**Abstract** (<350 words @ Functional Ecology)

1. Purpose

2. Methods

3. Results

4. Discussion

**Introduction**

The extent to which community assembly mechanisms are deterministic processes i.e. niche based is a central topic in ecology. There are two major hypothesized processes that predict contrasting but non-exclusive patterns in the trait and spatial distributions of coexisting species. Environmental filtering structures communities by only allowing only organisms capable of surviving local conditions to persist within the local species pool (Vellend 2010). This mechanism can lead to a convergence in ecological attributes or traits among species in that area. In deserts, traits that allow organisms to cope with environmental stressors such as water stress and temperature extremes are common. In contrast, limiting similarity predicts that competitive exclusion results among ecologically similar species (Abrams 1983). This mechanism can lead to a divergence in traits among coexisting species. These processes work together but at differing scales, at large scales habitat filtering is expected to be stronger than biotic interactions, whereas at the fine scales that interactions actually occur at, biotic interactions prevail.

Ants are ideal system to test assembly processes…

Functional traits reflect aspects of morphology, life history or physiology that enable an organism to be successful in an environment (Violle et al. 2007). Plant ecologists have incorporated the role of facilitation by foundation plants into trait-based community assembly theory (Schöb et al. 2012). In deserts, shrubs reduce environmental stress by reducing daytime heat and retaining warmth overnight, creating locally stable micro-climates (McIntire and Fajardo 2014). Shrub facilitation often increases the cover and biomass of annual plants (Holzapfel et al. 2006, Pugnaire et al. 2011), as well as arthropod species richness and abundance beneath the canopy (Liu et al. 2016, Braun and Lortie 2020, Braun et al. 2021). Surface complexity can act as a filter on ant traits, for example longer-legged ant species are more successful on rugose surfaces that short-legged species (Kaspari and Weiser 1999). Therefore, at fine scales, there is the capacity for foundation shrubs and ground-covering vegetation to alter taxonomic composition, trait distributions and coexistence patterns of the ant community.

The relative importance of environmental and biotic filtering mechanisms to can change along environmental gradients. The stress gradient hypothesis proposes that competitive interactions decrease in importance as environmental stress increases (Bertness and Callaway 1994). There is extensive empirical support for the stress gradient hypothesis in plant communities (Lortie and Callaway 2006, He et al. 2013), however, in animal communities, tests are relatively rare. Due to the ecological importance and dominance of ants in arid ecosystems, understanding how their communities may change under increasing stress and along climatic gradients is important research to predicting the outcome of global change on these ecosystems.

We studied desert ant community assembly. We hypothesized that desert ant communities are assembled through niche-based processes.

We expect the following the predictions:

1. There will be evidence for filtering on ant communities through the following responses to environmental gradients: a) alpha and beta diversity, b) trait convergence or divergence i.e. changes to functional dispersion, and c) specific trait by environment interactions (community weighted mean trait values)
2. Species that are more similarity in terms of their environmental niche will be more similar in terms of their traits
3. Effect size measure of the trait dispersion of co-occurring ants will decrease with stress (stress gradient).

community trait composition (community weight mean trait values) -> This is the db-RDA

