Tallo: A global tree allometry and crown architecture database

Tommaso Jucker1 | Fabian Jörg Fischer1 | Jérôme Chave2,3 | David A. Coomes4 | John Caspersen5 | Arshad Ali6 | Grace Jopaul Loubota Panzou7,8 | Ted R. Feldpausch9 | Daniel Falster10 | Vladimir A. Usoltsev11,12 | Stephen Adu-Bredu13 | Luciana F. Alves14 | Mohammad Aminpour15 | Ilondea B. Angoboy16 | Niels P. R. Anten17 | Cécile Antin18 | Yousef Askari19 | Rodrigo Muñoz20,21 | Narayanan Ayyappan22 | Patricia Balvanera23 | Lindsay Banin24 | Nicolas Barbier25 | John J. Battles25 | Hans Beeckman26 | Yannick E. Bocko8 | Ben Bond-Lamberty27 | Frans Bongers21 | Samuel Bowers28 | Thomas Brade28 | Michiel van Breugel29,30,31 | Arthur Chantrain7 | Rajeev Chaudhary32 | Jingyu Dai33 | Michele Dalponte34 | Kangbêni Dimobe35 | Jean-Christophe Domec36,37 | Jean-Louis Doucet7 | Remko A. Duursma38 | Moisés Enríquez20 | Karin Y. van Ewijk39 | William Farfán-Rios40 | Adeline Fayolle7 | Eric Forni41 | David I. Forrester42 | Hammad Gilani43 | John L. Godlee28 | Sylvie Gourlet-Fleury41 | Matthias Haeni44 | Jefferson S. Hall30 | Jie-Kun He45 | Andreas Hemp46 | José L. Hernández-Stefanoni47 | Steven I. Higgins48 | Robert J. Holdaway49 | Kiramat Hussain50 | Lindsay B. Hutley51 | Tomoaki Ichie52 | Yoshiko Iida53 | Hai-sheng Jiang45 | Puspa Raj Joshi54 | Hasan Kaboli55 | Maryam Kazempour Larsary56 | Tanaka Kenzo57 | Brian D. Kloeppe58,59 | Takashi Kohyama60 | Suwash Kunwar32,61 | Shem Kuyah62 | Jakub Kvasnica63 | Siliang Lin64 | Emily R. Lines65 | Hongyan Liu33 | Craig Lorimer66 | Jean-Joël Loumoto8 | Yadvinder Malhi67 | Peter L. Marshall58 | Eskil Mattsson69,70 | Radim Matula71 | Jorge A. Meave20 | Sylvanus Mensah72 | Xiangcheng Mi73 | Stéphane Momot18,74 | Glenn R. Moncrieff75,76 | Francisco Mora23 | Sarath P. Nissanka77 | Kevin L. O’Hara25 | Steven Pearce78 | Raphaël Pelissier18 | Pablo L. Peri79 | Pierre Ploton18 | Lourens Poorter21 | Mohsen Javanniri Pour80 | Hassan Pourbabaei56 | Juan Manuel Dupuy-Rada47 | Sabina C. Ribeiro81 | Casey Ryan28 | Anvar Sanaei82 | Jennifer Sanger78 | Michael Schlund83 | Giacomo Sellan84,85 | Alexander Shenkin67 | Bonaventure Sonké74 | Frank J. Sterck21 | Martin Svátek63 | Kentaro Takagi86 | Anna T. Trugman87 | Farman Ullah6,61 | Matthew A. Vadeboncoeur88 | Ahmad Valipour89 | Mark C. Vanderwel90 | Alejandra G. Vovides91 | Weiwei Wang73 | Li-Qiu Wang61 | Christian Wirth92,93 | Murray Woods94 | Wenhua Xiang95 |
1 | INTRODUCTION

Trees vary enormously in the size and shape of their crowns, and accurately capturing and describing this incredible variation in tree architecture is central to numerous lines of ecological research (Verbeeck et al., 2019). For instance, data capturing the relationship between the stem diameter, height and crown radius of trees have been used to test theory linking body size and metabolism across ecological scales (Coomes et al., 2012; Enquist et al., 2009; Shenkin et al., 2020), as well as exploring the ecological, environmental and evolutionary constraints that shape allometric scaling relationships of woody plants (Banin et al., 2012; Jucker et al., 2015; Lines et al., 2012; Loubota Panzou et al., 2021). These data also underpin efforts to develop more accurate and generalizable models for estimating forest biomass stocks and their uncertainties (Chave et al., 2014; Goodman et al., 2014; Jucker et al., 2017; Ploton et al., 2016). Moreover, tree height and crown size data are increasingly being used to bridge the gap between remote sensing and traditional field ecology (Aguirre-Gutiérrez et al., 2015; Jucker et al., 2017; Marconi et al., 2021), including facilitating the integration of remote sensing data into individual-based models of forest structure and dynamics (Fischer et al., 2019, 2020; Taubert et al., 2015). However, because basic properties of tree size, such as their height and crown dimensions, are challenging and time-consuming to measure accurately on the ground, limited access to curated tree crown architectural data is often a real barrier to progress in these fields.
Building on previous efforts to compile regional and global tree allometry databases (Falster et al., 2015; Feldpausch et al., 2011; Jucker et al., 2017; Loubota Panzou et al., 2021), here we bring together the world’s largest open access collection of trees for which stem diameter, height and/or crown radius have been measured—the Tallo database (Figure 1). Tallo includes nearly 500,000 georeferenced and taxonomically standardized records from more than 5000 tree species acquired at over 60,000 sites worldwide, including data from all major terrestrial biomes and some of the world’s largest ever recorded trees. After describing the key steps involved in the acquisition and standardization of the data, we showcase some of the potential applications of the Tallo database through a series of three case studies.

