Received: 25 February 2021 | Revised: 21 March 2022 | Accepted: 28 March 2022

DOI: 10.1111/geb.13564

Global Ecology and Biogeography | A Journal of Macroecology

WILEY

RESEARCH ARTICLE

Strong phylogenetic signals in global plant bioclimatic envelopes

Claire Harris[1,2] | Neil Brummitt[2,3] | Christina A. Cobbold[2,4] | Richard Reeve[1,2]

[1]Institute of Biodiversity, Animal Health and Comparative Medicine, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, UK

[2]Boyd Orr Centre for Population and Ecosystem Health, University of Glasgow, Glasgow, UK

[3]Natural History Museum, London, UK

[4]School of Mathematics and Statistics, University of Glasgow, Glasgow, UK

Correspondence
Richard Reeve, Institute of Biodiversity, Animal Health and Comparative Medicine, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, G12 8QQ, UK.
Email: richard.reeve@glasgow.ac.uk

Funding information
Engineering and Physical Sciences Research Council, Grant/Award Number: EP/M506539/1; Natural Environment Research Council, Grant/Award Number: NE/T004193/1 and NE/T010355/1

Handling Editor: Gregory Jordan

Abstract

Aim: The environmental preferences of species are an important facet of their response to changing conditions, and these have long been thought to exhibit phylogenetic conservatism. However, these bioclimatic envelopes have not previously been imputed from climate records at the date and location of occurrence, and the strength of their phylogenetic signal has not been studied at a broad scale. Here, we combine records from global climate reconstructions with contemporaneous plant occurrences for all available terrestrial plant species and test for phylogenetic niche conservatism in plant climatic traits.

Location: Global.

Time period: 1901–2018.

Major taxa studied: Terrestrial plants.

Methods: We used >100 million plant records from the Global Biodiversity Information Facility (GBIF) to produce distributions of bioclimatic envelopes for >200,000 species, using a range of climate variables. We matched species observations to historical climate reconstructions from the European Centre for Medium-Range Weather Forecasting (ECMWF) and compared this with WorldClim climate averages. We tested for phylogenetic signal in a supertree of plants using Pagel's λ. Finally, to investigate how well bioclimatic envelopes could be inferred for poorly known and rare species, we performed cross-validation by removing occurrence records for some common species to test how accurately their bioclimatic envelopes were estimated.

Results: We found extremely strong phylogenetic signals ($\lambda > 0.9$ in some cases) for climate variables from both climate datasets, including temperature, soil temperature, solar radiation and precipitation. We were also able to impute missing bioclimatic envelopes for artificially removed species, having a correlation with observed data of .7.

Main conclusions: We reconstructed plant climatic tolerances for >200,000 plant species historically recorded on GBIF using a technique that could be applied to any comparable biodiversity dataset. Although global information on most species is sparse, we explored methods for bias correction and data imputation, with positive results for both.

KEYWORDS
biodiversity, climate change, global, phylogenetic niche conservatism, plants, traits

1 | INTRODUCTION

The United Nations Framework Convention on Climate Change (UNFCC) Paris Agreement was ratified in 2015 in an effort to keep global temperature level rise well below 2°C above pre-industrial levels, with an ultimate goal of 1.5°C (UNFCCC, 2015). In 2018, the Intergovernmental Panel on Climate Change (IPCC) reported that we were likely to reach 1.5°C in the next few decades (2030–2052), with an estimated increase of *c*. 0.2°C per decade (Intergovernmental Panel on Climate Change, 2018). The impacts of such climate change are already being observed (Cheng et al., 2019; EASAC, 2018; Marvel et al., 2019; Slater et al., 2020) and set to increase in severity (Intergovernmental Panel on Climate Change, 2018), especially if mitigation measures against CO_2 emissions are insufficient (Intergovernmental Panel on Climate Change, 2018). There is evidence that historical and current CO_2 emissions are tracking representative concentration pathway (RCP) 8.5, the most extreme in terms of fossil fuel use of all RCPs used in climate modelling (Schwalm et al., 2020). Under the RCP8.5 scenario, global temperatures would be likely to increase in the range of 2.6–4.8°C above pre-industrial levels by 2100 (IPCC, 2014). However, according to the United Nations Environment Programme, even if all of the action plans put in place as a result of the Paris Agreement were followed, we could still see a rise of 3.2°C by the end of the century (Olhoff & Christensen, 2019).

Heat waves, droughts and extreme weather events, such as cyclones, are all expected to increase in frequency over the coming years and are thought to have a particular impact on the survival and diversity of plant communities (Jentsch & Beierkuhnlein, 2008; Reyer et al., 2013). These climate extremes, along with rising sea levels, are expected to contribute towards the loss of climatically suitable range for species in the future (Nunez et al., 2019; Warren et al., 2018). Some studies already report widespread climate-change-induced local extinctions in animals and plants, especially amongst tropical and subtropical regions when compared with temperate regions (Wiens, 2016). In a study of *c*. 70,000 plant species, 16% were estimated to lose >50% of their range if warming were to reach 2°C by 2100, compared with 8% if it were restricted to the lower 1.5°C scenario (Warren et al., 2018). These predictions paint a bleak picture for the future of plant biodiversity, particularly when an estimated fifth of assessed plant species were already classified as threatened with extinction and new reports suggest this figure could be as high as 40% (Antonelli et al., 2020; Brummitt et al., 2015).

Key to understanding the responses of species to climate change are their historical and current relationships with climate. According to traditional ecological theory, species are adapted for a particular set of environmental conditions, occupying a unique niche within the ecosystem (Hutchinson, 1957). The climate makes up a large part of this fundamental niche, the bioclimatic envelope, which is often modelled separately from other environmental variables, such as land cover or soil type, and from the biotic interactions that can limit ranges (Pearson & Dawson, 2003). Apart from the bioclimatic envelopes produced by species distribution models (SDMs), there are few other sources of information for plant climatic envelopes. Although physiological information on the climatic tolerances of species can be inferred through experiments, it is impractical to do so for the vast majority of plants on Earth (Araújo & Peterson, 2012). There is some information on plant species traits through databases, such as TRY (Kattge et al., 2011; Weigelt et al., 2019) or the Global Root Traits database (GRooT; Guerrero-Ramírez et al., 2020), but little relates to climatic variables, and what is available is either predominately categorical data or taxonomically or geographically limited. Given that it is widely established that climate is one of the biggest drivers of plant morphology and function, it has long been posited that the climate range of plants can be predicted through using information on plant functional traits (Box, 1996; Stahl et al., 2014; Woodward & Williams, 1987). Traits such as specific leaf area and woodiness are much more readily available than direct information on bioclimatic envelopes, but there is little consensus beyond small case studies on how to relate these traits to climate tolerance (Stahl et al., 2014).

Along with plant functional trait information, there is increasing incorporation of evolutionary history into the prediction of plant climate ranges. Although there is some evidence of a phylogenetic signal in climate variables, there are equally many discrepancies (Zhang et al., 2017). This has been hampered by a lack of resolved, global phylogenies and accompanying data on the climatic envelopes of species. It has been hypothesized that such climatic niches are phylogenetically conserved, a concept known as phylogenetic niche conservatism (PNC; Harvey & Pagel, 1991). There are a number of different factors that could contribute to a pattern of PNC, including evolutionary constraints on physiology and processes such as neutral drift and dispersal limitation (Crisp & Cook, 2012). Therefore, the mechanism by which we measure PNC is confounded by the number of different processes that could give rise to such a pattern (Crisp & Cook, 2012; Losos, 2011). In many studies, measurements of phylogenetic signal (PS), such as Blomberg's *K* (Blomberg et al., 2003) or Pagel's λ (Pagel, 1999), are also used to detect PNC, although there is still some debate about the link between PNC and phylogenetic signal.

Here, we extract information on the climate experienced by >200,000 plant species recorded in the Global Biodiversity Information Facility (GBIF) from long-term climate re-analysis datasets created by the European Centre for Medium-Range Weather Forecasting (ECMWF), in addition to climate normals (averaged for a given month over a 30 year period) created by

WorldClim. We then link these bioclimatic envelopes to a super-tree of >30,000 plant species Qian and Jin (2016) to detect phylogenetic signals in climatic tolerances for the full tree and the top 5,000 most common species in GBIF. Given that most species in GBIF are under-recorded, we also correct for these biases and test the capacity of the models to predict bioclimatic envelopes for missing data.

