




# Predicting intraspecific trait variation among California's grasses

Brody Sandel<sup>1</sup>  | Claire Pavelka<sup>1</sup> | Thomas Hayashi<sup>1</sup> | Lachlan Charles<sup>2</sup> | Jennifer Funk<sup>3</sup> | Fletcher W. Halliday<sup>4</sup>  | Gaurav S. Kandlikar<sup>5</sup>  | Andrew R. Kleinhesselink<sup>5</sup> | Nathan J. B. Kraft<sup>5</sup> | Lorelee Larios<sup>2</sup> | Tesa Madsen-McQueen<sup>6</sup> | Marko J. Spasojevic<sup>6</sup>

<sup>1</sup>Department of Biology, Santa Clara University, Santa Clara, CA, USA; <sup>2</sup>Department of Botany and Plant Sciences, University of California: Riverside, Riverside, CA, USA; <sup>3</sup>Department of Plant Sciences, University of California, Davis, CA, USA; <sup>4</sup>Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, CH, Switzerland; <sup>5</sup>Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA, USA and <sup>6</sup>Department of Evolution, Ecology, and Organismal Biology, University of California: Riverside, Riverside, CA, USA

## Correspondence

Brody Sandel

Email: bsandel@scu.edu

## Funding information

NSF, Grant/Award Number: #1309014 and DEB #1644641

Handling Editor: Meghan Avolio

## Abstract

1. Plant species can show considerable morphological and functional variation along environmental gradients. This intraspecific trait variation (ITV) can have important consequences for community assembly, biotic interactions, ecosystem functions and responses to global change. However, directly measuring ITV across many species and wide geographic areas is often infeasible. Thus, a method to predict spatial variation in a species' functional traits could be valuable.
2. We measured specific leaf area (SLA), height and leaf area (LA) of grasses across California, covering 59 species at 230 sampling locations. We asked how these traits change along climate gradients within each species and used machine learning to predict local trait values for any species at any location based on phylogenetic position, local climate and that species' mean traits. We then examined how much these local predictions alter patterns of assemblage-level trait variation across the state.
3. Most species exhibited higher SLA and grew taller at higher temperatures and produced larger leaves in drier conditions. The random forests predicted spatial variation in functional traits very accurately, with correlations up to 0.97. Because trait records were spatially biased towards warmer areas, and these areas tend to have higher SLA individuals within each species, species means of SLA were upwardly biased. As a result, using species means over-estimates SLA in the cooler regions of the state. Our results also suggest that height may be substantially under-predicted in the warmest areas.
4. *Synthesis.* Using only species mean traits to characterize the functional composition of communities risks introducing substantial error into trait-based estimates of ecosystem properties including decomposition rates or NPP. The high performance of random forests in predicting local trait values provides a way forward for estimating high-resolution patterns of ITV without a massive data collection effort.

## KEYWORDS

functional traits, grass, intraspecific trait variation, machine learning, Poaceae, specific leaf area

## 1 | INTRODUCTION

All organisms face the challenge of allocating limited resources to competing demands (Enquist & Niklas, 2002; Smith & Fretwell, 1974; Westoby et al., 2002) including growth, defense and reproduction (Heckman et al., 2019; Tuller et al., 2018). How an individual addresses these trade-offs influences its fitness in a particular environment as well as its influence on its environment (Suding et al., 2008). These allocation strategies are expressed as functional traits, which describe a phenotypic feature of an individual that influences its fitness in a particular environment (Violle et al., 2007). Species with traits that are poorly matched to their abiotic and biotic environments may be excluded from a community, a process often conceptualized as an environmental or biotic filter (e.g. Diaz et al., 1998; Kraft et al., 2015; Laughlin et al., 2012; Stahl et al., 2014). At the same time, the functional traits expressed by the species in an ecosystem can strongly influence ecosystem properties, such as productivity (Garnier et al., 2004; Pontes et al., 2007) and decomposition rates (Cornwell et al., 2008; Makkonen et al., 2012).

Most studies of functional traits in plants characterize each species with a single value for each trait, typically the arithmetic or geometric mean and often irrespective of the conditions where the traits were measured. However, individuals within a species can show marked intraspecific trait variation (ITV) reflecting both local adaptation and acclimation (Albert et al., 2011, 2012; Violle et al., 2012). Globally, ITV accounts for about 25% of trait variation within communities and 32% between communities (Siefert et al., 2015). ITV can have important implications for environmental filtering (Laughlin et al., 2012; Siefert, 2012), species coexistence (Hart et al., 2016; Lichstein et al., 2007), ecosystem properties (Crutsinger et al., 2006) and restoration (Baughman et al., 2019). Furthermore, if trait values within a species vary along environmental gradients, applying one mean trait value to that species at all locations risks misestimating that species' fitness in a particular environment as well as its influence on its environment across most of its distribution.

This potential misestimation between species' mean traits and the traits they express in a particular context is particularly consequential in the coupling of dynamic global vegetation models (DGVMs) to climate models (Bloomfield et al., 2018; Scheiter et al., 2013). These models use functional traits of plants including specific leaf area (SLA) and height to estimate their ecosystem impacts (Bonan et al., 2003; Sato et al., 2007; Yang et al., 2015). Uncertainty in these trait measurements can have large consequences for model outputs, such as the carbon sink of the terrestrial biosphere (Verheijen et al., 2015). There has been substantial progress on estimating geographic patterns of functional traits, but the role of ITV in vegetation models has largely been ignored (Butler et al., 2017; Berzaghi et al., 2020, but see Sakschewski et al., 2015). Therefore, there is a

need to understand ITV across climate gradients, but it is often infeasible to collect these kinds of measurements for large numbers of species and traits over the large geographic areas required.

In addition to being of practical importance, understanding the drivers of ITV also promises to advance our fundamental understanding of plant adaptation and strategies. Theory based on plant physiology has been developed to predict how particular environmental gradients should impose filters on plant community composition. For example, in conditions supporting a higher leaf area index (generally warmer and wetter climates), game theory predicts higher relative allocation to stem biomass and therefore greater plant heights and later reproduction (Falster & Westoby, 2003). Similarly, an economic model predicts that plants growing in drier environments should produce high concentrations of rubisco to maximize CO<sub>2</sub> uptake while limiting stomatal water loss (Wright et al., 2003). Thus, we would expect plants growing in drier environments to have higher leaf nitrogen concentrations, a pattern which has been detected locally (e.g. Cornwell & Ackerly, 2009) and globally (e.g. Bruelheide et al., 2018; Sandel et al., 2016; Wright et al., 2005).

While such theory has typically been tested at the species level, most theories predict congruent within-species responses. For example, if dry environments filter out species with low leaf N concentrations, it is likely that they also tend to filter out individuals within species that allocate relatively lower N concentrations to their leaves. Thus, quantifying trait variation within species across environmental gradients can provide a new test for these predictions. In most cases, it appears that the directions of ITV match those of interspecific functional turnover (Ackerly & Cornwell, 2007; Carlucci et al., 2015; Derroire et al., 2018; Kumordzi et al., 2015; Lepš et al., 2011; Sandel & Low, 2019), although exceptions exist (Kichenin et al., 2013; Lepš et al., 2011).

Our goal here is to characterize ITV among multiple species along climate gradients and to assess our ability to predict spatial variation in ITV. It is typically infeasible to collect detailed data on spatial patterns of ITV across many species, so a predictive model trained on whatever data are available is an attractive alternative. However, ITV predictions may transfer poorly among species (Lajoie & Vellend, 2015). For example, some species increase and others decrease their leaf N concentration along an elevational gradient in New Zealand (Kichenin et al., 2013). Similarly, environment-ITV relationships of grass species in a common garden experiment varied substantially among species (Roybal & Butterfield, 2019). The apparent idiosyncrasy of environment-ITV relationships across species presents a substantial challenge to any predictive model. One possible solution is to use species-level characteristics such as mean traits or phylogenetic relationships in a machine learning context to determine whether the variation among species can be attributed to characteristics of each species.

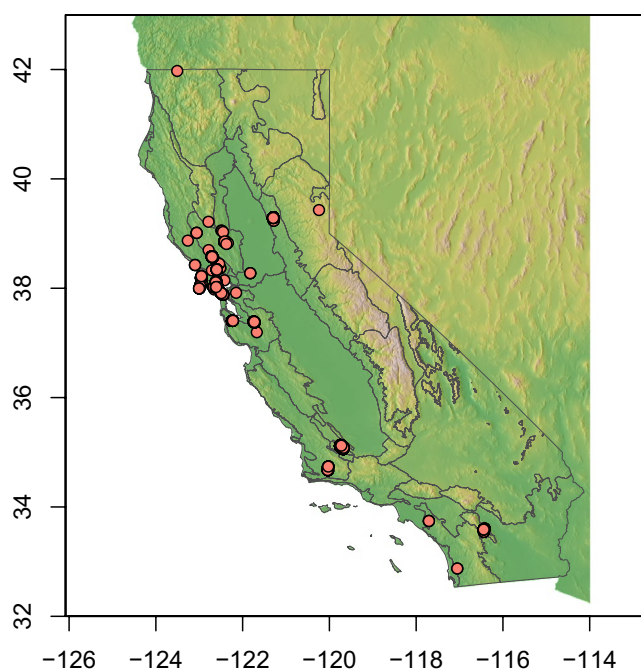
That is, we might expect two species with similar evolutionary histories and with similar mean functional traits to respond similarly to environmental gradients, while two distantly related species with very different functional strategies might not. We explore this possibility here.

As our focal group, we take California's grasses, a diverse and often-dominant group with high ecological and economic importance. There are roughly 450 species of wild grasses in the state, of which about half are naturalized non-native species. Grasslands cover 11% of the state (Davis et al., 1998) and span a wide range of climate conditions from hot, dry conditions in the Mojave Desert to Mediterranean conditions, to alpine meadows in the Sierra Nevada mountains. Specifically, we ask (a) how the traits of individuals within species vary along climate gradients, (b) how successfully ITV patterns can be predicted and (c) how including ITV influences patterns of functional assemblage composition along gradients.