**Methods**

*Field collection*

Ant communities (Hymenoptera: Formicidae) were sampled using pitfall traps at nine sites in the San Joaquin Valley, California, USA that span a distance of ~200 km from north to south (Figure 1). Each site was sampled once per month between July and September of 2020 for a total of 27 sampling instances (Appendix Table 1). The shrub species *Ephedra californica* (Ephedraceae) and *Atriplex* sp. (Chenopodiaceae) are the dominant perennial species at six of the sites and the remaining three sites are relatively open with few shrubs. White plastic drink cups (12.4 cm tall, 9 cm diameter) were placed with the top of the cup flush with the ground. In order to prevent vertebrate bycatch, 0.5-inch hardware cloth was placed horizontally within the trap and a piece of aluminum flashing was elevated three cm above the trap to shelter the trap. The traps were filled to a depth of three cm with 100% propylene glycol. Propylene glycol is a biodegradable, non-toxic preservative that does not evaporate and preserves DNA (Nakamura et al. 2020). At shrub sites, traps were placed at 12 pairs of shrub/open microsites and pairs were located at least 10 m apart. Shrub microsites were located beneath the canopy of a foundation shrub at the center of a 0.5 m quadrat placed just inside the dripline of the shrub. Open microsites were located randomly at least 2 m away from shrub microsites. At sites without shrubs in collections areas (i.e. within 500 -1000 m of collections), pitfalls traps were deployed every 10 m in open areas along two transects located at least 10 m apart. At each site, 24 pitfall traps were deployed continuously for 72 hours per sampling event. Each sampling location was georeferenced using a handheld GPS unit. The traps were deployed in different locations within the study site each month to avoid repeated measures. Throughout the season, 648 traps were deployed totaling 46 656 trap-hours (24 traps per site \* 3 sampling instances \* 72 hours). Site-level vegetation characteristics were measured each month by placing 0.5 m quadrats every 4 m in ten 25 m transects distributed around the site and recording the percent cover of ground-covering vegetation, rocky cover, woody cover and bare ground within the quadrat. Vegetation height was measured at the center of the quadrat.

*Lab work*

Ants were first identified to genus using Fisher & Cover (2007), and to species using AntWiki keys (www.antwiki.org). Representative individuals of each species were also barcoded by the Canadian Centre for DNA Barcoding (CCD). The resulting COI? sequences were compared to BINs in BOLD systems (https://boldsystems.org/). Two singletons, *Solenopsis molesta* and *Solenopsis aurea* were excluded from analyses. We measured the following morphological trait data from the ant specimens: Weber’s length, which is the length of the longest diagonal on the thorax and a measure of overall worker body size (Brown 1953); femur length, which relates to foraging ability in complex environment and thermoregulation (Feener Jr et al. 1988); mandible length, and head length and width, which relate to diet; scape length, which relates to chemosensory abilities; and eye size which relates to foraging behaviours (Weiser and Kaspari 2006). The latter six traits were divided by Weber’s body length to remove the impact of body size.

We measured each of the seven traits in up to six individuals per species, per site (265 individual ants) to capture the intraspecific variation in traits between populations. The majority of functional trait studies use traits measured at the species-level (de Bello et al. 2021), however traits captured at the population-level can better reflect environmental filtering (Jung et al. 2010, Violle et al. 2012) and may represent local adaptation (Kawecki and Ebert 2004). We dissected each ant and affixed them to microscope slides using Elmer’s glue. We placed each ant mount slide on top of a stage micrometer slide and took focus-bracketed photographs using a Canon 60D DSLR camera with a 60 mm macro lens and Canon EF 25 II extension tube. Helicon Focus software was used to combine the focus stacks into single images. We imported each composite image into ImageJ software and used the micrometer divisions within the image to calibrate the measurement scale within the software and then measured each trait using the software.

*Analysis*

We extracted the mean annual precipitation, mean annual temperature and maximum annual temperature for each site from WorldClim (Fick and Hijmans 2017). We extracted NDVI (normalized difference vegetation index) from EVIIRS (USGS EROS Visible Infrared Imaging Radiometer Suite). EVIIRS has a 375 m spatial resolution and 7 or 14 day temporal resolution (USGS, 2021). We used the NDVI measurement collected from the date closest to the field sampling date (Appendix Table 2). NDVI is often used as an estimate of plant productivity within animal ecology (Pettorelli et al. 2011). We extracted two below-ground measures from a recently published dataset of global soil temperature: mean annual soil temperature and the annual range of soil temperatures, both at the 5cm to 15 cm depth (Lembrechts et al. 2022).

