic trait values may be selected against in temperate species, this is not so for tropical species. Expressing extreme leaf metabolome chemistry is thus a viable life-history strategy in tropical environments, potentially due to more intense biotic interactions selecting for a diversity of defensive leaf metabolites and thus divergence of chemical strategies (53). Third, metabolic traits are less aggregated around nonextreme values than classical traits. Hence, although plants converge toward one of several viable classical trait strategies [e.g., woody and nonwoody; (7)], they can display a wider array of leaf metabolome strategies.

We performed PCAs on metabolic traits to characterize major axes of leaf metabolic specialization among species. Metabolic trait variation is consistent among tropical and temperate species, with 92% variation being explained by the first two PCs (Fig. 3). The first axis (tropical: 52%; temperate: 49%) separates species that produce more unsaturated, aromatic, nonpolar leaf metabolites from those that produce more saturated, polar, leaf metabolites. This axis likely describes a spectrum of leaf chemical defense strategies (12). Unsaturated aromatic metabolites, such as alkaloids, coumarins, and flavonoids (Fig. 2), are reactive and serve as toxins or antioxidants in response to stress (54–56). For instance, conjugated bond structures can interfere with protein function by binding covalently to sidechains (54), generate or quench oxidative stress (55), or absorb damaging wavelengths of light (56). By contrast, saturated nonaromatic metabolites can play different roles in leaf defense, both directly as toxins (57) and indirectly as signaling molecules (58). An alternative explanation is that the first axis reflects an investment (or not) into defense metabolites (12), which could arise due to the high cost of synthesizing aromatic metabolites (13). However, all plants must defend their leaves against stress, and we find that plants can maximize fitness at either end of this axis (Fig. 3). It is thus more plausible that the first axis represents a leaf chemical defense spectrum.

The second PCA axis (tropical: 40%; temperate: 43%) discriminates species that produce larger metabolites with many H-bond acceptors from those that produce smaller metabolites with fewer H-bond acceptors. This axis likely relates to leaf longevity because large metabolites, such as lignans, peptides, and waxes, form the basis of long-lived physical (59, 60) and storage (61) structures. Many of these structures also depend on strong intermolecular forces to form stable macromolecules or complexes (62, 63), which are enhanced by H-bond donors and acceptors (43). As for the first axis, plants can maximize fitness at either end of the second axis, with species at the positive end of the axis investing in metabolites indicative of a longer leaf life span. Hence, the second axis is probably a manifestation of the fast-slow continuum (64) at the leaf level. Alternatively, the second axis may reflect an investment (or not) into structurally complex metabolites for physical defense, such as lignans or waxes (59, 60). However, these metabolites act by creating rigid structures that lengthen the leaf life span anyway (65), so physical protection is one component of leaf longevity. Together, metabolic trait PCAs reveal two major axes of leaf metabolic specialization. These axes relate to metabolite bond saturation and polarity (i.e., a leaf chemical defense spectrum) as well as metabolite size and propensity for intermolecular interactions (i.e., leaf longevity), with many combinations yielding successful life-history strategies.
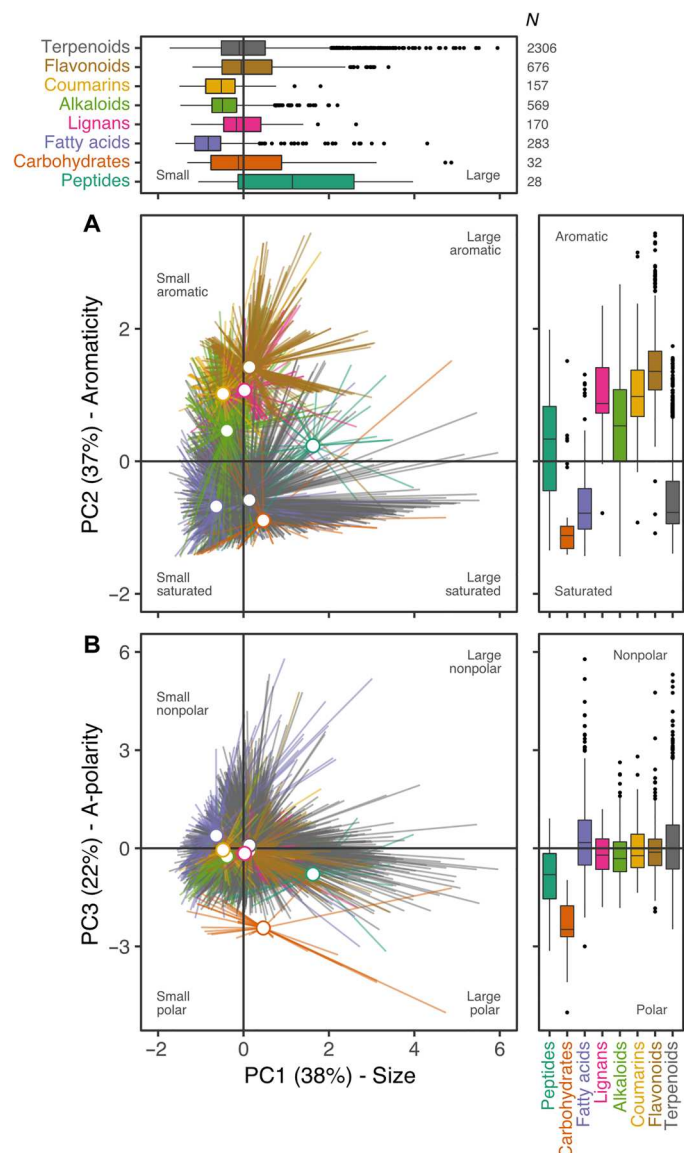
Fig. 2. Variation among and within eight metabolite families is described by five chemical properties. Biplots of PC1 scores and (**A**) PC2 scores or (**B**) PC3 scores from a PCA describing variation in five selected chemical properties (C hybridization, H-bond acceptors, molecular weight, polarity, and aromatic atoms) among all annotated leaf metabolites detected in 862 plant species ($N = 4292$; line ends: metabolites; centroids: means). The PCA is functionally equivalent to a PCA containing all 21 chemical properties (see main text), with axes separating leaf metabolites based on size (PC1), aromaticity (PC2), and a-polarity (PC3; see Fig. 1A, bottom). Colors separate metabolite families as indicated by marginal boxplot fills/labels [center: median; box: 25 to 75% quantiles; whiskers: $1.5 \times$ IQR (interquartile range); points: outliers beyond whiskers].

Axes of leaf metabolic specialization reveal hidden dimensions of plant form and function

The apparent consistency of metabolic trait variation with classical life-history theory (*7, 12, 13, 15, 16, 64, 65*) raises the possibility that metabolic traits describe aspects of plant form and function already captured by classical traits. We thus performed a PCA on metabolic traits plus the same widely used functional traits used above to test
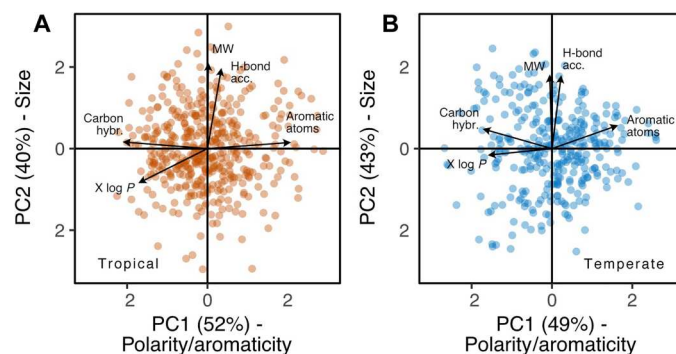

Fig. 3. Metabolic traits describe two axes of leaf metabolic specialization among plant species. Biplots of PC1 and PC2 scores for PCAs describing variation in five metabolic functional traits (arrows) between (**A**) tropical (brown; $N = 457$) and (**B**) temperate (blue; $N = 339$) plant species. Axes separate species based on the mean aromaticity versus polarity (inverse X log *P*) and carbon bond saturation (PC1) and mean size/complexity (PC2) of all annotated leaf metabolites present (Materials and Methods).

whether axes of leaf metabolic specialization are colinear with, or orthogonal to, trait-based life-history strategies. Considered alone, classical trait variation supports life-history theory (*7*), with the first two PCs (tropical: 49%; temperate: 51%) describing the leaf economics spectrum [SLA, leaf nitrogen, and leaf phosphorus; (*16*)] and variation indicative of plant longevity [plant height and seed mass; (*64*)] and hardiness [stem density and leaf carbon; (*15*)] (fig. S1). However, when combining classical and metabolic traits in the same PCA, the first four PCs (tropical: 66%; temperate: 67%) are equivalent to axes described by individual PCAs (Fig. 4A, lower marginal matrices). The first and third axes describe the first two axes of leaf metabolic specialization (Fig. 3), whereas the second and fourth axes describe the first two axes of classical trait specialization (fig. S1). Pairwise correlations among all traits further show that metabolic and classical traits do not covary (Fig. 4, A and B) but separate into six clusters describing unique aspects of either metabolic or classical trait variation. Hence, major axes of leaf metabolic specialization vary orthogonally to major axes of classical trait variation (fig. S4) and are at least equally relevant for describing variation in plant form and function among tropical and temperate plant species.