## 2 | MATERIALS AND METHODS

### 2.1 | Trait data

We measured traits of grasses across California (Figure 1). These measurements were made for various other projects, and therefore varied somewhat in their temporal and spatial extents and sampling methods. At each site, functional trait data were collected from 3 to 10 individuals of each species. We focus on three traits: specific leaf area (SLA), plant height and leaf area (LA). Height was measured as



**FIGURE 1** Map of sampling locations. The background map is a hill-shaded topographic map, showing the division into ecoregions as defined by the Jepson Flora Project (see text for details)

the length from the base of the plant to the highest, fully expanded leaf blade. If leaf blades were collected in the study, the blades were fully expanded, non-senescent and undamaged. Most of the studies excluded the ligule (Table S1). After collection, the leaf blades were oven-dried from 55 to 65°C for at least 48 hours before being weighed. For most samples, leaf area was obtained by photographing the leaf against a white background and using image analysis software such as ImageJ to determine the area. Some studies instead directly scanned leaf blades to determine area (Table S1).

In total, we obtained trait information for 1,854 unique site-by-species combinations, across 59 species and 230 unique sampling locations. Thirty-one species were recorded from at least 10 different sampling locations. SLA was available for nearly all of the site-by-species combinations (1,805), while LA was available for 1,371 and Height for 1,231. Summaries of sample sizes and trait means and standard deviations for each species are shown in Table S2.

### 2.2 | Climate data

We obtained 30-year normal 800m resolution rasters of annual precipitation (AP) and mean annual temperature (MAT) from PRISM (PRISM Climate Group, 2018). We then extracted the values of AP and MAT at each sampling location. We also considered seasonal averages of temperature and precipitation, but these were generally well correlated with their annual means ( $r > 0.73$ , except winter temperature [ $r = 0.56$ ] and summer precipitation [ $r = 0.40$ ]).

### 2.3 | Other trait data

To supplement our database of local grass trait measurements, we also used a global grass trait database compiled from literature review (see Sandel et al., 2016, and references in Appendix S1) and several trait databases including TRY (Kattge et al., 2011) and the Tundra Trait Team (Bjorkman et al., 2018). We aimed to obtain estimates of species mean SLA, Height and LA for as many California grass species as possible. These values would be used to derive estimates of spatial patterns of grass traits, even in the absence of detailed local measurements (see *Predicting ITV* below). Of 457 grass species in California, we were able to obtain mean SLA values for 264, Height values for 280 and LA for 202. Data coverage was slightly better for annual than perennial species (e.g. for SLA 67% coverage among annuals and 53% among perennials). In addition, some ( $n = 85$ ) of the georeferenced trait measurements in this database fall within California and were used as additional local trait records in this study (Table S1).

### 2.4 | Jepson distributional data and synonymy

For each grass species in California, we obtained distributional information from the Jepson Flora Project (Jepson Flora Project (eds.),

2019. Jepson eFlora, <http://ucjeps.berkeley.edu/eflora/> [accessed on 9/16/2019]). The description for each species includes its presence or absence in each of 35 ecoregions and the elevational range where it is known to occur. We divided each ecoregion of California into 100 m elevational bands using a 1km resolution DEM (from worldclim, Hijmans et al., 2005), producing 666 discrete elevational bands across the state. A species was considered to be present in a band if it is known from that region and occurs in at least some of the 100 m elevational range (e.g. a species with an elevational range from 0 to 1,250 m was considered to be present in a band spanning 1,200–1,300 m).

The Jepson Flora recognizes 461 grasses in California. We matched these names to the Kew Gardens GrassBase (Clayton et al., 2006 onwards) nomenclature, which in most cases is congruent with Jepson. However, some genera are treated differently. For example, GrassBase separates some members of *Stipa* (including the widespread *Stipa pulchra*) into several genera including *Nassella*, *Piptochaetium*, *Piptatherum* and *Ampelodesmos*, Jepson's members of the genus *Festuca* are divided into three genera in Kew (*Festuca*, *Lolium* and *Vulpia*), *Elymus* is divided into *Elymus*, *Leymus* and *Taeniatherum*, and some members of *Bouteloua* are assigned to the genus *Chondrosum*. In addition, some species divisions in Jepson are not recognized in GrassBase, such as *Panicum capillare* and *P. hillmanii*, which GrassBase treats together as *P. capillare*. Removing these duplicated names produced 459 distinct names. On the other hand, Kew recognizes *Gastridium ventricosum* and *Gastridium phleoides* as separate species, while *G. ventricosum* is treated by Jepson as a subspecies of *G. phleoides*, and Kew separates *Lolium perenne* and *Lolium multiflorum*, while Jepson treats them together as *Festuca perennis*. Because we therefore lack distributional information for *G. ventricosum* and *L. multiflorum*, we followed the Jepson delineation in those two cases and merged each species pair. Whenever species were merged, we took the union of all bands where each merged species was listed to occur. Thus, our final species list includes 457 grass species.

## 2.5 | Phylogeny

We sought to estimate the phylogenetic relationships among as many of these species as possible. Following the approach in Sandel and Tsirogiannis (2016), we began with a large tree of grasses by Edwards et al. (2010) containing 2,684 taxa. We matched the nomenclature to GrassBase, randomly added missing species within their genera, and removed all non-California species. In total, this produced a tree with 445 species (of 457 total species in California), of which 232 were originally found on the Edwards tree and 213 were grafted into their genera.

This resulting tree was neither dated nor ultrametric. To date interior nodes, we used a smaller phylogeny of grasses that had been dated with both macrofossil and microfossil evidence (Christin et al., 2014). We identified matching nodes between the two trees and applied the dates from the Christin et al. tree to our California

tree at any matched nodes. We then forced our tree to be ultrametric, constrained by these node dates, using penalized likelihood (Paradis, 2013; Sanderson, 2002), implemented in the *chronos()* function (package *APE* 3.0–11; Paradis et al., 2004). This resulting final tree has the same topology as the original California tree, but is dated and ultrametric.

Our goal is to use an individual's phylogenetic position as a predictor for its local trait values (see below for model details). To that end, we calculated a set of variables describing a species' phylogenetic position: phylogenetic eigenvector maps (PEM, Guénard et al., 2013), implemented in the R package *MPSEM* (Guénard, 2013). Each such PEM is a vector analogous to a principle component analysis (PCA) axis and summarizes one dimension of phylogenetic relationships among species. We retained the first five PEMs to use in the predictive modelling below. This approach does not assume phylogenetic conservation of traits or trait–environment relationships, but takes advantage of it if it exists.

## 2.6 | Analysis

### 2.6.1 | Trait–trait and trait–climate relationships

For each pair of traits, we computed the correlation between species mean values. We then computed this correlation within each species for which there were at least 10 pairwise complete local records.

For each species with sufficient data, we fit an ordinary least squares (OLS) linear model predicting variation in each of the three traits from annual precipitation (AP), mean annual temperature (MAT) and their interaction (AP:MAT). To be used, a species had to meet four criteria: trait records from at least 10 sampling locations, sampling locations spanning at least 2 degrees of MAT and 100 mm of AP, and a correlation between AP and MAT of no more than 0.8. These criteria were used to prevent fitting models with extreme slopes when all measurements were made over a narrow range of climate values, or when AP and MAT were highly collinear. Twenty-two species met all four criteria for at least one of the three focal traits. Note that across all samples, annual precipitation and annual temperature were hardly correlated ( $r = 0.03$ ).

All three trait values were log-transformed before analysis to improve linearity and normality of model residuals.

It is possible that some variation in sampling schemes among the datasets we collected would lead to different trait measures. To assess this possibility, we also fit a linear mixed effects model equivalent to the OLS regression described above, but also including the dataset as a random factor. We asked whether the coefficients from these models are similar to those from OLS regression models.

We then sought to understand variation in the AP–trait and MAT–trait relationships across species. We related among-species variation in coefficients to each species' taxonomic tribe, mean SLA, Height and LA, and whether the species was annual or perennial. To test the hypothesis that species with different

distributions in climate space might show different trait responses along environmental gradients, we also computed the mean MAT and mean AP across all Jepson bands where the species occurs within California. These predictor variables were related to the model coefficients relating each trait to AP, MAT and the AP:MAT interaction using linear models with a bidirectional stepwise model selection.

## 2.6.2 | Predicting ITV

Our goal was to build a predictive model to estimate the spatial pattern of ITV for any given grass species. For this purpose, we chose random forests, a machine learning approach that is well-suited to modelling complex interactions and nonlinear responses (Breiman, 2001; Cutler et al., 2007) and that is related to techniques that have been useful in relating plant traits to fitness (Pistón et al., 2019). We began by exploring model performance using different subsets of predictor variables, then built a simplified model to predict spatial variation in ITV for each grass species with a minimum of local information. We did not exclude any species from this model building.

We began by computing, for each site-by-species combination a 'delta-trait' value; this is the difference between the trait value of that species at that site and its mean overall trait value. Thus, each species has a mean delta-trait value of 0, and positive values indicate trait values higher at a particular location than the average trait value for the species. This was done to place the emphasis on training a model to explain ITV, rather than between-species trait differences.

We grouped predictor variables into five categories: climate (AP and MAT), species mean traits (SLA, Height and LA) and life span (annual or perennial), other local trait measurements (e.g. if predicting SLA, the LA and Height of each local population of that species), phylogenetic position (the first five PEMs) and species identity. A full model containing all five variable groupings would therefore be capable of modelling climate-ITV relationships that change as a function of species mean traits, or that vary among phylogenetic groupings.

We explored the possibility of replacing raw climate variables at each sampling location with delta-climate values, giving the difference between the MAT or AP at a location and the mean for a species across a range. This would model the possibility that species show different responses to climate depending on the mean climate where they occur. Random forests trained in this way displayed nearly identical predictive performance, so our further analysis focused only on models using raw climate values.

We trained models on different subsets of these variable groups and evaluated their performance. All combinations of these five groups were considered, producing 31 models ( $2^5 - 1$ , removing the empty model with no predictor variables). In each case, we randomly selected 80% of the data to train the model and reserved the remaining 20% to test predictive performance. We calculated four

measures of model performance. Two were the correlation between predicted and observed delta-trait values for the training dataset and the same for the testing dataset. Then, we converted predicted delta-trait value into predicted trait values by adding each delta-trait value to its species mean and computed the training and testing correlations between observed trait values and predicted trait values as the other two measures of model performance. We replicated this entire process 100 times and computed the average of each performance measure across these 100 replicates.

For each predictor variable group, we computed the average difference in each performance measure between each pair of models with and without that variable group. This serves as a measure of that variable group's overall importance.