2 | METHODS

2.1 | Data

We extracted the entire history of plant occurrences from the GBIF (GBIF, 2019), which totals *c*. 212 million individual records. The data were filtered for species with an accepted name in The Plant List (TPL, 2013), in order to remove marine species and fossils, which left *c*. 215,000 species and slightly >100 million records. The data were additionally cleaned for the presence of 3,441 botanical garden locations with available GPS coordinates, taken from Botanic Gardens Conservation International (BGCI, 2019), to eliminate records erroneously georeferenced to the locations at which the collections are stored. A buffer of 0.02 decimal degrees was taken around each botanical garden, and all points that lay within this area were excluded. The same procedure was performed for country centroids, which can be assigned erroneously to records during the georeferencing process if no locality information is available. We also explored filtering this data further for the top 100 recording institutions, excluding those likely to include garden or greenhouse observations. Given that we found comparable results, we retain the original data in the study (for the small differences observed, see Supporting Information Figure S1; Table S1). GBIF also makes no distinction between native and alien records in their data. However, given that we are interested in the climate envelopes that these species could tolerate, rather than their native ranges, information from introductions to other climates is very valuable to this end because it indicates their survival in new environments.

We downloaded historical climate reconstructions from the ECMWF for two re-analysis datasets: CERA-20C and ERA-Interim (Dee et al., 2011; Laloyaux et al., 2018). We used the aggregated monthly datasets made available through the ECMWF, which are means of these 3 h records, for 11 climatic variables, including temperature, precipitation and solar radiation (see Table 1). We replaced the final 39 years (1979–2018) with the more spatially resolved ERA-Interim for the same variables to create a complete time line for 1901–2018. Finally, we compared this with the WorldClim dataset (Fick & Hijmans, 2017). In contrast to the ECMWF data, WorldClim is much less temporally resolved and is averaged over a 30-year time period from 1970 to 2000 for every month in the year. For computational efficiency, we used a grid size of 10 min of arc, or *c*. 18 km at the equator (ranging in size down to 1 km at the poles). See Table 2 for an overview of the temporal and spatial resolutions of the datasets.

In order to correct for biases in plant occurrence data, we also used mean global-level enhanced vegetation index (EVI) as a proxy of plant density across the period 2001–2015 (Huete et al., 1999). The data were originally derived from the US National Aeronautics and Space Administration (NASA) Moderate Resolution Imaging Spectrometer (MODIS) dataset, with processing and gap filling in cloud cover performed by Gibson and Weiss (2015). We excluded Antarctica and all islands within the Arctic circle, including Greenland, owing to poor data quality.

2.2 | Data extraction and bias correction

Using the *c*. 215,000 plant species we obtained from GBIF and the climate reconstructions from ECMWF (and other sources), we extracted the plant bioclimatic envelopes using the timing and location of occurrences. This created a raw dataset of global bioclimatic envelopes, and we also explored several adjustments to account for biases in historical collection practices. We present the full compiled database as an output of this work.

2.2.1 | Raw data

For extracting profiles from ERA-Interim and CERA-20C data, we assumed that each of the individual plants recorded in GBIF, or a parent in the case of annuals, would have been present at the site for ≥1 year before collection. Thus, for every record available in

TABLE 1 Details of the climate variables downloaded from each dataset, where ✓ and × denote availability

Climate variable	Derived variables	Parameter	Unit	CERA-20C	ERA-Interim	WorldClim
Temperature at 2 m	Minimum	tmin	°C	✓	✓	✓
	Maximum	tmax	°C	✓	✓	✓
	Mean	tmean	°C	✓	✓	✓
Total precipitation	Mean	tp	mm	✓	✓	✓
Solar radiation	Mean	ssr	kW	✓	✓	✓
Soil temperature layers 1–4	Mean	stl	°C	✓	✓	×
Soil water volume layers 1–4	Mean	swvl	m^3	✓	✓	×

TABLE 2 Details of the climate and vegetation datasets used, including the range of years used and the spatial resolution

Dataset	Temporal range	Temporal resolution	Spatial resolution (km)	Reference
CERA-20C	1901–2010	Monthly	125	Laloyaux et al. (2018)
ERA-Interim	1979–2018	Monthly	80	Dee et al. (2011)
WorldClim	1970–2000	Monthly	20	Fick and Hijmans (2017)
		Average[a]		
EVI	2001–2015	Average[b]	80 and 20	Gibson and Weiss (2015)

[a]Climate normals, which are averaged over the entire time period for each month.
[b]Single average over the entire dataset.

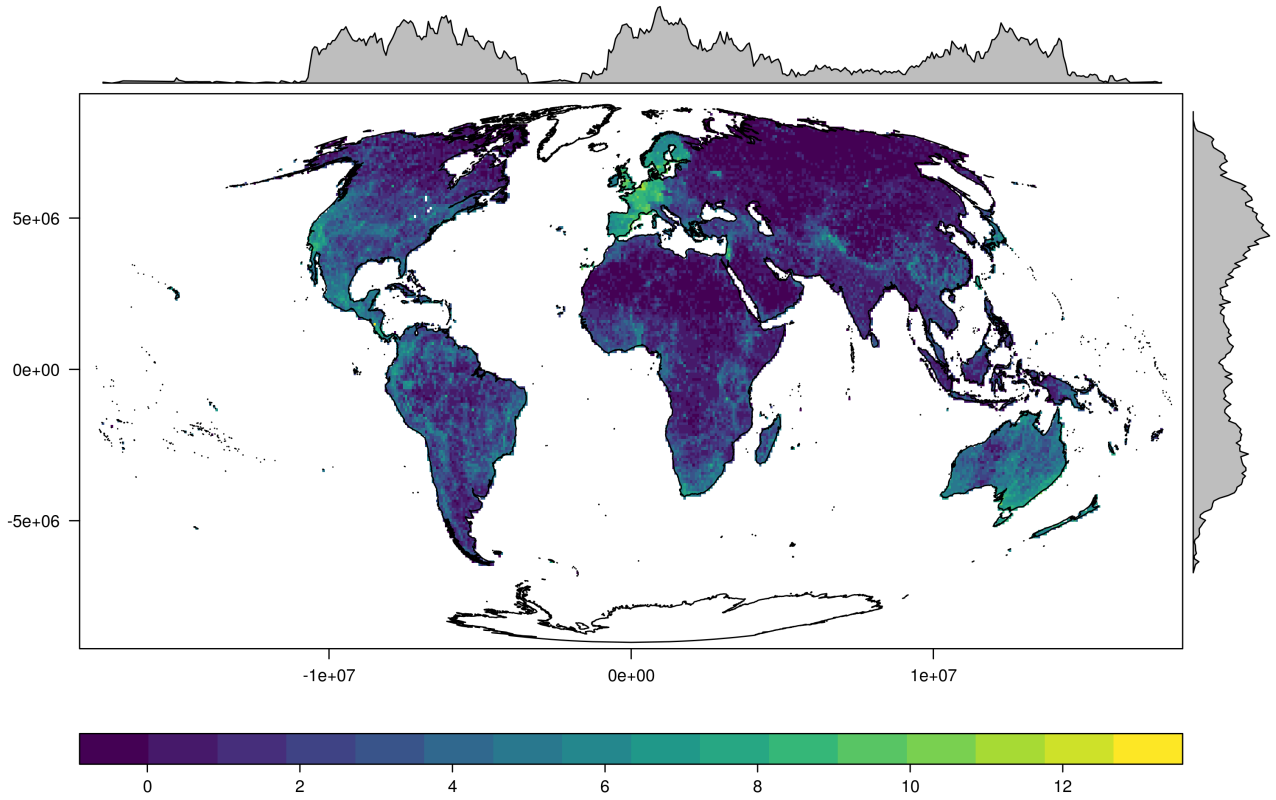

FIGURE 1 Natural log recording effort for historical plant observations in the Global Biodiversity Information Facility (GBIF) since 1901 after controlling for plant density using the enhanced vegetation index (EVI). White indicates areas excluded from this analysis, namely islands from the Arctic circle and the Antarctic land mass. Grey lines above and to the right of the graph indicate summed values of recording effort at each longitude and latitude, respectively. The map is plotted in the Mollweide equal area projection.

the filtered GBIF database, we extracted each of the variables in Table 1 ("climate variables") at the month and location in which they were recorded and the preceding 11 months as a profile of climates in which the species could survive. We averaged across the 12 months for every record and for each climate variable, with the exception of temperature, for which we also calculated minimum and maximum values (Table 1, "derived variables"). For WorldClim data, we extracted the entire 12 months of averaged data at the location where the plants were found. We then binned all records per species into 1,000 bins for each climate variable to produce a set of plant bioclimatic envelope profiles, to which we could then apply various corrections for bias in the underlying GBIF data.