All analyses used R version 4.1.0 (R Core Team 2022). Ants forage socially and this leads to clustering of individuals from the same colony (Longino and Coddington 2002). We converted the data from abundance to occupancy i.e. the proportion of pitfall traps at site that each ant species was detected in. We calculated site-level vegetation cover and height by taking the mean value of the quadrats. We calculated variation in vegetation cover by taking the standard deviation of cover within the quadrats.

We combined the environmental variables (mean annual precipitation, mean annual temperature and maximum annual temperature, NDVI, mean annual soil temperature, annual range of soil temperatures, mean percent vegetation cover, variation in vegetation cover and mean annual vegetation height) into a composite environmental gradient using PCA. Prior to the PCA all variables were standardized to a mean of zero and unit variance (vegan citation).

We calculated functional dispersion (FDisp) for each of the 27 sites using the population-level trait means using the dbFD functioning in the FD package (Laliberté et al. 2014). In order to determine whether the site is functionally clustered or dispersed relative to a randomly assemble community, we compared the observed communities to a randomly assembled. We generated 1000 random communities using the RandomizeMatrix function in the picante package (Kembel et al. 2010). The null models were generated using the independent swap algorithm, which holds rows sums i.e. species richness and column sums i.e. occupancy/abundance of the population constant while randomizing. FDisp was calculated for each of these 1000 matrices. Standardized effect size (SES) was then calculated for each site using the following formula: (Obs – MeanRandom)/SD Random) (Swenson 2014). Positive values of SES indicate that a community is functionally dispersed, whereas negative values indicate functional trait clustering (de Bello et al. 2021). A one sample t-test was used to determine if the mean value of SES across all the populations is significantly different than zero i.e. clustered or dispersed (de Bello et al. 2021). To determine if environmental filtering impacts SES scores along the environmental gradient, we regressed the SES scores against the first axis of the PCA (PC1).

We used variance partitioning to quantify the relative contribution of environmental factors, pure space i.e. dispersal limitations and history, and spatially structured environmental factors on variation in SES functional dispersion. By controlling for spatial auto-correlation, variance partitioning as able to quantify the component of variation resulting solely from environmental factors (Smith and Lundholm 2010). We converted our coordinates which were in lat/long WGS 1984 format to cartesian coordinates using the geoXY function in the SoDA package (Chambers 2020). We then created Moran’s eigenvector maps (MEMs) from the cartesian coordinates using the dbmem function in adespatial (Dray et al. 2018). The function varpart from vegan (Oksanen et al. 2010) was used using the SES values as the response, and the standardized environmental variable matrix i.e. the same matrix as the PCA input and the MEM matrix as predictors.

We calculated the community-weighted mean (CWM) trait values for each individual trait using the dbFD function in the FD package (Laliberté et al. 2014). In order to understand how the mean value of each trait is correlated with environmental gradient, we regressed the CWM for each against the PC1 axis.

We regressed species richness against the PC1 axis.Taxonomic beta-diversity?