Our ultimate goal was to build a model that can predict local trait measurements of a species from an unmeasured population. Two variable groups considered above are largely incompatible with that goal—other local trait measurements and species name. That is, it would not be particularly useful to have a model that could predict local SLA of a grass, but only if given the local Height and LA of that grass, or a model that could only work for the specific named species for which it had been trained. Thus, for our final predictive model, we built a random forest using only climate, species mean traits and phylogenetic position. Additionally, for the species-level traits we used only the mean of the focal trait and the species' life span so that the prediction for local trait values of one trait would not depend on the availability of another trait. These models have the potential to predict trait values for species in locations where its traits have not been measured if climate-ITV relationships are transferable among close phylogenetic relatives or among species with similar mean traits. As above, each model was trained on 80% of the data for each trait, and tested on the remaining 20%.

## 2.6.3 | ITV influences on patterns of grass functional composition

The above procedure produced three final trained random forests: one for each of the traits considered here. Given local AP and MAT measures, the phylogenetic position of a species, its life span and its mean SLA, LA or Height, these three models can then predict the local SLA, LA or Height expected for a species at a particular site.

Within each Jepson band, we computed the mean AP and MAT and extracted the PEMs and mean traits of all species occurring in that band. We then used the three trained random forests to obtain predictions for SLA, LA and Height of each species in each band. For each estimate, we also obtained a standard error using a jackknifing approach (Wager et al., 2014), implemented in the R package RANGER (Wright & Ziegler, 2017).

For each band, we computed measures of the mean and variation in traits of species present in that band. We used the geometric mean and geometric standard deviation, computed either using the local trait estimates of each grass species in that band, or using the species mean trait values.



### 3 | RESULTS

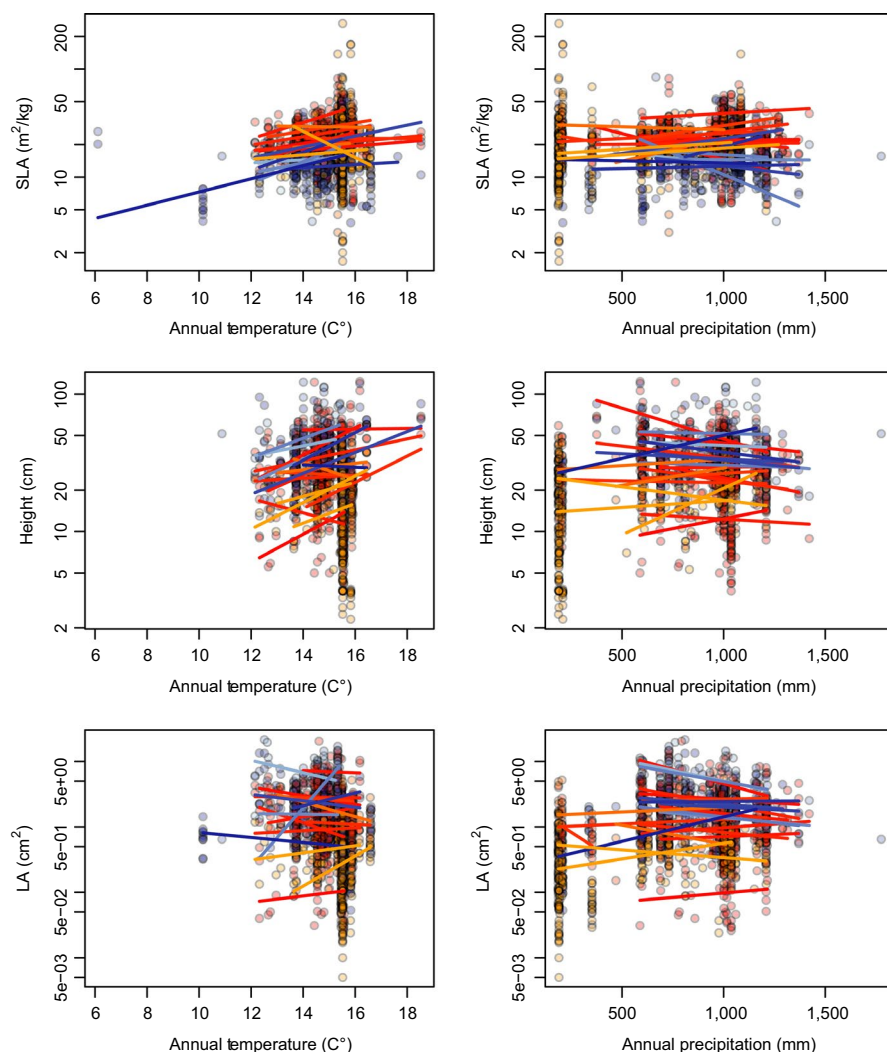
#### 3.1 | Trait-trait and trait-climate relationships

Among species, mean SLA and Height were negatively correlated ( $r = -0.29$ ,  $p = 0.03$ ) and LA and Height were positively correlated ( $r = 0.70$ ,  $p < 0.001$ ). SLA and LA were not significantly related ( $p = 0.46$ ). Within species, SLA and Height tended to be slightly positively correlated (in 20 of the 27 species with at least 10 records, and a median  $r = 0.13$ ). SLA and LA were generally positively correlated (in 25 of 29 species, median  $r = 0.27$ ) and LA and Height were positively correlated (in 25 of 27 species, median  $r = 0.57$ ).

There was substantial variation among species in the ITV responses to AP and MAT, although some responses were fairly consistent among species (Figure 2; Table 1). For example, a majority of species exhibited higher SLA values at higher temperatures (although *Avena fatua* and *Hordeum murinum* showed the opposite trend), grew taller at higher temperatures (with *Briza minor* being the main exception) and produced larger leaves in drier conditions (*Vulpia microstachys* was the exception). Interactions between AP and MAT were fairly common. For both Height and LA, the interactions tended to

be negative, indicating that the effects of AP on these traits tended to become more negative at higher values of MAT (and vice versa). There was wide variation in model  $R^2$ , ranging from 0.02 to 0.85, with mean  $R^2$  about 0.24 (Table 1). Annual species tended to have slightly lower  $R^2$  values than perennial species, by an average of 0.035.

In some cases, much of this variation among species could be explained by species mean traits and taxonomic tribes. Species with larger leaves had more negative AP-SLA relationships ( $R^2 = 0.22$ ). Perennials had more positive MAT-SLA relationships, and members of tribes Stipeae and Danthonieae tended to have more negative relationships than other groups ( $R^2 = 0.58$ ). The AP:MAT interaction effect on SLA tended to be more positive for small-statured grasses ( $R^2 = 0.23$ ). AP-Height relationships were most negative for large-leaved species in the Danthonieae ( $R^2 = 0.69$ ), and MAT-Height relationships were most positive for perennial species ( $R^2 = 0.25$ ). The AP:MAT interaction was negative for most species, but especially those with large leaves ( $R^2 = 0.36$ ). AP-LA relationships were especially negative for species with large mean leaf size, annuals and species from wet areas of the state ( $R^2 = 0.68$ ). None of the species-level descriptors were significantly related to MAT-LA coefficients, but the AP:MAT interaction effect was particularly negative for species



**FIGURE 2** Relationships between locally measured trait values for specific leaf area (SLA), Height and leaf area (LA) and mean annual temperature or annual precipitation. Each species is colour coded, with a fitted bivariate regression line. Blue symbols and lines represent perennial species, while reds and oranges indicate annuals

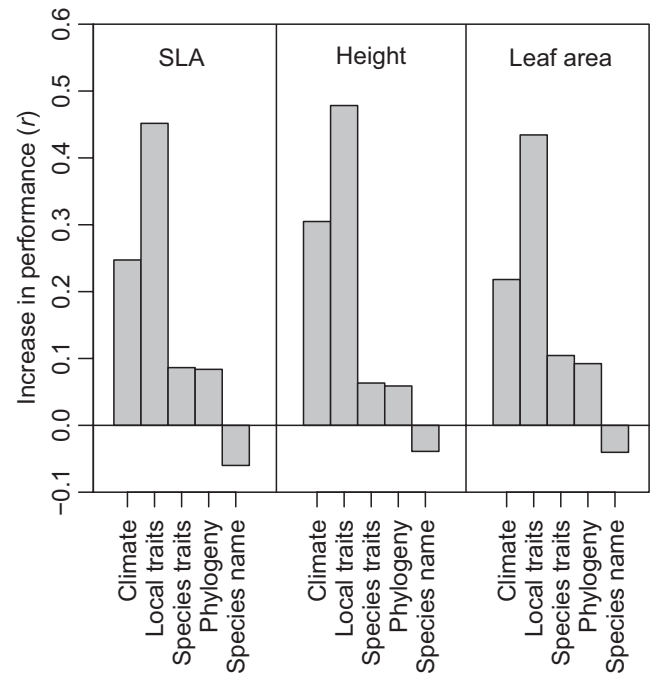
**TABLE 1** Standardized regression coefficients (and their standard errors) and  $R^2$  for multiple regressions predicting ITV of specific leaf area (SLA), Height and leaf area (LA) from local annual precipitation (AP), mean annual temperature (MAT) and their interaction (AP:MAT). The mean coefficients, number of species with positive or negative regression slopes (and the number of each that were significantly different from zero) and mean  $R^2$  are summarized below; bold face:  $p < 0.05$ , † denotes  $p < 0.05$  after a Bonferroni correction