2.2.2 | Adjusted data

Plant records available in GBIF suffer from extreme spatial and taxonomic biases (Figure 1). In particular, there is a concentration of higher recording effort in Western Europe, Eastern America and Australia, and recording is particularly low throughout the tropics, where there is the highest concentration of rare and endangered species. In order to control for such biases in the raw GBIF data, we performed several corrections to reduce bias in the corresponding plant species bioclimatic envelopes. The first was to recover the true abundance of plant species at different levels of climate variables. In order to do this, we weighted the binned species preference profiles by the ratio of global distribution of plant collections against a proxy

of plant abundance, the EVI, to reduce observation bias (hereafter "effort adjustment"; Supporting Information Figure S2b). The EVI has improved sensitivity to regions of high plant biomass in comparison to the more commonly used normalized difference vegetation index (NDVI) (Huete et al., 1999). The second correction was to recover species envelopes in an idealized landscape where all climates were equally available, by correcting for the observed biases in the global climate (hereafter "climate adjustment"; Supporting Information Figure S2c). This followed the observation that the climates experienced by plant species were strongly bimodal in some parameters, particularly temperature, because there were excess land areas with low (around freezing) and high (around 26°C) average monthly temperatures (Supporting Information Figure S3). Finally, to prepare the data for input into the phylogenetic models, we averaged across the climate bins, weighting by the number of records, to obtain an overall value per species for each climate variable and correction type. This resulted in three datasets: the raw averaged bioclimatic envelope per species; the bioclimatic envelope adjusted for effort; and the bioclimatic envelope adjusted for both effort and climate.

There is large variation in the climate variables and levels of plant data collection between different parts of the world, with a particular bias towards Western Europe (Figure 1). Making adjustments at the level of the individual grid square or as a spatial kernel is problematic, because there are large parts of the world for which there are no or very few plant records, which would need enormous corrections to achieve European sampling levels. We therefore also divided the data into six continental-scale areas [henceforth continents; South America, North America, Africa, Europe (excluding Russia), Asia (including Russia) and Australasia] and applied the adjustments separately to each. Thus, we have both a global and continental scale for each analysis.

2.3 | Phylogeny and estimation of evolutionary signal

There is much debate in the literature over the prevalence of PNC (Muünkemüller et al., 2015), whereby related species have more similar environmental niches. If we were able to detect a strong phylogenetic signal in the raw or adjusted bioclimatic variables that we extracted (and not with simulated data with no such relationship), then it would be a clear indication both that we were correctly identifying true traits for these species (see next paragraph) and that PNC is broadly present across the plant kingdom. To investigate this, we therefore tested for evidence of phylogenetic signals in the extracted bioclimatic parameters (Figure 2). Specifically, Qian et al. (2016) revised a supertree of vascular plants that comprises 31,389 species and corrects many taxonomic issues with the previously published version (Zanne et al., 2014). Of these species, 26,466 species had at least one corresponding record in the GBIF database, *c*. 84% of species in the supertree; therefore, we tested for a signal in the parameters in that 26,466-tip subtree. In order to test for a phylogenetic signal in the climate variables, we used Pagel's λ, which has been shown to be more robust against incomplete phylogenies and inaccurate branch-length information than Blomberg's *K* (Molina-Venegas & Rodríguez, 2017). Values of Pagel's λ closer to zero indicate that traits show little phylogenetic signal, whereas values approaching

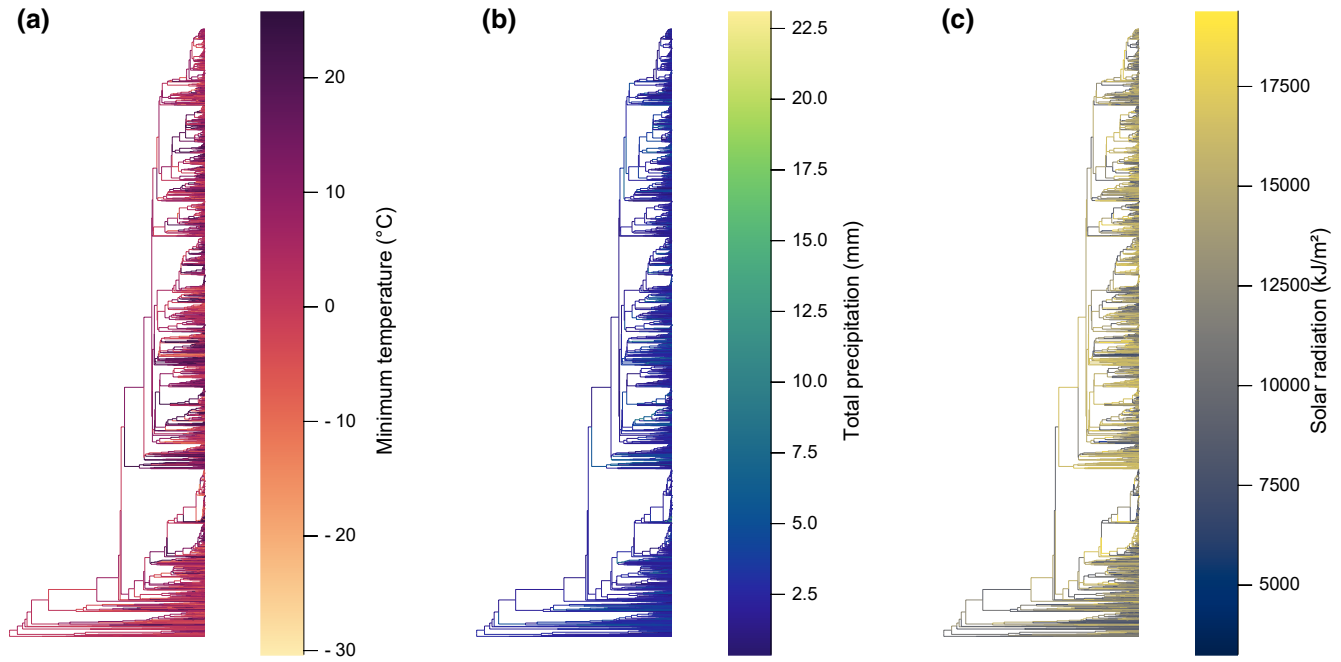

FIGURE 2 Evolutionary tree of the 5,000 plant species with the most records in the Global Biodiversity Information Facility (GBIF), along with their reconstructed climate profiles: (a) minimum temperature (in degrees Celsius); (b) mean monthly rainfall (in millimetres); and (c) mean monthly solar radiation (in kilowatts).

one indicate traits evolved along the phylogeny with a Brownian motion model, or random genetic drift (Harmon, 2019). We also simulated signal-free data by randomly permuting the tree tips to estimate what value of Pagel's λ would be inconsistent with the absence of a phylogenetic signal.

In addition to simply detecting the phylogenetic signal, we used Pagel's λ as a barometer for the success of both the parameter extraction and the adjustments described above, because random (changes to the) data should produce weak(er) responses in λ. High values of λ (or increases in λ) were viewed as an indication that a parameter was captured correctly (or a correction had been effective). We also applied the λ models to the continentally adjusted datasets. In every case, we randomized the tips to confirm that $\lambda \approx 0$ when no phylogenetic signal was present. For the fitting, we used the PhyloNetworks package v.0.11.0 (Solís-Lemus et al., 2017) in Julia v.1.5.2 (Bezanson et al., 2017). The code for these analyses is available under an open source license on GitHub (see data availability statement).

2.3.1 | Full tree and subtrees

Between the tree and bioclimatic envelopes, there was a crossover of >26,000 plant species. We ran analyses of Pagel's λ across all the climate variables for this full tree. We also performed the same analysis for individual taxonomic levels of the tree, including runs for each genus, family, order and class, when there were ≥50 species present. We excluded phyla from this analysis, however, because a single phylum encompasses essentially all species in the tree.

Given that there were not enough historical occurrence records to classify the bioclimatic envelopes of many of the species accurately, we also took various subsets of the data to compare the signals found.