*Large scale environmental niche overlap*

We conducted an additional set of analyses to explore the relationship between overall climatic niche overlap and trait similarity. We extracted occurrence data from the Global Ant Biodiversity Informatics (GABI) database (Guenard et al. 2017) for each of the 11 ant species in our study. We excluded the Brazilian occurrence points for *Dorymyrmex insanus* because they were the only points south of Panama and they are likely *D. pyramicus* (Cuezzo and Guerrero 2012). For each species, occurrence points were thinned to one point per raster cell to (reason for thinning with citation). We generated pseudo-absences by randomly sampling 10000 points using the randompoints function from the dismo package (Hijmans et al. 2017). For each species, the background sampling area was determined by buffering the minimum convex hull of the occurrences by 100 km. This step was done to ensure that the background points were relevant to the range of each species. Mean annual temperature, maximum annual temperature and mean annual precipitation rasters from Worldclim, and mean and range of annual soil temperature rasters were used as input into species distribution models. The two soil rasters were resampled using bilinear interpolation to match the cells and extent of the climate rasters. The rasters were clipped to the following extent, all of the north America and some Caribbean. Therefore, while the area for the background points was unique for each species, the area used for prediction i.e. the extent of the climate rasters was identical for all species to enable the comparisons of geographic area. We used the package ENMeval (Muscarella et al. 2017) to create species distribution models for each of the ant species. This package implements the MaxEnt algorithm. The approach used by ENMeval runs a series of models in succession, using different combinations of tuning parameters to avoid overfitting while maximizing goodness of fit. We used regularization values between 1 and 5, as well as the following feature classes: : linear, linear + quad-ratic, hinge, linear + quadratic + hinge, which correspond to settings described within the packages vignette (Muscarella et al. 2017). We used random k folds at a value of five to determine which partitions to hold back for model validation. We chose models from the range of candidate models by using the one with the lowest AIC value. We assessed the predictive performance of each model using the area under the receiver operating characteristic curve (AUC). For the resulting SDM prediction rasters, we calculated the geographic i.e. climatic niche overlap for each pair of species using the function calc.niche.overlap. This function give Schoener’s D, a similarity index. We converted Shoener’s D to a dissimilarity index but subtracting each value from one. We then computed Gower dissimilarity for each pair of species using the species level mean trait values for each trait. We used a Mantel test to test if the dissimilarity in traits is correlated with the dissimilarity in climatic niche overlap between species.

**Results**

A total of 15519 individual ants from 11 species were collected and identified. All species are native to California. The most abundant species are *Solenopsis xyloni*, the native Southern fire ant, and *Pheidole hyatti*, the big-headed ant. This pitfall trapping was part of a larger arthropod sampling project that included sweep-netting of open areas and shrubs, as well as malaise trapping. There were no additional species detected using the other methods. On average, species richness per site was x and ranged from x to x.

The first PCA axis explained 48.16% of the variation. Precipitation, max and mean annual temperature and vegetation height were the greatest environmental contributors to the axis. The composite environmental gradient was therefore cooler, wetter sites (negative PC1 values) to hotter, drier sites (positive PC values).

*Taxonomic diversity patterns*

Alpha diversity does not shift with PC1.

The majority of taxonomic beta diversity comes from the turnover component (species replacement) (87.9%). Nestedness or gains losses is 12.1%. Mantels showed that the nestedness component of betadiversity is related to dissimilarity in the environment (Mantel r: 0.164, p = 0.028), but not the turnover component (Mantel r: 0.082, p = 0.135).

*Assembly*

Mean SES functional dispersion (mean = 0.58 +/- ) was significantly greater than zero (t-test, t = 3.54, p = 0.002). The values of SESdisp decreased as PC1 increased (Radj = 0.165), (coef = -0.507, p = 0.019) but not PC2 (coef = 0.17, p = 0.4) (Figure 2). SESdisp was not different between sites with large foundation shrubs vs open areas (p = ~0.86).

*Variance partitioning*

Environmental predictors explained 19% of the variation in SESfdisp and spatially distributed? Environmental variables explained 32% of the variation (Figure 3). The contribution of pure space was zero.

*Community-weighted mean traits*

Weber’s body length (AdjR2 = 0.26, p = 0.004) and relative femur length (AdjR2 = 0.243, p = 0.005) both decrease along the PC1 gradient (Figure 4). The other traits are not linearly related to the PC1 gradient. The contribution of taxonomic identity to trait variation varied with the trait (Figure 5).