Leaf metabolism underpins energy production via photosynthesis (*66*) but, in doing so, yields tissue that is sensitive to light (*67*) and temperature (*68*) and is of high nutritional value to herbivores (*65*). Plants are thus under selective pressure to maximize photosynthetic rates while also protecting leaf tissue against damage. While this trade-off is consistent with life-history trade-offs derived from classical traits (*7, 15, 16*), it does not necessarily follow that individuals apply the same strategy in every organ to maintain fitness. We observe that species with "productive" trait values (e.g., high SLA and high leaf nitrogen) can produce leaves containing many complex aromatic metabolites, species with "hardy" trait values (e.g., high stem density and high leaf carbon) can produce leaves containing many simple polar metabolites, and species with trait values typical of a fast pace of life (e.g., low stature and small seeds) can produce leaves with many large stable or structural metabolites. Our findings thus question classical trait (*16*) and plant defense (*12*) theory that predicts relationships between the leaf
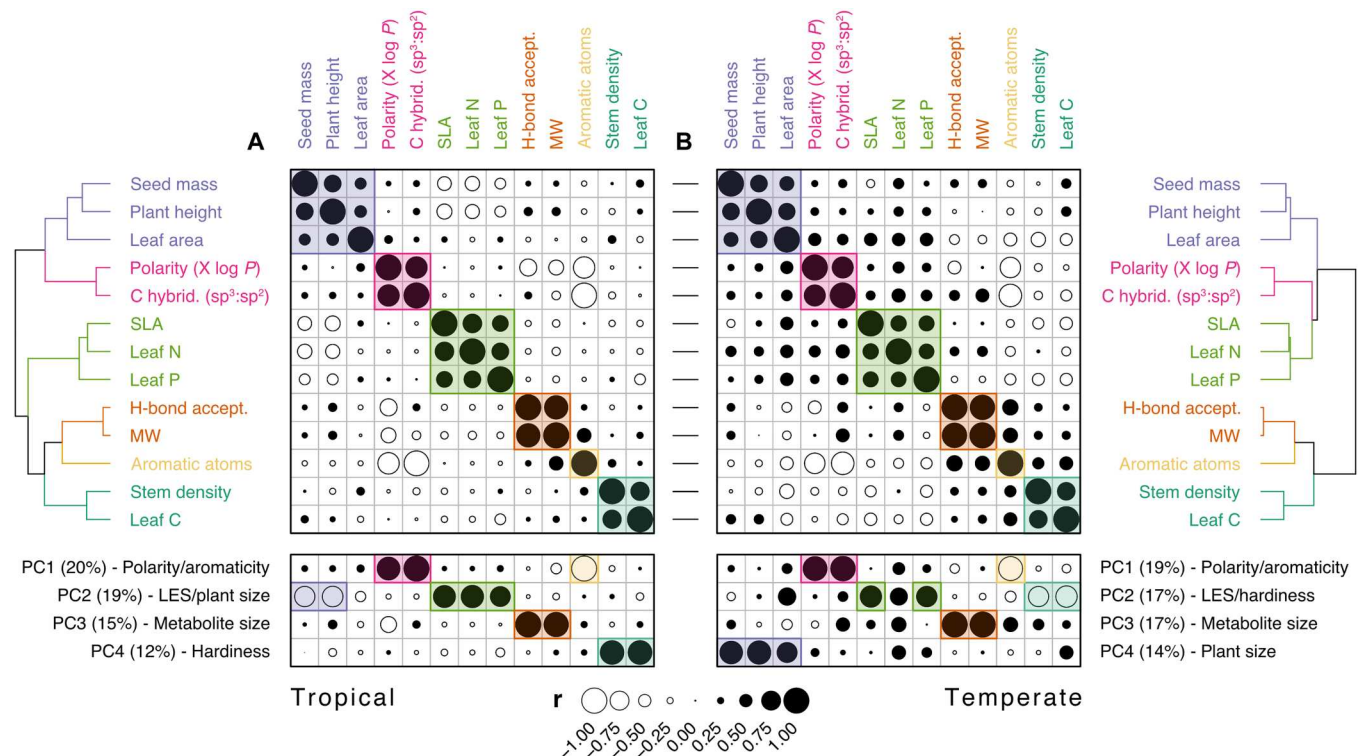
Fig. 4. Metabolic and classical traits describe unique dimensions of plant form and function. Correlation coefficients for pairwise Pearson correlations among metabolic (Fig. 3) and classical functional traits (fig. S1) in (**A**) tropical ($N = 457$) and (**B**) temperate ($N = 339$) species. Traits are ordered on the basis of correlation similarity (hierarchical clustering; Ward), with colors separating distinct clusters (maximal silhouette width: $K = 6$) and horizontal lines between matrices showing correspondence between tropical and temperate traits. Bottom panel shows correlations between traits and scores from the first four axes of a PCA of the same traits (tropical: 66%; temperate: 67%; LES, leaf economics spectrum). Circle sizes indicate coefficient strengths, while fills discriminate between positive (black) and negative (white) correlations.

chemical phenotype, plant productivity, and pace of life. In short, metabolic functional traits describe unique dimensions of plant life-history variation that are complementary to, and independent from, those captured by classical functional traits.

DISCUSSION

In summary, we show that leaf metabolite chemistry is highly structured and can be described by five chemical properties. Using chemical properties as metabolic functional traits reveals that plants vary on two axes of leaf metabolic specialization—a leaf chemical defense spectrum (i.e., bond saturation and polarity) and leaf longevity (i.e., size and propensity for intermolecular interactions). Axes of metabolic specialization are similar in tropical and temperate species, with many combinations of trait values yielding successful life-history strategies. Axes are also consistent with expectations from life-history theory that plants invest strategically in tissue defense (*12, 13, 16*) and between a fast versus slow pace of life (*7, 64*). Nevertheless, we find that metabolic trait variation is orthogonal to, not colinear with, classical trait variation. Metabolic functional traits thus reveal dimensions of plant life-history variation missed by existing functional traits. We note that metabolomics data were acquired differently for tropical and temperate species and were matched to open functional trait databases. Nevertheless, we find consistent variation in the leaf metabolomes

of 457 tropical and 339 temperate species that originate from distinct geographic distributions and evolutionary histories (fig. S1, A and B), as well as that encompass most documented functional trait variation (fig. S1G, red ribbons). We thus propose that distilling the leaf metabolome into metabolic functional traits captures macroecological patterns of metabolic variation widely across the plant kingdom, enhances the explanatory power of the functional trait concept, and offers a new set of tools for the discovery of species or genotypes with trait combinations adapted to societal needs.

MATERIALS AND METHODS
Sample collection and preparation
We used two existing datasets describing the leaf metabolomes of tropical and temperate plant species (as the unit of observation). The tropical dataset originated from a random subsample of tropical leaves from the Pierre Fabre Sample Library (*69*), which is an archive of ~17,000 plant samples from in situ communities around the globe that was originally collected for drug discovery and is a registered collection of the European Union (registration code: 03-FR-2020) (*70*). Samples were oven-dried (55°C for 3 days) and extracted using overnight maceration (8 g of sample and 80 ml of ethyl acetate). Extracts were dried once (Genevac, SP Industries, Warminster, PA, USA) and eluted [30 mg of extract plus

200 mg of silica, 1-g silica solid phase extraction cartridges, and 6 ml of dichloromethane (DCM) plus 85:15 DCM:MeOH] and then dried again and resuspended in dimethyl sulfoxide at a final concentration of 5 mg ml^{-1}. The temperate dataset originated from a study of landscape-scale ecological variation in phytochemical diversity (*37*), wherein leaves were sampled from in situ grassland communities across Switzerland. Briefly, leaves were dried (40°C for 5 days), milled (Retsch TissueLyser, QIAGEN, Hilden, Germany), and extracted (20 mg of tissue and 0.5 ml of 80:19.5:0.5 MeOH:H$_2$O:H$_2$CO$_2$). Extracts were homogenized (glass beads; 30 Hz for 3 min) and centrifuged (14,000 rpm for 3 min), and the supernatant was isolated for measurement.

Metabolomics measurements

Untargeted metabolomics was performed on all samples using ultra high-performance LC-MS (UHPLC-MS). Tropical extracts were analyzed using a Waters Acquity UHPLC system coupled to a Q-Exactive Focus MS (Thermo Fisher Scientific, Bremen, Germany) with a heated electrospray ionization (HESI-II) source in positive mode. Extracts (2 μl) were separated on an Acquity C18 column (50 mm by 2.1 mm, 1.7 μm; Waters, Milford, MA, USA) at a flow rate of 600 μl min^{-1} using a binary solvent system (A: H$_2$O plus 0.1% H$_2$CO$_2$; B: acetonitrile plus 0.1% H$_2$CO$_2$) and a linear gradient elution (5 to 100% B, 7 min) plus isocratic elution (100% B, 1 min). MS data were acquired in data-dependent acquisition mode, in which MS/MS scans were performed on the three most intense ions detected during repeated full MS scans [full scans = 35,000 full width at half maximum (FWHM) at mass/charge ratio (*m/z*) 200; MS/MS scans = 17,000 FWHM; MS/MS isolation window width = 1 Da; and normalized collision energy = 15, 30, and 45 units]. The MS was calibrated using a quality control (QC) mix containing caffeine, methionine-arginine-phenylalanine-alanine-acetate, SDS, sodium taurocholate, and Ultramark 1621 dissolved in acetonitrile/methanol/H$_2$O plus 1% H$_2$CO$_2$. Temperate extracts were analyzed using a Waters Acquity UHPLC system coupled to a SYNAPT G2 MS (Waters, Milford, MA, USA) with a heated electrospray ionization (HESI-II) source in positive mode. Extracts (2.5 μl) were separated on an Acquity C18 column (50 mm by 2.1 mm, 1.7 μm; Waters, Milford, MA, USA) at a flow rate of 600 μl min^{-1} using a binary solvent system (A: H$_2$O plus 0.05% H$_2$CO$_2$; B: acetonitrile plus 0.05% H$_2$CO$_2$) and a linear gradient elution (2 to 100% B, 6 min). MS data were acquired in data-independent acquisition mode, in which all precursor ions across the full mass range (85 to 1200 Da) were fragmented to yield MS/MS spectra. The MS was calibrated using a QC mix derived from all sample extracts.

Metabolomics data preprocessing

Tropical MS data were treated using MZMine (version 2.53) (*71*). Peak detection was performed using the centroid mass detector (MS noise = 1×10^4; MS/MS noise = 0) and the ADAP chromatogram builder [minimum scan group size = 5, min. group intensity = 1×10^4, min. highest intensity = 5×10^5, and *m/z* tolerance = 12 parts per million (ppm)] (*72*). Peaks were deconvoluted using the wavelets algorithm in ADAP with a signal to noise intensity window (SN) [S/N threshold = 10; min. feature height = 5×10^5, coefficient/area threshold = 130, peak duration range = 0 to 5 min, and retention time (RT) wavelet range = 0.01 to 0.03 min]. Isotopes were detected using the isotopes peak grouper (*m/z* tolerance = 12 ppm, absolute RT tolerance = 0.01 min absolute, and

maximum charge = 2), and, where present, the most intense isotope was chosen. Peaks were filtered to retain features having an MS/MS scan and a RT of between 0.5 and 8 min, following which they were aligned using the join aligner method (*m/z* tolerance = 40 ppm, absolute RT tolerance = 0.2 min, *m/z* weight = 2, and RT weight = 1). Temperate MS data were treated using MS-DIAL (*73*). Peaks were detected with a minimum height of 300, and data were collected with an MS1 tolerance of 0.05 Da and an MS2 tolerance of 0.01 Da from 100 to 1200 Da and 0.5 to 12 min. Peaks were deconvoluted with a sigma window value of 0.5 and an MS/MS abundance cutoff of 50. Last, raw features were aligned using the QC mix with a retention time tolerance of 0.1 min. Following this, we separately clustered tropical and temperate metabolic features into sets of spectrally similar "consensus" features using molecular networks constructed on the Global Natural Products Social (GNPS; precursor and MS/MS fragment ion mass tolerances = 0.02 Da) (*74*, *75*). We retained edges with a cosine score of more than 0.7 and with more than six matched peaks, permitted edges that connected two nodes only if each node appeared in the neighbor's "top 10" most similar node list, and, where necessary, limited the total size of a cluster to the 100 highest scoring edges. Last, we converted peak height/area data to presence-absence data by summing peak heights/areas at the consensus metabolic feature level and expressing them as a one (present) or zero (absent) (tropical: *N* = 7343; temperate: *N* = 6682).

Metabolite annotation

We annotated all metabolic features with chemical information by in silico spectral matching (*76*) and taxonomically informed scoring (*77*) using a theoretically fragmented version of the LOTUS resource (*78–80*), retaining matches with a minimum score of 0.7 and at least six matched peaks. We assigned consensus metabolic features with a single consensus chemical classification, which we derived by taking the most common chemical classification across individual features contributing to a consensus feature. We filtered data to retain those features that were annotated with a metabolite identifier [i.e., SMILES notation; tropical: *N* = 3356 (46%); temperate: *N* = 2227 (33%)]. We then derived the chemistry of annotated metabolites by using SMILES identifiers to query the Chemistry Development Kit (CDK) (*27*) with the package "rcdk" (*81*). The CDK comprises 51 chemical descriptors, but we focused on 21 chemical properties that describe metabolite constitution, geometry, and topology (table S1). We also derived values for annotated metabolites from the first three dimensions of the ChemGPS-NP tool (*28*), which is an existing framework for classifying natural products that separates molecules based on size, then aromaticity, and then lipophilicity (*31*).