2.3.2 | Subset tree

We fitted the λ models of Brownian motion to the 5,000 most common species ranked by number of occurrences, when we were confident that there were enough GBIF records to build a bioclimatic envelope accurately (i.e., ≥1,000 records). The data from these 5,000 species encompass *c*. 80% of the total GBIF data. To explore whether the signals were influenced by spatial autocorrelation between congenerics, we also calculated Pagel's λ on a subset of the data for species with occurrences on more than one continent and thus with the broadest range sizes (*c*. 4,000 species in total). In order to see how this signal varied across the phylogeny, we included a Moran's *I* phylogenetic correlogram and a plot of local Moran's *I* index for the climate variable with the strongest λ signal (minimum temperature) using the R package phylosignal (Keck et al., 2016). Local Moran's *I* index was calculated at the genus level owing to the size of the tree and the resulting computational and visualization challenges. We also examined the correlation between variables using Pearson correlation.

Finally, we used the top 5,000 most common species again to test how well we could impute missing values. We carried out a 10-fold cross-validation to reconstruct each of the climate variables using *ancestralStateReconstruction* from the Julia package PhyloNetworks (10% of the data at a time). This method predicts the expected values and variances of traits (the bioclimatic envelopes) for plant species with missing data by using known information for other species and estimating the evolutionary parameters of a Brownian motion model. We then explored the correlation of the resulting imputed climate profiles to the original data using Pearson correlation in order to quantify the performance of the reconstruction. We also compared the strength of the correlations against those from a tree with randomly shuffled tips. We were unable to perform this analysis on the full tree owing to computational constraints.

3 | RESULTS

Starting with examination of the raw ECMWF data, which was uncorrected for any type of sampling bias, we observed strong correlations between the climate variables. Temperature-related variables and water-related variables showed positive correlations within groups (.01–.99) and negative correlations between groups (−.09 to −.86). There was also a strong relationship between the raw climate variables calculated from ECMWF and those from WorldClim (.83–.97), approaching one in the case of mean and minimum temperatures.

3.1 | Phylogenetic signal in climate variables

The phylogenetic signal, Pagel's λ, was strong across the range of climate variables tested (Figure 3; Table 3). Randomizing the tips for this tree showed $\lambda < 0.01$, demonstrating that zero was a valid null. For the full tree, we found strong signals of >0.8 for most variables (Figure 3a). For the following results, we considered the subset tree only (Figure 3b), because the signal strength was very similar. Temperature, particularly minimum temperature and the soil temperature levels, had a signal of >0.9, and rainfall showed a similarly strong signal. Soil water volume, in contrast, performed relatively poorly, but still showed a signal of 0.7. However, this is still ordinarily considered to be a sign of a strong phylogenetic signal. The effort-adjusted data performed better than the raw data values, indicating that this effort adjustment was a valuable correction, except for a very low signal for soil water volume in the full tree. The subsequent climate correction reduced the signal at both the global and continental levels, suggesting that this correction was not functioning as intended. We excluded it from our further results, but we noted that it did improve the signals for soil water volume and precipitation when performed at the continental level (Supporting Information Figure S4; Table S2). The results for species with occurrences in more than one continent were

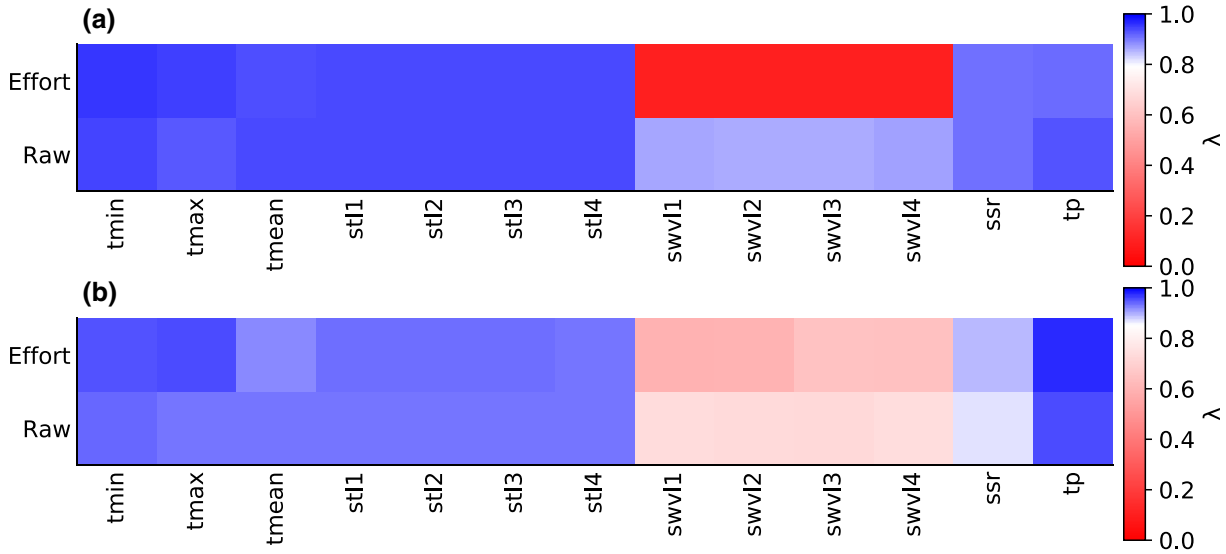

FIGURE 3 Phylogenetic signal, Pagel's λ, for each of the climate variables and levels of correction: Raw climate variables (Raw) and adjusted by effort (Effort). The colour scheme in the key is scaled to be centred on the mean λ value. The analyses used: (a) the full 26,466 species tree; and (b) a subset of the top 5,000 most common species.

TABLE 3 Phylogenetic signal, Pagel's λ, for each of the climate variables in the European Centre for Medium-Range Weather Forecasting (ECMWF) bioclimatic envelopes for both full and subset trees (for full parameter names, see Table 1), and for each of the levels of correction: Raw climate variables; and adjusted by effort

Climate variable	Full ECMWF		Subset ECMWF		WorldClim	
	Raw	Effort	Raw	Effort	Raw	Effort
tmin	0.945	0.953[a]	0.939	0.951[a]	0.940	0.958[a,b]
tmax	0.929	0.947[a]	0.930	0.954[a]	0.935[+]	0.952[a,b]
tmean	0.938	0.935	0.931	0.920	0.926	0.945[a,b]
stl1	0.938	0.94[a]	0.930	0.936[a]	–	–
stl2	0.938	0.94[a]	0.932	0.935[a]	–	–
stl3	0.939	0.94[a]	0.933	0.934[a]	–	–
stl4	0.939	0.939	0.933	0.933	–	–
swvl1	0.865	0.1	0.734	0.596	–	–
swvl2	0.862	0.1	0.730	0.599	–	–
swvl3	0.86	0.1	0.723	0.650	–	–
swvl4	0.868	0.1	0.738	0.645	–	–
Ssr	0.908	0.909[a]	0.871	0.892[a]	0.867	0.883[a]
Tp	0.932	0.913	0.956	0.975[a]	0.941	0.955[a]

Note: Corresponding numerical values of the phylogenetic signal from Figure 3.

[a]Values that are higher than for the raw data.

[b]Positive increases on the ECMWF analysis.

also very similar to the original analyses (Table S3). The Moran's *I* correlogram showed significant positive autocorrelation between minimum temperature and phylogenetic distance throughout the most recent 500 Ma (Supporting Information Figure S5). This signal weakened and became negative towards the root of the tree (>600 Ma). At the genus level, we found significant hotspots of phylogenetic signal amongst many of the genera, in addition to significant heterogeneity across the phylogeny (Supporting Information Figure S6). We also saw a great deal of variation in Pagel's λ across individual taxonomic levels of the tree (Supporting Information Figure S7), although average λ value tended to increase as we included more of the tree.

There was a strong positive correlation between the equivalent parameters extracted from WorldClim and ECMWF, hence it was unsurprising that there was a similar phylogenetic signal seen in these parameters despite the 30-year temporal averaging of the WorldClim dataset (Figure 3; Table 3). Again, temperature showed a strong signal, indicating that it is an evolutionary driver, and overall, the values for λ were slightly higher than for the corresponding ECMWF variables. Solar radiation and total precipitation showed

much the same response as in the previous analysis, but were slightly weaker than with ECMWF data.

3.2 | Imputation of bioclimatic envelopes

Given that there were strong phylogenetic signals in the parameters tested, we investigated how well we might be able to impute bioclimatic envelope data on rarer or less-observed species in GBIF. We ran cross-validations using data available from the 5,000 species for which the most data were available. Correlations and root mean square errors (RMSEs) between imputed and raw data were strong for most variables, particularly those that showed a strong phylogenetic signal in the previous analysis (Figure 4, imputed data). For instance, minimum temperature, the variable with the strongest λ value, showed a correlation between the imputed values and real data of .79, and an RMSE of 4.69°C (for a plot of real against imputed results for minimum temperature, see Supporting Information Figure S8). We randomized the tips of the tree in order to ground-truth the data and found a correspondingly low correlation between imputed and raw values (Figure 4, randomized imputed data).