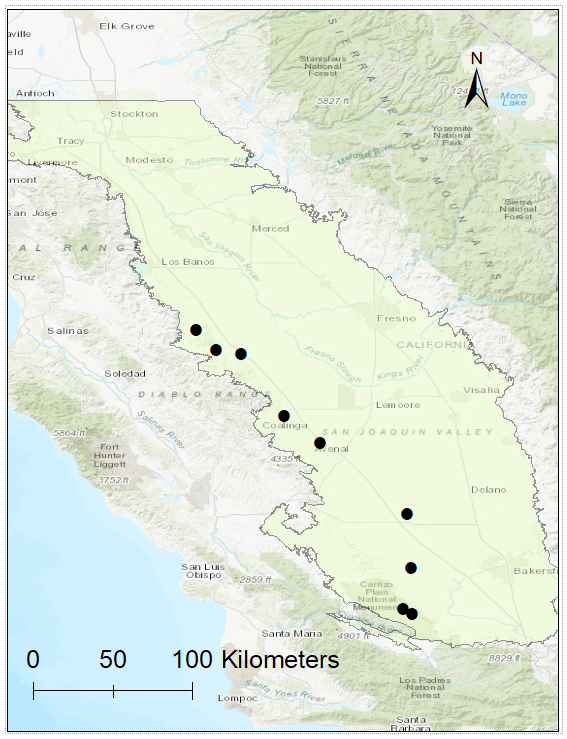


Figure 1: A map of the nine study areas.

Figure 1: PCA figures

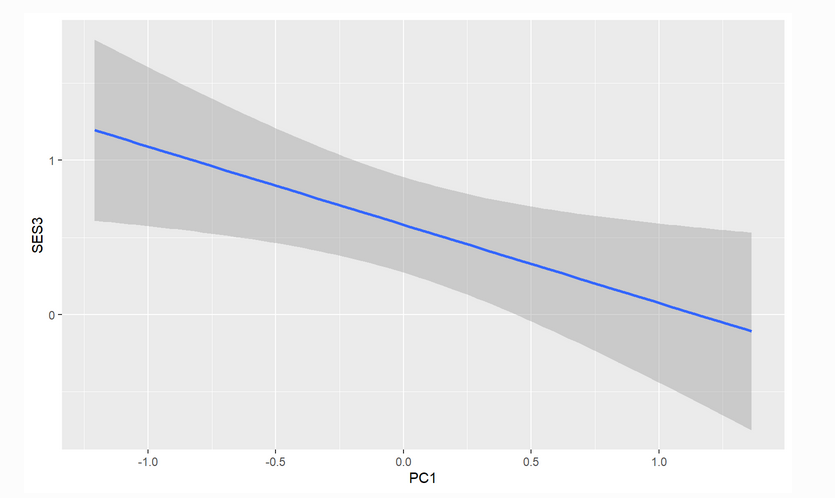
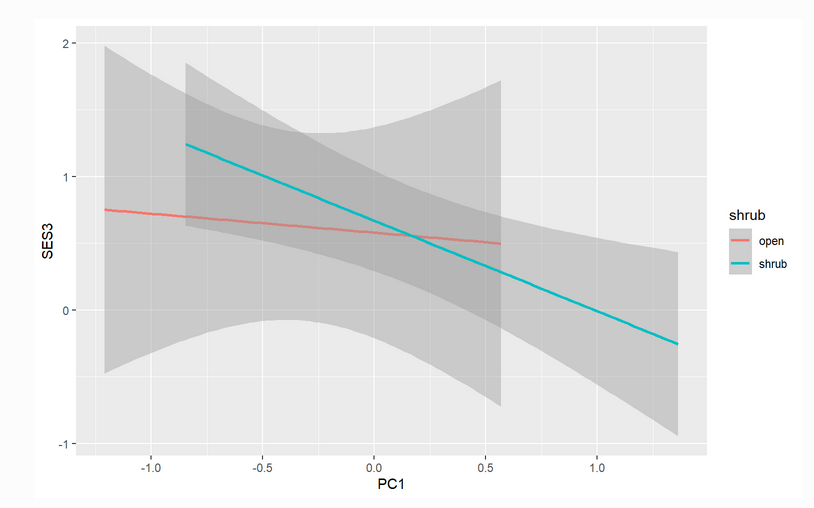
 