Classical functional traits

All further data processing, analysis and visualization were undertaken in R (*82*). For core functionality, we used the R packages "cowplot" (*83*), "data.table" (*84*), "drake" (*85*), "furrr" (*86*), "future" (*87*), "parallel" (*82*), "psych" (*88*), and "tidyverse" (*89*) (see below for packages pertaining to specific steps). Plant species names were cleaned against the Catalogue of Life and Encyclopedia of Life using the global name resolver in the package "taxize" (*90*), following which we added a taxonomic hierarchy using "taxonlookup" (*91*). We removed 14 tropical (ferns, cycads, and conifers) and nine temperate (ferns and trees) species that were phylogenetically distinct from others within their respective datasets to prevent a

small number of divergent species from overwhelming downstream analyses (fig. S2). We obtained data for classical functional traits with proven relevance to plant form and function from the open access version of the TRY database (version 5) (*92*)—namely, plant height (m), seed mass (mg), stem-specific density (mg mm^{-3}), leaf area (mm^2), specific leaf area (mm^2 mg^{-1}) and leaf carbon, and nitrogen and phosphorus concentrations (mg g^{-1}) (*7, 16, 93*). After removing clear outliers (reported error risk ≥4; values falling outside of documented limits), we calculated species-wise means and gap-filled missing trait values for all species present in the dataset (*N* = 46,905 species) using a two-step process. First, we performed a Bayesian hierarchical probabilistic matrix factorization imputation [package "BHPMF" (*94*)], which is a taxonomically constrained gap-filling approach that has been regularly applied to the TRY database (*7, 24*). The imputation was repeated 90 times, each time starting with a different combination of parameters (per-fold samples = 900 to 1000, cross-validation steps = 10 to 20, and burn-in steps = 10% data length) and filtering out extreme (>1.5 times the maximum observed value for a trait) or highly uncertain (coefficient of variation of >1) gap-filled values. We then calculated means of remaining gap-filled values across all iterations and used these to replace missing cases in original trait data. Second, we performed five iterations of multivariate imputation by chained equations [package "mice" (*95*)] on the partially gap-filled dataset and replaced remaining missing cases with the mean values from all iterations. Following this, we filtered fully gap-filled datasets for selected tropical and temperate species (see table S4 for details about missing data) using genus-level mean trait values where species-level data were not available (215 tropical species and 5 temperate species). Note that trait data coverage differed substantially between tropical and temperate species (table S4). While this undoubtedly reduced the strength of observed trait variation in tropical relative to temperate data, it is unlikely that the reduced resolution of the tropical data changed the nature of functional trait variation because artificially reducing the resolution of trait datasets to genus level did not change the resulting PC axes (fig. S3). Last, we accounted for non-normal distributions in plant height, seed mass, and leaf area data using log$_{10}$ transformations and summarized trait variation among tropical and temperate species separately using scores from the first two components of varimax-rotated PCAs performed on all traits (fig. S1).

Species distributions

We matched species names to Global Biodiversity Information Facility (GBIF) IDs using "rgbif" (*29*) and gathered all available species occurrences from the GBIF website (www.gbif.org; 29 September 2020; GBIF occurrence download: tropical DOI = 10.15468/dl.824a9z, temperate DOI = 10.15468/dl.6wappm). Working with each species individually, we filtered occurrences to include human observations and living or preserved specimens recorded from the year 1945 to present day and accompanied by either no individual counts or counts with a reasonable value (>100) (*96*). We used the packages "CoordinateCleaner" (*97*) to discard records falling in the sea, outside the stated country, or within 10 km of capital cities, GBIF headquarters, or known biodiversity institutions; records lying more than five times the interquartile range of the minimum distances to the nearest neighbor; and records containing known issues with longitude/latitude information (i.e., zeros, rounding errors, or conversion errors from other coordinate

systems) (*96, 97*). Following this, we manually removed a further 20 tropical and 4 temperate records representing single occurrences within a continent, leaving a total of 239,223 and 1,251,500 cleaned occurrences for tropical and temperate species, respectively. We plotted cleaned occurrences to visualize the geographical extents of tropical and temperate plant species (fig. S1, A and B), although it is important to emphasize that samples for metabolomics analyses were taken from within natural (not ornamental) distributions.

Leaf metabolite chemistry

We collated all unique annotated metabolites detected in any tropical or temperate species into a single dataset capturing overall variation in the chemical properties of the leaf metabolome (*N* = 4292). We tested for interrelatedness between metabolite chemical properties using pairwise Pearson correlations coupled with hierarchical clustering (Ward; Fig. 1A, top) and estimated the number of discrete clusters of chemical properties using the maximal average silhouette width (Fig. 1A, dendrogram) (*98*). We described major dimensions of leaf chemical variation using a varimax-rotated PCA performed on a subset of five representative chemical properties [one from each cluster: C hybridization ratio, H-bond acceptor count, molecular weight, a-polarity (X log *P*), and aromatic atom count]. The first three PC axes described 97% variation, which we anchored back to individual chemical properties using pairwise Pearson correlations (Fig. 1A, bottom). We tested whether the subset PCA was functionally equivalent to a full PCA performed on all 21 properties using two approaches. First, we performed pairwise Pearson correlations between subset versus full PC scores to determine whether the two PCAs separated metabolic features similarly on the same dimensions of chemical variation (Fig. 1B, left; see main text for statistics). Second, we used a Procrustes rotation test [package "vegan" (*99*); 999 permutations] to determine whether maximally rotated configurations of full versus subset PCAs were significantly (i.e., nonrandomly) associated (m_{12} = 0.04, r^2 = 0.98, *P* < 0.001). We also confirmed that the similarity among subset and full PCAs was not sensitive to which chemical properties were selected from each cluster. This was achieved by repeating Procrustes rotation tests and PC score correlations for PCAs performed on every combination of five chemical properties that covered all five clusters and plotting distributions of resulting coefficients (fig. S5). We performed pairwise Pearson correlations between subset PC scores and values from the first three dimensions of the ChemGPS-NP framework (*28*) to test for correspondence between the subset PCA and existing descriptors of biologically relevant chemical space (Fig. 1B, right; see main text for statistics). We visualized chemical variation among metabolites using biplots of PC scores from the subset PCA, grouped by biochemical class (Fig. 2). We also illustrated the range of chemical structures captured by the five selected chemical properties by using the PubChem sketcher (https://pubchem.ncbi.nlm.nih.gov//edit3/index.html) to predict chemical graphs of metabolites with the lowest and highest values for each property (Fig. 1C).

Plant-level metabolic functional traits

We selected C hybridization ratio, H-bond acceptor count, molecular weight, X log *P*, and aromatic atom count to convert into five plant-level metabolic functional traits. For each species (i.e., unit of observation), we calculated the mean value of a chemical property

considering all the metabolites present. In doing so, we generated datasets akin those capturing classical functional trait variation among tropical or temperate species but instead quantifying variation in the mean C hybridization ratio, H-bond acceptor count, molecular weight, a-polarity, and aromatic atom count of the annotated leaf metabolome among tropical or temperate species. We considered tropical and temperate datasets as separate entities for data analysis of trait variation, for four reasons. First, while in silico data treatment was identical for aligned MS data, sample collection and LC-MS measurements were not (see above), making it challenging to combine data on species-level metabolic variation without inflating differences between them. Second, unlike genomics and proteomics analyses, where only one type of molecule is isolated, untargeted metabolomics aims to simultaneously isolate thousands of molecules with a wide range of chemical properties (*17*, *100*). Technical decisions made during sample extraction and analysis discriminate for or against certain types of metabolites, which filters the resulting view of the metabolome. There is still no single accepted method for extracting samples or acquiring metabolomics data (*101*), so we chose here to use two datasets generated using different approaches to confirm that findings are not sensitive to the methods chosen. Third, classical functional trait and species occurrence data were sparser for tropical than temperate species (see the "*Classical functional traits*" section above), so we chose to keep datasets separate rather than artificially reduce the resolution of the temperate dataset. Last, maintaining a separation between the two datasets increased our interpretive power by providing the opportunity to characterize coupling between the metabolome and functional traits that held true not only within but also among tropical and temperate plant species.

Plant multidimensional space

Following (*7*), we explored constraints on the multidimensional space occupied by metabolic and classical functional traits separately for tropical and temperate species using an observed versus simulated hypervolume approach. Briefly, we calculated an *n*-dimensional convex hull volume [package "geometry"; (*102*)] representing the observed multidimensional trait space occupied by species, where *n* is the number of metabolic ($n = 5$) or classical ($n = 8$) functional traits. We compared the observed hypervolume to the same four simulated null hypervolumes used in (*7*), which tested the following null hypotheses: (i) Traits vary independently and have uniform distributions (i.e., each trait is a unique axis and no selection against extreme trait values); (ii) traits vary independently but have normal distributions (i.e., each trait is a unique axis and selection against extreme trait values); (iii) traits vary independently but distributions are as observed (i.e., each trait is a unique axis but no assumptions about selection on values); and (iv) traits covary as but have normal distributions (i.e., no assumptions about trait interdependence but selection against extreme values). We simulated null hypervolumes 999 times and used permutation tests [package "ade4"; (*103*)] to test for differences between the overall size of observed versus simulated hypervolumes and presented percent changes between observed and mean simulated hypervolume sizes (where positive = observed > simulated; table S2). We also calculated the "lumpiness" (i.e., aggregation around certain values) of species in multidimensional space by splitting *n*-dimensional space into 10 bins (i.e., 10^n cells), counting the number of species present in each cell, and calculating the minimum number of cells required to capture 10% of species (*7*). Again, we tested for differences between observed and simulated null hypervolumes using permutation tests and presented percent changes between observed and mean simulated hypervolumes (table S3).

Axes of leaf metabolic specialization

We characterized major axes of leaf metabolic specialization separately among tropical or temperate species using varimax-rotated PCAs performed on metabolic functional traits. The first two PCs explained 92% variation for both tropical and temperate species, respectively, which we displayed using biplots of PC scores (Fig. 3). We explored whether axes of leaf metabolic specialization (Fig. 3) were orthogonal to, or colinear with, axes of classical functional trait variation (Fig. 4) by doing varimax-rotated PCAs containing metabolic plus classical functional traits, again separately among tropical and temperate species. Four PCs were needed to explain 66 and 67% variation among tropical and temperate species, respectively, which we displayed in two ways. First, we created a biplot matrix of all pairwise combinations of the first four axes (fig. S4) to visualize the lack of overlap between metabolic and classical functional traits. Second, we performed Pearson correlations between PC scores and individual functional traits (Fig. 4, bottom panels). Last, we examined interdependence within and among metabolic and classical functional traits by performing pairwise Pearson correlations (Fig. 4, top panels), clustering correlation coefficients (hierarchical clustering; Ward), and identifying discrete clusters of traits using the maximal silhouette width (Fig. 4, dendrograms) (*98*). Dendrograms are organized on the basis of the outcome of a tanglegram that tested for correspondence between the clustering of tropical and temperate traits (Fig. 4, central connectors).

