How climate drives the functional characteristics of vegetation across the globe has been a key question in ecological research for more than a century. While functional information is available for a large portion of the global pool of plant species, we do not know how functional traits of the different species that co-occur in a community are combined, which is what...
Table 1 | Traits used in this study and their function in the community

| Trait | Description | Function | Expected correlation with macroclimate |
|-------|-------------|----------|----------------------------------------|
| Specific leaf area, Leaf area, Leaf fresh mass, Leaf N, Leaf P | Leaf economics spectrum: Thin, N-rich leaves with high turnover and high mass-based assimilation rates | Productivity, competitive ability | Very high |
| ↓ |  |  |  |
| Leaf dry matter content, Leaf N per area, Leaf C | Thick, N-conservative, long-lived leaves with low mass-based assimilation rates | Productivity, drought tolerance | Very high |
| ↓ |  |  |  |
| Stem specific density | Fast growth, Mechanical support, Longevity | Water use efficiency | High |
| ↓ |  |  |  |
| Conduit element length | Efficient water transport |  |  |
| ↓ |  |  |  |
| Stem conduit density | Safe water transport |  |  |
| ↓ |  |  |  |
| Plant height | Mean individual height of adult plants | Competitive ability | High |
| ↓ |  |  |  |
| Seed number per reproductive unit | Seed economics spectrum: Small, well-dispersed seeds | Dispersal, regeneration | Moderate |
| ↓ |  |  |  |
| Seed mass, Seed length, Dispersal unit length | Seeds with storage reserve to facilitate establishment and increase survival |  |  |
| ↓ |  |  |  |
| Leaf N/P ratio | P limitation (N/P > 15) | Nutrient supply | Moderate |
|  | N limitation (N/P < 10) |  |  |
|  | Leaf nitrogen isotope ratio (leaf δ¹⁵N) | Access to N derived from N₂ fixation, N supply via mycorrhiza | Moderate |

Traits are arranged according to the degree to which they should respond to macroclimatic drivers. ↓ in the trait column denotes opposing relationships; ↓ in the description column denotes trade-offs. For trait units, plot-level trait means and within-plot trait variance see Table 2.

Determines their joint effect on ecosystems. At the species level, Diaz et al. demonstrated that 74% of the global spectrum of six key plant traits determining plant fitness in terms of survival, growth and reproduction can be accounted for by two principal components. They showed that the functional space occupied by vascular plant species is strongly constrained by trade-offs between traits and converges on a small set of successful trait combinations, confirming previous findings. While these constraints evolve adaptively available ecological strategies for vascular plant species globally, they provide only limited insight into trait composition within communities. There are many reasons why plant composition within communities would produce very different patterns, and indeed much theory predicts this. However, it is still unknown to what extent community-level trait composition depends on local factors (macroclimate, fine-scale soil properties, disturbance regime, successional dynamics) and regional to global environmental drivers (macroclimate, coarse-scale soil properties). As ecosystem functions and services are ultimately dependent on the traits of the species composing ecological communities, exploring community trait composition at the global scale can advance our understanding of how climate change and other anthropogenic drivers affect ecosystem functioning.

So far, studies relating trait composition to the environment at continental to global extents have been restricted to coarse-grained species occurrence data (for example, presence in 1° grid cells). Such data capture neither biotic interactions (co-occurrence in large grid cells does not indicate local co-existence), nor local variation in environmental filters (for example, variation in soil, topography or disturbance regime within grid cells). In contrast, functional composition of ecological communities sampled at fine-grained vegetation plots—with areas of a few to a few hundred square metres—is the direct outcome of the interaction between both local and large-scale factors. Here, we present a global analysis of plot-level trait composition. We combined the ‘sPlot’ database, a new global initiative incorporating more than 1.1 million vegetation plots from over 100 databases (mainly forests and grasslands; see Methods), with 30 large-scale environmental variables and 18 key plant functional traits derived from TRY, a global plant-trait database (see Methods and Table 2). We selected these 18 traits because they affect different key ecosystem processes and are expected to respond to macroclimatic drivers (Table 1). In addition, they were sufficiently measured across all species globally to allow for imputation of missing values (see Methods). All analyses were confined to vascular plant species and included all vegetation layers in a community, from the canopy to the herb layer (see Methods).

We used this unprecedented fine-resolution dataset to test the hypothesis (Hypothesis 1) that plant communities show evidence of environmental or biotic filtering at the global scale, making use of the observed variation of plot-level trait means and means of within-plot trait variation across communities. Ecological theory suggests that community-level convergence could be interpreted as the result of filtering processes, including environmental filtering and biotic interactions. Globally, temperature and precipitation drive the differences in vegetation between biomes, suggesting strong environmental filtering that constrains the number of successful trait combinations and leads to community-level trait convergence. Similarly, biotic interactions may eliminate excessively divergent trait combinations. However, alternative functional trait combinations may confer equal fitness in the same environment. If plant communities show a global variation of plot-level trait means higher than expected by chance, and a lower than expected within-plot trait variation (see Fig. 1), this would support the view that environmental or biotic filtering are dominant structuring processes of community trait composition at the global scale. A consequence of strong community-level trait convergence, and thus low variation within plots with species trait values centred around the mean, would be that plot-level means will be similar to the trait values of the species in that plot. Hence, community mean trait values should mirror the trait values of individual species.
While Hypothesis 1 addresses the degree of filtering, it does not make a statement on the attribution of driving factors. The main drivers should correlate strongly (though not necessarily linearly\(^{20}\)) with plot-level trait means and within-plot trait variance. Identifying these drivers has the potential to fundamentally improve our understanding of global trait–environment relationships. We tested the hypothesis (Hypothesis 2) that there are strong correlations between global environmental drivers such as macroclimate and coarse-scale soil properties and both plot-level trait means and within-plot trait variances\(^{14,15,17,20–22}\) (see Table 1 for expected relationships and Supplementary Table 2 for variables used). Such evidence, although correliative, may contribute to the formulation of novel hypotheses to explain global plant trait patterns.