4 | DISCUSSION

The extraction of climate tolerance profiles for >200,000 species is the largest so far. The few previous studies that extracted this type of information from climate datasets have been limited in taxonomic or geographical scope (Curtis & Bradley, 2016; Feeley, 2015; Harbert & Nixon, 2015; Sparrius et al., 2018). This is a natural consequence of most studies focusing on certain regions or families of plants. Although ground-truthing the data for a project of this size is extremely challenging, the strong phylogenetic signal seen in almost all traits suggests that the extracted data are indeed a comprehensive picture of global species bioclimatic envelopes, at least for the more common species. In particular, controlling for effort in plant sampling improves the phylogenetic signals in both ECMWF and WorldClim data. We have also tested the possibility that we could infer traits for missing or rarer species in the tree using phylogenetic relationships, finding a good correlation between imputed and real values.

Most phylogenetic analyses of plant functional traits or habitat preferences focus on phenological variables (e.g., Basnett et al., 2019; Davies et al., 2013) or characteristics of the environment such as soil pH or nitrogen levels (e.g., Schreeg et al., 2010). Of the few studies that incorporate climate parameters, some report evidence of phylogenetic signal in climatic variables, such as temperature and precipitation, at the genus or family level (Steinbauer et al., 2016; Xu et al., 2019), whereas others report little or non-significant signals (Koski & Ashman, 2016; Li et al., 2017; Liu et al., 2015) at that level. Here, however, we see a very strong phylogenetic signal in almost all the climate variables considered, even sometimes in soil water volume, but at the level of the whole plant kingdom. This unusually strong result could simply reflect the lack of studies at such a geographical and taxonomic scale, reflecting in turn the lack of supertrees until recently. Indeed, running the analysis at the genus level showed that many genera have low or very low phylogenetic signal (Supporting Information Figure S7). For instance, for the large genus *Solanum*, we observed phylogenetic signal reduced to levels comparable to the studies cited above (Supporting Information Figure S9), which indicates that, for some genera, this scale is too restricted to detect such signals. We also see a great deal

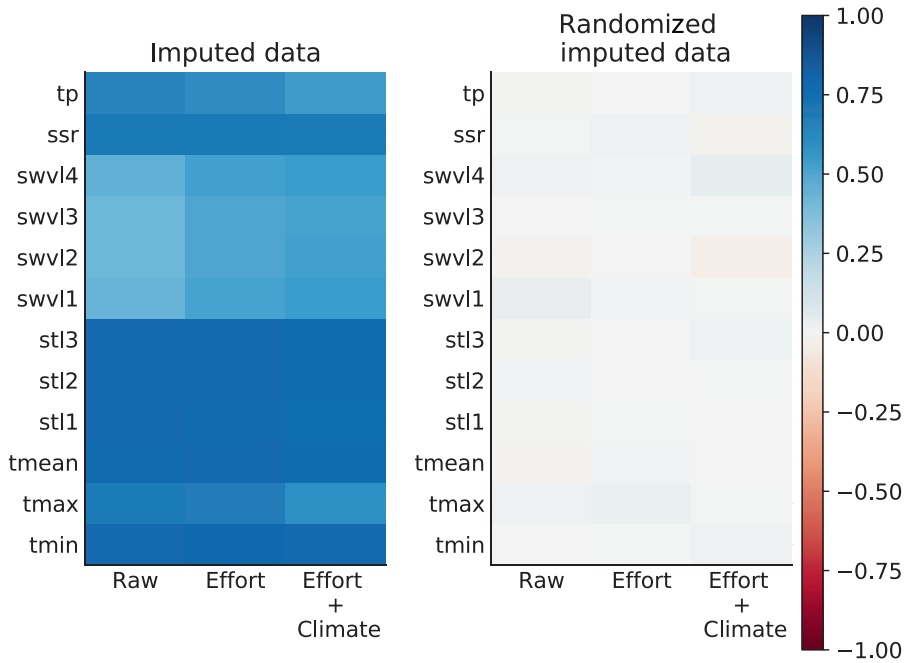

FIGURE 4 Pearson correlation between imputed and real data for each of the climate variables (for full parameter names, see Table 1).

of variation in phylogenetic signal when we look at individual genera, families, orders and classes, compared with the full tree. Some phylogenetic signals can be related to spatial autocorrelation between congenerics (Freckleton & Jetz, 2009), which would be of particular concern for rare and range-limited species. In this analysis, however, we consider the 5,000 most common species in GBIF occurrence records, for which >80% have ranges that span more than one continent, and supplementary analyses restricted to only those species with the broadest ranges yielded very similar results (Table S3).

The finding that minimum temperature produces the strongest signal is unsurprising, because numerous studies have suggested that responses to cold temperature extremes are an indicator of plant distributions and exhibit strong evolutionary conservatism (Currie et al., 2004; Qian et al., 2016; Woodward & Williams, 1987). This theory is often referred to as the tropical conservatism hypothesis (TCH) (Donoghue, 2008; Kerkhoff et al., 2014; Qian et al., 2016). In particular, freezing is thought to act as a strong selective force because it is often lethal, and mechanisms to withstand freezing are difficult to evolve (Qian et al., 2016). A Brownian motion model seems to capture this phylogenetic signal well, although we might expect this trait to respond better to discrete models of freezing tolerance. We also expected that solar radiation would show a strong phylogenetic signal because light is a limiting factor in plant growth, and we observed this, although there is little evidence in the literature of this being tested before. Minimum temperature also showed the strongest correlation with phylogenetic distance closest to the tips, although it remained significantly positive throughout the tree until *c*. 500 Ma. Using local Moran's *I*, we see that there is heterogeneity in autocorrelation for minimum temperature across genera, but there are still large clades showing strong signals across the tree, and from the autocorrelogram we expect that this index would show even stronger correlation across the plant kingdom at species level. We also observe strong heterogeneities in phylogenetic signal for minimum temperature, particularly at the genus and family levels (Supporting Information Figure S7).

In contrast, soil water volume produced a lower, but still strong, phylogenetic signal in most cases. There is evidence that water availability is as great a driver of species and phylogenetic diversity as temperature (Qian et al., 2016; Silvertown et al., 2015), and it is expected to act as a selective pressure on plant communities. It has been hypothesized that water-related traits in plants undergo rapid local evolution and are therefore much more labile than temperature adaptations (Arène et al., 2017). Indeed, Arène et al. (2017) have reported that the base temperature at which development takes place in plants has a strong phylogenetic signal, whereas the base water potential, a measure of water moisture, does not. Although there seems to be little explanation of this phenomenon in the literature, the strong evidence of a phylogenetic signal in average precipitation and the small positive correlation between these two variables suggests that water does play a role as a selective force in plant evolution (Brodribb et al., 2013, 2014; McAdam & Brodribb, 2012). Therefore, the reduced signal in soil water volume could instead indicate that the scale at which we extracted the data (80 km for ERA-Interim) is simply too coarse, particularly in areas such as Africa with high effort correction, because this correction weakened the signal uniquely for soil water volume in Table 3. However, perhaps more pertinently, ECMWF raise questions about the quality of soil moisture data in ERA-Interim [Copernicus Climate Change Service (C3S), 2021], stating that soil moisture values are only intended to "provide a qualitative picture of major anomalies", with plans to improve their estimates in ERA-5. Failure to identify strong and/or consistent phylogenetic signals in soil moisture variables might therefore simply reflect poor data quality.

The phylogenetic analyses we performed here were necessarily limited by the size of the tree (*c*. 26,000 species) and the quality of records available (either restricted to the top 5,000 species with ≥1,000 records or including species with very few records). However, the final dataset of bioclimatic envelopes includes >200,000 species with GBIF records in the past century, and we make this available as a resource for further exploration of the effect of climate change on plant species world-wide. We have already made use of such information to parameterize dynamic models of plant biodiversity across the continent of Africa for the past century (C. L. Harris, 2019). We expect that these bioclimatic envelopes could be used to drive other types of vegetation model, including dynamic global vegetation models (Scheiter et al., 2013) and forest gap models (Shugart et al., 2018), and could be used as a comparison to the output of species distribution models, which typically use WorldClim data. Information on the climates that plants can tolerate is important for understanding how different species might respond to future change, how well mitigation strategies might work, and the interaction between climate and other threats to biodiversity, such as invasive species and habitat loss. Although these bioclimatic envelopes are only a subsection of the fundamental niche that these plant species could occupy, we expect that their global scope and long temporal scales mean that their realized niches approach the fundamental in many cases (Araújo & Peterson, 2012).