Supplementary Materials

This PDF file includes:
Figs. S1 to S5
Tables S1 to S4
References

REFERENCES AND NOTES

1. S. Wang, S. Alseekh, A. R. Fernie, J. Luo, The structure and function of major plant metabolite modifications. *Mol. Plant* **12**, 899–919 (2019).
2. G. Faccio, Plant complexity and cosmetic innovation. *iScience* **23**, 101358 (2020).
3. H. J. Klee, D. M. Tieman, The genetics of fruit flavour preferences. *Nat. Rev. Genet.* **19**, 347–356 (2018).
4. B. M. Schmidt, D. M. Ribnicky, P. E. Lipsky, I. Raskin, Revisiting the ancient concept of botanical therapeutics. *Nat. Chem. Biol.* **3**, 360–366 (2007).
5. J. Sardans, A. Gargallo-Garriga, O. Urban, K. Klem, T. W. N. Walker, P. Holub, I. A. Janssens, J. Peñuelas, Ecometabolomics for a better understanding of plant responses and acclimation to abiotic factors linked to global change. *Metabolites* **10**, 239 (2020).
6. J. Sardans, A. Gargallo-Garriga, O. Urban, K. Klem, P. Holub, I. A. Janssens, T. W. N. Walker, A. Pesqueda, J. Peñuelas, Ecometabolomics of plant–herbivore and plant–fungi interactions: A synthesis study. *Ecosphere* **12**, e03736 (2021).
7. S. Díaz, J. Kattge, J. H. C. Cornelissen, I. J. Wright, S. Lavorel, S. Dray, B. Reu, M. Kleyer, C. Wirth, I. C. Prentice, E. Garnier, G. Bönisch, M. Westoby, H. Poorter, P. B. Reich, A. T. Moles, J. Dickie, A. N. Gillison, A. E. Zanne, J. Chave, S. J. Wright, S. N. Sheremet'ev, H. Jactel, C. Baraloto, S. Cerabolini, S. Pierce, B. Shipley, D. Kirkup, F. Casanoves, J. S. Joswig, A. Günther, V. Falczuk, N. Rüger, M. D. Mahecha, L. D. Gorné, The global spectrum of plant form and function. *Nature* **529**, 167–171 (2016).
8. M. Wink, Evolution of secondary metabolites from an ecological and molecular phylogenetic perspective. *Phytochemistry* **64**, 3–19 (2003).
9. M. C. Schuman, N. M. van Dam, F. Beran, W. S. Harpole, How does plant chemical diversity contribute to biodiversity at higher trophic levels? *Curr. Opin. Insect Sci.* **14**, 46–55 (2016).

10. A. E. Zanne, D. C. Tank, W. K. Cornwell, J. M. Eastman, S. A. Smith, R. G. FitzJohn, D. J. McGlinn, B. C. O'Meara, A. T. Moles, P. B. Reich, D. L. Royer, D. E. Soltis, P. F. Stevens, M. Westoby, I. J. Wright, L. Aarssen, R. I. Bertin, A. Calaminus, R. Govaerts, F. Hemmings, M. R. Leishman, J. Oleksyn, P. S. Soltis, N. G. Swenson, L. Warman, J. M. Beaulieu, Three keys to the radiation of angiosperms into freezing environments. *Nature* **506**, 89–92 (2014).

11. G. E. Hutchinson, *The Ecological Theater and the Evolutionary Play* (Yale Univ. Press, 1965).

12. P. Feeny, in *Biochemical Interaction Between Plants and Insects*, J. W. Wallace, R. L. Mansell, Eds. (Springer, 1976), pp. 1–40.

13. P. D. Coley, J. P. Bryant, F. S. Chapin, Resource availability and plant antiherbivore defense. *Science* **230**, 895–899 (1985).

14. P. Capdevila, T. W. N. Walker, F. Schrodt, R. C. Rodriguez Caro, R. Salguero-Gomez, Global patterns of plant form and function are strongly determined by evolutionary relationships. *bioRxiv*, 2023.01.13.523963 (2023).

15. J. P. Grime, K. Thompson, R. Hunt, J. G. Hodgson, J. H. C. Cornelissen, I. H. Rorison, G. A. F. Hendry, T. W. Ashenden, A. P. Askew, S. R. Band, R. E. Booth, C. C. Bossard, B. D. Campbell, J. E. L. Cooper, A. W. Davison, P. L. Gupta, W. Hall, D. W. Hand, M. A. Hannah, S. H. Hillier, D. J. Hodkinson, A. Jalili, Z. Liu, J. M. L. Mackey, N. Matthews, M. A. Mowforth, A. M. Neal, R. J. Reader, K. Reiling, W. Ross-Fraser, R. E. Spencer, F. Sutton, D. E. Tasker, P. C. Thorpe, J. Whitehouse, Integrated screening validates primary axes of specialisation in plants. *Oikos* **79**, 259 (1997).

16. I. J. Wright, P. B. Reich, M. Westoby, D. D. Ackerly, Z. Baruch, F. Bongers, J. Cavender-Bares, T. Chapin, J. H. C. Cornelissen, M. Diemer, J. Flexas, E. Garnier, P. K. Groom, J. Gulias, K. Hikosaka, B. B. Lamont, T. Lee, W. Lee, C. Lusk, J. J. Midgley, M.-L. Navas, U. Niinemets, J. Oleksyn, N. Osada, H. Poorter, P. Poot, L. Prior, V. I. Pyankov, C. Roumet, S. C. Thomas, M. G. Tjoelker, E. J. Veneklaas, R. Villar, The worldwide leaf economics spectrum. *Nature* **428**, 821–827 (2004).

17. T. W. N. Walker, J. M. Alexander, P. Allard, O. Baines, V. Baldy, R. D. Bardgett, P. Capdevila, P. D. Coley, B. David, E. Defossez, M. Endara, M. Ernst, C. Fernandez, D. Forrister, A. Gargallo-Garriga, V. E. J. Jassey, S. Marr, S. Neumann, L. Pellissier, J. Peñuelas, K. Peters, S. Rasmann, U. Roessner, J. Sardans, F. Schrodt, M. C. Schuman, A. Soule, H. Uthe, W. Weckwerth, J. Wolfender, N. M. Dam, R. Salguero-Gómez, Functional Traits 2.0: The power of the metabolome for ecology. *J. Ecol.* **110**, 4–20 (2022).

18. B. E. Sedio, J. D. Parker, S. M. McMahon, S. J. Wright, Comparative foliar metabolomics of a tropical and a temperate forest community. *Ecology* **99**, 2647–2653 (2018).

19. D. L. Forrister, M.-J. Endara, G. C. Younkin, P. D. Coley, T. A. Kursar, Herbivores as drivers of negative density dependence in tropical forest saplings. *Science* **363**, 1213–1216 (2019).

20. A. Gargallo-Garriga, J. Sardans, V. Granda, J. Llusià, G. Peguero, D. Asensio, R. Ogaya, I. Urbina, L. Van Langenhove, L. T. Verryckt, J. Chave, E. A. Courtois, C. Stahl, O. Grau, K. Klem, O. Urban, I. A. Janssens, J. Peñuelas, Different "metabolomic niches" of the highly diverse tree species of the French Guiana rainforests. *Sci. Rep.* **10**, 6937 (2020).

21. C. Scherling, C. Roscher, P. Giavalisco, E.-D. Schulze, W. Weckwerth, Metabolomics unravel contrasting effects of biodiversity on the performance of individual plant species. *PLOS ONE* **5**, e12569 (2010).

22. C. Violle, M.-L. Navas, D. Vile, E. Kazakou, C. Fortunel, I. Hummel, E. Garnier, Let the concept of trait be functional! *Oikos* **116**, 882–892 (2007).

23. P. B. Adler, R. Salguero-Gómez, A. Compagnoni, J. S. Hsu, J. Ray-Mukherjee, C. Mbeau-Ache, M. Franco, Functional traits explain variation in plant life history strategies. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 740–745 (2014).

24. H. Bruelheide, J. Dengler, O. Purschke, J. Lenoir, B. Jiménez-Alfaro, S. M. Hennekens, Z. Botta-Dukát, M. Chytrý, R. Field, F. Jansen, J. Kattge, V. D. Pillar, F. Schrodt, M. D. Mahecha, R. K. Peet, B. Sandel, P. van Bodegom, J. Altman, E. Alvarez-Dávila, M. A. S. A. Khan, F. Attorre, I. Aubin, C. Baraloto, J. G. Barroso, M. Bauters, E. Bergmeier, I. Biurrun, A. D. Bjorkman, B. Blonder, A. Čarni, L. Cayuela, T. Černý, J. H. C. Cornelissen, D. Craven, M. Dainese, G. Derroire, M. De Sanctis, S. Díaz, J. Doležal, W. Farfan-Rios, T. R. Feldpausch, N. J. Fenton, E. Garnier, G. R. Guerin, A. G. Gutiérrez, S. Haider, T. Hattab, G. Henry, B. Hérault, P. Higuchi, N. Hölzel, J. Homeier, A. Jentsch, N. Jürgens, Z. Kącki, D. N. Karger, M. Kessler, M. Kleyer, I. Knollová, A. Y. Korolyuk, I. Kühn, D. C. Laughlin, F. Lens, J. Loos, F. Louault, M. I. Lyubenova, V. Malhi, C. Marcenò, M. Mencuccini, J. V. Müller, J. Munzinger, I. H. Myers-Smith, D. A. Neill, Ü. Niinemets, K. H. Orwin, W. A. Ozinga, J. Penuelas, A. Pérez-Haase, P. Petřík, O. L. Phillips, M. Pärtel, P. B. Reich, C. Römermann, A. V. Rodrigues, F. M. Sabatini, J. Sardans, M. Schmidt, G. Seidler, J. E. Silva Espejo, M. Silveira, A. Smyth, M. Sporbert, J.-C. Svenning, Z. Tang, R. Thomas, I. Tsiripidis, K. Vassilev, C. Violle, U. Jandt, Global trait-environment relationships of plant communities. *Nat. Ecol. Evol.* **2**, 1906–1917 (2018).

25. G. B. De Deyn, J. H. C. Cornelissen, R. D. Bardgett, Plant functional traits and soil carbon sequestration in contrasting biomes. *Ecol. Lett.* **11**, 516–531 (2008).

26. S. Lavorel, J. Storkey, R. D. Bardgett, F. de Bello, M. P. Berg, X. Le Roux, M. Moretti, C. Mulder, R. J. Pakeman, S. Díaz, R. Harrington, A novel framework for linking functional diversity of plants with other trophic levels for the quantification of ecosystem services. *J. Veg. Sci.* **24**, 942–948 (2013).

27. E. L. Willighagen, J. W. Mayfield, J. Alvarsson, A. Berg, L. Carlsson, N. Jeliazkova, S. Kuhn, T. Pluskal, M. Rojas-Chertó, O. Spjuth, G. Torrance, C. T. Evelo, R. Guha, C. Steinbeck, The Chemistry Development Kit (CDK) v2.0: Atom typing, depiction, molecular formulas, and substructure searching. *J. Chem.* **9**, 33 (2017).

28. J. Rosén, A. Lövgren, T. Kogej, S. Muresan, J. Gottfries, A. Backlund, ChemGPS-NP(Web): Chemical space navigation online. *J. Comput. Aided Mol. Des.* **23**, 253–259 (2009).