Species	SLA			Height			LA		
	AP	MAT	AP:MAT	$R^2$	AP	MAT	AP:MAT	AP	$R^2$
<i>Aira caryophylllea</i>	0.019 (0.252)	0.499 (0.273)	0.555 (0.399)	0.104	-0.233 (0.267)	-0.035 (0.268)	-0.583 (0.399)	-0.177 (0.263)	0.073
<i>Avena barbata</i>	-0.033 (0.089)	0.109 (0.077)	-0.087 (0.085)	0.022	-0.404† (0.094)	0.428† (0.094)	-0.265 (0.133)	-0.241 (0.092)	0.165
<i>Avena fatua</i>	-1.17† (0.116)	-1.28† (0.148)	-0.34† (0.058)	0.847	-0.601 (0.377)	-1.036 (0.517)	-0.347 (0.231)	-0.745† (0.194)	0.523
<i>Briza maxima</i>	-0.521 (0.175)	0.413 (0.177)	-0.068 (0.244)	0.266	-0.156 (0.168)	0.422 (0.269)	0.034 (0.267)	-0.442 (0.202)	0.131
<i>Briza minor</i>	0.129 (0.151)	0.377 (0.156)	0.209 (0.214)	0.138	-0.343 (0.25)	-0.566 (0.197)	-0.579 (0.293)	-0.027 (0.147)	0.184
<i>Bromus carinatus</i>	-0.722 (0.347)	0.51 (0.306)	0.778 (0.698)	0.226	-0.138 (0.388)	0.161 (0.343)	-0.09 (0.774)	-0.643 (0.343)	0.364
<i>Bromus diandrus</i>	-0.053 (0.098)	0.412† (0.129)	0.311 (0.108)	0.105	-0.296 (0.108)	0.215 (0.146)	-0.035 (0.115)	-0.414† (0.098)	0.172
<i>Bromus hordeaceus</i>	0.366† (0.078)	0.203 (0.079)	0.266† (0.08)	0.165	-0.234 (0.097)	-0.191 (0.109)	-0.545† (0.135)	-0.218 (0.101)	0.117
<i>Bromus madritensis</i>	0.305 (0.148)	0.048 (0.144)	0.062 (0.193)	0.081	0.198 (0.172)	0.298 (0.166)	0.288 (0.25)	0.291 (0.146)	0.078
<i>Cynosurus echinatus</i>	0.588 (0.192)	0.351 (0.167)	0.568 (0.26)	0.261	-0.647† (0.194)	-0.175 (0.15)	-1.527† (0.267)	-0.208 (0.229)	0.275
<i>Danthonia californica</i>	-0.214 (0.15)	0.505 (0.171)	0.086 (0.24)	0.207	-0.778 (0.374)	0.738 (0.223)	0.314 (0.538)	-0.135 (0.199)	0.122
<i>Elymus glaucus</i>	-0.219 (0.169)	0.411 (0.167)	-0.478 (0.287)	0.279	0.092 (0.193)	0.497 (0.184)	-1.326 (0.373)	0.514† (0.142)	0.488
<i>Hordeum marinum</i>	0.374 (0.223)	-0.171 (0.242)	-0.205 (0.28)	0.115	0.32 (0.275)	-0.033 (0.363)	0.553 (0.377)	-1.014† (0.116)	0.828
<i>Hordeum murinum</i>	-0.017 (0.141)	-0.859 (0.299)	1.102 (0.352)	0.181	0.313 (0.213)	0.122 (0.402)	0.003 (0.44)	-0.602 (0.304)	0.162
<i>Lolium multiflorum</i>	0.416† (0.128)	0.258 (0.154)	0.044 (0.05)	0.143					
<i>Lolium perenne</i>	-0.084 (0.15)	0.025 (0.163)	0.157 (0.129)	0.034	-0.135 (0.135)	0.433 (0.15)	-0.035 (0.118)	0.092 (0.175)	0.030
<i>Nassella pulchra</i>	0.093 (0.082)	0.115 (0.097)	0.088 (0.054)	0.028	-0.191 (0.109)	0.084 (0.175)	-0.305 (0.136)	-0.045 (0.13)	0.075
<i>Phalaris aquatica</i>	-0.029 (0.578)	0.371 (0.332)	-0.127 (0.281)	0.395					

(Continues)

TABLE 1 (Continued)

Species	SLA			Height			LA		
	AP	MAT	AP:MAT	R <sup>2</sup>	AP	MAT	AP:MAT	R <sup>2</sup>	AP
<i>Poa secunda</i>	0.639 <sup>†</sup> (0.174)	1.589 <sup>†</sup> (0.274)	-0.772 <sup>†</sup> (0.21)	0.492	0.818 <sup>†</sup> (0.191)	0.27 (0.282)	0.028 (0.12)	0.490	-0.001 (0.318)
<i>Vulpia bromoides</i>					0.928 <sup>†</sup> (0.154)	0.224 (0.171)	0.393 (0.258)	0.750	
<i>Vulpia microstachys</i>	0.28 (0.121)	-0.208 (0.203)	-0.047 (0.295)	0.159	0.372 (0.165)	0.354 (0.304)	-0.042 (0.32)	0.154	0.503 <sup>†</sup> (0.106)
<i>Vulpia myuros</i>	0.215 (0.134)	0.062 (0.136)	0.311 (0.214)	0.109	-0.009 (0.162)	0.617 <sup>†</sup> (0.188)	-0.556 (0.293)	0.289	-0.018 (0.182)
Positive ( <i>p</i> < 0.05)	11 (6)	17 (8)	13 (4)		7 (3)	14 (5)	7 (0)		4 (1)
Negative ( <i>p</i> < 0.05)	10 (2)	4 (2)	8 (2)		13 (4)	6 (1)	13 (5)		15 (6)
Mean	0.017	0.178	0.115	0.208	-0.056	0.141	-0.231	0.280	-0.141
									-0.240
									-0.149
									0.241



**FIGURE 3** Improvements in model performance when adding variable groupings. Model performance was measured as the correlation between observed and predicted delta-trait values in the testing dataset. For each variable group, we take the mean performance of all models that included that variable group minus the mean performance for all models that excluded that variable. Climate variables were mean annual temperature and annual precipitation, local traits were local measures of specific leaf area (SLA), Height or leaf area (LA) at a site, excluding the predicted measures (e.g. models predicting SLA were trained on Height and LA), species traits were the overall species means of SLA, Height and LA, Phylogeny was the first five phylogenetic eigenvector maps, and species name is a categorical variable giving the species name

in warm areas of the state, with large average leaf sizes and in the *Danthonieaea* ( $R^2 = 0.67$ ).

Model coefficients from linear mixed effects models where dataset was treated as a random factor produced similar coefficient estimates (Figure S1). In most cases, the coefficients showed correlations >0.8, although the correlation for the MAT-Height relationship was lower due to three species with large coefficient uncertainties.

### 3.2 | Modelling ITV

Across all specifications of the random forest models, performance scores were very similar on the training and testing data subsets (on average, differing by <0.03, Table S3), suggesting little overfitting. Training and testing performance were particularly similar when species names were not used in the model, resulting in mean differences <0.01. When applied to the testing dataset, random forests containing all five predictor groups predicted values that were well



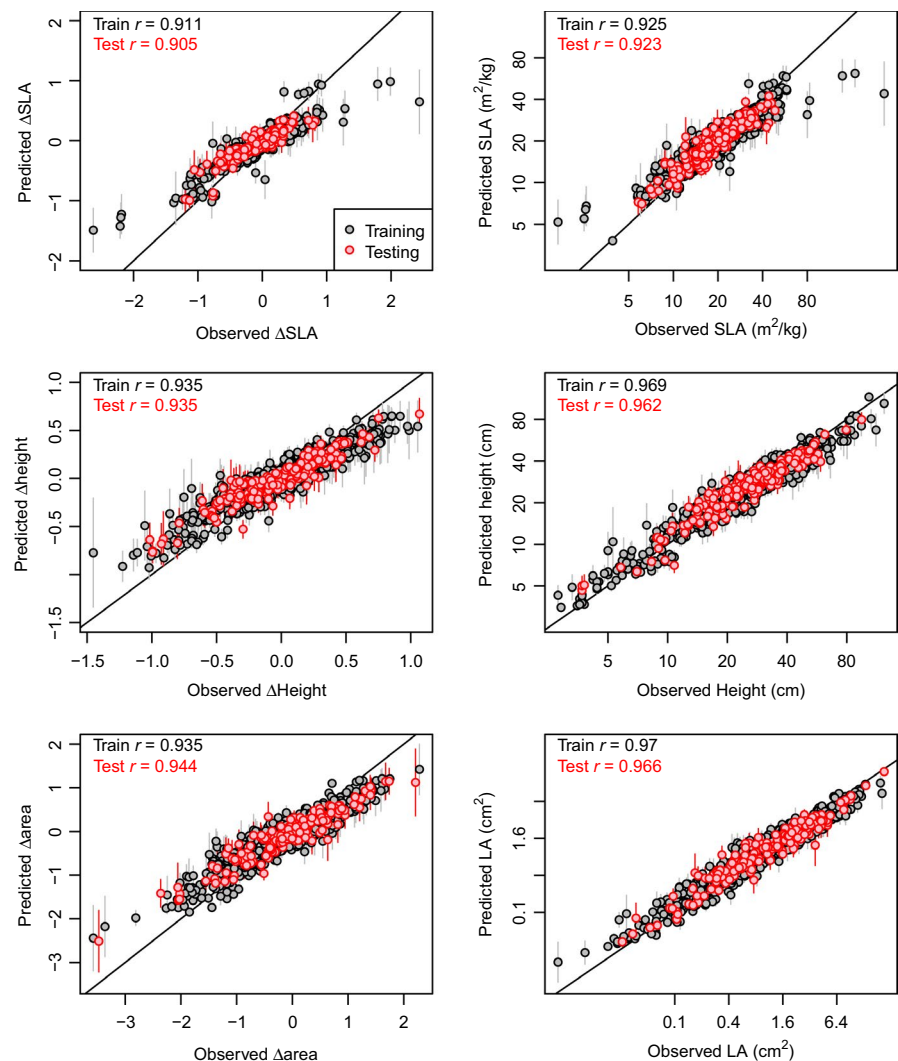
correlated with the observed trait values (for delta-SLA: 0.90, SLA: 0.91, delta-Height: 0.92, Height = 0.96, delta-LA: 0.93, LA: 0.96, Table S3). Across all subsets of variable groups, other local traits (values of the non-focal trait from the local population, e.g. when predicting SLA, the Height of the plants) and to a lesser extent climate were the most important groups for model performance (Figure 3). Species mean traits and phylogeny had smaller contributions to model fit, while species identity tended to hurt model predictive performance. The performance of one such random forest, excluding the species predictor variable, is shown in Figure 4. The correlation between observed and predicted values is strong for both training and testing datasets. However, the observed-predicted relationships deviated somewhat from the 1:1 line, particularly for the delta-trait predictions. Standard major axis (SMA) regression slopes were less than 1, ranging from 0.65 to 0.73 for delta-trait models and 0.82 and 0.92 for the final local trait predictions. These deviations indicate that these models tend to predict less extreme values for the most extreme trait observations.

A model including other local trait measurements and species names would be of limited use for predicting trait values of a plant

in an unmeasured location. In contrast, the climate of that location is readily available, and phylogenetic relationships are known for most species. Thus, we focused on a reduced model including just these two variable groups and two species-level traits: the species mean value for the focal trait and its life span. Removing species names from the model had little impact (Table S3), but removing other local traits reduced model performance (Figure 5). For example, predicted-observed correlations for SLA, Height and LA dropped to 0.85, 0.93 and 0.92. This likely reflects the fact that other local trait measurements can provide insight into local conditions that are not captured by our two broad climate predictors. Despite this modest reduction, model performance for this simplified model was still fairly high.