Figure 2: SESdisp regressed against PC1

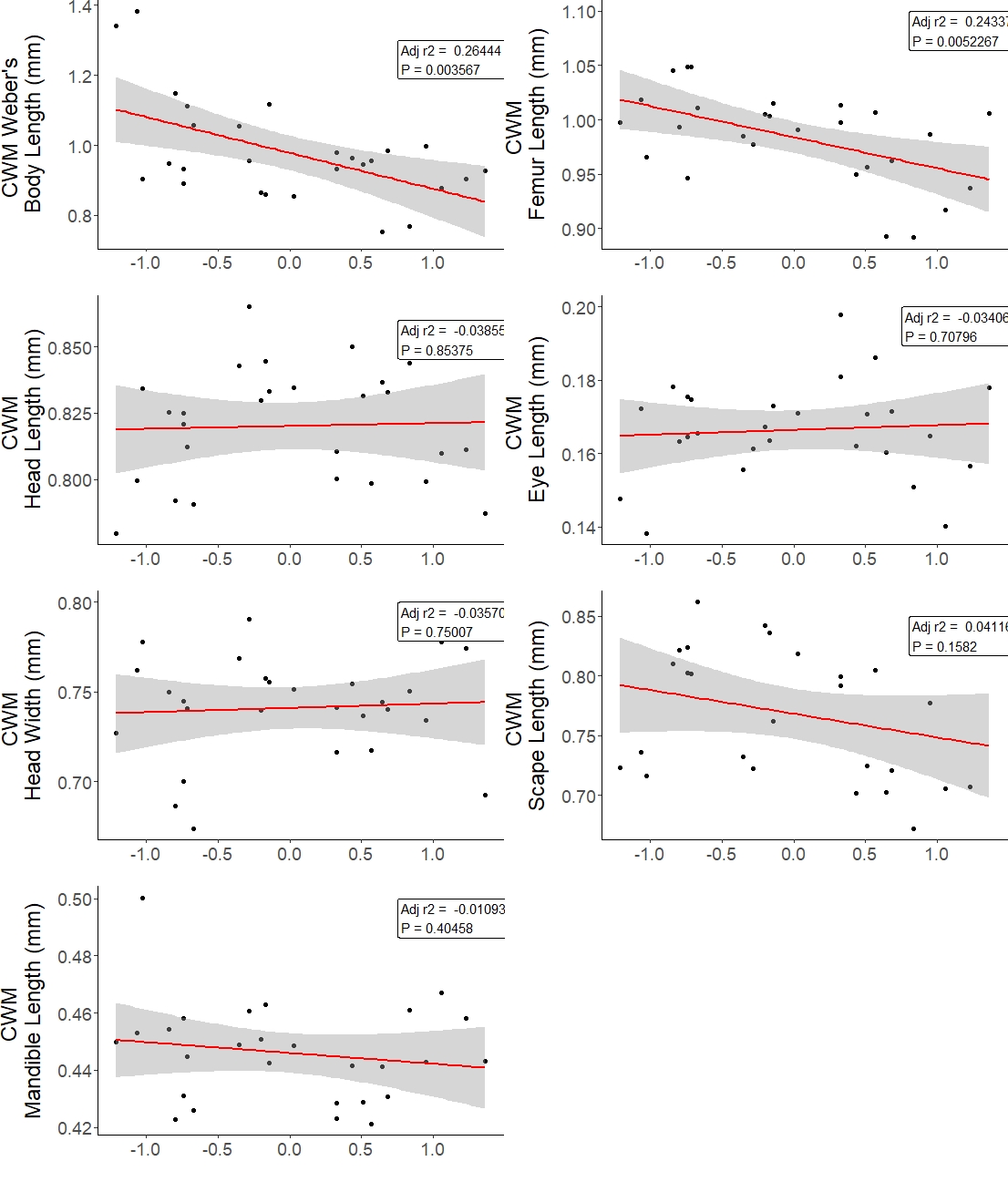


Figure 3: Linear regressions of community-weighted mean trait values and the composite environmental gradient (PC1). Sites with higher annual temperatures and lower annual precipitation are higher on the PC1 gradient. The shaded gray band represents the 95% confidence interval.