GBIF plant occurrence records span hundreds of years and at least half of the estimated 400,000 species world-wide. They also suffer from numerous taxonomic, geographical and temporal biases (Meyer et al., 2016). Although there are now automatic algorithms to correct for obvious simple errors in georeferencing (CODATA, 2020), such as swapped coordinate signs, further corrections are needed to account for more complex biases, such as global collection effort. We made several corrections to the data used in the present study, including the use of proxies for plant density to adjust the weight of different records, which resulted in a corresponding increase in phylogenetic signal. Given that there is no reason why phylogenetic signal should be boosted randomly by such corrections, as is evidenced when we randomize the tips, this approach is a candidate for selecting other, future analyses. There are limitations to this approach, especially for species with few or no georeferenced coordinates, for which data imputation will be necessary. We also explored correction for biases in both world-wide and continent-level climates, but our simple approaches failed to improve the phylogenetic signal. In light of this, we suggest that any climate correction should be applied

to species individually to account for their specific available climate. This would require detailed information on dispersal and other barriers to availability of climate, such as geographical boundaries, and as such, is beyond the scope of the present study. However, when trying to account for spatial autocorrelation between species in the dataset, we did find similarly strong phylogenetic signals when we filtered for the most widespread species (those found on more than one continent). Information on processes such as dispersal and the physiological limits of species would help further to disentangle the factors contributing to the strong phylogenetic signal we see here and the potential for PNC amongst global plant species.

Given the strong correlation between the ECWMF data and WorldClim, it is unsurprising that there is a similarly strong phylogenetic signal in temperature and precipitation in the WorldClim data. However, given that ECMWF has a much broader range of climate parameters, it remains the obvious choice for this type of analysis. The ideal next step for such research would be to incorporate all data points for all species, rather than using averages of climate bins for subsets of species in the dataset. Using phylogenetic mixed models for this would both facilitate the extraction of climate profiles directly from estimated distributions and account better for more of the heterogeneity in the data (de Villemereuil & Nakagawa, 2014). Importantly, the simplest versions of these models are equivalent to Pagel's λ, allowing comparison with this work (de Villemereuil & Nakagawa, 2014; Freckleton et al., 2002; Housworth et al., 2004). As yet, these analyses are far too computationally intensive for the *c*. 26,000 species in the dataset, never mind for the 200,000 species that could be added to a partially resolved supertree. Further investigation of the feasibility of data imputation is also necessary for using these phylogenetic signals to impute data for missing or rare species, although we find that there is no systematic under- or over-estimation of minimum temperature as tested in the present study (Supporting Information Figure S8).

5 | CONCLUSION

Global plant species will face many coming threats this century, the most devastating of which is likely to be climate change. Understanding the relationship between plants and their climate is fundamental both to prediction of their future ranges and to the implementation of effective conservation strategies. Here, we have demonstrated that information on plant bioclimatic envelopes could be extracted reliably using historical records and climate reconstructions, which could then be used in further analyses, and we provide these variables for all 200,000 species as a public resource alongside this publication. We found a very strong phylogenetic signal in many climatic parameters, including temperature, soil temperature, solar radiation and precipitation. This evolutionary signal was improved by the implementation of a correction for the bias in collection effort world-wide and can also be used to impute data for related missing species. Future analyses could explore further the evidence for niche conservatism for climatic parameters at the supertree level.

ACKNOWLEDGMENTS

C.H. was supported by an Engineering and Physical Sciences Research Council (EPSRC) studentship (EP/M506539/1 ref. 1654080). C.H., N.B., C.A.C. and R.R. were supported by Natural Environment Research Council (NERC) grants NE/T004193/1 and NE/T010355/1. We would like to thank Anna Harper, Roderic Page and Tiina Sarkinen for their insights into the paper.

DATA AVAILABILITY STATEMENT

The source code for the analysis is publicly available under an open source licence on GitHub at: https://github.com/boydorr/ClimatePref.jl (C. Harris, 2022). The data used in the analysis are already publicly available (from Qian et al., 2016 and GBIF). The trait data generated by the paper have been published through NERC EDS Environmental Information Data Centre (Harris et al., 2022).

ORCID

Claire Harris https://orcid.org/0000-0003-0852-2340
Neil Brummitt https://orcid.org/0000-0001-7769-4395
Christina A. Cobbold https://orcid.org/0000-0001-8814-7688
Richard Reeve https://orcid.org/0000-0003-2589-8091

REFERENCES

Antonelli, A., Fry, C., Smith, R. J., Simmonds, M. S. J., Kersey, P. J., Pritchard, H. W., Abbo, M. S., Acedo, C., Adams, J., Ainsworth, A. M., Allkin, B., Annecke, W., Bachman, S. P., Bacon, K., Bárrios, S., Barstow, C., Battison, A., Bell, E., Bensusan, K., … Zhang, B. G. (2020). *State of the world's plants and fungi 2020 (Tech. Rep.)*. Royal Botanic Gardens. https://doi.org/10.34885/172

Araújo, M. B., & Peterson, A. T. (2012). Uses and misuses of bioclimatic envelope modeling. *Ecology*, *93*(7), 1527–1539. https://doi.org/10.2307/23225219

Arène, F., Affre, L., Doxa, A., & Saatkamp, A. (2017). Temperature but not moisture response of germination shows phylogenetic constraints while both interact with seed mass and lifespan. *Seed Science Research*, *27*(2), 110–120. https://doi.org/10.1017/S0960258517000083

Basnett, S., Nagaraju, S. K., Ravikanth, G., & Devy, S. M. (2019). Inuence of phylogeny and abiotic factors varies across early and late reproductive phenology of Himalayan Rhododen drons. *Ecosphere*, *10*(1), 1–15. https://doi.org/10.1002/ecs2.2581

Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM Review*, *59*(1), 65–98. https://doi.org/10.1137/141000671

BGCI. (2019). *GardenSearch online database*. Botanical Gardens Conservation International. https://tools.bgci.org/garden_search.php

Blomberg, S. P., Garland, T., & Ives, A. R. (2003). Testing for phylogenetic signal in comparative data: Behavioral traits are more labile. *Evolution*, *57*(4), 717–745. https://doi.org/10.1111/j.0014-3820.2003.tb00285.x

Box, E. O. (1996). Plant functional types and climate at the global scale. *Journal of Vegetation Science*, *7*(3), 309–320. https://doi.org/10.2307/3236274

Brodribb, T. J., Bowman, D. M., Grierson, P. F., Murphy, B. P., Nichols, S., & Prior, L. D. (2013). Conservative water management in the widespread conifer genus Callitris. *AoB PLANTS*, *5*, 1–11. https://doi.org/10.1093/aobpla/plt052

Brodribb, T. J., McAdam, S. A., Jordan, G. J., & Martins, S. C. (2014). Conifer species adapt to low-rainfall climates by following one of

two divergent pathways. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(40), 14489–14493. https://doi.org/10.1073/pnas.1407930111

Brummitt, N., Bachman, S. P., Griffths-Lee, J., Lutz, M., Moat, J. F., Farjon, A., Donaldson, J. S., Hilton-Taylor, C., Meagher, T. R., Albuquerque, S., Aletrari, E., Andrews, A. K., Atchison, G., Baloch, E., Barlozzini, B., Brunazzi, A., Carretero, J., Celesti, M., Chadburn, H., … Nic Lughadha, E. M. (2015, aug). Green plants in the red: A baseline global assessment for the IUCN sampled red list index for plants. *Plos One*, *10*(8), e0135152. https://doi.org/10.1371/journal.pone.0135152

Cheng, L., Abraham, J., Hausfather, Z., & Trenberth, K. E. (2019). How fast are the oceans warming? *Science*, *363*(6423), 128–129. https://doi.org/10.1126/science.aav7619

CODATA (2020). The Committee on Data of the International Science Council. In H. Pfeiffenberger & P. Uhlir (Eds.), *Twenty-year review of GBIF (Tech. Rep.)*. CODATA. https://doi.org/10.35035/ctzm-hz97

Copernicus Climate Change Service (C3S). (2021). Monthly summaries of precipitation, relative humidity and soil moisture. *Climate Bulletin* https://climate.copernicus.eu/monthly-summaries-precipitation-relative-humidity-and-soil-moisture