2 | COMPILING THE DATABASE

2.1 | Data aggregation

We compiled data on trees for which stem diameter (D, in cm), total tree height (H, in m) and/or crown radius (CR, in m) were measured. D was measured at breast height (1.3 m aboveground) or otherwise just above buttress roots using either diameter tape or callipers. As is common practise, for multi-stemmed trees a single pooled value of D was calculated by summing the diameter values of all individual stems (Di) using the quadratic diameter: \( D = \sqrt{\sum D_i^2} \) (de Souza et al., 2021; Paul et al., 2016). While care was taken to identify records from multi-stemmed individual, it is possible that for records compiled from existing databases a small number of multi-stemmed trees were mistakenly treated as separate individuals.

H and CR—which we mostly derived from 2 to 8 orthogonal crown radii measurements or otherwise from crown projection areas—were measured using a variety of approaches, including laser or ultrasonic range finders, clinometers, as well as tape measures and telescopic poles for smaller trees. For a very small subset of trees with fully sun-exposed crowns, H and CR were measured using a combination of high-resolution aerial photos and airborne LiDAR (Cano et al., 2019). Previous work comparing tree height measurements derived using laser range finders and clinometers—the two most common methods used to take tree biometric measurements—has shown that the two approaches provide consistent estimates, with laser rangefinders allowing for greater precision but with a tendency to slightly underestimate total tree heights (Larjavaara & Muller-Landau, 2013).

In addition to the data on tree size and shape, we also recorded the latitude and longitude of the site where each tree was measured, and any available taxonomic information. Data were obtained from a range of sources, including the published literature, online databases and unpublished data collected by co-authors of this study (see Appendix S1 in Supporting Information for a complete list of data sources). In compiling these data, we excluded records from heavily managed and industrial tree plantations, as well as agroforestry systems. Care was also taken to avoid double counting trees, which could occur either because the same data had been obtained from multiple sources or because trees were measured more than once as part of successive forest inventories. Specifically, for data obtained from public databases we made sure none had been compiled from the same primary sources. Additionally, for the small subset of trees that had been measured more than once, we only retained data from the most recent census.

2.2 | Taxonomy

Species names were cross-referenced and harmonized against those of The Plant List (TPL; http://www.theplantlist.org), using a combination of the taxonstand package in R (Cayuela et al., 2012; R Core Development Team, 2021) and the Taxonomic Name Resolution Service (Boyle et al., 2013). While TPL has been static since 2013, it was chosen as a reference as it remains widely used in ecology. Future versions of Tallo will align to the World Flora Online (http://www.worldfloraonline.org) as this becomes the new standard for plant taxonomy. Taxa that did not match TPL were reviewed manually and any misspellings or synonyms that had not been automatically detected were corrected. The small number of species for which no direct match to TPL was found (n = 43, 0.8% of the total) were checked against the Global Tree Search database (GTS; https://tools.bgci.org/global_tree_search.php), a curated list of over 60,000 plant species that meet the IUCN’s Global Tree Specialist Group definition of a tree. In all, 11 species (0.2% of the total) did not match either TPL or GTS, but because all of these could be traced back to published records online, we retained them in the database. Finally, genus names were used to assign each tree to its family and group trees into major divisions of vascular plants (i.e. angiosperms and gymnosperms) following the classification of Kew Royal Botanic Gardens (http://data.kew.org/vpfg1992/genfile.html).

Having standardized taxonomic names, we then removed any records from species that did not meet our working definition of trees—perennial woody seed plants with a single dominant stem, that are self-supporting and undergo secondary growth. This included removing all ferns, palms, lianas, strangler figs, bamboos, pandans, as well as a number of shrub species that rarely exceed 2 m in height and are generally multi-stemmed.