### Results and discussion

Consistent with Hypothesis 1 and as illustrated in Fig. 1, global variation in plot-level trait means was much higher than expected by chance; all traits had positive standardized effect sizes (SESs), which were significantly \(>0\) for 17 out of 18 traits based on gap-filled data (mean SES = 8.06 standard deviations (s.d.), Table 2). This suggests that environmental or biotic filtering is a dominant force of community trait composition globally. Also as predicted by Hypothesis 1, within-plot trait variance was typically lower than expected by chance (mean SES = \(-1.76\) s.d., significantly \(<0\) for ten traits but significantly \(>0\) for three traits; Table 2). Thus, trait variation within communities may also be constrained by filtering.

### Table 2 | Traits, abbreviation, identifier in the Thesaurus Of Plant characteristics (TOP; www.top-thesaurus.org, ref.\(^{35}\)), units of measurement, observed values (obs.) SESs and significance (\(P\)) of SES for means and variances of both plot-level trait means (CWMs) and within-plot trait variances (CWVs)

| Trait                      | Abbreviation | TOP   | Unit       | Obs. | SES  | \(P\) | Mean Variance | Obs. | SES  | \(P\) | Mean Variance |
|----------------------------|--------------|-------|------------|------|------|------|--------------|------|------|------|--------------|
| Leaf area                  | LA           | 25    | mm\(^2\)   | 6.130| \(-9.75\) | *   | 1.691        | 12.53| *   | 1.565| \(-2.59\) |
| Specific leaf area         | SLA          | 50    | m\(^2\)kg\(^{-1}\) | 2.850| 9.89 | *   | 0.172        | 12.88| *   | 0.150| \(-1.33\) |
| Leaf fresh mass            | Leaf.fresh.mass | 35    | g          | \(-2.125\) | \(-13.28\) | *   | 1.395        | 10.83| *   | 1.520| \(-2.05\) |
| Leaf dry matter content    | LDMC         | 45    | g g\(^{-1}\) | \(-1.294\) | \(-5.67\) | *   | 0.101        | 11.52| *   | 0.130| 0.95 |
| Leaf C                     | LeafC        | 452   | mg g\(^{-1}\) | 6.116| \(-3.77\) | *   | 0.003        | 8.80 | *   | 0.002| \(-1.78\) |
| Leaf N                     | LeafN        | 462   | mg g\(^{-1}\) | 3.038| 4.22 | *   | 0.055        | 6.29 | *   | 0.063| \(-3.19\) |
| Leaf P                     | LeafP        | 463   | mg g\(^{-1}\) | 0.535| 9.57 | *   | 0.097        | 2.81 | *   | 0.117| \(-5.17\) |
| Leaf N per area            | LeafN.per.area | 481  | g m\(^{-2}\) | 0.251| \(-9.06\) | *   | 0.075        | 8.18 | *   | 0.099| \(-0.28\) |
| Leaf N/P ratio             | Leaf.N/P.Ratio | -    | g g\(^{-1}\) | 2.444| \(-11.95\) | *   | 0.040        | 0.40 | n.s. | 0.081| \(-2.74\) |
| Leaf \(\delta^{15}\)N     | Leaf.delta15N | -    | ppm        | 0.521| \(-3.58\) | *   | 0.254        | 6.68 | *   | 0.455| 2.82 |
| Seed mass                  | Seed.mass    | 103   | mg         | 0.407| \(-11.9\) | *   | 2.987        | 3.69 | *   | 2.784| \(-9.06\) |
| Seed length                | Seed.length  | 91    | mm         | 1.069| \(-4.51\) | *   | 0.294        | 5.50 | *   | 0.365| \(-4.67\) |
| Seed number per reproductive unit | Seed.num.rep.unit | -    | unit       | 6.179| 7.67 | *   | 2.783        | 4.40 | *   | 5.156| 1.44 |
| Dispersal unit length      | Disp.unit.length | 90   | mm         | 1.225| \(-2.51\) | *   | 0.343        | 6.50 | *   | 0.451| \(-3.21\) |
| Plant height               | Plant.height | 68    | m          | \(-0.315\) | \(-12.15\) | *   | 1.532        | 13.34| *   | 1.259| \(-9.01\) |
| Stem specific density      | SSD          | 286   | g cm\(^{-3}\) | \(-0.869\) | \(-14.93\) | *   | 0.041        | 13.15| *   | 0.058| 2.09 |
| Stem conduit density       | Stem.cond.dens | -    | mm\(^{-2}\) | 4.407| 15.08 | *   | 0.656        | 8.45 | *   | 0.975| \(-0.95\) |
| Conduit element length     | Cond.elem.length | -    | \(\mu m\) | 5.946| \(-7.09\) | *   | 0.182        | 9.14 | *   | 0.367| 7.12 |
| Mean SES                   | -            | 3.50  | 8.06       | \(-1.76\) | 1.25 |
| Mean absolute SES          | -            | 8.66  | 8.06       | 3.36 | 2.43 | 1908}
Trait correlations at the community level were relatively well captured by the first two axes of a principal component analysis (PCA) for both plot-level trait means and within-plot trait variances (Figs. 2 and 3). The dominant axes were determined by those traits with the highest absolute SESs of plot-level trait mean trait values (Table 2, mean of community-weighted means (CWMs)). The PCA of plot-level trait means (Fig. 2) reflects two main functional continua on which community trait values converge: one from short-stature, small-seeded communities such as grasslands or herbaceous vegetation to tall-stature communities with large, heavy diaspores such as forests (the size spectrum), and the other from communities with resource-acquisitive to those with resource-conservative leaves (that is the leaf economics spectrum). The high similarity between this PCA and the one at the species level by Díaz et al. is striking. Here, at the community level, based on 1.1 million plots, the same functional continua emerged as at the species level, based on 2,214 species. While the trade-offs between different traits at the species level can be understood from a physiological and evolutionary perspective, finding similar trade-offs between traits at the community level was unexpected, as species with opposing trait values can co-exist in the same community. In combination with our finding of strong trait convergence, these results reveal a strong parallel of present-day community assembly with individual species’ evolutionary histories.

