Late Quaternary climate legacies in contemporary plant functional composition


It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1111/gcb.14375

Publisher: Wiley-Blackwell

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.
Late Quaternary climate legacies in contemporary plant functional composition

<table>
<thead>
<tr>
<th>Journal:</th>
<th><em>Global Change Biology</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>GCB-18-0470.R1</td>
</tr>
<tr>
<td>Wiley - Manuscript type:</td>
<td>Primary Research Articles</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>n/a</td>
</tr>
</tbody>
</table>
| Complete List of Authors: | Blonder, Benjamin; Arizona State University, School of Life Sciences
Enquist, Brian; University of Arizona, Ecology and Evolutionary Biology
Graae, Bente; Norwegian University of Science and Technology, Department of Biology
Kattge, Jens; Max-Planck-Institute for Biogeochemistry, Organismic Biogeochemistry
Maitner, Brian; University of Arizona, Ecology and Evolutionary Biology
Morueta-Holme, Naia; University of Copenhagen
Ordonez, Alejandro
Simova, Irena; Charles University
Singarayer, Joy; University of Reading
Svenning, Jens-Christian; Aarhus University, Department of Bioscience - Ecoinformatics and Biodiversity
Valdes, Paul; University of Bristol, School of Geographical Sciences
Violle, Cyrille; CNRS, Centre d'Ecologie Fonctionnelle et Evolutive |
| Keywords: | functional diversity, functional trait, disequilibrium, lag, climate change, legacy, immigration, exclusion |
| Abstract: | The functional composition of plant communities is commonly thought to be determined by contemporary climate. However, if rates of climate-driven immigration and/or exclusion of species are slow, then contemporary functional composition may be explained by paleoclimate as well as by contemporary climate. We tested this idea by coupling contemporary maps of plant functional trait composition across North and South America to paleoclimate means and temporal variation in temperature and precipitation from the Last Interglacial (120 ka) to the present. Paleoclimate predictors strongly improved prediction of contemporary functional composition compared to contemporary climate predictors, with a stronger influence of temperature in North America (especially during periods of ice melting) and of precipitation in South America (across all times). Thus, climate from tens of thousands of years ago influences contemporary functional composition via slow assemblage dynamics. |
Title
Late Quaternary climate legacies in contemporary plant functional composition

Running head
Climate legacies in functional composition

Authors
Benjamin Blonder 1,2,3 *
Brian J. Enquist 4,5
Bente J. Graae 2
Jens Kattge 6,7
Brian S. Maitner 4
Naia Morueta-Holme 8
Alejandro Ordonez 9,10
Irena Simova 11,12
Joy Singarayer 13
Jens-Christian Svenning 9,14
Paul J Valdes 15
Cyrille Violle 16

Affiliations
1: Environmental Change Institute, School of Geography and the Environment, University of
Oxford, Oxford, United Kingdom
For Review Only

2: Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway

3: School of Life Sciences, Arizona State University, Tempe, Arizona, USA

4: Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona, United States

5: Santa Fe Institute, Santa Fe, New Mexico, United States

6: Max Planck Institute for Biogeochemistry, Jena, Germany

7: German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Germany

8: Center for Macroecology, Evolution and Climate, Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark

9: Section for Ecoinformatics & Biodiversity, Department of Bioscience, Aarhus University, Aarhus C, Denmark

10: School of Biological Sciences, Queens University, Belfast, Northern Ireland

11: Center for Theoretical Study, Charles University, Prague, Czech Republic

12: Department of Ecology, Faculty of Science, Charles University, Prague, Czech Republic

13: Department of Meteorology, University of Reading, Reading, United Kingdom

14: Center for Biodiversity Dynamics in a Changing World (BIOCHANGE), Aarhus University, Aarhus, Denmark

15: School of Geographical Sciences, University of Bristol, Bristol, United Kingdom

16: CNRS, CEFE, Université de Montpellier – Université Paul Valéry – EPHE, Montpellier, France
*: Corresponding author, bblonder@gmail.com, +1 480 965 6419, School of Life Sciences, Arizona State University, Tempe, Arizona, USA

**Key words**
Functional diversity, functional trait, disequilibrium, lag, climate change, legacy, immigration, exclusion, Holocene, Pleistocene

**Submission type**
Primary research article
Abstract

The functional composition of plant communities is commonly thought to be determined by contemporary climate. However, if rates of climate-driven immigration and/or exclusion of species are slow, then contemporary functional composition may be explained by paleoclimate as well as by contemporary climate. We tested this idea by coupling contemporary maps of plant functional trait composition across North and South America to paleoclimate means and temporal variation in temperature and precipitation from the Last Interglacial (120 ka) to the present. Paleoclimate predictors strongly improved prediction of contemporary functional composition compared to contemporary climate predictors, with a stronger influence of temperature in North America (especially during periods of ice melting) and of precipitation in South America (across all times). Thus, climate from tens of thousands of years ago influences contemporary functional composition via slow assemblage dynamics.
Introduction

Shifts in the functional composition of plant communities can indicate variation in ecosystem functioning and ecosystem services (Chapin et al., 2000, Díaz & Cabido, 2001, Hooper et al., 2005, Jetz et al., 2016). Forecasting the two components of functional composition, functional trait means (FM) and functional diversity (FD) (Villéger et al., 2008), is therefore of central interest. Insights into geographic variation in the contemporary functional composition of plant communities (Violle et al., 2014) comes from field surveys (Asner et al., 2014, Baraloto et al., 2010, De Bello et al., 2006), macroecological approaches (Campbell & McAndrews, 1993, Lamanna et al., 2014, Šímová et al., 2015, Swenson et al., 2012), and remote sensing approaches (Asner et al., 2017a, Asner et al., 2017b, Jetz et al., 2016). However, little is known about changes in these functional trait patterns over longer time scales (Blonder et al., 2014, Polly et al., 2011, Thuiller et al., 2008). There is also growing evidence that paleoclimate has directly and indirectly structured contemporary species composition and functional composition (Ordonez & Svenning, 2016, Svenning et al., 2015). It has been unclear how these paleoclimate effects on species composition translate to differences in functional composition, because even species assemblages in disequilibrium with contemporary climate may have equilibrium functional relationships with contemporary climate (Fukami et al., 2005).

A core hypothesis of plant functional ecology is that contemporary environments determine contemporary functional composition (Enquist et al., 2015, Grime, 1974, Raunkiær, 1907, Schimper, 1898, von Humboldt & Bonpland, 1807 (tr. 2009)). Many studies have shown relationships between FMs or FD and contemporary environmental variables, e.g. Cornwell and Ackerly (2009), Moles et al. (2014), suggesting equilibrium with contemporary environmental conditions is plausible. However, paleoclimate may also have had a strong influence on
contemporary functional composition at large spatial scales (Svenning et al., 2015). A mismatch could exist between contemporary climate and contemporary FMs and FD because of disequilibrium in species’ geographic ranges and lack of more appropriate species in the regional pool (Davis & Shaw, 2001, Enquist et al., 2015). Mechanisms that could lead to differing degrees of lagged responses of FMs and FD, and thus disequilibrium, include differential rates of exclusion and immigration driven by variation in dispersal limitation, longevity, and species interaction strengths that are associated with certain functional traits (Davis, 1984, Eiserhardt et al., 2015, Enquist et al., 2015, Svenning & Sandel, 2013, Webb, 1986). Evidence for disequilibrium in functional composition is growing. For example, instability in climate in the Late Quaternary may have influenced contemporary functional composition in Europe (Mathieu & Jonathan Davies, 2014, Ordonez & Svenning, 2015, Ordonez & Svenning, 2017, Svenning et al., 2015) and in the Americas (Ordonez & Svenning, 2016).

Paleoclimate influences on plant species composition are better known. For example, many tropical forests and temperate understory assemblages have compositions lagging contemporary climate changes at $10^1$-$10^3$ year timescales (Campbell & McAndrews, 1993, Cole et al., 2014, DeVictor et al., 2008, La Sorte & Jetz, 2012). At $10^3$-$10^5$ year timescales, the European flora (Svenning & Skov, 2007) and North American plant range size distributions (Morueta-Holme et al., 2013) show strong signals of slow recovery from cover of ice sheets due to late-Quaternary glaciation. At $10^5$-$10^6$ year timescales, African and Madagascan palm distributions can be predicted by Pliocene precipitation patterns (Blach-Overgaard et al., 2013, Rakotoarivonina et al., 2013). Last, at $10^6$-$10^7$ year timescales, Cenozoic climate change and land connectivity shifts have resulted in cold tolerance-driven extinction of some temperate trees.
(Eiserhardt et al., 2015), and have limited the dispersal and radiation of certain clades (Morley, 2011, Woodruff, 2010).

We first test a hypothesis (Hypothesis 0) that paleoclimate has additional predictive power for functional composition beyond that provided by contemporary climate. We do so by determining whether FMs or FD are best predicted by contemporary climate alone or by contemporary climate and paleoclimate together.