Crisp, M. D., & Cook, L. G. (2012). Phylogenetic niche conservatism: What are the underlying evolutionary and ecological causes? *New Phytologist*, *196*(3), 681–694. https://doi.org/10.1111/j.1469-8137.2012.04298.x

Currie, D. J., Mittelbach, G. G., Cornell, H. V., Field, R., Guégan, J. F., Hawkins, B. A., Kaufman, D. M., Kerr, J. T., Oberdorff, T., O'Brien, E., & Turner, J. R. (2004). Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecology Letters*, *7*(12), 1121–1134. https://doi.org/10.1111/j.1461-0248.2004.00671.x

Curtis, C. A., & Bradley, B. A. (2016). Plant distribution data show broader climatic limits than expert-based climatic tolerance estimates. *PLoS One*, *11*(11), 1–15. https://doi.org/10.1371/journal.pone.0166407

Davies, T. J., Wolkovich, E. M., Kraft, N. J., Salamin, N., Allen, J. M., Ault, T. R., Betancourt, J. L., Bolmgren, K., Cleland, E. E., Cook, B. I., Crimmins, T. M., Mazer, S. J., McCabe, G. J., Pau, S., Regetz, J., Schwartz, M. D., & Travers, S. E. (2013). Phylogenetic conservatism in plant phenology. *Journal of Ecology*, *101*(6), 1520–1530. https://doi.org/10.1111/1365-2745.12154

de Villemereuil, P., & Nakagawa, S. (2014). Chapter 11 general quantitative genetic methods for comparative biology. In L. Z. Garamszegi (Ed.), *Modern phylogenetic comparative methods and their application in evolutionary biology* (pp. 287–303). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-43550-2

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., … Vitart, F. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. https://doi.org/10.1002/qj.828

Donoghue, M. J. (2008). 13, a phylogenetic perspective on the distribution of plant diversity. In J. C. Avise, S. P. Hubbell, & J. Francisco (Eds.), *In the light of evolution, volume ii: Biodiversity and extinction*. (Vol. *II*, pp. 247–262). National Academies Press.

EASAC. (2018). *Extreme weather events in Europe. Preparing for climate change adaptation: an update on EASAC's 2013 study (Tech. Rep. No. March)*. EASAC Secretariat.

Feeley, K. (2015). Are we filling the data void? An assessment of the amount and extent of plant collection records and census data available for tropical South America. *PLOS One*, *10*(4), 1–17 . https://doi.org/10.1371/journal.pone.0125629

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315. https://doi.org/10.1002/joc.5086

Freckleton, R. P., Harvey, P. H., & Pagel, M. (2002). Phylogenetic analysis and comparative data: A test and review of evidence. *American Naturalist*, *160*(6), 712–726. https://doi.org/10.1086/343873

Freckleton, R. P., & Jetz, W. (2009). Space versus phylogeny: Disentangling phylogenetic and spatial signals in comparative data. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1654), 21–30. https://doi.org/10.1098/rspb.2008.0905

GBIF. (2019). *Global biodiversity information facility*. GBIF. https://www.gbif.org

Gibson, H., & Weiss, D. (2015). *Oxford MAP EVI: Malaria atlas project gap-filled enhanced vegetation index*. https://developers.google.com/earth-engine/datasets/catalog/Oxford_MAP_EVI_5km_Monthly#descriptio

Guerrero-Ramírez, N. R., Mommer, L., Freschet, G. T., Iversen, C. M., McCormack, M. L., Kattge, J., Poorter, H., van der Plas, F., Bergmann, J., Kuyper, T. W., York, L. M., Bruelheide, H., Laughlin, D. C., Meier, I. C., Roumet, C., Semchenko, M., Sweeney, C. J., van Ruijven, J., Valverde-Barrantes, O. J., … Weigelt, A. (2020). Global root traits (GRooT) database. *Global Ecology and Biogeography*, *1–13*, 25–37. https://doi.org/10.1111/geb.13179

Harbert, R. S., & Nixon, K. C. (2015). Climate reconstruction analysis using coexistence likelihood estimation (CRACLE): A method for the estimation of climate using vegetation. *American Journal of Botany*, *102*(8), 1277–1289. https://doi.org/10.3732/ajb.1400500

Harmon, L. J. (2019). *Phylogenetic comparative methods: Learning from trees*. CreateSpace Independent Publishing Platform.

Harris, C. (2022). *ClimatePref.jl. Zenodo*. https://doi.org/10.5281/zenodo.6452065

Harris, C., Brummitt, N., Cobbold, C., & Reeve, R. (2022). *Bioclimatic envelopes for global plant species*. NERC EDS Environmental Information Data Centre. https://doi.org/10.5285/ca339c86-3674-4030-b891-35326e71141e

Harris, C. L. (2019). *Simulating global plant biodiversity (Unpublished doctoral dissertation)*. University of Glasgow.

Harvey, P. H., & Pagel, M. (1991). *The comparative method in evolutionary biology*. Oxford University Press.

Housworth, E. A., Martins, E. P., & Lynch, M. (2004). The phylogenetic mixed model. *American Naturalist*, *163*(1), 84–96. https://doi.org/10.1086/380570

Huete, A., Justice, C., & van Leeuwen, W. (1999). *Modis vegetation index (Mod 13) algorithm theoretical basis document (Tech. Rep. No. 3)*. University of Arizona.

Hutchinson, G. E. (1957). Concluding remarks and future work. *Cold Spring Harbor Symposia on Quantitative Biology*, *22*, 415–427.

Intergovernmental Panel on Climate Change (2018). IPCC, 2018: Summary for policymakers. In V. Masson-Delmotte, et al. (Eds.), *Global warming of 1.5 c. an ipcc special report on the impacts of global warming of 1.5 c above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change* (pp. 1–32). World Meteorological Organization.

IPCC. (2014). *Climate change 2014: Synthesis report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (Tech. Rep.)*. IPCC.

Jentsch, A., & Beierkuhnlein, C. (2008). Research frontiers in climate change: Effects of extreme meteorological events on ecosystems. *Comptes Rendus - Geoscience*, *340*(9–10), 621–628. https://doi.org/10.1016/j.crte.2008.07.002

Kattge, J., Diaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Bonisch, G., Garnier, E., Westoby, M., Reich, P. B., Wright, I. J., Cornelissen, J. H. C., Violle, C., Harrison, S. P., van Bodegom, P. M., Reichstein, M., Enquist, B. J., Soudzilovskaia, N. A., Ackerly, D. D., Anand, M., … Wirth, C. (2011). TRY - a global database of plant traits. *Global Change Biology*, *17*(9), 2905–2935. https://doi.org/10.1111/j.1365-2486.2011.02451.x

Keck, F., Rimet, F., Bouchez, A., & Franc, A. (2016). Phylosignal: An R package to measure, test, and explore the phylogenetic signal. *Ecology and Evolution*, *6*(9), 2774–2780. https://doi.org/10.1002/ece3.2051

Kerkhoff, A. J., Moriarty, P. E., & Weiser, M. D. (2014). The latitudinal species richness gradient in New World woody angiosperms is consistent with the tropical conservatism hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(22), 8125–8130. https://doi.org/10.1073/pnas.1308932111

Koski, M. H., & Ashman, T. L. (2016). Macroevolutionary patterns of ultraviolet oral pig mentation explained by geography and associated bioclimatic factors. *The New Phytologist*, *211*(2), 708–718. https://doi.org/10.1111/nph.13921

Laloyaux, P., de Boisseson, E., Balmaseda, M., Bidlot, J. R., Broennimann, S., Buizza, R., Dalhgren, P., Dee, D., Haimberger, L., Hersbach, H., Kosaka, Y., Martin, M., Poli, P., Rayner, N., Rustemeier, E., & Schepers, D. (2018). CERA-20C: A coupled reanalysis of the twentieth century. *Journal of Advances in Modeling Earth Systems*, *10*(5), 1172–1195. https://doi.org/10.1029/2018MS001273

Li, D., Ives, A. R., & Waller, D. M. (2017). Can functional traits explain phylogenetic signal in community composition? *New Phytologist*, *214*(2), 607–618. https://doi.org/10.1111/nph.14397

Liu, H., Xu, Q., He, P., Santiago, L. S., Yang, K., & Ye, Q. (2015). Strong phylogenetic signals and phylogenetic niche conservatism in ecophysiological traits across divergent lineages of Magnoliaceae. *Scientific Reports*, *5*(12246), 1–12. https://doi.org/10.1038/srep12246