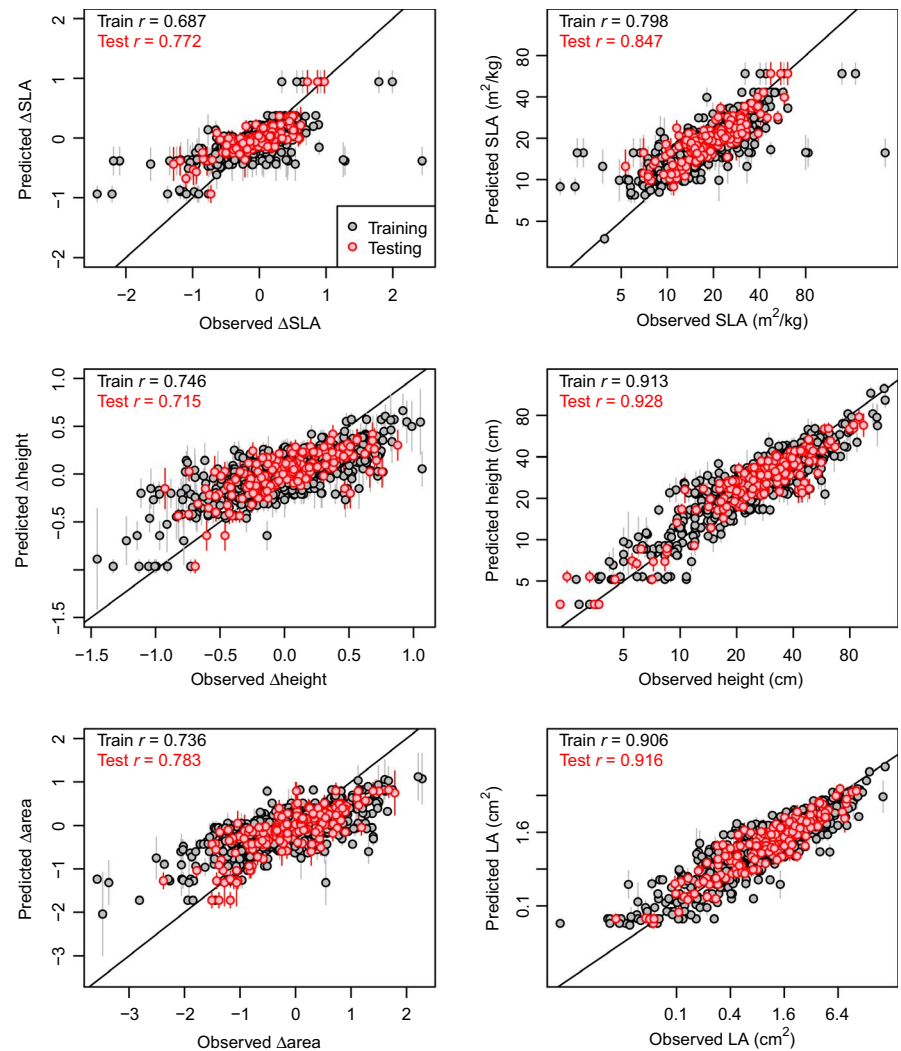
### 3.3 | Predicting ITV

These three final random forests can make predictions for the trait value expected for an unknown species in an unmeasured location, provided the climate conditions at that location and the



**FIGURE 4** Model fit for random forests predicting local trait values from climate, other local traits, species mean traits and phylogenetic position. Each point represents a sample of a grass species from a particular location. Error bars indicate standard errors for the predictions. Models predicting delta-trait values are attempting to predict deviation of an individual from its species mean (left column). Adding the species means to these predictions gives an overall estimate of the trait value for an individual (right column)

**FIGURE 5** Model fit for random forests using only mean traits and phylogeny and trained on the entire dataset. Each point represents a sample of a grass species from a particular location. Error bars indicate standard errors. Models predicting delta-trait values are attempting to predict deviation of an individual from its species mean (left column). Adding the species means to these predictions gives an overall estimate of the trait value for an individual (right column)



mean traits and phylogenetic relationships of the species. This allows us to map a species' expected ITV pattern across its entire distribution in the state (see Figures S2–S17, e.g. with 16 common species).

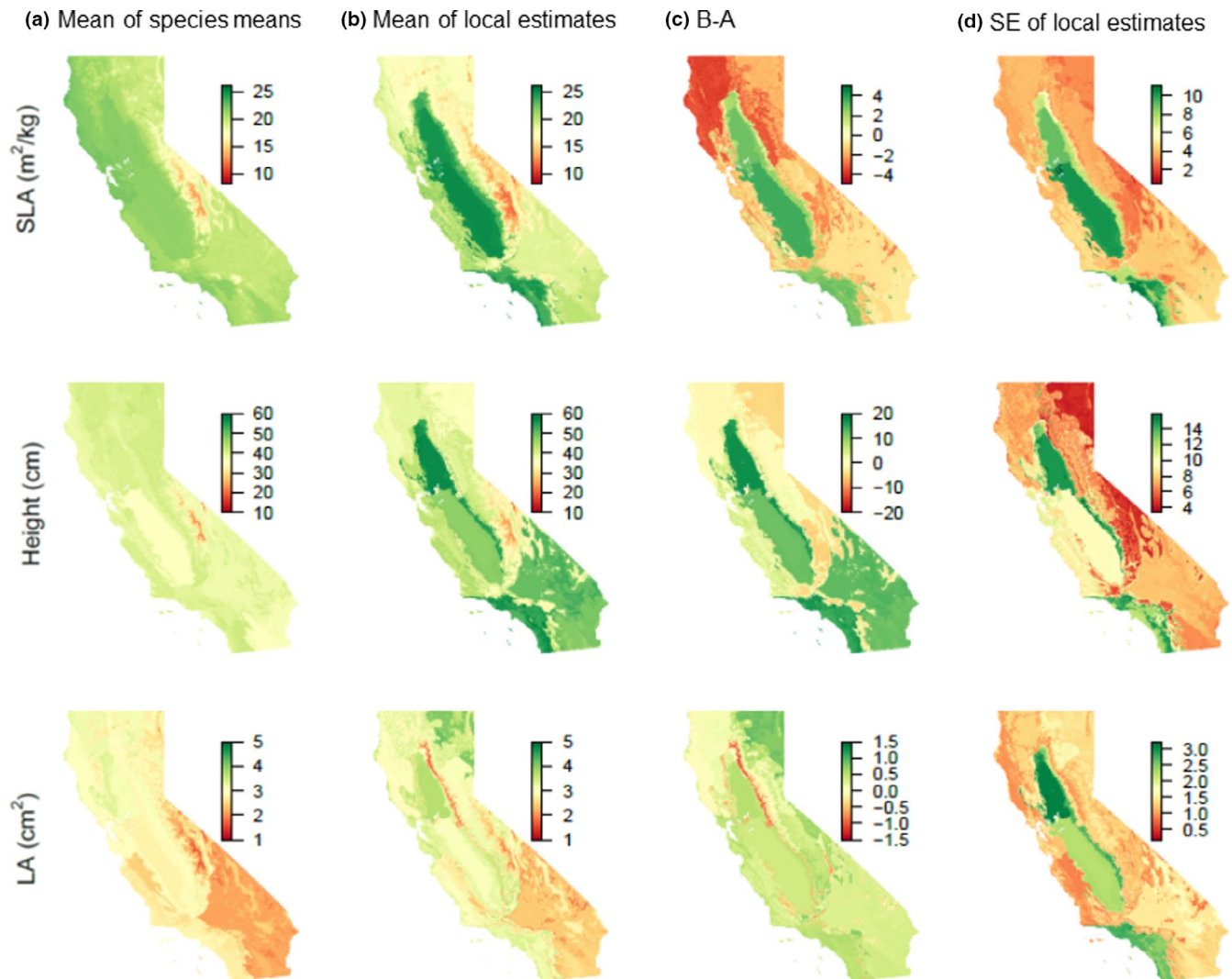
To investigate the importance of ITV in driving functional compositional changes across the state, we computed the mean trait value and standard deviation of all species occurring in each band. The trait values for each species were obtained in one of two ways: (a) as is traditional, by simply applying the species mean trait to each species, or (b), by using our predictive models to estimate the local trait values for each species.

Using local estimates of grass SLA led to lower estimates of band-mean SLA across most of the state, and in some cases produced a substantial change (Figure 6). For Height, a strong tendency to predict taller plants within each species in warmer areas produced a marked increase in the predicted band-mean Height in warm parts of the state, including the southern coast, Mojave Desert and Central Valley (Figures 6 and 7). Including ITV typically lead to a slight increase in estimated LA, particularly in drier parts of the state. Including ITV had a relatively minor effect on the variation of traits within assemblages, but consistently decreased it across the state

for LA. SLA variation decreased somewhat in most bands, while Height variation tended to increase (Figure S18).