We also test four hypotheses for how paleoclimate and contemporary climate could influence contemporary FMs and FD (Figure 1). The hypotheses explore fast vs. slow processes for exclusion and immigration of species under linear change in a mean climate value (Blonder et al., 2017). ‘Fast’ and ‘slow’ are terms used to indicate temporal rates of change in species composition and functional traits relative to the rate of climate change; mechanisms underlying exclusion and immigration could include ecological processes such as environmental filtering, competition, or dispersal or evolutionary processes such as speciation, adaptation, or extinction.

These hypotheses are thus relevant over intervals where changes in climate can be treated as linear. They also all assume an underlying linear trait-environment relationship that would be obtained in the equilibrium limit.
**Figure 1.** Four hypothetical scenarios for the relationship between contemporary functional traits and climate change. Inset panel shows the assumed equilibrium trait-environment relationship.

**A)** Hypothesis 1, fast exclusion and fast immigration: species will track contemporary climate, and there will be a strong contemporary climate mean – functional trait mean relationship. **B)** Hypothesis 2, slow exclusion but fast immigration: many species that were at one time suitable still remain part of the assemblage, and there will be a positive relationship between paleoclimate temporal variation and functional diversity. **C)** Hypothesis 3, fast exclusion but slow immigration: only species that were at all times suitable will be able to enter the assemblage, and there will be a negative relationship between paleoclimate temporal variation and functional diversity. **D)** Hypothesis 4, slow exclusion and slow immigration: species will fail to track contemporary climate, and there will be a strong paleoclimate mean – functional trait mean relationship.
In Hypothesis 1 (Figure 1A), if exclusion of species with inappropriate traits for a novel climate is fast and if immigration of more appropriate species is fast, then contemporary climate mean – contemporary FM relationships will exist. In Hypothesis 2 (Figure 1B), if exclusion of species with inappropriate traits is slow and if immigration of more appropriate species is fast, paleoclimate temporal variation – contemporary FD relationships will be positive because more species with appropriate traits are continually added to the assemblage without loss of other species. In Hypothesis 3 (Figure 1C), if exclusion of species with inappropriate traits is fast but if immigration of appropriate species is slow, then paleoclimate temporal variation – contemporary FD relationships will be negative because species with inappropriate traits become lost from an assemblage without replacement by other species. In Hypothesis 4 (Figure 1D), if exclusion is slow and if immigration is slow, then paleoclimate mean – contemporary FM relationships will exist because of temporally lagged losses and gains of suitable species.

These four hypotheses provide non-exclusive predictions of relationships between climate and functional trait patterns. More than one of these patterns could be simultaneously observed, depending on the dynamics of climate over a long period comprising multiple approximately linear changes. That is, predictions of relationships between e.g. paleoclimate variation and contemporary FD do not preclude observation of relationships between paleoclimate mean and contemporary FMs.

Here, we ask: 1) whether paleoclimate means and temporal variation improve predictions of contemporary FMs and FD (Hypothesis 0), and 2) which of the proposed hypotheses (Hypothesis 1 - Hypothesis 4) are consistent with empirical patterns of contemporary FMs and FD. We derived gridded maps of contemporary FMs and FD (as convex hull volume (Cornwell et al., 2006)) across the Americas by merging species-mean trait data with maps of species
distributions. We used five plant functional traits that are representative of major ecological
strategy axes (Díaz et al., 2016, Westoby & Wright, 2006), and predictive of species sorting
along environmental gradients (Moles et al., 2014, Simova et al., 2018, Šimová et al., 2015). We
then coupled these estimates with contemporary and paleoclimate maps at timescales spanning
the Last Interglacial (120 ka) to the present. We chose climate axes of mean annual temperature
and annual precipitation because of their established trait-environment relationships (Moles et
al., 2014), and their ability to be reconstructed by general circulation models.

**Materials and Methods**

*Species distribution maps*

We obtained occurrence data for New World plants from the BIEN database (version 3.0, access
date 26 February 2017, http://www.biendata.org) (Enquist et al., 2009, Enquist et al., in
preparation, Maitner et al., 2017). Following Morueta-Holme et al. (2013), we selected only data
that represented geo-validated and non-cultivated occurrences, and standardized all taxonomic
names (Boyle et al., 2013). Occurrence points were non-randomly distributed, with higher
observation densities in the continental United States and in Central America / northwestern
South America.

To reduce biases from spatial variation in sampling intensity, we estimated species’
geographic ranges using convex hulls (Elith & Leathwick, 2009). Convex hulls can be estimated
without using climate variables for niche modeling, avoiding any potential circularity in our
analyses that would be caused by (for example) a maximum entropy model calibrated on
contemporary climate variables. We generated range polygons from latitude/longitude
coordinates for species with more than three non-collinear observation points. For species with
three or fewer observations (6,886/74,491 species=9.2%), we assumed that the species was present only in the 100×100-km grid cell(s) containing the observation. We rasterized predictions over the Western Hemisphere on a 100×100-km grid cell equal area projection centered at 80°W, 15°N.

Functional trait data

We selected five functional traits representing major ecological strategy axes for growth, survival, and reproduction (Díaz et al., 2016, Westoby & Wright, 2006). These included specific leaf area, plant height, seed mass, stem specific density, and leaf nitrogen. Trait data were obtained from the TRY database (https://www.try-db.org, accession date 19 June 2013) (Kattge et al., 2011), covering 45,507 species (7,051 genera). A list of data references is in Table S1. Because many taxa were missing some observations of certain variables, a phylogenetic gap-filling approach (Schrodt et al., 2015) was used to estimate missing values; then for a fraction of taxa that were present in the occurrence data but not present in the TRY data (59,423 species, 3343 genera), missing values were filled with genus means estimated from the TRY data. This approach likely results in less bias than omitting data for species without exact matches to trait data.

We also categorized each species by its growth habit. Using a New World database (Engemann et al., 2016), we classified species as woody (29,676 species) or non-woody (44,324 species). Analyses were carried out for either all or only woody species to distinguish potentially different climate drivers on traits between growth forms (Díaz et al., 2016, Simova et al., 2018).

Functional trait mapping
We used the distribution maps to estimate the species composition within each grid cell. We then matched this species list against the functional trait data to estimate the distribution of log-transformed traits within each grid cell. To reduce undersampling biases, we then removed from the analysis all cells with richness of species with non-missing trait values less than 100 (a value chosen to be small, in this case representing the 7% quantile of the data, and which primarily removes extreme-latitude cells in Greenland and Ellesmere Island in the northern hemisphere, and Tierra del Fuego in the southern hemisphere) (Figure S1). This procedure was repeated for woody species and for all species.

To estimate FMs, we calculated the average trait value across all species occurring within the cell, based on overlapping range maps. To simplify these five axes, we calculated a ‘FM PC1’, defined as the score along the first principal component of the five mapped trait axes. This axis explained 83.5% of the variation in traits for the woody species subset and 74.5% of the variation for all species, and corresponds to increasing plant height, seed mass, and stem specific density, as well as decreasing SLA and leaf nitrogen content (Figure S2).

To estimate FD, we first calculated the five-dimensional convex hull volume across log-transformed values of each trait value occurring within the cell (Villéger et al., 2008). Second, we corrected this estimate because convex hull volumes often scale linearly with sample size, and because the fraction of species per grid cell with available trait measurements (‘trait coverage’) was variable (78% mean, 10% s.d). This value was sufficiently high to lead to limited bias in FM estimates, according to simulations (Borgy et al., 2017c). To correct for the sample size effect in FD, we built a null model. We calculated the convex hull volume of random samples of the full trait dataset with species richness varying from 100 to 10,000 in steps of 100 (‘FDtrue’), then subsampled each to a trait coverage value varying from 0.05 to 1.0 in steps of
0.05, and then recalculated the convex hull volume based on this subsample (‘FD$_{\text{observed}}$’). We repeated the convex hull volume calculation 10 times for each combination of species richness and trait coverage. We then built a linear model to predict FD$_{\text{true}}$ on the basis of linear terms of FD$_{\text{observed}}$, species richness, and trait coverage, as well as 2-way and 3-way interactions between these variables. This model explained 95.8% of the variation in FD$_{\text{true}}$. We therefore applied this model to FD$_{\text{observed}}$ in the empirical data to yield a corrected estimate of FD$_{\text{true}}$ (hereafter, FD) that accounted for variation in trait coverage.

FD$_{\text{observed}}$ and species richness are positively correlated, because as species richness increases within a grid cell, FD$_{\text{observed}}$ can only stay constant or increase. Thus, it may be difficult to separate effects of paleoclimate-related processes on FD from effects on species richness. To partially address this issue, we also repeat all analyses for another composite variable FD$_{\text{res}}$, defined as the residuals of a regression of the corrected estimate of FD (FD$_{\text{true}}$) on species richness. Thus positive values of FD$_{\text{res}}$ indicate FD values higher than expected based on a random assemblage with the same species richness, while negative values indicate values lower than expected.