2.3 | Geographical coordinates

Each tree in the database is associated with a set of geographical coordinates, recorded in decimal degrees of latitude and longitude. These range in precision between 1 and 3 decimal places (approximately 0.1–11 km at the equator), with the majority of trees geolocated with a precision of ≤1 km. To facilitate the integration of the Tallo database with other large-scale spatial datasets, we used the CoordinateCleaner package in R to flag and correct common issues known to affect georeferenced records obtained from online databases (Zizka et al., 2019).
FIGURE 1  Overview of the Tallo database, including (a) geographical coverage, (b–c) size range of sampled trees, (d) climatic range of the data and (e) taxonomic coverage in phylogenetic space. Panel (a) shows the total number of trees recorded in grid cells of approximately 200×200 km. In (b–d), the density of overlapping points is reflected by a colour gradient ranging from black (low point density) to yellow (high point density). Data on mean annual rainfall and temperature show in (d) were obtained from WorldClim2 database (Fick & Hijmans, 2017) at a spatial resolution of 30 arc-seconds (approximately 1 km). Panel (e) shows a phylogenetic tree constructed from all species in the Tallo database (n = 5163). Branch tips have been colour coded to reflect the number of trees sampled for each species and the position of several seed plant families on the tree has been labelled. The phylogenetic tree was generated using the V.PhyloMaker package in R (Jin & Qian, 2019), the backbone of which is a phylogeny of 79,881 taxa of seed plants developed by Smith and Brown (2018).
First, we checked for any coordinates that were either invalid or had zero values for longitude or latitude. Data from two locations with a longitude of zero were retained after being checked against primary sources. Next, we checked for coordinates that did not fall on land by overlaying them onto a map of the world’s coastlines obtained from the Natural Earth database at 1:10, 1:50 and 1:110 million scales (https://www.naturalearthdata.com). A small number of locations (n = 13 sites, 0.02% of the total) were located at sea at all three scales. These records were all from the Balearic Islands in Spain and were manually corrected to the nearest land point using the 1:10 million scale map as a reference. Lastly, we also checked that all coordinates aligned to high-resolution (30arc-seconds, approximately 1 km) gridded climate data from the WorldClim2 database (Fick & Hijmans, 2017).

2.4 | Data quality control

As a quality control measure, we first removed any trees recorded as dead or damaged and then filtered the database to exclude trees with $D < 1 \text{ cm}$ and $H < 1.3 \text{ m}$. We then used Mahalanobis distance as implemented in the OutlierDetection package in R to identify trees with unrealistically large or small $H$ and $CR$ values given the size of their trunk, their biome association and their functional group (see Appendix S2 for details). These outliers could be the result of a data entry error (e.g. shift in decimal place or mistaken conversion between m, cm and mm) or possibly reflect a tree with substantial damage to its crown which went unrecorded. In total, 508 trees were identified as outliers based on $H$ (Figure S1) and a further 490 based on $CR$ (Figure S2). These records were retained in the Tallo database but flagged as outliers, allowing them to be easily removed by users depending on the application.

3 | DATABASE OVERVIEW AND ACCESS

The Tallo database includes a total of 498,838 trees from 61,856 unique sites across the world for which $D$, $H$ and/or $CR$ have been measured (Figure 1a), including 311,326 trees where all three dimensions are recorded. These represent all major forested and non-forested biomes (Table 1) and cover a gradient of over 45°C in mean annual temperature and more than 7000 mm in annual rainfall (Figure 1d). Trees in the database span multiple orders of magnitude in size (Figure 1b–e and Table 1), with $D$ ranging from 1.0 to 770.0 cm, $H$ from 1.3 to 115.8 m and $CR$ from 0.05 to 24.25 m. The majority of trees (88%) are identified to species, and overall the Tallo database includes records for 5163 species, 1453 genera and 187 plant families (Figure 1e).

The version of the Tallo database described in this article is publicly archived on Zenodo under a CC-BY 4.0 licence so that it can be freely used, shared and modified so long as appropriate credit is given (https://doi.org/10.5281/zenodo.6637599). Major version updates will be periodically uploaded to Zenodo, in addition to which we will also maintain an up-to-date version of the database on GitHub (https://github.com/selva-lab-repo/TALLO). Tallo should be referenced by citing this paper and users are also encouraged to report the version of the database they have accessed and to cite the original data sources whenever possible. The database is stored as a csv file which contains the individual tree morphological data, the geographical coordinates of each tree, any available taxonomic information, an identifier flagging any records classified as outliers and a reference code linking to the source from which records were obtained. A look-up table with full bibliographical sources is provided separately in Table S1 and as a csv file on the GitHub repository. Additionally, metadata files with a detailed description of each field in Tallo database can also be found on GitHub.

4 | CASE STUDIES

To showcase a few of the possible applications of the Tallo database, we developed three case studies that explore a range of theoretical and applied questions in ecology related to tree allometry. To enable users to replicate and build on these examples, all R code and ancillary environmental data used in the case studies have been archived on the GitHub repository.
4.1 Case study 1: Testing the predictions of metabolic scaling theory across biomes and functional groups

Metabolic scaling theory (MST) makes a number of predictions about how different axes of tree size should scale against one another based on first principles of plant water transport, branching architecture and stem biomechanics (Anderson-Teixeira et al., 2015; Shenkin et al., 2020; West et al., 2009). Specifically, both $H$ and $CR$ are expected to be proportional to the $\frac{2}{3}$ power of $D$ (i.e. $H \propto D^{\frac{2}{3}}$ and $CR \propto D^{\frac{2}{3}}$), while $CR$ is assumed to scale isometrically with $H$ (i.e. $CR \propto H^2$). However, there is substantial evidence that real-world scaling relationships can depart substantially and systematically from the theoretical predictions of MST due to the environmental context in which a tree is growing (e.g. water availability, competition for light, browsing, disturbance regime), as well as its evolutionary history (Jucker et al., 2017; Lines et al., 2012; Moncrieff et al., 2011; Muller-Landau et al., 2006; Shenkin et al., 2020).