## 4 | DISCUSSION

Many recent studies indicate that intraspecific trait–environment relationships can differ among species (Lajoie & Vellend, 2015; Laughlin et al. 2012; Roybal & Butterfield, 2019). Our results corroborate this. For each of the six combinations of trait and climate variable, there was at least one species that showed a significant positive response and at least one with a significant negative response. On its face, this result is daunting for achieving generality. However, in some cases a large part of this variation among species may be attributed to phylogenetic position and species mean traits. For example, perennial species tended to have more positive temperature–Height relationships, and larger-leaved species tended to have more negative precipitation–leaf area relationships. With respect to the phylogeny, grasses in tribe Danthonieae often responded differently than other tribes (Figure S19). Thus, there is hope that, for sufficiently large



**FIGURE 6** Maps of assemblage mean trait values based on species means (column a), local estimates for each species (b) and the difference between them (c). Also shown are the mean standard error of local estimates (d), indicating areas with greater uncertainty in predictions for local species trait values

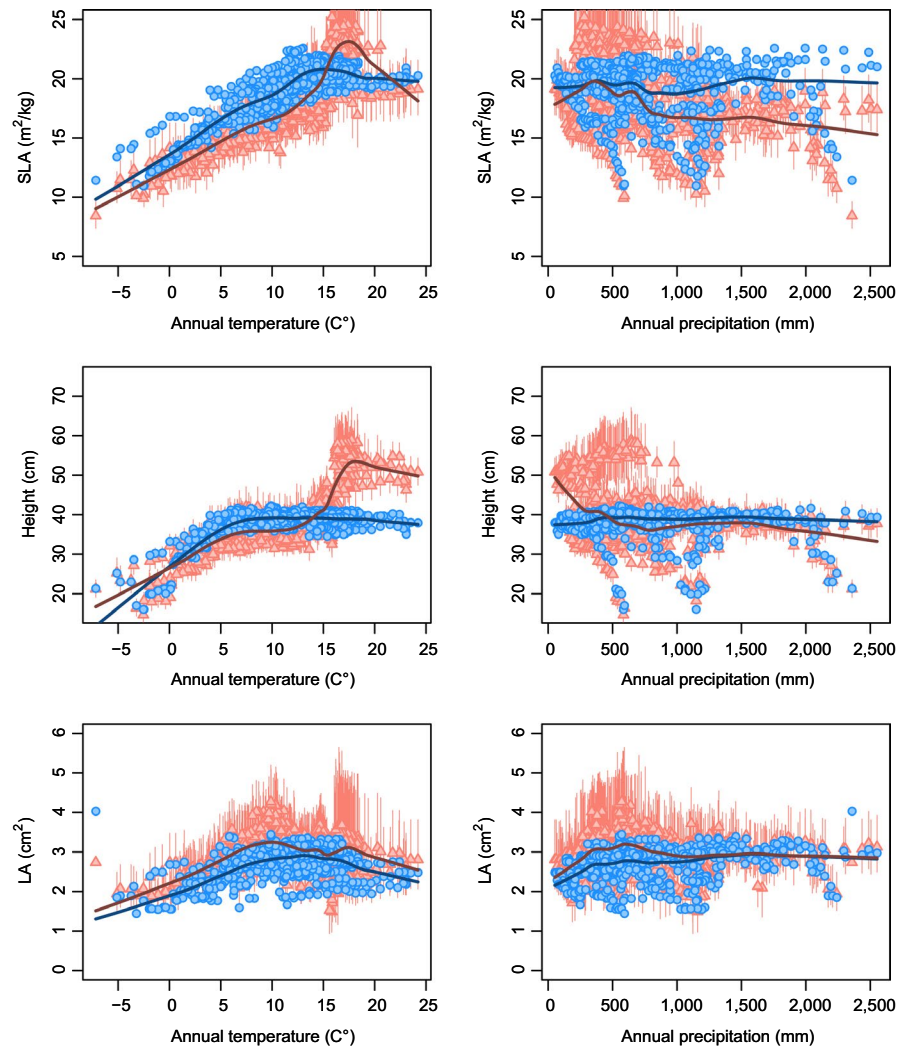
datasets, we can begin to understand the underlying drivers of among-species variation of ITV-climate responses. Building theoretical expectations for variation in these responses, and testing them with new independent datasets, is an important challenge for the future.

Our random forest models exhibited good predictive performance. This is especially true when other local trait measurements were included in predictions. In part, this was driven by covariation of Height and LA both among and within species. This probably reflects allometric relationships between these two traits, whereby taller individuals tend to have larger leaves. However, no such simple relationships existed between SLA and Height or LA, yet the local values of these traits also substantially improved SLA predictions. This may reflect the fact that local trait measurements may be influenced by local environmental conditions that are not captured by broad climate variables. Models still performed fairly well even without other local trait values, suggesting that they can be usefully applied to areas where no measurements exist.

The idiosyncratic trait-climate relationships observed within species are somewhat predictable. However, there is a limit to the flexibility of a random forest to model very different responses in functionally similar and phylogenetically closely related species. The species pair of *Vulpia myuros* and *Vulpia microstachys* provides a good example. These two species are fairly similar in their mean Height, SLA, LA and life span. They are also very close relatives. However, some of their trait-climate relationships differed and the random forest models were not able to predict these different responses (Figures S16 and S17); the spatial patterns of trait predictions for the two species were moderately to strongly correlated (SLA  $r = 0.60$ , Height  $r = 0.90$ , LA  $r = 0.53$ ). These species may be separable along other axes not studied here, or this may represent a limit to the predictability of trait-climate relationships across species.

The overall magnitude of ITV was high. For SLA, only 34% of the variation is attributable to species, but for Height and LA 63% and 68% were explained by species. Within one well-sampled species

**FIGURE 7** The relationships between annual temperature or precipitation and assemblage mean traits based on species means (blue circles) or local estimates (red triangles). For local estimates, the error bars show the mean standard error across all species' trait estimate in that location. Fitted lines are loess regressions



(*Avena barbata*,  $n = 184$ ), SLA varied from 5.8 to 81.7  $\text{m}^2/\text{kg}$ , Height varied from 15 to 106 cm and LA varied from 0.18 to 9.2  $\text{cm}^2$ . Thus, there is a large potential for ITV to drive important changes at the ecosystem level. For example, we estimate that using species means results in an overestimation of assemblage mean SLA by about 2–4  $\text{m}^2/\text{kg}$  for much of the state. Higher assemblage-level SLA is associated with a number of ecosystem properties, including higher flammability (Grootemaat et al., 2017) and faster decomposition rates (Liu et al., 2018), so failing to account for ITV in large-scale models of these processes will lead to biased estimates.

The relationships between mean annual temperature and trait values were largely congruent between and within species. This supports the idea that trait-climate relationships predicted at the inter-species level can often be transferred to the intra-species level (Sandel & Low, 2019). For example, warmer regions tended to contain species with higher mean SLA and Height, and within those species, the individuals with relatively high SLA and Height. Still, there was substantial variation among species, and some species displayed opposite ITV patterns. Patterns with respect to precipitation were less clear. As in a previous study of interspecific trait variation in California grasses (Sandel & Dangremond, 2012), we found fairly weak relationships between annual precipitation and trait means. In

contrast, woody plants in California show stronger relationships between precipitation and plant Height or SLA in some cases (Harrison et al., 2020). Interspecific trait-precipitation relationships were somewhat scattered, with some positive and negative relationships, but there was a hint of an overall tendency towards positive relationship between precipitation and SLA.

At a global scale, taller grass species with larger leaves tend to be found in warmer and wetter climates (Jardine et al., 2020; Sandel et al., 2016). Globally, high SLA has also been weakly associated with warmer temperatures (Sandel et al., 2016), although a different approach revealed essentially no relationship (Jardine et al., 2020). In contrast, within California we found that temperature was consistently more important than precipitation, and that SLA increases with temperature at both inter- and intraspecific levels. The relatively weak responses to precipitation may reflect variation in drought response strategies among species, such as avoidance versus tolerance strategies (Vaughn et al., 2011).

Some of this discrepancy might stem from the features of California's climate. Most of the samples used here are from Mediterranean climates, characterized by cool wet winters and warm dry summers. The growing season is in the winter and is limited by water availability—thus higher evaporative water losses



associated with higher temperatures lead to shorter growing seasons in grasslands (Chiariello, 1989). Low-SLA leaves have higher production costs for a given area but compensate with longer leaf life spans (Wright et al., 2004). However, this investment might be wasted in warmer regions with short growing seasons. Thus, warmer regions will tend to have high-SLA leaves which are cheaper but shorter-lived. However, while this effect might be expected to be strongest for annual species, which tend to pursue a drought-avoidance strategy (Vaughn et al., 2011), a species' life span was not a significant predictor of its ITV response to precipitation for any of the traits.

The relationships between precipitation and Height were somewhat surprising. Ignoring ITV and considering only the influence of turnover in species composition, there was nearly no AP–Height relationship. This contrasts with results for all plants at global and continental scales (Bruehlheide et al., 2018; Moles et al., 2009; Šimová et al., 2018), and with grasses at a global scale (Jardine et al., 2020; Sandel et al., 2016) and regional scale (Forrestel et al., 2017). However, it is consistent with a previous study of California grasses (Sandel & Dangremond, 2012) and grasses within the San Francisco Bay Area (Sandel & Low, 2019). Temperature, however, exerted strong positive influences on Height at both inter- and intraspecific levels. Thus, it is possible that the weak and inconsistent responses to precipitation were largely due to covariance with temperature within each species, rather than a direct effect of precipitation itself.

An important limitation of our approach here is that we used long-term climate averages to describe each site. However, plants may express different traits from year to year as conditions change, for example producing higher SLA leaves during a wet year (Dwyer et al., 2014). This might be particularly true for annual species. If this is the case, we should expect poorer model fits for annual species, and there was some indication that this was the case. In general, we expect that further efforts will benefit from refining the fairly coarse climate descriptions used here.

## 5 | CONCLUSIONS

Plant species can show remarkable variation in functional traits along climate gradients. Paired with uneven spatial sampling, this can lead to strongly biased estimates of species mean trait values and the misrepresentation of these species in ecosystem models. Patterns of trait variation within species may initially seem idiosyncratic, but can be fairly well explained by their phylogenetic position, functional strategies and climate. This predictability provides a means to correct biased species-level estimates, improving the precision of models that rely on such estimates.

## ACKNOWLEDGEMENTS

We are grateful to the many landowners who have supported work on their land, and to the numerous students and assistants who have contributed to the data collection. Special thanks to

Susan Harrison for contributing her grass trait measurements. L.L. was supported by the NSF Postdoctoral Research Fellowship in Biology Award #1309014. G.S.K., A.R.K. and N.J.B.K. were supported by NSF DEB #1644641. We have no conflict of interest to declare.

## AUTHORS' CONTRIBUTIONS

B.S., C.P. and T.H. conceived of this project and performed the data analysis; L.C., J.F., F.W.H., G.S.K., A.R.K., N.J.B.K., L.L., T.M.-M. and M.J.S. contributed data, and all authors wrote the manuscript.

## DATA AVAILABILITY STATEMENT

Data available from the Dryad Digital Repository <https://doi.org/10.5061/dryad.3r2280gfw> (Sandel et al., 2021).