Surprisingly, we found only limited support for Hypothesis 2. Community-level trait composition was poorly captured by global climate and soil variables. None of the 30 environmental variables accounted individually for more than 10% of the variance in the traits defining the main dimensions in Fig. 2 (Supplementary Fig. 2). The coefficients of determination were not improved when testing for non-linear relationships (see Methods). Using all 30 environmental variables simultaneously as predictors only accounted for 10.8% or 14.0% of the overall variation in plot-level trait means (cumulative variance, respectively, of the first two or all 18 constrained axes in a redundancy analysis (RDA)). Overall, our results show that similar global-scale climate and soil conditions can support communities that differ markedly in mean trait values and that different climates can support communities with quite similar mean trait values.

The ordination of within-plot variance of the different traits (Fig. 3) revealed two main continua. Variances of plant height and diaspore mass varied largely independently of variances of traits representing the leaf economics spectrum. This suggests that short and tall species can be assembled together in the same community independently from combining species with acquisitive leaves with species with conservative leaves. Global climate and soil variables accounted for even less variation on the first two PCA axes in within-plot trait variances than on the first two PCA axes.
in plot-level trait means. Only two environmental variables had $r^2 > 3\%$ (Supplementary Fig. 3), whether allowing for non-linear relationships (see Methods) or not, and, overall, macro-environment accounted for only 3.6% or 5.0% of the variation (cumulative variance, respectively, of the first two or all 18 constrained axes).

Removing species richness effects from within-plot trait variances did not increase the amount of variation explained by the environment (see Methods).

The findings of our study contrast strongly with studies where the variation in traits between species was calculated at the level of the species pool in large grid cells\textsuperscript{15,16}, suggesting that plot-level and grid cell-level trait composition are driven by different factors\textsuperscript{21}. Plot-level trait means and variances may both be predominately driven by local environmental factors, such as topography (for example north-facing versus south-facing slopes), local soil characteristics (for example soil depth and nutrient supply),\textsuperscript{3,14,24,25} disturbance regime (including land use\textsuperscript{26} and successional status\textsuperscript{2,27}) and biotic interactions\textsuperscript{18,19,28}, while broad-scale climate and soil conditions may only become relevant for the whole species pool in large grid cells. Such differences emphasize the importance of the effect of local environment on communities’ trait composition, which should be taken into account when interpreting the effect of environmental drivers on functional trait diversity, using data on either floristic pools or ecological communities.

We note that the strongest community-level correlations with environment were found for traits not linked to the leaf economics spectrum. Mean stem specific density increased with potential evapotranspiration (PET, $r^2 = 15.6\%$; Fig. 4a,b), reflecting the need to produce denser wood with increasing evaporative demand. Leaf N/P ratio increased with growing-season warmth (growing degree days above 5 °C, GDD5, $r^2 = 11.5\%$; Fig. 4d), indicating strong phosphorus limitation\textsuperscript{29} in most plots in the tropics and subtropics (Fig. 4c,d). This pattern was not brought about by a parallel increase in the presence of legumes, which tend to have relatively high N/P ratios; excluding all species of Fabaceae resulted in a very similar relationship with GDD5 ($r^2 = 10.0\%$). The global N/P pattern is consistent with results based on traits of single species related to mean annual temperature\textsuperscript{30}. We assume that the main underlying mechanism is the high soil weathering rate at high temperatures and humidity, which in the tropics and subtropics was not reset by Pleistocene glaciation. Thus, phosphorus limitation may weaken the relationships between productivity-related traits and macroclimate
For example, specific leaf area (SLA) may be low as a consequence of low nutrient availability in favourable climates as well as being low as a consequence of low temperature and precipitation with favourable nutrient supply. Overall, our findings are relevant for improving dynamic global vegetation models (DGVMs), which so far have used trait information only from a few calibration plots. The sPlot database provides much-needed empirical data on the community trait pool in DGVMs and identifies traits that should be considered when predicting ecosystem functions from vegetation, such as stem specific density and leaf N/P ratio.

Our results were surprisingly robust with respect to the selection of trait data, when comparing different plant formations, and when explicitly accounting for the uneven distribution of plots. Using the original trait values measured for the species from the TRY database for the six traits used by Díaz et al. (see Methods), resulted in the same two main functional continua and overall a highly similar ordination pattern (Supplementary Fig. 4) compared with using gap-filled data for 18 traits (Fig. 2). Community-level trait composition was also similarly poorly captured by global climate and soil variables. Single regressions of CWMs with all environmental variables revealed very similar patterns to those based on the full set of 18 gap-filled traits (Supplementary Fig. 5). Similarly, subjecting the CWMs based on six original traits to an RDA with all 30 environmental variables accounted for only 20.6% or 21.8% of the overall variation in CWMs (cumulative variance of the first two or all six constrained axes, respectively, Supplementary Fig. 4). These results clearly demonstrate that the imputation of missing trait values did not result in spurious artefacts that may have obscured community trait–environment relationships.

We also assessed whether the observed trait–environment relationships hold for forest and non-forest vegetation independently (see Methods). Both subsets confirmed the overall patterns in trait means (Supplementary Figs. 3–6). The variance in plot-level trait means explained by large-scale climate and soil variables was higher for forest than non-forest plots, probably because forests belong to a well-defined and rather resource-conservative formation, whereas non-forest plots encompass a heterogeneous mixture of different vegetation types, ranging from alpine meadows to semi-deserts, and tend to depend more on disturbance and management, which can strongly affect trait–environment relationships of communities. Finally, to test whether our findings depended on the uneven distribution of plots among the world’s different climates and soils, we...
repeated the analyses in 100 subsets of ~ 100,000 plots resampled in the global climate space (Supplementary Figs. 7 and 8). The analysis of the resampled datasets revealed the same patterns and confirmed the impact of PET and GDD5 on stem specific density and leaf N/P ratio, respectively. The correlations between trait means and environmental variables were, however, stronger in the resampled subsets, possibly because the resampling procedure reduced the overrepresentation of the temperate-zone areas with intermediate climatic values.