29. C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **46**, 3–26 (2001).

30. F. Lovering, J. Bikker, C. Humblet, Escape from flatland: Increasing saturation as an approach to improving clinical success. *J. Med. Chem.* **52**, 6752–6756 (2009).

31. J. Larsson, J. Gottfries, S. Muresan, A. Backlund, ChemGPS-NP: Tuned for navigation in biologically relevant chemical space. *J. Nat. Prod.* **70**, 789–794 (2007).

32. J. L. Funk, J. E. Larson, G. M. Ames, B. J. Butterfield, J. Cavender-Bares, J. Firn, D. C. Laughlin, A. E. Sutton-Grier, L. Williams, J. Wright, Revisiting the Holy Grail: Using plant functional traits to understand ecological processes. *Biol. Rev. Camb. Philos. Soc.* **92**, 1156–1173 (2017).

33. J. Firn, J. M. McGree, E. Harvey, H. Flores-Moreno, M. Schütz, Y. M. Buckley, E. T. Borer, E. W. Seabloom, K. J. La Pierre, A. M. MacDougall, S. M. Prober, C. J. Stevens, L. L. Sullivan, E. Porter, E. Ladouceur, C. Allen, K. H. Moromizato, J. W. Morgan, W. S. Harpole, Y. Hautier, N. Eisenhauer, J. P. Wright, P. B. Adler, C. A. Arnillas, J. D. Bakker, L. Biederman, A. A. D. Broadbent, C. S. Brown, M. N. Bugalho, M. C. Caldeira, E. E. Cleland, A. Ebeling, P. A. Fay, N. Hagenah, A. R. Kleinhesselink, R. Mitchell, J. L. Moore, C. Nogueira, P. L. Peri, C. Roscher, M. D. Smith, P. D. Wragg, A. C. Risch, Leaf nutrients, not specific leaf area, are consistent indicators of elevated nutrient inputs. *Nat. Ecol. Evol.* **3**, 400–406 (2019).

34. F. van der Plas, T. Schröder-Georgi, A. Weigelt, K. Barry, S. Meyer, A. Alzate, R. L. Barnard, N. Buchmann, H. de Kroon, A. Ebeling, N. Eisenhauer, C. Engels, M. Fischer, G. Gleixner, A. Hildebrandt, E. Koller-France, S. Leimer, A. Milcu, L. Mommer, P. A. Niklaus, Y. Oelmann, C. Roscher, C. Scherber, M. Scherer-Lorenzen, S. Scheu, B. Schmid, E.-D. Schulze, V. Temperton, T. Tscharntke, W. Voigt, W. Weisser, W. Wilcke, C. Wirth, Plant traits alone are poor predictors of ecosystem properties and long-term ecosystem functioning. *Nat. Ecol. Evol.* **4**, 1602–1611 (2020).

35. T. W. N. Walker, W. Weckwerth, L. Bragazza, L. Fragner, B. G. Forde, N. J. Ostle, C. Signarbieux, X. Sun, S. E. Ward, R. D. Bardgett, Plastic and genetic responses of a common sedge to warming have contrasting effects on carbon cycle processes. *Ecol. Lett.* **22**, 159–169 (2019).

36. P.-M. Allard, A. Gaudry, L.-M. Quirós-Guerrero, A. Rutz, M. Dounoue-Kubo, T. W. N. Walker, E. Defossez, C. Long, A. Grondin, B. David, J.-L. Wolfender, Open and reusable annotated mass spectrometry dataset of a chemodiverse collection of 1,600 plant extracts. *Gigascience* **12**, giac124 (2022).

37. E. Defossez, C. Pitteloud, P. Descombes, G. Glauser, P.-M. Allard, T. W. N. Walker, P. Fernandez-Conradi, J.-L. Wolfender, L. Pellissier, S. Rasmann, Spatial and evolutionary predictability of phytochemical diversity. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2013344118 (2021).

38. R. Mannhold, G. I. Poda, C. Ostermann, I. V. Tetko, Calculation of molecular lipophilicity: State-of-the-art and comparison of log P methods on more than 96,000 compounds. *J. Pharm. Sci.* **98**, 861–893 (2009).

39. H. Wiener, Structural determination of paraffin boiling points. *J. Am. Chem. Soc.* **69**, 17–20 (1947).

40. V. Sharma, R. Goswami, A. K. Madan, Eccentric connectivity index: A novel highly discriminating topological descriptor for structure—property and structure—activity studies. *J. Chem. Inf. Comput. Sci.* **37**, 273–282 (1997).

41. Y. Yang, H. Chen, I. Nilsson, S. Muresan, O. Engkvist, Investigation of the relationship between topology and selectivity for druglike molecules. *J. Med. Chem.* **53**, 7709–7714 (2010).

42. B. Kuhn, P. Mohr, M. Stahl, Intramolecular hydrogen bonding in medicinal chemistry. *J. Med. Chem.* **53**, 2601–2611 (2010).

43. Th. Zeegers-Huyskens, P. Huyskens, in *Intermolecular Forces*, P. L. Huyskens, W. A. P. Luck, T. Zeegers-Huyskens, Eds. (Springer Berlin Heidelberg, Berlin, Heidelberg, 1991), pp. 1–30.

44. P. B. Dervan, Molecular recognition of DNA by small molecules. *Bioorg. Med. Chem.* **9**, 2215–2235 (2001).

45. G. Padroni, N. N. Patwardhan, M. Schapira, A. E. Hargrove, Systematic analysis of the interactions driving small molecule-RNA recognition. *RSC Med. Chem.* **11**, 802–813 (2020).

46. M. Gao, J. Skolnick, A comprehensive survey of small-molecule binding pockets in proteins. *PLoS Comput. Biol.* **9**, e1003302 (2013).

47. R. Wang, Y. Fu, L. Lai, A new atom-additive method for calculating partition coefficients. *J. Chem. Inf. Comput. Sci.* **37**, 615–621 (1997).

48. A. K. Ghose, G. M. Crippen, Atomic physicochemical parameters for three-dimensional structure-directed quantitative structure-activity relationships I. Partition coefficients as a measure of hydrophobicity. *J. Comput. Chem.* **7**, 565–577 (1986).

49. P. A. Clemons, N. E. Bodycombe, H. A. Carrinski, J. A. Wilson, A. F. Shamji, B. K. Wagner, A. N. Koehler, S. L. Schreiber, Small molecules of different origins have distinct distributions of structural complexity that correlate with protein-binding profiles. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 18787–18792 (2010).

50. J. N. Israelachvili, The nature of van der waals forces. *Contemp. Phys.* **15**, 159–178 (1974).

51. D. Van Vranken, G. A. Weiss, *Introduction to Bioorganic Chemistry and Chemical Biology* (Garland Science, ed. 1, 2012).

52. E. Pichersky, R. A. Raguso, Why do plants produce so many terpenoid compounds? *New Phytol.* **220**, 692–702 (2018).

53. D. L. Forrister, M.-J. Endara, A. J. Soule, G. C. Younkin, A. G. Mills, J. Lokvam, K. G. Dexter, R. T. Pennington, C. A. Kidner, J. A. Nicholls, O. Loiseau, T. A. Kursar, P. D. Coley, Diversity and divergence: Evolution of secondary metabolism in the tropical tree genus *Inga*. *New Phytol.* **237**, 631–642 (2023).

54. R. M. LoPachin, D. S. Barber, T. Gavin, Molecular mechanisms of the conjugated α,β-unsaturated carbonyl derivatives: Relevance to neurotoxicity and neurodegenerative diseases. *Toxicol. Sci.* **104**, 235–249 (2008).

55. J. Widomska, W. K. Subczynski, Mechanisms enhancing the protective functions of macular xanthophylls in the retina during oxidative stress. *Exp. Eye Res.* **178**, 238–246 (2019).

56. M. H. Tran, D.-P. Phan, E. Y. Lee, Review on lignin modifications toward natural UV protection ingredient for lignin-based sunscreens. *Green Chem.* **23**, 4633–4646 (2021).

57. N.-H. Tan, J. Zhou, Plant cyclopeptides. *Chem. Rev.* **106**, 840–895 (2006).

58. G.-H. Lim, R. Singhal, A. Kachroo, P. Kachroo, Fatty acid- and lipid-mediated signaling in plant defense. *Annu. Rev. Phytopathol.* **55**, 505–536 (2017).

59. T. H. Yeats, J. K. C. Rose, The formation and function of plant cuticles. *Plant Physiol.* **163**, 5–20 (2013).

60. Q. Liu, L. Luo, L. Zheng, Lignins: Biosynthesis and biological functions in plants. *Int. J. Mol. Sci.* **19**, 335 (2018).

61. P. E. Staswick, Storage proteins of vegetative plant tissues. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **45**, 303–322 (1994).

62. T. O. Sandberg, C. Weinberger, J.-H. Smått, Molecular dynamics on wood-derived lignans analyzed by intermolecular network theory. *Molecules* **23**, 1990 (2018).

63. M. Pollard, F. Beisson, Y. Li, J. B. Ohlrogge, Building lipid barriers: Biosynthesis of cutin and suberin. *Trends Plant Sci.* **13**, 236–246 (2008).

64. R. Salguero-Gómez, O. R. Jones, E. Jongejans, S. P. Blomberg, D. J. Hodgson, C. Mbeau-Ache, P. A. Zuidema, H. de Kroon, Y. M. Buckley, Fast-slow continuum and reproductive strategies structure plant life-history variation worldwide. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 230–235 (2016).

65. P. D. Coley, in *Plant-Animal Interactions: Evolutionary Ecology in Tropical and Temperate Regions*, P. W. Price, T. M. Lewison, G. W. Fernandes, W. W. Benson, Eds. (John Wiley and Sons, 1991), pp. 54–69.

66. R. Sulpice, P. C. McKeown, Moving toward a comprehensive map of central plant metabolism. *Annu. Rev. Plant Biol.* **66**, 187–210 (2015).

67. S. B. Powles, Photoinhibition of photosynthesis induced by visible light. *Annu. Rev. Plant Physiol.* **35**, 15–44 (1984).

68. J. Berry, O. Bjorkman, Photosynthetic response and adaptation to temperature in higher plants. *Annu. Rev. Plant Physiol.* **31**, 491–543 (1980).

69. B. David, F. Aussiel, in *Handbook of Chemical and Biological Plant Analytical Methods*, K. Hostettmann, Ed. (Wiley, Chichester, West Sussex, ed. 1, 2014), p. 1176.

70. B. David, F. Ausseil, in *Encyclopedia of analytical chemistry: Applications, theory and instrumentation*, R. A. Meyers, Ed. (John Wiley & Sons, 2006), pp. 1–24.

71. T. Pluskal, S. Castillo, A. Villar-Briones, M. Oresic, MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **11**, 395 (2010).

72. O. D. Myers, S. J. Sumner, S. Li, S. Barnes, X. Du, One step forward for reducing false positive and false negative compound identifications from mass spectrometry metabolomics data: New algorithms for constructing extracted ion chromatograms and detecting chromatographic peaks. *Anal. Chem.* **89**, 8696–8703 (2017).