Contemporary climate data

We obtained contemporary climate predictions (1979-2013 AD averages) for mean annual temperature (MAT) and mean annual precipitation (MAP) from the CHELSA dataset version 1.2 (available at http://chelsa-climate.org/) (Karger et al., 2016). The climate dataset is based on a quasi-mechanistic statistical downscaling of the ERA (European Re-Analysis) interim global circulation model with a GPCC (Global Precipitation Climatology Centre) bias correction, and incorporating topoclimate (Karger et al., 2016). This approach avoids biases inherent to
interpolation between weather stations with uneven coverage of geographic and climate space. We then re-projected the 1-arcsecond resolution data to the same grid as the species distribution maps.

Paleoclimate data

We obtained paleoclimate data from the HadCM3 general circulation model. The HadCM3 model consists of a coupled atmospheric, ocean, and sea ice model with non-interactive vegetation, with an atmospheric resolution of 2.5° latitude × 3.75° longitude. The model was driven by variations in orbital configuration, greenhouse gases, ice-sheet topography, and coincident sea level changes and bathymetry since 120 ka. Simulations included the effects of abrupt “fresh-water” pulses and the resulting abrupt climate changes that occurred during at 17 ka (Heinrich event) and 13 ka (Younger Dryas). Boundary conditions and spin-up are fully described in Hoogakker et al. (2016), Singarayer and Valdes (2010). Data were available at time points beginning 0-120 ka in 1 kyr slices from 1–22 ka, in 2 kyr slices from 22–84 ka, and in 4 kyr slices from 84–120 ka (example time slices in Figure 2, all time slices in Figure S3, S4).
Figure 2. Example contemporary climate and HadCM3 general circulation model temporal mean values for annual temperature (MAT) and annual precipitation (MAP) as well as for temporal standard deviations of MAT and MAP for the present day (0 ka), and for intervals beginning at 13 ka (Younger Dryas), 21 ka (Last Glacial Maximum), and 120 ka (Last Interglacial). Colors are scaled and transformed (see Materials and Methods), with labels indicating values back-transformed to original units. The full analysis includes a larger number of temporal mean values at intervals spaced between 0 – 120 ka.
Model output was re-projected to the same coordinate system and resolution as the contemporary species distribution maps. This approach assumes a negligible impact of variation in sea level on the vast majority of pixels and is appropriate given that only contemporary functional composition data were available. Paleoclimate maps are close to contemporary climate maps during the Holocene, and diverge strongly during the Pleistocene, as measured by mean absolute deviation between contemporary and paleoclimate pixel values (Figure S5).

**Statistical analysis**

To prepare climate data for analysis, we first square-root transformed contemporary and paleoclimate MAP data to improve normality. We calculated a temporal mean value at \( x \) ka, for \( x \) in 0 to 120, as well as a temporal standard deviation at \( x \) ka within each grid cell using a moving window approach, i.e. over values within the interval \([x-k, x+k]\). These temporal standard deviations were then standardized by divided by the total temporal range of the moving window. Temporal standard deviations thus have units of either °C kyr\(^{-1}\) or mm kyr\(^{-1}\). We used a value of \( k=1 \) where possible, but \( k=4 \) in some cases where HadCM3 data had coarser resolution (i.e. closer to 120 ka). Edge cases at 0 and 120 ka were calculated treating out-of-range data as missing. Contemporary climate was used for values at 0 ka, while paleoclimate was used for values at 1-120 ka.

We then rescaled all contemporary and paleoclimate predictor variables by \( z\)-transforming each relative to their grand mean and standard deviation (over all pixels and years) for each variable type from the HadCM3 model (MAT and MAP mean values and temporal standard deviation of MAT and MAP). This approach standardizes values across both variable types and models relative to estimates of their ranges across study interval. Thus, a value of +1 in
a MAT layer indicates that this cell has a value that is 1 standard deviation larger than the mean value relative to all values seen in all locations over the 0-120 ka interval.

We used partial least squares (PLS) regression to determine the best predictors of FMs, FD, and FDres in independent analyses. We conducted PLS regressions separately for North America and South America (split at the Panama/Colombia border) because of their different glaciation histories (Ehlers et al., 2011). The PLS approach accounts for the statistical non-independence of large numbers of predictors by finding the rotation of the predictor matrix that best overlaps with the response vector, and identifies the latent factors (components) that correspond to these rotations (Geladi & Kowalski, 1986). The PLS components describe the independent contribution of each predictor variable to the response variable and are ordered by their explanatory capacity such that the first component (PLS1) by definition explains the most variation in the data. Thus the approach can identify independent effects of multiple correlated predictors (i.e. separating the effects of contemporary and paleoclimate, even if they are sometimes correlated with each other). We built models that simultaneously incorporated up to six classes of predictors: contemporary climate mean values, paleoclimate temporal mean values, and paleoclimate temporal standard deviations (metrics of paleoclimate variation) for each of MAT and MAP.

We also performed a separate set of PLS analyses in order to assess biases from climate changes occurring at times and locations where plants could not have grown. Although predicting ice sheet spatial coverage at each time and location would be ideal, we instead masked out pixels at all times and places where there was ice cover during the Last Glacial Maximum (21 ka) (corresponding to pixels in the black polygon in Figure 2I). This choice was motivated
by the currently limited knowledge of temporally-resolved ice sheet dynamics during the full
extent of study period (Kleman et al., 2013, Kleman et al., 2010).

We tested Hypothesis 0 for each of FMs, FD, or FDres by comparing root mean square
error of prediction (RMSEP) values for PLS models that included contemporary climate ($n = 2$
total predictors) and/or paleoclimate values ($n = 250$ total predictors). Because RMSEP
necessarily decreases with number of PLS components, we compared RMSEP values after fixing
the number of PLS components in each model. This approach is more appropriate than model
selection methods based on Akaike Information Criterion comparisons (Li et al., 2002) because
it is difficult to calculate degrees of freedom in PLS in order to correctly penalize likelihood
values (Krämer & Sugiyama, 2011).

In this PLS framework, Hypotheses 1–4 can be distinguished by regression of
contemporary FMs, FD, or FDres on contemporary climate mean values, paleoclimate mean
values over multiple times, and paleoclimate temporal variation over multiple times. We
assessed the importance of each PLS component via the percentage of variance explained by the
component. The effect of each variable at each time for FMs, FD, or FDres can be interpreted as
the PLS component’s loading coefficient explaining the most variance in each model, with
positive loading coefficients indicating that higher than average (over the 0-120 ka interval)
values of this predictor yield higher than average values of the response variable. We also
defined an overall effect for each class of predictor as the maximum absolute loading coefficient
for that predictor type along each axis across all times.

All analyses were carried out with the R statistical environment (version 3.3.3).

Occurrence data were obtained with the ‘BIEN’ package (Maitner et al., 2017). Map rescaling
and re-projection were carried out with the ‘raster’ (Hijmans & van Etten, 2014) and ‘maptools’
(Bivand & Lewin-Koh, 2013) packages. Convex hulls were calculated with the ‘geometry’
package (Habel et al., 2015). PLS regression was carried out within the ‘pls’ package (Mevik &
Wehrens, 2007).

Results

Contemporary functional trait patterns

Mapped FMs for all species for the five focal functional traits showed strong spatial gradients. Mean estimates of specific leaf area were highest in temperate/boreal North America (Figure
3A). Maximum plant height and seed mass were highest in the eastern Amazon basin (Figure
3B, 3C). Stem specific density was highest in the Amazon basin (Figure 3D). Leaf nitrogen
content was highest in western North America and the southern South America (Figure 3E), all
leading to similar latitudinal tropical-temperate-boreal gradients in FMs for PC1 (Figure 3F).

FD was high throughout the tropics and into southeastern North America (Figure 3G), and FD$_{res}$
was high in southeastern North America, Central America, and the Caribbean, as well as along
the northeastern and eastern coasts of South America (Figure 3H). Species richness was highest
in Central America and the western Amazon basin (Figure 3I). All of these results were
qualitatively consistent when restricted to woody species only (Figure S6).