Our findings have important implications for understanding and predicting plant community trait assembly. First, worldwide trait variation of plant communities is captured by a few main dimensions of variation, which are surprisingly similar to those reported by species-level studies5,9, suggesting that the drivers of past trait evolution, which resulted in the present-day species-level trait spectra1, are also reflected in the composition of today’s plant communities. If species-level trade-offs indeed constrain community assembly, then the present-day contrasts in trait composition of terrestrial plant communities should also have existed in the past and will probably remain, even for novel communities, in the future. Most species in our present-day communities evolved under very variable filtering conditions across the globe, with respect to temperature and precipitation regimes. Therefore, it can be assumed that future filtering conditions will result in novel communities that follow the same functional continua from short-stature, small-seeded communities to tall-stature communities with large, heavy diasporas and from communities with resource-acquisitive leaves to those with resource-conservative leaves. Second, the main plot-level vegetation trait continua cannot easily be captured by coarse-resolution environmental variables21. This brings into question both the use of simple large-scale climate relationships to predict the global spectra of plant assemblages13,15,16,22 and attempts to derive net primary productivity and global carbon and water budgets from global climate, even when employing powerful trait-based vegetation models31. The finding that within-plot trait variance is only very weakly related to global climate or soil variables points to the importance of i) local-scale climate or soil variables, ii) disturbance regimes and iii) biotic interactions for the degree of local trait dispersion11. Finally, our findings on the limited role of large-scale climate in explaining trait patterns and on the prevalence of phosphorus limitation in most plots in the tropics and subtropics call for including local variables when predicting community trait dispersion.20

**Fig. 4 | The two strongest relationships found for global plot-level trait means (CWMs) in the sPlot dataset.**

**a, b.** CWM of the natural logarithm of stem specific density (g cm\(^{-3}\)) as a global map, interpolated by kriging within a radius of 50 km around the plots using a grid cell of 10 km (a), and function of potential evapotranspiration (PET, \(r^2 = 0.156\)) (b). **c, d.** CWM of the natural logarithm of the N/P ratio (g g\(^{-1}\)) as a global kriging map (c) and function of the warmth of the growing season, expressed as growing degree days over a threshold of 5 °C (GDD5, \(r^2 = 0.115\)) (d). Plots with N/P ratios > 15 (of 2.71 on the log scale) tend to indicate phosphorus limitation29 and are shown above the broken line in red colour (90,979 plots, 8.16% of all plots). The proportion of plots with N/P ratios >15 increases with GDD5 \((r^2 = 0.895\) for a linear model on the log response ratio of counts of plots with N/P > 15 and ≤15 counted within bins of 500 GDD5).
patterns. Even under similar macro-environmental conditions, communities can vary greatly in trait means and variances, consistent with high local variation in species’ trait values. Future research on the functional response of communities to changing climate should incorporate the effect of local environmental conditions and biotic interactions in building reliable predictions of vegetation dynamics.

Methods

Vegetation data. The sPlot 2.1 vegetation database contains 1,121,244 plots with 23,586,216 species × plot observations, which is, records of a species in a plot (https://www.sdiv.de/en/sdiv/working_groups/wg_paul/splot.html). This database aims to compile plot-based vegetation data from all vegetation types worldwide, but with a particular focus on forest and grassland vegetation. Although the initial aim of sPlot was to achieve global coverage, the plots are very unevenly distributed, with most data coming from Europe, North America and Australia and an overrepresentation of temperate vegetation types (Supplementary Fig. 1).

For most plots (97.2%), information on the single species’ relative contribution to the sum of plants in the plot was available, expressed as cover, basal area, individual count, importance value or percentage frequency in subplots. For the other 2.8% (31,461 plots), for which only presence/absence (p/a) was available, we assigned equal relative abundance to the species (1 species richness). For plots with a mix of cover and p/a information (mostly forest plots, where for other layers, information had been added on a p/a basis: 8,524 plots), relative abundance was calculated by assigning the smallest cover value that occurred in a particular plot to all species with only p/a information in that plot. In most cases (98.4%), plot records in sPlot include full species lists of vascular plants. Bryophytes and lichens were additionally identified in 14% and 7% of plots, respectively. After removing plots without geographic coordinates and all observations on bryophytes and lichens, the database contained 22,195,966 observations on the relative abundance of vascular plant species in a total of 1,117,369 plots. The temporal extent of the data spans from 1885 to 2015, but >95% of vegetation plots were recorded later than 1900. Plot size was reported in 65.4% of plots. While forest plots had plot sizes ≥ 100 m², in most cases ≤ 1,000 m², non-forest plots typically ranged from 5 to 100 m².

Taxonomy. To standardize the nomenclature of species within and between sPlot and TRY (see below), we constructed a taxonomic backbone of the 121,661 names contained in the two databases. Prior to name matching, we ran a series of string manipulation routines in R, to remove special characters and numbers, as well as standardized abbreviations in names. Taxon names were parsed and resolved using Taxonomic Name Resolution Service version 4.0 (TNRIS; http://tnris.iplantcollaborative.org; accessed 20 September 2015), selecting the best match across the five following sources: i) The Plant List (version 1.1; http://www.theplantlist.org/; accessed 19 August 2015), ii) Global Compositional Checklist (GCC, http://compositional.landcareresearch.co.nz/Default.aspx; accessed 21 August 2015), iii) International Legume Database and Information Service (ILDIS, http://www.ildis.org/legumeweb/; accessed 21 August 2015), iv) Tropicos (http://www.tropicos.org/; accessed 19 December 2014), and v) USDA Plants Database (https://www.plants.usda.gov/wps/portal/usda/plants/; accessed 17 January 2015). We allowed for partial matching to the next highest taxonomic rank (genus or family) in cases where full taxon names could not be found. All names matched or converted from a synonym by TNRIS were considered accepted taxon names. In cases where no exact match was found (for example when alternative spelling corrections were reported), names with probabilities of ≥ 95% or higher were accepted and those with <95% were examined individually. Remaining non-matching names were resolved using the National Center for Biotechnology Information’s Taxonomy database (NCBI, http://www.ncbi.nlm.nih.gov/; accessed 25 October 2011) within TNRIS, or sequentially compared directly against The Plant List and Tropicos (Accessed September 2015). Names that could not be resolved against any of these lists were left as blanks in the final standardized name field. This resulted in a total of 86,760 resolved names, corresponding to 664 families, occurring in sPlot or TRY or both. Classification into families was carried out according to APGIII, and was used to identify non-vascular plant species (5.1% of the taxon names), which were excluded from the subsequent statistical analysis.