Using Tallo, we tested whether crown allometric scaling relationships can be reconciled with the predictions of MST or if instead they vary systematically and predictably among major plant lineages (i.e. angiosperms and gymnosperms) and biome types. We modelled $H$-$D$, $CR$-$D$ and $CR$-$H$ scaling relationships using a power-law function by fitting linear mixed-effects regressions to log–log transformed data and allowing both the normalization constant (intercept) and scaling exponent (slope) to vary among biomes for both angiosperm and gymnosperm lineages. Models were fit using the lme4 package in R and took the following general form:

$$\text{lmer(} \log(Y) - \log(X) + (\log(X)|\text{Biome:Lineage})$$

Additionally, we also tested whether biome-level scaling exponents varied in relation to the degree of aridity experienced by trees within a biome, quantified as the ratio between mean annual precipitation and potential evapotranspiration. Aridity data for this analysis were obtained at 30-arc second resolution from the Global Aridity Index and PET Database (Trabucco & Zomer, 2019) and matched to each tree based on its geographical coordinates.

We found that agreement with MST varied considerably among the two plant lineages and biomes, as well as the allometric relationship being examined (Figure 2). Overall, observed $H$-$D$ scaling exponents were lower than the $\frac{2}{3}$ predicted by MST (exponent estimate $\pm$95% confidence intervals $= 0.581 \pm 0.068$). This was especially true for angiosperm trees, whose $H$-$D$ scaling exponents were on average substantially lower than those of gymnosperms across biomes ($0.537$ and $0.637$, respectively). However, for both lineages departure from MST was most pronounced in arid biomes, whereas $H$-$D$ scaling exponents of trees growing in non-water limited ecosystems such as tropical and temperate forests were consistent with the predictions of MST (Figure 2a). This trend was reflected in a significantly positive correlation between a biome’s $H$-$D$ scaling exponent and the mean aridity index experience by trees with that biome (Pearson’s correlation coefficient, $\rho = 0.56$, $p = 0.017$).

A similar picture emerged for $CR$-$D$, where once again observed scaling exponents were on average lower than those predicted by MST (Figure 2b), although in this case 95% confidence intervals of the parameter estimate overlapped with $\frac{2}{3}$ ($0.620 \pm 0.062$). On average, gymnosperms had higher $CR$-$D$ scaling exponents than angiosperms ($0.648$ and $0.597$, respectively) and trees growing in drylands had $CR$-$D$ scaling exponents that were furthest from those predicted by MST (although in this case the correlation with aridity was not statistically significant; $\rho = 0.35$, $p = 0.15$). By contrast, $CR$-$H$ scaling relationships showed a much bigger departure from MST predictions, with an overall scaling exponent well below 1 ($0.695 \pm 0.077$). This was true for both angiosperms and gymnosperms ($0.708$ and $0.676$ on average.
Light on the role that both species turnover and intraspecific plasticity across aridity gradients from which the former are derived (Shenkin et al., 2020). This suggests that while MST may serve as a useful starting point for understanding scaling relationships between different axes of tree size, at least some of its underlying assumptions need to be revisited.

4.2 Case study 2: Plasticity in height–diameter scaling relationships along aridity gradients

Trees adapt their size and shape to match the environment in which they grow (Jucker et al., 2015; Kafuti et al., 2022; Lines et al., 2012). A classic example is the fact that trees tend to be shorter for a given stem diameter in drier environments (Chave et al., 2014; Husfloen et al., 2015; Jucker et al., 2017; Lines et al., 2012; Vieilledent et al., 2012), as taller trees are generally at greater risk of hydraulic failure due to embolism and this effect is exacerbated when access to water becomes progressively more limiting (Domec et al., 2008; McDowell & Allen, 2015; Olson et al., 2018; Stovall et al., 2019). However, it remains unclear to what extent the negative relationship between tree height and aridity is driven by turnover in species composition along environmental gradients (i.e. species with shallower H–D scaling relationships dominating arid environments and vice versa) as opposed to plasticity and local adaptation within species (Lines et al., 2012).

To answer this question, we selected a subset of species in the Tallo database that had been sampled at multiple sites spanning a gradient in aridity (defined and quantified in the same way as the previous case study). Specifically, we only included records for species that (i) were found at two or more distinct sites with at least 10 individual trees sampled at each site, (ii) were recorded at locations with at least a 20% difference in aridity index between their most arid and humid site and (iii) spanned a size range of at least 20 cm in stem diameter. This left us with 155,022 trees belonging to 342 species that (i) were found at two or more distinct sites with at least 10 individuals trees sampled at each site, (ii) were recorded at locations with at least a 20% difference in aridity index between their most arid and humid site and (iii) spanned a size range of at least 20 cm in stem diameter. This left us with 155,022 trees belonging to 342 species (303 angiosperms and 39 gymnosperms). Using these data, we tested how H–D relationships (modelled on a log–log scale) vary along aridity gradients in relation to both species turnover and within-species plasticity by fitting the following linear mixed-effects model in the lme4 package in R:

\[
\text{lmer} (\text{log}(H) - \text{log}(D) + AISP + \text{AIGMC} + (\text{log}(D) + \text{AIGMC})|\text{Species}))
\]

where \(AISP\) is the mean aridity index value of each species and \(AIGMC\) is the group-mean centred aridity index value of each tree (calculated as the difference between each tree’s aridity index value and \(AISP\), the mean value of its species). The \(AISP\) term in the model tests whether tree species found in more arid environments tend to be shorter, for a given stem diameter, than those from more humid regions. Instead, \(AIGMC\) tests whether individuals within a species growing at the arid-end of their distribution (where \(AIGMC < 0\)) are shorter than those at the humid-end (where \(AIGMC > 0\)). The effects of both \(\text{log}(D)\) and \(AIGMC\) on \(\text{log}(H)\) were allowed to vary among species (i.e. random intercept and slopes model) and a permutation approach was used to generate 95% confidence intervals for the random slope terms of the model. This allowed us to determine which tree species exhibited significantly negative or positive shifts in height in response to rising aridity.