Figure 3. Estimated plant species assemblage characteristics, based on data for all species.
Distributions of functional trait means (FMs) for five functional traits (each colored by log-
transformed values, with labels indicating values back-transformed to original units) are shown
for A) Specific leaf area, B) plant height, C) seed mass, D) stem specific density, and E) leaf
nitrogen per unit mass. F) First principal component of FMs. G) Functional diversity (FD;
convex hull volume of log$_e$-transformed values); H) FD$_{res}$, the residual of FD regressed on
species richness, and I) Species richness. The black polygon indicates the maximum ice sheet
extent during the Last Glacial Maximum.
Overall predictive power of paleoclimate

We found that models that incorporated paleoclimate and contemporary climate had higher predictive power than models that incorporated only contemporary climate (Figure 4). When comparing models with the same number of PLS components, the contemporary + paleoclimate models usually had equivalent or lower root mean square error of prediction (RMSEP) than the
contemporary climate models. For example, for FD calculated with data for all species and HadCM3 climate data, using 1 PLS component, RSMEP was 9% lower in North America and 20% lower in South America; when using data for woody species, RMSEP was 14% lower in North America and 20% lower in South America. Similar results held for all other response variables, other methodological choices, and 2 PLS components (Figure S7).
**Figure 4.** Predictive uncertainty in models for FD as measured by the cross-validated root mean squared error of prediction (RMSEP) for increasing numbers of PLS components. Y-axis units correspond to units of functional diversity (compare to Figure 3G). Results are for PLS regression models generated using trait data for all species and climate data from HadCM3. Orange lines indicate models using only contemporary climate predictors; blue lines, models using contemporary and paleoclimate predictors.
Paleoclimate and contemporary climate predictors of contemporary functional composition

We present results for the HadCM3 paleoclimate model using all species, as results are representative across all modeling choices.

For FD in North America, we found that the first PLS component explained 57% of the variation in the data (Figure 5A). This component represented large effects (> 0.1 in absolute standard deviations) for paleo MAT mean value (+0.11), paleo MAT temporal standard deviation (-0.20), and for paleo MAP temporal standard deviation (+0.11). There were no large effects from contemporary MAT or MAP mean values. These effects were strongest immediately after the Last Glacial Maximum (~20 ka) and the Last Interglacial (~120 ka).

For FD in South America, we found that the first PLS component explained 60% of the variation in the data (Figure 5B). This component represented large effects for contemporary MAP mean value (+0.13), paleo MAP mean value (+0.10), and paleo MAP temporal standard deviation (+0.28). There was no large effect from any MAT predictor. Paleo MAP temporal standard deviation was most important at time periods beginning at 17 ka and 13 ka, corresponding to abrupt change from a Heinrich event and the Younger Dryas, respectively.
Figure 5. Contemporary climate and paleoclimate effects on functional diversity (FD), for the first PLS component, for A) North America, and B) South America. Results are for models generated using trait data for all species and using climate data from HadCM3. Left subpanels indicate effect sizes (loading coefficients) for each model component at different times. Contemporary climate data are shown in triangles; paleoclimate values as dark lines and temporal standard deviations as lighter lines. Red indicates MAT, blue MAP. Right subpanel symbols indicate the maximum absolute effect for each variable class over time. A gray background rectangle indicates a significance threshold. Orange shading behind each panel indicates global atmospheric temperatures reconstructed by Bintanja et al. (2011), with deeper shading indicating warmer conditions. Results for FMs and FD_{res} are shown in Figure S8.
Results for FMs and FD\textsubscript{res} were similar to those for FD. One exception occurred in South America, where estimates for FD\textsubscript{res} were opposite in sign (Figure S8). Results for higher PLS components are not reported, as explained variation for each was individually low (e.g. at most 7\% for PLS2 across all response variables using the HadCM3 data and all species across response variables). Model residuals for North and South America for varying numbers of components are shown in Figure S9.

All of the above results were qualitatively similar when restricting data to woody-only species (Figure S10). Analyses were also qualitatively similar when excluding pixels covered by ice sheets at the Last Glacial Maximum. Results for these analyses are presented in Figure S11.

Discussion

We identified spatially and temporally variable effects of paleoclimate on contemporary functional trait patterns, independent from those of contemporary climate. Across methodological choices, functional composition was predicted in North America by paleo MAT mean values, paleo MAT temporal standard deviations, and paleo MAP temporal standard deviation, and in South America by paleo MAP mean values and paleo MAP mean values. Paleo MAT and MAP mean values had similar effects over time, while in North America MAT temporal standard deviation at the Last Glacial Maximum and Last Interglacial had strongest effects, and in South America MAP temporal standard deviation at the Younger Dryas and the 17\textsuperscript{ka} Heinrich event had strongest effects. Thus climate immediately after the Last Glacial Maximum appears to have left a strong legacy on contemporary functional composition. We also found that paleoclimate was a useful predictor of contemporary functional composition, supporting Hypothesis 0. Predictive errors for predicting FMs, FD, and FD\textsubscript{res} were lower when
paleoclimate variables were incorporated into regression models than when only including contemporary climate variables.

The PLS models support several of the hypotheses. Hypothesis 1 (a relationship between contemporary FMs and contemporary climate mean values, with fast immigration and fast exclusion) was supported in South America for MAP. Hypothesis 2 (a positive relationship between contemporary FD and paleoclimate temporal standard deviation, with fast immigration slow exclusion) was supported for MAP in North America and in South America. Hypothesis 3 (a negative relationship between contemporary FD and paleoclimate temporal standard deviation, with slow immigration and fast exclusion) was supported for MAT in North America. Hypothesis 4 (a relationship between contemporary FMs and paleoclimate mean values, with slow immigration and slow exclusion) was supported for MAP in North and South America.

Thus, all of the scenarios of Figure 1 received some support in either North or South America. The general implication is that processes of species immigration or exclusion can sometimes be slow, leading to spatial variation in colonization and extinction debts across these continents.

The results therefore do not map cleanly onto any one class of dynamics dominating at continental scales. Elucidating the details of these sometimes slow immigration and exclusion dynamics more precisely would require comparing time series of functional composition to time series of paleoclimate (Blonder et al., 2017). That approach contrasts with the approach taken in the present study, which compared time series of paleoclimate to a single time-point estimate of functional composition, and tested hypotheses most relevant for single linear climate changes. Time series data for functional composition are highly challenging to obtain from available paleoprosxies. However, such data would enable direct measurement of the rates and lags in temporal response of functional composition to climate variation.
Results in North America are consistent with limited dispersal after ice sheet retreat (Davis & Shaw, 2001, Morueta-Holme et al., 2013, Svenning et al., 2015), and on thermal tolerances that constrain species distributions in high-latitude environments (Hawkins et al., 2013, Körner, 2003, Morin & Lechowicz, 2011, Sakai & Weiser, 1973). The paleoclimate MAT signal seen in these data may be driven by cooling in temperate and boreal portions of the continent during the last glacial period that have caused regional extinctions and slow recolonization dynamics (Davis, 1984). These findings extend the spatial and temporal extent of analyses exploring glacial effects on biodiversity (Ordonez & Svenning, 2017), providing additional confidence that this period plays a key role in shaping contemporary biodiversity patterns.

Results in South America supported the role for paleoprecipitation variation in shaping contemporary biodiversity patterns in tropical areas (Blach-Overgaard et al., 2013, Göldel et al., 2015, Rakotoarinivo et al., 2013), possibly by survival and recolonization from refugia along hydrological gradients. Lower precipitation values and higher precipitation temporal variation in the Late Pleistocene in certain coastal regions of this continent have led to contemporary FD being lower than expected based on contemporary climate. The strong precipitation effects in South America caused by Northern hemisphere ice melting during the 17 ka Heinrich event and the Younger Dryas are consistent with strong cross-hemisphere telecoupling of climate during these intervals, in which ice sheets and ice melting in the Northern hemisphere caused atmospheric and ocean circulation changes, leading to changes in Southern hemisphere climate regimes (Clement & Peterson, 2008, Jones et al., 2018). This result suggests that other climate telecoupling may also drive initially unintuitive relationships between climate change and functional composition change.
The spatial uncertainties in our results are possibly large. Biases in trait data coverage could spatially bias our maps of FMs and FD if botanical collecting effort in certain areas were focused on certain taxonomic or functional groups (Borgy et al., 2017b). Because our maps of functional composition are broadly consistent with other estimates (Butler et al., 2017, Simova et al., 2018, van Bodegom et al., 2014), this is unlikely to be a major concern. Nevertheless, trait data and species occurrence are poor in some regions (e.g. the central Amazon, as well as southern South America). Thus, this approach is unlikely to be able to parse out sub-regional biodiversity patterns because of limitations in available data. The spatial resolution of paleoclimate simulations (>2° per grid cell) also limits parsing of sub-regional spatial patterns due to within-pixel climate heterogeneity (Stein et al., 2014). Nevertheless, the broad consistency of our findings across methodological choices gives some confidence in the generality of our conclusions.