73. H. Tsugawa, T. Cajka, T. Kind, Y. Ma, B. Higgins, K. Ikeda, M. Kanazawa, J. VanderGheynst, O. Fiehn, M. Arita, MS-DIAL: Data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nat. Methods* **12**, 523–526 (2015).

74. A. M. Frank, N. Bandeira, Z. Shen, S. Tanner, S. P. Briggs, R. D. Smith, P. A. Pevzner, Clustering millions of tandem mass spectra. *J. Proteome Res.* **7**, 113–122 (2008).

75. M. Wang, J. J. Carver, V. V. Phelan, L. M. Sanchez, N. Garg, Y. Peng, D. D. Nguyen, J. Watrous, C. A. Kapono, T. Luzzatto-Knaan, C. Porto, A. Bouslimani, A. V. Melnik, M. J. Meehan, W.-T. Liu, M. Crüsemann, P. D. Boudreau, E. Esquenazi, M. Sandoval-Calderón, R. D. Kersten, L. A. Pace, R. A. Quinn, K. R. Duncan, C.-C. Hsu, D. J. Floros, R. G. Gavilan, K. Kleigrewe, T. Northen, R. J. Dutton, D. Parrot, E. E. Carlson, B. Aigle, C. F. Michelsen, L. Jelsbak, C. Sohlenkamp, P. Pevzner, A. Edlund, J. McLean, J. Piel, B. T. Murphy, L. Gerwick, C.-C. Liaw, Y.-L. Yang, H.-U. Humpf, M. Maansson, R. A. Keyzers, A. C. Sims, A. R. Johnson, A. M. Sidebottom, B. E. Sedio, A. Klitgaard, C. B. Larson, C. A. Boya, D. Torres-Mendoza, D. J. Gonzalez, D. B. Silva, L. M. Marques, D. P. Demarque, E. Pociute, E. C. O'Neill, E. Briand, E. J. N. Helfrich, E. A. Granatosky, E. Glukhov, F. Ryffel, H. Houson, H. Mohimani, J. J. Kharbush, Y. Zeng, J. A. Vorholt, K. L. Kurita, P. Charusanti, K. L. McPhail, K. F. Nielsen, L. Vuong, M. Elfeki, M. F. Traxler, N. Engene, N. Koyama, O. B. Vining, R. Baric, R. R. Silva, S. J. Mascuch, S. Tomasi, S. Jenkins, V. Macherla, T. Hoffman, V. Agarwal, P. G. Williams, J. Dai, R. Neupane, J. Gurr, A. M. C. Rodríguez, A. Lamsa, C. Zhang, K. Dorrestein, B. M. Duggan, J. Almaliti, N. Bandeira, Sharing and community curation of mass spectrometry data with GNPS. *Nat. Biotechnol.* **34**, 828–837 (2016).

76. P.-M. Allard, T. Péresse, J. Bisson, K. Gindro, L. Marcourt, V. C. Pham, F. Roussi, M. Litaudon, J.-L. Wolfender, Integration of molecular networking and in-silico MS/MS fragmentation for natural products dereplication. *Anal. Chem.* **88**, 3317–3323 (2016).

77. A. Rutz, M. Dounoue-Kubo, S. Ollivier, J. Bisson, M. Bagheri, T. Saesong, S. N. Ebrahimi, K. Ingkaninan, J.-L. Wolfender, P.-M. Allard, Taxonomically informed scoring enhances confidence in natural products annotation. *Front. Plant Sci.* **10**, 1329 (2019).

78. A. Rutz, J. Bisson, P.-M. Allard, The LOTUS initiative for open natural products research: Frozen dataset union wikidata (with metadata) (Zenodo, 2021); https://doi.org/10.5281/zenodo.5794107.

79. P.-M. Allard, J. Bisson, A. Rutz, ISDB: In silico spectral databases of natural products (Zenodo, 2021), https://doi.org/10.5281/zenodo.5607264.

80. A. Rutz, M. Sorokina, J. Galgonek, D. Mietchen, E. Willighagen, A. Gaudry, J. G. Graham, R. Stephan, R. Page, J. Vondrášek, C. Steinbeck, G. F. Pauli, J.-L. Wolfender, J. Bisson, P.-M. Allard, The LOTUS initiative for open knowledge management in natural products research. *eLife* **11**, e70780 (2022).

81. R. Guha, Chemical informatics functionality in R. *J. Stat. Softw.* **18**, 1–16 (2007).

82. R Core Team, *R: A language and environment for statistical computing* (R Foundation for Statistical Computing, 2014).

83. C. O. Wilke, *cowplot: Streamlined Plot Theme and Plot Annotations for "ggplot2"* (R Project, 2019).

84. M. Dowle, A. Srinivasan, *data.table: Extension of `data.frame`* (R Project, 2021).

85. W. Michael Landau, The drake R package: a pipeline toolkit for reproducibility and high-performance computing. *JOSS* **3**, 550 (2018).

86. D. Vaughan, M. Dancho, *furrr: Apply Mapping Functions in Parallel using Futures* (R Project, 2021).

87. H. Bengtsson, A unifying framework for parallel and distributed processing in R using futures. arXiv:2008.00553 [cs.DC] (2021).

88. W. R. Revelle, *psych: Procedures for Personality and Psychological Research* (Northwestern University, 2019).

89. H. Wickham, M. Averick, J. Bryan, W. Chang, L. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. Pedersen, E. Miller, S. Bache, K. Müller, J. Ooms, D. Robinson, D. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, H. Yutani, Welcome to the tidyverse. *JOSS* **4**, 1686 (2019).

90. S. A. Chamberlain, E. Szöcs, taxize: Taxonomic search and retrieval in R. *F1000Res.* **2**, 191 (2013).

91. W. Cornwell, R. FitzJohn, M. Pennell, *taxonlookup: A dynamically-updating versioned taxonomic resource for vascular plants* (R Project, 2016).

92. J. Kattge, G. Bönisch, S. Díaz, S. Lavorel, I. C. Prentice, P. Leadley, S. Tautenhahn, G. D. A. Werner, T. Aakala, M. Abedi, A. T. R. Acosta, G. C. Adamidis, K. Adamson, M. Aiba, C. H. Albert, J. M. Alcántara, C. Carolina Alcázar, I. Aleixo, H. Ali, B. Amiaud, C. Ammer, M. M. Amoroso, M. Anand, C. Anderson, N. Anten, J. Antos, D. M. G. Apgaua, T.-L. Ashman, D. H. Asmara, G. P. Asner, M. Aspinwall, O. Atkin, I. Aubin, L. Baastrup-Spohr, K. Bahalkeh, M. Bahn, T. Baker, W. J. Baker, J. P. Bakker, D. Baldocchi, J. Baltzer, A. Banerjee, A. Baranger, J. Barlow, D. R. Barneche, Z. Baruch, D. Bastianelli, J. Battles, W. Bauerle, M. Bauters, E. Bazzato, M. Beckmann, H. Beeckman, C. Beierkuhnlein, R. Bekker, G. Belfry, M. Belluau, M. Beloiu, R. Benavides, L. Benomar, M. L. Berdugo-Lattke, E. Berenguer, R. Bergamin, J. Bergmann, M. B. Carlucci, L. Berner, M. Bernhardt-Römermann, C. Bigler, A. D. Bjorkman, C. Blackman, C. Blanco, B. Blonder, D. Blumenthal, K. T. Bocanegra-González, P. Boeckx, S. Bohlman, K. Böhning-Gaese, L. Boisvert-Marsh, W. Bond, B. Bond-Lamberty, A. Boom, C. C. F. Boonman, K. Bordin, E. H. Boughton, V. Boukili, D. M. J. S. Bowman, S. Bravo, M. R. Brendel, M. R. Broadley, K. A. Brown, H. Bruelheide, F. Brumnich, H. H. Bruun, D. Bruy, S. W. Buchanan, S. F. Bucher, N. Buchmann, R. Buitenwerf, D. E. Bunker, J. Bürger, S. Burrascano, D. F. R. P. Burslem, B. J. Butterfield, C. Byun, M. Marques, M. C. Scalon, M. Caccianiga, M. Cadotte, M. Cailleret, J. Camac, J. J. Camarero, C. Campany, G. Campetella, J. A. Campos, L. Cano-Arboleda, R. Canullo, M. Carbognani, F. Carvalho, F. Casanoves, B. Castagneyrol, J. A. Catford, J. Cavender-Bares, B. E. L. Cerabolini, M. Cervellini, E. Chacón-Madrigal, K. Chapin, F. S. Chapin, S. Chelli, S.-C. Chen, A. Chen, P. Cherubini, F. Chianucci, B. Choat, K.-S. Chung,