Trait data. Data for 18 traits that are ecologically relevant (Table 1) and sufficientlly covered across species were requested from TRY (version 3.0) on 10 August 2016. We applied gap-filling with Bayesian Hierarchical Probabilistic Matrix Factorization (BHPMF). We used the prediction uncertainties provided by BHPMF for each imputation to assess the quality of gap-filling, and removed all imputations with a coefficient of variation > 1. We obtained 18 gap-filled traits for 26,632 out of 58,065 taxa in sPlot, which corresponds to 45.9% of all species but 88.7% of all species × plot combinations. Trait coverage of the most frequent species was 77.2% and 96.2% for taxa that occurred in more than 100 or 1,000 plots, respectively. The gap-filled trait data comprised observed and imputed values on 632,938 individual plants, which we log-transformed and aggregated by taxon. For those taxa that were recorded at the genus level only, we calculated genus means. Of 22,195,966 records of vascular plant species with geographic reference, 21,172,989 (95.4%) refer to taxa for which we had gap-filled trait values. This resulted in 1,115,785 and 1,099,463 plots for which we had at least one taxon or two taxa with a trait value (99.5% and 98.1%, respectively, of all 1,121,244 plots), and for which trait means and variances could be calculated.

As some mean values of traits in TRY were based on a very small number of replications per species, which results in uncertainty in trait mean and variance calculations, we tested to what degree the trait patterns in the dataset might be caused by a potential removal of trait variation by imputation of trait values, and we additionally carried out all analyses using the original trait data on the same 632,938 individual plants instead of gap-filled data (Supplementary Table 1). The degree of trait coverage of species ranged between 7.0% and 58.0% for leaf fresh mass and plant height, respectively. Across all species, mean coverage of species with at least one trait value was 21.8%, compared with 45.9% for gap-filled trait data. Linking these trait values to the species occurrence data resulted in a coverage of species × plot observations with trait values between 7.6% and 96.6% for conduit element length and plant height, respectively (Supplementary Table 1), with a mean of 60.7% compared with 88.7% for those based on gap-filled traits. Using these original trait values to calculate CWM trait values (see below) resulted in a plot coverage of trait values between 48.2% and 100% for conduit element length and SLA, respectively. Across all plots, mean coverage of plots with original trait values was 89.3%, compared with 100% for gap-filled trait data (Supplementary Table 1).

We are aware that using species mean values for trait means excludes the possibility of accounting for intraspecific variance, which can also strongly respond to the environmental. Thus, using one single value for a species is a source of error in calculating trait means and variances.

Environmental data. We compiled 30 environmental variables (Supplementary Table 2). Macroclimate variables were extracted from CHELSA40,41, V1.1 (Climatologies at High Resolution for the Earth’s Land Surface Areas, www.chelsa-climate.org). CHELSA provides 19 bioclimatic variables equivalent to those used in WorldClim (www.worldclim.org) at a resolution of 30 arcsec (~ 1 km at the equator), averaging global climatic data from the period 1979–2013 and using a quasi-mechanistic statistical downscaling of the ERA-Interim reanalysis42. Variables reflecting growing-season warmth were growing degree days above 1 °C (GDD1) and 5 °C (GDD5), calculated from CHELSA data. We also compiled an index of aridity (AR) and a model for PET extracted from the Consortium of Spatial Information (CGIAR-CSI) website (www.cgiar-csi.org). In addition, 7 soil variables were extracted from the SOILGRIDS project (https://soilgrids.org/, licensed by ISRIC—World Soil Information), downloaded at 250 m resolution and then resampled using the 30 arcsec grid of CHELSA (Supplementary Table 2). We refer to these climate and soil data as environmental data.

Community trait composition. For every trait j and plot k, we calculated the plot-level trait mean as community-weighted mean according to:}

\[
CWM_{jk} = \sum_{i} p_{jk} x_{jk} / \sum_{i} p_{jk} = \sum_{i} p_{jk} x_{jk} / \sum_{i} p_{jk}
\]

where \(n_{jk}\) is the number of species sampled in plot \(k\), \(p_{jk}\) is the relative abundance of species \(i\) in plot \(k\), referring to the sum of abundances for all species with traits in the plot, and \(x_{jk}\) is the mean value of species \(i\) for trait \(j\). This computation was done for each of the 18 traits for 1,115,785 plots. The within-plot trait variance is given by community-weighted variance (CWMV):}

\[
CWMV_{jk} = \sum_{i} p_{jk} (x_{jk} - CWM_{jk})^2 / \sum_{i} p_{jk}
\]

CWMV is equal to functional dispersion as described by Raio quadratic entropy (FD\(_{\text{Q}}\))\(^2\), when using a squared Euclidean distance matrix \(d_{ki}^2\):}

\[
CWMV_{jk} = \sum_{i,j} p_{jk} (x_{jk} - x_{ij})^2 / \sum_{i,j} p_{jk}
\]

We had CWMV information for 18 traits for 1,099,463 plots, as at least two taxa were needed to calculate CWMV. We performed the calculations using the ‘data.table’ package in R.

Assessing the degree of filtering. To analyse how plot-level trait means and within-plot trait variances (based on gap-filled trait data) depart from random expectation, for each trait we calculated SEs for the variance in CWMVs and for the mean in CWMVs. Significantly positive SEs in variance of CWM and significantly negative ones in the mean of CWM can be considered a global-level measure of environmental or biotic filtering. To provide an indication of the global direction of filtering, we also report SEs for the mean of CWM trait values. Similarly,
to measure how much within-community trait dispersion varied globally, we also calculated SESs for the variance in CWV.