The temporal uncertainties in our results are probably smaller than the spatial uncertainties. The HadCM3 simulations included multi-millennial drivers of climate change (orbit, greenhouse gases, ice sheets), as well as the Heinrich event at 17 ka (Hemming, 2004) and the Younger Dryas event at 13 ka (Alley, 2000). Detailed simulations of similar events in deeper time were not available (e.g. the Heinrich event at ~45 ka (Hemming, 2004), or Dansgaard-Oschger millennial events that may increase the variability of temperature and precipitation, especially between 30 and 60 ka), but it is possible that these events also have large and persistent effects on contemporary functional composition. Regardless, these models provide some of the best available estimates of past climates, though independent paleo-proxy validation of predictions remain sparse, especially in South America (Harrison et al., 2014).
Non-climate factors may also be important drivers of functional composition over multiple timescales. For example, past human impacts on landscapes via active propagation, land clearance, or fire regimes (Bond & Keeley, 2005, Keeley et al., 2011) are widely acknowledged throughout tropical (Levis et al., 2017, Malhi, 2018, Ross, 2011) and temperate (Abrams & Nowacki, 2008, Borgy et al., 2017a, Feng et al., 2017, Nowacki & Abrams, 2008) regions. Soil and surficial geology may also be important in determining plant species distributions (Ordoñez et al., 2009). However, the mechanisms linking specific traits to different non-climate abiotic variables are not yet completely clear. Moreover, all of these variables remain difficult and controversial to estimate over time and space. While we were unable to include them in our analysis, there is likely scope to extend our approach as datasets improve.

Climate may also indirectly drive changes in functional composition through changes in species interactions. Megafauna had large impacts on plant assemblages. These impacts would have shifted after the extinction of many megafauna in North and South America during the late Pleistocene (Gill et al., 2009, Johnson, 2009). While humans are acknowledged to be a major driver of these extinctions (Lorenzen et al., 2011), many also were strongly linked to climate change during this period on these continents (Bartlett et al., 2016). Indeed, some of the changes in immigration and exclusion rates could have been driven indirectly by these organisms, e.g. reduction in seed dispersal services leading to slow immigration (Pires et al., 2018) (but see (van Zonneveld et al., 2018)), or reduced trampling leading to slow exclusion (Bakker et al., 2016). The temporal and spatial dynamics of megafaunal distributions remains poorly constrained by data, but such information may ultimately provide additional insight into climate-linked drivers of plant functional composition.
Our findings suggest that when predicting the future response of biodiversity to climate change, disequilibrium effects due to slow immigration or exclusion may be important. Statistical models based on the assumption that trait-environment relationships calibrated from contemporary climate data are at equilibrium (Laughlin et al., 2012, Shipley et al., 2006) could potentially be improved by incorporating paleoclimate predictors. Alternatively, it could be useful to include more mechanistically model processes of slow immigration and/or exclusion dynamics (Blonder et al., 2017, Fukami, 2015, Svenning et al., 2015). Such models, e.g. demography-constrained species distribution models (Zurell et al., 2016) or dynamic global vegetation models (van Bodegom et al., 2014), can represent disequilibrium dynamics that may result in nonlinear relationships between climate, paleoclimate, and functional traits.

The overall conclusion of our study is that functional trait patterns are predicted better by inclusion of paleoclimate than by contemporary climate alone, as seen via a Pleistocene temperature legacy in North America and a precipitation legacy in South America. While current functional composition may be well-adapted to contemporary environments, the high importance of paleoclimate suggests that the equilibrium assumption of functional ecology may be inappropriate for plant functional traits over $10^3$-$10^5$ yr timescales and continental spatial scales. The interplay between contemporary climate and paleoclimate drivers of biodiversity patterns will need to be better understood in order to accurately predict assemblage responses to future climate change.
Acknowledgements

BB was supported by a UK Natural Environment Research Council independent research fellowship (NE/M019160/1) and the Norwegian Research Council (KLIMAFORSK 250233). JCS was supported by the European Research Council (ERC-2012-StG-310886-HISTFUNC), and also considers this work a contribution to his VILLUM Investigator project (VILLUM Fonden grant 16549). NMH was supported by the Carlsberg Foundation and acknowledges the Danish National Research Foundation for support to the Center for Macroecology, Evolution and Climate. IS was supported by the Czech Science Foundation (16-26369S). CV was supported by the European Research Council (ERC) Starting Grant Project “Ecophysiological and biophysical constraints on domestication in crop plants” (Grant ERC-StG-2014-639706-CONSTRAINTS). The study was supported by the TRY initiative on plant traits (http://www.try-db.org), which is hosted, developed, and maintained at the Max Planck Institute for Biogeochemistry, and further supported by DIVERSITAS/Future Earth and the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig. This work was conducted as part of the Botanical Information and Ecology Network (BIEN) Working Group (PIs BJE, R. Condit, RK Peet, B Boyle, S Dolins and BM Thiers) supported by the National Center for Ecological Analysis and Synthesis, a center funded by the National Science Foundation (EF-0553768), the University of California, Santa Barbara, and the State of California. The BIEN Working Group was also supported by the iPlant collaborative and the National Science Foundation (DBI-0735191). We also thank all BIEN data contributors (see http://bien.nceas.ucsb.edu/bien/people/data-providers/ for a full list). Lotte Nymark Busch Jensen assisted with preparing Figure 1.
Statement of authorship

BB conceived the project and carried out analyses. JK provided trait data. JS, PV, and AO provided paleoclimate data. BJE and JCS provided species occurrence data. NMH contributed to species distribution modeling. All authors contributed to writing the manuscript. Authors were ordered alphabetically by last name after the first author.

Data accessibility

All georeferenced data products underlying this analysis are available in File S1 and also will be deposited in Dryad upon acceptance.
References

Abrams MD, Nowacki GJ (2008) Native Americans as active and passive promoters of mast and fruit trees in the eastern USA. The Holocene, 18, 1123-1137.


Enquist BJ, Sandel B, Boyle B et al. (in preparation) Plant diversity in the Americas is driven by climatic-linked differences in evolutionary rates and competitive displacement.


Hoogakker BaA, Smith RS, Singarayer JS et al. (2016) Terrestrial biosphere changes over the last 120 kyr. Climate of the Past, 12, 51-73.


Karger DN, Conrad O, Böhner J et al. (2016) CHELSA climatologies at high resolution for the earth's land surface areas (Version 1.1).


Raunkiær CC (1907) *Planterigets livsformer og deres betydning for geografien*, Kjøbenhavn og Kristiania, Gyldendalske boghandel, Nordisk forlag.


Supporting Information

File S1 – Processed raster maps (ASCII grid format) for FM, FD, FD_{res}, and contemporary and paleoclimatic means and temporal standard deviations. Includes metadata file (PDF format) explaining data variables, units, and provenance.
Figure S1. Summary of data coverage. A) Raw counts of occurrences for the BIEN3 database. B) Number of species for which trait data were available. C) Number of woody species for which trait data were available. Note log_{10} scale for all panels.
Figure S2. Principal component analysis of log-transformed trait values for A) all species and B) only woody species.
Figure S3. Paleotemperature predictions from the HadCM3 model for 0 – 120 ka. Values are reported in scaled coordinates relative to mean and standard deviation across all pixels and all times.
Figure S4. Paleoprecipitation predictions from the HadCM3 model for 0 – 120 ka. Values are reported in scaled coordinates relative to mean and standard deviation across all pixels and all times.
Figure S5. Deviation between present-day climate and paleoclimates at different past times for the HadCM3 model. Y-axis values indicate the mean absolute deviation between contemporary and paleoclimate pixel values in transformed coordinates (standard deviations relative to 0-120 ka ranges). Blue lines, mean annual precipitation; red lines, mean annual temperature.
**Figure S6.** Estimated plant species assemblage characteristics, based on data for only woody species. Compare caption to **Figure 3.**
Figure S7. Predictive ability of PLS models usually increases when including paleoclimate predictors as well as contemporary climate predictors. Bar height indicates percent decrease in RMSEP of each model (paleo. + contemp. relative to contemp. only) for different variables (bar colors – red, FD, green FDres, blue, FM (PC1)). Plots are shown for models for each continent, and for every combination of trait data (woody vs. all species) and number of PLS components (1 contemp. PLS axis vs. 1 paleo. + contemp. PLS axis, or 2 contemp. PLS axes vs 2 paleo. + contemp. PLS axes).
Figure S8. Contemporary climate and paleoclimate effects on contemporary FM (PC1) (A,B) and FDres (C,D) using trait data for all species and climate data from HadCM3. Compare caption to Figure 5.
Figure S9. Residuals of PLS regression model for FD. Results are based on trait data for all species and climate data from HadCM3. Panels indicate the number of PLS components included in the model ($n$) and the cross-validated root mean square error of prediction (RMSEP). Overpredicted values are shown in red and under-predicted values are shown in blue.
Figure S10. Contemporary climate and paleoclimate effects on contemporary FD (A,B), FMs (C,D) and FD\textsubscript{res} (E,F) using trait data for woody species and climate data from HadCM3. Compare caption to Figure 5.
Figure S11. Contemporary climate and paleoclimate effects on contemporary FD (A,B), FMs (C,D) and FD_{res} (E,F) using trait data for all species and climate data from HadCM3. In this analysis, locations under ice at the Last Glacial Maximum are wholly excluded from the analysis. Compare caption to Figure 5.
Table S1. Original data sources for trait data extracted from the TRY database.