## ORCID

Brody Sandel  <https://orcid.org/0000-0003-2162-6902>

Fletcher W. Halliday  <https://orcid.org/0000-0003-3953-0861>

Gaurav S. Kandlikar  <https://orcid.org/0000-0003-3043-6780>

## REFERENCES

- Ackerly, D. D., & Cornwell, W. K. (2007). A trait-based approach to community assembly: Partitioning of species trait values into within- and among-community components. *Ecology Letters*, 10(2), 135–145. <https://doi.org/10.1111/j.1461-0248.2006.01006.x>
- Albert, C. H., de Bello, F., Boulangeat, I., Pellet, G., Lavorel, S., & Thuiller, W. (2012). On the importance of intraspecific variability for the quantification of functional diversity. *Oikos*, 121(1), 116–126.
- Albert, C. H., Grassein, F., Schurr, F. M., Vieilledent, G., & Violle, C. (2011). When and how should intraspecific variability be considered in trait-based plant ecology? *Perspectives in Plant Ecology Evolution and Systematics*, 13(3), 217–225.
- Baughman, O. W., Agneray, A. C., Forister, M. L., Kilkenny, F. F., Espeland, E. K., Fiegner, R., Horning, M. E., Johnson, R. C., Kaye, T. N., Ott, J., & St. Clair, J. B. (2019). Strong patterns of intraspecific variation and local adaptation in Great Basin plants revealed through a review of 75 years of experiments. *Ecology and Evolution*, 9(11), 6259–6275.
- Berzaghi, F., Wright, I. J., Kramer, K., Oddou-Muratorio, S., Bohn, F. J., Reyer, C. P., Sabaté, S., Sanders, T. G., & Hartig, F. (2020). Towards a new generation of trait-flexible vegetation models. *Trends in Ecology & Evolution*, 35(3), 191–205. <https://doi.org/10.1016/j.tree.2019.11.006>
- Bjorkman, A. D., Myers-Smith, I. H., Elmendorf, S. C., Normand, S., Thomas, H. J., Alatalo, J. M., Alexander, H., Anadon-Rosell, A., Angers-Blondin, S., Bai, Y., & Baruah, G. (2018). Tundra trait team: A database of plant traits spanning the tundra biome. *Global Ecology and Biogeography*, 27(12), 1402–1411.
- Bloomfield, K. J., Cernusak, L. A., Eamus, D., Ellsworth, D. S., Colin Prentice, I., Wright, I. J., Boer, M. M., Bradford, M. G., Cale, P., Cleverly, J., & Egerton, J. J. (2018). A continental-scale assessment of variability in leaf traits: Within species, across sites and between seasons. *Functional Ecology*, 32(6), 1492–1506.
- Bonan, G. B., Levis, S., Sitch, S., Vertenstein, M., & Oleson, K. W. (2003). A dynamic global vegetation model for use with climate models: Concepts and description of simulated vegetation dynamics. *Global Change Biology*, 9(11), 1543–1566. <https://doi.org/10.1046/j.1365-2486.2003.00681.x>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- Bruehlheide, H., Dengler, J., Purschke, O., Lenoir, J., Jiménez-Alfaro, B., Hennekens, S. M., Botta-Dukát, Z., Chytrý, M., Field, R., Jansen, F.,



- & Kattge, J. (2018). Global trait–environment relationships of plant communities. *Nature Ecology & Evolution*, 2(12), 1906–1917.
- Butler, E. E., Datta, A., Flores-Moreno, H., Chen, M., Wythers, K. R., Fazayeli, F., Banerjee, A., Atkin, O. K., Kattge, J., Amiaud, B., Blonder, B., Boenisch, G., Bond-Lamberty, B., Brown, K. A., Byun, C., Campetella, G., Cerabolini, B. E. L., Cornelissen, J. H. C., Craine, J. M., ... Reich, P. B. (2017). Mapping local and global variability in plant trait distributions. *Proceedings of the National Academy of Sciences of the United States of America*, 114(51), E10937–E10946. <https://doi.org/10.1073/pnas.1708984114>
- Carlucci, M. B., Debastiani, V. J., Pillar, V. D., & Duarte, L. D. (2015). Between-and within-species trait variability and the assembly of sapling communities in forest patches. *Journal of Vegetation Science*, 26(1), 21–31. <https://doi.org/10.1111/jvs.12223>
- Chiariello, N. R. (1989). Phenology of California grasslands. In L. F. Huenneke, & H. A. Mooney (Eds.), *Grassland structure and function* (pp. 47–58). Springer.
- Christin, P. A., Spriggs, E., Osborne, C. P., Strömberg, C. A., Salamin, N., & Edwards, E. J. (2014). Molecular dating, evolutionary rates, and the age of the grasses. *Systematic Biology*, 63(2), 153–165. <https://doi.org/10.1093/sysbio/syt072>
- Clayton, W. D., Vorontsova, M. S., Harman, K. T., & Williamson, H. (2006 onwards). *GrassBase - The Online World Grass Flora*. Retrieved from <http://www.kew.org/data/grasses-db.html>
- Cornwell, W. K., & Ackerly, D. D. (2009). Community assembly and shifts in plant trait distributions across an environmental gradient in coastal California. *Ecological Monographs*, 79(1), 109–126.
- Cornwell, W. K., Cornelissen, J. H., Amatangelo, K., Dorrepaal, E., Eviner, V. T., Godoy, O., Hobbie, S. E., Hoorens, B., Kurokawa, H., Pérez-Harguindeguy, N., & Queded, H. M. (2008). Plant species traits are the predominant control on litter decomposition rates within biomes worldwide. *Ecology Letters*, 11(10), 1065–1071.
- Crutsinger, G. M., Collins, M. D., Fordyce, J. A., Gompert, Z., Nice, C. C., & Sanders, N. J. (2006). Plant genotypic diversity predicts community structure and governs an ecosystem process. *Science*, 313(5789), 966–968.
- Cutler, D. R., Edwards Jr., T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology*, 88(11), 2783–2792.
- Davis, F. W., Stoms, D. M., Hollander, A. D., Thomas, K. A., Stine, P. A., Odion, D., Borchert, M. I., Thorne, J. H., Gray, M. V., Walker, R. E., Warner, K., & Graae, J. (1998). *The California Gap Analysis Project - Final Report*. University of California.
- Derroire, G., Powers, J. S., Hulshof, C. M., Varela, L. E. C., & Healey, J. R. (2018). Contrasting patterns of leaf trait variation among and within species during tropical dry forest succession in Costa Rica. *Scientific Reports*, 8(1), 285.
- Diaz, S., Cabido, M., & Casanoves, F. (1998). Plant functional traits and environmental filters at a regional scale. *Journal of Vegetation Science*, 9, 113–122.
- Dwyer, J. M., Hobbs, R. J., & Mayfield, M. M. (2014). Specific leaf area responses to environmental gradients through space and time. *Ecology*, 95(2), 399–410.
- Edwards, E. J., Osborne, C. P., Strömberg, C. A., Smith, S. A., & C4 Grasses Consortium. (2010). The origins of C4 grasslands: Integrating evolutionary and ecosystem science. *Science*, 328(5978), 587–591.
- Enquist, B. J., & Niklas, K. J. (2002). Global allocation rules for patterns of biomass partitioning in seed plants. *Science*, 295(5559), 1517–1520.
- Falster, D. S., & Westoby, M. (2003). Plant height and evolutionary games. *Trends in Ecology & Evolution*, 18(7), 337–343.
- Forrestel, E. J., Donoghue, M. J., Edwards, E. J., Jetz, W., du Toit, J. C., & Smith, M. D. (2017). Different clades and traits yield similar grassland functional responses. *Proceedings of the National Academy of Sciences*, 114(4), 705–710.
- Garnier, E., Cortez, J., Billès, G., Navas, M.-L., Roumet, C., Debussche, M., Laurent, G., Blanchard, A., Aubry, D., Bellmann, A., Neill, C., & Toussaint, J.-P. (2004). Plant functional markers capture ecosystem properties during secondary succession. *Ecology*, 85(9), 2630–2637. <https://doi.org/10.1890/03-0799>
- Grootemaat, S., Wright, I. J., van Bodegom, P. M., & Cornelissen, J. H. (2017). Scaling up flammability from individual leaves to fuel beds. *Oikos*, 126(10), 1428–1438. <https://doi.org/10.1111/oik.03886>
- Guénard, G., Legendre, P., & Peres-Neto, P. (2013). Phylogenetic eigenvector maps: A framework to model and predict species traits. *Methods in Ecology and Evolution*, 4(12), 1120–1131. <https://doi.org/10.1111/2041-210X.12111>
- Harrison, S., Spasojevic, M. J., & Li, D. (2020). Climate and plant community diversity in space and time. *Proceedings of the National Academy of Sciences of the United States of America*, 117(9), 4464–4470.
- Hart, S. P., Schreiber, S. J., & Levine, J. M. (2016). How variation between individuals affects species coexistence. *Ecology Letters*, 19(8), 825–838. <https://doi.org/10.1111/ele.12618>
- Heckman, R. W., Halliday, F. W., & Mitchell, C. E. (2019). A growth–defense trade-off is general across native and exotic grasses. *Oecologia*, 191(3), 609–620. <https://doi.org/10.1007/s00442-019-04507-9>
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 25(15), 1965–1978. <https://doi.org/10.1002/joc.1276>
- Jardine, E. C., Thomas, G. H., Forrestel, E. J., Lehmann, C. E., & Osborne, C. P. (2020). The global distribution of grass functional traits within grassy biomes. *Journal of Biogeography*, 47(3), 553–565.
- Jepson Flora Project (eds.). (2019). *Jepson eFlora*. Retrieved from <http://ucjeps.berkeley.edu/eflora/>
- Kattge, J., Diaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Bönsch, G., Garnier, E., Westoby, M., Reich, P. B., Wright, I. J., & Cornelissen, J. H. C. (2011). TRY—a global database of plant traits. *Global Change Biology*, 17(9), 2905–2935.
- Kichenin, E., Wardle, D. A., Peltzer, D. A., Morse, C. W., & Freschet, G. T. (2013). Contrasting effects of plant inter- and intraspecific variation on community-level trait measures along an environmental gradient. *Functional Ecology*, 27(5), 1254–1261. <https://doi.org/10.1111/1365-2435.12116>
- Kraft, N. J., Adler, P. B., Godoy, O., James, E. C., Fuller, S., & Levine, J. M. (2015). Community assembly, coexistence and the environmental filtering metaphor. *Functional Ecology*, 29(5), 592–599. <https://doi.org/10.1111/1365-2435.12345>
- Kumordzi, B. B., Wardle, D. A., & Freschet, G. T. (2015). Plant assemblages do not respond homogeneously to local variation in environmental conditions: Functional responses differ with species identity and abundance. *Journal of Vegetation Science*, 26(1), 32–45. <https://doi.org/10.1111/jvs.12218>
- Lajoie, G., & Vellend, M. (2015). Understanding context dependence in the contribution of intraspecific variation to community trait–environment matching. *Ecology*, 96(11), 2912–2922. <https://doi.org/10.1890/15-0156.1>
- Laughlin, D. C., Joshi, C., van Bodegom, P. M., Bastow, Z. A., & Fulé, P. Z. (2012). A predictive model of community assembly that incorporates intraspecific trait variation. *Ecology Letters*, 15(11), 1291–1299. <https://doi.org/10.1111/j.1461-0248.2012.01852.x>
- Lepš, J., de Bello, F., Šmilauer, P., & Doležal, J. (2011). Community trait response to environment: Disentangling species turnover vs intraspecific trait variability effects. *Ecography*, 34(5), 856–863. <https://doi.org/10.1111/j.1600-0587.2010.06904.x>