M. Chytrý, D. Ciccarelli, L. Coll, C. G. Collins, L. Conti, D. Coomes, J. H. C. Cornelissen, W. K. Cornwell, P. Corona, M. Coyea, J. Craine, D. Craven, J. P. G. M. Cromsigt, A. Csecserits, K. Cufar, M. Cuntz, A. C. da Silva, K. M. Dahlin, M. Dainese, I. Dalke, M. D. Fratte, A. T. Dang-Le, J. Danihelka, M. Dannoura, S. Dawson, A. J. de Beer, A. De Frutos, J. R. De Long, B. Dechant, S. Delagrange, N. Delpierre, G. Derroire, A. S. Dias, M. H. Diaz-Toribio, P. G. Dimitrako- poulos, M. Dobrowolski, D. Doktor, P. Dřevojan, N. Dong, J. Dransfield, S. Dressler, L. Duarte, E. Ducouret, S. Dullinger, W. Durka, R. Duursma, O. Dymova, A. E.-Vojtkó, R. L. Eckstein, H. Ejtehadi, J. Elser, T. Emilio, K. Engemann, M. B. Erfanian, A. Erfmeier, A. Es- quivel-Muelbert, G. Esser, M. Estiarte, T. F. Domingues, W. F. Fagan, J. Fagúndez, D. S. Falster, Y. Fan, J. Fang, E. Farris, F. Fazlioglu, Y. Feng, F. Fernandez-Mendez, A. Ferrara, J. Ferreira, A. Fidelis, B. Finegan, J. Firn, T. J. Flowers, D. F. B. Flynn, V. Fontana, E. Forey, C. Forgiarini, L. François, D. Frank, C. Frenette-Dussault, G. T. Freschet, E. L. Fry, N. M. Fyllas, G. G. Mazzochini, S. Gachet, R. Gallagher, G. Ganade, F. Ganga, P. García- Palacios, V. Gargaglione, E. Garnier, J. L. Garrido, A. L. de Gasper, A. Gea-Izquierdo, D. Gibson, A. N. Gillison, A. Giroldo, M.-C. Glasenhardt, S. Gleason, M. Gliesch, E. Goldberg, B. Göldel, E. Gonzalez-Akre, J. L. Gonzalez-Andujar, A. González-Melo, A. González-Robles, B. J. Graae, E. Granda, S. Graves, W. A. Green, T. Gregor, N. Gross, G. R. Guerin, A. Günther, A. G. Gutiérrez, L. Haddock, A. Haines, J. Hall, A. Hambuckers, W. Han, S. P. Harrison, W. Hattingh, J. E. Hawes, T. He, P. He, J. M. Heberling, A. Helm, S. Hempel, J. Hentschel, B. Hérault, A.-M. Hereş, K. Herz, M. Heuertz, T. Hickler, P. Hietz, P. Higuchi, A. L. Hipp, A. Hirons, M. Hock, C. M. R. Pereira, M. Mack, V. Maire, A. Mäkelä, H. Mäkinen, A. C. M. Malhado, A. Mallik, P. Manning, S. Manzoni, Z. Marchetti, L. Marchino, V. Marcilio-Silva, E. Marcon, M. Marignani, L. Markesteijn, A. Martin, C. Martínez-Garza, J. Martínez-Vilalta, T. Mašková, K. Mason, N. Mason, T. J. Massad, J. Masse, I. Mayrose, J. M. Carthy, M. L. M. Cormack, K. M. Culloh, I. R. M. Fadden, B. J. M. Gill, M. Y. M. Partland, J. S. Medeiros, B. Medlyn, P. Meerts, Z. Mehrabi, P. Meir, F. P. L. Melo, M. Mencuccini, C. Meredieu, J. Messier, I. Mészáros, J. Metsaranta, S. T. Michaletz, C. Michelaki, S. Migalina, R. Milla, J. E. D. Miller, V. Minden, R. Ming, K. Mokany, A. T. Moles, A. Molnárth, J. Molofsky, M. Molz, R. A. Montgomery, A. Monty, L. Moravcová, A. Moreno-Martínez, M. Moretti, A. S. Mori, S. Mori, D. Morris, J. Morrison, L. Mucina, S. Mueller, C. D. Muir, S. C. Müller, F. Munoz, I. H. Myers-Smith, R. W. Myster, M. Nagano, S. Naidu, A. Narayanan, B. Natesan, L. Negoita, A. S. Nelson, E. L. Neuschulz, G. Niedrist, J. Ni, Ü. Niinemets, R. Nolan, H. Nottebrock, Y. Nouvellon, A. Novakovskiy; Nutrient Network, K. O. Nystuen, A. O'Grady, K. O'Hara, A. O'Reilly-Nugent, S. Oakley, W. Oberhuber, T. Ohtsuka, A. Oliveira, K. Öller, M. E. Olson, V. Onipchenko, Y. Onoda, R. E. Onstein, J. C. Ordonez, N. Osada, I. Ostonen, G. Ottaviani, S. Otto, G. E. Overbeck, W. A. Ozinga, A. T. Pahl, C. E. T. Paine, R. J. Pakeman, A. C. Papageorgiou, E. Parfionova, M. Pärtel, M. Patacca, S. Paula, J. Paule, H. Pauli, J. G. Pausas, B. Peco, J. Pe- nuelas, A. Perea, P. L. Peri, A. C. Petisco-Souza, A. Petraglia, A. M. Petritan, O. L. Phillips, S. Pierce, V. D. Pillar, J. Pisek, A. Pomogaybin, H. Poorter, A. Portsmuth, P. Poschlod, C. Potvin, D. Pounds, A. S. Powell, S. A. Power, A. Prinzing, G. Puglielli, P. Pyšek, V. Raevel, A. Rammig, J. Ransijn, C. A. Ray, P. B. Reich, M. Reichstein, D. E. B. Reid, M. Réjou-Méchain, V. R. de Dios, S. Ribeiro, S. Richardson, K. Riibak, M. C. Rillig, F. Riviera, E. M. R. Robert, S. Roberts, D. Robroek, A. Roddy, A. V. Rodrigues, A. Rogers, E. Rollinson, V. Rolo, C. Römermann, D. Ronzhina, C. Roscher, J. A. Rosell, M. F. Rosenfield, C. Rossi, D. B. Roy, S. Royer-Tardif, N. Rüger, R. Ruiz-Peinado, S. B. Rumpf, G. M. Rusch, M. Ryo, L. Sack, A. Saldaña, B. Salgado- Negret, R. Salguero-Gomez, I. Santa-Regina, A. C. Santacruz-García, J. Santos, J. Sardans, B. Schamp, M. Scherer-Lorenzen, M. Schleuning, B. Schmid, M. Schmidt, S. Schmitt, J. V. Schneider, S. D. Schowanek, J. Schrader, F. Schrodt, B. Schuldt, F. Schurr, G. S. Garvizu, M. Semchenko, C. Seymour, J. C. Sfair, J. M. Sharpe, C. S. Sheppard, S. Sheremetiev, S. Shio- dera, B. Shipley, T. A. Shovon, A. Siebenkäs, C. Sierra, V. Silva, M. Silva, T. Sitzia, H. Sjöman, M. Slot, N. G. Smith, D. Sodhi, P. Soltis, D. Soltis, B. Somers, G. Sonnier, M. V. Sørensen, E. E. Sosinski Jr., N. A. Soudzilovskaia, A. F. Souza, M. Spasojevic, M. G. Sperandii, A. B. Stan, J. Stegen, K. Steinbauer, J. G. Stephan, F. Sterck, D. B. Stojanovic, T. Strydom, M. L. Suarez, J.- C. Svenning, I. Svitková, M. Svitok, M. Svoboda, E. Swaine, N. Swenson, M. Tabarelli, K. Takagi, U. Tappeiner, R. Tarifa, S. Tauugourdeau, C. Tavsanoglu, M. T. Beest, L. Tedersoo, N. Thiffault, D. Thom, E. Thomas, K. Thompson, P. E. Thornton, W. Thuiller, L. Tichý, D. Tissue, M. G. Tjoelker, D. Y. P. Tng, J. Tobias, P. Török, S. Tarin, J. M. Torres-Ruiz, B. Tóthmérész, M. Treurnicht, V. Trivellone, F. Trolliet, V. Trotsiuk, J. L. Tsakalos, I. Tsiripidis, N. Tysklind, T.