Auger, S. 201. MSc thesis, Université de Sherbrooke, Sherbrooke (Quebec) L'importance de la variabilité interspécifique des traits fonctionnels par rapport à la variabilité intraspécifique chez les jeunes arbres en forêt mature.


Campetella, G; Botta-Dukát, Z; Wellstein, C; Canullo, R; Gatto, S; Chelli, S; Mucina, L; Bartha, S. 2011. Patterns of plant trait-environment relationships along a forest succession chronosequence. Agriculture, Ecosystems & Environment 145: 38-48


structure and function of Hawaiian Metrosideros polymorpha at high vs. low precipitation.

Craine JM, Towne EG, Ocheltree TW, Nippert JB (2012) Community traitscape of foliar

Penuelas, P. B. Reich, E. A. G. Schuur, W. D. Stock, P. H. Templer, R. A. Virgini


and within-site comparisons of structural and physiological characteristics and foliar nutrient
content of 14 tree species at a wet, fertile site and a dry, infertile site in Panama. Forest
Ecology and Management 238: 335-346.

variability of photosynthetic characteristics influences growth of eight tropical tree species at
two sites with contrasting precipitation in Panama. Forest Ecology and Management 261:
1643-1653.

Dainese M, Bragazza L. 2012. Plant traits across different habitats of the Italian Alps: a


Pakeman, R. J., Garnier, E., Lavorel, S., Ansquer, P., Castro, H., Cruz, P., Dolezal, J., Eriksson, 
O., Golodets, C., Kigel, J., Kleyer, M., Leps, J., Meier, T., Papadimitriou, M., Papanastasis, 
V. P., Quested, H., Quetier, F., Rusch, G., Sternberg, M., Theau, J.-P., Thébault, A & Vile, 
D. 2008. Impact of abundance weighting on the response of seed traits to climate and land 

Pakeman, R. J., J. Leps, M. Kleyer, S. Lavorel, E. Garnier, and V. Consortium. 2009. Relative 
climatic, edaphic and management controls of plant functional trait signatures. Journal of 
Vegetation Science 20: 148-159.

Schwarz, M., Santos, A.J.B., Aguilar, A., Czimczik, C.I., Gallo, J., Horna, V., Hoyos, E.J., 
Jimenez, E.M., Palomino, W., Peacock, J., Peña-Cruz, A., Sarmiento, C., Sota, A., Turriago, 
J.D., Villanueva, B., Vitzthum, P., Alvarez, E., Arroyo, L., Baraloto, C., Bonal, D., Chave, J., 
Costa, A.C.L., Herrera, R., Higuchi, N., Killeen, T., Leal, E., Luizão, F., Meir, P., 
Monteagudo, A., Neil, D., Núñez-Vargas, P., Peñuela, M.C., Pitman, N., Priante Filho, N., 
Prieto, A., Panfil, S.N., Rudas, A., Salomão, R., Silva, N., Silveira, M., Soares deAlmeida, 
Branch xylem density variations across the Amazon Basin. Biogeosciences 6: 545-568.


Paula, S., M. Arianoutsou, D. Kazanis, Ç. Tavsanoglu, F. Lloret, C. Buhk, F. Ojeda, B. Luna, J. 
M. Moreno, A. Rodrigo, J. M. Espelta, S. Palacio, B. Fernández-Santos, P. M. Fernandes, 

Ecology 90: 1420.


Sack, L., P. D. Cowan, N. Jaikumar, and N. M. Holbrook. 2003. The 'hydrology' of leaves: co-
ordination of structure and function in temperate woody species. Plant Cell and Environment
26: 1343-1356.

intracanopy leaf plasticity in temperate deciduous trees? American Journal of Botany 93:
829-839.


in a Quercus ilex Mediterranean forest. Forest Science 54: 513-522.

Sardans, J., J. Penuelas, P. Prieto, and M. Estiarte. 2008. Changes in Ca, Fe, Mg, Mo, Na, and S
content in a Mediterranean shrubland under warming and drought. Journal of Geophysical
Research 113:G3.

functional significance of forest diversity: A new long-term experiment with temperate tree


WSL.


Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C., Midgley, J.J., Navas, M.L.,
Niinemets, U., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V.I., Roumet,
Wright, S.J., Kitajima, K., Kraft, N.J., Reich, P.B., Wright, I.J., Bunker, D.E., Condit, R.,
Dalling, J.W., Davies, S.J., Diaz, S., Engelbrecht, B.M., Harms, K.E., Hubbell, S.P., Marks,
C.O., Ruiz-Jaen, M.C., Salvador, C.M. & Zanne, A.E. 2010. Functional traits and the
conductance of blue oak (Quercus douglasii) under prolonged summer drought and high
Phytophagy on phylogenetically isolated trees: why hosts should escape their relatives.
G. Swenson, M. C. Wiemann, and J. Chave. 2009 Global wood density database. Dryad
identifier: http://hdl.handle.net/10255/dryad.10235.
27 May 2018

Dear Dr. Penuelas,

Thank you for your effort handling our manuscript. We appreciate the thoughtful and constructive reviews that we received. We have now prepared a revised submission that addresses all of the reviewer points. In particular, we have:

- Removed the ECBilt-CLIO model from the paper, per advice from Reviewer #1;
- Redrawn most figures for enhanced clarity, with larger font sizes and legends;
- Better explained the biases inherent to the data, and justified our choices to minimize them, throughout the text;
- Extended discussion of megafauna and non-climate factors in driving our results.

A detailed response to the review comments follows below, with our responses in **boldface** type. We hope that our changes will be sufficient to render the manuscript acceptable for publication. Thank you again for your consideration.

Sincerely,
Benjamin Blonder, on behalf of my co-authors
Reviewer: 1

Comments to the Author

This study uses what appear to be the best available plant assemblage and climate reconstruction data to test whether palaeoclimate is 1) important in explaining contemporary plant community functional means and diversity, and 2) infers from those influences whether processes of plant functional community turnover are fast or slow.

Overall the study represents a thorough, comprehensive, and well-reasoned undertaking of an important question, which stands to influence a broad variety of global change topics. There are very few points on which I think the study could improve, although some recommendations are made below, including around breadth of explanations and reference to the literature, as well as points around ease of understanding and presentation of figures. A few issues around description of results require definite correction before publication can be considered.

We thank the reviewer for their interest in our work.

Specific Comments on Manuscript

Introduction –

Paragraphs 1 – 3: Suitably general in their statements, but literature cited is very plant-dominated, with reference to few other taxa. Discussion that this study is specific to plant FM & FD comes is not yet introduced. Some additional examples from other taxa would help reflect the generality of these statements – in particular I raise a point w.r.t. the discussion (see below) regarding late-Quaternary loss of megafauna, which may a suitable topic to cite here. Alternatively, the authors may wish to make it more explicit that this study and the cited works are principally in relation to plant communities earlier on the in this section.

We regret this lack of clarity. We have rewritten the first few paragraphs of the introduction to clarify that our results – and cited literature – are meant to primarily apply to plant assemblages.

Lines 104 – 129: I would like to commend the authors on the quality and clarity of writing in this section.

We appreciate this feedback.

Fig. 1: This explanatory figure was greatly appreciated during reading, however some minor changes to improve interpretability would be:
- The green ‘tree symbols’ differ in their opacity (and shape). Differing their colour (yellow, blue, red?) without differences in their opacities would ease interpretation in my opinion.

We appreciate the suggestion, but prefer to keep the existing scheme. While we agree that high contrast could be provided by using differing colors, we believe that changing opacity will reproduce better in black & white printouts of the figure. The differing shapes also already provide contrast. We are open to changing this upon further editorial advice.

- The charts illustrating +ve or -ve effects on FM & FD would be more easily understood if the graphs for FM and FD were entirely separated with white space between plot panels. Additionally, the ‘+’ and ‘-’ symbols would be better placed outside the plot-axis areas. I would also like to see the font size increased, and the use of annotation lines to allow for horizontal text would improve readability.

We have increased the white space between the right panels and the left panels. We have used some of this space to move the +/- symbols to a clearer position, and also added a 0 symbol. We have also increased the font size, but are not sure what is meant by annotation lines.

- The legend being above the figure (true also elsewhere) is atypical, but I expect this would be changed following publisher’s formatting.
We believe that this placement of the legend minimizes unused space in the figure, as some of this vertical header is needed to label the rightmost panels of the figure.

Lines 144-161: Again, the clarity of the explanations here should be commended.

Thank you!