We obtained SESs from 100 runs of randomizing trait values across all species globally for every trait. We calculated all CWM values and CWV with random trait values, but keeping all species abundances in plots. Thus, the results of randomization are independent from species co-occurrence structure of plots. For every trait, the SESs of the variance in CWM were calculated as the observed value of variance in CWM minus the mean variance in CWM of the random runs, divided by the standard deviation of the variance in CWM of the random runs.

We subjected all CWM values for forest and non-forest plots to PCA, RDA and bivariate linear regressions to environmental variables as described above.

The forest plots, in particular, confirmed the overall patterns with respect to variation in CWM explained by the first two PCA axes (60.5%) and the two orthogonal continua from small to large size and the leaf economics spectrum (Supplementary Fig. 4). The variance in CWV of the forest explained by macroclimate and soil conditions was much larger for the forest subset than for the total data, with the best relationship (leaf N/P ratio and the mean temperature of the coldest quarter, bio11) having \( r^2 = 0.369 \) and the next best ones (leaf N/P ratio and GDD1 and GDD5) close to this value with \( r^2 = 0.357 \) (Supplementary Fig. 7) with an overall variation in CWM values explained by environment of 25.3% (cumulative variance of the constrained axes in an RDA). The non-forest plots showed a stronger functional continua, but with a lower total amount of variation in CWM accounted for by the first two PCA axes (41.8%, Supplementary Fig. 8) and much lower overall variation explained by environment. For non-forests, the best correlation of any CWM trait with environment was that of volumetric content of coarse fragments in the soil (CoarseFrags, \( r^2 = 0.156 \)) with environment, the highest term were encountered between stem specific density and PET (CoarseFrags × environment correlations. Here, the strongest correlations were encountered among all environmental variables showed statistically significant correlations. Instead, we refrain from presenting any inference statistics, as with independent from species co-occurrences structure of plots. For every trait, the SESs for the variance in CWV were calculated as the observed value of variance in CWV minus the mean variance in CWV of the random runs, divided by the standard deviation of the variance in CWV of the random runs. Thus, the SESs for the mean in CWM, the mean in CWV and the variance in CWV were calculated accordingly. Tests for significance of SESs were obtained by fitting generalized Pareto-distribution of the most extreme random values and then estimating P values from this fitted distribution.

Vegetation trait–environment relationships. Of the 1,115,785 plots with CWM values, 1,114,304 (99.9%) had complete environmental information and coordinates. This set of plots was used to calculate single linear regressions of each of the 18 traits on each of the 30 environmental variables. We used the 'corplot' function in R to illustrate Pearson correlation coefficients (see Supplementary Figs. 1, 2, 4, 6 and 8) and, for the strongest relationships, produced bivariate graphs and mapped the global distribution of the CWM values using kriging interpolation in ArcGIS 10.2 (Fig. 1). We also tested for non-linear relationships with environment by including an additional quadratic term in the linear model and then reported coefficients of determination. As in the linear relationships of CWM with environmental variables, the strongest correlations were encountered between stem specific density and PET (CoarseFrags × environment correlations. Here, the strongest correlations were encountered among all environmental variables showed statistically significant correlations. Instead, we refrain from presenting any inference statistics, as with independent from species co-occurrences structure of plots. For every trait, the SESs for the variance in CWV were calculated as the observed value of variance in CWV minus the mean variance in CWV of the random runs, divided by the standard deviation of the variance in CWV of the random runs. Thus, the SESs for the mean in CWM, the mean in CWV and the variance in CWV were calculated accordingly. Tests for significance of SESs were obtained by fitting generalized Pareto-distribution of the most extreme random values and then estimating P values from this fitted distribution.

Resampling procedure in environmental space. To achieve a more even representation of plots across environmental space, we first subjected the same 30 global climate and soil variables, as described above, to a PCA using the climate space of the whole globe, irrespective of the presence of plots in this space, and scaling each variable to a mean of zero and a standard deviation of one. We used a 2.5 arcmin spatial grid, which comprised 8,384,404 terrestrial grid cells. We then counted the number of vegetation plots in the splot database that fell into each grid cell. For this analysis, we did not use the full set of 1,117,369 plots with trait information (see above), but only those plots that had a location accuracy of max. 3 km, resulting in a total of 799,400 plots. The resulting PCA scores based on the first two principal components (PC1–PC2) were rasterized to a 100 × 100 grid in PC1–PC2 environmental space, which was the most appropriate resolution according to a sensitivity analysis. This sensitivity analysis tested different grid resolutions, from a coarse-resolution bivariate space of 100 grid cells (10 × 10) to a very fine-resolution space of 250,000 grid cells (500 × 500), iteratively increasing the number of cells along each principal component by 10 cells. For each iteration, we computed the total number of plots from splot per environmental grid cell and plotted the median sampling effort (number of plots) across all grid cells versus the resolution of the PC1–PC2 space. We found that the curve flattens off at a bivariate environmental space of 100 × 100 grid cells, which was the resolution for which the median sampling effort stabilized at around 50 plots per grid cell. As a result, we resampled plots only in environmental cells with more than 50 plots (858 cells in total).

To optimize our resampling procedure within each grid cell, we used the heterogeneity-constrained random (HCR) resampling approach. The HCR approach selects the subset of vegetation plots for which those plots are the most dissimilar in their species composition while avoiding selection of plots representing peculiar and rare communities (outliers), thus providing a representative subset of plots from the resampled grid cell. We used the turnover component of the Jaccard’s dissimilarity index (\( j_{tu} \)) as a measure of dissimilarity. The \( j_{tu} \) index accounts for species replacement without being influenced by differences in species richness. Thus, it reduces the effects of any imbalances that may exist between different plots due to differences in plot richness. With the HCR approach, the original trait values were used to build a binary dissimilarity matrix by running 1,000 iterations, randomly selecting 50 plots out of the total number of plots available within that grid cell. Where the cell contained 50 or fewer plots, all were included and the resampling procedure was not run. This procedure thinned out over-sampled climate types, while retaining the full environmental gradient.

All 1,000 random draws of a given grid cell were subsequently sorted according to the decreasing mean of \( j_{tu} \) between pairs of vegetation plot and then resampled again according to the increasing variation in \( j_{tu} \) between pairs of vegetation plots. Ranks from both sortings were summed for each random draw, and the random draw with the lowest summed rank was considered as the most representative of the focal grid cell. Because of the randomized nature of the HCR approach, this procedure was repeated 100 times for each grid cell. This enabled us to produce 100 different subsamples out of the full sample of 799,400 vegetation plots subjected to the resampling procedure. Each of these 100 subsamples was finally subjected to ordinary linear regression, PCA and RDA as described above. We calculated the mean correlation coefficient across the 100 resampled datasets for each environmental variable with each trait.
To plot bivariate relationships, we used the mean intercept and slope of these relationships. PCA loadings of all 100 runs were stored and averaged. As different runs showed different orientation on the first PCA axes, we switched the signs of the axis loadings in some of the runs to make the 100 PCAs comparable to the reference PCA, based on the total dataset. Across the 100 resampled datasets, we then calculated the minimum and maximum loading for each of the two PCA axes and plotted the result as ellipsoids. We also collected the post hoc regression coefficients of PCA scores with the environmental variables in each of the 100 runs, switched the signs accordingly and plotted the correlations to PC1 and PC2 as ellipsoids. The result is a synthetic PCA of all 100 runs. To illustrate the coverage of plots in PCA space, we used plot scores of one of the 100 random runs. Similarly, the coefficients of determination obtained from the RDAs of these 100 resampled sets were averaged.

The mean PCA loadings across these 100 subsets (summarized in Supplementary Fig. 10) were fully consistent with those of the full dataset in Fig. 2, with the same two functional continua in plant size and diaspore mass (from bottom left to top right), and perpendicular to that, the leaf economics spectrum. The variation in CWM accounted for by the first two axes was on average 50.9% ± 0.04 s.d., and thus virtually identical with that in the total dataset. In contrast, the variation explained on average by macroclimate and soil conditions (26.5% ± 0.01 s.d. as the average cumulative variance of all 18 constrained axes in the RDAs across all 100 runs) was considerably larger than that for the total dataset, which is also reflected in consistently higher correlations between traits and environmental variables (Supplementary Fig. 11). The highest mean correlation was encountered for plants growing at PET (mean r² = 0.342 across 100 runs). PET was a better predictor for plant height than the precipitation of the wettest months (bio13, mean r² = 0.231), as had been suggested previously\(^1\). The correlation of PET with stem specific density (mean r² = 0.284) and warmth of the growing season (expressed as growing degree days above the threshold 5 °C, GDD5) with leaf N/P ratio (mean r² = 0.250) ranked among the best 12 correlations encountered out of all 540 trait–environment relationships, which confirms the patterns found in the whole dataset (compared with Fig. 1). Overall, the coefficients of determination were much closer to the ones reported from other studies with a global collection of a few hundred plots (r² values ranging from 36% to 53% based on multiple regressions of single traits with 5 to 6 environmental drivers\(^2\)).

**Reporting Summary.** Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

**Data availability**
The data contained in sPlot (the vegetation-plot data complemented by trait and environmental information) are available on request, by contacting any of the sPlot consortium members, for submission of a paper proposal. The proposals should follow the Governance and Data Property Rules of the sPlot Working Group, which are available on the sPlot website (www.idv.de/sPlot/).

Received: 2 March 2018; Accepted: 18 September 2018; Published online: 19 November 2018