- Lichstein, J. W., Dushoff, J., Levin, S. A., & Pacala, S. W. (2007). Intraspecific variation and species coexistence. *The American Naturalist*, 170(6), 807–818. <https://doi.org/10.1086/522937>
- Liu, G., Wang, L., Jiang, L., Pan, X., Huang, Z., Dong, M., & Cornelissen, J. H. (2018). Specific leaf area predicts dryland litter decomposition via two mechanisms. *Journal of Ecology*, 106(1), 218–229.
- Makonnen, M., Berg, M. P., Handa, I. T., Hättenschwiler, S., van Ruijven, J., van Bodegom, P. M., & Aerts, R. (2012). Highly consistent effects of plant litter identity and functional traits on decomposition across a latitudinal gradient. *Ecology Letters*, 15(9), 1033–1041.
- Moles, A. T., Warton, D. I., Warman, L., Swenson, N. G., Laffan, S. W., Zanne, A. E., Pitman, A., Hemmings, F. A., & Leishman, M. R. (2009). Global patterns in plant height. *Journal of Ecology*, 97(5), 923–932. <https://doi.org/10.1111/j.1365-2745.2009.01526.x>
- Paradis, E. (2013). Molecular dating of phylogenies by likelihood methods: A comparison of models and a new information criterion. *Molecular Phylogenetics and Evolution*, 67(2), 436–444.
- Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics*, 20, 289–290.
- Pistón, N., de Bello, F., Dias, A. T., Götzenberger, L., Rosado, B. H., de Mattos, E. A., Salguero-Gómez, R., & Carmona, C. P. (2019). Multidimensional ecological analyses demonstrate how interactions between functional traits shape fitness and life history strategies. *Journal of Ecology*, 107(5), 2317–2328. <https://doi.org/10.1111/1365-2745.13190>
- Pontes, L. D. S., Soussana, J. F., Louault, F., Andueza, D., & Carrere, P. (2007). Leaf traits affect the above-ground productivity and quality of pasture grasses. *Functional Ecology*, 21(5), 844–853. <https://doi.org/10.1111/j.1365-2435.2007.01316.x>
- PRISM Climate Group. (2018). Oregon State University. Retrieved from <http://prism.oregonstate.edu>, created 20 April, 2018.
- Roybal, C. M., & Butterfield, B. J. (2019). Species-specific trait-environment relationships among populations of widespread grass species. *Oecologia*, 189(4), 1017–1026. <https://doi.org/10.1007/s00442-019-04372-6>
- Sakschewski, B., von Bloh, W., Boit, A., Rammig, A., Kattge, J., Poorter, L., Peñuelas, J., & Thonicke, K. (2015). Leaf and stem economics spectra drive diversity of functional plant traits in a dynamic global vegetation model. *Global Change Biology*, 21(7), 2711–2725.
- Sandel, B., & Dangremond, E. M. (2012). Climate change and the invasion of California by grasses. *Global Change Biology*, 18(1), 277–289.
- Sandel, B., & Low, R. (2019). Intraspecific trait variation, functional turnover and trait differences among native and exotic grasses along a precipitation gradient. *Journal of Vegetation Science*, 30(4), 633–643.
- Sandel, B., Monnet, A. C., & Vorontsova, M. (2016). Multidimensional structure of grass functional traits among species and assemblages. *Journal of Vegetation Science*, 27(5), 1047–1060. <https://doi.org/10.1111/jvs.12422>
- Sandel, B., Pavelka, C., Hayashi, T., Charles, L., Funk, J., Halliday, F. W., Kandlikar, G. S., Kleinhesselink, A. R., Kraft, N. J. B., Larios, L., Madsen-McQueen, T., & Spasojevic, M. J. (2021). Data from: Predicting intraspecific trait variation among California's grasses. *Dryad Digital Repository*, <https://doi.org/10.5061/dryad.3r2280gfw>
- Sandel, B., & Tsirogiannis, C. (2016). Species introductions and the phylogenetic and functional structure of California's grasses. *Ecology*, 97(2), 472–483.
- Sanderson, M. J. (2002). Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. *Molecular Biology and Evolution*, 19(1), 101–109.
- Sato, H., Itoh, A., & Kohyama, T. (2007). SEIB-DGVM: A new Dynamic Global Vegetation Model using a spatially explicit individual-based approach. *Ecological Modelling*, 200(3–4), 279–307.
- Scheiter, S., Langan, L., & Higgins, S. I. (2013). Next-generation dynamic global vegetation models: Learning from community ecology. *New Phytologist*, 198(3), 957–969. <https://doi.org/10.1111/nph.12210>
- Siefert, A. (2012). Incorporating intraspecific variation in tests of trait-based community assembly. *Oecologia*, 170(3), 767–775.
- Siefert, A., Violle, C., Chalmandrier, L., Albert, C. H., Taudiere, A., Fajardo, A., Aarssen, L. W., Baraloto, C., Carlucci, M. B., Cianciaruso, M. V., & de L Dantas, V. (2015). A global meta-analysis of the relative extent of intraspecific trait variation in plant communities. *Ecology Letters*, 18(12), 1406–1419.
- Šimová, I., Violle, C., Svenning, J. C., Kattge, J., Engemann, K., Sandel, B., Peet, R. K., Wiser, S. K., Blonder, B., McGill, B. J., & Boyle, B. (2018). Spatial patterns and climate relationships of major plant traits in the New World differ between woody and herbaceous species. *Journal of Biogeography*, 45(4), 895–916.
- Smith, C. C., & Fretwell, S. D. (1974). The optimal balance between size and number of offspring. *The American Naturalist*, 108, 499–506. <https://doi.org/10.1086/282929>
- Stahl, U., Reu, B., & Wirth, C. (2014). Predicting species' range limits from functional traits for the tree flora of North America. *Proceedings of the National Academy of Sciences of the United States of America*, 111(38), 13739–13744. <https://doi.org/10.1073/pnas.1300673111>
- Suding, K. N., Lavelle, S., Chapin III, F. S., Cornelissen, J. H., Díaz, S., Garnier, E., Goldberg, D., Hooper, D. U., Jackson, S. T., & Navas, M. L. (2008). Scaling environmental change through the community-level: A trait-based response-and-effect framework for plants. *Global Change Biology*, 14(5), 1125–1140.
- Tuller, J., Marquis, R. J., Andrade, S. M., Monteiro, A. B., & Faria, L. D. (2018). Trade-offs between growth, reproduction and defense in response to resource availability manipulations. *PLoS ONE*, 13(8), e0201873.
- Vaughn, K. J., Biel, C., Clary, J. J., de Herralde, F., Aranda, X., Evans, R. Y., Young, T. P., & Savé, R. (2011). California perennial grasses are physiologically distinct from both Mediterranean annual and perennial grasses. *Plant and Soil*, 345(1–2), 37–46.
- Verheijen, L. M., Aerts, R., Brovkin, V., Cavender-Bares, J., Cornelissen, J. H., Kattge, J., & Van Bodegom, P. M. (2015). Inclusion of ecologically based trait variation in plant functional types reduces the projected land carbon sink in an earth system model. *Global Change Biology*, 21(8), 3074–3086. <https://doi.org/10.1111/gcb.12871>
- Violle, C., Enquist, B. J., McGill, B. J., Jiang, L., Albert, C. H., Hulshof, C., Jung, V., & Messier, J. (2012). The return of the variance: Intraspecific variability in community ecology. *Trends in Ecology & Evolution*, 27(4), 244–252. <https://doi.org/10.1016/j.tree.2011.11.014>
- Violle, C., Navas, M. L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., & Garnier, E. (2007). Let the concept of trait be functional! *Oikos*, 116(5), 882–892. <https://doi.org/10.1111/j.0030-1299.2007.15559.x>
- Wager, S., Hastie, T., & Efron, B. (2014). Confidence intervals for random forests: The jackknife and the infinitesimal jackknife. *The Journal of Machine Learning Research*, 15(1), 1625–1651.
- Westoby, M., Falster, D. S., Moles, A. T., Vesk, P. A., & Wright, I. J. (2002). Plant ecological strategies: Some leading dimensions of variation between species. *Annual Review of Ecology and Systematics*, 33, 125–159. <https://doi.org/10.1146/annurev.ecolsys.33.010802.150452>
- Wright, I. J., Reich, P. B., Cornelissen, J. H. C., Falster, D. S., Groom, P. K., Hikosaka, K., Lee, W., Lusk, C. H., Niinemets, Ü., Oleksyn, J., Osada, N., Poorter, H., Warton, D. I., & Westoby, M. (2005). Modulation of leaf economic traits and trait relationships by climate. *Global Ecology and Biogeography*, 14(5), 411–421. <https://doi.org/10.1111/j.1466-822x.2005.00172.x>
- Wright, I. J., Reich, P. B., & Westoby, M. (2003). Least-cost input mixtures of water and nitrogen for photosynthesis. *The American Naturalist*, 161(1), 98–111. <https://doi.org/10.1086/344920>

- Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J. H., Diemer, M., & Flexas, J. (2004). The worldwide leaf economics spectrum. *Nature*, 428(6985), 821–827.
- Wright, M. N., & Ziegler, A. (2017). ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software*, 77(1), 1–17.
- Yang, Y., Zhu, Q., Peng, C., Wang, H., & Chen, H. (2015). From plant functional types to plant functional traits: A new paradigm in modelling global vegetation dynamics. *Progress in Physical Geography*, 39(4), 514–535. <https://doi.org/10.1177/0309133315582018>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Sandel B, Pavelka C, Hayashi T, et al. Predicting intraspecific trait variation among California's grasses. *J Ecol.* 2021;109:2662–2677. <https://doi.org/10.1111/1365-2745.13673>