Line 170 (& elsewhere): I would recommend the authors try and more clearly distinguish their uses of ‘Ka’ depending on whether it is being used to mean ‘thousands of years ago’ or simply ‘thousand years’ – i.e. a time before present or just a period of time. Whilst I appreciate there is no uniform consensus on this matter, in this study ‘Ka’ seems to be used to mean both, and then elsewhere in the manuscript ‘Kyr’ is also used (see lines 298 & surrounding). The reader’s understanding would be better served by use of ‘kya’ for ‘thousand years ago’ and ‘kyr’ for ‘thousand years’, or similar distinguishing units meaning ‘before present’ and as a period unit of time. Additionally, I would highlight that the ‘K’(kilo) shouldn’t be capitalised.

We have standardized on lower case capitalization in all figures and the text. We have also carefully checked all uses of ka and kyr and found only one instance where the term was not used correctly. The text is now fixed.

Methods –

Line 208: ‘arbitrarily’ raises concerns here as a phrase. Can some context be given for this value of 100? How it relates to mean or median species richness across all the cells (and a measure of spread around such a figure), would be informative. I trust this choice of < 100 not to be a problem, but some comparative context would help in the explanation of this cut-off.

The reviewer is correct that the cut-off was chosen to be small, and to remove unwanted cells with poor data coverage or current permanent ice cover. It represents the 7% quantile of richness for the ‘all’ species case and the 29% for the ‘woody’ species case. We simply do not feel comfortable estimating trait distributions with a very low number of species with trait coverage and prefer to lose spatial coverage rather than proceed with biased estimates.

The empirical cumulative distribution function of richness for the ‘all’ species case is shown below. The vertical red line indicates the chosen cut-off. The cells that are removed by the analysis are shown in green in the rightmost map. (Top row, all species; bottom row, woody species only).
We have clarified the text to reflect the rationale for this choice.

We do want to note that the woody maps and results are slightly different in our revised submission compared to our original submission. In the original, we gap-filled traits for all species using data for only the woody subset, then applied these to all species for the functional composition estimates. We have decided it is better (and more consistent with how we wrote the methods originally) to gap-fill traits for all species, then apply to only species that are classified as woody. The net effect is that the richness per cell in the woody species case decreases (as we are now only including woody species, rather than all species estimated with woody species’ traits) and some of the functional composition metrics change slightly. There is however no qualitative change in analysis outcomes, as can be verified by comparing the main text figures in the revised and original submission.

General: I was pleased to see how the authors carried out their testing, in particular the use of residual FD in addition to their initial FD and FM measurements. The justification for their approaches was also eloquently explained.

Thank you!

Fig. 2: Colour bar annotations are difficult to make out – I recommend widening the colour bars, removing or reducing the thick black outline/box, and increasing adjacent font sizes (maybe by reducing decimal place precision).

We have widened the colorbars and reduced the precision of the labels in the legends.

Climate Model: The mismatched MAP values between the HadCM3 and ECBilt-Clio are a little concerning. Do the authors feel comfortable commenting on / assessing which model is likely to be more accurate? If the HadCM3 climate reconstruction is likely superior (lines 525-526 suggest so), I would encourage the authors to cut all analysis and mention of the ECBilt-Clio model, as it is so temporally restricted in comparison.

We believe the HadCM3 model is strictly superior, but originally included the ECBilt-CLIO model based on prior reviewer suggestions. We have now removed it entirely from the analysis.

Line 298 (& 303): Immediate clarity with variation in window sizes depending on reconstruction period, as soon as the window concept is introduced, would be appreciated. The earlier description of the climate models make the mention of this window size on L298 confusing, only for the clarification to come five lines later on L303.

We clarified this paragraph by re-arranging, indicating the normalization approach and the HadCM3 resolution issue earlier on.

312-334 – Notably well written & justified methodology.
Thank you.

352-355 – Citations for all packages would be appropriate (from a cursory check, at least one of the used packages has an easily accessed associated citation).

**We have added citations.**

Results –

Fig. 3: Same comments as Fig. 2 and other maps – larger colour scale bars, reduce outlining box line weight, increase font size.

**We have improved the font size in the legends.**

Fig. 4: More care needs to be taken with this figure and its explanation. “Blue lines indicate models using only contemporary <…> orange lines, models using contemporary & palaeo” this account in the legend is directly opposite to the actual illustrated legend / annotation on the graph. The correct colours can be inferred from the results, but this is a problematic error. Additionally, it is not clear to me why the orange line is so limited compared to the blue? The blue line spans the whole length of the X-axis range, whilst the orange stops after only 1 x-axis step. An explanation of this, or correction of the plotting error, is necessary.

**We regret the confusion. The color scheme was reversed immediately before submission and we neglected to update the caption. The reviewer is correct that orange reflects contemporary climate variables.**

The differences in x-axis range for blue vs. orange is intentional and correct. We have only two contemporary climate axes (temp and precip) while we have many more paleoclimate axes (temp & precip at 0 ka, 1 ka, 2 ka, etc.). Mathematically, the number of PLS components in the model can be up the number of predictor variables (i.e. in the same way that a principal component analysis has as many principal components as input variables). Thus the paleo + contemp model can potentially have many more PLS components than the contemp model. We clarified this by including a sentence on number of variables in the ‘We tested Hypothesis H0’ paragraph.

Additionally, the link to figure panel 3H must be clarified. I think I understand that this is testing prediction of FD, and therefore model evaluation is in the same units as residual-FD (which is shown in 3H). However a cursory reading could lead to thinking that what is being tested is FDres, mapped in fig 3H, rather than FD, mapped in figure 3G.

**We regret this imprecision, which was also caused by a panel reorganization before submission. We now write 'units of functional diversity (compare to Figure 3G)'.**

Discussion –

Overall the discussion is well written and there are next-to-no changes to suggest.

My main comment is around lines 533-542. I think the authors understate the importance of this work in its links to other climate-interacting processes determining plant assemblages. Some reference is made to megafaunal extinctions (Gill et al. 2009), but the phrasing suggests that these processes are separate from the effects of the palaeoclimate. I think that megafauna-mediated effects provide an interesting indirect mechanism for palaeoclimate to influence plant assemblages, and therefore FM & FD, and should be (at least briefly) discussed here.

For example, Pires et al. (2018), Gill (2014), Gill et al. (2012), and Doughty et al. (2009) all demonstrate that the loss of megafauna will influence plant assemblages into the long-term, with examples specifically from the Western Hemisphere and this study’s temporal period. Whilst human activity is undeniably a cause of megafaunal loss, Bartlett et al. (2016) showed that climate also had an important role in the loss of Pleistocene
megafauna – including in the Western Hemisphere. In particular I think it’s noteworthy that their study also used what appears to be the same (or a very similar version) of the HadCM3 reconstruction used by the authors of this study.

I think therefore that more can be made of megafaunal extinction or population decline as an additional mechanism behind the effect of palaeoclimate on contemporary functional assemblage. It may even provide insights into when immigration or exclusion are fast or slow processes.

I think the authors underestimate the relevance of their study in terms of its integration with this topic, and think with additional citations (a few of which I have provided here) and at least brief discussion, their findings would be even more impactful than they are currently presented to be.

We thank the reviewer for this point, and agree with it. We have added a new paragraph to discuss indirect effects of climate on species composition in much more depth.


Reviewers: 2

Comments to the Author
Dear authors,
see my comments to the editor and my detailed comments below:

In their manuscript “Late Quaternary climate legacies in contemporary plant functional composition” Blonder et al. show for plants across the Americas that contemporary functional trait composition of assemblages is linked to paleoclimate indicating that processes reorganizing assemblages have been slower than the changing climate. They test four hypotheses on how past climate, past climate variation and contemporary climate affect contemporary functional trait means and functional diversity. They find that climate closely after the Last Glacial Maximum has left a strong imprint on contemporary functional composition. Results for what mechanisms (slow or fast immigration and/or exclusion) are responsible for the observed patterns were different among the two continents showing the complexity of the processes behind the observed patterns and calling for future analyses based on time series of functional composition.

The manuscript is very well written, the methods are state of the art and well-described, and the results are presented in a nice way and discussed appropriately given the available body of literature. The topic is certainly of interest to a broad community of Macroecologists and beyond and the results are novel (given the functional perspective) and exciting. I therefore recommend considering the paper for publication in Global Change Biology. However, I have some concerns about how potential effects of quality issues with the trait and distribution data (which are discussed in the text) on the results are assessed and I suggest to perform rigorous sensitivity analyses. I therefore recommend a “major revision”. Unfortunately the online review system only allows to chose “minor revision without reassessment” and “reject and invite to resubmit”. Since I don’t want to participate in this game of artificially reducing the time from submission to publication, I chose “minor revision”. However, I urge the authors to take my comments seriously.

We thank the reviewer for their interest in our work.