**References**
1. Warming, E. Lehrbuch der ökologischen Pflanzengeographie – Eine Einführung in die Kenntnisse der Pflanzenwelt. Berlin, 1896.
2. Garnier, E. et al. Plant functional markers capture ecosystem properties during secondary succession. *Ecology* 85, 2630–2637 (2004).
3. Ordoñez, J. C. A global study of relationships between leaf traits, climate and soil measures of nutrient fertility. *Glob. Ecol. Biogeogr.* 18, 137–149 (2009).
4. Garnier, E., Navas, M.-L. & Grigulis, K. Plant Functional Diversity – Organism Traits, Community Structure, and Ecosystem Properties (Oxford Univ. Press, Oxford, 2016).
5. Diaz, S. et al. The global spectrum of plant form and function. *Nature* 529, 167–171 (2016).
6. Mora, A. T. et al. Global patterns in plant height. *J. Ecol.* 97, 923–932 (2009).
7. Wright, I. J. et al. The worldwide leaf economics spectrum. *Nature* 428, 821–827 (2004).
8. Reich, P. B. The world-wide ‘fast–slow’ plant economics spectrum: a traits manifesto. *J. Ecol.* 102, 275–301 (2014).
9. Adler, P. B. et al. Functional traits explain variation in plant life history strategies. *Proc. Natl Acad. Sci. USA* 111, 740–745 (2014).
10. Marks, C. O. & Lechowicz, M. J. Alternative designs and the evolution of functional diversity. *Am. Nat.* 167, 55–67 (2006).
11. Grime, J. P. Trait convergence and trait divergence in herbaceous plant communities: mechanisms and consequences. *J. Veg. Sci.* 17, 255–260 (2006).
12. Muscarella, R. & Uriarte, M. Do community-weighted mean functional traits reflect optimal strategies? *Proc. R. Soc. B* 283, 20152434 (2016).
13. Swenson, N. G. & Weiser, M. D. Plant geography upon the basis of functional traits: an example from eastern North American trees. *Ecology* 91, 2234–2241 (2010).
Articles
55. Garnier, E. et al. Towards a thesaurus of plant characteristics: an ecological
53. Lengyel, A., Chytrý, M. & Tichý, L. Heterogeneity-constrained random
52. Oksanen, J. et al. 
51. Friendly, M. Corrgrams: exploratory displays for correlation matrices. 
49. Hawkins, B. A. et al. Structural bias in aggregated species-level variables
46. Rao, C. R. Diversity and dissimilarity coefficients: a unified approach. 
45. Buzzard, V. et al. Re-growing a tropical dry forest: functional plant trait
44. Enquist, B. et al. Scaling from traits to ecosystems: developing a general trait
43. Instituto de Investigación Ambiental, La Plata, Argentina. 42. Laboratoire de Physiologie et Ecologie des Plantes, Université Laval, Québec, Canada.
41. Department of Soil Science and Geology, Cornell University, Ithaca, NY, USA. 40. Institute of Biology, Academy of Sciences of the Czech Republic, Prague, Czech Republic. 39. Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT, USA. 38. Department of Biology, University of Potsdam, Potsdam, Germany. 37. Department of Plant Physiology and Microbiology, University of Oslo, Oslo, Norway. 36. Department of Life Sciences, University of Bristol, Bristol, UK. 35. Department of Environmental Management, University of Ghent, Ghent, Belgium. 34. Institute of Ecology, University of Tartu, Tartu, Estonia. 33. Department of Life Sciences, University of York, York, UK. 32. Department of Biology, University of Iceland, Reykjavik, Iceland. 31. Department of Biology, University of Ljubljana, Ljubljana, Slovenia. 30. Department of Natural Sciences, University of North Dakota, Grand Forks, ND, USA. 29. Department of Biology, University of Minnesota, Minneapolis, MN, USA. 28. Department of Botany and Plant Biology, University of California, Davis, CA, USA. 27. Department of Botany, University of Florida, Gainesville, FL, USA. 26. Department of Botany, University of Washington, Seattle, WA, USA. 25. Department of Biology, Laval University, Quebec City, Quebec, Canada. 24. Department of Biology, University of Alberta, Edmonton, Alberta, Canada. 23. Department of Biology, University of Otago, Dunedin, New Zealand. 22. Department of Biology, University of Hawaii at Manoa, Honolulu, HI, USA. 21. Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada. 20. Department of Biology, University of Guelph, Guelph, Ontario, Canada. 19. Department of Biology, University of Victoria, Victoria, British Columbia, Canada. 18. Department of Biology, University of Tennessee, Knoxville, TN, USA. 17. Department of Biology, University of Wisconsin-Madison, Madison, WI, USA. 16. Department of Biology, University of Basel, Basel, Switzerland. 15. Department of Biology, University of Oxford, Oxford, UK. 14. Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. 13. Department of Biology, University of São Paulo, São Paulo, Brazil. 12. Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany. 11. Department of Evolutionary Ecology and Conservation, University of Edinburgh, Edinburgh, UK. 10. School of Biology, Georgia Institute of Technology, Atlanta, GA, USA. 9. Department of Ecology and Evolutionary Biology, Brown University, Providence, RI, USA. 8. College of Life and Environmental Sciences, University of Exeter, Exeter, UK. 7. Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, UK. 6. Department of Botany, University of Göttingen, Göttingen, Germany. 5. Department of Plant Biology, Swedish University of Agricultural Sciences, Umeå, Sweden. 4. Institute of Botany, University of Regensburg, Regensburg, Germany. 3. Department of Biology, University of Missouri-Columbia, Columbia, MO, USA. 2. Department of Biology, University of Wisconsin-Madison, Madison, WI, USA. 1. Department of Biology, University of Massachusetts, Amherst, MA, USA.

Acknowledgements
The sPlot was initiated by sDiv, the Synthesis Centre of the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, funded by the German Research Foundation (FZT 118) and is now a platform of iDiv. H.B., J.D., O.Pu., U.J., B.J.-A., J.K., D.C., F.M.S., M.W. and C.W. appreciate the direct funding through iDiv. For all further acknowledgements see the Supplementary Information.

Author contributions
H.B. and U.J. wrote the first draft of the manuscript, with considerable input by B.J.-A. and R.F. Most of the statistical analyses and the production of the graphs were carried out by H.B., H.B., O.Pu. and U.J. initiated sPlot as an iDiv working group and iDiv platform. J.De. compiled the plot databases globally. J.De., S.M.H., U.J., O.Pu. and F.J. harmonized vegetation databases. J.De. and B.J.-A. coordinated the sPlot consortium. J.K. provided the trait data from TRY. F.S. performed the trait data gap filling. O.Pu. produced the taxonomic backbone. B.J.-A., G.S. and E. Welk compiled environmental data and produced the global maps. S.M.H. wrote the Turbov3 software, which holds the sPlot database. J.L. and T.H. wrote the resampling algorithm. Many authors participated in one or more of the three sPlot workshops at iDiv where the sPlot initiative was conceived and planned, and evaluation of the data and first drafts were discussed. All other authors contributed data. All authors contributed to writing the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41559-018-0699-8.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to H.B.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2018
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
- Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted.
- Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated
- Clearly defined error bars
- State explicitly what error bars represent (e.g. SD, SE, CI)

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection

The data collection is described in detail.

Data analysis

All data analysis was carried out in R. All packages used are mentioned.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data contained in sPlot (the vegetation-plot data complemented by trait and environmental information) are available by request, through contacting any of the sPlot consortium members for submitting a paper proposal. The proposals should follow the Governance and Data Property Rules of the sPlot Working Group, which are available on the sPlot website (www.idiv.de/sPlot).
Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences
- Behavioural & social sciences
- Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| Sample size | Sample size is the number of plots compiled in this study, which are independent observations of species composition. In total, there are 1,121,244 plots. |
| Data exclusions | This is described in the Methods. Only plots were used with coordinates and with vascular plant species for which traits were available. |
| Replication | We repeated the analysis with stratified subsamples, which is described in detail. |
| Randomization | Traits were randomized globally to obtain standardized effect sizes, which is described in detail. |
| Blinding | Not relevant because samples were not subjected to different treatments. |

Reporting for specific materials, systems and methods

| Materials & experimental systems | Methods |
|---------------------------------|---------|
| n/a                             | n/a     |
| - Unique biological materials   | - Involved in the study |
| - Antibodies                    | - ChIP-seq |
| - Eukaryotic cell lines         | - Flow cytometry |
| - Palaeontology                 | - MRI-based neuroimaging |
| - Animals and other organisms   |         |
| - Human research participants   |         |