We found that aridity plays a key role in modulating the relationship between tree height and stem diameter, with trees growing in more arid environments generally much stouter than those from more humid climates (Figure 3). For example, a 30 cm diameter tree growing in a location where mean annual rainfall is only half of potential evapotranspiration (aridity index = 0.5) is on average 9.7 m shorter (~42%) than one growing where annual rainfall is double the evaporative demand (aridity index = 2). Standardized model coefficients for \(AISP\) (0.145 ± 0.026) and \(AIGMC\) (0.085 ± 0.021) were both significantly positive (\(p < .0001\)). This indicates that the strong effect of aridity on H–D scaling relationships is driven by a combination of both species turnover and intraspecific plasticity across aridity gradients, with the former playing a particularly important role.

While we found that decreasing aridity generally led to trees becoming more slender, this effect varied considerably among species. Most species (\(n = 241\), 70% of the total) tended to be taller at the humid end of their sampled distribution, with 44% exhibiting a significantly positive increase in height with decreasing aridity (i.e. lower bounds of the 95% confidence intervals of the random \(AIGMC\) slope > 0; blue arrows in Figure 3). However, we also observed a smaller proportion of species that exhibited the opposite trend (\(n = 38\), 11% of the total; red arrows in Figure 3). In relative terms, these were more likely to be gymnosperms (26% of species) than angiosperms (9% of species). Moreover, we found that species adapted to drier environments were generally more likely to respond positively to increased water availability in terms of investment in height growth compared to those from more humid climates (\(p = −0.18\), \(p = .0006\) when relating a species’ random \(AIGMC\) slope estimate to its \(AISP\)). Overall, our results confirm the importance of water availability in shaping H–D scaling relationships in trees and shed new light on the role that both species turnover and intraspecific plasticity play in driving these patterns.

4.3 Case study 3: Global maps of potential tree height under current and future climates

Large, tall trees play a disproportionately big role in shaping carbon cycling on land, as they store the vast majority of the aboveground biomass in a given patch of forest (Bastin et al., 2018; Lutz et al., 2018; Slik...
To trial this approach, we used the entire Tallo database to fit a multiple regression model in which we expressed tree height (\(\log(H)\); log-transformed) as a function of stem diameter (\(\log(D)\); log-transformed), aridity index (\(\log(\text{AI})\); log-transformed), precipitation seasonality (\(P_{\text{SEASON}}\)), mean annual temperature (\(T_{\text{MEAN}}\)), maximum annual temperature (\(T_{\text{MAX}}\)), and the interaction between \(\text{AI}\) and \(T_{\text{MEAN}}\):

\[
\log(H) = \log(D) + \log(\text{AI}) + P_{\text{SEASON}} + T_{\text{MEAN}} + T_{\text{MAX}} + \log(\text{AI})T_{\text{MEAN}}
\]

Climate data were obtained at 30-arc second resolution and assigned to each tree based on their geographical coordinates. \(\text{AI}\) values were taken from the Global Aridity Index and PET Database, while other climatic predictors were obtained from WorldClim2 (Fick & Hijmans, 2017). As our aim was to provide a proof of concept, climatic predictors were selected based on those identified by previous studies as playing a role in modifying \(H-D\) scaling relationships (Chave et al., 2014; Hulshof et al., 2015; Lines et al., 2012), rather than through an extensive model selection process.

The fitted model was used to generate spatially explicit predictions of tree height at global scale under both current and future climate scenarios at 5-arc minute spatial resolution (approximately 10 km) obtained from WorldClim2. Rather than using a single size threshold for defining ‘large trees’ across all ecosystem types, predictions were made using the 99th percentile stem diameter value of trees from each biome class in the Tallo database. These ranged from 33 cm in drylands to 95 cm in tropical rainforests. To explore how climate change might affect the distribution of tall trees, we used CMIP6 future climate projections for the period of 2061–2080 derived from the CNRM-ESM2-1 global climate model run under the shared socio-economic pathway (SSP) 245 (equivalent to RCP 4.5). To simplify the analysis, stem diameter values used to define ‘large trees’ for each biome type were kept the same under both current and future climate scenarios. As such, any changes in tree height predicted by the model will result entirely from changes in \(H-D\) scaling relationships along climatic gradients. Given that in many ecosystems tree size distributions are predicted to shift towards smaller-sized individuals under climate change (McDowell et al., 2020), any predicted declines in tree height may therefore be conservative. However, it is also important to note that our model predictions do not account for the effects of rising atmospheric CO\(_2\) on plant water use efficiency, which may offset some of the impacts of rising aridity on tree hydraulic function under global warming (Rifai et al., 2022).

Maps of potential tree height capture major transitions in ecosystem types (Figure 4a), with projected heights of large trees ranging between 4.7 and 69.4 m. Model predictions also capture several known hotspots of tall forests, such as those of Borneo and Southeast Asia (Banin et al., 2012; Jucker, Bongalov, et al., 2018 as well as temperate rainforests in Australia, New Zealand, the western coast of the United States, Chile and Norway (Scheffer et al., 2018). However, other regional trends in forest height are less well replicated. For instance, the map does not capture known east-to-west gradients in canopy height across the Amazon basin, highlighting how other drivers aside from climate—such as soils, wind, fire and herbivory—can play a key role in shaping geographical variation in forest vertical structure (Gorgens et al., 2021; Jucker, Bongalov, et al., 2018; Moncreeff et al., 2011).

In terms of projected changes in tree height in response to climate change, the height of large-diameter trees is expected...
to decrease by an average of 5.4% globally when CO$_2$ fertilization is not taken into account. However, projected changes varied substantially among biomes and biogeographical regions (Figure 4b), ranging anywhere between −20.1% and +18.8%. Trees in Mediterranean woodlands are predicted to show the strongest decreases in height, with an average projected height decrease of 5.4% globally.
loss of 12.5%. Tropical rain forest trees are also expected to decrease in height by 5.6% on average, but this trend is much more pronounced across Amazonia and the Neotropics (−7.9%) and Africa (−4.9%) compared to Southeast Asia (−3.0%) and Australasia (−0.8%). By contrast, high-latitude forests in the northern hemisphere are predicted to increase in height as the climate warms (Figure 4b). Overall, the Tallo database provides a new way to explore how global patterns of forest canopy structure can be reconciled with the processes that constrain the allometry of individual trees. For instance, model predictions could be compared to canopy height maps derived from remote sensing to identify areas of agreement and discrepancy between the two, providing new clues on the processes that shape variation in tree height across the world’s forests. Moreover, spatially explicit maps of potential tree height generated in this way could also be used to benchmark the outputs of dynamic global vegetation models.

5 | FUTURE DEVELOPMENTS

Looking ahead, we intend to continue curating and expanding the scope and scale of the Tallo database. In addition to increasing the geographical and taxonomic coverage of the database, we also plan to source new data capturing additional axes of crown size. In particular, we aim to incorporate data on crown depth for as many trees as possible. In addition to being interesting in its own right (Shenkin et al., 2020; Vermeulen, 2014), information on crown depth would allow users to calculate more realistic estimates of crown surface area and volume (Jucker et al., 2015; Loubota Panzou et al., 2021; Shenkin et al., 2020). Additionally, we also plan to augment the database by adding information on local competitive environment (e.g. stand basal area, tree density or cover), as it is well known that tree crown architecture is strongly influenced by competition for light with neighbouring trees (Jucker et al., 2015; Lines et al., 2012; Purves et al., 2007). As part of these efforts, we will also look to expand Tallo beyond its initial focus on seed plants with a single self-supporting dominant stem that undergoes secondary growth—better capturing multi-stemmed trees, as well as other life forms such as shrubs, tree ferns, palms and climbers.

Finally, as the database expands, we also plan to begin incorporating more data on crown dimensions and tree height derived from remote sensing platforms such as airborne and terrestrial laser scanning and structure-from-motion UAV photogrammetry. These emerging technologies allow for much more accurate and comprehensive measurements of different crown attributes (Disney, 2019), as well as capturing data on the crown dimensions of large, canopy dominant trees which tend to be disproportionately underrepresented in traditional field-based surveys (Fischer et al., 2019; Marconi et al., 2021). To this end, we strongly encourage users to help us improve the Tallo database by not only reporting any errors they may come across, but also contributing their own data to future releases.

AUTHORS’ CONTRIBUTIONS

T.J. conceived the idea for the Tallo database and led the aggregation of the data with the assistance of J.Ch., D.A.C., J.Ca., A.A., G.J.L.P., T.R.F., D.F. and V.A.U. and all co-authors contributed to the data. T.J. performed the analyses with the assistance of F.J.F. T.J. wrote the first draft of the manuscript, with all authors providing editorial input.

ACKNOWLEDGEMENTS

We are indebted to the countless researchers and field assistants who helped collect the field data compiled in the Tallo database and without whom this work would not have been possible. T.J. was supported by a UK NERC Independent Research Fellowship (grant: NE/S01537X/1). J.Ch. acknowledges an ‘Investissement d’Avenir’ grant managed by the Agence Nationale de la Recherche (CEBA grant: ANR-10-LABX-25-01 and TULIP grant: ANR-10-LABX-0041). A.A. is currently supported by Hebei University (grant: 521100221033) and was previously supported by the Jiangsu Science and Technology Special Project (grant: BX2019084) and Metasequoia Faculty Research Startup Funding at Nanjing Forestry University (grant: 163010230). G.J.L.P. was supported by projects DynAFFor (grant: CZZ1636.01D) and P3FAC (grant: CZZ1636.02D) and by the International Foundation for Science (grant: D/5822-1). T.R.F. was funded by NERC (grant: NE/NO11570/1). L.F.A. was supported by CAPES and ABC-CNPq (grant: 004/96). B.B.L. was supported by COMPASS-FME, a multi-institutional project supported by the U.S. Department of Energy, Office of Science, Biological and Environmental Research as part of the Environmental System Science Program. M.v.B. acknowledges funding from the Agua Salud Project, a collaboration between the Smithsonian Tropical Research Institute (STRI), the Panama Canal Authority (ACP) and the Ministry of the Environment of Panama (MiAmbiente), the Smithsonian Institution Forest Global Earth Observatory (ForestGEO), Heising-Simons Foundation, HSBC Climate Partnership, Stanley Motta, Small World Institute Fund, Frank and Kristin Levinson, the Hoch family, the U Trust, the Working Land and Seascapes Program of the Smithsonian, the National Science Foundation (grant: EAR-1360391), Singapore’s Ministry of Education and Yale–NUS College (grant: IG16-LR004).