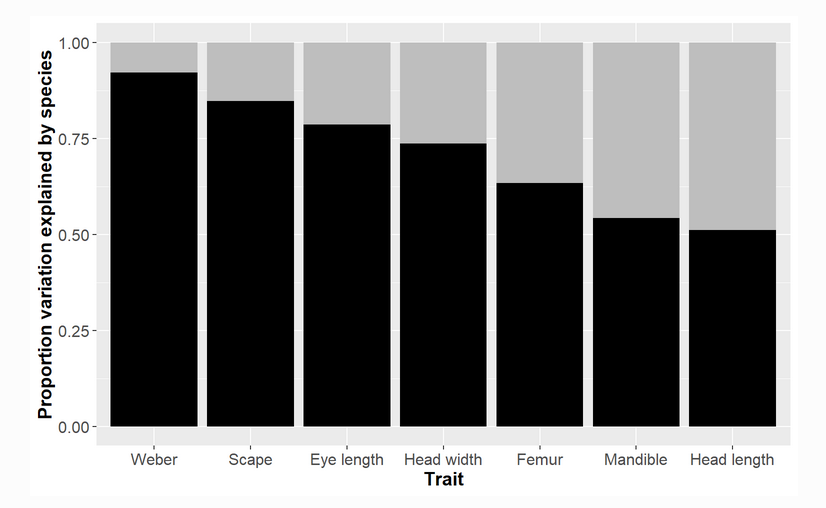
****

Figure 5: ITV and traits

**Discussion**

P1: Hypothesis, predictions. Overall assembly conclusions. Environmental and biotic filtering acts jointly.

P2: Assembly processes are scale-dependent

P3: Contrast stress gradient predictions for ants vs plants

P4: Discussion on implications for body size and femur length CWM along gradient

**Appendix**

Table 1: Summary of sampling dates and locations.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Site Name | Pitfall Trap Sampling dates | Coordinate of site centroid | | |
|  | |  | | |
| CaS | July 10 - 13 | | 35.11995 | 119.6283 | |
| Aug 2 - 5 | | 35.12000 | 119.6280 | |
| Sept 18 - 21 | | 35.11600 | 119.6240 | |
| CaSl | July 9 - 12 | | 35.09 | -119.574 | |
| Aug 2 - 5 | | 35.089 | -119.576 | |
| Sept 18 - 21 | | 35.065 | -119.539 | |
| SemiT | July 15 - 18 | | 35.658 | -119.612 | |
| Aug 6 - 9 | | 35.658 | -119.612 | |
| Sept 24 – 27 | | 35.658 | -119.61 | |
| Lokern | July 14 - 17 | | 35.354 | -119.584 | |
| Aug 3 - 6 | | 35.354 | -119.584 | |
| Sept 23 - 26 | | 35.355 | -119.588 | |
| PaPl | July 23 - 27 | | 36.698 | -120.799 | |
| Aug 12 - 15 | | 36.696 | -120.795 | |
| Sept 10 -13 | | 36.7 | -120.801 | |
| Aven | July 21 - 24 | | 36.094 | -120.197 | |
| Aug 8 - 11 | | 36.088 | -120.19 | |
| Sept 6 - 9 | | 36.0878 | -120.1912 | |
| Mov | July 28 - 31 | | 36.563 | -120.547 | |
| Aug 13 - 16 | | 36.562 | -120.545 | |
| Sept 12 - 16 | | 36.561 | -120.548 | |
| SiCr | July 20 - 23 | | 36.586 | -120.687 | |
| Aug 12- 15 | | 36.586 | -120.686 | |
| Sept 12 - 16 | | 36.586 | -120.688 | |
| Coal | July 16 - 19 | | 36.213 | -120.305 | |
| Aug 8 - 11 | | 36.212 | -120.304 | |
| Sept 6 – 9 | | 36.213 | -120.303 | |