M. Chytrý, D. Ciccarelli, L. Coll, C. G. Collins, L. Conti, D. Coomes, J. H. C. Cornelissen, W. K. Cornwell, P. Corona, M. Coyea, J. Craine, D. Craven, J. P. G. M. Cromsigt, A. Csecserits, K. Cufar, M. Cuntz, A. C. da Silva, K. M. Dahlin, M. Dainese, I. Dalke, M. D. Fratte, A. T. Dang-Le, J. Danihelka, M. Dannoura, S. Dawson, A. J. de Beer, A. De Frutos, J. R. De Long, B. Dechant, S. Delagrange, N. Delpierre, G. Derroire, A. S. Dias, M. H. Diaz-Toribio, P. G. Dimitrako-poulos, M. Dobrowolski, D. Doktor, P. Dřevojan, N. Dong, J. Dransfield, S. Dressler, L. Duarte, E. Ducouret, S. Dullinger, W. Durka, R. Duursma, O. Dymova, A. E.-Vojtkó, R. L. Eckstein, H. Ejtehadi, J. Elser, T. Emilio, K. Engemann, M. B. Erfanian, A. Erfmeier, A. Es-quivel-Muelbert, G. Esser, M. Estiarte, T. F. Domingues, W. F. Fagan, J. Fagúndez, D. S. Falster, Y. Fan, J. Fang, E. Farris, F. Fazlioglu, Y. Feng, F. Fernandez-Mendez, A. Ferrara, J. Ferreira, A. Fidelis, B. Finegan, J. Firn, T. J. Flowers, D. F. B. Flynn, V. Fontana, E. Forey, C. Forgiarini, L. François, D. Frank, C. Frenette-Dussault, G. T. Freschet, E. L. Fry, N. M. Fyllas, G. G. Mazzochini, S. Gachet, R. Gallagher, G. Ganade, F. Ganga, P. García-Palacios, V. Gargaglione, E. Garnier, J. L. Garrido, A. L. de Gasper, A. Gea-Izquierdo, D. Gibson, A. N. Gillison, A. Giroldo, M.-C. Glasenhardt, S. Gleason, M. Gliesch, E. Goldberg, B. Göldel, E. Gonzalez-Akre, J. L. Gonzalez-Andujar, A. González-Melo, A. González-Robles, B. J. Graae, E. Granda, S. Graves, W. A. Green, T. Gregor, N. Gross, G. R. Guerin, A. Günther, A. G. Gutiérrez, L. Haddock, A. Haines, J. Hall, A. Hambuckers, W. Han, S. P. Harrison, W. Hattingh, J. E. Hawes, T. He, P. He, J. M. Heberling, A. Helm, S. Hempel, J. Hentschel, B. Hérault, A.-M. Hereş, K. Herz, M. Heuertz, T. Hickler, P. Hietz, P. Higuchi, A. L. Hipp, A. Hirons, M. Hock, C. M. Hogan, K. Holl, O. Honnay, D. Hornstein, E. Hou, S. Hough-Snee, K. A. Hovstad, T. Ichie, B. Igić, E. Illa, M. Isaac, M. Ishihara, L. Ivanov, L. Ivanova, C. M. Iversen, J. Izquierdo, R. B. Jackson, B. Jackson, H. Jactel, A. M. Jagodzinski, U. Jandt, S. Jansen, T. Jenkins, A. Jentsch, J. R. P. Jespersen, G.-F. Jiang, J. L. Johansen, D. Johnson, E. J. Jokela, C. A. Joly, G. J. Jordan, G. S. Joseph, G. Junaedi, R. R. Junker, E. Justes, R. Kabzems, J. Kane, Z. Kaplan, K. Kattenborn, L. Kavelenova, E. Kearsley, A. Kempel, T. Kenzo, A. Kerkhoff, M. I. Khalil, N. L. Kinlock, W. D. Kissling, K. Kitajima, T. Kitzberger, R. Kjøller, T. Klein, M. Kleyer, J. Klimešová, J. Klipel, B. Kloeppel, S. Klotz, J. M. H. Knops, T. Kohyama, F. Koike, J. Kollmann, B. Komac, K. Komatsu, C. König, N. J. B. Kraft, K. Kramer, H. Kreft, I. Kühn, D. Kumarathunge, J. Kuppler, H. Kurokawa, Y. Kurosawa, S. Kuyah, J.-P. Laclau, B. Lafleur, E. Lallai, E. Lamb, A. Lamprecht, D. J. Larkin, D. Laughlin, Y. Le Bagousse-Pinguet, G. le Maire, P. C. le Roux, E. le Roux, T. Lee, F. Lens, S. L. Lewis, B. Lhotsky, Y. Li, X. Li, J. W. Lichstein, M. Liebergesell, J. Y. Lim, Y.-S. Lin, J. C. Linares, C. Liu, D. Liu, U. Liu, S. Livingstone, J. Llusià, M. Lohbeck, Á. López-García, G. Lopez-Gonzalez, Z. Lososová, F. Louault, B. A. Lukács, P. Lukeš, Y. Luo, M. Lussu, S. Ma, C. M. R. Pereira, M. Mack, V. Maire, A. Mäkelä, H. Mäkinen, A. C. M. Malhado, A. Mallik, P. Manning, S. Manzoni, Z. Marchetti, L. Marchino, V. Marcilio-Silva, E. Marcon, M. Marignani, L. Markesteijn, A. Martin, C. Martínez-Garza, J. Martínez-Vilalta, T. Mašková, K. Mason, N. Mason, T. J. Massad, J. Masse, I. Mayrose, J. M. Carthy, M. L. M. Cormack, K. M. Culloh, I. R. M. Fadden, B. J. M. Gill, M. Y. M. Partland, J. S. Medeiros, B. Medlyn, P. Meerts, Z. Mehrabi, P. Meir, F. P. L. Melo, M. Mencuccini, C. Meredieu, J. Messier, I. Mészáros, J. Metsaranta, S. T. Michaletz, C. Michelaki, S. Migalina, R. Milla, J. E. D. Miller, V. Minden, R. Ming, K. Mokany, A. T. Moles, A. Molnárth, J. Molofsky, M. Molz, R. A. Montgomery, A. Monty, L. Moravcová, A. Moreno-Martínez, M. Moretti, A. S. Mori, S. Mori, D. Morris, J. Morrison, L. Mucina, S. Mueller, C. D. Muir, S. C. Müller, F. Munoz, I. H. Myers-Smith, R. W. Myster, M. Nagano, S. Naidu, A. Narayanan, B. Natesan, L. Negoita, A. S. Nelson, E. L. Neuschulz, G. Niedrist, J. Ni, Ü. Niinemets, R. Nolan, H. Nottebrock, Y. Nouvellon, A. Novakovskiy; Nutrient Network, K. O. Nystuen, A. O'Grady, K. O'Hara, A. O'Reilly-Nugent, S. Oakley, W. Oberhuber, T. Ohtsuka, A. Oliveira, K. Öller, M. E. Olson, V. Onipchenko, Y. Onoda, R. E. Onstein, J. C. Ordonez, N. Osada, I. Ostonen, G. Ottaviani, S. Otto, G. E. Overbeck, W. A. Ozinga, A. T. Pahl, C. E. T. Paine, R. J. Pakeman, A. C. Papageorgiou, E. Parfionova, M. Pärtel, M. Patacca, S. Paula, J. Paule, H. Pauli, J. G. Pausas, B. Peco, J. Pe-nuelas, A. Perea, P. L. Peri, A. C. Petisco-Souza, A. Petraglia, A. M. Petritan, O. L. Phillips, S. Pierce, V. D. Pillar, J. Pisek, A. Pomogaybin, H. Poorter, A. Portsmuth, P. Poschlod, C. Potvin, D. Pounds, A. S. Powell, S. A. Power, A. Prinzing, G. Puglielli, P. Pyšek, V. Raevel, A. Rammig, J. Ransijn, C. A. Ray, P. B. Reich, M. Reichstein, D. E. B. Reid, M. Réjou-Méchain, V. R. de Dios, S. Ribeiro, S. Richardson, K. Riibak, M. C. Rillig, F. Riviera, E. M. R. Robert, S. Roberts, D. Robroek, A. Roddy, A. V. Rodrigues, A. Rogers, E. Rollinson, V. Rolo, C. Römermann, D. Ronzhina, C. Roscher, J. A. Rosell, M. F. Rosenfield, C. Rossi, D. B. Roy, S. Royer-Tardif, N. Rüger, R. Ruiz-Peinado, S. B. Rumpf, G. M. Rusch, M. Ryo, L. Sack, A. Saldaña, B. Salgado-Negret, R. Salguero-Gomez, I. Santa-Regina, A. C. Santacruz-García, J. Santos, J. Sardans, B. Schamp, M. Scherer-Lorenzen, M. Schleuning, B. Schmid, M. Schmidt, S. Schmitt, J. V. Schneider, S. D. Schowanek, J. Schrader, F. Schrodt, B. Schuldt, F. Schurr, G. S. Garvizu, M. Semchenko, C. Seymour, J. C. Sfair, J. M. Sharpe, C. S. Sheppard, S. Sheremetiev, S. Shio-dera, B. Shipley, T. A. Shovon, A. Siebenkäs, C. Sierra, V. Silva, M. Silva, T. Sitzia, H. Sjöman, M. Slot, N. G. Smith, D. Sodhi, P. Soltis, D. Soltis, B. Somers, G. Sonnier, M. V. Sørensen, E. E. Sosinski Jr., N. A. Soudzilovskaia, A. F. Souza, M. Spasojevic, M. G. Sperandii, A. B. Stan, J. Stegen, K. Steinbauer, J. G. Stephan, F. Sterck, D. B. Stojanovic, T. Strydom, M. L. Suarez, J.-C. Svenning, I. Svitková, M. Svitok, M. Svoboda, E. Swaine, N. Swenson, M. Tabarelli, K. Takagi, U. Tappeiner, R. Tarifa, S. Tauugourdeau, C. Tavsanoglu, M. T. Beest, L. Tedersoo, N. Thiffault, D. Thom, E. Thomas, K. Thompson, P. E. Thornton, W. Thuiller, L. Tichý, D. Tissue, M. G. Tjoelker, D. Y. P. Tng, J. Tobias, P. Török, S. Tarin, J. M. Torres-Ruiz, B. Tóthmérész, M. Treurnicht, V. Trivellone, F. Trolliet, V. Trotsiuk, J. L. Tsakalos, I. Tsiripidis, N. Tysklind, T. Umehara, V. Usoltsev, M. Vadeboncoeur, J. Vaezi, F. Valladares, J. Vamosi, P. M. van Bodegom, M. van Breugel, E. Van Cleemput, M. van de Weg, S. van der Merwe, F. van der Plas, M. T. van der Sande, M. van Kleunen, K. Van Meerbeek, K. A. Vanselow, A. Vårhammar, L. Varone, M. Y. V. Valderrama, K. Vassilev, M. Vellend, E. J. Veneklaas, H. Verbeeck, K. Verheyen, A. Vibrans, I. Vieira, J. Villacís, C. Violle, P. Vivek, K. Wagner, M. Waldram, A. Waldron, A. P. Walker, M. Waller, G. Walther, H. Wang, F. Wang, W. Wang, H. Watkins, J. Watkins, U. Weber, J. T. Weedon, L. Wei, P. Weigelt, E. Weiher, A. W. Wells, C. Wellstein, E. Wenk, M. Westoby, A. Westwood, P. J. White, M. Whitten, M. Williams, D. E. Winkler, K. Winter, C. Womack, I. J. Wright, S. J. Wright, J. Wright, B. X. Pinho, F. Ximenes, T. Yamada, K. Yamaji, R. Yanai, N. Yankov, B. Yguel, K. J. Zanini, A. E. Zanne, D. Zelený, Y.-P. Zhao, J. Zheng, J. Zheng, K. Ziemińska, C. R. Zirbel, G. Zizka, I. C. Zo-Bi, G. Zotz, C. Wirth, TRY plant trait database - enhanced coverage and open access. *Glob. Chang. Biol.* **26**, 119–188 (2020).

93. J. P. Grime, Vegetation classification by reference to strategies. *Nature* **250**, 26–31 (1974).

94. F. Schrodt, J. Kattge, H. Shan, F. Fazayeli, J. Joswig, A. Banerjee, M. Reichstein, G. Bönisch, S. Díaz, J. Dickie, A. Gillison, A. Karpatne, S. Lavorel, P. Leadley, C. B. Wirth, I. J. Wright, S. J. Wright, P. B. Reich, BHPMF - a hierarchical Bayesian approach to gap-filling and trait prediction for macroecology and functional biogeography. *Glob. Ecol. Biogeogr.* **24**, 1510–1521 (2015).

95. S. van Buuren, K. Groothuis-Oudshoorn, mice: Multivariate imputation by chained equations in *R. J. Stat. Softw.* **45**, 1–67 (2011).

96. C. Maldonado, C. I. Molina, A. Zizka, C. Persson, C. M. Taylor, J. Albán, E. Chilquillo, N. Rønsted, A. Antonelli, Estimating species diversity and distribution in the era of big data: To what extent can we trust public databases? *Glob. Ecol. Biogeogr.* **24**, 973–984 (2015).

97. A. Zizka, D. Silvestro, T. Andermann, J. Azevedo, C. Duarte Ritter, D. Edler, H. Farooq, A. Herdean, M. Ariza, R. Scharn, S. Svanteson, N. Wengström, V. Zizka, A. Antonelli, Coor-dinateCleaner: Standardized cleaning of occurrence records from biological collection databases. *Methods Ecol. Evol.* **10**, 744–751 (2019).

98. D. S. Maynard, L. Bialic-Murphy, C. M. Zohner, C. Averill, J. van den Hoogen, H. Ma, L. Mo, G. R. Smith, A. T. R. Acosta, I. Aubin, E. Berenguer, C. C. F. Boonman, J. A. Catford, B. E. L. Cerabolini, A. S. Dias, A. González-Melo, P. Hietz, C. H. Lusk, A. S. Mori, Ü. Niinemets, V. D. Pillar, B. X. Pinho, J. A. Rosell, F. M. Schurr, S. N. Sheremetev, A. C. da Silva, Ê. Sosinski, P. van Bodegom, E. Weiher, G. Bönisch, J. Kattge, T. W. Crowther, Global relationships in tree functional traits. *Nat. Commun.* **13**, 3185 (2022).

99. J. Oksanen, F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, H. Wagner, *vegan: Com-munity Ecology Package* (R Project, 2019).

100. W. Weckwerth, Metabolomics in systems biology. *Annu. Rev. Plant Biol.* **54**, 669–689 (2003).

101. H. Uthe, N. M. van Dam, M. R. Hervé, M. Sorokina, K. Peters, A. Weinhold, in *Plant Me-tabolomics in full swing* (Elsevier, 2021), vol. 98 of *Advances in botanical research*, pp. 163–203.

102. K. Habel, R. Grasman, R. B. Gramacy, P. Mozharovskyi, D. C. Sterrat, *geometry: Mesh Gen-eration and Surface Tessellation* (R, 2022).

103. S. Dray, A.-B. Dufour, The ade4 package: Implementing the duality diagram for ecologists. *J. Stat. Softw.* **22**, 1–20 (2007).

104. R. Salguero-Gómez, Applications of the fast-slow continuum and reproductive strategy framework of plant life histories. *New Phytol.* **213**, 1618–1624 (2017).

105. H. Kusano, H. Li, H. Minami, Y. Kato, H. Tabata, K. Yazaki, Evolutionary developments in plant specialized metabolism, exemplified by two transferase families. *Front. Plant Sci.* **10**, 794 (2019).

106. R. T. Bush, F. A. McInerney, Leaf wax n-alkane distributions in and across modern plants: Implications for paleoecology and chemotaxonomy. *Geochim. Cosmochim. Acta* **117**, 161–179 (2013).

107. L. F. Maia, V. E. De Oliveira, H. G. M. Edwards, L. F. C. De Oliveira, The diversity of linear conjugated polyenes and colours in nature: Raman spectroscopy as a diagnostic tool. *ChemPhysChem* **22**, 231–249 (2021).

108. J. K. Hoober, L. L. Eggink, M. Chen, Chlorophylls, ligands and assembly of light-harvesting complexes in chloroplasts. *Photosyn. Res.* **94**, 387–400 (2007).

109. D. F. Veber, S. R. Johnson, H.-Y. Cheng, B. R. Smith, K. W. Ward, K. D. Kopple, Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* **45**, 2615–2623 (2002).

110. P. Ertl, B. Rohde, P. Selzer, Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *J. Med. Chem.* **43**, 3714–3717 (2000).

111. K. Palm, P. Stenberg, K. Luthman, P. Artursson, Polar molecular surface properties predict the intestinal absorption of drugs in humans. *Pharm. Res.* **14**, 568–571 (1997).

112. G. A. Jeffrey, W. Saenger, *Hydrogen bonding in biological structures* (Springer Berlin Hei-delberg, 1991).