J.D. and H.L. were supported by the National Natural Science Foundation of China (grants: 41790422 and 42161140008). K.D. was supported by the African Forest Forum, and the DAAD within the framework of ClimapAfrica (Climate Research for Alumni and Postdocs in Africa) with funds of the Federal Ministry of Education and Research of Germany (grant: 91785431). E.R.L. was supported by a UKRI Future Leaders Fellowship (grant: MR/T019832/1). E.M. was supported by Swedish Energy Agency (grant: 35586-1). J.A.M. was supported by Consejo Nacional de Ciencia y Tecnología (CONACYT; grant: CB-2009-01-128136) and Universidad Nacional Autónoma de México (DGAPA-PAPIIT; grants: IN218416 and IN217620). J.M.D.-A. acknowledges funding from Reinforcing REDD+ and the South–South Cooperation Project, CONAFOR and USFS. S.C.R. acknowledges funding from FAPEMIG (grant:
CAG2327-07), DAAD/CAPES and CNPq. G.S. acknowledges funding by Manchester Metropolitan University’s Environmental Science Research Centre. M.S. was funded by a grant from the Ministry of Education, Youth and Sports of the Czech Republic (grant: INTER-TRANSFER LTT19018). A.T.T. acknowledges funding from the NSF (grant: 2003205). M.A.Z. thanks the MAPA-Spain for granting access to the Spanish Forest Inventory data. We thank Prof Kristina Anderson-Teixeira, Dr Anping Chen and an anonymous reviewer for their feedback which helped us improve our paper. Dr Abd Rahman Kassim, who contributed data to this project, sadly passed away before this paper was completed.

DATA AVAILABILITY STATEMENT
The version of the Tallo database described in this paper is permanently archived on Zenodo (https://doi.org/10.5281/zenodo.6637599) and an updated version of Tallo is also maintained on GitHub (https://github.com/selva-lab-repo/TALLO). Both repositories contain a metadata file describing each field of the Tallo database and a look-up table with the full list of bibliographical sources from which records were obtained. R code and ancillary data needed to replicate the three case studies presented in this paper can be found on the GitHub repository.

ORCID
Tommaso Jucker https://orcid.org/0000-0002-0751-6312
Jérôme Chave https://orcid.org/0000-0002-7766-1347
David A. Coomes https://orcid.org/0000-0002-8261-2582
Arshad Ali https://orcid.org/0000-0001-9966-2917

REFERENCES
Aguiar-Gutiérrez, J., Rifai, S., Shenkin, A., Oliveras, I., Bentley, L. P., Svátek, M., Girardin, C. A. J., Both, S., Riutta, T., Berenguer, E., Kissling, W. D., Bauman, D., Raab, N., Moore, S., Farfan-Rios, W., Figueiredo, A. E. S., Reis, S. M., Ndong, J. E., Ondo, F. E., ... Malhi, Y. (2021). Pantropical modelling of canopy functional traits using Sentinel-2 remote sensing data. Remote Sensing of Environment, 252, 112122. https://doi.org/10.1016/J.RSE.2020.112122

Anderson-Teixeira, K. J., McGarvey, J. C., Muller-Landau, H. C., Park, J. Y., Gonzalez-Akre, E. B., Herrmann, V., Bennett, A. C., So, C. V., Bourg, N. A., Thompson, J. R., McMahon, S. M., & McShea, W. J. (2015). Size-related scaling of tree form and function in a mixed-agist forest. Functional Ecology, 29(12), 1587-1602. https://doi.org/10.1111/1365-2435.12470

Atkins, J. W., Walter, J. A., Stovall, A. E. L., Fahey, R. T., & Gough, C. M. (2022). Power law scaling relationships link canopy structural complexity and height across forest types. Functional Ecology, 36(3), 713-726. https://doi.org/10.1111/1365-2435.13983

Banin, L., Feldpaus, T. R., Phillips, O. L., Baker, T. R., Lloyd, J., Affum-Baffoe, K., Arets, E. J. M., Berry, N. J., Bradford, M., Brienen, R. J. W., Davies, S., Drescher, M., Higuchi, N., Hilbert, D. W., Hladik, A., Iida, Y., Salim, K. A., Kassim, A. R., King, D. A., ... Lewis, S. L. (2012). What controls tropical forest architecture? Testing environmental, structural and floristic drivers. Global Ecology and Biogeography, 21, 1179–1190. https://doi.org/10.1111/1466-8238.12078 x