Table 2: Scene ID

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| eVIIRS NDVI | EVUSS20200630202007066 | EVIIRS NDVI 375M | 154.78 MiB | Complete |
| eVIIRS NDVI | EVUSS20200707202007136 | EVIIRS NDVI 375M | 155.12 MiB | Complete |
| eVIIRS NDVI | EVUSS20200714202007206 | EVIIRS NDVI 375M | 154.11 MiB | Available |
| eVIIRS NDVI | EVUSS20200714202007276 | EVIIRS NDVI 375M | 169.38 MiB | Complete |
| eVIIRS NDVI | EVUSS20200721202007276 | EVIIRS NDVI 375M | 154.01 MiB | Complete |
| eVIIRS NDVI | EVUSS20200728202008036 | EVIIRS NDVI 375M | 154.96 MiB | Available |
| eVIIRS NDVI | EVUSS20200728202008106 | EVIIRS NDVI 375M | 170.24 MiB | Complete |
| eVIIRS NDVI | EVUSS20200804202008106 | EVIIRS NDVI 375M | 155.96 MiB | Available |
| eVIIRS NDVI | EVUSS20200811202008176 | EVIIRS NDVI 375M | 155.71 MiB | Available |
| eVIIRS NDVI | EVUSS20200811202008246 | EVIIRS NDVI 375M | 168.62 MiB | Available |
| eVIIRS NDVI | EVUSS20200818202008246 | EVIIRS NDVI 375M | 153.81 MiB | Available |
| eVIIRS NDVI | EVUSS20200825202008316 | EVIIRS NDVI 375M | 153.74 MiB | Available |
| eVIIRS NDVI | EVUSS20200825202009076 | EVIIRS NDVI 375M | 167.25 MiB | Available |
| eVIIRS NDVI | EVUSS20200901202009076 | EVIIRS NDVI 375M | 152.77 MiB | Available |
| eVIIRS NDVI | EVUSS20200908202009146 | EVIIRS NDVI 375M | 147.03 MiB | Available |
| eVIIRS NDVI | EVUSS20200908202009216 | EVIIRS NDVI 375M | 162.40 MiB | Available |
| eVIIRS NDVI | EVUSS20200915202009216 | EVIIRS NDVI 375M | 150.21 MiB | Available |
| eVIIRS NDVI | EVUSS20200922202010056 | EVIIRS NDVI 375M | 162.86 MiB | Available |

*From <*[*https://dds.cr.usgs.gov/queue/orderlist/BulkDownload/20220822\_104216/1*](https://dds.cr.usgs.gov/queue/orderlist/BulkDownload/20220822_104216/1)*>*

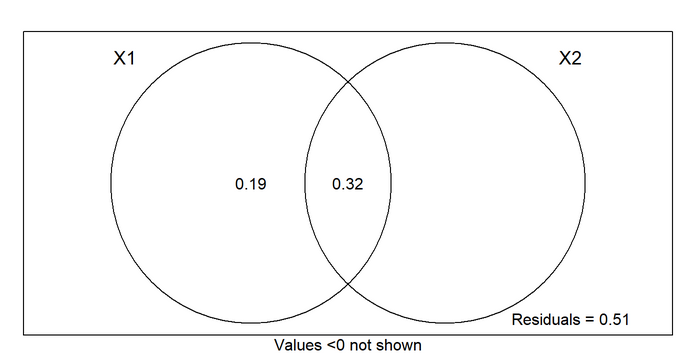


Figure 3: Variance partitioning

Supplemental Information

Table S1: Population level trait measurements

Literature Cited

USGS EROS Archive - Vegetation Monitoring - EROS Visible Infrared Imaging Radiometer Suite (eVIIRS) , https://doi.org/10.5066/P9Q3B2A7