Losos, J. B. (2011). Seeing the forest for the trees: The limitations of phylogenies in comparative biology. *American Naturalist*, *177*(6), 709–727. https://doi.org/10.1086/660020

Marvel, K., Cook, B. I., Bonfils, C. J., Durack, P. J., Smerdon, J. E., & Williams, A. P. (2019). Twentieth-century hydroclimate changes consistent with human inuence. *Nature*, *569*(7754), 59–65. https://doi.org/10.1038/s41586-019-1149-8

McAdam, S. A., & Brodribb, T. J. (2012). Stomatal innovation and the rise of seed plants. *Ecology Letters*, *15*(1), 1–8. https://doi.org/10.1111/j.1461-0248.2011.01700.x

Meyer, C., Weigelt, P., & Kreft, H. (2016). Multidimensional biases, gaps and uncertainties in global plant occurrence information. *Ecology Letters*, *19*(8), 992–1006. https://doi.org/10.1111/ele.12624

Molina-Venegas, R., & Rodríguez, M. (2017). Revisiting phylogenetic signal; strong or negligible impacts of polytomies and branch length information? *BMC Evolutionary Biology*, *17*(1), 1–10. https://doi.org/10.1186/s12862-017-0898-y

Muünkemüller, T., Boucher, F. C., Thuiller, W., & Lavergne, S. (2015). Phylogenetic niche conservatism - common pitfalls and ways forward. *Functional Ecology*, *29*(5), 627–639. https://doi.org/10.1111/1365-2435.12388

Nunez, S., Arets, E., Alkemade, R., Verwer, C., & Leemans, R. (2019). Assessing the impacts of climate change on biodiversity: Is below 2°C enough? *Climatic Change*, *154*(3–4), 351–365. https://doi.org/10.1007/s10584-019-02420-x

Olhoff, A., & Christensen, J. (Eds.). (2019). *Emissions gap report 2019*. United Nations Environment Programme.

Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature*, *401*(6756), 877–884. https://doi.org/10.1038/44766

Pearson, R. G., & Dawson, T. P. (2003). Predicting the impacts of climate change on the distribution of species: Are bioclimate envelope models useful? *Global Ecology and Biogeography*, *12*(5), 361–371. https://doi.org/10.1046/j.1466-822X.2003.00042.x

Qian, H., Field, R., Zhang, J. L., Zhang, J., & Chen, S. (2016). Phylogenetic structure and ecological and evolutionary determinants of species richness for angiosperm trees in forest communities in China. *Journal of Biogeography*, *43*(3), 603–615. https://doi.org/10.1111/jbi.12639

Qian, H., & Jin, Y. (2016). An updated megaphylogeny of plants, a tool for generating plant phylogenies and an analysis of phylogenetic community structure. *Journal of Plant Ecology*, *9*(2), 233–239. https://doi.org/10.1093/jpe/rtv047

Reyer, C., Leuzinger, S., Rammig, A., Wolf, A., Bartholomeus, R. P., & Bonfante, A. (2013). A plant's perspective of extremes: Terrestrial plant responses to changing climatic variability. *Global Change Biology*, *19*(1), 75–89. https://doi.org/10.1111/gcb.12023.A

Scheiter, S., Langan, L., & Higgins, S. I. (2013). Next-generation dynamic global vegetation models: Learning from community ecology. *New Phytologist*, *198*(3), 957–969. https://doi.org/10.1111/nph.12210

Schreeg, L. A., Kress, W. J., Erickson, D. L., & Swenson, N. G. (2010). Phylogenetic analysis of local-scale tree soil associations in a lowland moist tropical forest. *PLoS One*, *5*(10), e13685. https://doi.org/10.1371/journal.pone.0013685

Schwalm, C. R., Glendon, S., & Duffy, P. B. (2020). RCP8.5 tracks cumulative CO_2 emissions. *Proceedings of the National Academy of Sciences*, *117*(33), 19656–19657. https://doi.org/10.1073/pnas.2007117117

Shugart, H. H., Wang, B., Fischer, R., Ma, J., Fang, J., Yan, X., Huth, A., & Armstrong, A. H. (2018). Gap models and their individual-based relatives in the assessment of the consequences of global change. *Environmental Research Letters*, *13*(3), 1–17. https://doi.org/10.1088/1748-9326/aaaacc

Silvertown, J., Araya, Y., & Gowing, D. (2015). Hydrological niches in terrestrial plant com munities: A review. *Journal of Ecology*, *103*(1), 93–108. https://doi.org/10.1111/1365-2745.12332

Slater, T., Hogg, A. E., & Mottram, R. (2020). Ice-sheet losses track high-end sea-level rise projections. *Nature Climate Change*. *10*, 879–881. https://doi.org/10.1038/s41558-020-0893-y

Solś-Lemus, C., Bastide, P., & Ané, C. (2017). PhyloNetworks: A package for phylogenetic net works. *Molecular Biology and Evolution*, *34*(12), 3292–3298. https://doi.org/10.1093/molbev/msx235

Sparrius, L., van den Top, G., & van Swaay, C. (2018). An approach to calculate a species temperature index for ora based on open data. *Gorteria*, *40*(1), 73–78.

Stahl, U., Reu, B., & Wirth, C. (2014). Predicting species' range limits from functional traits for the tree ora of North America. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(38), 13739–13744. https://doi.org/10.1073/pnas.1300673111

Steinbauer, M. J., Field, R., Fernández-Palacios, J. M., Irl, S. D., Otto, R., Schaefer, H., & Beierkuhnlein, C. (2016). Biogeographic ranges do not support niche theory in radiating Canary Island plant clades. *Global Ecology and Biogeography*, *25*(7), 792–804. https://doi.org/10.1111/geb.12425

TPL. (2013). *The plant list version 1.1*. Published on the Internet; http://www.theplantlist.org/

UNFCCC. (2015). *The Paris agreement (Tech. Rep.)*. Author FCCC/CP/2015/L.9/Rev.1.

Warren, R., Price, J., Graham, E., Forstenhaeusler, N., & VanDerWal, J. (2018, may). The projected effect on insects, vertebrates, and plants of limiting global warming to 1.5 C rather than 2°C. *Science*, *360*(6390), 791–795. https://doi.org/10.1126/science.aar3646

Weigelt, P., Koönig, C., & Kreft, H. (2019). GIFT–A Global Inventory of Floras and Traits for macroecology and biogeography. *Journal of Biogeography*, *1–28*, 16–43. https://doi.org/10.1111/jbi.13623

Wiens, J. J. (2016). Climate-related local extinctions are already widespread among plant and animal species. *PLoS Biology*, *14*(12), 1–18. https://doi.org/10.1371/journal.pbio.2001104

Woodward, F. I., & Williams, B. G. (1987). Climate and plant distribution at global and local scales. *Vegetatio*, *69*(1–3), 189–197. https://doi.org/10.1007/BF00038700

Xu, X., Dimitrov, D., Shrestha, N., Rahbek, C., & Wang, Z. (2019). A consistent species richness–climate relationship for oaks across the Northern Hemisphere. *Global Ecology and Biogeography*, *28*(8), 1051–1066. https://doi.org/10.1111/geb.12913

Zanne, A. E., Tank, D. C., Cornwell, W. K., Eastman, J. M., Smith, S. A., FitzJohn, R. G., McGlinn, D. J., O'Meara, B. C., Moles, A. T., Reich, P. B., Royer, D. L., Soltis, D. E., Stevens, P. F., Westoby, M., Wright, I. J., Aarssen, L., Bertin, R. I., Calaminus, A., Govaerts, R., … Beaulieu, J. M. (2014). Three keys to the radiation of angiosperms

into freezing environments. *Nature*, *506*(7486), 89–92. https://doi.org/10.1038/nature12872

Zhang, C., Yang, J., Sha, L., Ci, X., Li, J., Cao, M., Brown, C., Swensonb, N. G., & Lin, L. (2017). Lack of phylogenetic signals within environmental niches of tropical tree species across life stages. *Scientific Reports*, *7*, 1–5. https://doi.org/10.1038/srep42007

BIOSKETCH

The authors are all affiliated to the Boyd Orr Centre for Population and Ecosystem Health at the University of Glasgow. Their research covers the measurement and modelling of biodiversity, with particular focus on the plant kingdom in the face of climate and land-use change.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Harris, C., Brummitt, N., Cobbold, C. A., & Reeve, R. (2022). Strong phylogenetic signals in global plant bioclimatic envelopes. *Global Ecology and Biogeography*, *31*, 2191–2203. https://doi.org/10.1111/geb.13564