Bastin, J. F., Rutishauser, E., Kellner, J. R., Saatchi, S. Pélissier, R., Héralut, B., Slik, F., Bogaert, J., De Cannièere, C., Marshall, A. R., Poulsen, J., Alvarez-Loyaya, P., Andrade, A., Angbong-Basia, A., Araújo-Murakami, A., Arroyo, L., Ayyappan, N., de Azevedo, C. P., Banki, O., ... Zebaze, D. (2018). Pan-tropical prediction of forest structure from the largest trees. Global Ecology and Biogeography, 27(11), 1366–1383. https://doi.org/10.1111/geb.12803

Bennett, A. C., McDowell, N. G., Allen, C. D., Anderson-Teixeira, K. J., Trenberth, K. E., Nepstad, D. C., Tohver, I. M., Ray, D., Moutinho, P., Cardinot, G., Phillips, O. L., Lindenmayer, D. B., Laurance, W. F., Franklin, J. F., Lutz, J. A., Larson, A. J., Swanson, M. E., Freund, J. A., Wullschleger, S. D., ... Haack, R. A. (2015). Larger trees suffer most during drought in forests worldwide. Nature Plants, 1(10), 15139. https://doi.org/10.1038/nplants.2015.139

Boyle, B., Hopkins, N., Lu, Z., Raygoza Garay, J. A., Mozzerhin, D., Rees, T., Matasci, N., Narro, M. L., Piel, W. H., McKay, S. J., Lowry, S., Freeland, C., Peet, R. K., & Enquist, B. J. (2013). The taxonomic name resolution service: An online tool for automated standardization of plant names. BMC Bioinformatics, 14(1), 16. https://doi.org/10.1186/1471-2105-14-16

Cano, I. M., Muller-Landau, H. C., Joseph Wright, S., Bohlman, S. A., & Pacala, S. W. (2019). Tropical tree height and crown allometries for the Barro Colorado nature monument, Panama: A comparison of alternative hierarchical models incorporating interspecific variation in relation to life history traits. Biogeosciences, 16(4), 847-862. https://doi.org/10.5194/bg-16-847-2019

Cayuela, L., Granzow-de la Cerda, I., Albuquerque, F. S., & Golicher, D. J. (2012). Taxonstand: An R package for species names standardisation in vegetation databases. Methods in Ecology and Evolution, 3(6), 1078–1083. https://doi.org/10.1111/j.2041-210X.2012.00232.x

Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M. S., Delitti, W. B. C., Duque, A., Eid, T., Fearnside, P. M., Goodman, R. C., Henry, M., Martinez-Yrizar, A., Mugasha, W. A., Muller-Landau, H. C., Mencuccini, M., Nelson, B. W., Ngomanda, A., Nogueira, E. M., Ortiz-Malavassi, E., ... Viellendt, G. (2014). Improved allometric models to estimate the aboveground biomass of tropical trees. Global Change Biology, 20, 3177–3190. https://doi.org/10.1111/gcb.12629

Coomes, D. A., Holdaway, R. J., Kobe, R. K., Lines, E. R., & Allen, R. B. (2012). A general integrative framework for modelling woody biomass production and carbon sequestration rates in forests. Journal of Ecology, 100(1), 42–64. https://doi.org/10.1111/j.1365-2745.2011.01920.x

de Frenne, P., Lenoir, J., Luoto, M., Scheffers, B. R., Zellweger, F., Aalto, J., Ashcroft, M. B., Christiansen, D. M., Decoq, G., de Pauw, K., Govaert, S., Greiser, C., Gril, E., Hampe, A., Jucker, T., Klinges, D. H., Koelmeijer, I. A., Lembrechts, J. J., Marrec, R., ... Hylander, K. (2021). Forest microclimates and climate change: Importance, drivers and future research agenda. Global Change Biology, 27(11), 2279–2297. https://doi.org/10.1111/GCB.15569

de Souza, C. R., Coelho de Souza, F., Maia, V. A., de Aguiar-Campos, N., Coelho, P. A., Farrapo, C. L., Santos, A. B. M., Araújo, F. C., Gianasi, F. M., Paula, G. G. P., Morel, J. D., Fagundes, N. C. A., Garcia, P. O., Santos, P. F., Silva, W. B., Fontes, M. A. L., & Santos, R. M. (2021). Tropical forests structure and diversity: A comparison of methodological choices. Methods in Ecology and Evolution, 12(10), 2017–2027. https://doi.org/10.1111/2041-210X.13670

Disney, M. (2019). Terrestrial LiDAR: A three-dimensional revolution in how we look at trees. New Phytologist, 222(4), 1736–1741. https://doi.org/10.1111/nph.15517

Domec, J. C., Lachenbruch, B., Meinzer, F. C., Woodruff, D. R., Warren, J. M., & McCulloh, K. A. (2008). Maximum height in a conifer is associated with conflicting requirements for xylem design. Proceedings of the National Academy of Sciences of the United States of America, 105(33), 12069–12074. https://doi.org/10.1073/PNAS.0710454105

Enquist, B. J., West, G. B., & Brown, J. H. (2009). Extensions and evaluation of a general quantitative theory of forest structure and dynamics. Proceedings of the National Academy of Sciences of the United States of America, 106(17), 7046–7051. https://doi.org/10.1073/pnas.0812303106