Comments to the authors:
Major points
My only main doubt about this paper concerns if potential biases in the trait and distribution data might have affected the results. The maps of trait means and FD look extremely smooth (Fig. 3), which I would guess is due to the quite low availability of actual trait (a lot imputed) and distribution data (convex hulls sometimes around few occurrences)? Also, some patterns look a bit artificial (almost straight horizontal and vertical lines) in some of the plots (e.g. Fig. 3 G. Functional Diversity and H. FD residual). Can you explain this? I would therefore like to see sensitivity analyses on how gap filling for traits, genus means for species without trait records and species with few distribution records influence the results. It would also be helpful to show spatial coverage patterns for the trait and distribution data.

The reviewer is correct in noticing that the distribution data are coarse, and sometimes include what are likely artifacts of the modeling method. Part of this effect comes from the greater quantity of occurrence data available in the United States relative to Canada (the upper horizontal line) and the lower quantity of data available in southern South America (the southern angled line). Some species also may have coastal distributions, which can lead to artifacts in range maps when a convex hull method is used. We acknowledge these issues, but feel that they are inevitable ‘costs’ of a SDM approach that does not calibrate predictions based on contemporary climate data. For example, a MaxEnt type model would produce species distribution (& thus functional diversity) maps without many of these artifacts. The downside is some circularity, as a map of FD produced by calibration on contemporary climate data will surely show that contemporary climate is a good predictor of FD.

In a previous iteration of this manuscript at a different journal, we included MaxEnt type models for the FD calculations. We show a few example figures from this analysis below. While some of these ‘sharp edge’ artifacts are now gone, the overall spatial patterns are often quite similar.

It is also important to remember that the overall analysis is constrained by the spatial resolution of the paleoclimate data, which is at a nominal 2° resolution, but in practice has lower resolution due to the spatial structure of the model. As such, many features in the FD and richness maps are effectively ‘blurred out’ by the PLS analysis. For this reason, we decided against trying to take more complex modeling approaches.
Regarding the data coverage issues (both for species with too few occurrence points, and for species with no trait data), the reviewer has requested sensitivity analyses, or alternate analyses with removing these data points. We have thought long and hard about this issue when preparing this project, and ultimately believe that carrying them out would not be helpful. In the absence of complete data, any methods taken to fill in data will potentially introduce biases – and something as simple as removing data may in fact have a much larger bias than the gap-filling and space-filling methods we have elected to use. In this dataset, 24042 of the 74,000 species are represented by 10 or fewer observations, and most of these species are tropical – thus, losing these species would represent a very large bias in the dataset. In another in-progress study led by co-author Enquist, it is shown that most of these database-rare species are actually rare and have restricted ranges according to interviews with expert botanists.

Above, histogram of occurrence records per species; below, heatmap of occurrences over space (note the log-scale z axis). Additionally, trait data for 59,423 species out of 74,000 was not available. The trait coverage is higher in tropical regions.
Omitting cells with low trait coverage would also introduce a large spatial bias. Our resampling approach is, we believe, the best way to deal with the impacts of low trait coverage on functional diversity.

We therefore feel that the approach we have taken is the least bad of all the bad options, given the currently available data resources for plant macroecology. Removing rare species, and removing species without trait data would lead to highly sparse and biased estimates. Our trait maps and richness maps are approximately congruent with other recent mapping efforts (e.g. Jetz’s late-2000s work, or Butler et al. in PNAS this year, which the lead author is a co-author on), but are more appropriate for this application because of the lack of circularity.

We hope that this extended response is convincing to the reviewer. We have clarified the methods text to reflect the bulk of these points. We have included the trait coverage and occurrence point coverage figures in the resubmission.

I could imagine that due to the smooth spatial patterns in the response variables (Trait means and FD metrics) the spatially smooth paleo climate data performs better than the high resolution and spatially more heterogeneous contemporary climate data. Maybe the different resolutions and methods how contemporary climate and paleoclimate are derived are partly responsible for the findings (btw. Chelsa now also offers LGM climate at high resolution). It is striking that contemporary and Paleoclimate were not strongly correlated (lines 289-290). Maybe the coarser resolution of the Paleoclimate layer fits better to the coarse spatial scale of the distribution data?

We agree that methodological differences in how the paleoclimate and contemporary climate datasets are created could drive some of the findings. We appreciate the suggestion for the higher resolution CHELSA dataset, but for this application we would need that high resolution at every time point back to 120 ka, which is to our understanding not yet (or perhaps not ever) feasible.

While the reviewer argues that contemporary and paleoclimate axes are not strongly correlated, we believe that our supplementary figure shows otherwise, also the definition of ‘strongly’ is certainly debatable. Rather than report the Pearson correlation between these maps, which could be high even if the actual values are down or up biased, we report the mean absolute deviation between them. These values are less than 0.5 standard deviations throughout the Holocene, which we think reflects strong correlation, and then diverge during the late Pleistocene, as expected. We have clarified the main text to better explain when and when not there is evidence for matching between these variables.

We also note that based on feedback from the first reviewer, we now no longer include the ECBilt-CLIO analyses in the paper.

Minor points:
Line 63: delete one of the two “to”s

Fixed.
Line 79: rather “little” than “less”?

**Changed.**

Line 81: “paleoclimate has structured contemporary” sounds too obvious to me. The question rather is, how much of it is still visible, right?

We prefer to keep this phrase as-is – a great deal of species distribution modeling assumes that there is no influence of paleoclimate on species ranges – so while we agree that it is obvious paleoclimate should matter, we do not think everyone would agree with this statement.

Line 82: Also secondary effects due to climate change like changes in sea levels (See literature examples for islands)

We agree this is reasonable, but our downstream analyses do not have the ability to account for variation in sea level. We now write ‘climate has directly & indirectly’ as a compromise.

Lines 83-85: Is this something tackled here?

We believe it is – all of our conceptual models focus on fast vs slow (i.e. lagged) dynamics of organisms. To clarify we now write, “It has been unclear how these paleoclimate effects on species composition translate to differences in functional composition, because even species assemblages in disequilibrium with contemporary climate may have equilibrium functional relationships with contemporary climate”.

Lines 101: Do you want to state that the influence is still visible?

To clarify, we write ‘contemporary functional composition’ instead of ‘functional composition’.

Line 145 “and also”?

We clarified conditional phrasing throughout this paragraph.

Line 170: “[120 Ka [thousands of years ago]]” Not clear to me

Ka is meant to define ‘thousand years ago’. We removed this definition as we think it is well-understood by most readers.

Line 185: What does “collinear observations” mean?

**Collinear is a standard mathematical term referring to points that fall on the same line ([https://en.wikipedia.org/wiki/Collinearity](https://en.wikipedia.org/wiki/Collinearity)).** We clarify by writing ‘observations’ now as ‘observation points’.

Line 235-238: Species richness

We do not understand what change we should make to the text here, as species richness is already used throughout the sentence. We did find one ‘richness’ and changed it to ‘species richness’ on the following line.

Line 307 “type type” -> “type”

**Fixed.**

Line 454: “appears”

**Fixed.**
Line 455: Why “additionally”? Aren’t the previous sentences saying the same?

We removed the ‘additionally’.

Line 465: “, ,”

Fixed.

Figure S3 I think you could remove white space and make the maps larger if you only show one legend for all of them and change the location of the titles.

We prefer to keep the legend on each panel to enable easy comparison of colors. However we have moved the titles to reduce white space.
Reviewer: 3

Comments to the Author
It was a real pleasure to read this manuscript. The scope of the analysis is very impressive, the analytical methods appear well chosen and meticulously applied. Moreover the detection of lag effects on contemporary trait composition is of great interest because it means that the pool exposed to future global change and that deliver trait-associated ecosystem functions may exhibit on-going slow dynamics as it continues to equilibrate to current climate space.

We than the reviewer for their interest.

My only question centres on their assumption that the traits selected are strong carriers of the climate signal. For example Ordoñez et al (2009) report weak relationships between MAT, MAP and SLA, leaf N (mass and area basis in their Fig 2) but inferred clear interactions and main effects of soil nutrients. Also Wright et al (2005) reported weak correlations but did conclude that they were strong enough to show a biogeographic influence of climate. Thus if the traits selected are also strongly related to other abiotic conditions then it seems possible that the strong correlation between temporal paleo-climate variation and trait variation could be slightly artefactual. I admit that a mechanism for this is not obvious. For example it would require that the 100x100km cells that had high temporal climate variability also had high contemporary abiotic variation giving rise to higher functional diversity. Even if this were not the case then separate test of individual traits ought to show that traits more strongly related to climate showed a stronger paleo-climate signal and vice versa. I am not suggesting the authors revise their analysis but a greater level of comment would be useful. The authors openly allude to some of these difficulties at bottom of page 29 but I think a deeper consideration is needed.

We appreciate this point, which was also raised (in the context of megafauna drivers of traits) by another reviewer. In response we have extensively revised and extended the page the reviewer mentions. We now highlight the importance of soil as a possible driver of traits, and also the indirect role of megafauna. However we note that in both cases, we do not yet have the time-series data available to determine the role of these variables relative to climate. We are also optimistic that such analyses will become possible in the near future.

Refs: